AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

UNIVERSITÉ DE LORRAINE

École doctorale IAEM Lorraine

Contrôle acoustique d'un parc éolien

THÈSE

présentée et soutenue publiquement le 25 septembre 2018

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention informatique)

par

Baldwin DUMORTIER

Composition du jury

Rapporteurs :	Patrick Danès	Professeur, Université Paul Sabatier, Toulouse.
	Christophe Gonzales	Professeur, Université Pierre et Marie Curie, Paris.
Examinateurs :	Marianne Clausel	Professeur, Université de Lorraine, IECL, Nancy.
	Mathieu Lagrange	Chargé de recherche, Laboratoire des Sciences du Numérique de Nantes, Nantes.
Directeurs de thèse :	Emmanuel Vincent	Directeur de recherche, Inria, LORIA, Nancy.
	Madalina Deaconu	Chargée de recherche, Inria, IECL, Nancy.

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Mis en page avec la classe thesul.

« La musique donne une âme à nos coeurs et des ailes à la pensée »
- Platon.

Remerciements

Je tiens tout d'abord à remercier mes encadrants de thèse, Emmanuel Vincent et Madalina Deaconu, qui m'ont beaucoup appris et qui ont été d'un soutien sans pareil pendant ces années de thèse. Leurs compétences techniques, leur clairvoyance et leur sympathie se sont avérés indispensables à ce travail.

Je remercie ensuite les collègues chercheurs et amis de l'académie pour leur aide, leurs conseils et leur soutien. Je pense en particulier à Yann Salaün, Nathan Souviraà-Labastie, Antoine Liutkus, Dung Tran, Ken Deguernel, Mathieu Fontaine, Denis Jouvet, et de façon générale aux membres des deux équipes de recherche que j'ai pu côtoyer, Multispeech et Tosca. Je remercie également tous les autres membres du laboratoire qui ont permis le bon déroulement de cette thèse.

Je remercie également Venathec d'avoir financé ma thèse et ses employés pour m'avoir initié à l'ingénierie acoustique et sur le plan humain. Je pense en particulier à Jérémy Schild, Patrice Cornu, Mickaël Favre-Felix, Eric Marchal et Stéphane Griffiths.

Je remercie affectueusement mes amis et anciens colocataires rencontrés à Nancy durant ces années loin de ma ville natale. Je pense à particulier à Rémy et son caractère trempé, à Lulu ma nounou de cœur, à l'élégante et adorable Diane, à Andréa et sa nonchalance, à Bilia et sa joie de vivre mais aussi au soleil, essentiel à mon équilibre mental et trop peu présent dans cette région.

Pour finir, je remercie mes parents, mon petit frère et toute ma famille, qui bien que loin, sont toujours aux petits soins. Je remercie mes amis de Toulouse et d'ailleurs qui m'ont manqué mais ont été d'un bon soutien.

À mes parents et à mon frère.

Sommaire

Partie I Introduction **1**

Chapitre 1
Contexte

1.1 Contexte général . . . 3
1.2 Problème du choix d'un critère de la gêne sonore . . . 3
1.3 Réglementation dans le monde . . . 4
1.4 Réglementation Française . . . 5
1.4.1 Réglementation et normes en vigueur en France . . . 5
1.4.2 Le décret ICPE . . . 6
1.4.3 Mesures normatives . . . 6
1.4.4 Classes homogènes et indicateur de niveau . . . 8
1.4.5 Incertitudes sur les mesures . . . 8
1.5 Bridage des éoliennes et contrôle acoustique . . . 11
1.6 Méthode d'évaluation et de bridage en France . . . 12
1.6.1 Déroulement général d'une étude . . . 12
1.7 Limitations . . . 14
1.8 Projet iEar et cadre de la thèse . . . 14
1.9 Problématique et contributions . . . 15
1.10 Plan du manuscrit . . . 17

Chapitre 2
État de l'art

2.1 Commande discrète des systèmes . . . 19
2.1.1 Calcul d'une commande déterministe optimale . . . 19
2.1.2 Calcul d'une commande optimale sous incertitudes . . . 26
2.2 Estimation d'un modèle paramétrique pour du contrôle à temps discret . . . 27
2.2.1 Modèles temps-série et réseaux de neurones . . . 28

2.2.2 Inférence récursive bayésienne 29
2.2.3 Filtres de Kalman itérés 36
2.2.4 Autres filtres basés sur des méthodes d'optimisation 37
2.2.5 Filtrage particulaire et méthodes d'ensemble 37
2.3 Apprentissage des paramètres 39
2.3.1 Apprentissage des paramètres des modèles ARIMA 40
2.3.2 Apprentissage des paramètres des filtres bayésiens récursifs par filtrage bayésien 40
2.3.3 Apprentissage des paramètres des filtres Bayésiens récursifs : estimation par maximisation de la vraisemblance 41

Partie II Formulation du problème 45

Chapitre 3
Formulation Générale

3.1 Notations et analyse du problème 47
3.1.1 Analyse des variables du problèmes 48
3.2 Formulation originelle du problème de contrôle stochastique 53
3.2.1 Fonction objectif 53
3.2.2 Contraintes 53
3.2.3 Formulation complète 54
3.3 Simplification du problème et terminologie 54

Partie III Problème déterministe 57

Chapitre 4
Problème déterministe à contraintes instantanées

4.1 Introduction 59
4.2 Formulation 59
4.2.1 Généralisation du problème 60
4.3 Algorithme branch and bound 62
4.3.1 Algorithme général 62
4.3.2 Variantes autour de l'algorithme général 65
4.3.3 Relaxation continue 66
4.4 Algorithme de gradient projeté 68

4.4.1 Introduction de l'algorithme de gradient linéaire par morceaux sur un exemple simple . . . 68
4.4.2 Description de l'algorithme de gradient . . . 72
4.4.3 Éléments de preuve de l'optimalité de l'algorithme de gradient projeté 75
4.4.4 Heuristique gloutonne d'initialisation . . . 75
4.5 Évaluation expérimentale . . . 76
4.5.1 Données utilisées . . . 77
4.5.2 Résultats . . . 77

Chapitre 5
Problème déterministe à contraintes sur les médianes temporelles

5.1 Formulation . . . 81
5.2 Heuristiques proposées . . . 81
5.2.1 Démarche expérimentale . . . 81
5.2.2 Algorithme glouton de contrôle non-prédictif . . . 82
5.3 Évaluation expérimentale . . . 84
5.3.1 Description des variantes des algorithmes de contrôle testés . . . 84
5.3.2 Protocole expérimental . . . 85
5.3.3 Résultats . . . 86
5.3.4 Interprétation des résultats pour le plan statique et pour les méthodes à émergence instantanée . . . 88
5.4 Conclusion sur le problème déterministe basé sur le critère des médianes temporelles . . . 91

Partie IV Problème stochastique 93

Chapitre 6
Introduction

6.1 Problème et hypothèses . . . 95
6.2 Démarche . . . 96
6.2.1 Estimation des incertitudes et suivi des variables cachées . . . 96
6.2.2 Optimisation stochastique . . . 97
6.3 Schéma général d'optimisation . . . 97

Chapitre 7
Estimation des incertitudes et recalibration statistique du modèle

7.1 Définition d'un modèle, suivi des variables cachées et estimation des paramètres 99

7.2 Modèle d'incertitude 100
7.3 Etape M simplifiée : estimation des paramètres du modèle sous la connaissance de l'état caché 104
7.3.1 Estimateur de vraisemblance conditionnellement au bruit résiduel . . 104
7.3.2 Estimation par linéarisation de Taylor à l'ordre 1 106
7.3.3 Estimation par calcul des moments 107
7.3.4 Calcul du maximum de la log-vraisemblance 109
7.3.5 Évaluation expérimentale 110
7.4 Étape E : estimation des états cachées sous la connaissance des paramètres des incertitudes 112
7.4.1 Modèle général 112
7.4.2 Méthodes de filtrage 113
7.4.3 Évaluation expérimentale 113
7.5 Problème général : apprentissage conjoint par un algorithme Espérance-Maximisation (EM) 117
7.5.1 Algorithme EM proposé 117
7.5.2 Évaluation expérimentale 119

Chapitre 8
Optimisation stochastique

8.1 Formulation du problème 123
8.2 Caractérisation des solutions admissibles 124
8.2.1 Représentation graphique des solutions admissibles 124
8.2.2 Approximation par une méthode de Monte-Carlo 124
8.2.3 Approximation basée sur la fonction de répartition d'une loi gaussienne 125
8.3 Évaluation et comparaison de la méthode de Monte-Carlo et de l'approximation proposée 127
8.3.1 Valeurs de tests 127
8.3.2 Résultats 127
8.4 Adaptation de l'algorithme branch and bound pour le problème stochastique 129

Partie V Conclusion et perspectives **131**

Chapitre 9
Conclusion et perspectives

9.1 Travail effectué 133

9.1.1 Solution déterministe . . . 133
9.1.2 Solution stochastique . . . 134
9.2 Perspectives concernant la définition et l'apprentissage des paramètres . . . 135
9.2.1 Redéfinition des classes homogènes par des méthodes de clustering . . 135
9.2.2 Utilisation d'une méthode de gradient pour l'apprentissage des paramètres basée sur la méthode des moments proposée . . . 135
9.2.3 Ajout d'une loi a priori sur les incertitudes et estimation des biais du modèle . . . 136
9.2.4 Paramétrisation de l'incertitude de séparation de sources en fonction du rapport signal-à-bruit . . . 137
9.2.5 Modèle de prédiction . . . 137
9.3 Perspectives concernant le calcul de la commande optimale . . . 138
9.3.1 Utilisation d'un modèle acoustique par bandes de fréquences . . . 138
9.3.2 Apprentissage de la contrainte dans le problème *chance constraint programing* . . . 138

Annexes

Annexe A
Formules Mathématiques

A.1 Formules mathématiques de passage d'une loi log-gaussienne à une loi gaussienne 139

Annexe B
Détails des résultats de l'algorithme EM

Bibliographie **143**

Table des figures

1.1 Pondération A. 7
1.2 Exemple d'un découpage en 32 classes homogènes standardisées. 9
1.3 Schéma général du système. 15

2.1 Représentation graphique de l'espace des solutions d'un problème MINLP convexe. 21
2.2 Représentation graphique de l'espace des solutions d'un problème LP. 22
2.3 Représentation graphique du partitionnement d'un algorithme *branch and bound* 23
2.4 Représentation graphique du parcours d'un algorithme *branch and bound*. 24
2.5 Représentation graphique de l'utilisation du gradient pour construire une enveloppe convexe à contraintes linéaires. 25
2.6 Illustration du fonctionnement du filtre de Kalman. 31
2.7 Propagation d'une distribution gaussienne à travers la fonction $f(x) = 1 + x^3$. . . 34

4.1 Exemple d'arbre des solutions admissibles pour 4 éoliennes et 5 modes de bridage. 63
4.2 Exemple de parcours sur un arbre pour 4 éoliennes et 5 modes de bridage, en partant d'une solution initiale de 2000 kW. 64
4.3 Exemple de fonction $\mathfrak{f}$. 67
4.4 Représentation graphique des bridages possibles de l'exemple. 69
4.5 Représentation graphique des contraintes acoustiques données par $\tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}$. . . 70
4.6 Représentation graphique des bridages admissibles. 70
4.7 Représentation graphique des itérations de l'algorithme de gradient sur l'exemple. 71
4.8 Représentation de la projection du gradient au point $\tilde{\mathbf{x}}^{(3)}$ sur la contrainte. 71
4.9 Représentation des projections du gradient au point $\tilde{\mathbf{x}}^{(5)}$ sur les contraintes. . . . 72
4.10 Distribution des erreurs relatives pour l'algorithme glouton. 78
4.11 Distribution des erreurs relatives pour l'algorithme *branch and bound* (variante retenue). 79
4.12 Évolution temporelle des erreurs relatives pour l'algorithme *branch and bound* (variante retenue) pour les échantillons ayant un temps de calcul supérieur à 5 s. 79

5.1 Productions obtenues pour le problème déterministe. 86
5.2 Pertes de production vis à vis du plan *full power* pour le problème déterministe. 86
5.3 Nombre de fois où l'émergence réglementaire est dépassée pour le problème déterministe. 87
5.4 Dépassements maximaux (dB) pour le problème déterministe. 87
5.5 Distribution du bruit résiduel et des émergences instantanées engendrées par le plan statique. 90
5.6 Distribution du bruit résiduel et des émergences instantanées engendrées par l'algorithme instantanée V1. 90

5.7 Distribution du bruit résiduel et des émergences instantanées engendrées par l'algorithme instantanée V3. 91

7.1 Niveaux résiduels observés lors d'une étude acoustique en fonction de la vitesse du vent. 102
7.2 Évolution temporelle du niveau résiduel en un point. 102
7.4 Histogramme des évolutions du résiduel sur des données mesurées. 103
7.5 Coupe de la fonction de vraisemblance pour le problème où les valeurs terrain sont $\sigma_x = 1,5$, $\sigma_a = 1,5$ et $\sigma_{r\bullet} = 1,5$. 109
7.6 Résultats d'estimation des paramètres d'incertitude par maximisation de la log-vraisemblance. 110
7.7 Histogramme de la racine des erreurs au carré normalisée pour les deux approximations. 111
7.8 RNSE en fonction de la valeur terrain de l'incertitude à estimer. 111
7.9 Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman étendu dans le cas $J = 5$, $\sigma = 4,5$. 115
7.10 Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman *unscented* dans le cas $J = 5$, $\sigma = 4,5$. 116
7.11 Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman *central difference* dans le cas $J = 5$, $\sigma = 4,5$. 116
7.12 Erreur RMSE de l'estimation de l'émergence par le filtre *central difference Kalman filter* en fonction de l'écart-type des incertitudes du modèle (σ_x et σ_a) et de l'écart-type de l'incertitude de la séparation de sources ($\sigma_{r\bullet}$). 117
7.13 Erreur RMSE de l'estimation de l'émergence par le filtre *central difference Kalman filter* en fonction du nombre d'éoliennes sur le parc et du nombre de points de mesure. 117

8.1 Représentation de la loi jointe de $[l_t^T \; r_t^T]^T$ (courbes bleues foncées), de q (courbe rouge) et des solutions admissibles (aire bleue sous la courbe). 125
8.2 Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 10$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$. 128
8.3 Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 50$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$. 128
8.4 Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 32$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$. 128
8.5 Comparaison entre les valeurs de $\widehat{m}_l^{\max}$ obtenues par Monte-Carlo, et celles obtenues par l'approximation proposée en fonction de $\widehat{m}_r$. 129
8.6 Histogramme des erreurs entre les valeurs de $\widehat{m}_l^{\max}$ obtenues par Monte-Carlo, et celles obtenues par l'approximation proposée. 129

Première partie

Introduction

1

Contexte

1.1 Contexte général

L'énergie éolienne a le vent en poupe. Sa production représente aujourd'hui environ 3 % de la production d'électricité mondiale et son utilisation augmente de manière exponentielle. Cependant son utilisation connaît de nombreux détracteurs à cause notamment de l'utilisation de terres rares (aimants permanents contenant néodyme, dyprosium et praséodyme), du danger pour les oiseaux ou du coût de démantèlement et de recyclage des différents composants des éoliennes. Dans ce contexte, l'acoustique est un des sujets majeurs de polémique, la gêne sonore allant jusqu'à être accusée d'avoir un impact néfaste sur la santé à cause des infrasons que les éoliennes génèrent [Bolin et al., 2011]. De ce fait, de nombreuses réglementations acoustiques ont été mises en place dans le monde quant à la distance minimale entre les riverains et les éoliennes à proximité et quant à l'impact sonore maximal autorisé.

Cependant, limiter la gêne sonore nécessite d'avoir au préalable étudié plusieurs étapes clés : il faut pouvoir d'une part définir une notion de gêne sonore et proposer un critère pour la caractériser, et d'autre part être capable de mesurer ce critère et de mettre en place une solution permettant de limiter la gêne. De prime abord, toutes ces étapes peuvent sembler anodines, cependant aucune de ces étapes n'est réellement résolue aujourd'hui dans le contexte de l'éolien qui, en plus des problématiques classiques d'acoustique, a les siennes qui lui sont propres.

1.2 Problème du choix d'un critère de la gêne sonore

À titre d'explication, partons de zéro pour présenter les lacunes d'un critère de gêne sonore trop simpliste. En première approximation on pourrait partir du principe qu'il suffit d'utiliser un sonomètre bien calibré, de le placer au niveau des habitations, et de mesurer le niveau acoustique total moyen sur un pas de temps ni trop long ni trop court afin de caractériser la gêne sonore. En pratique, l'utilisation d'un tel critère présente de nombreux problèmes. Pour commencer, le niveau mesuré est peu robuste sur de courts intervalles de temps (1 seconde par exemple). Lors de l'apparition d'un impact sonore brusque du fait d'une activité humaine par exemple, les niveaux mesurés augmentent vite et sont alors non représentatifs. Il est donc nécessaire de filtrer les valeurs mesurées afin d'éviter d'introduire dans la mesure des aléas non-reproductibles et non-représentatifs. En France par exemple, les indicateurs acoustiques sont les médianes des niveaux d'énergie d'une seconde mesurés par les sonomètres sur des intervalles de 10 min. Cependant, même en ayant des mesures représentatives obtenues par ce type de filtrage, les niveaux mesurés évoluent rapidement dans le temps, même au sein d'une seule journée et pour deux raisons.

Tout d'abord, l'émission sonore des éoliennes et la propagation du son des éoliennes jusqu'aux habitations varient fortement avec les conditions météorologiques. La première dépend principalement de la vitesse de rotation des pales des éoliennes et donc directement de la vitesse du vent au niveau du rotor des éoliennes. La seconde est directement et fortement impactée par le vent, le gradient de température et l'humidité de l'air. De ce fait et même en fonctionnement normal, la pression acoustique reçue au niveau des habitations peut varier de 20 dB (décibels) au sein d'une même journée [Larsson & Israelsson, 1991; Larsson & Öhlund, 2014].

D'autre part, les sonomètres ne mesurent pas seulement la pression acoustique issue des éoliennes mais la pression acoustique issue de l'environnement, qui dépend aussi des autres sources sonores de l'environnement. De ce fait, même en l'absence de bruit éolien, le niveau acoustique varie fortement, même dans un environnement rural et perçu comme « calme » d'environ 30 dB dans une journée.

Dans ce contexte, les acousticiens utilisent les termes de *niveau ambiant* pour désigner la pression acoustique de la totalité du signal acoustique perçu en un point donné, comprenant à la fois le signal acoustique des éoliennes et de l'environnement et appellent *niveau résiduel*, le niveau du signal acoustique perçu provenant de l'environnement sans la contribution acoustique des éoliennes. Enfin, ils nomment *niveau particulier*, le niveau que généreraient les éoliennes seules, en l'absence de toute autre source sonore [1].

On peut se rendre compte dès lors que le bruit ambiant constitue une très mauvaise mesure de la gêne sonore. Du point de vue de la gêne sonore, il est évident qu'un environnement sonore affichant à un instant donné un niveau ambiant d'environ 40 dB mais issu d'un niveau résiduel de 20 dB et un niveau particulier de 40 dB n'est pas équivalent et bien plus gênant qu'un environnement sonore qui affiche à instant donné un niveau ambiant lui aussi d'environ 40 dB mais issu d'un niveau résiduel de 40 dB et d'un niveau particulier de 20 dB.

Par conséquent, on pourrait être tenté d'utiliser le niveau particulier comme critère de gêne sonore mais on se heurte alors à deux nouveaux écueils. Premièrement, le bruit particulier ne peut pas être mesuré et il est difficile à estimer. Même à l'aide des meilleures simulations numériques, une forte incertitude apparaît sur son estimation, la propagation dépendant des conditions météorologiques, instables par nature. Secondement, le bruit particulier n'est pas non plus une bonne mesure de la gêne sonore. D'un point de vue psycho-acoustique, il apparaît différents phénomènes liés à la perception humaine qui ont été étudiés longuement. La proportion entre le niveau particulier et le niveau résiduel influe sur la perception : un niveau résiduel élevé proportionnellement au niveau particulier provoque un phénomène de masquage qui atténue la gêne relative perçue [Fégeant, 1999; Pedersen & Halmstad, 2003].

1.3 Réglementation dans le monde

En pratique, les états ont souvent opté pour des réglementations simples afin de s'affranchir partiellement de la complexité du problème, qui ne rendent que partiellement compte de la gêne sonore.

Dans certains cas, comme dans la législation française, on tient compte de la réponse fréquentielle de l'oreille en utilisant une pondération (pondération A) sur le signal, et on considère comme critère de conformité le critère d'*émergence* qui est la soustraction entre le niveau ambiant et le niveau résiduel en dB. L'émergence acoustique représente le gain acoustique apporté par la

1. En acoustique, le terme « niveau particulier » désigne de manière générale le niveau imputé à la source acoustique objet de l'étude seule dans l'environnement [Norme, 1996]

Pays/institution	Limite de jour (dBA)		Limite de nuit (dBA)	
	émergence	ambiant	émergence	ambiant
World Health Organization	$+\infty$	50	$+\infty$	45
Australie	5	35	5	35
France	5	35	3	35
Allemagne	$+\infty$	60 ou 55 ou 50	$+\infty$	45 ou 40 ou 35
Royaume-Uni	5	35 ou 40	5	35 ou 40
États-Unis	$+\infty$	40	$+\infty$	40

TABLEAU 1.1 – Législation acoustique sur les éoliennes dans le monde.

source nuisible, ici le bruit des éoliennes, sur l'environnement existant et prétend donc mesurer la perception relative du bruit généré par les éoliennes.

Par exemple, une émergence de 3 dB revient à dire que les éoliennes génèrent autant de bruit que le bruit initial de d'environnement.

À l'inverse, certains pays, comme les États-Unis, utilisent encore le critère de niveau ambiant même si, comme expliqué précédemment, il constitue un critère de limitation encore plus éloigné de la gêne sonore réelle.

De manière générale, si tous les pays ont leurs spécificités, on peut néanmoins observer que les législations ont pour point commun d'utiliser le niveau ambiant, l'émergence acoustique ou un mélange de ces deux critères. Le Tableau 1.1 présente les seuils utilisés pour un certain nombre de pays [Schild & Chavand, 2015].

1.4 Réglementation Française

Dans la suite de la thèse, on se place dans le cadre de la législation française qui est une des législations les plus restrictives [Schild & Chavand, 2015]. De ce fait, la thèse ne perd pas en généralité dans la mesure où toutes les propositions avancées pourraient être réadaptées pour les autres législations qui en pratique, utilisent toutes un critère de conformité réglementaire basé principalement sur l'émergence ou le bruit ambiant. Dans la suite, on résume donc les principaux éléments de la législation française.

1.4.1 Réglementation et normes en vigueur en France

Voici les principales normes et réglementations en France :

— Décret ICPE (Installation Classée pour la Protection de l'Environnement) 28/08/2011. Contient l'essentiel des règles à respecter. Il s'agit d'un décret général qui repose sur trois critères : le critère d'émergence, le critère de limite de propriété et le critère de tonalité marquée (expliqués par la suite).

— Norme PR NF S 31-114 [2]. En projet mais applicable. Il s'agit de la norme principale et spécifique à l'éolien qui décrit comment doivent être réalisées les mesures et dans quelles conditions.

— Norme NF S 31-010 et NF S 31-S110. Ces deux normes historiques et plus générales traitant de l'acoustique dans l'environnement (pas seulement sur l'éolien) décrivent elles aussi comment les mesures de niveaux acoustiques doivent être réalisées. Même si elles

2. Notation standardisée : PR : projet, NF : norme française, S : industrie diverses/acoustique.

ne s'appliquent pas directement au domaine de l'éolien, elles sont à l'origine de la norme précédente. La norme PR NF S 31-114 peut être considérée comme une adaptation directe de ces normes pour le cas spécifique de l'éolien.

1.4.2 Le décret ICPE

La plupart des normes citées dans la partie précédente concernent la manière d'effectuer les mesures. Ici on présente le décret ICPE qui précise les règles à respecter. Lorsque la réglementation n'est pas respectée, l'administration demande à l'exploitant de missionner un cabinet d'acoustique. Le plus gros risque à considérer pour l'exploitant est l'arrêt définitif ou temporaire des machines s'il ne peut pas respecter ces critères d'un point de vue technique.

Le critère d'émergence

Il s'agit de la différence (en dB) maximale admissible entre le niveau ambiant pondéré A et le niveau résiduel pondéré A. L'émergence est fondamentalement une mesure de la perception du bruit d'une installation par un humain. Cette perception est modélisée comme le rapport en échelle linéaire entre la pression acoustique ambiante et la pression acoustique résiduelle, toutes deux filtrées par la pondération A ou comme une différence en dB. Le critère ICPE limite l'émergence, à proximité des lieux d'habitation, à un seuil maximal fixé à +5 dBA la journée, et à +3 dBA la nuit. Il est à noter que ce critère n'est plus considéré si le bruit ambiant est inférieur à un autre seuil fixé à ce jour à 35 dBA. De ce fait, la valeur du dépassement d'émergence considérée est alors le minimum entre le dépassement d'émergence et le niveau ambiant auquel on a retranché 35 dBA.

Critère de valeur limite à proximité des éoliennes

Ce critère indique que le niveau sonore maximal ambiant sur un périmètre de l'éolienne est de 70 dBA de jour et 60 dBA de nuit. Ce périmètre dépend de la hauteur de moyeu et du rotor et est défini comme suit : « le plus petit polygone dans lequel sont inscrits les disques de centre aérogénérateur et de rayon R », le rayon R étant fonction des dimensions des éoliennes du parc considéré. Ce critère est en fait une adaptation pour le domaine éolien du critère de limite de propriété de l'ICPE. En pratique, on ne s'intéressera pas à ce critère dans les travaux de la thèse, qui est en général respecté.

Critère de tonalité marquée

Le critère de tonalité marquée cherche à limiter la prépondérance de certaines fréquences. Sur toute bande fréquentielle considérée de taille 1/3 d'octave, le L_{eq} [3] ne doit pas être « trop »supérieur aux L_{eq} issus des 4 bandes d'1/3 d'octave les plus proches. Quantitativement, il s'agit de +10 dB maximum entre 50 et 315 Hz et +5 dB maximum entre 400 et 8000 Hz. En pratique, s'il y a présence de tonalité marquée, cela signifie que l'éolienne émet un sifflement gênant lié à un défaut de fabrication et l'éolienne doit être alors réparée.

1.4.3 Mesures normatives

Le décret ICPE décrit précédemment se rattache à des normes qui décrivent les règles nécessaires de mesurage à respecter pour la validité des mesures. La principale est la norme PR NF S

3. voir 1.4.3

31-114 qu'on résume succinctement dans la suite.

Mesures acoustiques

Les études sont réalisées à partir d'indicateurs statistiques issus des mesures acoustiques, dont les paramètres sont définis par la norme. Les mesures des différents niveaux sonores mesurés par chaque microphone doivent être synchrones et mesurées par pas d'1 seconde.
Pour comprendre, voici quelques définitions en acoustique :

- Niveau sonore en dB L : Le niveau sonore L en dB correspond au rapport en dB du carré de la pression acoustique efficace p et d'un niveau arbitraire de référence $p_0 = 2,10^{-5}$Pa, c'est-à-dire $L = 10\log_{10}(\frac{p^2}{p_0^2})$
- Niveau sonore équivalent L_{eq} : Pour caractériser un niveau sonore variant dans le temps, on utilise la notion de niveau sonore équivalent (ou niveau sonore moyen). Cela revient à intégrer le niveau sonore L sur une période de temps.
- Niveau sonore équivalent pondéré A, $L_{\mathrm{A},eq}$: La réponse de l'oreille humaine étant dépendante de la fréquence, on utilise une pondération fréquentielle dite « pondération A ». Cela revient à appliquer un filtre fréquentiel qui approxime grossièrement la réponse de l'oreille avant de calculer le niveau sonore équivalent. L'unité de mesure est appelée dBA.
- Niveau fractile L_n (ou $L_{\mathrm{A},n}$) : Le niveau fractile (ou niveau fractile pondéré A) correspond au niveau sonore équivalent en dB (ou dBA) qui a été dépassé n % du temps, sur un intervalle de temps fixé.

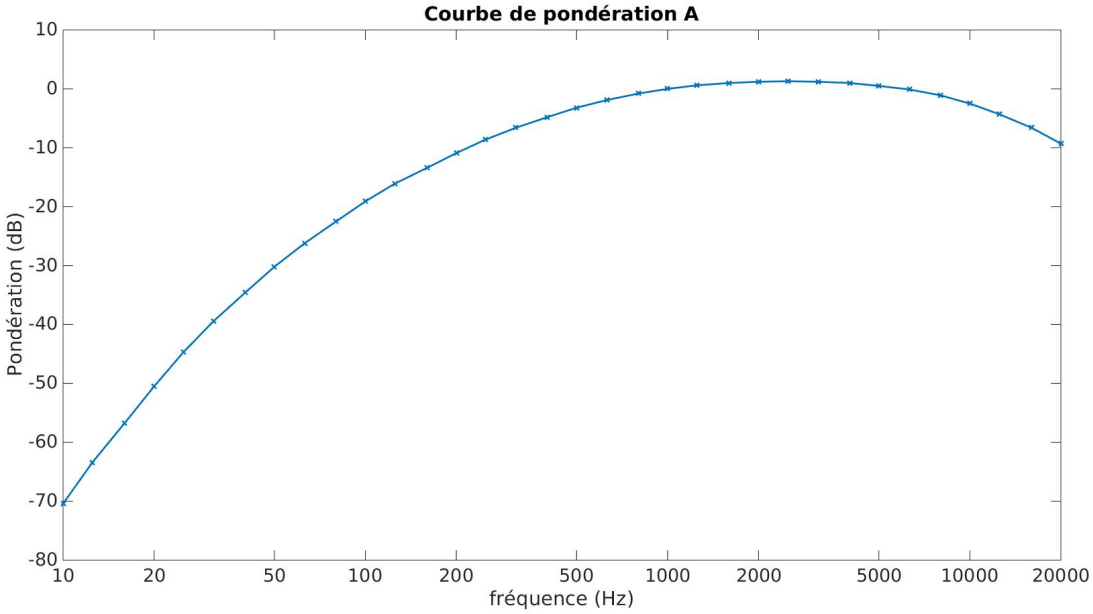


FIGURE 1.1 – Pondération A.

D'autre part la norme précise que le niveau de bruit à considérer sont les $L_{A,50}$ sur des intervalles de 10 min des $L_{A,eq}$ d'1 sec.

Mesures de vent

Lors d'une étude, les niveaux de bruit sont couplés à la vitesse du vent, mesurée au même moment. La vitesse du vent considérée est la vitesse moyenne sur des intervalles de 10 min, mesurée au niveau du rotor de l'éolienne et exprimée de façon standardisée à partir d'une extrapolation utilisant un profil de vent logarithmique à une hauteur de 10 m sur un sol de rugosité de référence de 0,05 m. L'équation (1.1) donne le calcul de la vitesse standardisée.

$$v_{\text{standard}} = v_{\text{rotor}} \frac{\log \frac{h_{\text{rotor}}}{0{,}05}}{\log \frac{10}{0{,}05}} \tag{1.1}$$

La variable v_{rotor} est la vitesse mesurée au niveau du rotor, h_{rotor} est la hauteur du rotor par rapport au sol, et v_{standard} est la vitesse standardisée.

1.4.4 Classes homogènes et indicateur de niveau

Les conditions acoustiques variant rapidement et fortement dans divers contextes industriels, les normes françaises ont défini en acoustique une notion de **classe homogène**. La classification sert à répertorier les mesures acoustiques en sous-ensembles observés dans des conditions acoustiques semblables et donc supposées reproductibles.

Les classes homogènes sont de natures diverses, dépendent au cas par cas des conditions de mesure et ne se limitent pas à celles présentées dans la thèse. Cependant, on retiendra dans un but de simplification et dans le cas spécifique de l'éolien, les classes homogènes définies par un partitionnement des données sur l'heure de la journée, sur la direction et sur la vitesse du vent. Cela revient à dire que pour des vitesses de vent, des directions, et des heures identiques, les mesures de niveau particulier et résiduel ont une distribution censée être concentrée.

En pratique, on découpe pour les heures deux classes avec comme bornes 7h et 22h, pour les directions des secteurs angulaires de 60 ° et pour les vitesses de vent des intervalles de taille 1 m/s centrés sur les valeurs entières de vitesse de vent. Les mesures sont répertoriées par classe et les différentes classes sont alors étudiées indépendamment. La Figure 1.2 présente graphiquement un exemple d'un tel découpage.

Un *indicateur de niveau* est alors calculé par classe homogène, et par type d'indicateur (ambiant ou résiduel). Cet indicateur est calculé comme la médiane des $L_{\text{A},50}$ répertoriés dans chaque classe. L'indicateur est considéré comme valide s'il est calculé à partir d'au moins 10 valeurs. Dans le cas contraire, la norme spécifie que l'on peut interpoler/extrapoler les indicateurs sous certaines conditions. On note L_{amb} et L_{res} respectivement les indicateurs de niveau ambiant et résiduel dans une classe homogène donnée.

1.4.5 Incertitudes sur les mesures

La norme NFS 31-114 aborde les incertitudes de mesure sur les niveaux de bruits absolus et sur les différences de niveaux (émergences). On considère deux types d'incertitudes différentes. D'une part, les incertitudes dues à la distribution statistique de l'indicateur dans une classe donnée, appelées incertitudes de type A. D'autre part, les incertitudes dues à des erreurs métrologiques, appelées incertitudes de type B.

Incertitudes sur les niveaux de bruit

On note L un indicateur de bruit (désignant indifféremment l'indicateur L_{amb} ou L_{res}). Pour **chaque indicateur** (ambiant et résiduel), pour chaque classe homogène et pour chaque classe

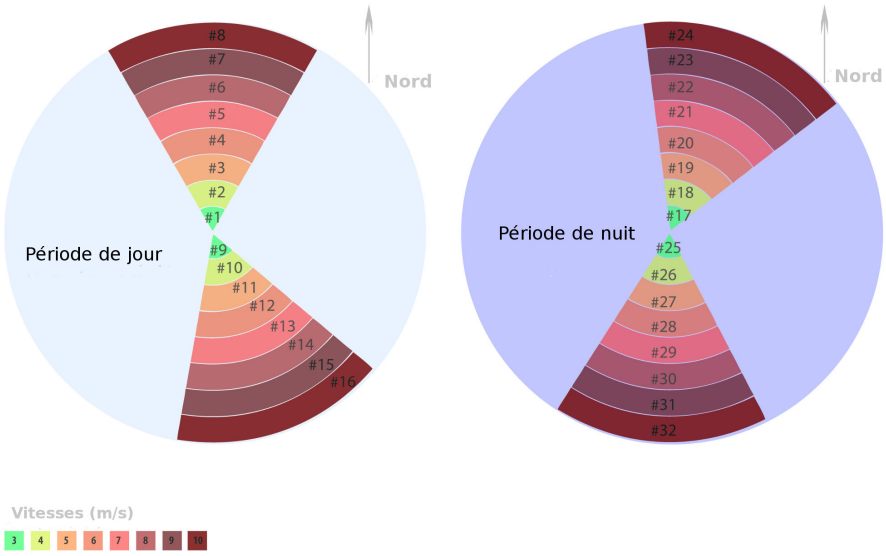


FIGURE 1.2 – Exemple d'un découpage en 32 classes homogènes standardisées.

de vent, on a :

— L'incertitude $U_A(L)$ de type A. Il s'agit de l'incertitude liée à la variation naturelle des niveaux acoustiques dans une classe homogène. Elle est calculée à l'aide d'une formule définie par la norme. Pour N niveaux L_i (les L_i sont les $L_{A,50m}$ permettant de calculer L) d'une classe homogène pour un indicateur de niveau sonore, on a :

$$U_A(L) = 1,858 \times t(L) \times \frac{\text{mediane}|L_i - \text{mediane}(L_i)|}{\sqrt{N-1}} \tag{1.2}$$

où $t(L)$ est un correctif pour les petits échantillons défini par $t(L) = \frac{2N-2}{2N-3}$.

— L'incertitude $U_B(L)$ de type B. C'est l'incertitude liée à la mesure sono-métrique. Elle est définie de la manière suivante :

$$U_B(L) = \sqrt{\sum_k (U_{Bk}(L))^2}. \tag{1.3}$$

où les $U_{Bk}(L)$ sont les composantes de l'incertitude métrologique. Elles sont données, ainsi que leur valeur type dans le tableau suivant (issu de la norme PR NF S 31 114) :

	Composante	Incertitude type	Condition
U_{B1}	Calibrage	± 0,20 dB ; ± 0,20 dBA	Durée maximale entre 2 calibrages : 15 jours
U_{B2}	Appareillage	± 0,20 dB ; ± 0,20 dBA	
U_{B3}	Directivité	± 0,52 dBA	Direction de référence du microphone : verticale
U_{B4}	Linéarité en fréquence et pondération fréquentielle	± 1,05 dBA	
U_{B5}	Température et humidité	± 0,15 dB ; ± 0,15 dBA	
U_{B6}	Pression statique pour une classe homogène	± 0,25 dB ; ± 0,25 dBA	
U_{B7}	Incertitude liée au vent	dépend du niveau ambiant mesuré dans les classes de vents les plus proches	

L'incertitude totale $U_C(L)$ s'écrit alors :

$$U_C(L) = \sqrt{(U_A(L))^2 + (U_B(L))^2}. \tag{1.4}$$

Incertitudes sur les différences de niveaux

De la même manière, pour un indicateur d'**émergence**, noté E, on a également des incertitudes de types A et B qui elles-mêmes dépendent des incertitudes de l'indicateur ambiant L_{amb} et de l'indicateur résiduel L_{res}. Elles se calculent comme suit :

— L'incertitude $U_A(E)$ de type A de l'émergence est calculée à partir des incertitudes de type A, $U_A(L_{\text{amb}})$ du niveau ambiant, et $U_A(L_{\text{res}})$ du niveau résiduel de la façon suivante :

$$U_A(E) = \sqrt{(U_A(L_{\text{amb}}))^2 + (U_A(L_{\text{res}}))^2}. \tag{1.5}$$

— L'incertitude de $U_B(E)$ de type B de l'émergence s'écrit de la même manière que celles des niveaux ambiants et résiduels, mais avec des valeurs différentes pour ses composantes. En effet, pour certaines composantes de l'incertitude métrologique, certaines erreurs se compensent ou se cumulent, ce qui donne le tableau suivant :

	Composante	Incertitude type	Condition
U_{B1}	Calibrage	négligeable	Durée maximale entre 2 calibrages : 15 jours
U_{B2}	Appareillage	négligeable	
U_{B3}	Directivité	négligeable	Direction de référence du microphone : verticale
U_{B4}	Linéarité en pondération fréquentielle	$\pm 1,05 \times \sqrt{2 - 2,10^{\frac{-E}{10}}}$ dBA	
U_{B5}	Température et humidité	± 0,22 dBA	
U_{B6}	Pression statique pour une classe homogène	± 0,24 dBA	
U_{B7}	Incertitude liée au vent	négligeable	

De la même manière, l'incertitude générale U_C s'écrit alors :

$$U_C(E) = \sqrt{(U_A(E))^2 + (U_B(E))^2}. \tag{1.6}$$

Utilisation des incertitudes

Les incertitudes sont utilisées de la manière suivante actuellement. Tout d'abord, l'incertitude générale U_C, sur les niveaux et sur les émergences ne doit pas dépasser un seuil, noté S_{QL}, actuellement fixé à 3 dB. En effet, les incertitudes de mesure ne doivent pas être du même ordre que les contraintes acoustiques, auquel cas il est difficile de prouver que le système respecte la réglementation. Dans le cas où ce seuil est dépassé, il faut faire de nouvelles mesures. D'autre part, lorsque ce seuil n'est pas dépassé, on utilise les incertitudes pour corriger les valeurs des indicateurs. Les incertitudes de mesure avantagent l'exploitant lorsque les incertitudes de type A sont faibles au regard des incertitudes de type B. Cela revient à dire que, plus on observe une faible variation des mesures dans une classe donnée, plus on a le droit d'utiliser les erreurs métrologiques présumées en faveur de l'exploitant. Quantitativement, pour un indicateur I (de niveau ambient, résiduel ou d'émergence), on utilise sa valeur corrigée I^* donnée par la formule suivante :

$$I^* = I - K.U_C + K.\max(U_C - \min(U_B, 1.9), 0) \times \frac{S_{QL}}{(S_{QL} - \min(U_B, 1.9))}. \tag{1.7}$$

Le terme $K.U_C$ où K est un facteur d'élargissement, pris actuellement égal à 1, avantage l'exploitant en diminuant la valeur de l'indicateur I considéré (niveau ambiant ou émergence). Le terme $K.\max(U_C - \min(U_B, 1.9), 0) \times \frac{S_{QL}}{(S_{QL} - \min(U_B, 1.9))}$, quant à lui, vient contrebalancer le terme précédent, d'autant plus que l'incertitude de type A est conséquente. Dans le cas où celle-ci serait nulle, ce terme le serait aussi. Dans tous les cas, les incertitudes ne désavantagent jamais l'exploitant car on a toujours $I^* \leq I$.

1.5 Bridage des éoliennes et contrôle acoustique

D'autres complications surviennent lors de la mise en place d'une solution de réduction du niveau acoustique. Dans de tels environnements, il n'est pas possible d'ériger des murs ou d'autres isolants acoustiques et la réduction du niveau passe alors par le bridage individuel des éoliennes. Les éoliennes modernes sont à ce titre pourvues de modes de bridage qui agissent sur la prise au vent des pales et sur la démultiplication afin de ralentir la vitesse de rotation et de diminuer le niveau acoustique. Ces modes, souvent en nombre réduit, permettent d'atténuer l'émission acoustique des éoliennes jusqu'à environ -5 dB mais réduisent fortement la production d'électricité, occasionnant ainsi de fortes pertes économiques aux exploitants. Le Tableau 1.2 présente un exemple des caractéristiques acoustiques et électriques d'une éolienne type, en fonction du mode de bridage et de la vitesse du vent mesurée à 10 mètres.

Puissance acoustique (dBA)	3 m/s	4 m/s	5 m/s	6 m/s	7 m/s	8 m/s	9 m/s	10 m/s
pleine puissance	88,9	95,6	101,2	104,8	106,5	105,9	104,8	103,9
mode 1	92,7	96,6	100,7	103,5	104,4	104,2	103,8	103,3
mode 2	92,8	96,8	100,4	102,2	102,4	102,3	102,1	102
mode 3	92,2	94,9	98,7	101,1	100,8	101	101	100,8
mode 4	92,7	96,8	100,3	102,1	102,9	103,9	103,9	103,3
mode 5	92,1	94,8	98,3	100,1	99,5	99,1	98,4	97,8

Puissance électrique (kW)	3 m/s	4 m/s	5 m/s	6 m/s	7 m/s	8 m/s	9 m/s	10 m/s
pleine puissance	171	456	955	1671	2551	3134	3293	3329
mode 1	171	464	949	1655	2463	3017	3244	3294
mode 2	171	465	945	1586	2207	2683	2979	3122
mode 3	171	452	929	1511	2012	2453	2761	2900
mode 4	171	465	944	1576	2275	2966	3249	3296
mode 5	171	452	926	1448	1878	2195	2349	2414

TABLEAU 1.2 – Tableaux des caractéristiques acoustiques et électriques des modes de bridage d'une éolienne type selon la vitesse du vent mesurée à 10 mètres. Le modèle de l'éolienne n'est pas divulgué pour des raisons de confidentialité. Notons également qu'il arrive que les valeurs de puissance acoustique/électrique du mode pleine puissance soient inférieures à celle d'un autre modes à certaines vitesses. On considère alors qu'il s'agit d'une erreur de mesure, et les modes de valeurs supérieurs ne sont alors pas considérés à ces vitesses.

Les exploitants font alors face à un problème de maximisation de la production des éoliennes de leur parc sous la contrainte du respect de la réglementation au niveau des habitations. Or, comme mentionné précédemment, le niveau résiduel et la propagation acoustique évoluent fortement et rapidement. Contrôler le bruit ambiant, le bruit particulier ou bien l'émergence acoustique est un véritable défi.

Actuellement, les études acoustiques sont réalisées avec des sonomètres pendant une durée limitée de quelques semaines au maximum et on bride les éoliennes de la manière suivante :

— Les modes de bridage affectés aux éoliennes sont pré-calculés à partir de mesures effectuées pendant une courte période de temps, qui est sélectionnée afin d'être représentative des conditions acoustiques annuelles et à partir de simulations acoustiques basées sur des modèles numériques complexes.
— Le bridage mis en place ensuite n'est pas directement asservi à des capteurs acoustiques mais aux conditions météorologiques (sens du vent, vitesse du vent) qui sont censées être corrélées aux conditions acoustiques. On parle de *bridage statique*.

En pratique, cette façon de faire ne permet pas de s'adapter aux conditions exceptionnelles et aux évolutions à long terme des conditions acoustiques. Par exemple, si une campagne de mesurage est réalisée en hiver, le bruit particulier ou l'émergence ont toutes les chances d'avoir des valeurs très différentes en été, même en moyenne : météorologie différente et événements exceptionnels (grillons l'été, animaux, bruits de chantiers, circulation automobile ...). Ceci est d'autant plus vrai que les plans de bridage, qui ne sont asservis actuellement que par le vent, sont calculés sur 20 ans ce qui induit une forte probabilité d'erreur à terme. Par exemple, si dans les années qui suivent, la circulation automobile à proximité augmente à cause de la démographie, cette manière de procéder ne peut pas en tenir compte, et la perte électrique peut devenir conséquente pour l'exploitant.

1.6 Méthode d'évaluation et de bridage en France

Dans cette partie, on présente le déroulement général d'une étude acoustique en France.

1.6.1 Déroulement général d'une étude

Il existe actuellement deux types d'études : les études pré-implantation et les études post-implantation. Les premières visent à faire l'étude d'un parc à venir (donc pas encore construit au moment de l'étude), tandis que les secondes visent à vérifier la conformité de parcs déjà existants. Un plan de bridage est calculé par les ingénieurs acousticiens dans le cas où il y a risque de dépassement prévisionnel pour les études pré-implantation, et dans le cas où il y a un dépassement réel pour les études de post-implantation/conformité.

Étude pré-implantation

Il s'agit d'étudier l'impact acoustique d'un parc éolien **avant sa construction.** Voici comment se déroule une étude dans le cas de pré-implantation :

— Tout d'abord, des sonomètres sont placés à des points stratégiques, à l'extérieur (sur des lieux de vie : jardin, terrasse, *etc.*) des habitations. Un mat de mesure de vent peut également être installé, si le site n'en est pas équipé.
— Des mesures de bruit résiduel et des mesures de vent sont alors réalisées sur environ 15 jours. Les données brutes récupérées sont les $L_{\mathrm{A,eq}}$ 1 s pour les mesures acoustiques, et les vitesses et directions du vent pour chaque intervalle de 10 min.
— On crée alors les classes homogènes en séparant les mesures de jour et de nuit, puis en classant les mesures en fonction des directions de vent par secteurs de 60 ° et pour finir en classant les mesures par vitesse de vent mesurée à 10 mètres.
— Pour chaque classe homogène on calcule les indicateurs de niveau résiduel. Comme expliqué précédemment, il s'agit des médianes des $L_{\mathrm{A,50}}$ par classe homogène et par classe de vent.

— A l'aide d'un logiciel de simulation, on simule la propagation pour calculer le niveau particulier. On utilise la topographie du terrain, la position des éoliennes, des points de mesure (sonomètres virtuels), les caractéristiques du sol (rugosité, absorbance, *etc.*), les caractéristiques du vent (vitesse et direction) et d'autre paramètres pertinents afin de calculer la propagation. Cela permet de connaître l'atténuation énergétique du signal sonore de chaque éolienne sur chaque point de mesure. On peut alors calculer la contribution de chaque éolienne sur chaque sonomètre, en utilisant les puissances acoustiques d'émission des éoliennes (fournies par le constructeur).
— S'il y a dépassement, il faut alors ajuster le bridage des éoliennes. Pour cela, on utilise un calcul d'optimisation sous contraintes afin de construire un bridage optimal en considérant les points suivants :
  — D'une part, on utilise les données électriques et acoustiques, fournies par le constructeur, qui donnent la puissance acoustique d'émission et la puissance de production de chaque éolienne, par mode de bridage et par vitesse de vent.
  — D'autre part, on utilise les données de propagation issues de la simulation numérique. Il s'agit d'une matrice en dB à ajouter à la puissance acoustique d'émission afin de calculer la pression acoustique perçue par chaque microphone pour chaque éolienne. [4]
  — On considère l'indicateur de bruit résiduel mesuré pour calculer l'émergence.
  — Un algorithme itératif permet alors de calculer un nouveau plan de bridage.

  Un rapport complet est ensuite envoyé au client.

Étude post-implantation

Dans le cas des études post-implantation, la manière de procéder est similaire à la différence que l'on possède pour le niveau ambiant des valeurs mesurées en plus des valeurs simulées. Si la manière de procéder diffère aussi peu, c'est pour des raisons de coût, car mesurer séparément l'impact de chaque éolienne en chaque point de mesure impliquerait des campagnes de mesurage beaucoup plus longues ce qui entraînerait des pertes élevées pour l'exploitant. Il y a néanmoins quelques différences, voici comment l'on procède :

— Ici aussi, des sonomètres et un mat de mesure de vent sont placés à des points stratégiques, à l'extérieur (sur des lieux de vie) des habitations.
— En revanche, afin de mesurer les niveaux résiduels l'exploitant effectue des cycles marche/arrêt de son parc éolien.
— Les mesures de bruits résiduel, ambiant et les mesures de vent sont alors réalisées sur 15 jours également.
— On classe les mesures par classe homogène et par type de mesure (résiduel ou ambiant).
— Pour chaque classe homogène on calcule alors, non seulement les indicateurs de bruits résiduels, mais aussi les indicateurs ambiants.
— Les indicateurs permettent alors de statuer d'un risque de dépassement des seuils réglementaires.
— Si un dépassement est observé, on simule numériquement la propagation. Si les niveaux ambiants ainsi simulés sont trop différents des niveaux mesurés (différence supérieure à 2 dB environ), on cherche à recaler le modèle de propagation. On modifie les paramètres de la simulation (humidité, température, vent, *etc.*) afin de faire correspondre les mesures à la simulation.

4. En effet, une fois les données de propagations fixées (vent, météo, topographie, *etc.*) la différence en dB entre la puissance émise par une source sonore et la pression acoustique issue de cette source en un point donné est supposée constante.

— Une fois le modèle recalé et s'il y a dépassement, on calcule à nouveau un plan de bridage comme précédemment.

Un rapport peut être alors rendu au client.

1.7 Limitations

Aujourd'hui et en résumé, l'ingénierie acoustique des parcs éoliens est insatisfaisante pour les raisons suivantes :

— **Problème simultané de la mesure du critère de conformité réglementaire et de l'incertitude des variables acoustiques** : Comme présenté précédemment, les critères de conformité réglementaire proposés sont soit simples à mesurer mais peu pertinents (dans le cas du bruit ambiant) ou bien plus pertinents mais difficiles à mesurer en raison des différentes sources acoustiques de l'environnement et de leur variabilité.

— **Problème d'optimisation et de contrôle temporel du niveau particulier** : Même dans le cas où on est capable de mesurer correctement le critère de conformité réglementaire à contrôler, le schéma actuel de bridage proposé ne permet pas un contrôle correct du critère de conformité. Les bridages actuels ne sont asservis qu'aux conditions météorologiques et ne prennent pas en compte l'évolution des conditions météorologiques et environnementales qui changent sans cesse et peuvent entraîner soit des dépassements des seuils réglementaires et donc éventuellement de la gêne acoustique réelle, soit de lourdes pertes économiques.

1.8 Projet iEar et cadre de la thèse

Afin de résoudre les problèmes actuels de mesure du critère de conformité réglementaire et d'asservissement du bridage des éoliennes à celle-ci, Venathec SAS, bureau d'études en acoustique, a lancé le projet de recherche et développement iEar dont le but final est de créer un système commercial asservi en temps-réel, basé sur un ensemble de capteurs et sur un serveur de contrôle afin de tenir compte en permanence de l'état acoustique de l'environnement et d'ajuster dynamiquement le bridage des différentes éoliennes en fonctionnement.

La Figure 1.3 représente le schéma de fonctionnement prévu du produit.

Des sonomètres seront postés à certains endroits stratégiques, à côté des habitations, de manière permanente afin de prendre en compte les changements de niveau sonore dus aux éoliennes.

Ces informations sont ensuite envoyées à la iEar box, le serveur, qui les traite afin d'établir un plan de bridage toutes les 10 minutes qui est ensuite envoyé au SCADA (Supervisory Control And Data Acquisition) éolien, c'est-à-dire au système de contrôle des éoliennes. En outre, la iEar Box bénéficie également d'informations provenant du SCADA éolien qui lui transmet des mesures météorologiques et de production électrique en temps réel.

Sur le plan scientifique, la solution est basée sur deux traitements successifs :

1. Un système de séparation de sources qui estime le niveau de bruit particulier et le niveau de bruit résiduel à partir du signal ambiant. La recherche menée à ce sujet est en dehors du cadre de la thèse.

2. Un système de contrôle des éoliennes qui, à partir des estimées du bruit particulier et du bruit résiduel, calcule un plan de bridage optimal. **C'est cette partie qui constitue le sujet de thèse**.

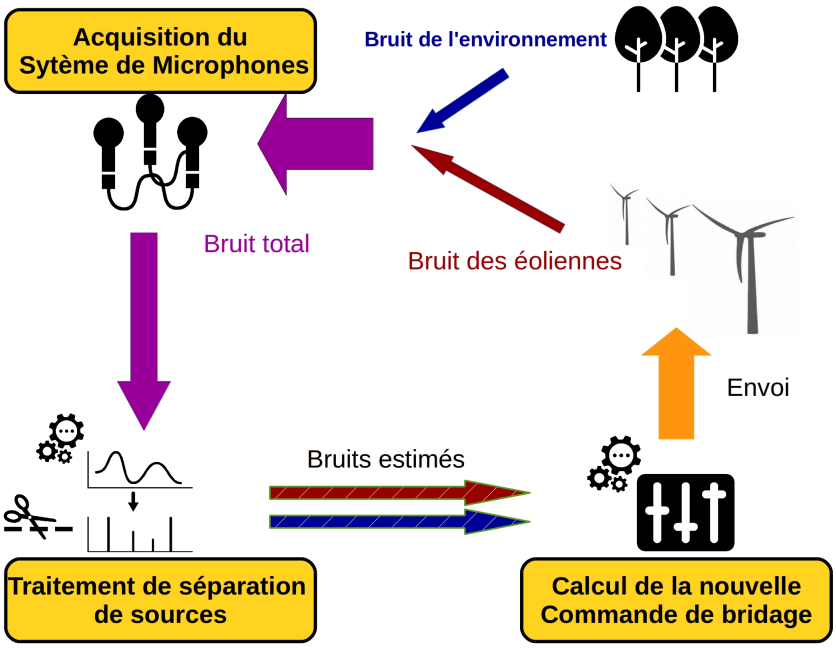


FIGURE 1.3 – Schéma général du système.

On suppose donc dans la suite qu'un système de mesure sonore multi-microphones est fourni ainsi qu'un système de séparation de sources de type considéré comme une « boite noire » permettant d'estimer au mieux le niveau du bruit particulier et du bruit résiduel. Sous ces hypothèses, **le but de la thèse est de concevoir un algorithme de contrôle robuste qui maximise la production électrique sous la contrainte du respect de la réglementation**.

À ce sujet, si l'utilisation des contraintes basées sur des médianes décrites dans la norme française PR NF S 13-114 est étudiée au Chapitre 5, on s'appuie sur un critère d'émergence instantanée dans les autres chapitres et ce pour plusieurs raisons :

- iEar implique une stratégie de bridage différente de la stratégie actuelle qui contraint les modes de bridage à être choisis de manière statique pour chaque classe homogène.
- La législation évolue rapidement. La norme PR NF S 13-114 en application est en cours de ré-écriture.
- Le critère d'émergence instantané est plus pérenne car il est historiquement celui utilisé dans les autres domaines de l'industrie comme indiqué dans les normes NF S 31-010 et NF S 31-S110.
- Le projet a la volonté de créer un système générique, qui s'émancipe du critère basé sur des médianes afin de pouvoir éventuellement s'exporter à l'étranger.

Ainsi, par la suite, on s'inscrira dans le cadre réglementaire français qui est le marché principal de Venathec, tout en conservant un caractère général dans les choix proposés.

1.9 Problématique et contributions

Le projet iEar a l'ambition d'utiliser des outils jusqu'alors non utilisés dans le domaine de l'acoustique en proposant un système de contrôle qui s'appuie sur des capteurs acoustiques permanents et de l'asservissement. Ceci a ouvert des problématiques de recherche dédiées autour

de la séparation de source et du contrôle.

La thèse explore le domaine du contrôle dans ce contexte et a pour finalité de proposer un algorithme de contrôle complet sur la base des connaissances dans le domaine de l'acoustique des éoliennes.

Durant la thèse, on s'est intéressé à l'adaptation des méthodes d'optimisation numérique et de contrôle afin de subvenir aux différents besoins techniques et scientifiques du projet.

Formalisation mathématique du problème et conception du schéma général de contrôle

Avant iEar, aucun système à notre connaissance n'a proposé un schéma d'asservissement des parcs éoliens à des capteurs acoustiques. De ce fait, le projet part de zéro en ce qui concerne l'algorithme de contrôle. Une première contribution a été de formaliser et de construire un schéma complet de contrôle. On a proposé un ensemble de notations, un schéma de contrôle, une formulation du problème et des idées spécifiques d'ingénierie dans le cadre du transfert de la thèse.

Inférence des variables du problème : incertitude sur la séparation et sur le modèle

De nombreuses variables acoustiques et météorologiques sont nécessaires afin de brider au mieux les éoliennes, et la méthode actuelle ne permet pas un contrôle fin du bruit perçu au niveau des habitations pour les raisons expliquées dans la partie 1.7. Pour améliorer ce contrôle, iEar propose l'utilisation d'un système de séparation de sources afin d'estimer la pression acoustique reçue au niveau des habitations.

Cependant, cette estimation ne suffit pas à elle seule à améliorer le contrôle. En supposant que l'estimation fournie par la séparation de sources fonctionne bien, il est nécessaire d'estimer les contributions individuelles de chaque éolienne sur la pression particulière en un point donné sans quoi il n'est pas possible de savoir quelles sont les éoliennes à brider ou à débrider.

Deuxièmement, même les meilleurs systèmes de séparation de sources sont entachés d'erreurs [Vincent et al., 2006], ce qui par voie de conséquence peut entraîner de fortes erreurs dans le contrôle et donc des dépassements ou des pertes de production électrique.

Pour finir, la commande de bridage est appliquée avec un retard, ce qui entraîne une incertitude supplémentaire, les conditions acoustiques et météorologiques évoluant rapidement.

La difficulté de l'estimation dans ce contexte réside dans l'absence de *vérité terrain*. La connaissance du niveau particulier et résiduel est toujours entachée d'une incertitude potentiellement grande.

Une seconde contribution a donc été de proposer une estimation statistique des incertitudes sur les différentes variables du problème. On propose donc un modèle de recalibration basé sur des incertitudes suivant des lois gaussiennes et estimées par du *filtrage de Kalman* non-linéaire.

Apprentissage et recalibration en ligne

Les éoliennes ont une durée de vie d'environ 20 ans et les conditions de propagations et le bruit résiduel sont donc susceptibles d'évoluer à moyen et long terme. Par exemple, l'humidité agit sur la propagation qui évolue au fil des saisons.

De ce fait, une autre problématique est celle de l'apprentissage des incertitudes du problème. On propose donc, comme autre contribution, un modèle d'apprentissage pour la recalibration automatique des paramètres du système basé sur un algorithme EM (espérance-maximisation).

Optimisation pour le calcul d'une commande optimale, maximisant la puissance électrique d'un parc, sous la contrainte acoustique

Finalement, lorsque les variables et les paramètres du système sont connus, le calcul de la commande de bridage optimale se présente comme un calcul de maximisation de la puissance électrique sous la contrainte du respect de la réglementation. La difficulté est la gestion de l'explosion combinatoire, les variables de choix étant un mode de bridage discret pour chaque éolienne du parc. On propose pour résoudre ce problème, un algorithme *branch and bound* dédié à ce problème, initialisé à partir d'une heuristique gloutonne et reposant sur l'utilisation d'une relaxation continue.

1.10 Plan du manuscrit

On présente tout d'abord un état de l'art autour des problématiques et des contributions de la thèse au chapitre 2. Dans la partie II, on introduit la formulation du problème en présentant les notations, les terminologies proposées et une formulation mathématique du problème avec les contraintes basées sur des médianes. Dans la partie III, on s'intéresse au problème allégé des incertitudes, avec au chapitre 4 la formulation de l'algorithme *branch and bound* proposé, puis une adaptation pour les contraintes basées sur des médianes temporelles au chapitre 5. Dans la partie IV, on s'intéresse aux incertitudes du système en proposant une formulation du problème au chapitre 6, puis une proposition d'estimation et d'apprentissage au chapitre 7, ainsi qu'une ré-adaptation de l'algorithme déterministe au chapitre 8. Pour finir, la partie V présente les conclusions et les perspectives scientifiques dans le cadre de ce projet.

2

État de l'art

À notre connaissance, il n'existe actuellement aucun système de contrôle acoustique pour les parcs éoliens. À ce titre, on présente dans la suite un état de l'art général autour du contrôle en l'illustrant par quelques applications dans le domaine de l'éolien. On divise ainsi l'état de l'art en deux grandes parties.

Dans la partie 2.1, on présente les **techniques d'optimisation utilisées pour le calcul d'une commande optimale** des systèmes discrets. On décrit tout d'abord dans la sous-partie 2.1.1. les techniques d'optimisation pour le calcul d'une commande optimale dans le cas instantané et déterministe puis dans la sous-partie 2.1.2 le calcul dans le cas instantané en présence d'incertitudes.

Les parties 2.2 et 2.3 concernent l'**estimation du modèle** où on présente une revue des techniques d'estimation des modèles paramétriques des systèmes de contrôle à temps discret. Le partie 2.2 introduit les modèles utilisés dans le but de prédire ou d'assurer un suivi des variables d'état avec un état de l'art sur les séries temporelles dans la sous-partie 2.2.1 puis un état de l'art sur le filtrage récursif bayésien dans la sous-partie 2.2.2. La partie 2.3 constitue, quant à elle, un état de l'art sur l'apprentissage des paramètres des modèles des séries temporelles et des filtres définis précédemment dans la partie 2.2.

2.1 Commande discrète des systèmes

On s'intéresse au système de calcul ne nécessitant qu'un calcul instantané, ne tenant compte que de l'instant présent t ou de sa prédiction à très court-terme $t+1$ ou $t+2$, et pour des variables de choix discrètes. On présente ainsi dans la suite, un état de l'art sur le calcul d'une commande optimale sous ces hypothèses. Dans ce cas, le problème de contrôle s'exprime alors comme un problème d'optimisation combinatoire sous contraintes.

2.1.1 Calcul d'une commande déterministe optimale

Pour le calcul d'une commande d'un système déterministe on s'intéresse dans la suite au problème suivant :

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ f(\mathbf{x}) \\ g_j(\mathbf{x}) \leq 0, \quad \forall j \in \{1, ..., J\} \end{cases} \tag{2.1}$$

où $\mathbf{x}$ est un vecteur de taille I à valeurs dans un ensemble discret $\mathcal{X} \subset \mathbb{R}^I$, f la fonction objectif de $\mathbb{R}^I$ dans $\mathbb{R}$, et $g_j(x)$ un ensemble de J fonctions de $\mathbb{R}^I$ dans $\mathbb{R}$. Dans le cadre du calcul d'une commande optimale, les données du problème $\mathcal{X}$, f, et g_j sont entièrement déterminées par la

connaissance des variables d'état $\mathbf{y}_t$ à un instant t et dont un état de l'art est présenté dans la partie 2.2.

Dans la littérature, on considère en général une classe plus large de problèmes d'optimisation dits *Mixed-Integer Nonlinear Program* (MINLP) [D'Ambrosio & Lodi, 2011; Bonami et al., 2012; Burer & Letchford, 2012], qui s'écrivent de la manière suivante :

$$\begin{cases} \max_{\mathbf{x}_\mathrm{d}\in\mathcal{X}_\mathrm{d},\mathbf{x}_\mathrm{c}\in\mathcal{X}_\mathrm{c}} \ f(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) \\ g_j(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) \leq 0, \quad 1 \leq j \leq J \end{cases} \tag{2.2}$$

où $\mathbf{x}_\mathrm{d}$ est un vecteur de taille I_d à valeurs dans un ensemble discret $\mathcal{X}_\mathrm{d} \subset \mathbb{R}^{I_\mathrm{d}}$, $\mathbf{x}_\mathrm{c}$ est un vecteur de taille I_c à valeurs dans un ensemble continu $\mathcal{X}_\mathrm{c} \subset \mathbb{R}^{I_\mathrm{c}}$, f la fonction objectif de $\mathbb{R}^{I_\mathrm{c}} \times \mathbb{R}^{I_\mathrm{d}}$ dans $\mathbb{R}$, et $g_j(x)$ un ensemble de j fonctions de $\mathbb{R}^{I_\mathrm{c}} \times \mathbb{R}^{I_\mathrm{d}}$ dans $\mathbb{R}$. On note g la fonction telle que :

$$g(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) = \begin{bmatrix} g_1(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) \\ g_2(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) \\ \cdots \\ g_J(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}) \end{bmatrix}. \tag{2.3}$$

On distingue généralement quatre sous-catégories de problèmes :

1. *Linear Program* (LP) : les fonctions f et g sont linéaires, et toutes les variables de choix sont continues ($I_\mathrm{d} = 0$).
2. *Nonlinear Program* (NLP) : les fonctions f ou g ne sont pas linéaires, et toutes les variables de choix sont continues ($I_\mathrm{d} = 0$).
3. *Mixed Integer Linear Program* (MILP) : les fonctions f et g sont linéaires et les variables de choix se divisent en variables continues et discrètes ($I_\mathrm{d} > 0$, $I_\mathrm{c} \geq 0$).
4. *Mixed Integer Nonlinear Program* (MINLP) : problème le plus général où f ou g ne sont pas linéaires et les variables de choix se divisent en variables continues et discrètes ($I_\mathrm{d} > 0$, $I_\mathrm{c} \geq 0$).

Ces problèmes se divisent eux-mêmes en **problèmes convexes** et en **problèmes non convexes**.

Solutions admissibles et convexité

Dans le domaine de l'optimisation sous contrainte, on appelle *solutions admissibles* les valeurs de $\mathbf{x}_\mathrm{c}, \mathbf{x}_\mathrm{d}$ qui respectent les contraintes du problème posé comme illustré dans la Figure 2.1. Par opposition, on appellera dans la suite *solutions possibles* l'ensemble des valeurs pouvant être prises par les variables $\mathbf{x}_\mathrm{c}$ et $\mathbf{x}_\mathrm{d}$. Pour finir on appelle, si elle existe, *solution optimale*, la solution globale du problème d'optimisation.

Notons que l'existence et l'unicité de la solution ne sont pas toujours avérées, cependant un certain nombre de théorèmes classiques de mathématiques permettent bien souvent de les déterminer. La compacité de l'espace des solutions admissibles (c'est-à-dire le fait que cet espace est fermé et borné en dimension finie) et la continuité de la fonction objectif f, permettent d'affirmer qu'un maximum global existe par le théorème des bornes de Weierstrass [Lange, 2004, p.105]. D'autre part lorsque le problème est convexe, c'est-à-dire que l'ensemble des solutions admissible est convexe et que la fonction à maximiser est concave (ou que la fonction à minimiser est convexe) alors un maximum global existe [Lange, 2004, p.105]. Pour finir, la dérivabilité des fonctions f et g permet de définir des conditions suffisantes et nécessaires pour l'existence des maxima globaux et locaux. Les conditions de *Karush-Kuhn-Tucker* [Boyd & Vandenberghe, 2009, p.243] donnent ainsi des conditions sur le *lagrangien* du problème d'optimisation.

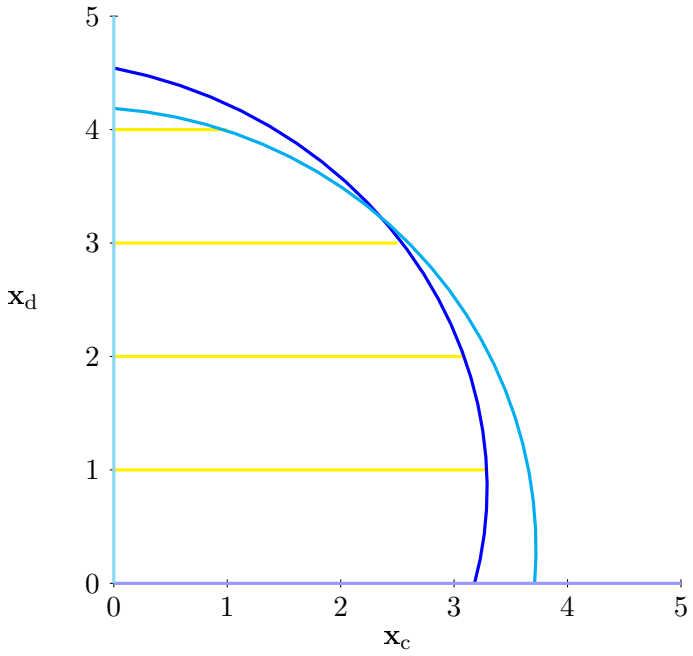


FIGURE 2.1 – Représentation graphique de l'espace des solutions d'un problème MINLP convexe. Les courbes en bleu représentent les contraintes, et les segments colorés en jaune les solutions admissibles.

Cas continu

Dans le cas continu et linéaire, le problème (2.2) devient

$$\begin{cases} \max_{\mathbf{x}_\mathrm{c} \in \mathcal{X}_\mathrm{c}} \ \mathbf{f}^T \mathbf{x}_\mathrm{c} \\ \mathbf{G}\mathbf{x}_\mathrm{c} \leq 0 \\ \mathbf{x}_\mathrm{c} \geq 0 \end{cases} \tag{2.4}$$

où $\mathbf{f}$ est un vecteur et $\mathbf{G}$ une matrice. Dans ce cas, l'ensemble des solutions admissibles est convexe comme illustré dans la Figure 2.2.

Si ce problème peut être résolu par les méthodes du gradient projeté [Goldstein, 1964; Levitin & Polyak, 1966] ou des multiplicateurs de Lagrange [Bertsekas, 2014], il existe toutefois quelques algorithmes dédiés. Tout d'abord, on peut citer l'algorithme du simplexe [Dantzig et al., 1955; Padberg, 2013] qui permet d'obtenir astucieusement la solution en parcourant l'ensemble des solutions admissibles en suivant une projection du gradient sur une intersection de $n-1$ hyperplans. Cette méthode était une des méthodes les plus populaires, mais elle est aujourd'hui concurrencée par les algorithmes de points intérieurs [Wright, 2005].

Pour les problèmes d'optimisation continus non-linéaires sous contrainte, il est difficile d'être exhaustif car de nombreuses méthodes existent. On peut néanmoins citer la méthode des points intérieurs, la programmation quadratique séquentielle [Bonnans, 1994], la méthode *active set* [Lawson & Hanson, 1995], les méthodes à régions de confiance [Coleman & Li, 1996] ou encore la méthode du gradient conjugué [Goldfarb & Lapidus, 1968].

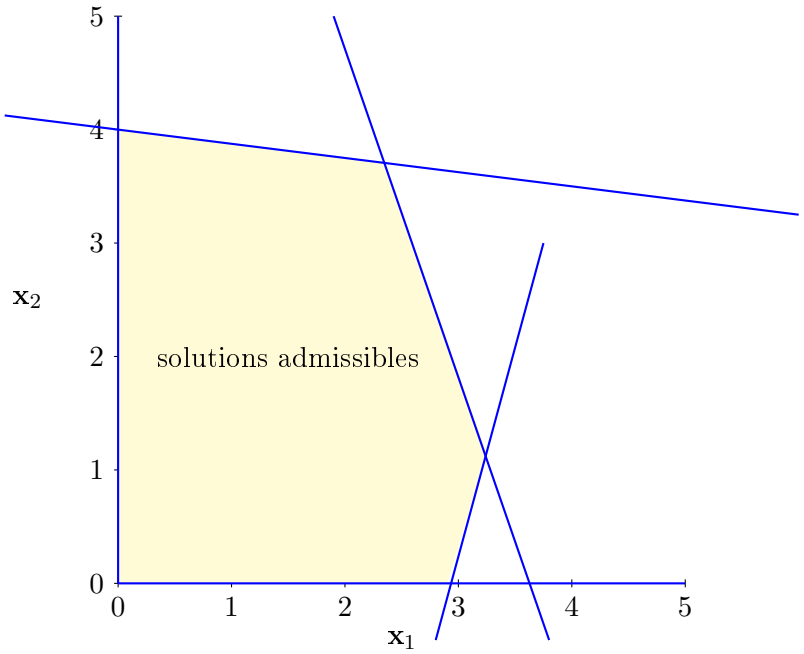


FIGURE 2.2 – Représentation graphique de l'espace des solutions d'un problème LP. Les courbes en bleu représentent les contraintes, et la surface colorée les solutions admissibles.

Cas discret convexe

Contrairement au cas continu, il n'est en général pas possible de résoudre les problèmes combinatoires sans énumération des solutions, notamment parce qu'il n'est pas possible de se reposer directement sur le calcul différentiel. Les problèmes MINLP sont d'ailleurs *NP-complets* [Bonami et al., 2012] ce qui implique qu'il n'y a pas de garantie à ce jour qu'un algorithme puisse résoudre un problème de manière optimale en temps polynomial, même si en pratique cela peut-être le cas.

Différents algorithmes sont décrits dans la littérature scientifique tels que les algorithmes *branch and bound* [Dakin, 1965; Gupta & Ravindran, 1985], *outer-approximation* [Duran & Grossmann, 1986], *generalized benders decomposition* [Geoffrion, 1972], *extended cutting plane* [Quesada & Grossmann, 1992], ou bien des méthodes hybrides [Bonami et al., 2008; Abhishek et al., 2010]. Le point commun de ces algorithmes est d'utiliser des **relaxations**, sous-problèmes plus simples construits itérativement qui, une fois résolus, permettent d'établir des bornes inférieures et supérieures des solutions. Chaque relaxation permet ainsi d'éliminer une partie des solutions et de diminuer le nombre de possibilités jusqu'à obtenir une solution optimale. Les relaxations transforment le problème principal en changeant les contraintes, en transformant les variables discrètes en variables continues, ou bien en transformant complètement le problème (fonction objectif, contraintes).

L'algorithme *branch and bound* [Dakin, 1965; Gupta & Ravindran, 1985] est certainement l'approche la plus générale. Il se base sur l'utilisation d'un arbre de recherche qui partitionne [5] successivement l'ensemble des solutions admissibles en sous-ensembles plus petits (au sens de l'inclusion). Par exemple, on peut le construire de sorte que, pour chaque nœud père, ses deux nœuds fils partitionnent le sous-ensemble qu'il représente en deux sous-espaces définis de part

5. De manière plus générale, on peut également utiliser un recouvrement.

et d'autre d'un hyperplan. Ce principe est illustré par la Figure 2.3, où l'espace des solutions admissibles est partitionné successivement deux fois, à l'aide de deux hyperplans. D'autre, part,

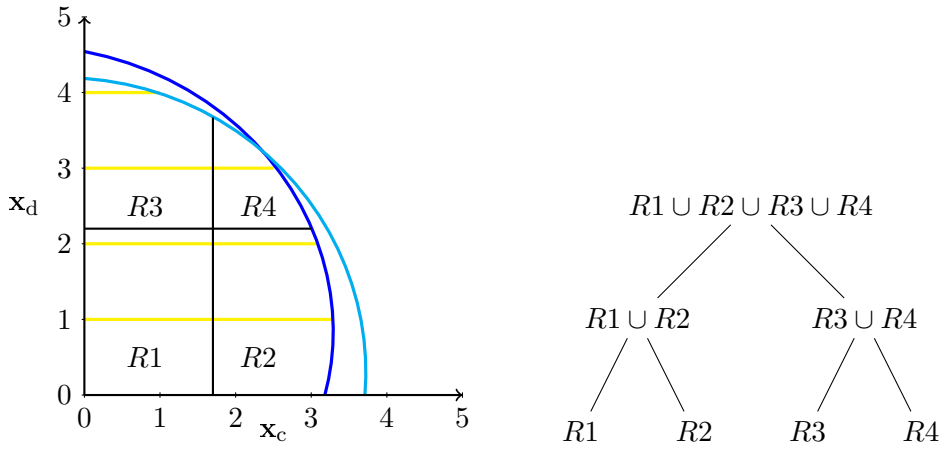


FIGURE 2.3 – Représentation graphique du partitionnement d'un algorithme *branch and bound*. L'espace des solutions admissibles est partitionné en deux à l'aide d'un hyperplan, puis en quatre à l'aide d'un second hyperplan. L'arbre des solutions représente ces deux divisions successives.

on attribue à chaque nœud un majorant de la fonction objectif des solutions qu'il contient. Typiquement, ce majorant est obtenu en transformant le problème discret en problème continu et en restreignant le sous-ensemble admissible à l'ensemble défini par le nœud. En reprenant la Figure 2.3, le majorant du nœud définissant l'ensemble $R1$, se calculerait en résolvant

$$\max_{(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d})\in R1} f(\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d}). \tag{2.5}$$

Ensuite, on parcourt à chaque itération l'arbre des solutions intelligemment en comparant à chaque nœud la valeur de la solution actuelle au majorant du nœud de l'arbre et en utilisant la propriété suivante :

Propriété d'élagage de l'algorithme *branch and bound* : Pour une solution admissible $\mathbf{x}^*$ de valeur $f(\mathbf{x}^*)$, il est impossible de trouver une solution $\mathbf{x}$ de valeur $f(\mathbf{x})$ et telle que $f(\mathbf{x}) > f(\mathbf{x}^*)$ dans les sous-ensembles dont le majorant est inférieur à $f(\mathbf{x}^*)$.

Plus précisément, si la valeur de la solution est plus grande que le majorant du nœud analysé, ce dernier est supprimé. Lors du parcours de l'arbre, seules deux possibilités existent alors :

1. Soit le nœud et sa descendance sont éliminés.
2. Soit on continue de parcourir l'arborescence de ce nœud.

Lorsqu'on atteint une nouvelle feuille de l'arbre [6], une nouvelle solution meilleure que la précédente est obtenue. On continue ensuite le parcours de l'arbre à partir de cette nouvelle valeur. L'algorithme garantit donc l'optimalité dans la mesure où, lorsqu'il termine, toutes les solutions ont été soit parcourues, soit éliminées. La Figure 2.4 représente graphiquement le parcours de l'arbre en reprenant l'exemple de la Figure 2.3. La difficulté de mise en œuvre d'un tel algorithme se situe à la fois dans le découpage de l'espace des solutions, dans le parcours de l'arbre

6. Nœud de profondeur maximale, définissant une solution.

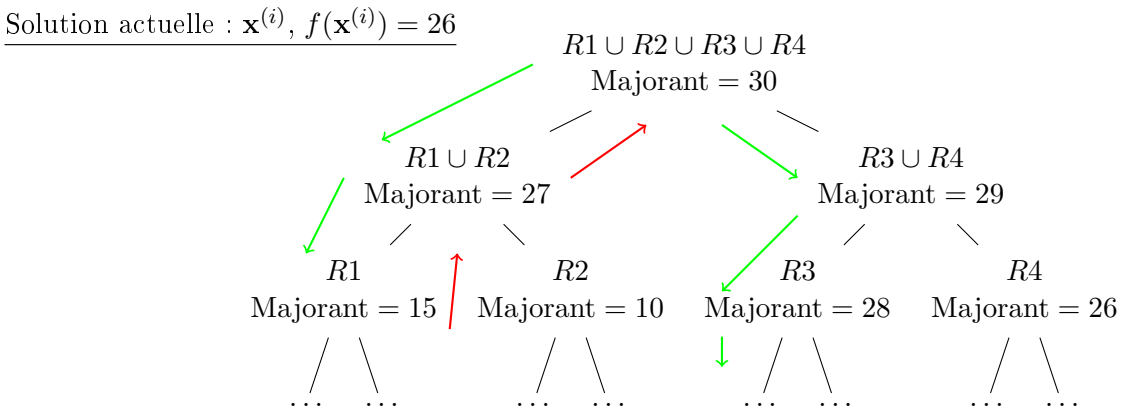


FIGURE 2.4 – Représentation graphique du parcours d'un algorithme *branch and bound*. En début de parcours, on suppose qu'on connaît une solution non-optimale du problème qui donne une valeur de 26 pour la fonction objectif. Tant que le majorant est plus grand que cette valeur, on continue de s'enfoncer dans l'arborescence de l'arbre (flèches vertes). Lorsque le majorant est inférieur, on supprime le nœud courant et on remonte au nœud père (flèches rouges). Lorsqu'on atteindra une feuille de l'arbre (solution admissible), la valeur de la fonction f sera ainsi nécessairement supérieure.

de recherche que génère le découpage, et dans la finesse du majorant qui permet l'élimination des sous-ensembles ne contenant pas de meilleure solution. En effet, une majoration trop grossière génère un parcours de l'arbre plus conséquent, mais une majoration fine requiert un temps de calcul supérieur pour attribuer un majorant pour chaque nœud.

Les autres algorithmes utilisés classiquement sont des méthodes itératives qui n'utilisent pas d'arbres de décision, mais un découpage de l'espace des solutions et des relaxations du problème. L'algorithme *outer-approximation* [Duran & Grossmann, 1986] considère à chaque itération, un demi-espace défini par un hyperplan lui-même défini comme l'hyperplan orthogonal au gradient du point actuel :

$$g(\mathbf{x}_\mathrm{c}^{(i)}, \mathbf{x}_\mathrm{d}^{(i)}) + \nabla g(\mathbf{x}_\mathrm{c}^{(i)}, \mathbf{x}_\mathrm{d}^{(i)}) \begin{pmatrix} \mathbf{x}_\mathrm{c} - \mathbf{x}_\mathrm{c}^{(i)} \\ \mathbf{x}_\mathrm{d} - \mathbf{x}_\mathrm{d}^{(i)} \end{pmatrix} \leq 0 \tag{2.6}$$

où $(\mathbf{x}_\mathrm{d}^{(i)}, \mathbf{x}_\mathrm{d}^{(i)})$ est la solution à l'itération i.

On construit alors à partir de ces deux sous-espaces deux problèmes de relaxation : l'équation (2.6) permet de définir un problème MILP, dont la solution est un minorant de la solution du problème MINLP, et un problème NLP, obtenu en fixant les variables discrètes $\mathbf{x}_\mathrm{d}$ à la solution courante dont la solution est un majorant du problème MINLP. Le premier problème permet de calculer une nouvelle solution et le second permet de définir un critère d'arrêt de l'algorithme : l'algorithme s'arrête lorsque les deux bornes (les solutions des deux problèmes) sont plus petites qu'un seuil à définir. L'algorithme *generalized benders decomposition* [Geoffrion, 1972] est relativement similaire, mais la construction des relaxations diffère légèrement. En fixant les variables discrètes $\mathbf{x}_\mathrm{d}$, on obtient d'abord un problème NLP sur les variables $\mathbf{x}_\mathrm{c}$, puis on inverse en fixant les variables continues $\mathbf{x}_\mathrm{c}$ avec la solution du problème précédent et, à l'aide d'une projection on obtient un problème MILP. Cette alternance est itérée jusqu'à convergence. L'algorithme *extended cutting plane* [Westerlund & Pettersson, 1995] fonctionne quant à lui en utilisant une suite de relaxations dont l'ensemble des solutions admissibles inclus celui du problème principal. La solution de chaque relaxation viole par conséquent une partie des contraintes. À partir de la

contrainte la plus distante en norme euclidienne, on construit ensuite un nouveau problème en linéarisant cette contrainte et en l'ajoutant à la relaxation précédente. L'algorithme itère de cette façon jusqu'à ce qu'une solution admissible soit obtenue. Les algorithmes *LP/NLP based branch and bound* et les algorithmes de Bonami et al. [2008] et Abhishek et al. [2010], constituent quant à eux des méthodes hybrides des algorithmes précédents.

Cas discret non convexe

Dans le cas non convexe, les algorithmes présentés jusqu'ici peuvent servir d'heuristiques mais ne garantissent pas l'optimalité du problème. Notamment, la non-convexité des contraintes rend difficile la relaxation du problème. Dans le cas de l'algorithme *branch and bound*, la non-convexité du problème entraîne en général la non-convexité des relaxations. Or la plupart des algorithmes d'optimisation sous contraintes non-linéaires dans le cas continu ne garantissent qu'une solution locale en général. Si on obtient des solutions non globales pour ces problèmes, les nœuds ne sont plus correctement majorés et présentent le risque d'être supprimés à tort.

On retrouve le même type de problème avec les algorithmes *generalized benders decomposition*, *outer-approximation* et *extended cutting plane*. En effet, comme illustré dans la Figure 2.5, la convexité des solutions admissibles permet d'inclure l'ensemble des solutions admissibles dans un sous-ensemble plus grand au sens de l'inclusion, lors de la linéarisation des contraintes. Dans le cas non convexe, une partie des solutions admissibles peut se retrouver en dehors de cet ensemble, rendant ainsi caduques les garanties d'optimalité de ces algorithmes.

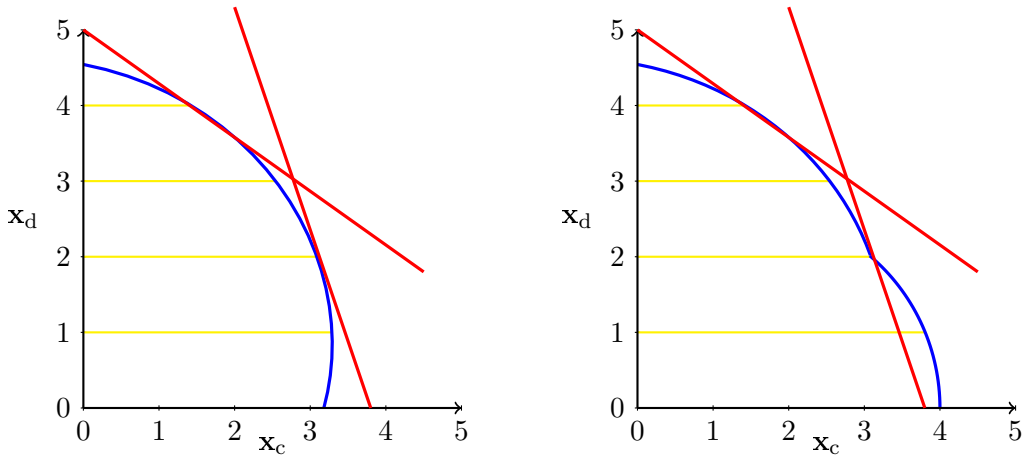


FIGURE 2.5 – Représentation graphique de l'utilisation du gradient pour construire une enveloppe convexe à contraintes linéaires. Lors de l'utilisation d'hyperplans perpendiculaires au gradient de g pour construire une enveloppe convexe (droites rouges), une partie des solutions risque d'être « oubliée » dans le cas non convexe.

Une première stratégie consiste à reformuler le problème en problème convexe, ceci n'est malheureusement valable que pour certains types de problèmes *ad hoc*. La stratégie plus générale consiste à trouver une enveloppe convexe au problème d'optimisation. Pour ce faire, on transforme généralement le problème à l'aide de l'astuce suivante, permettant de faire « rentrer » la fonction

objectif dans l'ensemble des contraintes :

$$\begin{cases} \max_{(\mathbf{z},\mathbf{x}_\mathrm{c},\mathbf{x}_\mathrm{d})} \ \mathbf{z} \\ \mathbf{z} - f(\mathbf{x}_\mathrm{c}, \mathbf{x}_\mathrm{d}) \leq 0 \\ g(\mathbf{x}_\mathrm{c}, \mathbf{x}_\mathrm{d}) \leq 0 \end{cases} \tag{2.7}$$

où $\mathbf{z}$ est une variable auxiliaire.

En trouvant des fonctions f_conv et g_conv convexes qui majorent respectivement f et g, il est alors parfois possible de résoudre le problème sans difficulté. En effet, lorsqu'un optimum global d'une enveloppe convexe est trouvé, si cette solution appartient à l'ensemble non convexe, elle est alors optimale pour celui-ci. Si on ne trouve pas de solution par cette méthode, il est alors nécessaire de raffiner le problème en utilisant, un algorithme *branch and bound* afin de diviser l'ensemble des solutions admissibles et de trouver des enveloppes convexes plus fines. Différents algorithmes ont été proposés à ce sujet, construisant des solutions optimales d'un coût potentiellement élevé : *spatial branch and bound* [Smith & Pantelides, 1999], *branch and reduce* [Ryoo & Sahinidis, 1996], *alpha-branch and bound* [Androulakis et al., 1995], *branch and cut* [Kesavan & Barton, 2000].

2.1.2 Calcul d'une commande optimale sous incertitudes

Lorsque les variables $\mathbf{x}$, f et g du problème d'optimisation sont incertaines, on modélise l'incertitude par l'ajout de variables aléatoires. Dans le cadre d'une optimisation robuste, on considère le problème *chance constraint programing* [Li et al., 2008] :

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ \mathbb{E}(f(\mathbf{x})) \\ p(g(\mathbf{x}, \mathbf{w}) \leq \mathbf{0}) \geq 1 - \varepsilon \end{cases}$$

où $\mathbf{w}$ est vecteur aléatoire de loi connue et ε est un paramètre de confiance que l'on se donne sur le respect des contraintes. En effet, il n'est en général pas possible d'assurer le respect des contraintes à 100% et le paramètre ε présente alors le taux acceptable de violation de celle-ci. Cette formulation est utilisée en finance et en ingénierie et est naturelle dans la mesure où il s'agit de maximiser un critère tout en assurant une fiabilité donnée pour le respect des contraintes. La stratégie générale pour résoudre ces problèmes est la transformation ou la relaxation du problème en un problème déterministe, potentiellement équivalent, mais qui est généralement non-linéaire. Même lorsque le problème de départ est linéaire, la transformation du problème ne l'est généralement pas [Li et al., 2008]. Cependant, le cas où le problème s'écrit

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ \mathbb{E}(f(\mathbf{x})) \\ p(g_j(\mathbf{x}, \mathbf{w}) \leq 0) \geq 1 - \varepsilon_j, \ \forall j \in \{0, ..., J\}, \end{cases} \tag{2.8}$$

la résolution est souvent plus simple. Cette formulation est appelée *single chance constraint programing* par opposition au problème (2.8) appelé *joint chance constraint programing*. Les problèmes *single chance constraint programing* ne prennent pas en compte les corrélations et s'avèrent en général plus conservatifs.

Pour transformer ce problème en problème déterministe, on utilise généralement la fonction de répartition. Dans le cas linéaire gaussien par exemple, le problème *single chance constraint*

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ \mathbb{E}(f(\mathbf{x})) \\ p\left((\mathbf{g}_{\mathrm{x},j}^T \mathbf{x} + \mathbf{g}_{\mathrm{w},j}^T \mathbf{w}) \leq 0\right) \geq 1 - \varepsilon_j, \ \forall j \in \{0, ..., J\} \end{cases} \tag{2.9}$$

où $\mathbf{w}$ suit une loi gaussienne de moyenne $\mathbf{m_w}$ et de variance $\mathbf{\Sigma_w}$, peut se transformer en

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ \mathbb{E}(f(\mathbf{x})) \\ \mathbf{g}_{\mathrm{x},j}\mathbf{x} \leq \Phi^{-1}_{\mathbf{g}_{\mathrm{w},j}\mathbf{w}}(\varepsilon), \ \forall j \in \{0,...,J\} \end{cases} \tag{2.10}$$

où $\Phi_{g_{\mathrm{w},j}\mathbf{w}}$ désigne la fonction de répartition de $\mathbf{g}_{\mathrm{w},j}\mathbf{w}$. En effet, la loi marginale du terme $\mathbf{g}_{\mathrm{w},j}\mathbf{w}$ est une loi gaussienne (combinaison linéaires de variables aléatoires gaussiennes). De ce fait, pour une valeur ε_j donnée, l'inégalité de contrainte peut se transformer de manière déterministe en une autre inégalité en utilisant la fonction de répartition de $g_{\mathrm{w},j}\mathbf{w}$ qui donne alors directement une marge à respecter. Dans ce cas, le problème obtenu est linéaire par rapport à la variable $\mathbf{x}$, et peut-être résolu de manière similaire au problème déterministe. En revanche, dans le cas *joint chance constraint*, ce problème devient non-linéaire. En effet, même dans le cas où les variables incertaines sont indépendantes entre elles, le problème s'écrit

$$\begin{cases} \max_{\mathbf{x}\in\mathcal{X}} \ \mathbb{E}(f(\mathbf{x})) \\ \prod_{j=1}^{N} \Phi_{g_{\mathrm{w},j}\mathbf{w}}(g_{\mathrm{x},j}\mathbf{x}) \leq \varepsilon \end{cases} \tag{2.11}$$

et dans ce cas le problème se complexifie et doit être résolu par un algorithme non-linéaire.

Dans le cas plus général où les variables sont corrélées, une intégration multivariée est nécessaire, ce qui s'avère difficile à calculer en pratique. On peut citer la méthode de Li et al. [2002] qui utilise un schéma numérique d'intégration dans le cas gaussien multivarié. On peut également citer la méthode de Li et al. [2004] qui consiste à trouver les valeurs de ε telles que le problème soit faisable.

Dans le cas d'un problème non-linéaire, les problèmes *chance constraint* restent à ce jour des problème complexes et non résolus [Li et al., 2008]. La stratégie consiste ici encore à se ramener à un problème déterministe. Les méthodes de Wendt et al. [2002] construisent une relation entre les variables aléatoires d'entrée (w) et leur effet sur les variables en sortie ($g(\mathbf{x},\mathbf{w})$). D'autres méthodes utilisent quant à elles, une transformation du problème en un problème déterministe de grande taille, basé sur un échantillonnage de Monte-Carlo [Ahmed & Shapiro, 2008; Pagnoncelli et al., 2009].

2.2 Estimation d'un modèle paramétrique pour du contrôle à temps discret

Le système industriel à l'origine de la thèse est un système basé sur des pas de temps discrets et on présente à ce sujet un état de l'art sur l'estimation des modèles paramétriques.

Dans la suite, on suppose un pas de temps discret, noté t, et un vecteur d'état noté $\mathbf{y}_t$ de taille n, contenant l'ensemble des variables d'intérêt décrivant l'état du système évoluant au cours du temps. On suppose également que les variables d'état contenues dans $\mathbf{y}_t$ sont incertaines et sont de ce fait dépendantes à la fois de toutes les variables d'état déjà passées, des vecteurs de commande de contrôle $\mathbf{c}_1,...,\mathbf{c}_t$, et des incertitudes qui sont représentées par des variables aléatoires notées $\mathbf{w}_{\mathbf{y}_1},...,\mathbf{w}_{\mathbf{y}_t}$. La dépendance complète est modélisée par une fonction f, qui vérifie

$$\mathbf{y}_{t+1} = f(\mathbf{y}_{1:t}, \mathbf{c}_{1:t+1}, \mathbf{w}_{\mathbf{y}_1:\mathbf{y}_t}, \boldsymbol{\theta}_\mathbf{y}) \tag{2.12}$$

où $\boldsymbol{\theta}_\mathbf{y} = \{\theta_{y,1},...,\theta_{y,K}\}$ est un ensemble fini de K paramètres. Pour une suite de vecteurs quelconque $\mathbf{x}_1$,...,$\mathbf{x}_t$,...,$\mathbf{x}_T$, la notation $\mathbf{x}_{1:t}$ désigne, ici et dans la suite, l'ensemble des variables $\mathbf{x}_i$ pour $i \in \{1,...,t\}$. L'équation (2.12), très générale, est nommée **équation d'état** et modélise entièrement l'évolution d'un système à temps discret. Par exemple, $\mathbf{y}_t$ peut contenir les variables

de position géographique d'une voiture sans conducteur et la fonction f décrit alors un modèle physique régissant l'évolution de la voiture en fonction de la commande d'accélération et de l'aléa lié à la qualité du sol, l'absorption de l'air, la fiabilité du moteur, *etc.*

Dans la plupart des applications, le vecteur $\mathbf{y}_t$ est non accessible et souvent seul un vecteur $\mathbf{z}_t$ de taille m, dépendant du vecteur d'état l'est. Pour reprendre l'exemple précédent, il peut s'agir de la position radar ou GPS de la voiture, ou bien les deux à la fois (concaténées dans ce vecteur) qui indiquent indirectement sa position. Cette observation est généralement entachée d'une incertitude $\mathbf{w}_{\mathbf{z}_t}$ indépendante de $\mathbf{w}_{\mathbf{y}_t}$ qui lui est propre. Toujours dans l'exemple présenté, il peut s'agir de l'erreur intrinsèque aléatoire au système GPS, ou à la détection radar.

Pour modéliser cette autre dépendance, on introduit une fonction h qui vérifie

$$\mathbf{z}_t = h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\theta}_\mathbf{z}) \tag{2.13}$$

où $\boldsymbol{\theta}_\mathbf{z} = \{\theta_{z,1}, ..., \theta_{z,M}\}$ est un ensemble fini de M paramètres. L'équation (2.13) est appelée **équation d'observation**. Les équations (2.12) et (2.13) définissent un système de contrôle, appelé **représentation d'état** [Kailath et al., 2000].

Afin de pouvoir effectivement contrôler le système, il apparaît en premier lieu nécessaire d'estimer les variables d'état au cours du temps et les paramètres du système. On différencie dans la suite trois problèmes d'estimation :

- Problème 1 - L'inférence de $\mathbf{y}_t$ dans le cas où le vecteur d'état est directement observé. On ne considère ici que l'équation (2.12). On parle ici d'un problème de prédiction dit à un horizon h où on cherche à estimer la probabilité d'un état futur $\mathbf{y}_{t+h}$ à partir des états passés $\mathbf{y}_{1:t}$. On utilise généralement pour ce type de problème des séries temporelles [Brockwell & Davis, 2016], ou des réseaux de neurones [Odom & Sharda, 1990].
- Problème 2 - L'inférence de $\mathbf{y}_t$ avec les équations (2.12) et (2.13). Le vecteur $\mathbf{y}_t$ n'est pas directement observé. Il s'agit du problème d'estimation de la probabilité $p(\mathbf{y}_t|\mathbf{z}_{1:t})$ appelée aussi inférence ou filtrage récursif bayésien [Bergman, 1999]. Concrètement, il s'agit de faire du suivi ou *tracking* en utilisant les observations passées $\mathbf{z}_{1:t}$ pour estimer au « mieux » le vecteur $\mathbf{y}_t$.
- Problème 3 - Le problème d'estimation des paramètres du système, $\boldsymbol{\theta}_\mathbf{y}$ et $\boldsymbol{\theta}_\mathbf{z}$ [Ljung, 1998], appelé apprentissage du modèle et qui nécessite généralement la capacité à résoudre au préalable l'un ou l'autre des deux problèmes précédents.

2.2.1 Modèles temps-série et réseaux de neurones

Dans le cas de prédiction de $\mathbf{y}_t$, on utilise le plus souvent des modèles ARMA (*autoregressive moving average*) ou ARIMA (*autoregressive integrated moving average*) [Brockwell & Davis, 2016], qui s'écrivent de la manière suivante :

$$\mathbf{y}_t = \sum_{i=1}^{i_\alpha} \alpha_i \mathbf{y}_{t-i} + \sum_{i=0}^{i_\beta} \beta_i \mathbf{w}_{t-i} \tag{2.14}$$

où les lois des $\mathbf{w}_t$ sont supposées indépendantes et identiquement distribuées et généralement de loi gaussienne centrée. Les polynômes formels $(\alpha_i)_{i\geq 0}$ et $(\beta_i)_{i\geq 0}$ déterminent entièrement le modèle et constituent respectivement les paramètres auto-régressifs (AR) et de moyenne glissante (MA). Ils constituent les paramètres à apprendre.

Les méthodes de prédiction sont généralement comparées au prédicteur dit de *persistance*, qui consiste à prédire les valeurs futures comme égales à la dernière valeur mesurée, c'est-à-dire

$$\hat{\mathbf{y}}_{t+h} = \mathbf{y}_t, \tag{2.15}$$

où $\hat{\mathbf{y}}_{t+h}$ est la prédiction à l'horizon h de $\mathbf{y}_t$.

Dans le contexte de l'éolien, de tels modèles on été utilisés pour la prédiction de la vitesse du vent à court-terme [Chou & Corotis, 1981; Schwartz & Milligan, 2002; Costa et al., 2008; Giebel et al., 2011]. On peut citer notamment la méthode de Brown et al. [1984] qui constitue une des premières tentatives de prédiction par un modèle ARMA, après avoir transformé les données de vent pour obtenir des distributions gaussiennes pour les vitesses de vent classées par heure de la journée. Au début des années 90, de nombreuses méthodes similaires ont été proposées [Daniel & Chen, 1991; Kamal & Jafri, 1997], ainsi que des modèles basés sur des réseaux de neurones [Kariniotakis et al., 1996; Mohandes et al., 1998; Alexiadis et al., 1999]. Des résultats relativement similaires ont été obtenus pour les deux familles de modèles comme décrit dans l'article de Tang et al. [1991]. Les auteurs ont démontré à cette époque que les résultats obtenus par des réseaux de neurones et des modèles ARIMA à historique long (beaucoup de coefficients) étaient relativement similaires, mais que les réseaux de neurones ne s'avéraient plus efficaces que pour des modèles à historique limité. D'autre part, ils ont conclu que les réseaux de neurones, moyennant une topologie bien spécifique, pouvaient être une alternative viable et robuste aux modèles ARIMA.

Plus récemment, des évolutions des modèles ARIMA ont été proposés pour ce type d'application, avec par exemple la méthode de Kavasseri & Seetharaman [2009] qui utilise des modèles ARIMA fractionnels, ou encore la méthode de Schwartz & Milligan [2002]. Ces méthodes ont permis finalement de dépasser l'estimateur de *persistance*.

Pour finir, notons que, une approche plus récente et moins usuelle concerne les réseaux de neurones afin de faire de la prédiction à court terme. On ne s'attardera pas sur le sujet et un bon état de l'art sur la prédiction à court-terme de la vitesse du vent à partir de réseaux de neurones peut se trouver en [Giebel et al., 2011].

2.2.2 Inférence récursive bayésienne

Estimateur du maximum de vraisemblance et estimation bayésienne

En statistique, la méthode du maximum de vraisemblance permet d'estimer des paramètres $\boldsymbol{\theta}$ de la loi d'une variable aléatoire $\mathbf{z}$, à partir de N observations $\mathbf{z}_i$, $i \in \{1, ..., N\}$ en maximisant la fonction de vraisemblance [Bishop, 2007] $\mathcal{L}(\boldsymbol{\theta}|\mathbf{z}_1, ..., \mathbf{z}_N)$ définie par

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{z}_1, ..., \mathbf{z}_N) = p(\mathbf{z}_1, ..., \mathbf{z}_N|\boldsymbol{\theta}) \tag{2.16}$$

où $p(\cdot)$ désigne une mesure de probabilité de loi paramétrique connue. Cette méthode d'estimation est convergente et asymptotiquement efficace, mais peut être biaisée [Wasserman, 2013].

Par opposition, l'estimation statistique bayésienne [Berger, 2013] est une méthode d'inférence statistique qui consiste à pondérer l'estimation du maximum de vraisemblance par une connaissance probabiliste *a priori* sur le paramètre à estimer par le biais de la loi de Bayes [Bayes, 1763] :

$$p(\boldsymbol{\theta}|\mathbf{z}_1, ..., \mathbf{z}_N) = \frac{p(\mathbf{z}_1, ..., \mathbf{z}_N|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{z}_1, ..., \mathbf{z}_N)}. \tag{2.17}$$

On appelle la probabilité $p(\boldsymbol{\theta}|\mathbf{z}_1, ..., \mathbf{z}_N)$ loi *a posteriori* (après observation) qu'on oppose ainsi à la loi *a priori* $p(\boldsymbol{\theta})$ (avant observation). La loi *a posteriori* se calcule donc comme le produit de la vraisemblance et de la loi *a priori* divisée par $p(\mathbf{z}_1, ..., \mathbf{z}_N)$. En pratique, le terme $p(\mathbf{z}_1, ..., \mathbf{z}_N)$ n'est qu'une valeur de normalisation puisque lors de l'estimation on observe une seule réalisation de ces variables. Cette valeur de normalisation est calculée en marginalisant la loi jointe $p(\mathbf{z}_1, ..., \mathbf{z}_N, \boldsymbol{\theta})$

sur $\boldsymbol{\theta}$:

$$p(\mathbf{y}_1, ..., \mathbf{y}_N) = \int p(\mathbf{z}_1, ..., \mathbf{z}_N, \boldsymbol{\theta}) d\boldsymbol{\theta} \tag{2.18}$$

$$= \int p(\mathbf{z}_1, ..., \mathbf{z}_N | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2.19}$$

Notons alors que dans cette méthode d'inférence, les paramètres $\boldsymbol{\theta}$ sont eux-même considérés comme des variables aléatoires. Rien n'empêche alors dans l'équation de remplacer $\boldsymbol{\theta}$ par une variable aléatoire $\mathbf{y}$ ce qui permet de voir cette méthode comme une manière d'exploiter l'observation d'un vecteur aléatoire $\mathbf{z}$ pour améliorer les connaissances d'un autre vecteur $\mathbf{y}$, non indépendant.

Filtrage bayésien récursif

La remarque précédente illustre parfaitement le principe du filtrage bayésien récursif [Bergman, 1999] qui applique ce procédé d'estimation récursivement sur chaque nouvelle observation en utilisant le résultat de l'estimation précédente pour calculer une nouvelle loi *a priori*.

En considérant les équations (2.12) et (2.13), en supposant que l'on connaît une loi *a priori* sur $\mathbf{y}_0$ et sous une hypothèse de chaîne de Markov [Stewart, 1994] sur l'équation d'état, c'est-à-dire

$$\forall t \; p(\mathbf{y}_t | \mathbf{y}_{1:t-1}) = p(\mathbf{y}_t | \mathbf{y}_{t-1}), \tag{2.20}$$

l'équation (2.12) devient

$$\mathbf{y}_t = f(\mathbf{y}_{t-1}, \mathbf{c}_{t-1}, \mathbf{w}_{\mathbf{y}_{t-1}}, \boldsymbol{\theta}_{\mathbf{y}}). \tag{2.21}$$

On peut ré-écrire la loi jointe $p(\mathbf{z}_{1:t}, \mathbf{y}_{1:t})$ de la manière suivante :

$$p(\mathbf{z}_{1:t}, \mathbf{y}_{1:t}) = p(\mathbf{y}_1) \prod_{i=1}^{t} p(\mathbf{z}_i | \mathbf{y}_i) \prod_{i=2}^{t} p(\mathbf{y}_i | \mathbf{y}_{i-1}). \tag{2.22}$$

À partir de cette formulation, il est possible de générer un schéma récursif d'estimation qui permet de calculer à chaque instant t la loi *a posteriori* $p(\mathbf{y}_t | \mathbf{z}_t)$ à partir de $\mathbf{z}_t$ et de $p(\mathbf{y}_{t-1} | \mathbf{z}_{t-1})$. En effet, à partir de l'équation d'état, on calcule dans un premier temps $p(\mathbf{y}_t | \mathbf{z}_{t-1})$ à partir de l'équation suivante :

$$p(\mathbf{y}_t | \mathbf{z}_{1:t-1}) = \int p(\mathbf{y}_t | \mathbf{y}_{t-1}) p(\mathbf{y}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{y}_{t-1}. \tag{2.23}$$

Cette équation constitue une prédiction de la valeur de $\mathbf{y}_t$ et joue alors le rôle de loi *a priori*. Puis, la distribution *a posteriori* est calculée à l'aide de la loi de Bayes et de la réalisation de $\mathbf{z}_t$:

$$p(\mathbf{y}_t | \mathbf{z}_{1:t}) = \frac{p(\mathbf{z}_t | \mathbf{y}_t) p(\mathbf{y}_t | \mathbf{z}_{1:t-1})}{p(\mathbf{z}_t | \mathbf{z}_{1:t-1})}. \tag{2.24}$$

De cette manière, il devient ainsi possible d'effectuer un suivi de la variable d'état $\mathbf{y}_t$ à partir des observations $\mathbf{z}_t$, à chaque instant.

Cas linéaire et gaussien : le filtre de Kalman

Le premier exemple de filtre récursif bayésien est le filtre de Kalman [Kalman et al., 1960]. Dans ce cadre on suppose la forme particulière suivante pour l'équation d'état et l'équation d'observation :

$$\begin{cases} \mathbf{y}_{t+1} = \mathbf{F}\mathbf{y}_t + \mathbf{G}\mathbf{c}_t + \mathbf{D}\mathbf{w}_{\mathbf{y}_t} \\ \mathbf{z}_t = \mathbf{H}\mathbf{y}_t + \mathbf{E}\mathbf{w}_{\mathbf{z}_t} \end{cases} \tag{2.25}$$

où $\mathbf{F}$, $\mathbf{G}$, $\mathbf{D}$, $\mathbf{H}$ et $\mathbf{E}$ sont des matrices qui définissent une forme linéaire pour les fonctions f et h. On choisit en général des lois gaussiennes multivariées pour les variables $\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_t}$ même si on est pas en pratique limité à cette loi et les équations de Kalman sont valables pour n'importe quelle distribution où la moyenne et la variance existe. En effet, les équations ont été à l'origine construites sans l'utilisation d'un cadre gaussien et de la règle de Bayes, mais comme le meilleur estimateur linéaire récursif de la moyenne et de la covariance selon un critère de moindres carrés.

Les équation de prédiction (2.23) et de mise à jour (2.24) se résument alors à un calcul de la moyenne et de la covariance *a posteriori* :

1. **Prédiction** :

$$\mathbf{m}_{\mathbf{y}_{t|t-1}} = \mathbf{F}\mathbf{m}_{\mathbf{y}_{t-1|t-1}} + \mathbf{G}\mathbf{c}_{t-1} + \mathbf{D}\mathbf{m}_{\mathbf{w}_\mathbf{y}}, \tag{2.26}$$

$$\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}} = \mathbf{F}\mathbf{\Sigma}_{\mathbf{y}_{t-1|t-1}}\mathbf{F}^T + \mathbf{D}\mathbf{\Sigma}_{\mathbf{w}_\mathbf{y}}\mathbf{D}^T. \tag{2.27}$$

2. **Mise à jour** :

$$\mathbf{m}_{\mathbf{z}_{t|t-1}} = \mathbf{H}\mathbf{m}_{\mathbf{y}_{t|t-1}} + \mathbf{E}\mathbf{m}_{\mathbf{w}_\mathbf{z}}, \tag{2.28}$$

$$\mathbf{\Sigma}_{\mathbf{z}_{t|t-1}} = \mathbf{H}\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}\mathbf{H}^T + \mathbf{E}\mathbf{\Sigma}_{\mathbf{w}_\mathbf{z}}\mathbf{E}^T, \tag{2.29}$$

$$\mathbf{K}_t = \mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}\mathbf{H}\mathbf{\Sigma}^{-1}_{\mathbf{z}_{t|t-1}}, \tag{2.30}$$

$$\mathbf{m}_{\mathbf{y}_{t|t}} = \mathbf{m}_{\mathbf{y}_{t|t-1}} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_{t|t-1}}), \tag{2.31}$$

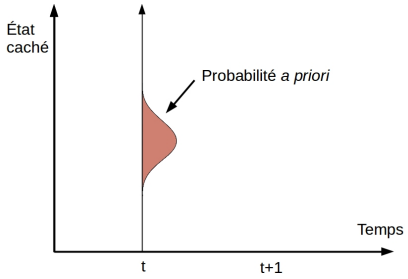
$$\mathbf{\Sigma}_{\mathbf{y}_{t|t}} = (\boldsymbol{I} - \mathbf{K}_t\mathbf{H})\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}. \tag{2.32}$$

Pour un vecteur quelconque $\mathbf{x}$ et des temps $t1$ et $t2$, $t2 \leq t1$, $\mathbf{m}_{\mathbf{x}_{t1|t2}}$ et $\mathbf{\Sigma}_{\mathbf{y}_{t1|t2}}$ désignent respectivement la moyenne et la covariance de la variable $\mathbf{x}$ au temps $t1$ sachant $\mathbf{z}_{1:t2}$.

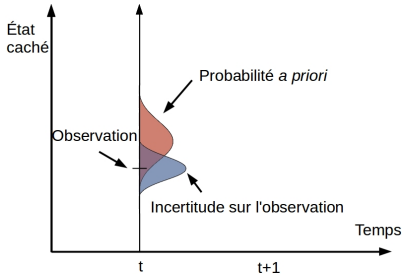
Ces équations permettent le filtrage et sont implémentées dans l'ordre. Les équations de prédiction consistent à propager la distribution *a posteriori*, $\mathbf{y}_{t-1}|\mathbf{z}_{t-1}$ obtenue lors **du filtrage précédent** à travers l'équation d'état (2.12). Dans le cas du filtrage de Kalman, l'équation (2.26) définit la propagation de la moyenne et l'équation (2.27) celle de la variance et s'obtiennent par simple calcul de moments. Les moments obtenus sont alors les moments de **la nouvelle loi** *a priori*. Les équations de mise à jour se divisent en trois parties. Tout d'abord les équations (2.28) et (2.29), similaires à (2.26) et (2.27), calculent la moyenne et la variance de $\mathbf{z}_t|\mathbf{y}_{t-1}$ (vraisemblance) en propageant la distribution à travers l'équation d'observation (2.13). Notons, que pour les quatre équations (2.26), (2.27), (2.28) et (2.29), c'est la linéarité de la représentation d'état qui permet de propager simplement les moments. L'équation (2.30), permet de calculer le gain de Kalman $\mathbf{K}_t$ afin d'effectuer la mise à jour bayésienne à proprement parler. Notons à ce sujet qu'une autre formulation équivalente peut être utilisée pour son calcul [7] :

$$K_t = \mathbf{\Sigma}_{\mathbf{y}_{t|t-1}\mathbf{z}_{t|t-1}}\mathbf{\Sigma}^{-1}_{\mathbf{z}_{t|t-1}}. \tag{2.33}$$

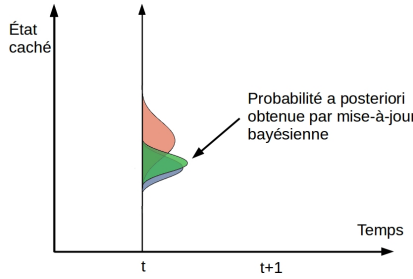
7. Cette autre formulation, un peu plus générale, est importante et permet de justifier certaines approximation dans la suite de l'état de l'art.



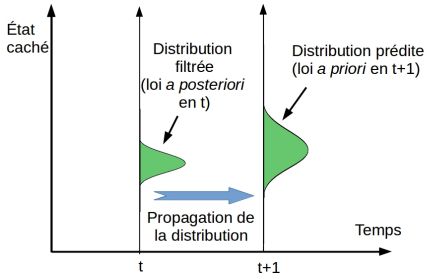
(a) **Avant mise-à-jour**. À l'instant t, on suppose que le modèle fournit une loi *a priori* pour l'état caché $\mathbf{y}_t$.



(b) **Observation**. À l'instant t, on observe une valeur pour $\mathbf{z}_t$ qui renseigne la valeur de $\mathbf{y}_t$ à une incertitude près.



(c) **Mise-à-jour**. On utilise la mise-à-jour bayésienne (équation (2.24)) pour calculer une distribution *a posteriori*.



(d) **Prédiction**. La distribution *a posteriori* est propagée via l'équation de prédiction (équation (2.23)) et servira de loi *a priori* pour l'itération suivante.

FIGURE 2.6 – Illustration du fonctionnement du filtre de Kalman.

Cette matrice peut être également vue comme la pseudo-inverse de Moore-Penrose [Ben-Israel & Greville, 2003], solution formelle du problème des moindres-carrés ordinaires dans le cas linéaire. Les équations (2.31) et (2.32) sont les équations de mise à jour à partir du gain de Kalman et de l'observation $\mathbf{z}_t$. Elles définissent la moyenne et la covariance *a posteriori* et seront utilisées lors de la prochaine itération.

Notons également que les paramètres $\boldsymbol{\theta}_\mathbf{y}$ et $\boldsymbol{\theta}_\mathbf{z}$, sont respectivement les coefficients des matrices $\mathbf{F}$, $\mathbf{G}$ et $\mathbf{D}$ et des paramètres de la loi de $\mathbf{w}_\mathbf{y}$ d'une part, et les coefficients des matrices $\mathbf{H}$ et $\mathbf{E}$ et des paramètres de la loi de $\mathbf{w}_\mathbf{z}$ d'autre part. Les paramètres des lois sont notés $\boldsymbol{m}_{\mathbf{w}_\mathbf{y}}$,$\boldsymbol{\Sigma}_{\mathbf{w}_\mathbf{y}}$, $\boldsymbol{m}_{\mathbf{w}_\mathbf{z}}$, $\boldsymbol{\Sigma}_{\mathbf{w}_\mathbf{z}}$ et sont respectivement les moyennes et matrices de covariances des lois de $\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_t}$.

Le filtre de Kalman a été utilisé intensivement dans de nombreuses applications dans l'aéronautique au sens large [Schmidt, 1981], le traitement du signal et de la parole [Paliwal & Basu, 1987], en économétrie [Bewley, 1994] ou bien même pour du monitoring dans des domaines divers (chimie [Brown, 1986], applications médicales [Smith & West, 1983]).

De ce fait, de nombreuses extensions ont été proposées afin de pouvoir gérer les problèmes de grande taille, prendre en compte les cas où les bruits ne sont pas gaussiens et où les relations entre les variables ne sont pas linéaires, comme dans le cas général, décrit par les équations (2.12) et (2.13).

Filtre de Kalman étendu

L'extension la plus naturelle est le filtre de Kalman étendu dont on peut trouver une bonne description dans le livre de Haykin et al. [2001]. Ce filtre est utilisé pour gérer le cas plus général où f et h ne sont pas linéaires et où on s'intéresse à la moyenne et à la covariance des incertitudes.

L'idée est de linéariser les fonctions f et h par un développement de Taylor à l'ordre 1 autour des valeurs moyennes des différentes variables :

$$\begin{aligned} f(\mathbf{y}_t, \mathbf{c}_t, \mathbf{w}_{\mathbf{y}_t}, \boldsymbol{\theta}_\mathbf{y}) =& f(\mathbf{m}_{\mathbf{y}_{t-1|t-1}}, \mathbf{c}_t, \mathbf{m}_{\mathbf{w}_{\mathbf{y}_{t-1|t-1}}}, \boldsymbol{\theta}_\mathbf{y}) \\ &+ \left.\frac{\partial f}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{m}_{\mathbf{y}_{t-1|t-1}}} (\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t-1|t-1}}) \\ &+ \left.\frac{\partial f}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{m}_{\mathbf{w}_{\mathbf{y}_{t-1|t-1}}}} (\mathbf{w}_{\mathbf{y}_t} - \mathbf{m}_{\mathbf{w}_{\mathbf{y}_{t-1|t-1}}}), \end{aligned} \tag{2.34}$$

$$\begin{aligned} h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\gamma}) =& h(\mathbf{m}_{\mathbf{y}_{t|t-1}}, \mathbf{m}_{\mathbf{w}_{\mathbf{z}_{t|t-1}}}, \boldsymbol{\gamma}) \\ &+ \left.\frac{\partial h}{\partial \mathbf{y}}\right|_{\mathbf{y}_t=\mathbf{m}_{\mathbf{y}_{t|t-1}}} (\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t|t-1}}) \\ &+ \left.\frac{\partial h}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{m}_{\mathbf{w}_{\mathbf{z}_{t|t-1}}}} (\mathbf{w}_{\mathbf{z}_t} - \mathbf{m}_{\mathbf{w}_{\mathbf{z}_{t|t-1}}}). \end{aligned} \tag{2.35}$$

Ensuite, on applique les équations de Kalman à la représentation d'état linéarisée :

1. **Prédiction** :

$$\mathbf{m}_{\mathbf{y}_{t|t-1}} = f(\mathbf{m}_{\mathbf{y}_{t-1|t-1}}, \mathbf{c}_t, \mathbf{m}_{\mathbf{w}_{\mathbf{y}_{t-1|t-1}}}, \boldsymbol{\theta}_\mathbf{y}), \tag{2.36}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}_{t|t-1}} = \mathbf{F}_t \boldsymbol{\Sigma}_{\mathbf{y}_{t-1|t-1}} \mathbf{F}_t^T + \mathbf{D}_t \boldsymbol{\Sigma}_{\mathbf{w}_\mathbf{y}} \mathbf{D}_t^T. \tag{2.37}$$

2. **Mise à jour** :

$$\mathbf{m}_{\mathbf{z}_{t|t-1}} = h(\mathbf{m}_{\mathbf{y}_{t|t-1}}, \mathbf{m}_{\mathbf{w}_{\mathbf{z}_{t|t-1}}}, \boldsymbol{\gamma}), \tag{2.38}$$

$$\boldsymbol{\Sigma}_{\mathbf{z}_{t|t-1}} = \mathbf{H}_t \boldsymbol{\Sigma}_{\mathbf{y}_{t|t-1}} \mathbf{H}_t^T + \mathbf{E}_t \boldsymbol{\Sigma}_{\mathbf{w}_\mathbf{z}} \mathbf{E}_t^T, \tag{2.39}$$

$$\mathbf{K}_t = \boldsymbol{\Sigma}_{\mathbf{y}_{t|t-1}} \mathbf{H}_t \boldsymbol{\Sigma}_{\mathbf{z}_{t|t-1}}^{-1}, \tag{2.40}$$

$$\mathbf{m}_{\mathbf{y}_{t|t}} = \mathbf{m}_{\mathbf{y}_{t|t-1}} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_{t|t-1}}), \tag{2.41}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}_{t|t}} = (\boldsymbol{I} - \mathbf{K}_t \mathbf{H}_t) \boldsymbol{\Sigma}_{\mathbf{y}_{t|t-1}}, \tag{2.42}$$

où $\mathbf{F}_t = \left.\frac{\partial f}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{m}_{\mathbf{y}_{t-1|t-1}}}$, $\mathbf{D}_t = \left.\frac{\partial f}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{m}_{\mathbf{w}_{\mathbf{y}_{t-1|t-1}}}}$, $\mathbf{H}_t = \left.\frac{\partial h}{\partial \mathbf{y}}\right|_{\mathbf{y}_t=\mathbf{m}_{\mathbf{y}_{t|t-1}}}$ et $\mathbf{E}_t = \left.\frac{\partial h}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{m}_{\mathbf{w}_{\mathbf{z}_{t|t-1}}}}$.

On retrouve une adaptation pour les équations de prédiction et de mise à jour des équations de Kalman. L'expression de la prédiction de la moyenne dans l'équation (2.36) consiste au simple calcul de l'image de la moyenne par f, et la propagation de la variance se fait en remplaçant les matrices par la jacobienne de F dans l'équation (2.37). De la même façon, les équations (2.38) et (2.39) de mise à jour s'écrivent en utilisant l'image de la moyenne par la fonction h et l'utilisation de la jacobienne de h. Le gain de Kalman écrit en (2.40) se calcule aussi également en utilisant la linéarisation de h qui constitue une approximation de l'équation (2.33). Finalement, les mise à jours sont effectuées en utilisant ce gain de Kalman approché dans (2.41) et (2.42).

Cette extension, très utilisée, est simple à implémenter mais peut mener dans de nombreux cas à des problèmes de divergence [Ljung, 1979][Boutayeb et al., 1997]. La raison principale est la méthode de linéarisation autour de la moyenne qui peut présenter, en cas de forte non-linéarité, une bien trop mauvaise méthode pour propager les variables aléatoire à travers celle-ci.

En effet, tout comme dans le cas linéaire, on cherche à propager les incertitudes à travers les fonctions f et h, c'est-à-dire à calculer la distribution de probabilité des variables aléatoires

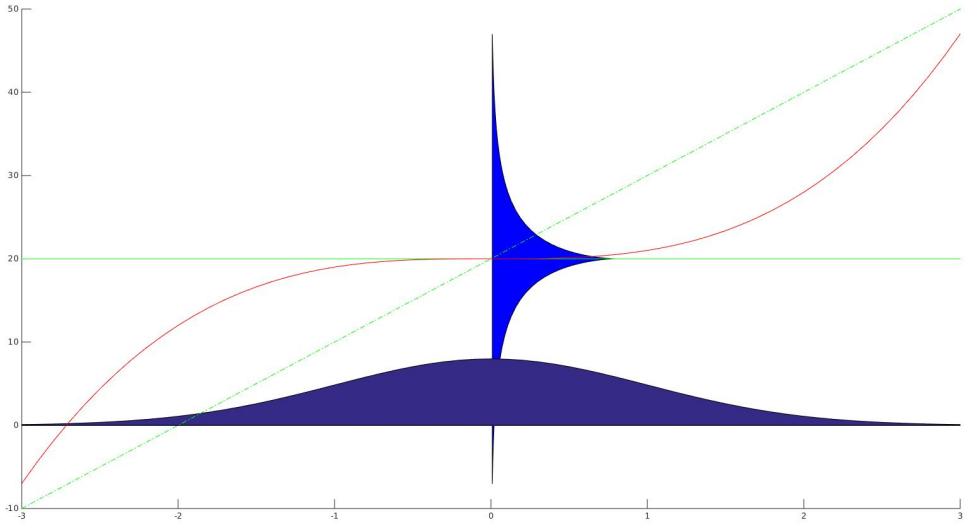


FIGURE 2.7 – Propagation d'une distribution gaussienne (représentée en bleu foncé) à travers la fonction $f(x) = 1 + x^3$ (représentée en rouge). La distribution en bleu clair obtenue est représentée sur l'axe des ordonnées. En linéarisant autour de la moyenne de la gaussienne, (*i.e.* en 0), on obtient l'approximation $f(x) \simeq f(0) + f'(0)x = 1$, représentée en vert dans la figure. En utilisant cette approximation pour propager la distribution, on a $\mathbb{E}[f(x)] \simeq \mathbb{E}[f(0) + f'(0)x] = 1$ et $\mathrm{Cov}[f(x)] \simeq \mathrm{Cov}[f(0) + f'(0)x] = 0$, soit une gaussienne de variance nulle.

$f(\mathbf{y}_{1:t}, \mathbf{c}_{1:t+1}, \mathbf{w}_{\mathbf{y}_1:\mathbf{y}_t}, \boldsymbol{\theta}_\mathbf{y})$ et $h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\gamma})$ à partir des distributions *a priori* des variables $\mathbf{y}_{1:t}$, $\mathbf{w}_{\mathbf{y}_1:\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_1:\mathbf{z}_t}$.

Cependant, en linéarisant la fonction non linéaire autour de la moyenne par un développement de Taylor, on peut se retrouver confronté au problème représenté sur la Figure 2.7. Cet exemple permet de comprendre intuitivement que linéariser localement, en ne considérant que la dérivée au voisinage de la fonction, ne permet pas toujours de propager correctement la distribution.

Filtres de Kalman basés sur la linéarisation statistique : filtre unscented et central difference

Pour contourner le problème posé par la linéarisation du filtre de Kalman étendu, on remarque qu'il serait préférable de linéariser, non pas par un développement de Taylor, mais en utilisant une autre fonction linéaire, comme la corde représentée en pointillés verts dans la Figure 2.7. donnée en exemple. C'est à partir de ce constat, qu'une méthode astucieuse a été proposée par Julier & Uhlmann [1997]. Cette méthode utilise une faible quantité d'échantillons déterministes afin de propager la moyenne et la covariance. L'idée est de choisir judicieusement un petit nombre de points, qui, une fois propagés à travers la fonction f ou h permettent de définir indirectement une linéarisation plus adaptée à la propagation d'une distribution gaussienne.

La méthode consiste d'abord à propager $\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{y}_t$ à travers la fonction f. On concatène dans cette première étape les vecteurs aléatoires en un unique vecteur aléatoire $\bar{\mathbf{y}}_t$ de taille L :

$$\bar{\mathbf{y}}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_{\mathbf{y}_t} \end{bmatrix}. \tag{2.43}$$

$\bar{\mathbf{y}}_t$ est une gaussienne multivariée dont on note la moyenne et la covariance *a priori* respectivement par $\bar{\mathbf{m}}_{\mathbf{y}_{t-1|t-1}}$ et $\bar{\mathbf{\Sigma}}_{\mathbf{y}_{t-1|t-1}}$. Similairement, on remplace f dans l'équation d'évolution par la fonction $\bar{f}$ qui prend en entrée $\bar{\mathbf{y}}_t$ et $\mathbf{c}_t$. Ensuite, on génère un nombre fini et déterministe de points η_i, appelés par les auteurs de la méthode *sigma-points*, à partir de la moyenne et de la matrice de covariance et des formules suivantes :

$$\boldsymbol{\eta_0} = \bar{\mathbf{m}}_{\mathbf{y}_{t-1|t-1}}, \tag{2.44}$$

$$\boldsymbol{\eta_i} = \bar{\mathbf{m}}_{\mathbf{y}_{t-1|t-1}} + \left(\sqrt{(i+\kappa)\bar{\mathbf{\Sigma}}_{\mathbf{y}_{t-1|t-1}}}\right)_i, \ i \in \{1, ..., L\}, \tag{2.45}$$

$$\boldsymbol{\eta_i} = \bar{\mathbf{m}}_{\mathbf{y}_{t-1|t-1}} - \left(\sqrt{(i+\kappa)\bar{\mathbf{\Sigma}}_{\mathbf{y}_{t-1|t-1}}}\right)_{i-L}, \ i \in \{L+1, ..., 2*L\}, \tag{2.46}$$

où κ est un paramètre réel, où la racine carrée représente la factorisation de Cholesky, et où la notation $(\mathbf{M})_i$ désigne la i-ème colonne de M. On calcule ensuite l'image $\boldsymbol{\xi}_i$ de chaque *sigma-point* $\boldsymbol{\eta}_i$ par f :

$$\boldsymbol{\xi}_i = f(\boldsymbol{\eta}_i, \mathbf{c}_t). \tag{2.47}$$

La moyenne et la covariance propagées sont calculées à l'aide de la moyenne et la variance empirique des $\boldsymbol{\xi}_i$ pondérées par des poids W^i_{mean} etW^i_{cov} :

$$\bar{\mathbf{m}}_{\mathbf{y}_{t|t-1}} = \sum_{i=0}^{2L} W^i_{\text{mean}} \boldsymbol{\xi}_i, \tag{2.48}$$

$$\bar{\mathbf{\Sigma}}_{\mathbf{y}_{t|t-1}} = \sum_{i=0}^{2L} W^i_{\text{cov}} \ [\boldsymbol{\xi}_i - \bar{\mathbf{m}}_{\mathbf{y}_{t|t-1}}][\boldsymbol{\xi}_i - \bar{\mathbf{m}}_{\mathbf{y}_{t|t-1}}]^{\mathrm{T}}. \tag{2.49}$$

Bien que ça ne soit pas nécessairement utile pour la fonction f, on peut également calculer la covariance croisée

$$\bar{\mathbf{\Sigma}}_{\mathbf{y}_{t|t-1}\mathbf{y}_{t-1|t-1}} = \sum_{i=0}^{2L} W^i_{\text{cov}} \ [\boldsymbol{\eta}_i - \bar{\mathbf{m}}_{\mathbf{y}_{t-1|t-1}}][\boldsymbol{\xi}_i - \bar{\mathbf{m}}_{\mathbf{y}_{t|t-1}}]^{\mathrm{T}}. \tag{2.50}$$

Les poids W^i_{mean} et W^i_{cov} permettent de calculer indirectement une linéarisation plus adaptée et sont définis par :

$$W^i_{\text{mean}} = \frac{\lambda}{\lambda + L}, \tag{2.51}$$

$$W^i_{\text{cov}} = \frac{\lambda}{\lambda + L} + (1 - \alpha^2 + \beta), \tag{2.52}$$

$$W^0_{\text{mean}} = W^0_{\text{cov}} = \frac{\lambda}{2(\lambda + L)}, \tag{2.53}$$

où λ, α et β sont des paramètres de la méthode, dont les valeurs jouent sur l'efficacité et seront discutés plus tard.

En appliquant exactement le même procédé, on peut ensuite calculer les distributions propagées par la fonction h et calculer de la même façon $\mathbf{m}_{\mathbf{z}_{t|t-1}}$, $\mathbf{\Sigma}_{\mathbf{z}_{t|t-1}}$ et la matrice de covariance croisée $\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}\mathbf{z}_{t|t-1}}$.

Dans le cas de la fonction h, la covariance croisée est cette fois utile car elle permet de calculer une matrice $\mathbf{K}_t$ similaire au gain de Kalman en utilisant l'équation (2.33).

La matrice $\mathbf{K}_t$ permet de faire les mises à jour à partir de l'observation $\mathbf{z}_t$ comme dans le filtre classique :

$$\mathbf{m}_{\mathbf{y}_{t|t}} = \mathbf{m}_{\mathbf{y}_{t|t-1}} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_{t|t-1}}), \tag{2.54}$$

$$\mathbf{\Sigma}_{\mathbf{y}_{t|t}} = \mathbf{\Sigma}_{\mathbf{y}_{t|t-1}} - \mathbf{K}_t\mathbf{\Sigma}_{\mathbf{z}_{t|t-1}}\mathbf{K}_t^{\mathrm{T}}. \tag{2.55}$$

Cette méthode de filtrage présente l'avantage qu'il n'est pas nécessaire de calculer les dérivées des fonctions non-linéaires pour une complexité algorithmique similaire au filtre étendu [Wan & Van Der Merwe, 2000]. De plus, cette méthode s'avère souvent plus performante : moyennant un choix correct des paramètres κ, α, β et λ l'approximation *unscented* estime correctement la moyenne et la covariance jusqu'à l'ordre 2 du développement limité de la fonction non-linéaire, voire jusqu'à l'ordre 3 si la distribution effective du problème est symétrique comme montré originellement par Julier [2002]. Le choix des paramètres κ, α, β et λ est pris de la façon suivante : Ils doivent vérifier $\kappa \geq 0$, $0 \leq \alpha \leq 1$, $\beta \geq 0$ et $\lambda = \alpha^2(L + \kappa) - L$. Le paramètre κ n'est pas critique mais permet d'assurer la semi-positivité des matrices de covariance et un bon choix de départ est 0. Le paramètre α est un paramètre d'échelle généralement petit, et le paramètre β est un paramètre de poids permettant d'incorporer des ordres plus élevés et qui doit être choisi en fonction de la symétrie de la distribution. La valeur $\beta = 2$ est optimale pour des distributions gaussiennes [Julier, 1998].

La méthode *unscented Kalman* a depuis été utilisée avec succès dans de nombreuses applications, se montrant la plupart du temps plus efficace et robuste que le filtre étendu. On peut citer l'estimation de paramètres aérodynamiques [Chowdhary & Jategaonkar, 2010], l'ingénierie des matériaux [Wu & Smyth, 2007], la navigation [Zhang et al., 2005; Shin & El-Sheimy, 2004] ou encore le monitoring [Ghahremani & Kamwa, 2011] et le contrôle de générateurs électriques des éoliennes [Vepa, 2011]. On peut également citer la publication de Kandepu et al. [2008] qui ont montré expérimentalement des performances améliorées pour le filtre *unscented* sur quatre cas applicatifs.

Pour finir, notons que des méthodes très proches, appelées *central difference filter* et *divided difference filter* ont été également proposées par Ito & Xiong [2000] et NøRgaard et al. [2000]. Ces méthodes diffèrent par l'utilisation de *sigma-points* de poids différents. Dans la thèse de Van Der Merwe [2004], ces méthodes ont été unifiées et désignées généralement sous le nom de *sigma-points Kalman filters*. Plus récemment, d'autres extensions ont été proposées comme celle de Ponomareva et al. [2010] qui, moyennant l'utilisation d'un nombre de *sigma-points* plus élevés, permettent une approximation des distributions propagées à des ordres plus élevés.

2.2.3 Filtres de Kalman itérés

Les filtres de Kalman itérés peuvent être établis, en sortant du cadre statistique et se plaçant du point de vue de l'optimisation [Skoglund et al., 2015]. Pour cela, considérons à nouveau la distribution *a posteriori* définie par l'équation (2.17). De façon similaire à l'estimateur du maximum de vraisemblance, on considère désormais l'estimateur du maximum *a posteriori* $\boldsymbol{\theta}_{\mathrm{map}}$ (MAP), qui est défini de façon générale comme

$$\boldsymbol{\theta}_{\mathrm{map}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{z}_{1:N}). \tag{2.56}$$

Dans le cas du filtrage bayésien récursif non-linéaire et gaussien, à l'instant t et en considérant l'équation (2.24), on a alors

$$\boldsymbol{\theta}_{\mathrm{map}} = \arg\max_{\mathbf{y}_t} p(\mathbf{y}_t|\mathbf{z}_{1:t}) \tag{2.57}$$

$$= \arg\max_{\mathbf{y}_t} \frac{p(\mathbf{z}_t|\mathbf{y}_t)p(\mathbf{y}_t|\mathbf{z}_{1:t-1})}{p(\mathbf{z}_t|\mathbf{z}_{1:t-1})} \tag{2.58}$$

$$= \arg\max_{\mathbf{y}_t} \frac{1}{(2\pi)^{\frac{N_\mathbf{z}}{2}}|\mathbf{\Sigma}_{\mathbf{w}_{\mathbf{z}_t}}|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{z}_t - h(\mathbf{y}_t, 0))^T \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{z}_t}}^{-1}(\mathbf{z}_t - h(\mathbf{y}_t, 0))) \tag{2.59}$$

$$\times \frac{1}{(2\pi)^{\frac{N_\mathbf{y}}{2}}|\mathbf{\Sigma}_{\mathbf{w}_{\mathbf{y}_t}}|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t-1|t-1}})^T \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{y}_t}}^{-1}(\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t-1|t-1}})) \tag{2.60}$$

$$\times \text{cst}(\mathbf{z}_t) \tag{2.61}$$

où $\text{cst}(\mathbf{z}_t) = \frac{1}{p(\mathbf{z}_t|\mathbf{z}_{t-1})}$ qui est une constante lorsque $\mathbf{z}_t$ est réalisé. Pour trouver le maximum de la fonction à maximiser, on considère en général le logarithme de la probabilité et on inverse le signe. Le problème peut s'écrire alors :

$$\boldsymbol{\theta}_{\text{map}} = \arg\min_{\mathbf{y}_t} -\log p(\mathbf{y}_t|\mathbf{z}_t) \tag{2.62}$$

$$= \arg\min_{\mathbf{y}_t} \frac{1}{2}(\mathbf{z}_t - h(\mathbf{y}_t, 0))^T \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{z}_t}}^{-1}(\mathbf{z}_t - h(\mathbf{y}_t, 0)) \tag{2.63}$$

$$+ \frac{1}{2}(\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t-1|t-1}})^T \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{y}_t}}^{-1}(\mathbf{y}_t - \mathbf{m}_{\mathbf{y}_{t-1|t-1}}). \tag{2.64}$$

C'est un problème de moindres carrés non-linéaires où on note ici la fonction de coût $V(\mathbf{y}_t)$ et qui s'écrit :

$$\boldsymbol{\theta}_{\text{map}} = \arg\min_{\mathbf{y}_t} V(\mathbf{y}_t) \tag{2.65}$$

$$= \arg\min_{\mathbf{y}_t} \frac{1}{2}(\mathbf{x}_t - g(\mathbf{y}_t))^T \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{x}_t}}^{-1}(\mathbf{x}_t - g(\mathbf{y}_t)) \tag{2.66}$$

où $\mathbf{x}_t = [z_t \ \mathbf{m}_{\mathbf{y}_{t-1|t-1}}]^T$, $g(\mathbf{y}_t) = [h(\mathbf{y}_t, 0) \ \mathbf{y}_t]^T$ et $\mathbf{\Sigma}_{\mathbf{w}_{\mathbf{x}_t}} = \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{z}_t}} & 0 \\ 0 & \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{z}_t}} \end{bmatrix}$.

On peut alors résoudre le problème de filtrage en utilisant des méthodes d'optimisation classiques (méthode du gradient, méthode de Newton, *etc*).

À ce titre, on peut noter que la mise à jour de Kalman dans l'équation (2.31) correspond à faire un unique « pas » de la méthode de Gauss-Newton. C'est ce qu'ont montré Bell & Cathey [1993]. En définissant la fonction $l(\mathbf{y}_t) = \mathbf{\Sigma}_{\mathbf{w}_{\mathbf{x}_t}}^{-\frac{1}{2}}(\mathbf{x}_t - g(\mathbf{y}_t))$ et en notant $J(\mathbf{y}_t) = \left.\frac{\partial l(\mathbf{y})}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}_t}$ sa matrice jacobienne, on a

$$V(\mathbf{y}_t) = l(\mathbf{y}_t)^T l(\mathbf{y}_t) \tag{2.67}$$

et le pas de Gauss-Newton s'écrit alors [Wedderburn, 1974]

$$\mathbf{y}_t^{(i+1)} = \mathbf{y}_t^{(i)} + \left(J(\mathbf{y}_t^{(i)})^T J(\mathbf{y}_t^{(i)})\right)^{-1} J(\mathbf{y}_t^{(i)})^T l(\mathbf{y}_t^{(i)}), \tag{2.68}$$

où l'exposant (i) représente une itération d'un pas de la méthode de Gauss-Newton. En développant (2.68) et en utilisant un lemme d'inversion de matrice par blocs [Skoglund et al., 2015], on peut arriver à l'expression utilisée usuellement pour le filtre de Kalman itéré :

$$\mathbf{y}_t^{(i+1)} = \mathbf{y}_t^{(i)} + \left(\mathbf{m}_{\mathbf{y}_{t|t-1}} - \mathbf{y}_t^{(i)} + \mathbf{K}_t^{(i)}\left(\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_{t|t-1}} - \mathbf{H}_t^{(i)}\left(\mathbf{m}_{\mathbf{y}_{t|t-1}} - \mathbf{y}_t^{(i)}\right)\right)\right), \tag{2.69}$$

où $\mathbf{y}_t^{(0)} = \mathbf{m}_{\mathbf{y}_{t|t-1}}$, $\mathbf{H}_t^{(i)} = \left.\frac{\partial h(\mathbf{y},0)}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}_t}$ et $\mathbf{K}_t^{(i)} = \mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}\mathbf{H}_t^{(i)}(\mathbf{H}_t^{(i)}\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}\mathbf{H}_t^{(i)T} + \mathbf{\Sigma}_{\mathbf{w}_\mathbf{z}})^{-1}$.

On constate d'ailleurs que pour $i = 1$, on retombe bien sur l'expression de la mise à jour de Kalman donnée en (2.31). Une expression similaire permet de calculer la matrice de covariance :

$$\mathbf{\Sigma}_{\mathbf{y}_{t|t}} = (\mathbf{I} - \mathbf{K}_t^{(i)}\mathbf{H}_t^{(i)})\mathbf{\Sigma}_{\mathbf{y}_{t|t-1}}. \tag{2.70}$$

Le filtre de Kalman itéré a été utilisé avec succès dans différentes applications [Gu et al., 2007; Fang & Gong, 2010] où les filtres de Kalman étendu ou *unscented* n'étaient pas efficaces.

2.2.4 Autres filtres basés sur des méthodes d'optimisation

D'autres méthodes de filtrage ont été proposées en se calquant sur des méthodes d'optimisation. Notamment, dans la méthode de Moriyama et al. [2003], les auteurs ont proposé d'utiliser une méthode de recherche linéaire. L'équation de mise à jour devient alors

$$\mathbf{y}_t^{(i+1)} = \mathbf{y}_t^{(i)} + \alpha_t^i \left(\mathbf{m}_{\mathbf{y}_{t|t-1}} - \mathbf{y}_t^{(i)} + \mathbf{K}_t^{(i)} \left(\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_{t|t-1}} - \mathbf{H}_t^{(i)} \left(\mathbf{m}_{\mathbf{y}_{t|t-1}} - \mathbf{y}_t^{(i)}\right)\right)\right) \tag{2.71}$$

où le paramètre α_t^i est choisi tel que $0 < \alpha_t \leq 1$ et de sorte à ce que la fonction de coût $V(\mathbf{y}_t)$ soit minimale.

On peut citer également la méthode de Bellaire et al. [1995] basée sur l'algorithme de Levenberg-Marquardt [Moré, 1978] ou l'algorithme quasi-Newton [Skoglund et al., 2015].

Pour finir, notons que d'autres auteurs ont proposé des adaptations du filtre de Kalman itéré en se basant sur les algorithmes à *sigma-points* tels que Sibley et al. [2006] qui comparent cette méthode avec le filtre itéré standard.

2.2.5 Filtrage particulaire et méthodes d'ensemble

Le filtrage particulaire, appelé aussi méthode de Monte-Carlo séquentielle, constitue une méthode plus générale pour résoudre le problème de filtrage bayésien récursif en présence de fortes non-linéarités ou de distributions fortement non-gaussiennes dans la mesure où on ne s'intéresse pas nécessairement aux moments de la distribution mais à approximer la distribution complète. On se place donc dans la suite dans le cas général des équations (2.12) et (2.13), où les bruits ne sont pas nécessairement gaussiens et les fonctions f et g ne sont pas linéaires. Le principe est d'approximer les distributions de probabilité par un ensemble de tirages aléatoires [Metropolis & Ulam, 1949] afin de les propager plus facilement à travers les non-linéarités du système. La méthode d'ensemble, quant à elle, constitue une méthode hybride entre le filtre de Kalman et le filtre à particules. Dans la suite, on présente brièvement ces méthodes.

Filtre à particules

Le filtre à particules [Liu & Chen, 1998; Arulampalam et al., 2002] consiste à approximer la distribution $\mathbf{y}_{0:t}|\mathbf{z}_{1:t}$ à chaque instant t par N tirages aléatoires $\mathbf{y}_t^n$, $n \in \{1, ..., N\}$:

$$p(\mathbf{y}_{0:t}|\mathbf{z}_{1:t}) \approx \sum_{n=1}^{N} w_t^n \delta(\mathbf{y}_t - \mathbf{y}_t^n), \tag{2.72}$$

où δ est la distribution de Dirac et où w_t^n est un poids à définir et censé représenter, du mieux possible, la densité de probabilité au voisinage du point afin de reconstituer la distribution. De la même façon que pour le filtre *unscented*, les échantillons, appelés aussi particules, sont ensuite propagés à travers le modèle (fonction f) et l'équation de mesure (fonction h) pour approximer la

distribution *a posteriori*. En pratique, il n'est pas possible d'effectuer des tirages de la distribution $p(\mathbf{y}_{0:t}|\mathbf{z}_{1:t})$. De ce fait, pour constituer une solution utilisable, on utilise donc une distribution dite d'importance $q(\mathbf{y}_{0:t}|\mathbf{z}_{1:t})$ pour laquelle on sait effectuer des tirages et les poids w_t^n sont alors définis par :

$$w_t^n = \frac{p(\mathbf{y}_{0:t}^n|\mathbf{z}_{1:t})}{q(\mathbf{y}_{0:t}^n|\mathbf{z}_{1:t})}. \tag{2.73}$$

Dans le cas récursif bayésien, on suppose également que la distribution d'importance peut se factoriser de la manière suivante :

$$q(\mathbf{y}_{0:t}|\mathbf{z}_{1:t}) = q(\mathbf{y}_t|\mathbf{y}_{0:t-1}, \mathbf{z}_{1:t})q(\mathbf{y}_{0:t-1}|\mathbf{z}_{1:t-1}). \tag{2.74}$$

À partir des équations (2.73) et (2.74), on peut mettre à jour les poids une fois les particules propagées à chaque étape à partir de l'équation suivante :

$$w_{t+1}^n \propto w_t^n \, \frac{p(\mathbf{z}_t|\mathbf{y}_t^n)p(\mathbf{y}_t^n|\mathbf{y}_{t-1}^n)}{q(\mathbf{y}_t^n|\mathbf{z}_t)}. \tag{2.75}$$

En normalisant les poids w_{t+1}^n, on retombe sur l'équation (2.72). Cette méthode est appelée *sequential importance sampling* et constitue la méthode de filtrage particulaire la plus simple.

La mise en œuvre de cet algorithme nécessite certaines précautions. Tout d'abord, le choix de la distribution d'importance est primordial. Doucet et al. [2000] ont montré que le choix optimal est celui qui minimise la variance des poids w_t^n et est le suivant :

$$q(\mathbf{y}_t|\mathbf{y}_{1:t-1}\mathbf{z}_{1:t}) = p(\mathbf{y}_t|\mathbf{y}_{t-1}, \mathbf{z}_t). \tag{2.76}$$

En pratique, ce choix de distribution est complexe à implémenter et bien souvent [Gordon et al., 1993; Avitzour, 1995; Kitagawa, 1996] on choisit plutôt

$$q(\mathbf{y}_t|\mathbf{y}_{1:t-1}\mathbf{z}_{1:t}) = p(\mathbf{y}_t|\mathbf{y}_{t-1}). \tag{2.77}$$

L'équation de mise à jour des poids devient alors

$$w_{t+1}^n = w_t^n \; p(\mathbf{z}_t|\mathbf{y}_t^n), \tag{2.78}$$

ce qui revient à dire qu'à chaque étape le tirage d'une particule n'est fait qu'à partir de la vraisemblance du modèle, sans tenir compte des observations les plus récentes. Cette formulation, bien que pratique, peut engendrer une variance élevée.

Un autre problème récurrent est le problème de dégénérescence [Doucet et al., 2001]. En effet, au bout de quelques itérations et parmi l'ensemble des particules choisies au départ, la plupart des particules ont un poids négligeable. Pour remédier à ce problème, on effectue en général une étape de sélection et de ré-échantillonnage [Douc & Cappé, 2005], où les particules de faible poids sont supprimées et où de nouveaux échantillons sont tirés. Parmi les stratégies de ré-échantillonnage, on peut citer les méthodes *sampling-importance resampling* (SIR) [Rubin, 1987], *residual resampling* [Higuchi, 1997] et *minimum variance sampling* [Crisan & Doucet, 2002].

Filtre d'ensemble

Le filtre d'ensemble est un filtre à mi-chemin entre le filtre à particules et le filtre de Kalman et constitue une variante stochastique du filtre de Kalman. Le principe ici est de tirer des échantillons de la loi de $\mathbf{y}_{t-1}|\mathbf{z}_{t-1}$ ainsi que le même nombre d'échantillons pour les bruits aléatoires

$\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_t}$, et de propager ces échantillons à travers les fonctions f et h. D'une manière relativement similaire au filtre *unscented*[8] on peut reconstruire approximativement les moments d'ordre 1 et 2 à partir de ces échantillons en calculant les moments empiriques. Pour les N échantillons $\mathbf{y}_{t|t}^{(n)}$ $n \in \{1, ..., N\}$ tirés de la loi de $\mathbf{y}_{t-1}|\mathbf{z}_{t-1}$ et les N échantillons $\mathbf{w}_{\mathbf{y}}^{(n)}$ tirés de la loi de $\mathbf{w}_{\mathbf{y}}$, l'étape de prédiction se calcule en propageant les échantillons à travers l'équation d'état :

$$\mathbf{y}_{t|t-1}^{(n)} = \mathbf{F}\mathbf{y}_{t-1|t-1}^{(n)} + \mathbf{G}\mathbf{c}_t + \mathbf{D}\mathbf{w}_{\mathbf{y}}^{(n)}. \tag{2.79}$$

La mise à jour est ensuite effectuée sur chaque échantillon propagé et s'écrit de la façon suivante :

$$\mathbf{y}_{t|t}^{(n)} = \mathbf{y}_{t|t-1}^{(n)} + \mathbf{K}_t \left(\mathbf{z}_t - \left(\mathbf{H}_{t-1}\mathbf{y}_{t|t-1}^{(n)} - \mathbf{w}_{\mathbf{z}}^{(n)} \right) \right) \tag{2.80}$$

où $\mathbf{K}_t$ est obtenue à l'aide du calcul des matrices de covariances empirique et de l'équation (2.33). Ces particules « *a posteriori* » peuvent être utilisées pour recalculer les moments de la loi *a posteriori* et servir pour l'itération suivante avec éventuellement un ré-échantillonnage. Notons que cette méthode et ses extensions sont utilisées notamment dans des problèmes de grande taille où l'échantillonnage permet une réduction significative de la dimension du problème [Crow & Wood, 2003; Brusdal et al., 2003; Keppenne & Rienecker, 2003; Natvik & Evensen, 2003] mais aussi dans des problèmes non-linéaires sous des bruits gaussiens en adaptant les équations.

2.3 Apprentissage des paramètres

Jusqu'ici, on a présenté l'inférence des variables d'état du problème, en supposant connus les paramètres du système. Cependant, ces paramètres sont souvent inconnus. De ce fait, on présente dans la suite un état de l'art des méthodes utilisées conjointement aux problèmes d'inférence présentés afin d'estimer les paramètres des systèmes. En général, on traite indifféremment l'estimation des paramètres des lois de probabilité (paramètres de loi des bruits $\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_t}$) et l'estimation des autres paramètres du modèle (paramétrant les fonctions non-linéaires f et g par exemple). Dans cette partie, on note désormais $\boldsymbol{\theta}$ l'ensemble des paramètres du système.

2.3.1 Apprentissage des paramètres des modèles ARIMA

L'apprentissage des paramètres des modèles ARIMA se fait en général en résolvant un problème d'optimisation non-linéaire potentiellement complexe. La difficulté intrinsèque de l'estimation dépend cependant du modèle sélectionné (AR, MA, ARMA ou ARIMA). Dans le cas plus simple d'un modèle auto-régressif pur ($\beta_i = 0$ dans l'équation (2.14)), l'apprentissage des paramètres peut se faire par une méthode des moments (équation de Yule-Walker [Brockwell & Davis, 2016]) ou en maximisant le maximum de vraisemblance.

On se place ici dans le cas univarié et on pose $\mathbf{a} = [\alpha_1\ \alpha_2\ ...\ \alpha_{i_\alpha}]$ et $\sigma_\tau^2 = \mathbb{E}((\mathbf{y}_t - \mathbb{E}(\mathbf{y}_t))(\mathbf{y}_{t-\tau} - \mathbb{E}(\mathbf{y}_{t-\tau})))$. On rappelle que i_α est le nombre de paramètres du modèle AR. L'équation de Yule-Walker s'écrit alors :

$$\boldsymbol{\rho} = \boldsymbol{\Gamma}\mathbf{a}, \text{ avec } \boldsymbol{\rho}_t = \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ ... \\ \sigma_{i_\alpha}^2 \end{bmatrix} \text{ et } \boldsymbol{\Gamma}_t = \begin{bmatrix} \sigma_0^2 & \sigma_1^2 & ... & \sigma_{i_\alpha-1}^2 \\ \sigma_1^2 & \sigma_0^2 & ... & \sigma_{i_\alpha-2}^2 \\ \vdots & \vdots & & \vdots \\ \sigma_{i_\alpha-1}^2 & \sigma_{i_\alpha-2}^2 & ... & \sigma_0^2 \end{bmatrix}. \tag{2.81}$$

8. À noter toutefois qu'historiquement, le filtre d'ensemble est apparu un peu avant le filtre *unscented*

Cette équation s'obtient en posant $\mathbf{Y}_t = [\mathbf{y}_t\ \mathbf{y}_{t-1}\ ...\ \mathbf{y}_{t-i_\alpha}]^T$ puis en calculant la covariance de $\mathbf{y}_t\mathbf{Y}_t$ et en utilisant l'équation (2.14). En inversant le système et en utilisant les variances et covariances empiriques, on peut alors apprendre les paramètres α_i.

L'autre façon de procéder est de considérer la vraisemblance du modèle ou de façon équivalente pour un bruit gaussien, le problème des moindres carrés suivant :

$$\arg_{\mathbf{a}} \min ||\mathbf{v} - \mathbf{M}\mathbf{a}||_2, \text{ avec } \mathbf{v} = [\mathbf{y}_t, ..., \mathbf{y}_{i_\alpha+1}]^T \text{et } \mathbf{M} = \begin{bmatrix} \mathbf{y}_{t-1} & \mathbf{y}_{t-2} & \cdots & \mathbf{y}_{t-i_\alpha} \\ \mathbf{y}_{t-2} & \mathbf{y}_{t-3} & \cdots & \mathbf{y}_{t-(i_\alpha-1)} \\ \vdots & \vdots & ... & \vdots \\ \mathbf{y}_{i_\alpha} & \mathbf{y}_{i_\alpha-1} & \cdots & \mathbf{y}_1 \end{bmatrix}. \quad (2.82)$$

La solution du problème aux moindres carrés s'écrit alors :

$$\mathbf{a} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{v}. \quad (2.83)$$

Dans le cas plus général des modèles ARMA ou ARIMA, l'équation de Yule-Walker ne permet plus l'estimation et le maximum de vraisemblance génère un problème non-linéaire. Pour estimer les paramètres, différentes méthodes ont été utilisées : algorithme de Hannan-Rissanen [Hannan & Rissanen, 1982], filtre de Kalman [Kohn & Ansley, 1986], ou optimisation numérique du maximum de vraisemblance.

2.3.2 Apprentissage des paramètres des filtres bayésiens récursifs par filtrage bayésien

L'apprentissage des paramètres lors des filtres récursifs bayésiens n'est pas nouveau, mais il constitue un sujet de recherche actif, notamment dans le cas du filtrage particulaire [Tulsyan et al., 2013]. On distingue globalement deux classes de méthodes, nommées en général **estimateurs (récursifs) bayésiens** et **estimateurs de vraisemblance** [Kantas et al., 2009].

Dans la première classe, celle des estimateurs bayésiens, l'idée est de considérer les paramètres comme des variables aléatoires au même titre que les variables d'état en leur conférant une loi *a priori*. Ces variables sont alors traitées par un filtre récursif bayésien [Van Der Merwe & Wan, 2001]. Il existe plusieurs variantes, notamment où on alterne entre le filtrage sur les variables d'état et les paramètres (*dual filtering*) ou bien en intégrant directement les paramètres dans l'équation d'état, engendrant alors une nouvelle représentation d'état, de taille supérieure, créant ainsi un système « augmenté » (*joint filtering*). Dans ce cas la représentation d'état s'écrit

$$\begin{cases} \begin{bmatrix} \mathbf{y}_t \\ \boldsymbol{\theta}_{\mathbf{y},t} \end{bmatrix} = \begin{bmatrix} f(\mathbf{y}_{t-1}, \mathbf{c}_{t-1}, \mathbf{w}_{\mathbf{y}_{t-1}}, 0) \\ \boldsymbol{\theta}_{\mathbf{y},t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \xi_t \end{bmatrix} \\ \mathbf{z}_t = h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\theta}_{\mathbf{z}}). \end{cases} \quad (2.84)$$

Cette approche a l'avantage d'unifier le problème d'estimation des paramètres et celui du filtrage, mais présente néanmoins le risque d'altérer la distribution recherchée, les paramètres $\boldsymbol{\theta}$ variant alors au cours du temps [Kantas et al., 2009]. De plus, la méthode est fortement dépendante de la distribution *a priori* proposée pour le paramètre $\boldsymbol{\theta}$ ce qui peut conduire à des problèmes et notamment à une dégénérescence accrue dans le cas du filtre particulaire. Plusieurs variantes ont été proposées pour compenser ces problèmes : *artificial dynamics* [Kitagawa, 1998], *resampling move* [Andrieu et al., 1999; Fearnhead, 2002] et *particle MCMC* [Andrieu et al., 2010].

2.3.3 Apprentissage des paramètres des filtres Bayésiens récursifs : estimation par maximisation de la vraisemblance

La seconde classe de méthode vise à maximiser la vraisemblance des paramètres. La vraisemblance est en général calculée sur l'observation complète, même si des variantes en temps-réel ont été proposées [Kantas et al., 2009]. La fonction de log-vraisemblance s'écrit :

$$\mathcal{L}_z(\boldsymbol{\theta}) = \log p(\mathbf{z}_{1:T}|\boldsymbol{\theta}) \tag{2.85}$$

$$= \log \int_{\mathbf{y}_{1:T}} p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) d\mathbf{y}_{1:T}, \tag{2.86}$$

où $\mathbf{z}_{1:T} = \{\mathbf{z}_1 ... \mathbf{z}_T\}$ et $\mathbf{y}_{1:T} = \{\mathbf{y}_1 ... \mathbf{y}_T\}$ sont respectivement l'ensemble des observations et l'ensemble des états cachés entre $t = 1$ et $t = T$.

Le problème vient du fait qu'on ne connaît pas la distribution de $\mathbf{y}_{1:T}$ sur laquelle intégrer. Pour pouvoir maximiser la log-vraisemblance marginale $\mathcal{L}_z(\boldsymbol{\theta})$, deux méthodes principales existent. La première est la méthode de descente de gradient de la log-vraisemblance marginale [Doucet & Tadić, 2003; Poyiadjis et al., 2011]. Elle nécessite que l'intégrande soit suffisamment régulière ($\mathcal{C}^1$ et de dérivée majorée par une fonction intégrable) afin de permettre d'inverser l'ordre de l'intégration et de la dérivation :

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}_z(\boldsymbol{\theta}) = \int \nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta}) d\mathbf{y}_{1:T}. \tag{2.87}$$

Cette formule est connue sous le nom d'identité de Fisher et permet de calculer le gradient en utilisant la distribution filtrée $p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta})$ et le modèle $p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})$. Les paramètres peuvent alors être mis à jour, en suivant l'algorithme de gradient classique :

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \lambda_k \nabla_{\boldsymbol{\theta}} \mathcal{L}_z(\boldsymbol{\theta}), \tag{2.88}$$

où λ_k est un scalaire d'échelle du pas du gradient. Par exemple et dans le cas du filtrage particulaire, cette méthode peut-être implémentée en utilisant la distribution de filtrage pour approximer $p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta})$ où les particules sont pondérées par les valeurs $\log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})$, afin de calculer numériquement l'intégrale de l'équation (2.87) [Andrieu et al., 2004].

Une seconde méthode consiste à utiliser l'algorithme espérance-maximisation (EM) [Ghahramani & Hinton, 1996; Ghahramani & Roweis, 1999]. C'est un cas particulier de l'algorithme Majorisation-Minimisation (MM) [Hunter & Lange, 2004], qui se base sur l'utilisation d'une fonction dite auxiliaire afin de simplifier le problème d'optimisation. Cette fonction auxiliaire est construite de sorte à minorer la fonction objectif (ou majorer la fonction de coût dans le cas d'un problème de minimisation). En recherchant le maximum de cette fonction, on a ainsi une garantie d'accroître la valeur de la fonction objectif. Dans le cas de l'algorithme EM, la fonction auxiliaire est construite à l'étape E, en inférant l'état caché à partir du paramètre $\boldsymbol{\theta}^{(i)}$ courant, puis la fonction auxiliaire est maximisée et $\boldsymbol{\theta}^{(i+1)}$ est ainsi obtenu. L'algorithme fonctionne grâce à une inégalité due à Jensen [Bishop, 2007] : pour une distribution quelconque $Q(\mathbf{y}_{1:T})$ de $\mathbf{y}_{1:T}$ on a

$$\log \int_{\mathbf{y}_{1:T}} p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) d\mathbf{y}_{1:T} = \log \int_{\mathbf{y}_{1:T}} Q(\mathbf{y}_{1:T}) \frac{p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})}{Q(\mathbf{y}_{1:T})} d\mathbf{y}_{1:T} \tag{2.89}$$

$$\geq \int_{\mathbf{y}_{1:T}} Q(\mathbf{y}_{1:T}) \log \left(\frac{p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})}{Q(\mathbf{y}_{1:T})} \right) d\mathbf{y}_{1:T} \tag{2.90}$$

$$= \int_{\mathbf{y}_{1:T}} Q(\mathbf{y}_{1:T}) \log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) d\mathbf{y}_{1:T} \tag{2.91}$$

$$- \int_{\mathbf{y}_{1:T}} Q(\mathbf{y}_{1:T}) \log Q(\mathbf{y}_{1:T}) d\mathbf{y}_{1:T} \tag{2.92}$$

$$= \mathcal{F}_{\text{opt}}(Q, \theta) - \mathcal{F}_{\text{entr}}(Q) \tag{2.93}$$

$$= \mathcal{F}(Q, \theta). \tag{2.94}$$

La fonction $\mathcal{F}(Q, \theta)$ est un minorant de $\log \mathcal{L}_z(\boldsymbol{\theta})$ quelle que soit la distribution $Q(\mathbf{y}_{1:T})$. Dans le cadre du filtrage récursif bayésien, la distribution Q est obtenue par filtrage (Kalman, Kalman étendu, filtre à particules, *etc.*) et constitue l'étape E de l'algortihme, puis sa maximisation, l'étape M. En initialisant les paramètres à une valeur $\boldsymbol{\theta}^{(0)}$, on obtient le schéma itératif suivant :

1. Étape E : calcul de la distribution $p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta}^{(i)})$ par filtrage bayésien récursif.
2. Étape M : mise à jour des paramètres $\boldsymbol{\theta}^{(i+1)}$ en résolvant

$$\boldsymbol{\theta}^{(i+1)} = \arg\max_{\boldsymbol{\theta}^{(i)}} \int \log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}^{(i+1)}) p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta}^{(i)}) d\mathbf{y}_{1:T}. \tag{2.95}$$

Ces étapes sont itérées jusqu'à l'obtention de la convergence. Selon les cas, on peut préférer utiliser l'algorithme de gradient ou l'algorithme EM. L'algorithme de gradient peut être plus rapide, et souvent préférable quand l'étape M de l'algorithme EM ne peut être résolue analytiquement. *A contrario*, l'algorithme EM est plus facile à gérer quand l'étape M peut-être calculée analytiquement et peut apporter des garanties supplémentaires. Notamment dans le cas linéaire et gaussien, $\mathcal{F}_{\text{opt}}(Q, \boldsymbol{\theta})$ peut s'exprimer analytiquement [Mader et al., 2014] comme

$$\begin{aligned}\mathcal{F}_{\text{opt}}(Q, \boldsymbol{\theta}) = &-\frac{T}{2} log(\boldsymbol{\Sigma}_{\mathbf{w_z}}) - \frac{1}{2}\text{tr}(\boldsymbol{\Sigma}_{\mathbf{w_z}}{}^{-1} \mathbb{E}_Q(\sum_{t=1}^{T} (\mathbf{z}_t - \mathbf{H}\mathbf{y}_t)(\mathbf{z}_t - \mathbf{H}\mathbf{y}_t)^T)) \\ &- \frac{T}{2} log(\boldsymbol{\Sigma}_{\mathbf{w_y}}) - \frac{1}{2}\text{tr}(\boldsymbol{\Sigma}_{\mathbf{w_y}}{}^{-1} \mathbb{E}_Q(\sum_{t=2}^{T} (\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1} - \mathbf{G}\mathbf{c}_t)(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1} - \mathbf{G}\mathbf{c}_t)^T))\end{aligned} \tag{2.96}$$

et une mise-à-jour formelle est alors possible car les espérances de cette expression peuvent se calculer à partir des moments obtenus lors du filtrage.

À ce sujet, Xu & Jordan [1996] ont démontré que, dans le cas gaussien, les pas de l'algorithme du gradient et de l'algorithme EM diffèrent à une projection linéaire près. Pour finir, il est à noter que les méthodes de gradient et EM convergent en général vers un maximum local et sont plus coûteuses en temps de calcul que les méthodes bayésiennes [Kantas et al., 2009].

Deuxième partie

Formulation du problème

3

Formulation Générale

La formulation du problème de contrôle acoustique d'un parc éolien constitue la première contribution de la thèse où un travail d'appropriation du contexte technique et scientifique et un rattachement à l'état de l'art est proposé. Dans ce chapitre on définit ainsi formellement le problème général de contrôle acoustique d'un parc éolien.

Ce problème a été formulé au départ par Venathec SAS de la manière suivante : « maximiser la production électrique totale en satisfaisant la réglementation française actuelle ». Notons à ce sujet que la réglementation française évolue rapidement, et le problème présenté perd dans ce chapitre son caractère général. Cependant, dans la suite du manuscrit, on verra qu'un bon nombre de résultats ont un caractère général que pour la seule réglementation française, notamment l'algorithme *branch and bound* proposé, et l'estimation et le suivi du modèle par filtrage et apprentissage.

Dans ce chapitre, on présente donc tout d'abord dans la partie 3.1, les notations introduites dans ce contexte applicatif, pour ensuite présenter la formulation originelle du problème dans le cas particulier français 3.2.

3.1 Notations et analyse du problème

On présente tout d'abord les notations proposées pour modéliser le problème de contrôle, établies à partir de la compréhension de l'ingénierie acoustique et dorénavant mises en forme dans un cadre rattachable à l'état de l'art du domaine du contrôle. Dans chaque sous-partie, on présente les notations du problème et une explication brève de celles-ci, établies à partir du modèle acoustique utilisé dans l'ingénierie acoustique.

Dans tout le reste de la thèse, on considère les conventions suivantes :

— Les grandeurs acoustiques sont données en dB pondérés A (dBA), c'est-à-dire comme les niveaux d'énergie en décibels des signaux pondérés fréquentiellement par la pondération A présentée dans la partie 1.4.3, qui constitue une modélisation de l'oreille humaine. Dans la suite, on parlera ainsi indifféremment de décibels ou de décibels pondérés A sans contradiction ou perte de généralité dans la mesure où tout se passe comme si tous les signaux acoustiques étaient préalablement traités par un filtre « A » avant d'être mesurés en décibels.

— Pour une grandeur acoustique X en dBA, on note $\tilde{X}$ la valeur correspondante en pression acoustique au carré. On a donc : $X = 10\log_{10}(\frac{\tilde{X}}{p_0^2})$, où p_0 est la valeur de référence de pression acoustique ($p_0 = 2,10^{-5}$Pa).

— On considère sur un parc éolien donné, I éoliennes indexées par $i \in \{1, ..., I\}$ et J points de contrainte indexés par $j \in \{1, ..., J\}$. Les points de contrainte sont les points géographiques autour du parc éolien où se situent les habitations et où la contrainte acoustique doit être respectée. Il s'agit également des points où des sonomètres (microphones de mesure) sont placés.
— L'unité de temps est indexée par t. Dans la pratique, cette unité de temps correspond à 10 min de mesure. Le choix de cette unité de temps est standard dans le domaine éolien [Giebel et al., 2011]. Il s'agit dans ce projet de la plus petite unité de temps considérée.
— Pour toutes les notations utilisées, on adopte les conventions suivantes :
  — les scalaires sont en **minuscules**
  — les vecteurs sont en **minuscules** et en **gras**,
  — les matrices sont en **majuscules** et en **gras**.
  Par exemple, la puissance acoustique d'émission de l'éolienne i à l'instant t se note $x_{t,i}$, le vecteur des émissions à ce même instant se note $\mathbf{x}_t$ et la matrice des atténuations $\mathbf{A}_t$.

3.1.1 Analyse des variables du problèmes

Variables d'état

Afin de pouvoir contrôler le système, il est nécessaire de définir l'ensemble des variables d'état qui évoluent dans le temps et qui sont nécessaires au calcul d'une commande optimale, c'est-à-dire correspondant à la variable $\mathbf{y}_t$ dans l'état de l'art établi au chapitre 2. Bien entendu, plusieurs choix sont possibles, et le modèle qu'on propose est justifié dans la suite de la thèse. Dans ce vecteur d'état, on ne fait figurer que les variables acoustiques. On utilise pour ces variables les notations suivantes :

— $\mathbf{x}_t = [x_{t,i}]$ de taille $I \times 1$: puissance instantanée d'émission acoustique de chaque éolienne i moyenne entre t et $t+1$, dont la valeur dépend du mode de bridage sélectionné et du vent,
— $\mathbf{A}_t = [a_{t,i,j}]$ de taille $I \times J$: atténuation acoustique liée à la propagation entre chaque éolienne i et chaque point contraint j moyenne entre t et $t+1$ (dépend de la vitesse du vent, des conditions météo, *etc.*),
— $\mathbf{r}_t = [r_{t,j}]$ de taille $J \times 1$: pression acoustique de bruit résiduelle pour chaque point contraint j moyenne entre t et $t+1$.

En ce qui concerne les puissances électriques de production, elles ne figurent pas dans les variables d'état. On considère ces variables comme étant fonction de l'émission acoustique des éoliennes. Pour éviter une redondance, on ne les inclut pas directement dans les variables d'état. On les note néanmoins :

— $\mathbf{p}_t = [p_{t,i}]$ de taille $I \times 1$: puissance électrique en kW de chaque éolienne i moyenne entre les temps t et $t+1$.

Certaines variables, comme la vitesse du vent, ne sont pas considérées non plus comme des variables d'état dans le modèle proposé. Elles sont utilisées néanmoins dans le modèle, mais de façon indirecte considérée comme connues. L'utilisation de ces variables dans les variables d'état, pourra faire l'objet de développements ultérieurs (voir partie V).

Variables acoustiques intermédiaires

Les variables $\mathbf{x}_t$ et $\mathbf{A}_t$ permettent de définir complètement les niveaux particuliers partiels $l_{t,i,j}$ et totaux $l_{t,j}$ au niveau des habitations. Tout d'abord, l'équation

$$l_{t,i,j} = x_{t,i} + a_{t,i,j} \tag{3.1}$$

décrit simplement l'atténuation de puissance acoustique entre la source sonore et le point de réception. Ce modèle simple est utilisé fréquemment en acoustique et notamment dans le domaine de l'éolien [Rogers et al., 2006]. Le niveau particulier total est donné par la somme logarithmique [Hansen & Sehrndt, 2001]

$$l_{t,j} = 10 \log_{10} \left(\sum_{i=1}^{I} 10^{\frac{l_{t,i,j}}{10}} \right), \tag{3.2}$$

ce qui est valide étant donnée l'absence de corrélation entre les signaux. De manière similaire, on calcule le niveau ambiant $b_{t,j}$ comme la somme logarithmique du niveau particulier $l_{t,j}$ et du niveau résiduel $r_{t,j}$:

$$b_{t,j} = 10 \log_{10} \left(10^{\frac{l_{t,j}}{10}} + 10^{\frac{r_{t,j}}{10}} \right). \tag{3.3}$$

On définit alors ces variables comme des variables intermédiaires de calcul :

- $\mathbf{b}_t = [b_{t,j}]$ de taille $J \times 1$: bruit ambiant perçu à chaque point contraint j moyen entre t et $t+1$,
- $\mathbf{l}_t = [l_{t,j}]$ de taille $J \times 1$: pression acoustique instantanée totale de réception sur chaque point contraint j moyenne entre t et $t+1$,
- $\mathbf{L}_t = [l_{t,i,j}]$ de taille $I \times J$: pression acoustique acoustique particulière de réception pondérée A de chaque éolienne i sur chaque point contraint j moyenne entre t et $t+1$.

Commandes de bridage

Un des intérêts principaux de décrire le niveau particulier comme la somme d'une émission et d'une atténuation, se situe vis-à-vis de la commande du système. L'émission acoustique est dépendante à la fois des conditions météorologiques (vitesse du vent principalement), mais aussi de la commande de bridage de chaque éolienne alors que l'atténuation ne dépend que des conditions météorologiques. Chaque éolienne peut être affectée d'une commande discrète différente, dont les caractéristiques dépendent du type de l'éolienne et qui est fixé au cours d'un pas de temps. On les note :

- $\mathbf{c}_t = [c_{t,i}]$ de taille $I \times 1$: modes de bridage sélectionnés pour chaque éolienne i entre t et $t+1$.

Le nombre de commandes utilisable pour chaque éolienne dépend du constructeur de l'éolienne et de son modèle. Différentes éoliennes peuvent se trouver sur un même parc mais, dans la suite, on considérera que toutes les éoliennes ont le même nombre de modes de bridage. En effet, les résultats peuvent s'étendre sans difficulté avec des éoliennes différentes sur le parc.

Les constructeurs éoliens fournissent en général des courbes d'émission acoustique et de production électrique. En théorie, les valeurs d'émission et de production dépendent de variables comme la vitesse du vent, la température, l'humidité et la pression. Cependant, en acoustique, seule la vitesse du vent est prise en compte [Norme, 1996]. D'autre part, pour une vitesse de vent donnée, la production électrique est une fonction croissante de l'émission acoustique. On définit alors les données du problème suivantes :

— $\mathcal{C} = \{1, 2, ..., C\}$: ensemble des C entiers qui désignent les modes de bridage des éoliennes d'un parc. 1 désigne le mode *full power* (puissance acoustique d'émission et puissance de production maximale) et C désigne l'arrêt des éoliennes. Les bridages $\{2, ..., C-1\}$, désignent les modes de bridage par valeurs décroissantes de puissance de production/d'émission. $\mathcal{C}$ est l'ensemble des valeurs possibles pour $c_{t,i}$.

— $\mathfrak{p}$: fonction qui représente les données constructeur sur l'énergie de production et dépend de la vitesse du vent et du mode de bridage. D'un point de vue mathématique, il s'agit d'une fonction définie par :

$$\begin{aligned} \mathfrak{p} : \mathcal{C} \times \mathcal{M} \times \Omega &\to \mathbb{R} \\ (c, v, \omega) &\mapsto \mathfrak{p}(c, v, \omega). \end{aligned}$$

Dans le cadre général, $\mathfrak{p}(c, v)$ est une variable aléatoire désignant la puissance électrique. On utilise la variable $\omega \in \Omega$ pour prendre en compte l'aspect aléatoire du problème. Pour ne pas alourdir les notations on n'écrira plus cette variable dans le problème. Dans sa simplification déterministe, $\mathfrak{p}(c, v)$ est un réel. $\mathfrak{p}_v$ et $\mathfrak{p}_c$ désignent les fonctions partielles de $\mathfrak{p}$ définies par $\mathfrak{p}_v = \mathfrak{p}(., v)$ et $\mathfrak{p}_c = \mathfrak{p}(c, .)$.

— $\mathfrak{x}$: fonction qui représente les données sur la puissance acoustique d'émission et dépend de la vitesse vent et du mode de bridage. D'un point de vue mathématique, il s'agit d'une fonction définie par :

$$\begin{aligned} \mathfrak{x} : \mathcal{C} \times \mathcal{M} \times \Omega &\to \mathbb{R} \\ (c, v, \omega) &\mapsto \mathfrak{x}(c, v, \omega). \end{aligned}$$

Dans le cadre général, $\mathfrak{x}(c, v)$ est une variable aléatoire réelle désignant la puissance acoustique. On utilise la variable $\omega \in \Omega$ pour prendre en compte l'aspect aléatoire du problème. Pour ne pas alourdir les notations on n'écrira plus cette variable dans le problème. D'autre part, dans sa simplification déterministe, $\mathfrak{x}(c, v)$ est un réel. $\mathfrak{x}_v$ et $\mathfrak{x}_c$ désignent les fonctions partielles de $\mathfrak{x}$ définies par $\mathfrak{x}_v = \mathfrak{x}(., v)$ et $\mathfrak{x}_c = \mathfrak{x}(c, .)$.

— $\mathfrak{F}$: fonction qui associe la puissance théorique de production $\mathbf{p}_t$ à partir d'un vecteur d'émission exprimé en échelle linéaire $\tilde{\mathbf{x}}_t$. Cette fonction dépend de la vitesse du vent, supposée connue à chaque instant t.

Variables météorologiques

Les variables d'état acoustiques et la production électrique dépendent de la commande de bridage et des conditions météorologiques. L'émission et la propagation dépendent de la vitesse et la direction du vent, de la température, de l'humidité et de la pression. Dans une moindre mesure, on constate également une dépendance du niveau résiduel à la vitesse du vent. Ces variables ne sont pas considérées dans notre modèle comme des variables d'état, mais comme des variables mesurées et influant sur la valeur et les incertitudes. On utilise les notations suivantes :

— $\mathbf{v}_t = [v_{t,i}]$: vitesse du vent de chaque éolienne à hauteur de nacelle entre t et $t+1$,

— $\mathbf{d}_t = [d_{t,i}]$: direction du vent de chaque éolienne entre t et $t+1$.

Classes des mesures

On a présenté brièvement, dans le chapitre 1, la manière de procéder dans le domaine de l'ingénierie acoustique lors des études réglementaires. On isole en général les conditions environnementales influant les variables acoustiques, de sorte que le niveau acoustique ne soit fonction

que d'une commande qu'on maîtrise. En pratique ce découpage se constitue de classes homogènes, c'est-à-dire d'une partition discrète des conditions environnementales. Dans l'éolien en France, il s'agit des périodes de la journée, de la vitesse et de la direction du vent :

— $\mathcal{P} = \{p_1, p_2\} = \{\texttt{jour}, \texttt{nuit}\}$: ensemble des classes sur les périodes de la journée. La classe `jour` (notée p_1) définit les mesures entre 7h et 22h, et la classe `nuit` (noté p_2), les mesures entre 22h et 7h.

— $\mathcal{D} = \{d_1, d_2\}$: ensemble des classes sur les secteurs de vent. Dans la pratique, 2 ou 3 classes aux maximum sont considérées, les directions de vent étant en général fortement réparties sur un petit nombre de directions. Chaque élément d_k définit un secteur de vent de 60 degrés.

— $\mathcal{M} = \{m_3, m_4, ..., m_{10}\}$: ensemble des classes sur le module de la vitesse de vent. Chaque élément m_k définit un intervalle centré sur la valeur k (en m/s) et de taille 1 m/s. En général, on considère les vitesses de 3 m/s à 10 m/s.

Notons que dans le cadre du contrôle acoustique, on considère une partition complète des valeurs possibles, afin notamment de prendre en compte toutes les directions d'angle possibles (et pas seulement 2 ou 3 secteurs d'angles de 60 degrés). Ce découpage reste néanmoins caractéristique de la législation française, et on a vu qu'il existe en pratique une dépendance à d'autres variables, comme la température. Pour conserver un caractère général on utilise la notation suivante pour décrire l'ensemble des classes homogènes :

— $\mathcal{O}$: ensemble des classes homogènes. Chaque élément de $\mathcal{O}$ est noté o_k. Dans le cas de la législation française, on considère $\mathcal{O} = \mathcal{P} \times \mathcal{D} \times \mathcal{M}$ et les éléments o_k sont indexés linéairement.

Ces classes homogènes définissent ainsi et de manière générale un découpage des conditions de l'environnement pour lesquelles on suppose que la distribution des variables acoustique est peu étalée (pour une commande donnée).

Incertitudes sur le modèle

Dans le cadre d'un contrôle robuste, on propose de définir ici des variables d'incertitude sur les variables d'état.

Pour deux vecteurs aléatoires $\mathbf{x}$ et $\mathbf{y}$ quelconques (resp. deux variables aléatoires x et y), on note $\mathbb{E}(\mathbf{x}) = \mathbf{m_x}$ son vecteur espérance (resp. m_x son espérance) et $\mathrm{Cov}(\mathbf{x}) = \mathbf{\Sigma_x}$ sa matrice de covariance et $\mathrm{Cov}(\mathbf{x}, \mathbf{y}) = \mathbf{\Sigma_{x,y}}$ la matrice de covariance croisée entre $\mathbf{x}$ et $\mathbf{y}$ (resp. σ_x^2 sa variance, et $\sigma_{x,y}$ la covariance entre x et y).

On note également $\mathbf{w_x}$ le vecteur des incertitudes associées au vecteur aléatoire gaussien $\mathbf{x}$ vérifiant $\mathbf{w_x} = \mathbf{x} - \mathbf{m_x}$. Dans le cas où $\mathbf{x}$ est gaussien, ce vecteur suit donc la loi $\mathbf{w_x} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_x})$.

On considère donc dans la suite les 3 variables d'incertitude suivantes :

— $\mathbf{w}_{\mathbf{x}_t}$: vecteur des incertitudes sur $\mathbf{x}_t$,

— $\mathbf{w}_{\mathbf{a}_t}$: vecteur des incertitudes sur $\mathbf{a}_t$,

— $\mathbf{w}_{\mathbf{r}_t}$: vecteur des incertitudes sur $\mathbf{r}_t$.

Variables de mesures et incertitudes de mesure

Outre les incertitudes sur le modèle, on considère, dans le modèle présenté, des incertitudes propres à la mesure des variables acoustiques. Dans le paradigme choisi pour le contrôle, on considère deux types de mesures :

— $\mathbf{b}_t^{*T}$: vecteur des mesures sonométriques aux J points de mesure moyennes entre t et $t+1$,

— $\mathbf{r}_t^{*\bullet T}$: vecteur des estimations du bruit résiduel aux J points de mesure moyennes entre t et $t+1$.

Le système de contrôle se base en effet sur une estimation par séparation de sources qui estime le niveau particulier et le niveau résiduel. En outre, le sonomètre permet de mesurer simplement le niveau ambiant par mesure directe du signal acoustique. Ces variables sont notées différemment, car elles diffèrent des valeurs réelles du niveau ambiant et du niveau résiduel via une incertitude :

— $\mathbf{w}_{\mathbf{b}_t^{*T}}$: incertitude sonométrique (liée à la mesure du sonomètre).
— $\mathbf{w}_{\mathbf{r}_t^{*\bullet T}}$: incertitude de l'estimation du résiduel, due à l'incertitude sonométrique et l'incertitude de la séparation de source.

Concaténation de variables et matrice de propagation vectorisée

Pour inclure l'ensemble des coefficients de $\mathbf{A}_t$ dans le vecteur des variables d'état, on adopte une notation supplémentaire :

— $\mathbf{a}_t = vec(\mathbf{A}_t)$. C'est-à-dire que $\mathbf{a}_t$ est le vecteur désignant la juxtaposition verticale des colonnes de $\mathbf{A}_t$.

On utilisera ainsi les notations suivantes pour le vecteur d'état et le vecteur des observations :

— $\mathbf{y}_t$: vecteur d'état des variables acoustiques. En particulier, dans la partie IV, on considérera $\mathbf{y}_t = [\mathbf{x}_t^T \mathbf{a}_t^T \mathbf{r}_t^T]^T$ de taille $(I + I \times J + J) \times 1$.
— $\mathbf{z}_t$: vecteur des observations. En particulier, dans la partie IV, on considérera $\mathbf{z}_t = [\mathbf{b}_t^{*T} \mathbf{r}_t^{*\bullet T}]^T$ de taille $2J \times 1$, c'est-à-dire le vecteur des observations acoustiques issu de la concaténation des vecteurs de mesure de bruit ambiant et d'estimation du résiduel.

Seuils de limitation et seuils réglementaire

Dans le cadre du contrôle, on utilise des seuils représentant la limite du bruit particulier à ne pas dépasser à un instant t. Dans le cas général, on note ce seuil par :

— $\mathbf{q}_t$: vecteur de taille $J \times 1$ des seuils de bruit particulier à ne pas dépasser en chaque point j.

Dans le cas particulier français, on considère les seuils d'émergence et de bruit ambiant présentés au chapitre 1. Comme ils dépendent de la période de la journée actuellement, on ajoute une dépendance à la classe homogène o.

— $S_{\mathrm{e},o}$: seuil d'émergence pour les mesures de classe o. Il s'agit d'un seuil réglementaire que l'estimateur réglementaire de l'émergence ne doit pas dépasser. Il vaut 3 dBA pour des mesures de nuit, et 5 dBA pour des mesures de jour.
— $S_{\mathrm{s},o}$: seuil de déclenchement du critère d'émergence pour les mesures de classe o. Il s'agit d'un autre seuil réglementaire pour l'émergence. Il vaut 35 dBA. Si l'estimateur réglementaire du bruit ambiant est inférieur à ce seuil, le critère d'émergence n'est pas considéré.

Notation simplifiée lorsque l'indice temporel t est fixé

Pour le problème déterministe, on retire l'indice temporel lorsque le contexte le permet. $\mathbf{p}_t$ devient $\mathbf{p}$, $\mathbf{l}_t$ devient $\mathbf{l}$, et ainsi de suite afin d'alléger les notations.

3.2 Formulation originelle du problème de contrôle stochastique

On propose dans cette partie de formuler le problème comme un problème de contrôle stochastique. On se situe dans le cas de la législation française, mais on montrera que certains résultats restent valides dans un cadre plus général dans la suite du manuscrit.

3.2.1 Fonction objectif

Le but du contrôle acoustique des parcs éoliens est de maximiser la production électrique de l'ensemble du parc dans le temps, c'est-à-dire l'énergie produite par l'ensemble des éoliennes sur une période de temps $[0, T]$ qu'on précisera par la suite. En effet, le but est avant tout économique, et de ce fait, on ne s'intéresse pas à la production d'une éolienne en particulier à un instant précis, mais à la production de l'ensemble des éoliennes pendant une journée, une semaine, ou une année. De ce fait, la fonction à optimiser peut s'écrire comme la somme des espérances mathématiques de la production électrique de chaque éolienne depuis l'instant courant jusqu'à un horizon T. Dans la mesure où la production électrique $p_{t,i}$ ne peut-être connue de manière exacte, on la considère en effet comme une variable aléatoire dont la loi de probabilité dépend du choix $\mathbf{c}_t$ pris à l'instant t définissant le bridage et des variables d'état prédites du système à l'instant t. La fonction objectif à maximiser à un instant t_0 s'écrit donc

$$\sum_{t=t_0}^{t=t_0+T-1} \sum_{i=1}^{I} \mathbb{E}(p_{t,i}). \tag{3.4}$$

3.2.2 Contraintes

Le système doit maximiser la production sous la contrainte acoustique. Cependant il est difficile d'adapter le critère légal français directement pour le contrôle. Le critère est calculé lors des études acoustiques sur la base de cycles d'arrêts et redémarrage des machines. Or, comme décrit dans le chapitre 1, le critère légal se base sur des statistiques, calculées après avoir fini la campagne de mesurage : « Pour chaque habitation (point de mesure) et pour chaque classe, lorsque l'indicateur de bruit ambiant recalé est supérieur à 35 dBA, la différence entre l'indicateur de bruit ambiant recalé et l'indicateur de bruit résiduel recalé ne doit pas excéder 5 dBA de jour et 3 dBA de nuit ». Remarquons que, lorsqu'on parle de chaque habitation, on ne considère en pratique qu'un nombre de points relativement faible, les J habitations les plus exposées autour du site. On rappelle également que les indicateurs de bruit ambiant et résiduel sont respectivement les médianes empiriques des bruits ambiants et résiduels considérés par classe homogène et sont calculées avec un minimum de 10 mesures. Le nombre de mesures utilisées en pratique dépend de la campagne de mesurage qui s'étend en général sur une ou deux semaines. La taille de la fenêtre glissante N devra être choisie en tenant compte de cet aspect. Notons pour finir que le recalage des indicateurs est donné par une formule définie dans la norme NFS 31-114 [Norme, 1996]. Cependant, dans le cadre de ce projet, les indicateurs utilisés ne sont pas les indicateurs recalés, les incertitudes étant considérées plus précisément au chapitre 7.

On classe ainsi l'ensemble des mesures correspondant à chaque classe o et un indicateur de bruit ambiant et de bruit résiduel est calculé comme la médiane des mesures observées au sein de chaque classe. On vérifie ensuite que le critère légal est respecté dans chaque classe. Afin de déterminer un critère de contrainte pour le contrôle acoustique, on a formulé le problème de la manière suivante : lors d'un contrôle acoustique traditionnel, impliquant des cycles marche-arrêt, le critère légal doit être respecté. En d'autres termes, il convient de définir un critère de contrainte de telle sorte que le critère légal soit vérifié *a posteriori*.

On considère donc, à un instant t_1, la distribution $b_{j,t_1-N+1:t_1}$ constituée des $b_{j,t}$ pour $t \in \{t_1 - N + 1, ..., t_1\}$ et $o_t = o_{t_1}$. On note la médiane de cette distribution $\hat{b}_{[t_1-N+1,t_1],j}$, ou encore $\hat{b}_{t_1,j}$ (en omettant N supposé fixé). Les contraintes peuvent alors s'écrire :

$$\begin{aligned} &\forall t \in \{t_0, ..., t_0 + T - 1\}, \forall o \in \mathcal{O}, \forall j \in \{1, ..., J\}, \\ &\hat{b}_{t,j} - \hat{r}_{t,j} < S_{e,o} \\ &\text{ou } \hat{b}_{t,j} < S_{s,o}. \end{aligned} \tag{3.5}$$

3.2.3 Formulation complète

En résumé, le problème général au temps t_0 pour la classe o peut alors s'écrire :

$$\begin{cases} \max_{(\mathbf{c}_{t_0}, ..., \mathbf{c}_{t_0+T-1}) \in \mathcal{C}^{I \times T}} \sum_{t=t_0}^{t=t_0+T-1} \sum_{i=1}^{I} \mathbb{E}(\mathfrak{p}_v(c_{t,i})) \\ \forall t \in \{t_0, ..., t_0 + T - 1\}, \forall j \in \{1, ..., J\}, \\ \hat{b}_{t,j} - \hat{r}_{t,j} < S_{e,o} \\ \text{ou } \hat{b}_{t,j} < S_{s,o}. \end{cases} \tag{3.6}$$

Il s'agit ainsi d'un problème de contrôle stochastique à horizon fini [Bertsekas, 1976]. Implicitement, cette formulation met en avant un concept clé des problèmes de contrôle : la décision prise à l'instant présent t_0 a un impact sur les états futurs du système entre $t_0 + 1$ et $t_0 + N$. En effet, l'état du système à un instant donné est dépendant des choix du passé. L'optimisation électrique du système jusqu'à l'horizon considéré ne peut alors *a priori* pas être traitée comme l'optimisation individuelle de chaque instant. La résolution d'un tel problème nécessite la résolution de plusieurs sous-problèmes. Tout d'abord, un modèle de prédiction est nécessaire afin de pouvoir considérer les états futurs du système mais aussi la connaissance la plus précise du modèle. En effet, même sans parler de prédiction, les modèles numériques étant imparfaits et les sources d'influences sur les variables acoustiques nombreuses, l'estimation des biais et de la distribution des erreurs est nécessaire. Ensuite, vient la résolution du problème à proprement parler. Il s'agit d'un problème d'optimisation combinatoire. L'arbre des solutions augmente à la fois avec le nombre d'éoliennes sur le parc, le nombre de modes de bridage et l'horizon du problème.

3.3 Simplification du problème et terminologie

Dans la suite, des versions simplifiées du problème sont étudiées. La complexité du problème général amène à poser des hypothèses simplificatrices. Pour désigner ces problèmes, on utilise le vocabulaire suivant :

— **Problème stochastique** qui s'oppose à **problème déterministe** : Les incertitudes de mesure et la réactivité des machines sont à prendre en compte dans le problème général. De ce fait, le problème général est dit stochastique, ce qui signifie mathématiquement qu'un certain nombre de variables sont modélisées par des variables aléatoires. Au contraire, la simplification déterministe du problème consiste à supposer que les valeurs sont exactes et on les modélise alors comme des variables réelles. Dans ce problème, ce sont les variables acoustiques, qui seront « mesurées » par un système de séparation de sources, mais

également les variables de production électrique ou d'émission acoustique ou les variables météorologiques (vent, température, *etc.*).

— **Problème à contraintes sur les médianes temporelles** qui s'oppose à **problème instantané** : Le problème général a des contraintes acoustiques (3.5) qui s'écrivent à partir des médianes des distributions des mesures acoustiques. Une simplification du problème, appelée problème instantané, consiste à remplacer ces contraintes par des contraintes plus simples qui portent sur les valeurs instantanées des mesures acoustiques.

— **Problème prédictif** qui s'oppose à **problème non-prédictif** : Afin d'anticiper les changements d'état du système, on pourra s'appuyer sur des prédictions des valeurs futures des paramètres météorologiques et environnementaux. Cependant, il convient d'étudier la nécessité de ces prédictions et on désigne alors par problème non-prédictif la simplification du problème sans utilisation de prédictions.

Dans la suite, on commence par résoudre le problème déterministe instantané (qui est nécessairement non-prédictif également). Ce problème constitue l'objet du chapitre 4. Ensuite, on s'intéresse au problème à contraintes sur les médianes temporelles, mais toujours déterministe et non-prédictif. Ce second problème constitue l'objet du chapitre 5.
Le problème stochastique et prédictif est étudié dans la partie IV.

Troisième partie

Problème déterministe

4

Problème déterministe à contraintes instantanées

4.1 Introduction

Dans cette partie, on s'intéresse au cas du problème déterministe où on ne considère aucune incertitude de mesure sur les bruits particuliers et résiduels. Les valeurs sont considérées exactes. En particulier, on ne considère aucune incertitude sur les puissances d'émission et de production des éoliennes. On a une unique valeur d'émission acoustique et de production électrique pour un mode et une vitesse de vent perçue pour chaque éolienne. D'autre part, on ne considère aucune incertitude de propagation acoustique dans une classe homogène donnée.

Dans ce premier chapitre, on considère un critère d'émergence instantanée qui considère la résolution du problème, non pas comme un problème à horizon fini portant sur les estimateurs basés sur des médianes temporelles des émergences mais comme un problème où le but est de maximiser la production du parc à un instant donné, contraint par les émergences instantanées à cet instant. On développe donc une formulation théorique du problème, un algorithme optimal et flexible en temps de calcul, une preuve de son optimalité et une vérification expérimentale.

On verra ultérieurement, dans le chapitre 5 la possibilité de contrôle d'un parc éolien sans utilisation de prédiction et sans incertitudes mais en intégrant les contraintes basées sur des médianes temporelles. On y présentera en outre une analyse du compromis entre la production électrique et le respect de le contrainte réglementaire.

4.2 Formulation

Dans cette partie, on considère le problème le plus simple, sans incertitudes et où les contraintes s'écrivent :

$$\forall t \in \{t_0, ..., t_0 + T - 1\}, \forall o \in \mathcal{O}, \forall j \in \{1, ..., J\}, b_j - r_j < S_{\mathrm{e},o} \text{ ou } b_j < S_{\mathrm{s},o}. \tag{4.1}$$

Le problème déterministe est un problème d'optimisation discrète dont le but est de trouver les modes de bridage optimaux en terme de production électrique tout en respectant les contraintes d'émergence instantanée. On se place donc à temps t fixé (on omet alors les indices temporels). Le vecteur des niveaux résiduels $\mathbf{r}$ à cet instant est supposé connu et estimé par séparation de sources, supposée exacte. D'autre part, les conditions météorologiques sont supposées mesurées exactement par le SCADA éolien donc par conséquent la classe homogène o à cet instant est

supposée connue. Le problème d'optimisation déterministe instantané peut alors s'écrire de la manière suivante :

$$\begin{cases} \max_{\mathbf{c}\in\mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_i) \\ \forall j \in \{1,...,J\}\ b_j - r_j \leq S_{\mathrm{e},o} \text{ ou } b_j \leq S_{\mathrm{s},o}. \end{cases} \tag{4.2}$$

Dans ce problème, la fonction objectif (la production électrique) et les termes situés à gauche dans les inégalités ne dépendent plus que de la commande. Concernant la fonction objectif, il s'agit de la puissance électrique totale du parc à l'instant t qui ne dépend que de la commande puisque la vitesse du vent est connue et fixée. La puissance peut être alors déterminée à partir des données constructeurs. Concernant les contraintes, le niveau ambiant et l'émergence acoustique ne dépendent également que de la commande. En effet, on peut calculer ces quantités à partir de l'émission acoustique $\mathbf{x}$, de l'atténuation due à la propagation $\mathbf{A}$, du niveau résiduel $\mathbf{r}$ et des équations acoustiques (3.1), (3.2) et (3.3) présentées dans la sous-partie 3.1.1. Seule l'émission acoustique, donnée par les données constructeurs de façon similaire à la puissance électrique, dépend de la commande. En effet, la propagation est entièrement déterminée par la classe homogène o et obtenue par simulation numérique, et le résiduel est supposé estimé de façon exacte par le système de séparation de source.

Notons qu'un problème relativement similaire en ingénierie acoustique concerne le calcul de plans de bridage. Cependant, il n'existe pas à notre connaissance de formalisme pour la résolution de ce problème. Les algorithmes utilisés sont des heuristiques qui ne prennent généralement pas en compte la puissance électrique comme paramètre d'optimisation et maximisent la somme des émissions acoustiques. Il n'y a donc en pratique aucune preuve théorique ou expérimentale de l'optimalité des algorithmes.

L'algorithme développé pour résoudre ce problème constitue une base de départ pour le problème à contraintes sur les médianes temporelles, dans la mesure, où comme on le verra **il permet en outre d'optimiser la puissance électrique d'un parc pour des contraintes de bruit particulier maximal quelconque en chaque point.**

4.2.1 Généralisation du problème

Afin d'obtenir une formulation mathématique plus adaptée du problème, on transforme le problème (4.2) en opérant un changement de variables pour les grandeurs acoustiques jusqu'alors exprimées en dB, en les écrivant sous leur forme naturelle en $\mathrm{W/m^2}$. Le problème (4.2) peut alors se ré-écrire

$$\begin{cases} \max_{\mathbf{c}\in\mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_i) & \\ \quad \forall j \in \{1,...,J\} & 10\log_{10}\left(\frac{\sum_i^I \tilde{l}_{i,j}+\tilde{r}_j}{p_0^2}\right) - r_j \leq S_{\mathrm{e},o} \\ & \text{ou } 10\log_{10}\left(\frac{\sum_i^I \tilde{l}_{i,j}+\tilde{r}_j}{p_0^2}\right) \leq S_{\mathrm{s},o} \end{cases} \tag{4.3}$$

$$\Leftrightarrow \begin{cases} \max_{\mathbf{c}\in\mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_i) & \\ \quad \forall j \in \{1,...,J\} & \sum_{i=1}^{I} \tilde{l}_{i,j} \leq p_0^2 10^{\frac{S_{\mathrm{e},o}+r_j}{10}} - \tilde{r}_j \\ & \text{ou } \sum_{i=1}^{I} \tilde{l}_{i,j} \leq p_0^2 10^{\frac{S_{\mathrm{s},o}}{10}} - \tilde{r}_j. \end{cases} \tag{4.4}$$

On pose alors $\mathbf{q} = [q_j]$ tel que $\forall j \in \{1, ..., J\}, q_j = \max(p_0^2 10^{\frac{S_{e,o}+r_j}{10}} - \tilde{r}_j, p_0^2 10^{\frac{S_{s,o}}{10}} - \tilde{r}_j)$. On a :

$$(4.2) \Leftrightarrow \begin{cases} \max_{\tilde{\mathbf{x}} \in \tilde{\mathfrak{x}}(\mathcal{C})^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_i) \\ \\ \quad \forall j \in \{1, ..., J\} \qquad \sum_{i=1}^{I} \tilde{a}_{i,j} \tilde{x}_i \leq q_j \end{cases} \tag{4.5}$$

$$\Leftrightarrow \begin{cases} \max_{\tilde{\mathbf{x}} \in \tilde{\mathfrak{x}}(\mathcal{C})^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_i) \\ \\ \quad \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}. \end{cases} \tag{4.6}$$

On obtient une formulation linéaire pour les contraintes. Pour autant, la forme du problème n'est pas encore entièrement satisfaisante dans la mesure où la dépendance entre les contraintes et la variable à optimiser n'apparaît pas de manière explicite. On propose donc de considérer comme variable de choix la variable $\tilde{\mathbf{x}}$ à la place de la variable $\mathbf{c}$. Cette considération permet de reformuler le problème sans perte de généralité, puisque ces deux variables sont liées par une bijection, la puissance électrique étant une fonction croissante de l'émission acoustique lorsque la vitesse du vent est fixée. Pour ce faire, on définit au préalable la fonction $\mathfrak{f}$ qui associe la puissance acoustique en W à la puissance de production de chaque éolienne par

$$\begin{aligned} \mathfrak{f} : \mathbb{R} &\to \mathbb{R} \\ \tilde{x}_i &\mapsto \mathfrak{p}_v(\tilde{\mathfrak{x}}_v^{-1}(\tilde{x}_i)). \end{aligned} \tag{4.7}$$

On définit également la puissance de production totale $\mathfrak{F}$, comme la fonction vérifiant

$$\forall \tilde{\mathbf{x}} \in \tilde{\mathfrak{x}}_v(\mathcal{C})^I, \mathfrak{F}(\tilde{\mathbf{x}}) = \sum_{i=1}^{I} \mathfrak{f}(\tilde{x}_i). \tag{4.8}$$

En remplaçant dans (4.6) on obtient

$$\begin{cases} \max_{\tilde{\mathbf{x}} \in \tilde{\mathfrak{x}}_v(\mathcal{C})^I} \sum_{i=1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}, \end{cases} \tag{4.9}$$

ou encore

$$\begin{cases} \max_{\tilde{\mathbf{x}} \in \tilde{\mathfrak{x}}_v(\mathcal{C})^I} \mathfrak{F}(\tilde{\mathbf{x}}) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}. \end{cases} \tag{4.10}$$

Cette formulation est particulièrement intéressante puisqu'elle permet de situer le problème dans un cadre théorique défini permettant de se rattacher plus facilement à l'état de l'art. Une vérification expérimentale a montré que la fonction $\mathfrak{f}$ est non-linéaire et concave (un exemple est donné dans la suite du chapitre). On peut alors affirmer que ce problème fait partie d'une classe de problèmes d'optimisation appelés problèmes de programmation entière non-linéaire convexe (*Convex Integer Non-linear Programming*) dans la littérature scientifique. Plus précisément, on peut noter qu'il s'agit d'un problème de maximisation non-linéaire sous contraintes linéaires.

D'autre part, avec ce formalisme, on constate que le problème appartient à une classe plus large de problèmes. Si on change la valeur de la variable $\mathbf{q}$, la forme du problème et la résolution de celui-ci restent inchangées. De ce fait, **le problème à émergence instantanée se ramène à un problème plus général d'optimisation avec des contraintes arbitraires de bruit particulier maximal en chaque point**. Cette propriété est exploitée dans le chapitre 5.

4.3 Algorithme branch and bound

Pour résoudre le problème d'optimisation présenté, on a développé un algorithme *branch and bound*. Cet algorithme est le plus général de l'état de l'art (voir chapitre 2) et présente plusieurs avantages. Tout d'abord, lorsqu'il termine et si le calcul de majorant est correct, l'algorithme garantit l'optimalité. Bien entendu, étant donné le caractère *NP-complet* du problème, le temps de calcul est potentiellement grand. Cependant, l'algorithme est *scalable* : il peut être arrêté avant d'arriver à son terme en fonction du temps de calcul imparti. Il construit, comme présenté dans l'état de l'art, une suite de solutions admissibles (qui respectent les contraintes) dont l'optimalité est croissante (la production s'améliore à chaque nouveau bridage calculé). On peut ainsi adapter le temps de calcul aux besoins. L'algorithme proposé prend en compte les valeurs réelles de puissance électrique et aussi les différents types d'éoliennes sur un même parc et différents vecteurs de vent.

4.3.1 Algorithme général

On présente la version créée pour le problème déterministe instantané. Deux variantes principales sont proposées. Ce qui est énoncé dans cette partie est commun aux deux variantes : les différences sont détaillées dans la partie suivante.

Construction de l'arbre

L'algorithme se base sur le parcours d'un arbre contenant initialement la totalité des solutions admissibles du problème. Le choix de l'arborescence de l'arbre est simple et constitue un découpage logique des solutions : chaque nouveau nœud constitue une unique possibilité de bridage pour une éolienne donnée. En profondeur, on trouve les différentes éoliennes du problème, et en largeur les modes possibles comme représenté dans la Figure 4.1.

Du point de vue de l'implémentation informatique, l'arbre est construit au fur et à mesure du déroulement de l'algorithme. Si cela n'a pas déjà été fait, on construit les fils du nœud courant, en ne conservant que les modes de bridage qui ne violent pas les contraintes acoustiques *a priori*. En d'autres termes, si la solution partielle que représente un potentiel nœud à créer viole les contraintes avant même d'avoir considéré ses nœuds fils, ce nœud n'est pas construit.

Majoration des noeuds et initialisation de l'algorithme

Deux variantes ont été construite. La première est une méthode lente mais exacte et jouant alors le rôle de vérité terrain pour la seconde qui constitue la méthode retenue. Elles diffèrent par leur initialisation et par leur méthode de majoration. Elles sont détaillées dans la partie 4.3.2

Variante « vérité terrain » :
- Solution initiale : plan de bridage *full-off*, c'est-à-dire $\forall i \in \{1, ..., I\}, c_i = C$.
- Majoration : à chaque nœud, le majorant est calculé comme la somme des puissances électriques des éoliennes déjà déterminées par ce nœud, plus celles du bridage *a priori* admissible de puissance maximale pour chacune des éoliennes restantes.

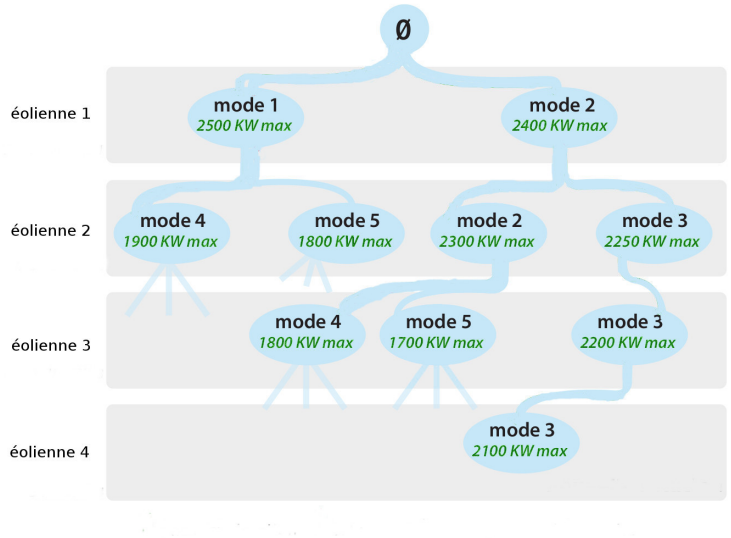


FIGURE 4.1 – Exemple d'arbre des solutions admissibles pour 4 éoliennes et 5 modes de bridage.

Variante « retenue » :
- Solution initiale : plan de bridage solution de l'algorithme glouton décrit dans la partie 4.4.4.
- Majoration : à chaque nœud, le majorant est calculé comme la somme des puissances électriques des éoliennes déjà déterminées par ce nœud, plus la puissance issue de la solution du problème de relaxation continue pour les éoliennes restantes, décrit à la partie 4.3.3.

Parcours de l'arbre

En ce qui concerne l'ordre de parcours, on choisit de le parcourir en profondeur, avec l'arbre ordonné de la manière suivante :

- En profondeur, les différentes éoliennes à brider sont classées par ordre décroissant d'impact acoustique en ordonnant les éoliennes i par rapport à leur valeur $\max_j a_{i,j}$.
- En largeur, les différents bridages admissibles pour l'éolienne considérée sont classés par ordre décroissant de majoration.

En considérant cet ordre-ci, l'arbre est parcouru en utilisant la technique de majoration présentée dans la partie 2.1.1 pour supprimer les branches où on a la certitude de ne pas trouver une meilleure solution que la solution courante. De ce fait, l'algorithme n'omet aucune solution, toutes les branches de l'arbre étant soit parcourues, soit élaguées ce qui garantit son optimalité.

Du point de vue de l'implémentation, les branches vont être physiquement supprimées en mémoire une fois examinées afin de ne pas les parcourir à nouveau et de fixer un critère d'arrêt simple : l'algorithme se termine quand l'arbre est réduit à sa racine. La Figure 4.2 présente un exemple de parcours sur un arbre pour 4 éoliennes et 5 modes de bridage, en partant d'une solution initiale de 2000 kW.

Le calcul des majorants des nœuds et l'initialisation de l'algorithme sont décrits dans la partie 4.3.2.

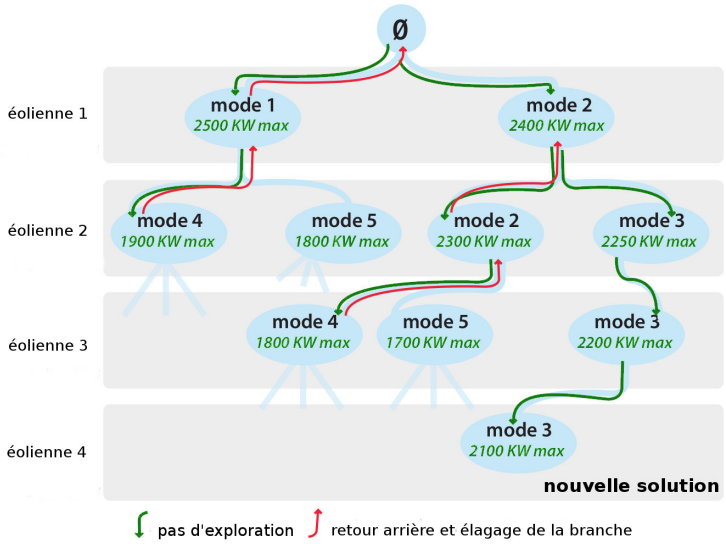


FIGURE 4.2 – Exemple de parcours sur un arbre pour 4 éoliennes et 5 modes de bridage, en partant d'une solution initiale de 2000 kW.

Pseudo code de l'algorithme

Voici les variables nécessaires à la description de cet algorithme :

— `noeud_courant` : variable pointant vers le nœud courant. Initialement, elle pointe vers la racine de l'arbre.

— `solution_temporaire` : variable contenant notamment les modes de bridage des éoliennes et la production électrique de la solution actuelle.

— `arbre` : variable pointant vers l'arbre. Au début de l'algorithme, cet arbre n'est constitué que d'un nœud, la racine de l'arbre.

Le pseudo code de l'algorithme est donné par l'Algorithme 1. Une autre description plus générale des algorithmes *branch and bound* peut être trouvée dans [Sakarovitch, 1984, P.135].

À chaque itération, on regarde d'abord le type de nœud sur lequel `noeud_courant` est positionné. S'il s'agit d'une feuille, cela signifie qu'on a trouvé une nouvelle solution et il faut alors la noter dans `solution_temporaire`, puis supprimer cette feuille et remonter à son père. Au contraire si `noeud_courant` n'est pas une feuille, il faut alors faire une analyse pour déterminer s'il faut s'enfoncer dans l'arborescence du nœud ou au contraire s'il faut élaguer l'arborescence et remonter au père du nœud.

Pour ce faire, il est tout d'abord nécessaire de construire les fils de `noeud_courant` si cela n'a pas déjà été fait auparavant. Auquel cas, on construit les fils de `noeud_courant`, on calcule leur majorant respectif, et on les ordonne par ordre décroissant.

Si la construction des fils de `noeud_courant` a déjà été faite par le passé ou que l'on a fini de construire ses fils, on peut alors passer à l'analyse du nœud :

— Si le premier fils de `noeud_courant` a un majorant plus grand que la production électrique de `solution_temporaire`, on s'enfonce alors dans l'arbre et ce fils devient alors la nouvelle valeur de `noeud_courant`.

— Si le premier fils de `noeud_courant` a un majorant plus faible que la production, il est suffisant de supprimer `noeud_courant` et de remonter à son père. En effet, si le premier fils de `noeud_courant` a un majorant plus faible que la production de `solution_temporaire`,

Algorithm 1 Algorithme *branch and bound* appliqué aux éoliennes

function BRANCHANDBOUND
  Initialisation de la racine de `arbre`
  `noeud_courant` pointe vers `arbre`
  Initialisation de la variable `solution_temporaire` par une solution admissible.
  while racine de `arbre` n'est pas majorée **OU** que `arbre` est non vide **do**
    if `noeud_courant` est une feuille **then**
      Noter la solution temporaire dans `solution_temporaire`
      Supprimer `noeud_courant` et remonter à son père.
    else if les fils de `noeud_courant` n'ont pas encore été construits **then**
      Construire les fils de `noeud_courant`
      Calculer les majorants des fils de `noeud_courant`
      Ordonner les fils en largeur par ordre décroissant de majorant
    end if
    if production de `solution_temporaire` < majorant du premier fils de `noeud_courant` **then**
      Affecter à `noeud_courant` son fils
    else
      Supprimer `noeud_courant` et remonter à son père.
    end if
  end while
end function

il faut non seulement élaguer le premier fils, mais aussi tous ses frères (c'est-à-dire tous les autres fils de `noeud_courant`) car ceux-ci ont tous un majorant inférieur à celui du premier fils, étant triés par ordre décroissant de majorant. De ce fait, il n'est plus possible d'espérer obtenir une solution depuis `noeud_courant`. Pour cette raison, on supprime alors directement `noeud_courant`.

4.3.2 Variantes autour de l'algorithme général

Variante « vérité terrain »

Cette version est la plus simple et est utilisée comme vérité terrain : sa simplicité de mise en œuvre et son optimalité en font la version de référence pour les tests. L'initialisation se base sur le plan de bridage *full-off*, c'est-à-dire $\forall i \in \{1, ..., I\}, c_i = C$. La majoration des nœuds est basée sur la condition suffisante de non-admissibilité suivante :

Condition suffisante de non admissibilité d'une valeur de bridage $c_i = k$ pour l'éolienne $i > I_0$ où les bridages des éoliennes d'indices $\{1, ..., I_0\}$ sont déjà déterminés :
$\exists j \in \{1, ..., J\},\ p_0^2 10^{\frac{\tilde{\mathfrak{r}}(c_i)+a_{i,j}}{10}} > q_j - \sum_{i=1}^{I_0} p_0^2 10^{\frac{l_{i,j}}{10}}$

Cette condition de non-admissibilité permet de construire les fils de chaque nœud, mais aussi d'établir un majorant pour les nœuds de l'arbre. Lorsqu'on examine une branche depuis un nœud quelconque de l'arbre, on examine en fait l'ensemble des solutions admissibles où on a déjà contraint le mode de bridage pour un nombre fini d'éoliennes. La condition suffisante ci-dessus, permet d'éliminer rapidement et sans gros calculs un certain nombre de bridages impossibles pour les éoliennes qui ne sont pas encore fixées. Il reste alors pour chacune de ces éoliennes, des

bridages *a priori* admissibles [9] et on peut calculer le majorant du nœud de la manière suivante :

Majoration d'un noeud donné où les bridages des éoliennes d'indices $[1, I_0]$ sont déjà déterminés :
$\sum_{i=1}^{I} \mathfrak{p}_v(c_i) \leq \sum_{i=1}^{I_0} \mathfrak{p}_v(c_i) + \sum_{i=I_0+1}^{I} \max_{c \in \mathcal{C}_{\text{apriori},i}} (\mathfrak{p}_v(c))$

où $\mathcal{C}_{\text{apriori},i}$ désigne les bridages *a priori* admissibles pour l'éolienne i.

Variante « retenue »

La seconde variante est plus complexe, car elle utilise deux autres algorithmes respectivement pour initialiser et majorer les noeuds de l'arbre. Elle est inspirée d'un algorithme décrit par exemple par Teghem [2012, P.201]. Dans la littérature, on retrouve des approches similaires proposées par Fang et al. [2012]; Mathur et al. [1983]; Hochbaum [1995]; Borchers & Mitchell [1994]; Bonami et al. [2011].

L'idée générale est de majorer des nœuds de l'arbre en résolvant le problème de relaxation continue pour les éoliennes restantes à déterminer, la solution de la relaxation continue ayant une valeur de production électrique toujours supérieure à celle du problème discret. Elle consiste à résoudre le problème comme si les variables de choix (bridage) étaient continues (voir partie 4.3.3).

La spécificité de l'algorithme que nous proposons provient de l'utilisation d'une heuristique gloutonne dédiée pour initialiser la solution initiale, et d'un algorithme de gradient projeté qui est adapté au problème de relaxation continue posé. L'algorithme glouton proposé est une heuristique qui n'est pas complètement optimale mais constitue une bonne base de départ, car très rapide et souvent proche de la solution optimale (voir partie 4.5). La description de l'algorithme glouton est donnée dans la partie 4.4.4. L'algorithme de gradient projeté proposé permet de majorer les nœuds de l'arbre en résolvant la relaxation continue. La relaxation continue est un problème d'optimisation similaire au problème discret mais dont les modes de bridage sont continus. L'algorithme proposé est une variante de l'algorithme du gradient projeté adapté au caractère linéaire par morceaux de la relaxation de $\mathfrak{f}$ utilisée (voir 4.4).

4.3.3 Relaxation continue

Relaxation continue proposée

Pour majorer les noeuds de l'arbre, on exploite une relaxation continue qui transforme les valeurs de bridage discrètes en valeurs continues. Ce prolongement est fait en supposant que les éoliennes ont une infinité de modes de bridage possibles, dont les valeurs de production/émission sont comprises entre celles du mode arrêt et du mode *full power*. Le problème continu peut s'écrire

$$\begin{cases} \max_{\tilde{\mathbf{x}} \in [0, \tilde{\mathbf{r}}(1)]^I} \sum_{i=1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}. \end{cases} \tag{4.11}$$

La variable à optimiser est alors $\tilde{\mathbf{x}}$ qui est désormais à valeurs continues. Pour étendre continûment la fonction $\mathfrak{f}$, on fait le choix d'interpoler $\mathfrak{f}$ entre chaque paire de points successifs par des droites affines. Un exemple de fonction $\mathfrak{f}$ étendue est donné dans la Figure 4.3.

9. Il s'agit des bridages qui semblent au premier abord compatibles avec les éoliennes déjà fixées par le nœud. Par exemple, si pour 5 éoliennes, le bridage des éoliennes 1 et 2 est fixé, les bridages *a priori* admissibles de l'éolienne 3 sont les modes de bridage qui n'occasionnent pas de dépassement acoustique sur les points lorsque les éoliennes non-fixées (ici les éoliennes 4 et 5) sont considérées à l'arrêt. De même, les bridages *a priori* admissibles de l'éolienne 4 sont ceux qui n'occasionnent pas de dépassements lorsque les éoliennes 3 et 5 sont considérées à l'arrêt.

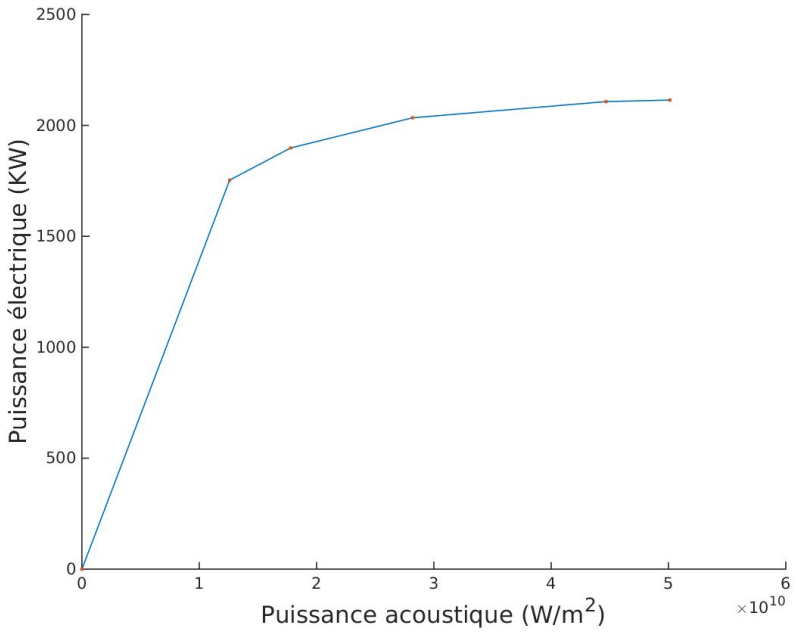


FIGURE 4.3 – Exemple de fonction $\mathfrak{f}$.

On a pu vérifier expérimentalement que la fonction objectif à maximiser est strictement concave. Cela rend le problème d'optimisation strictement convexe : il admet alors un ou plusieurs maxima globaux, et aucun maximum local. L'intérêt d'utiliser une fonction continue par morceaux vient du fait qu'il s'agit de la relaxation continue concave la plus étroite. Ainsi, le majorant obtenu est théoriquement le plus petit qu'on puisse trouver par relaxation continue.

Exploitation de la relaxation pour majorer les noeuds de l'arbre des solutions

La relaxation proposée est utilisée dans la seconde variante de l'algorithme *branch and bound* afin de résoudre le problème discret. En effet, comme expliqué dans le chapitre 2, on peut employer la propriété suivante afin d'utiliser la solution de la relaxation continue pour majorer les nœuds de l'algorithme *branch and bound*[10] :

Propriété utilisée de la relaxation continue : la valeur de la fonction objectif (production électrique) de la solution continue est toujours meilleure que celle de la solution discrète.

Considérons maintenant une solution intermédiaire de profondeur I_0 et qui donc détermine le bridage de I_0 éoliennes. Sachant que I_0 éoliennes sont fixées, on peut déduire de nouvelles contraintes acoustiques. On obtient alors un nouveau problème pour les $I - I_0$ éoliennes restantes

10. Cette propriété est intuitive et simple à démontrer. On peut, par exemple, commencer par utiliser la continuité de la fonction et la compacité de l'ensemble des solutions admissibles (discret et continu) afin de prouver qu'un maximum existe pour chacun des ensembles. L'ordre entre les maxima de chacun de ces deux ensembles peut alors être déduit par propriété de conservation de l'inclusion d'une fonction quelconque.

en posant

$$\tilde{\mathbf{x}}_{\text{det}} = \begin{pmatrix} \tilde{\mathfrak{x}}_v(c_1) \\ \tilde{\mathfrak{x}}_v(c_2) \\ \dots \\ \dots \\ \tilde{\mathfrak{x}}_v(c_{I_0}) \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} \text{ et } \tilde{\mathbf{x}}_{\text{free}} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \tilde{\mathfrak{x}}_v(c_{I_0+1}) \\ \tilde{\mathfrak{x}}_v(c_{I_0+2}) \\ \dots \\ \dots \\ \tilde{\mathfrak{x}}_v(c_I) \end{pmatrix},$$

et en restreignant le problème au vecteur $\tilde{\mathbf{x}}_{\text{free}}$:

$$\begin{cases} \max_{(\tilde{x}_{I_0+1},\dots,\tilde{x}_I)\in[0,\tilde{\mathfrak{x}}_v(1)]^{(I-I_0)}} \sum_{i=I_0+1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{free}} \leq \mathbf{q} - \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{det}}. \end{cases} \tag{4.12}$$

En résolvant ce problème à l'aide d'un algorithme d'optimisation sous contraintes (méthode du gradient projeté, méthode des multiplicateurs de Lagrange, *etc.*), on obtient en sommant les puissances obtenues par la relaxation et celles des I_0 éoliennes fixées un majorant pour ce nœud.

4.4 Algorithme de gradient projeté

Pour résoudre la relaxation continue, on propose d'utiliser un algorithme de gradient projeté. On construit une variante de l'algorithme usuel qui exploite les particularités du problème d'optimisation des éoliennes, c'est-à-dire avec une fonction objectif qui s'écrit comme $\max_{\tilde{\mathbf{x}}} \mathfrak{F}(\tilde{\mathbf{x}}) = \max_{\tilde{\mathbf{x}}} \sum_i \mathfrak{f}(\tilde{x}_i)$ où $\mathfrak{f}$ est une fonction continue croissante concave et affine par morceaux et des contraintes linéaires de type $\tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}$. On montrera dans la partie 4.4.3 **que cet algorithme converge vers le maximum global dans notre cas**.

4.4.1 Introduction de l'algorithme de gradient linéaire par morceaux sur un exemple simple

Données de l'exemple

Considérons un problème comprenant deux éoliennes ($I = 2$), deux points de mesure ($J = 2$) et 5 modes de fonctionnement ($C = 5$). On donne au problème les données (factices et à but didactique) suivantes.

- On considère les valeurs suivantes de puissance d'émission et de production des machines : $\tilde{\mathfrak{x}}_v(\mathcal{C}) = \{0, 1, 2, 3, 4\}$ et $\mathfrak{p}_v(\mathcal{C}) = \{0, 8, 12, 14, 15\}$. On pose $\tilde{\mathfrak{x}}_v(1) = 4$, $\tilde{\mathfrak{x}}_v(2) = 3$, $\tilde{\mathfrak{x}}_v(3) = 2$, *etc.* On pose également, $\mathfrak{p}_v(1) = 15$, $\mathfrak{p}_v(2) = 14$, $\mathfrak{p}_v(3) = 12$, *etc.*
- On considère que $\tilde{x}_1$ et $\tilde{x}_2$ peuvent prendre toutes les valeurs de l'intervalle $[\tilde{\mathfrak{x}}_v(C), \tilde{\mathfrak{x}}_v(1)] = [0, 4]$. On étend ensuite la fonction $\mathfrak{f}$ sur l'intervalle $[0, 4]$ à partir des valeurs de $\mathfrak{p}_v(\mathcal{C})$ en interpolant linéairement :

$$\mathfrak{f} : \mathbb{R} \to \mathbb{R}$$
$$\tilde{x} \mapsto \tfrac{\mathfrak{p}_v(k)-\mathfrak{p}_v(k+1)}{\tilde{\mathfrak{x}}_v(k)-\tilde{\mathfrak{x}}_v(k+1)} \times (\tilde{x} - \tilde{\mathfrak{x}}_v(k+1)) + \mathfrak{p}_v(k+1) \text{ avec } \tilde{\mathfrak{x}}_v(k+1) \leq \tilde{x} < \tilde{\mathfrak{x}}_v(k).$$

- On pose $\tilde{\mathbf{A}}^T = \begin{bmatrix} 3 & 1 \\ \frac{1}{8} & 1 \end{bmatrix}$ et $\mathbf{q} = \begin{bmatrix} 11 \\ 5 \end{bmatrix}$.

Solutions admissibles

Les plans de bridage possibles [11] pour $\tilde{\mathbf{x}}$ forment l'ensemble $[0,4] \times [0,4]$. C'est l'ensemble des points contenus dans le carré délimité par les axes rouges dans la Figure 4.4. Sur cette même figure, un quadrillage représente les changements de gradient. En effet, $\mathfrak{f}$ étant affine par morceaux, le gradient de la fonction $\mathfrak{F}$ [12] change brutalement d'un carré du quadrillage à un autre. De ce fait, le gradient de $\mathfrak{F}$ est constant dans chacun des carrés de la figure et est représenté par les flèches en vert sur le schéma. Les contraintes acoustiques définies par $\tilde{\mathbf{A}}^T$ et $\mathbf{q}$ définissent les bridages admissibles qui sont les points se trouvant à la fois dans le carré rouge et sous les axes bleus dans la Figure 4.5. En effet, le respect de la contrainte $\tilde{\mathbf{A}}^T\mathbf{x} \leq \mathbf{q}$ nécessite d'être sous les deux droites bleues, respectivement d'équation $\tilde{x}_2 = 11 - 3\tilde{x}_1$ et $\tilde{x}_2 = 5 - \frac{1}{8}\tilde{x}_1$. Visuellement, on peut voir que cela revient à redéfinir l'ensemble des points possibles par le quadrilatère de la Figure 4.6.

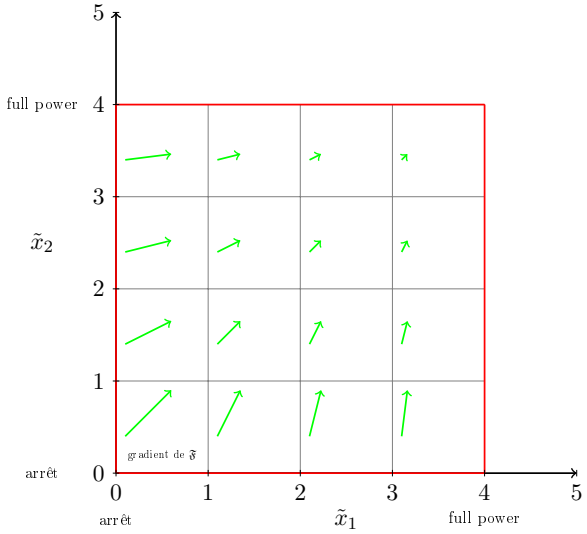


FIGURE 4.4 – Représentation graphique des bridages possibles de l'exemple.

On voit alors graphiquement qu'on obtient un polyèdre convexe représentant le domaine des solutions admissibles.

Déroulement de l'algorithme de gradient projeté

L'algorithme proposé consiste à se déplacer pas à pas dans ce polyèdre en partant de $\tilde{\mathbf{x}}^{(0)} = (0,0)$ et en suivant des directions augmentant la valeur de la fonction $\mathfrak{F}$ en suivant le gradient ou son projeté sur un sous-ensemble de contraintes. Les pas de l'algorithme sont représentés dans la Figure 4.7. Voici une description du fonctionnement :

— On commence au point $\tilde{\mathbf{x}}^{(0)} = (0,0)$ à l'initialisation et on se déplace selon un vecteur colinéaire au gradient jusqu'au prochain changement de gradient, c'est-à-dire au point $\tilde{\mathbf{x}}^{(1)}$ dans la figure.

— $\tilde{\mathbf{x}}^{(1)}$ se situe à la frontière de plusieurs carrés/sous-espaces, dans lesquels le gradient est constant. Plusieurs possibilités s'offrent alors pour le pas suivant. Cependant, un seul de

11. *NB* : bridages possibles et non pas admissibles, on ne tient pas compte des contraintes acoustiques.
12. On rappelle que $\mathfrak{F}(\tilde{\mathbf{x}}) = \sum_{i=1}^{I} \mathfrak{f}(\tilde{x}_i)$.

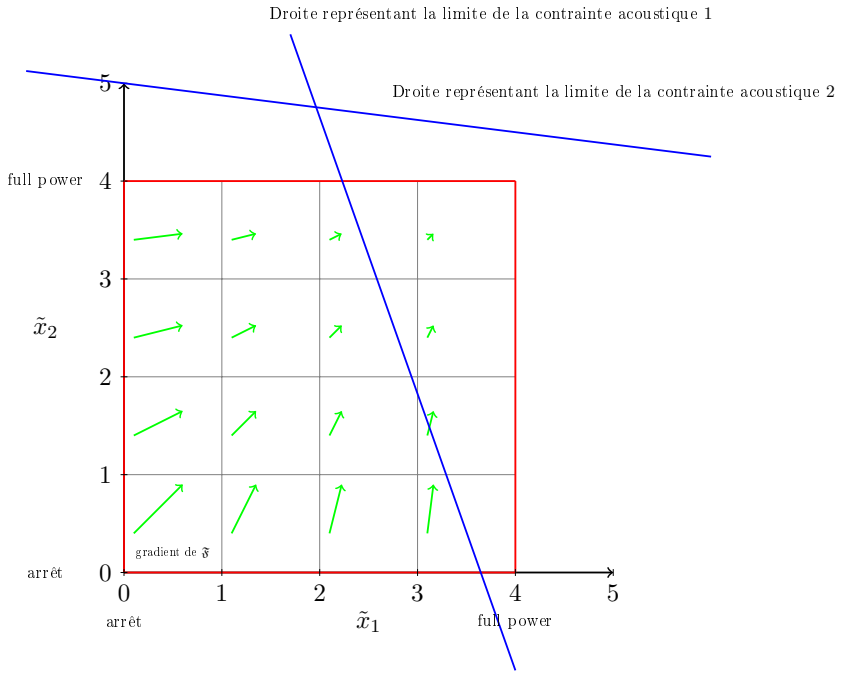


FIGURE 4.5 – Représentation graphique des contraintes acoustiques données par $\tilde{\mathbf{A}}^T\tilde{\mathbf{x}} \leq \mathbf{q}$.

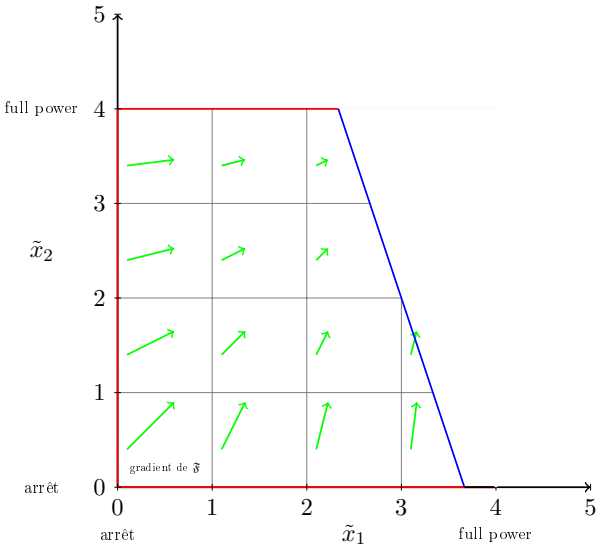


FIGURE 4.6 – Représentation graphique des bridages admissibles.

ces carrés permet un pas suivant la direction du gradient. On va donc choisir ledit carré pour un nouveau pas de $\tilde{\mathbf{x}}^{(1)}$ pour aller en $\tilde{\mathbf{x}}^{(2)}$.

— De nouveau, on va d'abord choisir le carré permettant de suivre le gradient. Cependant, cette fois-ci on va être obligé de s'arrêter sur la frontière d'une des contraintes en $\tilde{\mathbf{x}}^{(3)}$.

— Depuis $\tilde{\mathbf{x}}^{(3)}$, il n'est plus possible de suivre la direction du gradient, mais $\tilde{\mathbf{x}}^{(3)}$ n'est pas le point optimal du problème pour autant. Pour ce faire, on calcule la projection du gradient au point $\tilde{\mathbf{x}}^{(3)}$ sur la contrainte 1 jusqu'à la prochaine contrainte ou changement de gradient (voir Figure 4.8). On peut alors continuer jusqu'en $\tilde{\mathbf{x}}^{(4)}$ où le gradient change.

— De nouveau, en $\tilde{\mathbf{x}}^{(4)}$ on considère le projeté du nouveau gradient sur la contrainte 1 et l'on continue jusqu'en $\tilde{\mathbf{x}}^{(5)}$.
— Le point $\tilde{\mathbf{x}}^{(5)}$ se situe sur l'intersection des deux droites définissant la frontière de deux contraintes : la première est une contrainte acoustique, et la seconde une contrainte de puissance maximale. De ce fait, on essaie les pas définis par les projetés du gradient sur chacune des droites. Comme le montre la Figure 4.9, aucun de ces projetés ne permet de continuer sans violer les contraintes. On en déduit alors que l'on est arrivé sur le plan de bridage optimal.

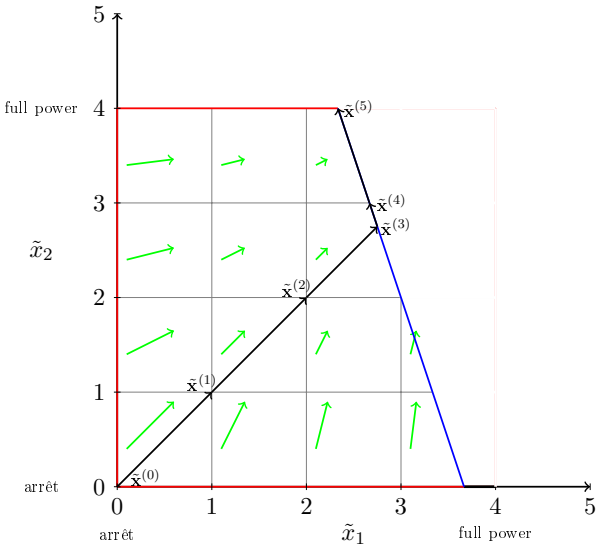


FIGURE 4.7 – Représentation graphique des itérations de l'algorithme de gradient sur l'exemple.

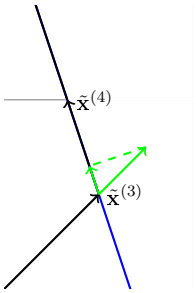


FIGURE 4.8 – Représentation de la projection du gradient au point $\tilde{\mathbf{x}}^{(3)}$ sur la contrainte.

Notons que pour un nombre d'éoliennes supérieur ou égal à 3, cet algorithme se complexifie. En dimension 3 par exemple, le polyèdre est un volume délimité non plus par des droites, mais par des plans. De ce fait, il faut alors « tester » quels sont les déplacements possibles en considérant toutes les combinaisons de plans sur lesquels le point se trouve. D'autre part, le gradient changeant sur les frontières de changement de régime, il peut arriver qu'aucune des directions testées ne soit admissible sans pour autant se situer sur le point optimal. Il faut alors considérer les projections du gradient sur ces frontières (bord des carrés dans la figure).

Par exemple, si le point actuel se trouve sur deux plans en même temps, l'un définissant un changement de régime, et l'autre une contrainte, il faut alors tester trois directions : le gradient projeté sur l'intersection des deux plans (qui est une droite), et sur chacun des deux plans

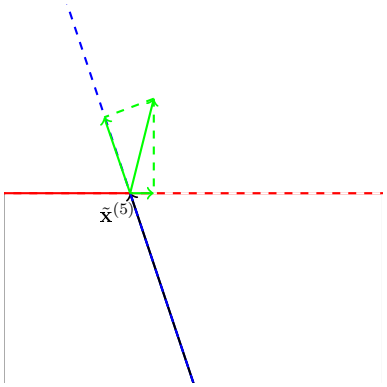


FIGURE 4.9 – Représentation des projections du gradient au point $\tilde{\mathbf{x}}^{(5)}$ sur les contraintes.

séparément.

Si on généralise à un problème de dimension I quelconque, toutes ces frontières sont représentées par des hyperplans et il faut alors tester les projetés du gradient sur toutes les combinaisons d'hyperplans définissant ces frontières, c'est-à-dire 2^k combinaisons où k est le nombre d'hyperplans sur lesquels se trouve le point. Dans la pratique, il n'est pas possible de le faire, et une heuristique doit être choisie. On choisit dans la suite de tester les projetés sur toutes les intersections possibles de $k-1$ contraintes.

4.4.2 Description de l'algorithme de gradient

On rappelle que le problème s'écrit de la manière suivante :

$$\begin{cases} \max_{\tilde{\mathbf{x}} \in [0, \tilde{\mathfrak{x}_v}(1)]^I} \sum_{i=1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q} \end{cases} \tag{4.13}$$

avec $\mathfrak{f}$ une fonction concave affine par morceaux.

Notations supplémentaires

Comme introduit dans la partie 4.4.1, différents sous-espaces décrivant les solutions sont parcourus. Pour les décrire formellement, on introduit les notations suivantes :

- On définit l'ensemble $\mathcal{X}_{\text{poss}} = [0, \tilde{\mathfrak{x}_v}(1)]^I$. C'est l'ensemble des plans de bridage continus possibles.
- On définit l'ensemble $\mathcal{X}_{\text{acou}} = \{\tilde{\mathbf{x}} \in \mathcal{X}_{\text{poss}} | \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq \mathbf{q}\}$. C'est l'ensemble des bridages continus respectant les contraintes acoustiques.
- On définit l'ensemble $\mathcal{X}_{\text{adm}} = \mathcal{X}_{\text{poss}} \cap \mathcal{X}_{\text{acou}}$. C'est l'ensemble des bridages continus admissibles.
- On définit $\forall \mathbf{c} \in \mathcal{C},\ \mathcal{X}_{\text{poss}}(\mathbf{c}) = \{\tilde{\mathbf{x}} \in \mathbb{R}^I | \forall i \in \{1, .., I\}, \tilde{\mathfrak{x}}_v(c_i) \leq \tilde{x}_i \leq \tilde{\mathfrak{x}}_v(c_i - 1)\}$. $\mathcal{X}_{\text{poss}}(\mathbf{c})$ désigne un sous-espace de bridage continu possible.
- De même on définit $\forall \mathbf{c} \in \mathcal{C}, \mathcal{X}_{\text{adm}}(\mathbf{c}) = \mathcal{X}_{\text{adm}} \cap \mathcal{X}_{\text{poss}}(\mathbf{c})$. $\mathcal{X}_{\text{adm}}(\mathbf{c})$ désigne un sous-espace de bridage admissible.

Propriétés

Par construction, on a alors les propriétés suivantes :

— Les ensembles $\mathcal{X}_{\text{poss}}(\mathbf{c})$ forment un recouvrement de $\mathcal{X}_{\text{poss}}$, c'est-à-dire $\mathcal{X}_{\text{poss}} = \cup_{\mathbf{c}\in\mathcal{C}}\mathcal{X}_{\text{poss}}(\mathbf{c})$
— De même, les ensembles $\mathcal{X}_{\text{adm}}(\mathbf{c})$ forment un recouvrement de $\mathcal{X}_{\text{adm}}$.
— De ce fait, $\forall\tilde{\mathbf{x}} \in \mathcal{X}_{\text{adm}}, \exists\ \mathbf{c} \in \mathcal{C}, \tilde{\mathbf{x}} \in \mathcal{X}_{\text{adm}}(\mathbf{c})$. Notons néanmoins qu'un plan de bridage $\tilde{\mathbf{x}}$ peut appartenir à plusieurs sous-espaces de bridage. Dans ce cas, cela signifie qu'une ou plusieurs composantes de $\tilde{\mathbf{x}}$ possèdent une valeur discrète de bridage.
— La fonction $\mathfrak{F}$ est dérivable en tout point $\tilde{\mathbf{x}}$ intérieur des ensembles $\mathcal{X}_{\text{adm}}(\mathbf{c})$. De plus, on peut remarquer que ce gradient est constant et strictement positif pour tous les points intérieurs du sous-espace considéré et vérifie

$$\boldsymbol{\nabla}\mathfrak{F}(\mathcal{X}_{\text{adm}}(\mathbf{c})) = \begin{pmatrix} \frac{\mathfrak{f}(\mathfrak{r}_v(c_1))-\mathfrak{f}(\mathfrak{r}_v(c_1-1))}{\mathfrak{r}_v(c_1)-\mathfrak{r}_v(c_1-1)} \\ \frac{\mathfrak{f}(\mathfrak{r}_v(c_2))-\mathfrak{f}(\mathfrak{r}_v(c_2-1))}{\mathfrak{r}_v(c_2)-\mathfrak{r}_v(c_2-1)} \\ \cdots \\ \frac{\mathfrak{f}(\mathfrak{r}_v(c_I))-\mathfrak{f}(\mathfrak{r}_v(c_I-1))}{\mathfrak{r}_v(c_I)-\mathfrak{r}_v(c_I-1)} \end{pmatrix}. \tag{4.14}$$

Construction de l'algorithme

On part d'une valeur initiale $\tilde{\mathbf{x}}^{(0)}$ respectant les contraintes. A chaque itération n on choisit un pas tel que $\mathfrak{F}(\tilde{\mathbf{x}}^{(n+1)}) > \mathfrak{F}(\tilde{\mathbf{x}}^{(n)})$. Pour garantir la croissance de la valeur de $\mathfrak{F}$ à chaque pas, il est suffisant d'avoir $\langle\tilde{\mathbf{x}}^{(n+1)} - \tilde{\mathbf{x}}^{(n)}, \boldsymbol{\nabla}\mathfrak{F}(\tilde{\mathbf{x}}^{(n)})\rangle > 0$ et $\|\tilde{\mathbf{x}}^{(n+1)} - \tilde{\mathbf{x}}^{(n)}\|_2$ suffisamment petit pour ne pas changer de gradient ou dépasser une contrainte (acoustique ou de production maximale). On suit alors la direction du gradient $\boldsymbol{\nabla}\mathfrak{F}(\tilde{\mathbf{x}})$ ou bien la direction du projeté orthogonal du gradient sur un ensemble de combinaisons de contraintes et/ou changements de régime lorsque cela est nécessaire. Cela revient à dire que $\tilde{\mathbf{x}}^{(n+1)} = \tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{u}^{(n)}$ où $\mathbf{u}^{(n)}$ est la direction du pas, correspondant au gradient ou au gradient projeté, et $\lambda^{(n)}$ est un réel déterminant la taille du pas. Par convention, on considère une direction unitaire, c'est-à-dire que $\|\mathbf{u}^{(n)}\|_2 = 1$ et donc $\lambda = \|\tilde{\mathbf{x}}^{(n+1)} - \tilde{\mathbf{x}}^{(n)}\|_2$.

Le pseudo code de l'algorithme est donné ci-dessous (Algorithme 2).

De fait, à chaque itération n, on atteint un certain nombre de nouvelles contraintes et/ou nouveaux changements de régimes. On calcule alors les projections du gradient sur les différentes intersections possibles de contraintes et changements de régimes. Ces projections sont les directions possibles pour le prochain pas. On va alors boucler sur ces directions, jusqu'à trouver une direction admissible, c'est-à-dire une direction $\mathbf{u}^{(n)}$ telle qu'il existe $\lambda^{(n)} > 0$ et $\mathbf{c} \in \mathcal{C}$ vérifiant $\langle\lambda^{(n)}\mathbf{u}^{(n)}, \boldsymbol{\nabla}\mathfrak{F}(\tilde{\mathbf{x}}^{(n)})\rangle > 0$, $\tilde{\mathbf{x}}^{(n)} \in \mathcal{X}_{\text{adm}}(\mathbf{c})$ et $(\tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{a}^{(n)}) \in \mathcal{X}_{\text{adm}}(\mathbf{c})$.

D'un point de vue calculatoire, pour projeter le gradient sur une intersection de contraintes, qui sont toutes des hyperplans de l'espace de dimension I, on utilise la pseudo-inversion de *Moore-Penrose*. Si on note $\mathbf{M}$ la matrice contenant en colonnes les vecteurs normaux des hyperplans formant l'intersection sur laquelle on désire projeter un vecteur $\mathbf{u}$, le projeté $\mathbf{u}_p$ peut se calculer de la manière suivante [Donchenko & Kirichenko, 2003, p. 855] :

$$\mathbf{u}_p = \mathbf{u} - \mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{u}. \tag{4.15}$$

Cette formule est utilisée dans l'algorithme pour calculer les projetés du gradient sur les intersections de contraintes et/ou changements de régimes. Ces projetés constituent à chaque itération l'ensemble des directions possibles pour le pas suivant.

En ce qui concerne le meilleur pas $\lambda^{(n)}$ pour une direction $\mathbf{u}_p^{(n)}$, il peut être déterminé en résolvant le sous-problème

$$\begin{cases} \max_{\lambda>0}\ \mathfrak{F}(\tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{u}^{(n)}) \\ \mathbf{M}(\tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{u}^{(n)}) \leq \mathbf{q}_{\text{adm}(\mathbf{c})} \end{cases} \tag{4.16}$$

Algorithm 2 Algorithme de gradient

function GRADIENT
Initialisation des variables du problème.
Initialisation de la solution initiale $\tilde{x}^{(0)} = (0, 0, ..., 0)$ de taille I
Initialisation du gradient : `grad` de taille I
Initialisation de la première direction : `grad_p_choisi` = `grad`
Calcul de `lambda` à partir du gradient qui correspond au pas amenant sur la première contrainte ou changement de régime.
while avancer est possible **do**
Avancer de `lambda` $*$ `grad_p_choisi`
if on arrive sur un/plusieurs nouveau(x) changement(s) de régime **then**
calculer le nouveau gradient `grad`
projeter `grad` sur les contraintes actives et les changements de régimes actifs
calculer les projetés `projetes_possible` issus des projections du vecteur résultant sur toutes les intersections possibles de changement de régimes rencontrés.
end if
if on arrive sur un ou plusieurs nouveaux changements de contraintes **then**
projeter `grad` sur les changements de régimes actifs (et pas sur les contraintes actives)
calculer les projetés `projetes_possible` issus des projections du vecteur résultant sur toutes les intersections possibles de contraintes
end if
while il y a des vecteurs dans `projetes_possible` et que `grad_p_choisi` non choisi **do**
calculer `lambda_courant` du projeté courant `p_courant`
if `lambda_courant` $\neq 0$ **then**
on choisit la direction à suivre `grad_p_choisi` = `p_courant`.
on garde le pas du gradient : `lambda` = `lambda_courant`
end if
end while
Mettre à jour les contraintes et les changements de régimes actifs
Mettre à jour les nouvelles contraintes rencontrées et/ou les nouveaux changements de régimes rencontrés
end while
end function

où $\mathbf{q}_{\text{adm}(\mathbf{c})}$ représente les contraintes définissant $\mathcal{X}_{\text{adm}}(\mathbf{c})$. Or, pour $\mathbf{c} \in \mathcal{C}$ et $\lambda^{(n)} \in \mathbb{R}$ tels que $\tilde{\mathbf{x}}^{(n)} \in \mathcal{S}_{\text{adm}}(\mathbf{c})$ et $(\tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{u}^{(n)}) \in \mathcal{S}_{\text{adm}}(\mathbf{c})$, comme le sous-espace $\mathcal{X}_{\text{adm}}(\mathbf{c})$ est linéaire, on a

$$\mathfrak{F}(\tilde{\mathbf{x}}^{(n)} + \lambda^{(n)}\mathbf{u}_p^{(n)}) = \mathfrak{F}(\tilde{\mathbf{x}}^{(n)}) + \lambda^{(n)}\langle \boldsymbol{\nabla}\mathfrak{F}(\tilde{\mathbf{x}}^{(n)}), \mathbf{u}_p^{(n)}\rangle. \tag{4.17}$$

D'autre part, $\mathbf{u}_p^{(n)}$ étant un projeté orthogonal du gradient sur une intersection d'hyperplans, on a

$$\langle \boldsymbol{\nabla}\mathfrak{F}(\tilde{\mathbf{x}}^{(n)}), \mathbf{u}_p^{(n)}\rangle \geq 0. \tag{4.18}$$

Le meilleur pas pour $\lambda^{(n)}$ est alors la valeur maximale telle que la contrainte du problème (4.16) est respectée. Si on note $\mathbf{y} = \mathbf{M}\mathbf{u}_p^{(n)}$ et $\mathbf{z} = \mathbf{q}_{\text{adm}(\mathbf{c})} - \mathbf{M}\tilde{\mathbf{x}}^{(n)}$, on a alors pour meilleur pas $\lambda^{(n)} = \min_{\{i\in\{1,\dots,I\}|y_i>0\}} \frac{z_i}{y_i}$. Si on ne trouve aucun pas strictement positif pour chacune des directions possibles, l'algorithme termine.

4.4.3 Éléments de preuve de l'optimalité de l'algorithme de gradient projeté

L'algorithme de gradient converge théoriquement vers le maximum global. On propose ici des éléments de preuve de ce fait. Il existe au moins un maximum global, il n'existe pas de maxima locaux qui ne soient pas globaux et les maxima globaux se trouvent sur la frontière topologique de l'ensemble des solutions admissibles [13]. Ces propriétés du problème peuvent être montrées à l'aide de théorèmes classiques de mathématiques dont la formulation peut être trouvée par exemple dans le livre de Lange [2004].

La compacité de l'espace des solutions admissibles [14] et la continuité de la fonction $\mathfrak{F}$ permettent d'affirmer qu'un maximum global existe par le théorème des bornes, aussi appelé théorème de Weierstrass [Lange, 2004, proposition 2.5.3, p.105]. La convexité des contraintes et la concavité de la fonction $\mathfrak{F}$ permettent de montrer qu'il n'existe pas d'autres maxima locaux [15] [Lange, 2004, Proposition 5.4.1, p.105]. Tous les points intérieurs du problème admettent une dérivée directionnelle dans toutes les directions. Si on note $\boldsymbol{\nabla}_{\mathbf{z}}\mathfrak{F}(\tilde{\mathbf{x}})$ la dérivée directionnelle de $\tilde{\mathbf{x}}$ selon la direction $\mathbf{z}$ alors le maximum $\tilde{\mathbf{x}}^*$ vérifie $\forall \tilde{\mathbf{x}} \in \mathcal{X}_{\text{adm}}, \langle \boldsymbol{\nabla}_{(\tilde{\mathbf{x}}-\tilde{\mathbf{x}}^*)}\mathfrak{F}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}} - \tilde{\mathbf{x}}^*\rangle \leq 0$ [Lange, 2004, proposition 5.4.1, p.105]. Or, étant donnée la positivité du gradient, cette condition ne peut-être vérifiée que sur la frontière de $\mathcal{X}_{\text{adm}}$. On en déduit que le maximum se trouve sur la frontière de $\mathcal{X}_{\text{adm}}$.

Ajoutons également deux autres remarques propres au problème dues à la nature affine par morceaux de $\mathfrak{f}$: le nombre de sous-espaces de bridage $\mathcal{X}_{\text{adm}}$ est limité et majoré par $(C-1)^I$ et chaque sous-espace de bridage a au plus $J + 2 \times I$ frontières. Pour un point à l'intérieur de celui-ci, on teste donc à l'itération correspondante au maximum $2^{J+2\times I}$ directions. De ce fait, l'algorithme converge vers la solution globale puisque le nombre de pas maximal est fini.

4.4.4 Heuristique gloutonne d'initialisation

L'algorithme glouton présenté ici est utilisé dans l'algorithme *branch and bound* comme solution d'initialisation et est une heuristique qui a vocation à donner un temps de calcul inférieur à la seconde. L'algorithme est dit « glouton »(*greedy*), car à chaque itération, on ne bride qu'une

13. La frontière topologique correspond à la formulation mathématique des points au « bord » du domaine. En d'autres termes, si on note $\mathbf{x}_{\max}$ la solution du problème, affirmer qu'elle se trouve sur la frontière topologique de $\mathcal{X}_{\text{adm}}$ revient à dire : $\exists j \in \{1,\dots,J\}, \sum_{i=1}^{I} \tilde{a}_{i,j} \times \tilde{x}_{\max,i} = q_j$ ou $\exists i \in \{1,\dots,I\}, \tilde{x}_{\max,i} = \tilde{\mathfrak{x}}_v(1)$ ou $\tilde{x}_{\max,i} = \tilde{\mathfrak{x}}_v(C)$.

14. En dimension finie, les ensembles compacts sont les ensembles fermés et bornés.

15. NB : En revanche, le maximum global n'est pas nécessairement unique.

unique éolienne à la fois en recherchant le choix optimal pour cette itération, sans garantie toutefois que la solution soit optimale une fois que l'algorithme termine.

Plus précisément, cette méthode consiste à partir du plan de bridage *full-power* (puissance acoustique et électrique maximale) et à brider progressivement en choisissant à chaque itération une seule éolienne à brider « d'un cran », c'est-à-dire avec le mode de fonctionnement de puissance tout juste inférieure. À chaque itération, on choisit l'éolienne qui a le meilleur « rapport perte acoustique/électrique » c'est-à-dire celle qui va permettre la plus grande perte acoustique aux points où il y a dépassement de contraintes acoustiques, proportionnellement à la perte électrique occasionnée.

À noter que lorsqu'on calcule la perte acoustique occasionnée sur les dépassements, on ne considère pas la perte acoustique totale, mais seulement la perte sur les points où la contrainte n'est pas respectée. Le pseudo-code de l'algorithme est donné ci-après (Algorithme 3).

Algorithm 3 Algorithme glouton

function GLOUTON
initialiser `c` tel que $\forall i, c_i = 1$
while un dépassement existe sur un des points **do**
`depassements` = liste des points où il existe un dépassement
for chaque i entre 1 et n **do**
$\texttt{critere}(i) = \frac{\tilde{d}_{\text{diff},i}}{p_{\text{diff},i}}$ où $\tilde{d}_{\text{diff},i}$ est la perte totale de dépassement sur l'ensemble des points de `depassements` et où $p_{\text{diff},i}$ est la perte totale de puissance électrique lorsqu'on passe de c_i à $c_i + 1$.
end for
$\texttt{i_choix} = \arg\max(\texttt{critere})$
$\texttt{c(i_choix)} = \texttt{c(i_choix)} + 1$
end while
end function

Pour i donné, on calcule $\texttt{critere}(i)$ de la manière suivante :

— $\tilde{d}_{\text{diff},i} = \sum_{j \in \texttt{depassements}} \tilde{a}_{i,j} \left[\tilde{\mathfrak{x}}_v(c_i) - \max\left(\tilde{\mathfrak{x}}_v(c_i+1), \tilde{\mathfrak{x}}_v(c_i) - \frac{\tilde{d}_j}{\tilde{a}_{i,j}} \right) \right]$ où $\tilde{d}_j$ est le dépassement en W/m^2 en j, c'est-à-dire la composante j du vecteur $\tilde{\mathbf{d}} = \mathbf{q} - \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}$. Le terme $\tilde{d}_{\text{diff},i}$ désigne la perte de dépassement occasionnée par le bridage de l'éolienne i. La complexité de la formule pour $d_{\text{diff},i}$ provient du cas où l'éolienne i annule un dépassement en un ou plusieurs points. Dans ce cas on ne comptabilise pour ces points qu'une partie de la perte d'énergie sonore, à savoir l'énergie nécessaire à annuler le dépassement.

— $p_{\text{diff},i} = \mathfrak{p}(c_i) - \mathfrak{p}(c_i + 1)$. $p_{\text{diff},i}$ est la perte électrique occasionnée par le bridage de l'éolienne i.

4.5 Évaluation expérimentale

On a construit un algorithme d'optimisation déterministe dont on souhaite déterminer l'optimalité et le temps de calcul afin de pouvoir vérifier la possibilité de l'utiliser dans un contexte industriel. Les contraintes sont, pour un tel algorithme, de permettre un calcul optimal et de nécessiter un temps de calcul inférieur à 10 minutes qui est la durée d'un cycle de calcul dans le domaine de l'éolien. On présente donc dans cette partie une évaluation expérimentale de l'algorithme *branch and bound* et de l'algorithme d'initialisation glouton.

4.5.1 Données utilisées

On a utilisé des données existantes pour tester l'algorithme. Ces données ont été extraites du serveur de Venathec SAS et proviennent de fichiers de type tableur utilisés comme entrées par l'algorithme de calcul de plans de bridage de Venathec SAS. Elles proviennent de 28 parcs éoliens français différents, dont les noms et la localisation exacte ne sont pas données pour des raisons de confidentialité industrielle.

De ces parcs, on obtient alors 1210 cas à traiter correspondant aux différentes classes homogènes, et donc autant d'appels aux différents algorithmes.

Dans la suite on compare respectivement l'algorithme glouton et l'algorithme *branch and bound* (variante retenue), à la vérité terrain.

4.5.2 Résultats

Les résultats de l'algorithme glouton et *branch and bound* (variante retenue) sont résumés dans le Tableau 4.1.

algorithme	**glouton**	**branch and bound**
temps de calcul maximal (s)	$0,25$ s	240 s
maximum non trouvé	32 %	$1,5$ %
perte relative maximale	30 %	$2,3$ %
perte relative moyenne	$1,5$ %	$0,09$ %
Excès acoustique maximum	$0,00$ dB	$0,00$ dB
Excès acoustique moyen	$-0,86$ dB	$-0,84$ dB

TABLEAU 4.1 – Résultats sur l'optimalité et la vitesse d'exécution des algorithmes glouton et *branch and bound*. L'algorithme a été codé en langage Matlab et chaque calcul a été mené en utilisant un processeur CPU unique (sans utilisation de *multithreading*) de 2,27 GHz.

Commentaires sur l'algorithme glouton

On peut voir que l'algorithme glouton est rapide (0,25 s au maximum dans les expériences) et optimal dans 70 % des cas. Quand ce n'est pas le cas, l'erreur relative par rapport à la puissance optimale donnée par l'algorithme *branch and bound* « vérité terrain » peut aller jusqu'à 30 %.

Néanmoins, la distribution des erreurs relatives, donnée par la Figure 4.10 est relativement bonne. On peut notamment observer que l'erreur est inférieure à 4 % dans 90 % des cas.

L'algorithme glouton constitue de ce fait une bonne base de départ pour l'algorithme *branch and bound*.

Commentaires sur la variante retenue de l'algorithme *branch and bound*

L'algorithme proposé utilise comme base de départ la solution de l'algorithme glouton. À partir de celui-ci, il va améliorer les solutions obtenues jusqu'à la solution optimale.

La Figure 4.11 présente la distribution des erreurs pour cet algorithme. Comme prévu, il permet de trouver le plan de bridage optimal dans l'immense majorité des cas.

Cependant, on peut observer que dans un nombre réduit de cas (une vingtaine sur l'ensemble des 1210 expériences), contrairement à ce que la théorie prévoit, l'erreur relative maximale peut aller jusqu'à $2,3$ %. Ceci s'explique par le résultat de l'algorithme de gradient pour le problème continu, qui n'est pas toujours optimal en raison de l'impossibilité pratique de tester les

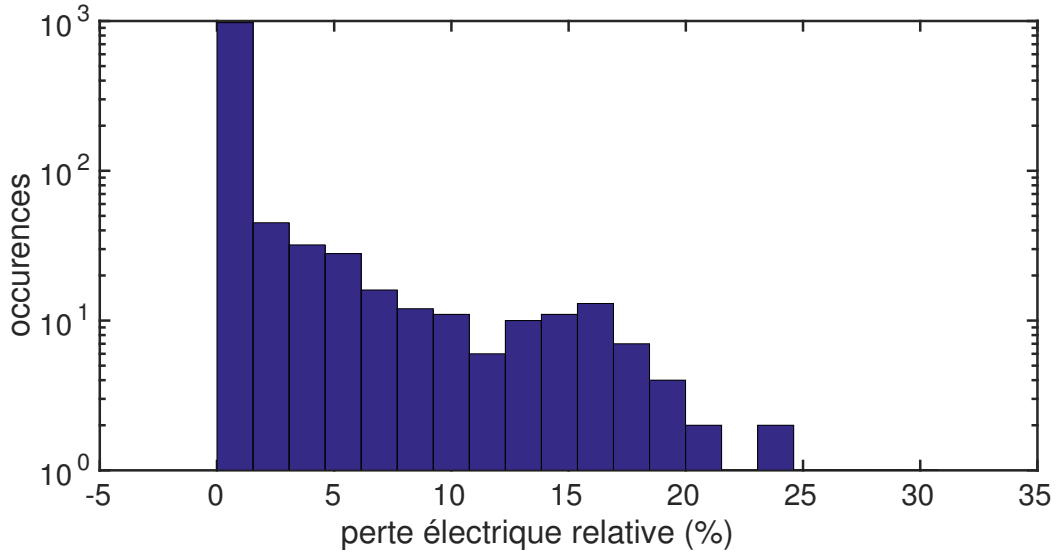


FIGURE 4.10 – Distribution des erreurs relatives pour l'algorithme glouton.

projections du gradient sur toutes les combinaisons de contraintes. Cependant, l'erreur relative maximale reste faible, et une erreur relative non-nulle n'est constatée que dans $1,5$ % des cas.

Pour finir, étant donné le temps de calcul qui peut aller jusqu'à 240 s, on peut se demander quelle est la vitesse temporelle de l'évolution de l'erreur afin de déterminer si l'algorithme ne converge pas avant que l'algorithme ne termine. La Figure 4.12 présente l'évolution temporelle de l'erreur relative pour les problèmes qui terminent en plus de 5 s, en présentant le pire cas à chaque instant et les quantiles sur ces différents tests.

On peut alors observer qu'au bout de 2 min environ, l'algorithme converge pour tous les cas étudiés.

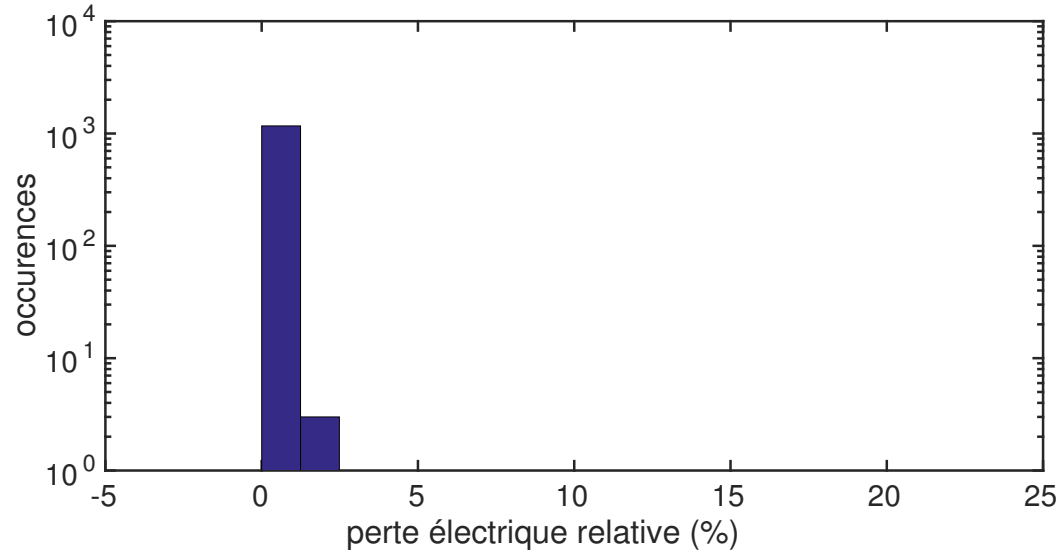


FIGURE 4.11 – Distribution des erreurs relatives pour l'algorithme *branch and bound* (variante retenue).

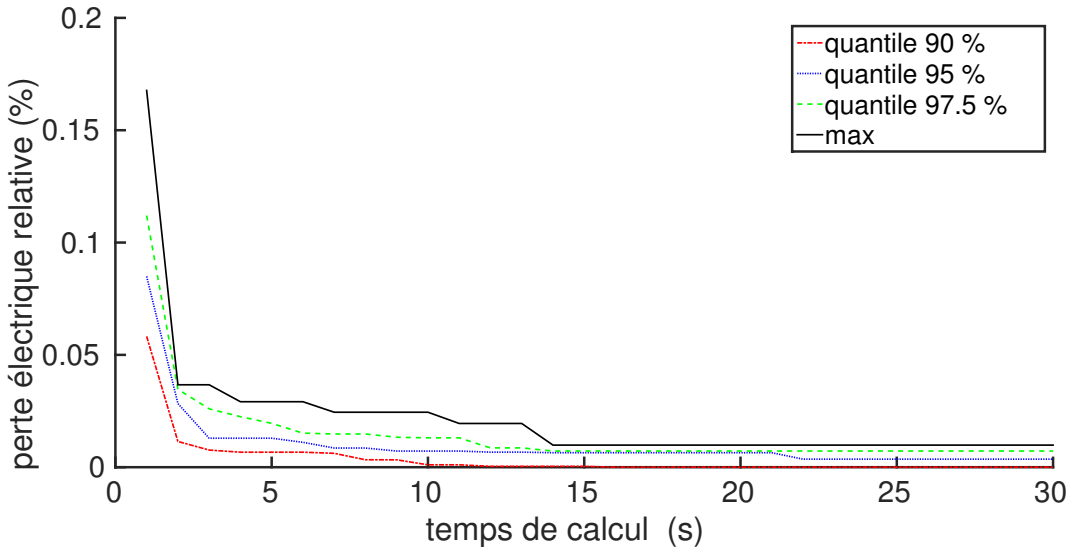


FIGURE 4.12 – Évolution temporelle des erreurs relatives pour l'algorithme *branch and bound* (variante retenue) pour les échantillons ayant un temps de calcul supérieur à 5 s.

5

Problème déterministe à contraintes sur les médianes temporelles

5.1 Formulation

Dans ce chapitre, on conserve les hypothèses déterministes du chapitre 4 mais on considère les contraintes sur les médianes temporelles, c'est-à-dire portant sur les valeurs médianes des distributions du bruit résiduel et des distributions du bruit ambiant pour chaque classe homogène o, comme définies dans le chapitre 3, plutôt que les valeurs instantanées de ces bruits. Voici la formulation mathématique du problème au temps t_0 du problème :

$$\begin{cases} \max_{(\mathbf{c}_{t_0},...,\mathbf{c}_{t_0+T-1})\in\mathcal{C}^{I\times T}} \sum_{t=t_0}^{t_0+T-1}\sum_{i=1}^{I} \mathfrak{p}_v(c_{t,i}) \\ \forall t \in \{t_0,...,t_0+T-1\}, \forall j \in \{1,...,J\}, \\ \hat{b}_{[t-N+1,t],j} - \hat{r}_{[t-N+1,t],j} < S_{e,o} \\ \text{ou } \hat{r}_{[t-N+1,t],j} < S_{s,o}. \end{cases} \tag{5.1}$$

Dans la suite, on simplifie la notation pour les médianes $\hat{b}_{[t-N+1,t],j}$ en $\hat{b}_{t,j}$, le paramètre N étant fixé en pratique à la durée d'une étude, soit 2 semaines de mesures environ.

L'intérêt de ce problème consiste à voir comment on peut adapter simplement, sans prédiction ni incertitudes, les résultats du chapitre 4 aux contraintes légales françaises.

5.2 Heuristiques proposées

5.2.1 Démarche expérimentale

Afin d'alléger les notations, on se place à classe homogène fixée sans perte de généralité, et on retire alors l'indice o des notations générales.

Pour résoudre le problème, la démarche globale est d'adapter l'utilisation de l'algorithme déterministe précédent, en définissant le bruit particulier maximum $\mathbf{q}$ de sorte à vérifier les contraintes sur les médianes temporelles. En effet, on a montré dans la partie 4.2.1 que l'algorithme *branch and bound* développé permet de calculer le plan de bridage optimal sous des contraintes du type $\tilde{\mathbf{A}}^T\tilde{\mathbf{X}} \leq \mathbf{q}$ avec $\mathbf{q}$ quelconque.

Dans la suite du chapitre nous proposons différents algorithmes basés sur 2 paradigmes : l'approche instantanée et l'approche gloutonne.

Approche instantanée

L'approche instantanée limite l'émergence instantanée à chaque instant exactement de la même façon que dans le chapitre 4. On s'attend en effet à ce que limiter l'émergence instantanée permette *a fortiori* de limiter l'émergence basée sur des médianes temporelles. Plusieurs variantes sont testées.

Approche gloutonne

L'approche gloutonne consiste à calculer à chaque instant t et pour chaque point j le niveau de bruit particulier maximal $q_{t,j}$ que l'on peut se permettre pour que le critère légal soit toujours respecté [16].

La première approche ne nécessite pas d'explication supplémentaire, le chapitre 4 étant suffisant à sa compréhension. Le détail des différentes variantes testées est donné dans la partie 5.3. L'approche gloutonne pour le problème est décrite dans la suite, avant de présenter les résultats.

5.2.2 Algorithme glouton de contrôle non-prédictif

Utiliser l'algorithme restreignant l'émergence instantanée apparaît de prime abord trop restrictif. En effet, **restreindre l'émergence instantanée revient à ajouter des contraintes supplémentaires au problème de contrôle**. De ce fait on décrit et propose ici une approche gloutonne.

Remarques préliminaires sur l'évolution de la médiane d'une distribution statistique

L'algorithme est basé sur le principe qu'il est possible d'anticiper l'évolution de la médiane $\hat{x}_N$ de l'ensemble d'échantillons $x_{1:N}$ par rapport à celle de l'ensemble $x_{1:N-1}$ lors de l'ajout de la valeur x_N. Prenons un exemple pour commencer. Supposons que l'on dispose d'un ensemble $x_{1:10} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, et donc de médiane $\hat{x}_{10} = 5,5$. Considérons maintenant l'ajout d'un nouvel élément $x_{11} \notin \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ et déterminons $\hat{x}_{11}$ la médiane du nouvel ensemble $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, x_{11}\}$. Il apparaît alors 3 cas possibles pour la valeur de cette nouvelle médiane $\hat{x}_{11}$:

— Premier cas : $x_{11} \leq 5$. Dans ce cas on observe que $\hat{x}_{11} = 5$.
— Deuxième cas : $x_{11} \geq 6$. Dans ce cas on observe que $\hat{x}_{11} = 6$.
— Troisième et dernier cas : $5 < x_{11} < 6$. Dans ce cas on observe que $\hat{x}_{11} = x_{11}$.

On peut alors résumer ces trois cas par $\hat{x}_{11} = \max(\min(x_{11}, 6), 5)$. Les éléments 5 et 6 constituent les bornes des valeurs possibles pour $\hat{x}_{11}$. D'autre part, ces éléments peuvent être déterminés respectivement comme $\max_x\{x \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} | x < \hat{x}_{10}\}$ et $\min_x\{x \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} | x > \hat{x}_{10}\}$. On les note respectivement x_{10}^{prev} et x_{10}^{next}.

On a donc généralement la formule

$$\hat{x}_N = \max(\min(x_N, \hat{x}_{N-1}^{\text{next}}), \hat{x}_{N-1}^{\text{prev}}). \tag{5.2}$$

16. Le terme de glouton n'a donc pas le même sens que dans le chapitre 4.

Cette formule n'est vraie que dans le cas où le nombre d'éléments de l'ensemble $x_{1:N}$ est pair. S'il est impair, la formule diffère légèrement et on a

$$\hat{x}_N = \frac{\max(\min(x_N, \hat{x}_{N-1}^{\text{next}}), \hat{x}_{N-1}^{\text{prev}}) + \hat{x}_{N-1}}{2}. \tag{5.3}$$

De cette manière, il est possible de connaître la valeur de la nouvelle médiane $\hat{b}_{t,j}$ en fonction de l'échantillon passé $b_{t-N:t-1,j}$ et de la valeur actuelle du bruit ambiant $b_{t,j}$ et ceci de manière déterministe. On est alors capable de calculer en t la valeur maximale du bruit ambiant $b_{t,j}$ telle que $\hat{b}_{t,j}$ respecte la contrainte d'émergence. On peut résoudre alors le problème en utilisant l'algorithme déterministe avec une valeur de $\mathbf{q}_t$ adaptée à l'instant t.

Reformulation déterministe du problème

L'algorithme glouton consiste à choisir le plan de bridage qui maximise le gain entre les instants $t_0 - 1$ et t_0 sans considérer l'impact d'un tel choix sur les instants futurs.
Le problème à résoudre peut alors se reformuler

$$\begin{cases} \max_{(c_{t_0,1}, c_{t_0,2}, ..., c_{t_0,I}) \in \mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_{t_0,i}) \\ \forall j \in \{1, ..., J\}, \hat{b}_{t_0,j} \leq \hat{b}_{t_0,j}^{\max}, \end{cases} \tag{5.4}$$

la commande $\mathbf{c}_t$ n'influant pas sur la variable $\hat{\mathbf{r}}_{t_0}$. On note $\hat{b}_{t_0,j}^{\max} = \max(S_{e,o} + \hat{r}_{t_0,j}, S_{s,o})$. Le problème ressemble alors au problème déterministe instantané, à la différence que le paramètre contraint est une médiane. Il faut donc d'abord déterminer la variation de $\hat{b}_{t,j}$ en fonction de la commande $\mathbf{c}_t$. Pour ce faire, on utilise la valeur de $\hat{b}_{t-1,j}$ qui est issue de la mesure à l'instant précédent, et de l'échantillon $b_{t-N:t-1,j}$.

Résolution gloutonne entre $t_0 - 1$ et t_0 pour un point de mesure unique

On considère pour commencer le problème à un point de mesure unique j_0, et pour simplifier, on se place dans le cas où $b_{t_0-N:t_0-1,j_0}$ a un nombre pair d'éléments. On peut utiliser les remarques préliminaires de la partie précédente afin de résoudre ce problème. Tout comme dans la partie précédente, on note $\hat{b}_{t_0-1,j_0}^{\text{prev}}$ et $\hat{b}_{t_0-1,j_0}^{\text{next}}$ les éléments de l'échantillon passé $b_{t_0-N:t_0-1,j_0}$ situés juste avant et juste après $\hat{b}_{t-1,j_0}$. En utilisant l'équation (5.2) et le fait que, pour résoudre le problème en t_0, il est nécessaire et suffisant d'avoir $\hat{b}_{t_0,j_0} \leq \hat{b}_{t_0,j_0}^{\max}$, on obtient alors trois cas possibles, donnant chacun un problème déterministe à résoudre :

— **Cas 1** : $\hat{b}_{t_0-1,j_0}^{\text{prev}} < \hat{b}_{t_0,j_0}^{\max} < \hat{b}_{t_0-1,j_0}^{\text{next}}$. Le problème (5.1) devient alors :

$$\begin{cases} \max_{(c_{t_0,1}, c_{t_0,2}, ..., c_{t_0,I}) \in \mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_{t_0,i}) \\ b_{t_0,j_0} \leq \hat{b}_{t_0,j_0}^{\max}. \end{cases} \tag{5.5}$$

On résout ce problème avec l'algorithme déterministe *branch and bound*.

— **Cas 2** : $\hat{b}_{t_0,j_0}^{\max} > \hat{b}_{t_0-1,j_0}^{\text{next}}$. Le problème (5.1) devient alors :

$$\max_{(c_{t_0,1}, c_{t_0,2}, ..., c_{t_0,I}) \in \mathcal{C}^I} \sum_{i=1}^{I} \mathfrak{p}_v(c_{t_0,i}). \tag{5.6}$$

Le problème est trivial puisque $(c_{t_0,1}, c_{t_0,2}, ..., c_{t_0,I}) = (1, 1, ..., 1)$ (plan *full power*) convient.

— **Cas 3** : $\hat{b}_{t_0,j_0}^{\max} < \hat{b}_{t_0-1,j_0}^{\text{prev}}$ alors le problème (5.1) n'est pas soluble au temps t_0. On choisira dans ce cas une heuristique afin d'espérer pouvoir résoudre le problème dans un futur proche.

On note $\hat{b}^{\text{choice}}_{t_0,j_0}$ la valeur choisie pour la résolution, à savoir $\hat{b}^{\text{choice}}_{t_0,j_0} = \hat{b}^{\max}_{t_0,j_0}$ pour le premier cas, $\hat{b}^{\text{choice}}_{t_0,j_0} = +\infty$ pour le deuxième cas, et une valeur à choisir de façon heuristique pour le dernier.

Finalement, pour se ramener à la formulation du problème déterministe du chapitre 4, on écrit alors, dans les 3 cas :

$$\begin{cases} \max_{(c_{t_0,1},c_{t_0,2},...,c_{t_0,n})\in\mathcal{C}^n)} \sum_{i=1}^{I} \mathfrak{p}_v(c_{t_0,i}) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} \leq q_{j_0} \end{cases} \tag{5.7}$$

avec $q_{j_0} = 10^{\frac{\hat{b}^{\text{choice}}_{t_0,j_0}}{10}} - \tilde{r}_{j_0}$.

À noter toutefois que, si q_{j_0} est négatif, dans le troisième cas, le critère d'émergence ne pourra pas être respecté en t_0.

Résolution gloutonne entre $t_0 - 1$ et t_0 pour un nombre de points quelconque

De la même façon, il est possible de résoudre le cas général, où le nombre de points de mesure J est quelconque avec un traitement similaire à celui fait pour un point de mesure. Pour ce faire, on considère les distributions $b_{t_0-N:t_0-1,j}$ de chaque point j et on commence par déterminer $\hat{b}^{\text{prev}}_{t_0-1,j}$, $\hat{b}^{\text{next}}_{t_0-1,j}$ et $\hat{b}^{\max}_{t_0-1,j}$ pour chacune d'elles. On génère alors un problème déterministe comme précédemment en appliquant une procédure similaire à celle pour un point unique. On se ramène alors au problème suivant :

$$\begin{cases} \max_{(c_{t_0,1},c_{t_0,2},...,c_{t_0,I})\in\mathcal{C}^I} & \sum_{i=1}^{I} \mathfrak{p}_v(c_{t_0,i}). & & \\ \forall j \in \{1,...,J\}, & b_{t_0,j} \leq \hat{b}^{\max}_{t_0,j} & \text{si } \hat{b}^{\max}_{t_0,j} \in [\hat{b}^{\text{prev}}_{t_0,j}, \hat{b}^{\text{next}}_{t_0,j}], \\ & b_{t_0,j} \leq \hat{b}^{\text{prev}}_{t_0,j} & \text{si } \hat{b}^{\max}_{t_0,j} < \hat{b}^{\text{prev}}_{t_0,j}. \end{cases} \tag{5.8}$$

Le problème est résolu par l'algorithme *branch and bound* en posant $\forall j \in \{1,...,J\}, q_j = 10^{\frac{\hat{b}^{\text{choice}}_{t_0,j}}{10}} - \tilde{r}_j$.

5.3 Évaluation expérimentale

5.3.1 Description des variantes des algorithmes de contrôle testés

Variantes de l'algorithme glouton

Trois variantes de l'algorithme glouton et trois variantes de l'algorithme à émergence instantanée ont été testées. Concernant l'algorithme glouton, on propose 3 variantes V1, V2 et V3 qui diffèrent par le choix de $\hat{b}^{\text{choice}}_{t,j}$ selon les situations comme représenté dans le Tableau 5.1. Les choix diffèrent dans le cas $\hat{b}^{\max}_{t,j} < \hat{b}^{\text{prev}}_{t-1,j}$ qui ne peut être résolu à l'instant t, et le cas $\hat{b}^{\max}_{t,j} > \hat{b}^{\text{next}}_{t-1,j}$ dont les contraintes sont respectées quel que soit le plan de bridage. Des choix différents sont présentés afin de tenter de prendre en compte l'évolution des distributions au cours du temps.

Variantes de l'algorithme instantané

Concernant l'algorithme instantané, on propose trois autres variantes notées V1, V2 et V3. La variante V1 de l'algorithme à émergence instantanée correspond au problème déterministe instantané du chapitre 4. C'est à dire qu'on résout le problème déterministe instantané à chaque instant en choisissant $\mathbf{q} = 10^{\max(S_{e,o}+\mathbf{r}_t, S_{s,o})} - \tilde{\mathbf{r}}_t$.

	$\hat{b}^{\max}_{t,j} < \hat{b}^{\text{prev}}_{t-1,j}$	$\hat{b}^{\text{prev}}_{t-1,j} < \hat{b}^{\max}_{t,j} < \hat{b}^{\text{next}}_{t-1,j}$	$\hat{b}^{\max}_{t,j} > \hat{b}^{\text{next}}_{t-1,j}$
Glouton V1	$\hat{b}^{\max}_{t,j}$	$\hat{b}^{\max}_{t,j}$	$+\infty$
Glouton V2	$\hat{b}^{\max}_{t,j}$	$\hat{b}^{\max}_{t,j}$	$\hat{b}^{\max}_{t,j}$
Glouton V3	$\hat{b}^{\text{prev}}_{t,j}$	$\hat{b}^{\max}_{t,j}$	$+\infty$

TABLEAU 5.1 – Choix de $\hat{b}^{\text{choice}}_{t,j}$ pour chaque variante testée de l'algorithme glouton en fonction de $\hat{b}^{\max}_{t,j}$, $\hat{b}^{\text{prev}}_{t-1,j}$ et $\hat{b}^{\text{next}}_{t-1,j}$.

La variante V2 de l'algorithme à émergence instantanée utilise le vecteur $\boldsymbol{\delta}$ des dépassements finaux de la première variante. On choisit alors $\mathbf{q} = 10^{\max(S_{e,o}+\mathbf{r}_t+\boldsymbol{\delta}, S_{s,o})} - \tilde{\mathbf{r}}_t$. Cette utilisation des dépassements finaux est une tentative pour recaler les dépassements finaux obtenus. L'idée est que les plans de bridage calculés engendrent une émergence strictement inférieure à l'émergence réglementaire. En élargissant légèrement la contrainte, on peut alors espérer obtenir plus finement un plan « tout juste réglementaire ».

La variante V3 de l'algorithme à émergence instantanée est similaire à la variante V1. Tout comme pour la variante V1, on calcule un plan de bridage à partir du critère d'émergence instantané $\mathbf{q} = 10^{\max(S_{e,o}+\mathbf{r}_t, S_{s,o})} - \tilde{\mathbf{r}}_t$. Cependant, ce plan calculé n'est effectivement appliqué que dans le cas où il génère une production supérieure à celle du plan statique conforme du parc (plan de bridage usuel, contrôlé par les conditions météorologiques uniquement).

5.3.2 Protocole expérimental

Les données utilisées pour les expérimentations sont les données de simulation acoustique, les mesures temporelles des niveaux résiduels et les données constructeurs extraites d'études de Venathec. Les données de simulation acoustique donnent l'information de la propagation (c'est-à-dire la matrice $\mathbf{A}$) et les données constructeurs renseignent sur les modes de bridage. Pour chaque algorithme à tester, on applique une procédure générale à partir de ces données. De plus, afin d'évaluer les algorithmes proposés, on applique également cette procédure au plan *full-power* et au plan statique conforme. La procédure appliquée est la suivante :

1. Les données du parc sont extraites et triées par classes homogènes (période de la journée, secteur de vent, vitesse de vent).
2. Seuls les niveaux résiduels mesurés sont conservés, et les valeurs sont synchronisées par date.
3. Pour chaque classe homogène, une expérience indépendante est lancée.
4. Pour chaque expérience et pour chaque méthode, on simule le contrôle du parc en parcourant temporellement les valeurs de résiduel. À chaque instant t, on procède alors de la façon suivante :
   — On calcule la distribution des bruits résiduels au temps t.
   — À partir de cette distribution, on calcule le plan de bridage à appliquer au temps t.
   — On stocke la production obtenue ainsi que le bruit ambiant généré.
   — On calcule le dépassement à partir de la médiane du bruit ambiant et de celle du bruit résiduel.

On évalue les résultats des expériences à partir des critère suivants :
— production électrique,
— perte de production vis-à-vis de la production du plan *full power*,

— nombre de fois où la contrainte réglementaire est dépassée sur la durée totale (calcul des médianes $\hat{b}_{t,j}$ et $\hat{r}_{t,j}$ pour chaque point j à l'aide des distributions de bruit résiduel et de bruit ambiant puis vérification de la contrainte réglementaire),
— dépassements finaux (valeur des dépassements acoustiques à partir des médianes) à l'instant final.

5.3.3 Résultats

Pour les expérimentations, les données de 3 parcs éoliens ont été utilisées. Le nom de ces parcs n'est pas donné pour des raisons de confidentialité. Pour chaque parc, deux classes homogènes ont été étudiées (un secteur de vent, période de jour et de nuit). Les résultats sont donnés dans les Tableaux 5.1, 5.2, 5.3 et 5.4. Pour les deux premiers tableaux, la couleur verte indique un gain de production vis-à-vis du plan statique, et la couleur rouge une perte de production. Pour les dépassements, les dépassements élevés (supérieurs à 0,5 dB) apparaissent en rouge.

Production (MW.h)	Parc 1		Parc 2		Parc 3	
	jour	nuit	jour	nuit	jour	nuit
Plan full-power	1000	807	107	133	158	850
Plan statique	977	464	107	48	1580	850
Glouton V1	930	307	107	59	1580	846
Glouton V2	521	181	98	33	1060	281
Glouton V3	911	311	107	53	1580	850
Instantané V1	908	529	103	57	1580	840
Instantané V2	927	528	104	60	1580	848
Instantané V3	985	558	107	61	1580	850

FIGURE 5.1 – Productions obtenues pour le problème déterministe.

Perte de production (%)	Parc 1		Parc 2		Parc 3	
	jour	nuit	jour	nuit	jour	nuit
Plan full-power	0	0	0	0	0	0
Plan statique	2,2	42,4	0	63,6	0	0
Glouton V1	7,0	61,9	0	55,2	0	0,4
Glouton V2	47,8	77,5	7,8	74,9	32,9	67
Glouton V3	8,8	61,4	0	59,5	0	0,4
Instantané V1	9,2	34,4	3,8	56,3	0	1,2
Instantané V2	7,3	34,5	2,9	54,1	0	0,2
Instantané V3	1,4	30,7	0	53,6	0	0

FIGURE 5.2 – Pertes de production vis à vis du plan *full power* pour le problème déterministe.

On peut remarquer que les performances des méthodes gloutonnes sont globalement moins bonnes que la simple application du plan statique. Elles génèrent en général une perte de production, allant jusqu'à 77,5 % vis-à-vis du plan *full power* pour le parc 2 en période de nuit. Concernant le respect de la contrainte sur les médianes temporelles, les méthodes gloutonnes V1 et V3 occasionnent des dépassements comparables au bridage statique (inférieurs à 0,5 dB) dans certains cas (parc 1 et 2 de jour, et parc 3) ou des dépassements supérieurs (parc 1 et 2 de nuit). La méthode gloutonne V2, quant à elle, ne présente aucun dépassement (valeur négative), mais

Nombre de dépassements	Parc 1		Parc 2		Parc 3	
	jour	nuit	jour	nuit	jour	nuit
Plan full-power	188	297	1	159	0	4
Plan statique	144	109	1	49	0	4
Glouton V1	87	217	1	78	0	1
Glouton V2	46	14	0	11	0	1
Glouton V3	129	248	1	95	0	1
Instantané V1	0	14	0	1	0	0
Instantané V2	146	97	0	43	0	4
Instantané V3	146	101	1	50	0	4

FIGURE 5.3 – Nombre de fois où l'émergence réglementaire est dépassée pour le problème déterministe.

Dépassement maximum	Parc 1		Parc 2		Parc 3	
	jour	nuit	jour	nuit	jour	nuit
Plan full-power	1,34	7,17	-0,09	6,57	-4,20	-0,26
Plan statique	0,00	0,09	-0,09	-0,03	-4,20	-0,26
Glouton V1	0,02	2,37	-0,09	1,62	-4,20	-0,26
Glouton V2	-1,03	-0,08	-0,12	-0,30	-4,23	-0,47
Glouton V3	0,20	2,37	-0,09	1,52	-4,20	-0,26
Instantané V1	-0,19	0,20	-0,30	-0,03	-4,20	-0,26
Instantané V2	-0,15	0,00	-0,12	0,24	-4,20	-0,26
Instantané V3	0,02	0,09	-0,09	-0,03	-4,20	-0,26

FIGURE 5.4 – Dépassements maximaux (dB) pour le problème déterministe.

s'avère être la plus conservatrice et la moins productive des méthodes. De ce fait, aucune des méthodes gloutonnes ne convient pour le contrôle.

Concernant les méthodes instantanées, on peut remarquer que toutes les méthodes présentent des dépassements acoustiques semblables ou plus faibles que ceux du plan statique. De plus la production est largement meilleure, bien que les méthodes V1 et V2 entraînent une perte dans la moitié des cas étudiés (parc 1 de jour, parc 2 de jour et parc 3 de nuit). Seule la méthode instantanée V3 présente à la fois de meilleures performances électriques et un dépassement similaire ou plus faible par rapport au plan statique. Cette méthode permet ainsi de récupérer plus de 10% de production électrique par rapport au plan statique (parc 1 et 2 de nuit).

En résumé, pour un contrôle non prédictif sous les contraintes de médianes temporelles, seules les méthodes instantanées ont permis dans ces expériences d'égaliser ou d'améliorer le dépassement. Parmi elles, seule la méthode instantanée V3 permet d'améliorer systématiquement la production sans détériorer le dépassement sur les médianes temporelles.

Interprétation des résultats pour les méthodes gloutonnes

Les mauvaises performances des méthodes gloutonnes peuvent être en partie expliquées par un phénomène « d'emballement » qui apparaît lorsque les conditions de niveau résiduel changent brutalement. Expliquons ce phénomène par un exemple. Considérons un exemple sur un horizon de taille 10 :

1. Supposons que les 4 premières occurrences ($t = 1$ à $t = 4$) de bruit résiduel valent 37 dBA.
2. Supposons qu'à partir de $t = 5$ toutes se trouvent autour de 33 dBA jusqu'en $t = 10$.
3. Supposons que le bruit particulier maximal (plan *full power*) vaut $l_{\text{full}} = 37$ dBA.

On peut alors constater que les algorithmes gloutons (variantes V1 et V3) ne vont pas permettre pour cet exemple un contrôle réglementaire :

— Entre $t = 1$ et $t = 4$, la médiane valant 37 dBA, les algorithmes gloutons vont proposer le plan *full-power*, l'émergence est alors inférieure à 3 dBA : $b_1 = b_2 = b_3 = b_4 = 10\log_{10}(10^{\frac{37}{10}} + 10^{\frac{37}{10}}) \simeq 40$ dBA.

— Entre $t = 5$ et $t = 7$, la médiane du bruit résiduel vaut toujours 37 dBA, de ce fait, les algorithmes gloutons vont proposer le plan *full-power*, car l'émergence réglementaire est respectée pour ce plan. En effet, la médiane du bruit ambiant vaut pour ces valeurs 40 dBA. On a alors $b_5 = b_6 = b_7 = 10\log_{10}(10^{\frac{37}{10}} + 10^{\frac{33}{10}}) \simeq 38,5$ dBA.

— En $t = 8$ la médiane du résiduel commence à chuter et vaut 35 dBA. L'émergence réglementaire se trouve subitement augmentée et les algorithmes gloutons arrêtent alors les éoliennes afin de compenser. On a alors $b_8 = r_8 = 33$ dBA.

— De même en $t = 9$ et $t = 10$, la médiane du résiduel continue de chuter et vaut 33 dBA. La médiane du bruit ambiant, cependant reste toujours élevée, et égale à 39,3 et 38,5 dBA respectivement pour $t = 9$ et $t = 10$.

Cet exemple illustre le problème pour ces deux méthodes, qui en choisissant des valeurs trop élevées de bruit ambiant, peuvent créer des situations qui deviennent incontrôlables.

5.3.4 Interprétation des résultats pour le plan statique et pour les méthodes à émergence instantanée

Concernant les méthodes instantanées, on propose une interprétation des résultats en raisonnant sur la distribution du bruit résiduel et sur la méthode de contrôle utilisée. On cherche ici à expliquer les résultats des plans à émergence instantanée en les comparant au plan statique. Pour ce raisonnement, on rappelle que la propagation est supposée constante dans une classe homogène donnée, et donc le bruit particulier est considéré constant pour une classe homogène et un plan de bridage fixés. Dans la suite, on se place en un point de mesure unique.

On considère les émergences instantanées induites par la méthode considérée (statique, ou émergence instantanée) et par les mesures de bruit résiduel d'un système entre $t = 1$ et $t = T$, où T est suffisamment grand (de l'ordre de la cinquantaine par exemple). Pour simplifier le discours on suppose également $N = T$.

Les Figures 5.5, 5.6 et 5.7 représentent graphiquement l'émergence instantanée associée aux valeurs de résiduel. Pour chacune de ces trois figures, on représente la distribution du bruit résiduel. Il est à noter que la forme de la distribution peut être quelconque, et que seules certaines valeurs clés (la médiane $\hat{r}_{[1,T]}$ et l_{full}) ont leur importance dans cette interprétation.

La coloration représente l'émergence instantanée associée aux valeurs de bruit résiduel :

— en rouge, les valeurs de bruit pour lesquelles l'émergence instantanée est supérieure à 3 dBA,

— en jaune, les valeurs de bruit pour lesquelles l'émergence instantanée est largement inférieure à 3 dBA,

— en vert, les valeurs de bruit pour lesquelles l'émergence instantanée vaut environ 3 dBA.

Pour finir, les doubles flèches sous les distributions indiquent les plages de bruit résiduel pour lesquelles on obtient un gain et/ou une perte de production par rapport au plan statique dans ces plages.

Notons que la variante V2 n'est pas représentée, dans la mesure où elle n'est pas considérée comme une véritable méthode de contrôle.

Effet du plan statique sur la distribution de l'émergence

La Figure 5.5 représente l'effet du bridage statique sur la distribution d'émergence. Sur cet exemple, on suppose en $t = T$ une distribution statistique du bruit résiduel passé ayant pour médiane $\hat{r}_{[1,T]=35}$. Le plan statique est calculé à la manière d'un bureau d'études acoustique en considérant que le bruit ambiant maximum doit être de 38 dBA et en considérant 35 dBA comme bruit résiduel. Le plan statique génère un bruit particulier constant supposé optimal (meilleur plan statique possible). En chaque instant t, le bruit particulier vaut $l_{\text{static}} = 10\log_{10}(10^{\frac{38}{10}} - 10^{\frac{35}{10}}) \simeq 35$ dBA. Pour le bruit particulier à l'instant t, on a $b_t = 10\log_{10}(10^{\frac{l_{\text{static}}}{10}} + 10^{\frac{r_t}{10}})$. On note alors que le bruit ambiant est une fonction croissante de la valeur de r_t. **On notera en outre qu'à chaque instant t la médiane du bruit ambiant est égale au bruit ambiant associé à la médiane du bruit résiduel** [17].

Un constat similaire peut être fait sur l'émergence instantanée. On a $e_t = 10\log_{10}(10^{\frac{l_{\text{static}}}{10}} + 10^{\frac{r_t}{10}}) - r_t$. L'émergence instantanée e_t est donc une fonction décroissante du bruit résiduel. De façon similaire, à chaque instant t, l'émergence réglementaire est égale à l'émergence instantanée associée à la médiane du bruit résiduel. Ce dernier point permet de comprendre les nombreux dépassements qui sont occasionnés lors des tests effectués. Si en $t_0 < T$, la médiane du bruit résiduel en t_0 est inférieure à la médiane finale en $t = T$, c'est-à-dire $\hat{r}_{[1,t_0]} < \hat{r}_{[1,T]}$, alors un dépassement de l'émergence réglementaire est observé. En effet, comme énoncé précédemment, l'émergence réglementaire est égale à l'émergence instantanée associée à la médiane du bruit résiduel. Or, l'émergence instantanée associée à une valeur de résiduel inférieure à $\hat{r}_{[1,T]}$ est nécessairement supérieure à 3 dB. Ceci entraîne un dépassement pour l'émergence réglementaire.

Ce point important souligne l'intérêt des bureaux d'étude à effectuer des mesures dans les périodes où le résiduel est considéré faible. En effet, une condition nécessaire et suffisante pour que le critère basé sur des médianes temporelles soit respecté est de calculer le plan de bridage pendant la période où la médiane du bruit résiduel est au plus bas [18].

Effet du contrôle par l'algorithme instantané V1 sur la distribution de l'émergence

L'algorithme instantané tente de maintenir autant qu'il le peut une émergence instantanée de 3 dBA, lorsque le bruit résiduel est supérieur à 32 dBA, et de maintenir le bruit ambiant à 35 dBA lorsque le bruit résiduel est inférieur à 32 dBA. L'effet sur la distribution d'émergence est représenté sur la Figure 5.6. Cependant, à partir d'une certaine valeur de résiduel ($l_{\text{full-power}}$ dans la figure), l'émergence est inférieure à 3 dBA. En effet, le bruit particulier atteint un maximum noté $l_{\text{full-power}}$, qui est le bruit particulier des éoliennes lorsque le plan est en *full-power*. Pour une valeur de résiduel supérieure au bruit particulier maximum $l_{\text{full-on}}$, le parc fonctionne alors en plan *full-power*. Sur le plan électrique, on obtient par rapport au plan statique une perte électrique associée aux bruits résiduels inférieurs à $\hat{r}_{[1,T]}$, et un gain pour ceux qui lui sont supérieurs. De ce fait, le gain/perte totale est dépendant de la forme de la distribution du parc considéré. Comme pour le plan statique **à l'instant t, l'émergence réglementaire est égale à l'émergence**

17. Cette propriété est intuitive : $b = \mathfrak{f}(r)$ avec $\mathfrak{f}$ fonction croissante, il vient que pour tout $r < \hat{r}$, $\mathfrak{f}(r) < \mathfrak{f}(\hat{r})$. Il y a donc bien 50% des occurrences de bruit ambiant qui sont inférieures à $\mathfrak{f}(\hat{r})$ qui est donc la médiane de la distribution des bruits ambiants.

18. La fenêtre temporelle pendant laquelle est calculée cette médiane est de ce fait importante, la variance de la médiane statistique étant dépendante de celle-ci.

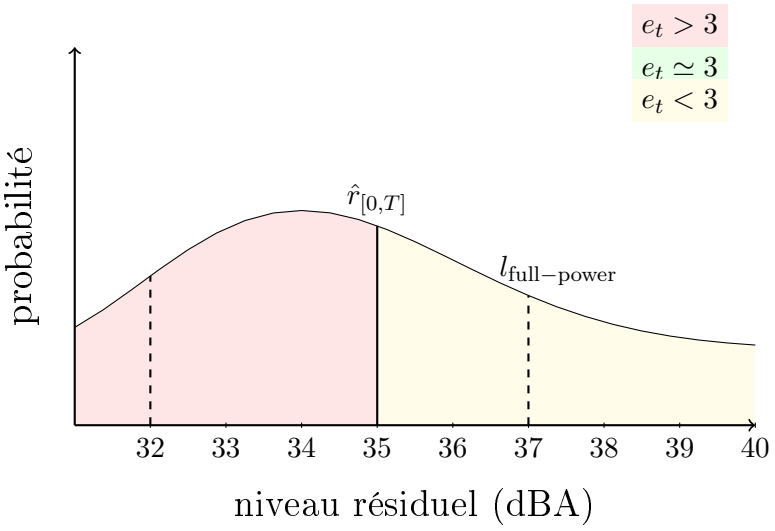


FIGURE 5.5 – Distribution du bruit résiduel et des émergences instantanées engendrées par le plan statique.

instantanée associée à la médiane du bruit résiduel. Si en $t_0 < T$, la médiane du bruit résiduel en t_0 est inférieure à la médiane finale en $t = T$, c'est-à-dire $\hat{r}_{[1,t_0]} < \hat{r}_{[1,T]}$, l'émergence réglementaire est, en raison de cette propriété, invariante.

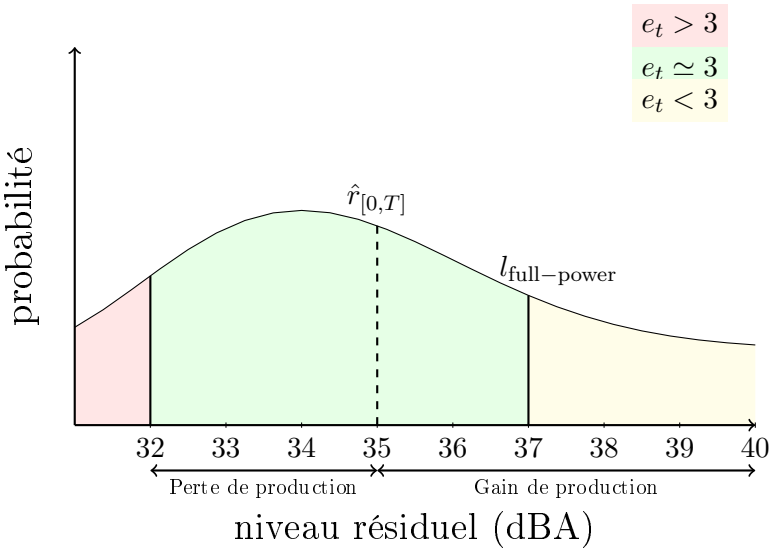


FIGURE 5.6 – Distribution du bruit résiduel et des émergences instantanées engendrées par l'algorithme instantanée V1.

Effet du contrôle par l'algorithme instantané V3 sur la distribution de l'émergence

L'algorithme instantané V3 est une variante hybride entre l'utilisation du plan statique et la variante 1. L'effet sur la distribution d'émergence est représenté sur la Figure 5.7. L'émergence instantanée est régulée à 3 dBA uniquement pour les résiduels supérieurs à $\hat{r}_{[1,T]}$. De ce fait,

aucune perte électrique n'est alors à déplorer pour les résiduels inférieurs à $\hat{r}_{[1,T]}$. Autrement dit, cette méthode consiste à rehausser l'émergence instantanée engendrée par le plan statique lorsque celle-ci est inférieure à 3 dBA. De ce fait, les mêmes dépassements de l'émergence réglementaire sont observés, mais un gain électrique est obtenu.

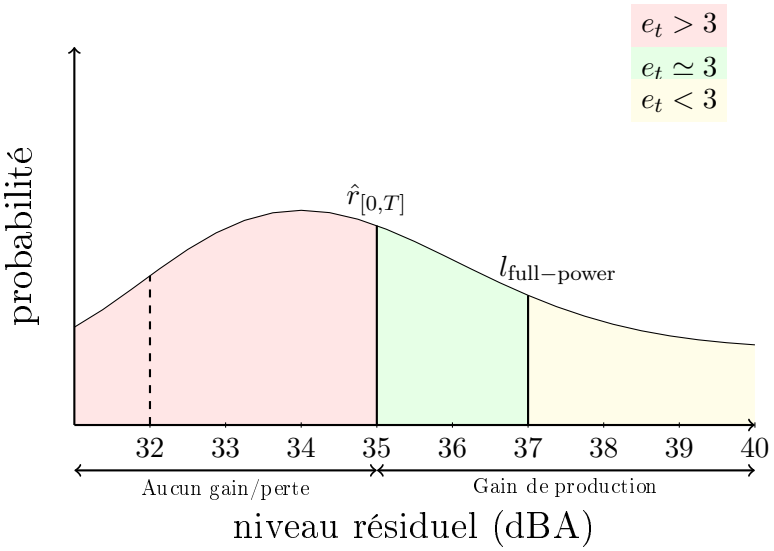


FIGURE 5.7 – Distribution du bruit résiduel et des émergences instantanées engendrées par l'algorithme instantanée V3.

5.4 Conclusion sur le problème déterministe basé sur le critère des médianes temporelles

Limiter l'émergence instantanée à 3 dBA de nuit et 5 dBA de jour paraissait de prime abord trop restrictif. De ce fait, des méthodes non-prédictives gloutonnes ont tout d'abord été proposées. Cependant, l'expérimentation de ces méthodes sur 3 parcs a montré que ces méthodes n'étaient pas performantes. Au contraire, un contrôle limitant l'émergence instantanée s'est avéré plus prometteur dans le cas déterministe et en l'absence de prédiction. Contrôler l'émergence instantanée, en essayant de la ramener à 3 ou 5 dBA en permanence a permis lors de ces tests de respecter plus fidèlement la réglementation et de bénéficier parfois d'un gain électrique allant jusqu'à 8% (parc 1 de nuit). En revanche, si l'amélioration acoustique est systématique, dans certaines situations (en fonction de la distribution du résiduel) des pertes électriques peuvent être observées (jusqu'à 7% sur le parc 1 de jour).

En outre, on a montré qu'en utilisant un fonctionnement hybride utilisant le plan statique par défaut et un plan dynamique rehaussant l'émergence instantanée quand celle-ci est trop faible, une augmentation systématique de la production électrique (jusqu'à 12 % sur le parc 1 de nuit) sans dégradation ou amélioration du respect du critère basé sur des médianes temporelles par rapport au plan statique simple a été obtenue dans tous les cas. Cependant, bien que cette méthode hybride soit intéressante d'un point de vue théorique, il s'agit d'une astuce mathématique qui n'a pas beaucoup d'intérêt en pratique. En effet, cette méthode ne fait en réalité qu'augmenter le dépassement acoustique réel puisque le plan de bridage appliqué génère des niveaux particuliers toujours supérieurs ou égaux à ceux générés par le plan statique. Cette méthode permet toute-

fois de montrer la complexité du critère de conformité actuel et de la limitation pratique qu'il engendre pour un contrôle acoustique. De ce fait et dans toute la suite, on considère un critère instantané, qui s'extrait en partie de la réglementation actuelle. Notons toutefois que l'utilisation des médianes préconisés dans la norme PRS 31-114 est en réalité une adaptation pratique liée au domaine spécifique de l'éolien née de l'incapacité à mesurer finement l'émergence instantanée (voir partie 1.4.1). S'extraire de l'utilisation des médianes et utiliser une contrainte instantanée reste toutefois conforme aux normes plus générale de l'acoustique NF S 31-010 et NF S 31-S110.

Quatrième partie

Problème stochastique

6

Introduction

6.1 Problème et hypothèses

Désormais, on s'intéresse à une version plus générale du problème d'optimisation, à savoir au **problème stochastique (non-prédictif) à contraintes instantanées**. En d'autres termes, on s'intéresse à la gestion des incertitudes dans l'algorithme de décision, tout en conservant un caractère instantané aux contraintes à considérer. Le choix de travailler sur cet axe se justifie scientifiquement sur plusieurs constats.

Contrainte instantanée

Tout d'abord, concernant les contraintes, on a montré dans le chapitre 5 qu'il était possible d'adapter la résolution du problème à contraintes instantanées afin de résoudre le problème à contraintes sur des médianes temporelles. Cependant, cette adaptation est basée sur une astuce mathématique qui permet d'augmenter la production électrique vis-à-vis du plan de bridage statique (actuellement en place sur les parcs), en satisfaisant la contrainte prévue par la loi française mais en augmentant le niveau acoustique moyen. En effet, le bruit particulier obtenu avec cette méthode est égal ou supérieur à celui obtenu par le plan statique. De ce fait, on laisse le problème à contraintes sur les médianes temporelles de côté dans la suite et on considère des contraintes instantanées. On peut également justifier ce choix par le fait que le critère instantané est dans l'absolu un critère de conformité réglementaire plus naturel et plus général. Les possibilités qu'amène un système de contrôle en temps réel changent la manière de traiter un tel problème.

Prédiction

Ensuite, l'aspect non-prédictif est en partie lié au problème stochastique. On verra d'ailleurs que dans les méthodes présentées on considère une forme de prédiction légère, en écrivant le processus comme une marche aléatoire. Cependant, on ne se concentrera pas particulièrement sur cet élément car l'utilisation d'un critère instantané et la bonne réactivité du système rendent pour le moment non indispensable une prédiction à moyen terme (sur quelques heures). En outre, les méthodes actuelles de prédiction à court terme ne permettent pas une amélioration nette sur un horizon de 10 ou 20 minutes, le prédicteur de persistance présenté dans l'état de l'art ayant dans ce cas des performances comparables [Giebel et al., 2011].

Incertitudes

En revanche la gestion des incertitudes est indispensable car les modèles acoustiques sont actuellement peu précis et nécessitent d'être recalés d'une part (erreurs sur l'émission acoustique des éoliennes et/ou sur le calcul de propagation), et le système de séparation de sources iEar est par nature entaché d'une erreur d'estimation d'autre part.

Formulation

Dans la suite, on se place donc sous les hypothèses d'un système instantanément réactif, sans retard d'exécution lorsqu'une éolienne est commandée et sous des contraintes instantanées similaires à celles du chapitre 4 mais où les différentes variables du système sont entachées d'erreurs. On y reviendra plus précisément dans la suite. Mathématiquement, on s'intéresse donc au problème *chance constraint programming* suivant :

$$\begin{cases} \max_{\mathbf{c}_t \in \mathcal{C}} \sum_{i=1}^{I} \mathbb{E}(p_{t,i} | c_{t,i}, \mathbf{v}_t) \\ \forall j \in \{1, ..., J\}, p(b_{t,j} - r_{t,j} < S_e \text{ ou } b_{t,j} < S_s) > 1 - \varepsilon. \end{cases} \tag{6.1}$$

Dans cette formulation, ε est une valeur arbitraire, positive et proche de zéro, qui sera paramétrable dans un système réel, et qui définit la probabilité de non-respect toléré de la contrainte acoustique.

6.2 Démarche

Dans la suite, on décrit la démarche adoptée afin de construire un modèle de contrôle stochastique, qui s'appuie sur les techniques présentées dans l'état de l'art au chapitre 2.

6.2.1 Estimation des incertitudes et suivi des variables cachées

Variables à estimer

On a montré dans la partie III que la connaissance déterministe des données de production et d'émission (données de la fonction $\mathfrak{F}$), de propagation (matrice $\mathbf{A}$) et de résiduel (vecteur $\mathbf{r}$) permet de réaliser le contrôle acoustique des éoliennes par un algorithme d'optimisation. Résoudre le problème stochastique nécessite l'estimation de ces mêmes variables, auxquelles on adjoint une incertitude à estimer. La première étape consiste donc à proposer un schéma d'estimation de ces variables aléatoires et de leurs paramètres.

Fusion de données

Pour estimer les variables acoustiques nécessaires à la résolution du problème, on dispose fondamentalement de deux sources d'information.

La première est une étude acoustique, calculée pour chaque parc. Elle comporte les valeurs $\mathbf{x}_t$ issues données constructeurs, de la matrice $\mathbf{A}_t$ calculée par simulation numérique pour chaque classe homogène et de $\mathbf{r}_t$ calculé comme médiane des mesures dans chaque classe homogène.

La seconde source d'information est constituée des données acquises par les sonomètres et récupérées en temps réel pendant le fonctionnement du parc. Il s'agit d'une part de la mesure du bruit ambiant et d'autre part de l'estimation du bruit particulier et/ou résiduel par séparation de sources.

Ces deux sources d'information peuvent se contredire si on se place dans un cadre déterministe, les incertitudes du modèle et des mesures étant différentes. Il est donc naturel de chercher à utiliser conjointement ces deux sources de données dans un cadre probabiliste afin d'augmenter la précision. On parle alors de fusion de données et l'intérêt de cette approche est triple. D'une part, il n'est pas possible de connaître l'ensemble des variables acoustiques en utilisant seulement les données renvoyées par les sonomètres, la séparation de sources ne calculant pas les contributions individuelles de chaque éolienne ($l_{i,j,t}$). D'autre part, dans certaines situations, il est probable que l'utilisation simultanée de la mesure du bruit ambiant et du modèle acoustique puisse offrir une meilleure précision que la séparation seule en ce qui concerne l'estimation du bruit résiduel [19]. Pour finir, il s'agit d'un problème non-supervisé : il n'est pas possible de connaître la valeur terrain du bruit résiduel afin d'évaluer la qualité de la séparation directement. En effet, parler de séparation n'a de sens que lorsque les éoliennes sont en fonctionnement et que le niveau du résiduel est aléatoire et non connu. Afin d'évaluer la qualité de la séparation, il est nécessaire d'utiliser une approche indirecte, en exploitant un modèle sur le bruit résiduel et sur le système de séparation. En effet, la question n'est pas de savoir si la séparation fonctionne ou non mais quelle est la précision qu'elle apporte dans le cadre du contrôle.

En d'autres termes, il est important de noter qu'il n'existe pas d'approche de mesure ou d'estimation des variables d'état qui soit *a priori* plus fiable qu'une autre. On doit utiliser une approche globale en exploitant toutes les données à notre disposition.

La définition d'un modèle d'incertitude, l'estimation des variables cachées et des paramètres du modèle d'incertitude constitue l'objet du chapitre 7.

6.2.2 Optimisation stochastique

Une fois les incertitudes du problème estimées, on propose d'utiliser un algorithme d'optimisation pour résoudre le problème *chance constrained programing* et construire une commande optimale. Il s'agit d'une adaptation de l'algorithme *branch and bound* déterministe proposé dans le chapitre 4. Ceci constitue l'objet du chapitre 8.

6.3 Schéma général d'optimisation

Afin de faciliter la lecture des deux chapitres suivants, on donne ici un bref descriptif d'une itération de l'algorithme complet de contrôle sous incertitudes :

1. **Fusion de données**. À partir des nouvelles observations lues en entrée à l'instant t (bruit ambiant, estimation du résiduel par séparation) et des paramètres du modèle d'incertitudes estimés jusqu'alors, on met à jour les connaissances sur l'état caché sous la forme d'une distribution de probabilité *a posteriori* sachant cette nouvelle mesure. Il s'agit de l'étape de fusion de données [20] évoquée précédemment où on exploite les meilleures informations accessibles afin d'obtenir la meilleure précision sur les variables du système.
2. **Prédiction** : À partir du modèle d'évolution proposé et de la distribution *a posteriori* de l'état caché, on prédit la distribution de l'état caché à l'instant $t+1$. Les paramètres de cette distribution seront utilisés par l'algorithme d'optimisation.

19. Cas où le rapport signal-à-bruit en dBA entre le niveau résiduel et le niveau particulier est négatif.
20. C'est également l'étape du filtrage de Kalman appelée classiquement "mise à jour " dans la théorie des filtres bayésiens, voir chapitre 2.

3. **Optimisation** : On calcule la commande optimale à partir des paramètres de la distribution prédite de l'état caché afin de résoudre le problème en prenant une marge suffisante. Cette marge dépendra de la tolérance souhaitée (voir équation (6.1)).
4. **Apprentissage** : On utilise un algorithme EM sur des données gardées en mémoire afin de mettre à jour les paramètres du modèle d'incertitude. Cet apprentissage utilise les dernières observations passées et gardées en mémoire jusqu'à un horizon donné. De ce fait, l'apprentissage peut-être effectué en parallèle du calcul d'optimisation.

7

Estimation des incertitudes et recalibration statistique du modèle

7.1 Définition d'un modèle, suivi des variables cachées et estimation des paramètres

Dans ce chapitre, on étend le modèle déterministe à un modèle stochastique par l'ajout d'incertitudes et par un modèle temporel sur le niveau résiduel, résumé par une représentation d'état. En d'autres termes, on décrit le modèle par le système d'équations :

$$\begin{cases} \mathbf{y}_{t+1} = f(\mathbf{y}_t, \mathbf{c}_{t+1}, \mathbf{w}_{\mathbf{y}_t}, \boldsymbol{\theta}_{\mathbf{y}}) \\ \mathbf{z}_t = h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\theta}_{\mathbf{z}}). \end{cases} \tag{7.1}$$

Dans le contexte de la thèse, on définit l'état caché $\mathbf{y}_t$ comme la concaténation des variables acoustiques d'état :

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{a}_t \\ \mathbf{r}_t \end{bmatrix} \tag{7.2}$$

où on rappelle que $\mathbf{x}_t$ est le vecteur des émissions, $\mathbf{a}_t$ est la matrice vectorisée des atténuations et $\mathbf{r}_t$ le vecteur des niveaux résiduels.

De la même façon, on définit les observations $\mathbf{z}_t$ comme la concaténation des variables acoustiques mesurées :

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{b}_t^* \\ \mathbf{r}_t^{*\bullet} \end{bmatrix} \tag{7.3}$$

où on rappelle que $\mathbf{b}_t^*$ est le vecteur des niveaux ambiants mesurés par les sonomètres et $\mathbf{r}_t^{*\bullet}$ est le vecteur des niveaux résiduels estimés par séparation de source.

Les variables $\mathbf{w}_{\mathbf{y}_t}$ et $\mathbf{w}_{\mathbf{z}_t}$, sont respectivement les incertitudes associées au modèle et aux mesures, qui suivent des distributions paramétriques de paramètres $\boldsymbol{\theta}_{\mathbf{y}}$ et $\boldsymbol{\theta}_{\mathbf{z}}$. Les fonctions f et h définissent les relations mathématiques entre les différentes variables. La définition du modèle est décrite dans la partie 7.2.

Ensuite on construit un algorithme EM, basé lui même sur du filtrage récursif bayésien, afin d'assurer le suivi des états cachés $\mathbf{y}_t$ et estimer les paramètres $\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\mathbf{y}}, \boldsymbol{\theta}_{\mathbf{z}}\}$. On présente le schéma d'estimation dans l'ordre suivant :

1. Dans la partie 7.3, on présente une estimation des paramètres $\boldsymbol{\theta}$ et une évaluation expérimentale, en supposant la connaissance des états cachés, ce qui constitue **une simplification de l'étape M** de l'algorithme EM. Il s'agit d'une étape préliminaire dans cette partie qui permet de s'assurer expérimentalement que le problème est suffisamment bien posé et que l'on est capable de résoudre le problème d'estimation des paramètres.
2. Dans la partie 7.4, on présente *a contrario* le suivi des états cachés $\mathbf{y}_t$ par filtrage récursif bayésien ainsi qu'une évaluation expérimentale en supposant connus les paramètres $\boldsymbol{\theta}$, ce qui constitue l'**étape E** de l'algorithme EM. Cette étape permet quand à elle de vérifier la faisabilité du suivi de l'état caché par un filtrage récursif bayésien.
3. Finalement, on présente dans la partie 7.5 l'**algorithme EM complet** ainsi construit et une évaluation expérimentale.

7.2 Modèle d'incertitude

Le problème étant non-supervisé, on doit faire des hypothèses sur la nature des incertitudes. On propose donc dans la suite un modèle paramétrique où les incertitudes sont des variables aléatoires gaussiennes. On s'intéresse en effet, par simplicité et pour des raisons historiques en acoustique, à la moyenne et à la variance des incertitudes. De ce fait, la loi gaussienne constitue un bon choix dans la mesure où il s'agit de la loi de probabilité qui fait le moins d'hypothèses sur les incertitudes, étant donnée qu'elle maximise l'entropie [Bishop, 2007].

Dans une telle modélisation, il est courant de considérer que la moyenne des incertitudes correspond au biais sur le modèle et sur la mesure. Dans toute la suite, on ne s'intéresse qu'à l'estimation de la variance car on suppose que le modèle a été bien calibré à l'origine de sorte à être non-biaisé. L'étude du biais fait l'objet des perspectives de la thèse, mentionnées au chapitre 9. La moyenne des incertitudes est donc supposée égale à la valeur déterministe dans toute la suite, et les paramètres $\boldsymbol{\theta}$ à estimer se résument aux paramètres de variance du modèle d'incertitude.

Incertitudes sur le modèle acoustique

Comme on dispose d'un modèle pour l'émission et la propagation, on suppose dans la suite que les variables suivent en moyenne les valeurs issues des tables constructeurs et des simulations données pour la classe homogène courante $\mathbf{o}_t$ mais qu'elles sont toutes entachées d'une erreur aléatoire.

Le vecteur d'émission $\mathbf{x}_t$ est supposé incertain à l'instant t et suivant une loi gaussienne. Le vecteur des moyennes $\mathbf{m}_\mathbf{x}$ qui dépend de $\mathbf{c}_t$ et de o_t et est supposé connu et fixé d'après les données constructeur. La matrice de covariance $\boldsymbol{\Sigma}_\mathbf{x}$ est supposée indépendante de la commande $\mathbf{c}_t$ et éventuellement dépendante de la classe o_t. De plus, on suppose que les x_i ne sont pas corrélés entre eux et que tous les x_i ont une variance identique [21] σ_x^2, c'est-à-dire

$$\mathbf{x}_t = \mathbf{m}_{\mathbf{x}_t} + \mathbf{w}_{\mathbf{x}_t} \tag{7.4}$$

avec

$$\mathbf{w}_{\mathbf{x}_t} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{x}) \tag{7.5}$$

et

$$\boldsymbol{\Sigma}_\mathbf{x} = \sigma_x^2 \boldsymbol{I}_I \tag{7.6}$$

21. NB : on ne considère pas les x_i identiques mais seulement leur variance.

où $\boldsymbol{I}_I$ désigne la matrice identité de taille I. L'hypothèse de non-corrélation est simplificatrice, dans la mesure où les erreurs pourraient être corrélées à l'incertitude sur la vitesse du vent par exemple. Cependant, on supposera que celle-ci est négligeable.

De façon similaire, on considère que les éléments $a_{t,i,j}$ de la matrice de propagation $\mathbf{A}_t$ ont une moyenne connue par simulation et déterminée par la classe courante o_t et sont entachés chacun d'une incertitude gaussienne centrée :

$$\mathbf{a}_t = \mathbf{m}_{\mathbf{a}_t} + \mathbf{w}_{\mathbf{a}_t} \tag{7.7}$$

avec

$$\mathbf{w}_{\mathbf{a}_t} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{a}) \tag{7.8}$$

et

$$\mathbf{\Sigma}_\mathbf{a} = \sigma_a^2 \boldsymbol{I}_{IJ} \tag{7.9}$$

où on rappelle que $\mathbf{a}_t$ désigne la vectorisation de la matrice $\mathbf{A}_t$.

En effet, comme pour $\mathbf{x}_t$, on supposera que les composantes de $\mathbf{A}_t$ sont des variables aléatoires de variance identique et non-corrélées entre elles. La non-corrélation des incertitudes sur les variables de propagation peut se justifier par l'éloignement relatif des points de mesure en pratique qui se traduit par des chemins acoustiques différents.

En résumé, les variables d'état sont modélisées comme des variables gaussiennes indépendantes, et dont la moyenne dépend de la classification et/ou de la commande de contrôle. Pour ces variables, deux paramètres sont à déterminer : σ_x et σ_a. On supposera en outre que ces variables sont décorrélées dans le temps.

Cas particulier du résiduel

Dans les études acoustiques, les classes homogènes sont usuellement utilisées pour décrire les niveaux résiduels, dans la mesure où une corrélation entre le niveau résiduel et la vitesse du vent existe. Cependant, comme présenté dans la Figure 7.1 issue d'une étude acoustique, la variance des indicateurs résiduels est élevée et ne permet pas de modéliser les évolutions temporelles finement.

A contrario, on peut constater qu'en pratique il existe des similarités temporelles pour les valeurs du résiduel d'un jour à un autre comme dans l'exemple de la Figure 7.2. On observe sur cette figure l'existence probable de tendances à long terme des niveaux résiduels. Notamment, on distingue généralement une période de nuit et une période de jour marquée, comme présenté dans la Figure 7.3, avec des transitions fortes prononcées en début et parfois en fin de journée. Cependant, comme présenté dans l'état de l'art au chapitre 2, il est en général difficile d'obtenir une amélioration significative avec des modèles de prédiction à long terme par rapport à l'estimateur de persistance sur un horizon de l'ordre de 10 minutes [Giebel et al., 2011]. D'autre part, comme présenté sur la Figure 7.4, la distribution de l'évolution $r_t - r_{t-1}$ suit en pratique une distribution vraisemblablement unimodale et centrée.

On propose donc dans la suite d'utiliser un modèle temporel simple de marche aléatoire pour le niveau résiduel. La valeur à chaque instant ne dépend que de l'instant précédent et on considère que la variance de l'erreur d'évolution est bornée :

$$\mathbf{r}_t = \mathbf{r}_{t-1} + \mathbf{w}_{\mathbf{r}_t}. \tag{7.10}$$

La prédiction fine de l'évolution temporelle fait partie des perspectives ultérieures de la thèse.

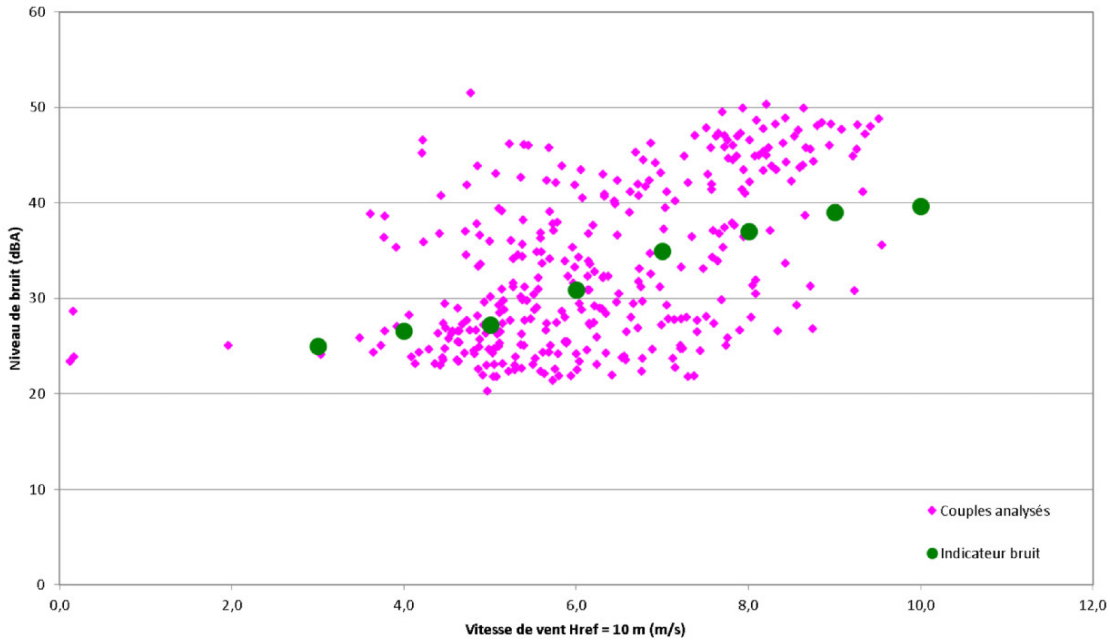


FIGURE 7.1 – Niveaux résiduels observés lors d'une étude acoustique en fonction de la vitesse du vent (losanges roses). Les indicateurs retenus lors de l'étude (ronds verts) ne permettent pas un contrôle optimal du parc éolien.

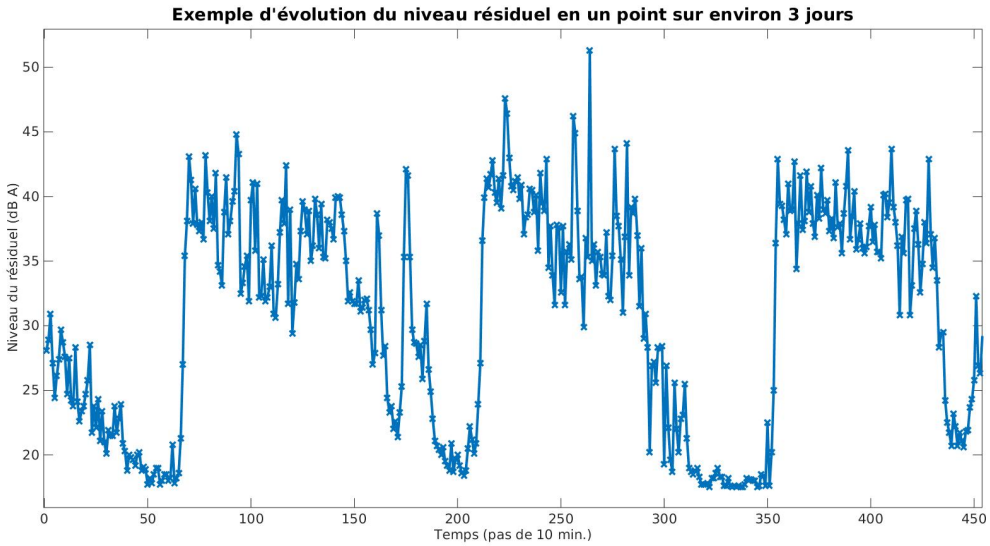


FIGURE 7.2 – Évolution temporelle du niveau résiduel en un point.

Incertitudes sur la mesure et la séparation de sources.

Concernant la mesure, on considère, comme mentionné dans le chapitre 3, deux types de mesures. La mesure sonométrique $\mathbf{b}_t^*$ mesure le bruit ambiant total $\mathbf{b}_t$ à une incertitude de

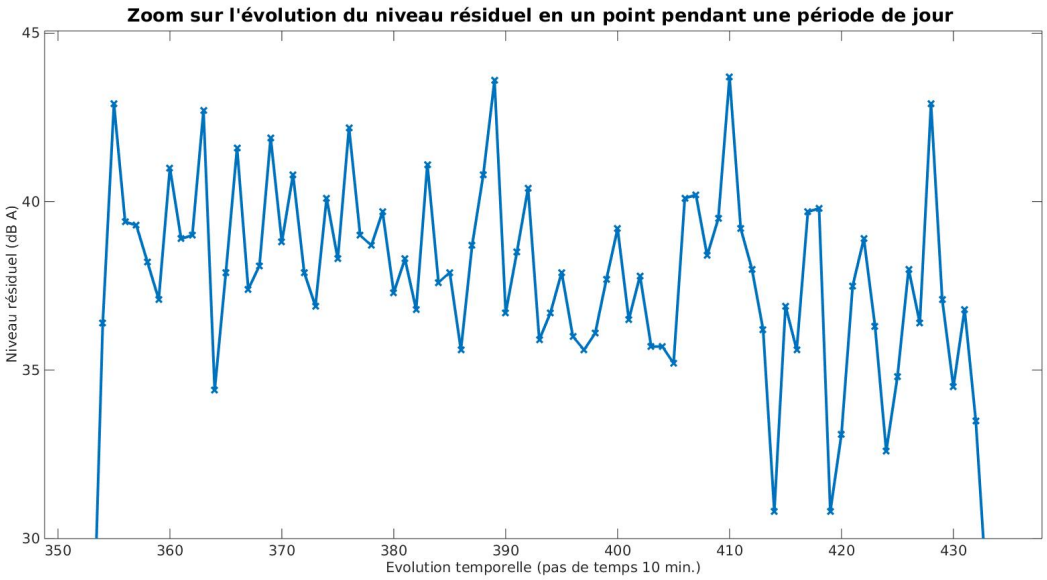


FIGURE 7.3 – Exemple de l'évolution temporelle du niveau résiduel en un point. Zoom sur une période de jour.

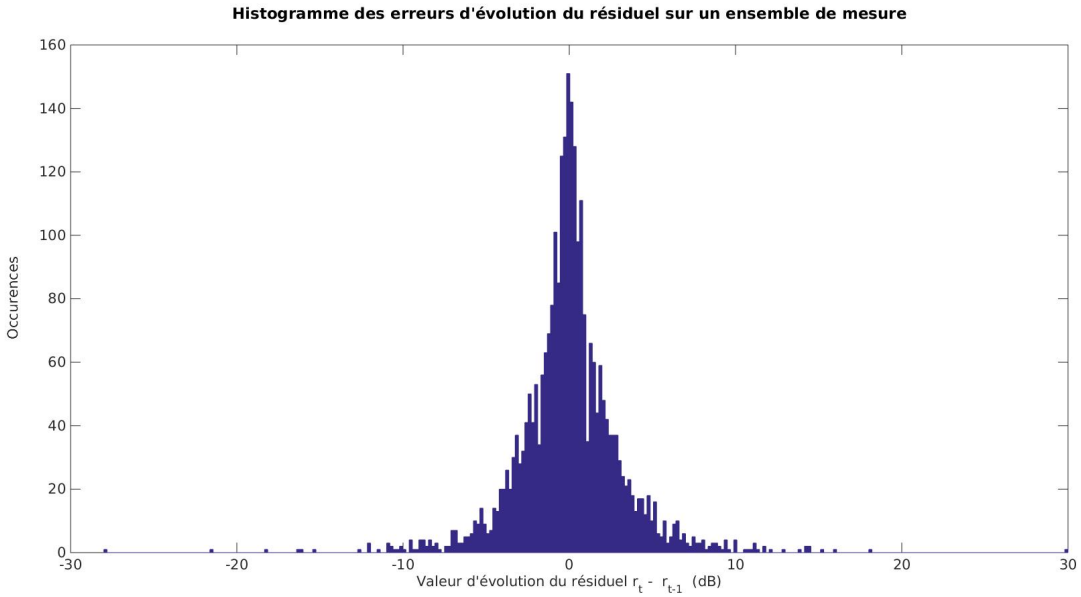


FIGURE 7.4 – Histogramme des évolutions du résiduel sur des données mesurées. Notons que cette distribution est centrée (moyenne de 0 dB) et symétrique (asymétrie statistique $\frac{\mu_3}{\mu_2^{\frac{3}{2}}} = -0,14$, où μ_i est le moment centré d'ordre i).

l'ordre d'1 dB près [Norme, 1996]. On la modélise ici comme une incertitude gaussienne :

$$\mathbf{b}_t^* = \mathbf{b}_t + \mathbf{w}_{\mathbf{b}_t^*} \tag{7.11}$$

avec

$$\mathbf{w}_{\mathbf{b}_t^*} \sim \mathcal{N}(\mathbf{0}, \sigma_{b^*}^2 \boldsymbol{I}_J). \tag{7.12}$$

D'autre part, on estime également le niveau résiduel par séparation de sources. On suppose que le système estime correctement le bruit résiduel en moyenne mais qu'il souffre à chaque instant d'une erreur de variance à déterminer. Cette hypothèse est *a priori* réductrice car la qualité de de la séparation dépend usuellement du rapport signal-à-bruit [Tran et al., 2015] et il est possible que le système soit biaisé (mauvaise calibration). Cependant, ce modèle constitue une première étape et un modèle paramétré par le rapport signal-à-bruit *a priori* pourrait être utilisé éventuellement pour étendre le travail présenté. On modélise la corrélation de cette incertitude avec celle de la mesure sonométrique en posant

$$\mathbf{r}_t^{*\bullet} = \mathbf{r} + \mathbf{w}_{\mathbf{r}_t^*} + \mathbf{w}_{\mathbf{r}_t^\bullet}. \tag{7.13}$$

On décompose en effet l'incertitude en deux composantes. On a d'une part une composante suivant la loi gaussienne

$$\mathbf{w}_{\mathbf{r}_t^\bullet} \sim \mathcal{N}(\mathbf{0}, \sigma_{r^\bullet}^2 \boldsymbol{I}_J), \tag{7.14}$$

qui constitue l'incertitude propre à l'algorithme de séparation. On a d'autre part une composante qui est parfaitement corrélée à la mesure sonométrique et qui vérifie donc

$$\mathbf{w}_{\mathbf{r}_t^*} = \mathbf{w}_{\mathbf{b}_t^*}. \tag{7.15}$$

Ces deux composantes, étant de natures différentes, sont supposées non-corrélées entre elles :

$$\mathrm{Cov}(\mathbf{w}_{\mathbf{r}_t^\bullet}, \mathbf{w}_{\mathbf{b}_t^*}) = \mathbf{0}. \tag{7.16}$$

Les matrices de covariance sont par nature diagonales : les erreurs de mesure et de séparation sont propres à chaque microphone et donc indépendantes.

Notons que le paramètre σ_{b^*} est supposé connu grâce aux valeurs d'incertitudes fournies par le fabriquant des sonomètres et seul $\sigma_{r^\bullet}$ est estimé.

7.3 Étape M simplifiée : estimation des paramètres du modèle d'incertitude sous la connaissance de l'état caché

Dans cette partie, on se concentre dans un premier temps sur l'estimation des paramètres σ_x, σ_a, $\sigma_r^\bullet$ du modèle lorsque le niveau du bruit résiduel est supposé connu. Il s'agit d'une simplification du problème général.

On suppose qu'on dispose d'un ensemble de données observées sur une période de temps $\{1, ..., T\}$ de l'ordre d'une journée par exemple. D'autre part, dans le cadre du problème simplifié, **on suppose que $\mathbf{r}_t$ est connu de manière déterministe** (ou bien encore que les lois de probabilité sont conditionnelles à $\mathbf{r}_t$, sans toutefois l'écrire explicitement afin d'alléger les écritures). On propose dans cette partie d'estimer les incertitudes des variables du problème. On verra dans les parties 7.4 et 7.5 comment estimer la distribution du vecteur $\mathbf{r}_t$ et comment ré-adapter l'estimation décrite ici.

7.3.1 Estimateur de vraisemblance conditionnellement au bruit résiduel

On cherche à déterminer les trois paramètres d'incertitude du modèle, σ_x, σ_a et $\sigma_{r^\bullet}$, respectivement sur l'émission, la propagation et la séparation de sources, par estimation du maximum de vraisemblance.

Pour ce faire, on exprime le problème de maximisation de la vraisemblance des T vecteurs d'observation $\mathbf{z}_t = [\mathbf{b}_t^* \ \mathbf{r}_t^{*\bullet}]^T$ en fonction de σ_x, σ_a et $\sigma_{r^\bullet}$:

$$\max_{(\mathcal{L}_\sigma)} \ p(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T^* | \mathbf{c}_1, o_1, \mathbf{c}_2, o_2, ..., \mathbf{c}_T, o_T, \sigma_d, \\ \mathcal{L}_m = \{\mathbf{m}_{\mathbf{x}_{c,o}}, \mathbf{m}_{\mathbf{A}_o} | (c,o) \in \mathcal{C} \times \mathcal{O}\}, \mathcal{L}_r = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_T, \}, \mathcal{L}_\sigma = \{\sigma_x, \sigma_a, \sigma_{r^\bullet}\}). \quad (7.17)$$

On rappelle qu'on optimise uniquement sur l'espace $\mathcal{L}_\sigma$ car on suppose que $\mathcal{L}_r$ et $\mathcal{L}_m$ sont connus. Étant donnée un vecteur aléatoire $\mathbf{z}$ de loi gaussienne multivariée, la vraisemblance de $\mathbf{z}$ peut être facilement exprimée à partir de sa moyenne $\mathbf{m}_\mathbf{z}$ et sa matrice de covariance $\mathbf{\Sigma}_\mathbf{z}$:

$$p(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T | \mathbf{m}_{\mathbf{z}_1}, ..., \mathbf{m}_{\mathbf{z}_T}, \mathbf{\Sigma}_{\mathbf{z}_1}, ..., \mathbf{\Sigma}_{\mathbf{z}_T}) = \prod_{t=1}^{T} \frac{1}{(2\pi)^{\frac{2J}{2}}} \frac{1}{|\mathbf{\Sigma}_{\mathbf{z}_t}|^{\frac{1}{2}}} \\ \times \exp\left[-\frac{1}{2} (\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_t})^T \mathbf{\Sigma}_{\mathbf{z}_t}^{-1} (\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_t}) \right]. \quad (7.18)$$

Comme souvent, il est plus simple de considérer le logarithme de la vraisemblance, qui ne change pas la valeur du maximum :

$$\log p(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T | \mathbf{m}_{\mathbf{z}_1}, ..., \mathbf{m}_{\mathbf{z}_T}, \mathbf{\Sigma}_{\mathbf{z}_1}, ..., \mathbf{\Sigma}_{\mathbf{z}_T}) = -\frac{2JN}{2} \log(2\pi) \\ - \frac{1}{2} \sum_{t=1}^{T} \log |\mathbf{\Sigma}_{\mathbf{z}_t}| + (\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_t})^T \mathbf{\Sigma}_{\mathbf{z}_t}^{-1} (\mathbf{z}_t - \mathbf{m}_{\mathbf{z}_t}). \quad (7.19)$$

Le problème se ramène alors à exprimer le vecteur $\mathbf{m}_\mathbf{z}$ et la matrice $\mathbf{\Sigma}_\mathbf{z}$ en fonction de σ_x, σ_a et $\sigma_{r^\bullet}$.

Calcul du vecteur moyen et de la matrice de covariance

À partir des hypothèses formulées dans la partie (7.2), on peut écrire :

$$\mathbf{z} = \begin{bmatrix} \mathbf{b}^* \\ \mathbf{r}^{*\bullet} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{r} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_{\mathbf{b}^*} \\ \mathbf{w}_{\mathbf{b}^*} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_{\mathbf{r}^\bullet} \end{bmatrix}. \quad (7.20)$$

On peut alors calculer le vecteur moyen et la matrice de covariance. Le vecteur moyen se calcule comme

$$\mathbf{m}_\mathbf{z} = \mathbb{E}\left(\begin{bmatrix} \mathbf{b} \\ \mathbf{r} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_{\mathbf{b}^*} \\ \mathbf{w}_{\mathbf{b}^*} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_{\mathbf{r}^\bullet} \end{bmatrix} \right) \quad (7.21) \\ = \begin{bmatrix} \mathbf{m}_\mathbf{b} \\ \mathbf{r} \end{bmatrix}$$

(on rappelle que $\mathbf{r}$ est connu et déterminé). La matrice de covariance, quant-à-elle, peut se calculer par la relation

$$\mathrm{Cov}\left(\begin{bmatrix} \mathbf{b}^* \\ \mathbf{r}^{*\bullet} \end{bmatrix} \right) = \begin{bmatrix} \mathrm{Cov}(\mathbf{b}^*, \mathbf{b}^*) & \mathrm{Cov}(\mathbf{r}^{*\bullet}, \mathbf{b}^*) \\ \mathrm{Cov}(\mathbf{b}^*, \mathbf{r}^{*\bullet}) & \mathrm{Cov}(\mathbf{r}^{*\bullet}, \mathbf{r}^{*\bullet}) \end{bmatrix} \quad (7.22)$$

où par indépendance de $\mathbf{b}$ et $\mathbf{w}_{\mathbf{b}^*}$, on a

$$\mathrm{Cov}(\mathbf{b}^*, \mathbf{b}^*) = \mathrm{Cov}(\mathbf{b}) + \mathrm{Cov}(\mathbf{w}_{\mathbf{b}^*}), \quad (7.23)$$

par indépendance de $\mathbf{w_{b^*}}$ et $\mathbf{w_{r^\bullet}}$, on a

$$\mathrm{Cov}(\mathbf{r}^{*\bullet}, \mathbf{r}^{*\bullet}) = \mathrm{Cov}(\mathbf{w_{b^*}}) + \mathrm{Cov}(\mathbf{w_{r^\bullet}}), \tag{7.24}$$

et par indépendance entre $\mathbf{A}$ et $\mathbf{x}$, on a

$$\begin{aligned}
\mathrm{Cov}(\mathbf{r}^{*\bullet}, \mathbf{b}^*) = \mathrm{Cov}(\mathbf{b}^*, \mathbf{r}^{*\bullet}) &= \mathbb{E}(\mathbf{b}^*\mathbf{r}^{*\bullet T}) - \mathbb{E}(\mathbf{b}^*)\mathbb{E}(\mathbf{r}^{*\bullet})^T \\
&= \mathbb{E}((\mathbf{b} + \mathbf{w_{b^*}})(\mathbf{r} + \mathbf{w_{b^*}} + \mathbf{w_{r^\bullet}})^T) - \mathbb{E}(\mathbf{b})\mathbf{r}^T \\
&= \mathbb{E}(\mathbf{b}\mathbf{r}^T + \mathbf{b}\mathbf{w}_{\mathbf{b}^*}^T + \mathbf{b}\mathbf{w}_{\mathbf{r}^\bullet}^T + \mathbf{w_{b^*}}\mathbf{r}^T + \mathbf{w_{b^*}}\mathbf{w}_{\mathbf{b}^*}^T + \mathbf{w_{b^*}}\mathbf{w}_{\mathbf{r}^\bullet}^T) - \mathbb{E}(\mathbf{b})\mathbf{r}^T \\
&= \mathbb{E}(\mathbf{b})\mathbf{r}^T + \mathbf{0} + \mathbf{0} + \mathbf{0} + \mathbb{E}(\mathbf{w_{b^*}}\mathbf{w}_{\mathbf{b}^*}^T) + \mathbf{0} - \mathbb{E}(\mathbf{b})\mathbf{r}^T \\
&= \mathrm{Cov}(\mathbf{w_{b^*}}).
\end{aligned} \tag{7.25}$$

En résumé, on obtient l'expression

$$\begin{aligned}
\mathbf{\Sigma_z} &= \begin{bmatrix} \mathrm{Cov}(\mathbf{b}) + \mathrm{Cov}(\mathbf{w_{b^*}}) & \mathrm{Cov}(\mathbf{w_{b^*}}) \\ \mathrm{Cov}(\mathbf{w_{b^*}}) & \mathrm{Cov}(\mathbf{w_{b^*}}) + \mathrm{Cov}(\mathbf{w_{r^\bullet}}) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{\Sigma_b} + \sigma_{b^*}^2 \boldsymbol{I}_J & \sigma_{b^*}^2 \boldsymbol{I}_J \\ \sigma_{b^*}^2 \boldsymbol{I}_J & \sigma_{b^*}^2 \boldsymbol{I}_J + \sigma_{r^\bullet}^2 \boldsymbol{I}_J \end{bmatrix}.
\end{aligned} \tag{7.26}$$

Il reste à exprimer $\mathbf{m_b}$ et $\mathbf{\Sigma_b}$ en fonction de σ_x, σ_a et $\sigma_{r^\bullet}$. Pour cela, on présente deux méthodes différentes : une méthode par linéarisation de Taylor, et une méthode d'approximation des moments basées sur des formules de passage de la loi normale à la loi log-normale.

7.3.2 Estimation par linéarisation de Taylor à l'ordre 1

La vraisemblance peut être approchée en linéarisant autour des moyennes de $\mathbf{x}$ et $\mathbf{a}$ (vectorisation de $\mathbf{A}$). Cette manière de procéder est similaire à la linéarisation utilisée pour le filtre de Kalman étendu qui est utilisé dans l'étape E de l'algorithme EM (voir partie 7.4). On pose h_b la fonction qui à $[\mathbf{x}\ \mathbf{a}]^T$ associe $\mathbf{b}$:

$$\mathbf{b} = h_b\left(\begin{bmatrix}\mathbf{x}\\ \mathbf{a}\end{bmatrix}\right). \tag{7.27}$$

On définit également $\mathbf{J}_{h_b}$ comme la matrice jacobienne de h_b prise en $\mathbf{m_y}$:

$$\mathbf{J}_{h_b} = \begin{bmatrix} \frac{\partial h_b}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\mathbf{m_x}} & \frac{\partial h_b}{\partial \mathbf{a}}\Big|_{\mathbf{a}=\mathbf{m_a}} \end{bmatrix}. \tag{7.28}$$

En utilisant un développement de Taylor tronqué au premier ordre de cette relation au point $\mathbf{m}_y$, on a l'approximation

$$\mathbf{b} \simeq h_b\left(\begin{bmatrix}\mathbf{m_x}\\ \mathbf{m_a}\end{bmatrix}\right) + \mathbf{J}_{h_b}\begin{bmatrix}\mathbf{w_x}\\ \mathbf{w_a}\end{bmatrix}, \tag{7.29}$$

dont il découle alors que

$$\mathbf{m_b} \simeq h_b\left(\begin{bmatrix}\mathbf{m_x}\\ \mathbf{m_a}\end{bmatrix}\right) \tag{7.30}$$

et

$$\mathbf{\Sigma_b} \simeq \mathbf{J}_{h_b}\begin{bmatrix}\mathbf{\Sigma_x} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma_a}\end{bmatrix}\mathbf{J}_{h_b}^T. \tag{7.31}$$

7.3.3 Estimation par calcul des moments

Une autre manière de procéder est d'utiliser une méthode de calcul des moments. Cette méthode est basée sur les formules de calcul des moments données dans l'annexe A, pour calculer les moments d'ordre 1 et 2 de l'exponentielle d'une variable aléatoire gaussienne suivant alors des loi log-normale. Les variables aléatoires étant additives en échelle linéaire (sommation énergétique) il devient alors possible d'exprimer les moments de $\mathbf{b}$ en fonction de ceux de $\mathbf{x}$ et $\mathbf{A}$. On procède en trois étapes : on exprime les moments de $\tilde{\mathbf{x}}$, $\tilde{\mathbf{A}}$ et $\tilde{\mathbf{r}}$ en fonction de ceux de $\mathbf{x}$, $\mathbf{A}$ et $\mathbf{r}$, puis les moments de $\tilde{\mathbf{b}}$ en fonction de ceux de $\tilde{\mathbf{x}}$, $\tilde{\mathbf{A}}$ et $\tilde{\mathbf{r}}$ et pour finir les moments de $\mathbf{b}$ en fonction de ceux de $\tilde{\mathbf{b}}$.

Conversion des variables d'état en échelle linéaire

On commence par exprimer les différentes variables en échelle linéaire. Pour $\mathbf{r}$, il suffit d'utiliser la définition du dB car ce n'est pas un vecteur aléatoire :

$$\tilde{r}_j = p_0^2 \exp\left(\frac{r_j}{k}\right) \tag{7.32}$$

où $k = \frac{10}{\log(10)}$. Pour $\mathbf{x}$ et $\mathbf{A}$ qui sont des variables aléatoires, on utilise les formules de passage définies dans l'annexe A. Pour $\tilde{\mathbf{x}}$, on obtient

$$m_{\tilde{x}_i} = p_0^2 \exp\left(\frac{m_{x_i}}{k} + \frac{\sigma_x^2}{2k^2}\right) \tag{7.33}$$

$$\sigma_{\tilde{x}_{i_1,i_2},\tilde{x}_{i_1,i_2}} = 0 \tag{7.34}$$

si $i_1 \neq i_2$, et

$$\sigma_{\tilde{x}_i}^2 = m_{\tilde{x}_i}^2 \left[\exp\left(\frac{\sigma_x^2}{k^2}\right) - 1\right] \tag{7.35}$$

sinon. De la même façon, pour $\mathbf{a}$, on obtient

$$m_{\tilde{a}_{i,j}} = p_0^2 \exp\left(\frac{m_{a_{i,j}}}{k} + \frac{\sigma_a^2}{2k^2}\right), \tag{7.36}$$

$$\sigma_{\tilde{a}_{i_1,j_1},\tilde{a}_{i_2,j_2}} = 0 \tag{7.37}$$

si $i_1 \neq i_2$ ou $j_1 \neq j_2$, et

$$\sigma_{\tilde{a}_{i,j}}^2 = m_{\tilde{a}_{i,j}}^2 \left[\exp\left(\frac{\sigma_a^2}{k^2}\right) - 1\right] \tag{7.38}$$

sinon.

Calcul des moments du bruit ambiant en fonction des variables d'état en échelle linéaire

On exploite ensuite les relations entre ces variables et $\tilde{\mathbf{b}}$:

$$\tilde{\mathbf{b}} = \tilde{\mathbf{A}}^T \tilde{\mathbf{x}} + \tilde{\mathbf{r}}. \tag{7.39}$$

On en déduit

$$\mathbf{m}_{\tilde{\mathbf{b}}} = \mathbf{M}_{\tilde{\mathbf{A}}}^T \mathbf{m}_{\tilde{\mathbf{x}}} + \tilde{\mathbf{r}}, \tag{7.40}$$

où $\mathbf{M}_{\tilde{\mathbf{A}}}$ est la matrice contenant la moyenne des $\tilde{a}_{i,j}$ et

$$\begin{aligned}\mathbf{\Sigma}_{\tilde{\mathbf{b}}} =&\mathbb{E}((\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}+\tilde{\mathbf{r}})(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}+\tilde{\mathbf{r}})^T)-\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}+\tilde{\mathbf{r}})\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}+\tilde{\mathbf{r}})^T\\ =&\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\tilde{\mathbf{A}}+\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}\tilde{\mathbf{r}}^T+\tilde{\mathbf{r}}\tilde{\mathbf{x}}^T\tilde{\mathbf{A}}+\tilde{\mathbf{r}}\tilde{\mathbf{r}}^T)\\ &-\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})^T-\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})\tilde{\mathbf{r}}^T-\tilde{\mathbf{r}}\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})^T-\tilde{\mathbf{r}}\tilde{\mathbf{r}}^T\\ =&\mathbb{E}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\tilde{\mathbf{A}})+\mathbb{E}(\tilde{\mathbf{A}}^T)\mathbb{E}(\tilde{\mathbf{x}})\tilde{\mathbf{r}}^T+\tilde{\mathbf{r}}\mathbb{E}(\tilde{\mathbf{x}}^T)\mathbb{E}(\tilde{\mathbf{A}})+\tilde{\mathbf{r}}\tilde{\mathbf{r}}^T\\ &-\mathbb{E}(\tilde{\mathbf{A}})^T\mathbb{E}(\tilde{\mathbf{x}})\mathbb{E}(\tilde{\mathbf{x}})\mathbb{E}(\tilde{\mathbf{A}})^T-\mathbb{E}(\tilde{\mathbf{A}})^T\mathbb{E}(\tilde{\mathbf{x}})\tilde{\mathbf{r}}^T-\tilde{\mathbf{r}}\mathbb{E}(\tilde{\mathbf{x}})^T\mathbb{E}(\tilde{\mathbf{A}})-\tilde{\mathbf{r}}\tilde{\mathbf{r}}^T\\ =&\mathrm{Cov}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})\end{aligned} \tag{7.41}$$

car $\tilde{\mathbf{A}}$, $\tilde{\mathbf{x}}$ sont indépendants et $\tilde{\mathbf{r}}$ est déterministe. Pour le développement suivant, on note $\mathrm{Cov}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})[j_1,j_2]$, avec $(j_1,j_2)\in\{1,\dots,J\}\times\{1,\dots,J\}$, les indices de la matrice $\mathrm{Cov}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})$. On a alors deux cas :

1. Si $j_1=j_2$, on a :

$$\begin{aligned}\mathrm{Cov}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})[j_1,j_2] &= \mathbb{E}\left[\left(\sum_{i=1}^{I}\tilde{a}_{i,j_1}\tilde{x}_i\right)^2\right]-\left(\sum_{i=1}^{I}m_{\tilde{a}_{i,j_1}}m_{\tilde{x}_i}\right)^2\\ &= \sum_{i=1}^{I}(\sigma^2_{\tilde{a}_{i,j_1}}\sigma^2_{\tilde{x}_i}+m^2_{\tilde{a}_{i,j_1}}\sigma^2_{\tilde{x}_i}+\sigma^2_{\tilde{a}_{i,j_1}}m^2_{\tilde{x}_i}).\end{aligned}$$

2. Si $j_1\neq j_2$, on a :

$$\begin{aligned}\mathrm{Cov}(\tilde{\mathbf{A}}^T\tilde{\mathbf{x}})[j_1,j_2] &= \mathbb{E}\left[\left(\sum_{i=1}^{I}\tilde{a}_{i,j_1}\tilde{x}_i\right)\left(\sum_{i=1}^{I}\tilde{a}_{i,j_2}\tilde{x}_i\right)\right]-\left(\sum_{i=1}^{I}m_{\tilde{a}_{i,j_1}}m_{\tilde{x}_i}\right)\left(\sum_{i=1}^{I}m_{\tilde{a}_{i,j_2}}m_{\tilde{x}_i}\right)\\ &= \sum_{i=1}^{I}m_{\tilde{a}_{i,j_1}}m_{\tilde{a}_{i,j_2}}\sigma^2_{\tilde{x}_i}.\end{aligned}$$

En résumé :

$$\mathrm{Cov}(\tilde{\mathbf{A}}\tilde{\mathbf{x}})=\mathbf{M}^T_{\tilde{\mathbf{A}}}\mathbf{\Sigma}_{\tilde{\mathbf{x}}}\mathbf{M}_{\tilde{\mathbf{A}}}+\mathrm{diag}\left(\sum_{i=1}^{I}\sigma^2_{a_{i,j}}(\sigma^2_{\tilde{x}_i}+m^2_{\tilde{x}_i})\right)_{1\leq j\leq J}. \tag{7.42}$$

Calcul des moments du bruit ambiant en échelle logarithmique

On peut alors récupérer l'expression approchée de $\mathbf{m_b}$ et $\mathbf{\Sigma_b}$ à partir des formules de passage inverses définies dans l'annexe A :

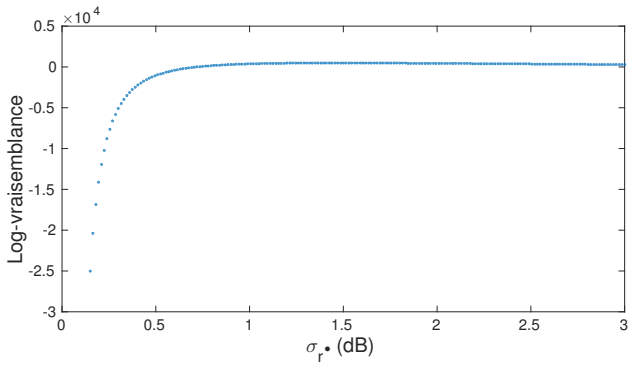
$$m_{b_j}\simeq k\times\left[\log\left(\frac{m_{\tilde{b}_j}}{p_0^2}\right)-\frac{1}{2}\log\left(\frac{\sigma_{\tilde{b}_j,\tilde{b}_j}}{m^2_{\tilde{b}_j}}+1\right)\right] \tag{7.43}$$

et

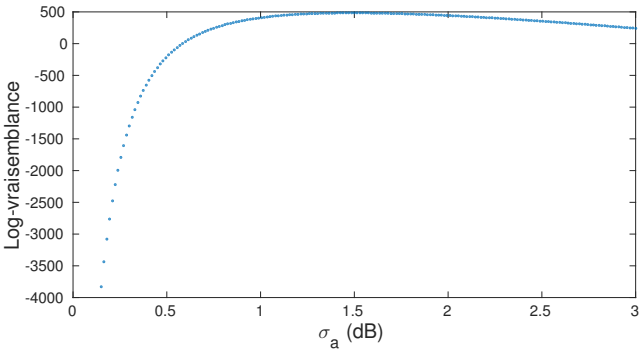
$$\sigma_{b_{j_1},b_{j_2}}\simeq k^2\times\left[\log\left(\frac{\sigma_{\tilde{b}_{j_1},\tilde{b}_{j_2}}}{m_{\tilde{b}_{j_1}}m_{\tilde{b}_{j_2}}}+1\right)\right]. \tag{7.44}$$

7.3.4 Calcul du maximum de la log-vraisemblance

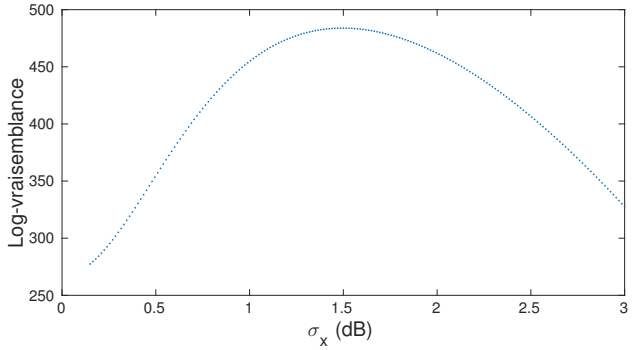
On cherche ensuite à estimer les trois paramètres d'incertitude par la maximisation de la log-vraisemblance. Les premières expériences semblent montrer que la log-vraisemblance est convexe. La Figure 7.5 montre pour un exemple donné des coupes de la fonction de log-vraisemblance (où on fait varier sur chaque sous-figure un seul des paramètres, les autres étant fixés à la valeur terrain). Chaque coupe est convexe. On peut voir cependant, d'après la forme de ces courbes, que le maximum est moins marqué pour certains paramètres comme $\sigma_r^\bullet$, et donc potentiellement plus difficile à déterminer.



(a) Log-vraisemblance en fonction de l'écart-type sur la séparation.



(b) Log-vraisemblance en fonction de l'écart-type sur la propagation.



(c) Log-vraisemblance en fonction de l'écart-type sur l'émission.

FIGURE 7.5 – Coupe de la fonction de vraisemblance pour le problème où les valeurs terrain sont $\sigma_x = 1,5$, $\sigma_a = 1,5$ et $\sigma_{r\bullet} = 1,5$. Sur chaque figure, seul un paramètre varie, les autres étant fixés à la valeur terrain.

Pour estimer les paramètres, on a calculé le maximum de vraisemblance à l'aide de la fonction *fminunc* de Matlab à laquelle on fournit l'expression du gradient de la log-vraisemblance.

7.3.5 Évaluation expérimentale

Dans cette partie, on propose une évaluation faite avec Matlab des méthodes proposées sur des données d'une étude en utilisant des incertitudes simulées par des tirages aléatoires.

Données

Les données utilisées sont issues d'une étude acoustique faite par Venathec SAS qui présente 6 points de mesure et 5 éoliennes. Afin de pouvoir évaluer le calcul d'estimation des incertitudes, les erreurs des variables sont simulées à l'aide de tirages aléatoires gaussiens indépendants. On crée de cette manière 125 séries/problèmes différents en considérant le produit cartésien de 5 valeurs d'écart-type (valeurs variant entre $0,5$ et $4,5$ dB par pas de 1 dB) pour les trois paramètres d'incertitude (σ_x, σ_a et $\sigma_{r\bullet}$). L'erreur métrologique est considérée négligeable pour le moment. Les calculs sont ensuite effectués en supposant que les niveaux résiduels $\mathbf{r}$ sont connus à chaque instant.

Résultats

Concernant l'évaluation, on note $\hat{\sigma}_x$, $\hat{\sigma}_a$, et $\hat{\sigma}_{r\bullet}$ les valeurs obtenues. Pour mesurer la qualité d'estimation, on définit la racine de l'erreur quadratique (RSE) en dB par :

$$RSE = \sqrt{(\hat{\sigma}_x - \sigma_x)^2 + (\hat{\sigma}_a - \sigma_a)^2 + (\hat{\sigma}_{r\bullet} - \sigma_{r\bullet})^2}. \tag{7.45}$$

De la même façon, on définit la racine de l'erreur quadratique normalisée (RNSE) comme :

$$RNSE = \sqrt{\frac{(\hat{\sigma}_x - \sigma_x)^2 + (\hat{\sigma}_a - \sigma_a)^2 + (\hat{\sigma}_{r\bullet} - \sigma_{r\bullet})^2}{\sigma_x^2 + \sigma_a^2 + \sigma_{r\bullet}^2}}. \tag{7.46}$$

On présente alors dans le tableau suivant, la moyenne, le maximum et le minimum des métriques RSE et RNSE des expériences menées pour les deux méthodes d'approximation de la vraisemblance.

	méthode des moments	linéarisation
RSE moyen	**0,53**	7,03
RSE min	**0,034**	0,046
RSE max	**1,53**	32,90
RNSE moyen	**0,10**	1,19
RNSE min	**0,0067**	0,013
RNSE max	**0,27**	5,58

FIGURE 7.6 – Résultats d'estimation des paramètres d'incertitude par maximisation de la log-vraisemblance.

Ces résultats montrent qu'il est possible d'estimer les variance des incertitudes lorsque les niveaux résiduels sont connus. Cependant, la méthode utilisant l'approximation des moments est incontestablement plus performante que la version linéarisée.

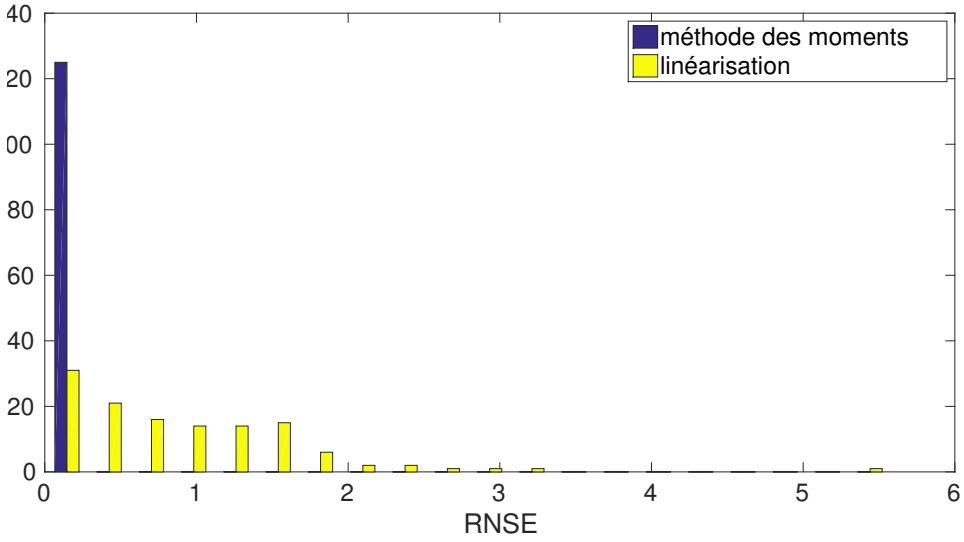


FIGURE 7.7 – Histogramme de la racine des erreurs au carré normalisée pour les deux approximations.

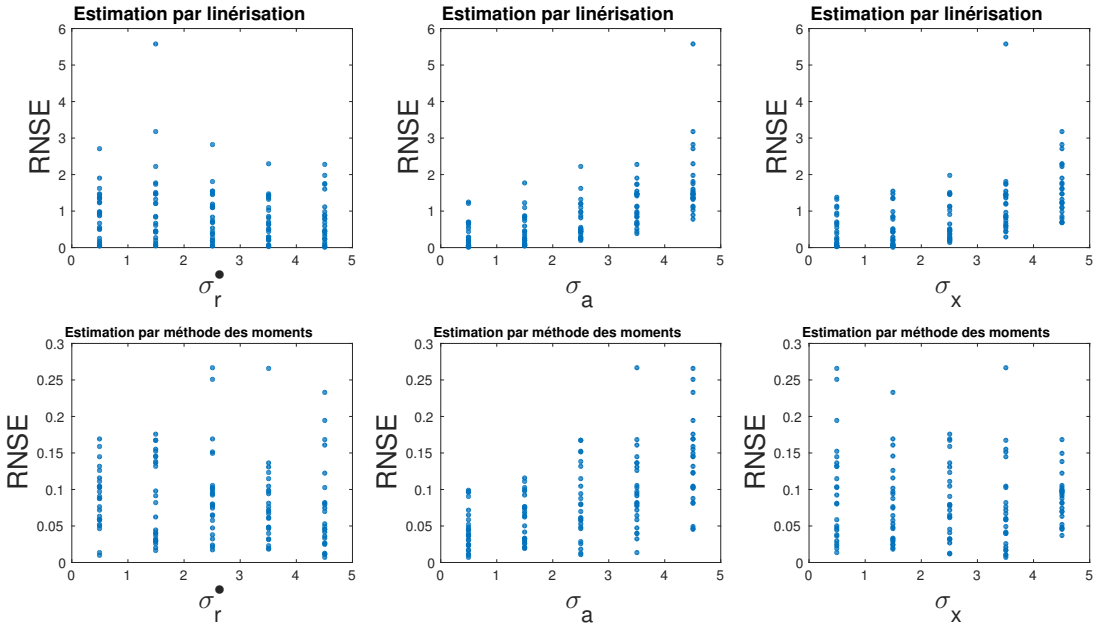


FIGURE 7.8 – RNSE en fonction de la valeur terrain de l'incertitude à estimer.

L'histogramme des erreurs RNSE de la Figure 7.7 montre graphiquement l'étalement des erreurs pour les deux méthodes et la différence de robustesse des deux méthodes. Pour finir, on voit sur la Figure 7.8 que l'erreur relative d'estimation et donc la qualité de l'estimateur ne montre pas de dépendance forte à la variance de l'incertitude.

7.4 Étape E : estimation des états cachées sous la connaissance des paramètres des incertitudes

Dans la suite, on présente l'utilisation d'un filtre[22] de Kalman adapté au problème qui va permettre d'estimer le bruit résiduel. Il s'agit d'un problème complètement différent par rapport à celui développé dans la partie précédente puisqu'on suppose cette fois les paramètres du modèle d'incertitude connus (moyennes et variances de la propagation, de l'émission et du résiduel). Le filtre de Kalman permet alors d'obtenir une estimation de l'état caché $\mathbf{y}_t = [\mathbf{x}_t\ \mathbf{a}_t\ \mathbf{r}_t]^T$ et cela sous la forme d'une distribution de probabilité corrigée par les observations.

Nous proposons dans la suite d'utiliser un modèle de filtrage récursif bayésien basé sur le modèle de marche aléatoire pour le niveau résiduel proposé dans la partie 7.2 ainsi que plusieurs filtres différents.

7.4.1 Modèle général

On propose d'utiliser un filtrage de Kalman non-linéaire afin de calculer à chaque instant la distribution *a posteriori* de $\mathbf{y}_t$ en utilisant le modèle d'évolution (7.10) et le modèle de mesure déjà proposé précédemment. Notons que l'on suppose un bruit d'évolution uniquement pour le résiduel (et pas pour $\mathbf{x}_t$ ou $\mathbf{A}_t$) pour simplifier et comme évoqué dans la partie 7.2. Néanmoins, rien n'empêche d'écrire le modèle plus général suivant :

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{a}_t \\ \mathbf{r}_t \end{bmatrix} = \begin{bmatrix} \mathbf{F}_\mathbf{x}\mathbf{x}_{t-1} + \mathbf{G}_\mathbf{x}\mathbf{m}_{\mathbf{x}_t} + \mathbf{w}_{\mathbf{x}_t} \\ \mathbf{F}_\mathbf{a}\mathbf{a}_{t-1} + \mathbf{G}_\mathbf{a}\mathbf{m}_{\mathbf{a}_t} + \mathbf{w}_{\mathbf{a}_t} \\ \mathbf{F}_\mathbf{r}\mathbf{r}_{t-1} + \mathbf{G}_\mathbf{r}\mathbf{m}_{\mathbf{r}_t} + \mathbf{w}_{\mathbf{r}_t} \end{bmatrix}, \tag{7.47}$$

$$b^*_{t,j} = 10\log_{10}\left(\sum_{i=1}^{I} 10^{\frac{a_{t,i,j}}{10}} 10^{\frac{x_{t,i}}{10}} + 10^{\frac{r_{t,j}}{10}}\right) + w_{b^*_{t,j}}, \tag{7.48}$$

$$r^{*\bullet}_{t,j} = r_{t,j} + w_{r^\bullet_{t,j}} + w_{b^*_{t,j}}. \tag{7.49}$$

L'équation d'évolution est écrite ici comme un processus ARMAX [Bertsekas, 2014] où $\mathbf{y}_t = [\mathbf{x}_t^T\ \mathbf{a}_t^T\ \mathbf{r}_t^T]^T$, $\mathbf{z}_t = [\mathbf{b}_t^{*T}\ \mathbf{r}_t^{*\bullet T}]^T$, $\mathbf{w}_{\mathbf{y}_t} = [\mathbf{w}_{\mathbf{x}_t}^T\ \mathbf{w}_{\mathbf{a}_t}^T\ \mathbf{w}_{\mathbf{r}_t}^T]^T$, $\mathbf{w}_{\mathbf{z}_t} = [\mathbf{w}_{\mathbf{b}_t^*}^T\ (\mathbf{w}_{\mathbf{b}_t^*}^T + \mathbf{w}_{\mathbf{r}_t^\bullet}^T)]^T$, $\mathbf{F}_\mathbf{x} = \mathbf{0}$, $\mathbf{F}_\mathbf{a} = \mathbf{0}$, $\mathbf{F}_\mathbf{r} = \boldsymbol{I}$, $\mathbf{G}_\mathbf{x} = \boldsymbol{I}$, $\mathbf{G}_\mathbf{a} = \boldsymbol{I}$, $\mathbf{G}_\mathbf{r} = \mathbf{0}$.

On note également le modèle de manière synthétique comme

$$\begin{cases} \mathbf{y}_{t+1} = \mathbf{M}\mathbf{y}_t + \mathbf{G}\mathbf{m}_t + \mathbf{w}_\mathbf{y} \\ \mathbf{z}_t = h_t(\mathbf{y}_t) + \mathbf{w}_\mathbf{z}. \end{cases} \tag{7.50}$$

Notons également, qu'en plus des paramètres σ_x σ_a et σ_{r*}, on considère dans cette partie les paramètres σ_{r_1}, σ_{r_2},...,σ_{r_J} qui constituent les écarts-types des niveaux résiduels pour le modèle d'évolution en chaque point j. Cet ajout est dû au niveau résiduel qui n'est plus supposé connu (contrairement à la partie 7.3) et qui suit un modèle d'évolution de marche aléatoire. Les points sont supposés suffisamment éloignés autour du parc pour supposer que la covariance du niveau

22. NB : la notion de filtrage est quelque peu différente de celle utilisée usuellement dans le domaine du traitement du signal. Ici, le terme filtrage a un sens statistique et désigne les techniques qui permettent d'obtenir à partir d'un état de connaissance sur des variables aléatoires (leur probabilité *a priori*) et des mesures (des réalisations), un état de connaissance affiné (probabilité *a posteriori*). Le terme de filtrage n'est cependant pas inadapté car des liens forts existent entre le filtrage statistique et le filtrage issu du traitement du signal.

résiduel entre les points de mesures est nulle deux à deux. En résumé, dans le modèle de filtrage on considère les matrices de covariances

$$\mathbf{\Sigma_y} = \begin{bmatrix} \sigma_x^2 \mathbf{I}_I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_a^2 \mathbf{I}_{IJ} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathrm{diag}(\sigma_{r_j}^2)_{1 \leq j \leq J} \end{bmatrix} \tag{7.51}$$

et

$$\mathbf{\Sigma_z} = \begin{bmatrix} \sigma_b^{*2} \mathbf{I}_J & \sigma_b^{*2} \mathbf{I}_J \\ \sigma_b^{*2} \mathbf{I}_J & \sigma_b^{*2} \mathbf{I}_J + \sigma_r^{\bullet 2} \mathbf{I}_J \end{bmatrix}. \tag{7.52}$$

7.4.2 Méthodes de filtrage

Nous avons évalué plusieurs filtres de Kalman non-linéaires :

— Filtre de Kalman étendu [Haykin et al., 2001]. Ce filtre est l'extension la plus simple et la plus utilisée du filtre de Kalman linéaire. Cette extension repose sur un développement de Taylor à l'ordre 1 afin de propager la moyenne et la covariance à travers les non-linéarités.

— Filtre de Kalman *unscented* [Julier & Uhlmann, 1997; Wan & Van Der Merwe, 2000]. Ce filtre est basé sur une autre forme de linéarisation dite « statistique ». Elle consiste à approximer les distributions gaussiennes par un petit nombre de points, appelés *sigma-points*. Ces points sont propagés ensuite à travers les non-linéarités et permettent de reconstruire la moyenne et la covariance.

— Filtre de Kalman *central difference* [Van Der Merwe, 2004]. Ce filtre est très similaire au filtre *unscented* mais la sélection des *sigma-points* diffère.

Notons que l'approximation du filtre de Kalman étendu est similaire à l'approximation utilisée dans la partie 7.3. En revanche, il n'existe pas à notre connaissance de filtre de Kalman basé sur la méthode des moments proposée dans cette même partie. De ce fait nous mettons cette méthode de côté dans la suite.

7.4.3 Évaluation expérimentale

Données

Pour l'évaluation expérimentale, nous utilisons environ 19 heures de données issues d'une étude acoustique comprenant 5 points de mesure et effectuée sur un parc éolien constitué de 6 éoliennes dont le nom n'est pas donné pour des raisons de confidentialité.

Dans la mesure où la vérité terrain n'est pas connue, les incertitudes sont générées par des tirages aléatoires gaussiens afin de constituer les données de test. Pour les niveaux résiduels, nous avons considéré comme vérité terrain les valeurs mesurées par les sonomètres lorsque les éoliennes sont à l'arrêt. Concernant l'émission et la propagation, les valeurs constructeurs d'émission et les valeurs d'atténuation acoustique obtenues par simulation numérique sont utilisées pour les moyennes $\mathbf{m}_{\mathbf{x}_t}$ et $\mathbf{m}_{\mathbf{a}_t}$. La génération des données est réalisée en utilisant l'équation (7.47) et en tirant $\mathbf{w}_{\mathbf{b}_t^*}$ et $\mathbf{w}_{\mathbf{r}_t^\bullet}$ selon des lois aléatoires gaussiennes décorrélées temporellement et spatialement.

Deux ensembles de tests sont générés. Le premier ensemble est constitué de 125 cas qui ont en commun d'avoir tous 6 éoliennes et 5 points de mesure mais qui diffèrent entre eux par l'utilisation d'un triplet $(\sigma_x, \sigma_a, \sigma_{r\bullet})$ différent pour la génération des incertitudes. Les écarts-types sont pris dans $\{0,5; 1,5; 2,5; 3,5; 4,5\}$ et les cas de test sont générés en considérant tous les triplets du produit cartésien de ces valeurs. Ce premier ensemble de données est utilisé afin de constater l'influence des paramètres d'incertitude sur le filtrage. Le second ensemble, quant à lui, est constitué de 30 cas de test qui ont en commun d'avoir des incertitudes générées à partir

du même triplet $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (3,5; 3,5; 3,5)$. Ils diffèrent cependant par le nombre d'éoliennes et de points de mesure conservés dans chaque cas. Dans chaque cas, un sous-ensemble de points et d'éoliennes est sélectionné de sorte à obtenir un nombre d'éoliennes variant de 1 à 6 et un nombre de points de 1 à 5. Ce second ensemble de données est utilisé pour déterminer l'influence du nombre d'éoliennes et de points de mesure sur le filtrage.

Pour les deux ensembles, l'écart-type de l'incertitude de mesure sonométrique $\mathbf{w}_{\mathbf{b}^*}$ est fixé à 1,5 dB, qui est la valeur considérée par la norme PR NF S 31-114. D'autre part, les paramètres $\sigma_{r_1}, \sigma_{r_2}, \ldots, \sigma_{r_J}$ sont appris directement à partir des valeurs terrain des niveaux résiduels.

Résultats

Pour évaluer les méthodes, nous considérons d'une part l'erreur RMSE en dB entre le niveau résiduel estimé $\hat{\mathbf{r}}_t$ et la valeur terrain $\mathbf{r}_t$, et d'autre part la log-vraisemblance de la vérité terrain par rapport à la distribution *a posteriori* supposée gaussienne de moyenne $\hat{\mathbf{r}}_t$ et de covariance $\hat{\mathbf{\Sigma}}_{\mathbf{r}_t}$. Afin de mesurer l'effet du filtrage sur l'émergence, nous évaluons également l'erreur RMSE et la log-vraisemblance entre l'émergence terrain et la distribution estimée de l'émergence. Cette dernière est calculée en propageant la moyenne $\hat{\mathbf{y}}_t$ et la covariance $\hat{\mathbf{\Sigma}}_{\mathbf{y}_t}$ estimées à travers la fonction h en utilisant la même méthode d'approximation (linéarisation ou *sigma-points*) que celle du filtre testé.

Les résultats pour le niveau résiduel et l'émergence sont présentés dans les Tableaux 7.1 et 7.2. Notons que l'erreur RMSE est plus importante pour l'estimation du niveau résiduel que pour l'estimation de l'émergence. Ceci est du au fait que l'incertitude de mesure $\mathbf{w}_{\mathbf{b}_t^*}$ s'annule lorsqu'on soustrait $\mathbf{r}_t$ à $\mathbf{b}_t$ pour calculer l'émergence. Il apparaît également que tous les filtres utilisés permettent de diminuer l'incertitude de façon substantielle par rapport à l'estimateur par séparation de sources sans utilisation de filtrage. En effet, l'erreur RMSE est plus de deux fois plus faible que sans filtrage pour le niveau résiduel et trois à quatre fois plus faible pour l'émergence.

Des trois méthodes utilisées, c'est le filtre *central difference Kalman filter* qui minimise le mieux l'erreur RMSE. Cependant, le filtre *unscented Kalman filter* donne une RMSE comparable tout en étant le filtre qui calcule la distribution *a posteriori* ayant la meilleure valeur de log-vraisemblance pour l'émergence. Le filtre *unscented* est donc le filtre le plus consistant pour l'émergence.

	RMSE	log-vraisemblance
Estimation par séparation (sans filtrage)	3,24	$\mathbf{2{,}50 \times 10^5}$
Filtre étendu	1,63	$2,42 \times 10^5$
Filtre *unscented*	1,63	$2,41 \times 10^5$
Filtre *central difference*	**1,61**	$2,40 \times 10^5$

TABLEAU 7.1 – Erreur RMSE et log-vraisemblance de l'estimation du niveau résiduel sur le premier ensemble de test.

L'intérêt des méthodes à *sigma-points* est également renforcé par une analyse des évolutions temporelles des filtrages qui sont représentés sur les Figures 7.9, 7.10 et 7.11.

On constate sur ces figures que le filtre étendu a tendance à surestimer beaucoup plus le niveau résiduel et donc à sous-estimer l'émergence, ce qui pourrait avoir pour conséquence d'entraîner un dépassement des niveaux réglementaires s'il était utilisé dans un système de contrôle.

D'autre part, notons que la diminution de l'incertitude par rapport à l'estimation de la séparation de sources seule provient de la capacité du filtre de Kalman à fusionner les incertitudes.

	RMSE	log-likelihood
Estimation par séparation (sans filtrage)	2,87	$\mathbf{4{,}27{\times}10^4}$
Filtre étendu	1,01	$-0,69 \times 10^5$
Filtre *unscented*	0,79	$\mathbf{1{,}54{\times}10^5}$
Filtre *central difference*	**0,76**	$-0,1 \times 10^5$

TABLEAU 7.2 – Erreur RMSE et log-vraisemblance de l'estimation de l'émergence sur le premier ensemble de test.

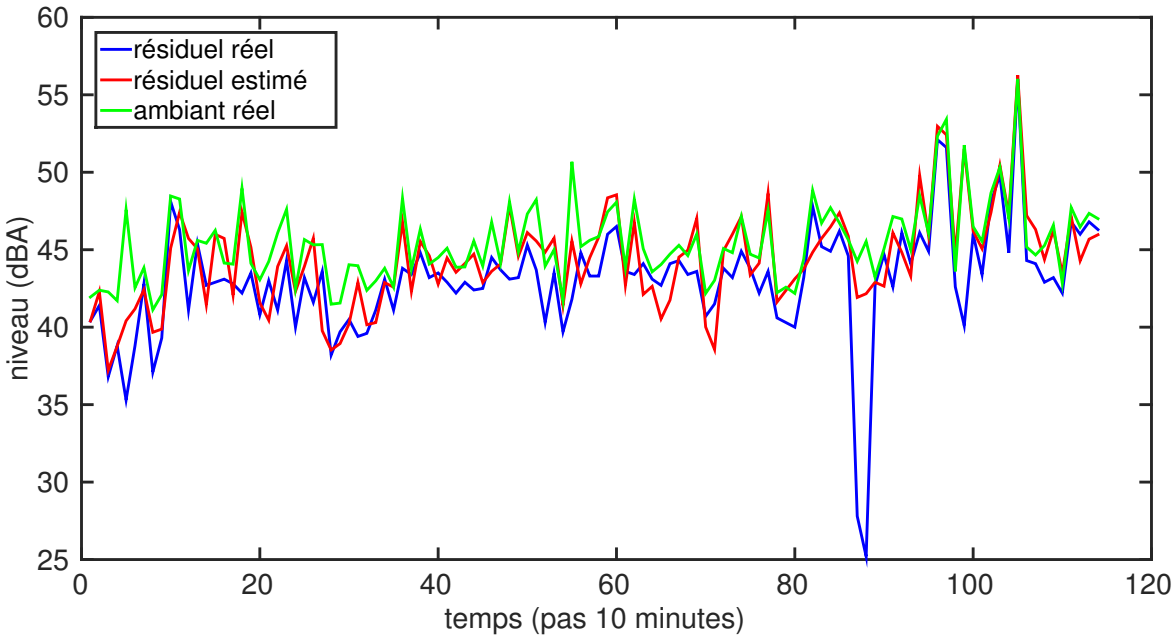


FIGURE 7.9 – Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman étendu dans le cas $J = 5$, $\sigma = 4,5$.

La Figure 7.12 représente la dépendance entre l'erreur RMSE d'estimation par rapport aux paramètres d'incertitude du modèle (σ_x et σ_a) et de la séparation de sources ($\sigma_r^\bullet$). Comme on peut le constater, l'erreur RMSE croît lorsque l'écart-type de l'incertitude du modèle augmente et est toujours inférieure à l'erreur RMSE de l'incertitude sur la séparation de sources, sauf aux points $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (3,5; 3,5; 0,5)$ et $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (4,5; 4,5; 0,5)$ où l'approximation de linéarisation statistique devient mauvaise. Par exemple, pour $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (0,5; 0,5; 0,5)$, l'erreur RMSE *a posteriori* vaut 0,2 dB, pour $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (3,5; 3,5; 3,5)$, elle vaut 0,9 dB, et pour $(\sigma_x, \sigma_a, \sigma_{r\bullet}) = (0,5; 0,5; 4,5)$, elle vaut 0,5 dB.

Le deuxième ensemble de test a permis de mettre en évidence l'amélioration relative apportée par l'utilisation de plusieurs points de mesure comparativement au nombre d'éoliennes sur un parc. Les résultats sont donnés graphiquement dans la Figure 7.13. On peut constater que plus le nombre d'éoliennes est grand, plus l'erreur RMSE augmente. Ceci est du à l'augmentation du nombre de variables d'incertitude indépendantes. *A contrario*, plus le nombre de points de mesure est grand, plus l'erreur RMSE diminue. En effet, bien que les incertitudes des variables x_i et $a_{i,j}$ s décorrélées deux à deux, les valeurs de niveau particulier sont corrélées spatialement entre les différents points de mesure. De ce fait, l'augmentation du nombre de points permet de

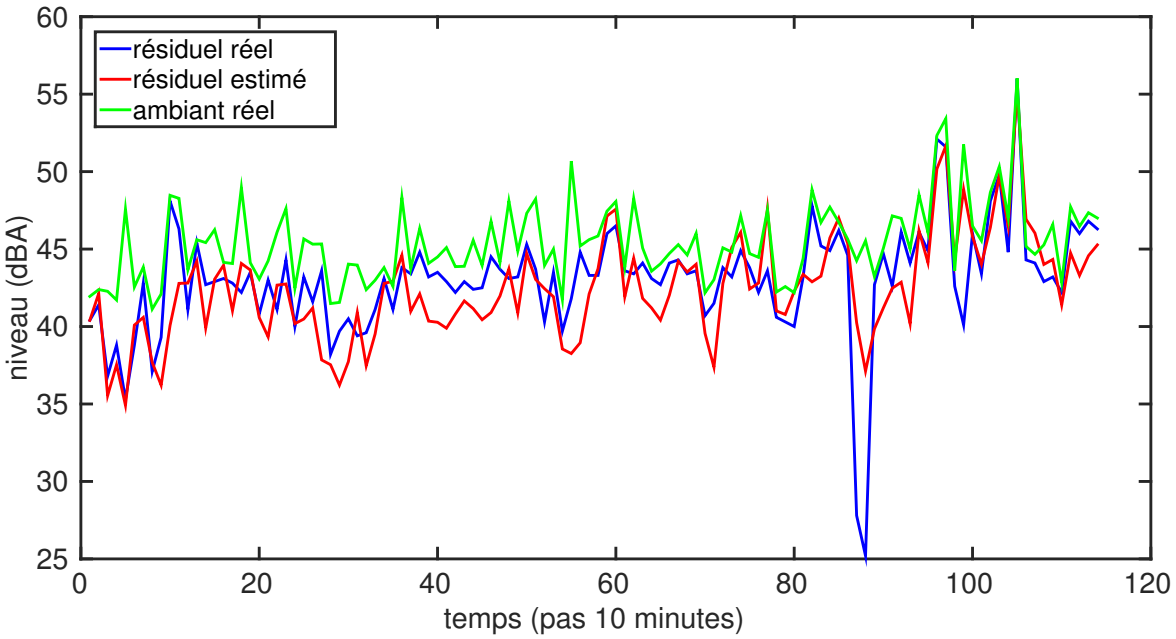


FIGURE 7.10 – Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman *unscented* dans le cas $J = 5$, $\sigma = 4,5$.

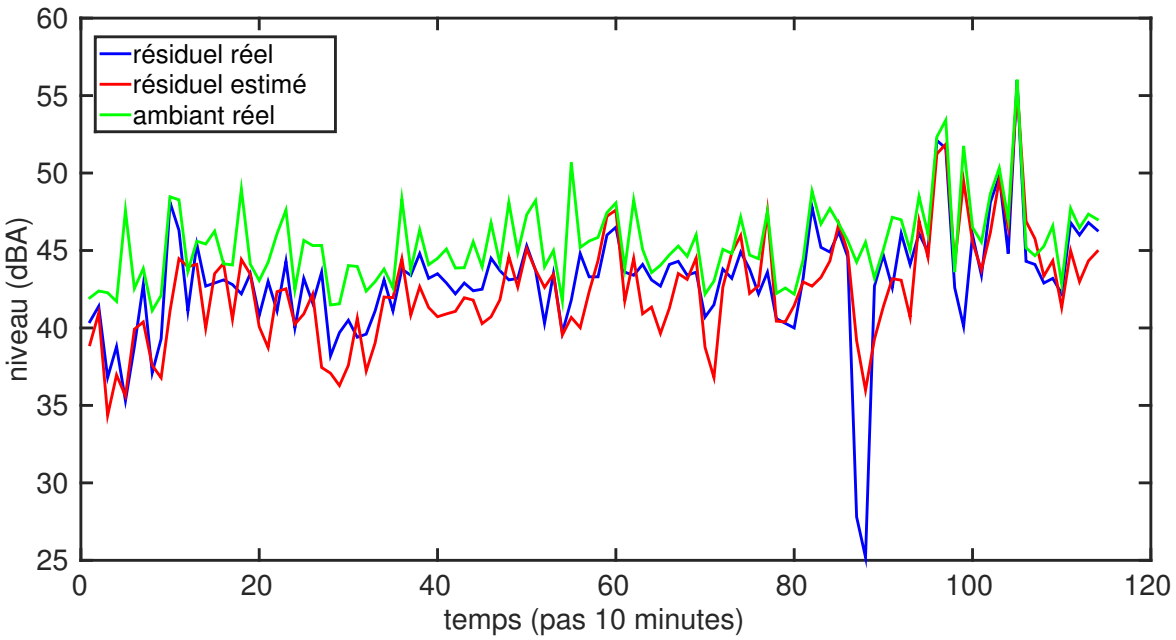


FIGURE 7.11 – Évolution temporelle de l'estimation du niveau résiduel par le filtre de Kalman *central difference* dans le cas $J = 5$, $\sigma = 4,5$.

diminuer l'incertitude *a posteriori*. Par exemple, pour 2 éoliennes et 5 points de mesure, l'erreur vaut 0,5 dB, pour 5 éoliennes et 1 points de mesure, l'erreur vaut 1,7 dB et pour 3 éoliennes et 3 points de mesure l'erreur vaut 0,7 dB.

Pour conclure, les évaluations expérimentales ont montré l'efficacité des filtres de Kalman non-linéaires et en particulier des filtres basés sur des *sigma-points*. Si d'autres filtres plus complexes

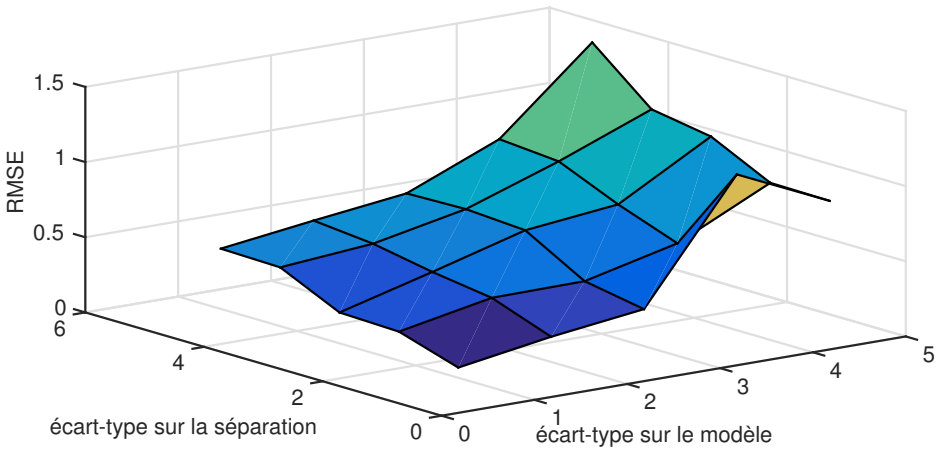


FIGURE 7.12 – Erreur RMSE de l'estimation de l'émergence par le filtre *central difference Kalman filter* en fonction de l'écart-type des incertitudes du modèle (σ_x et σ_a) et de l'écart-type de l'incertitude de la séparation de sources ($\sigma_{r\bullet}$).

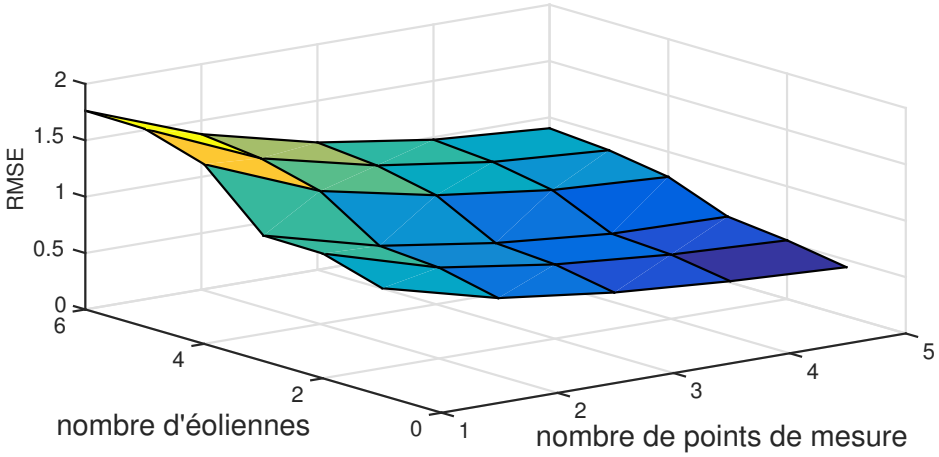


FIGURE 7.13 – Erreur RMSE de l'estimation de l'émergence par le filtre *central difference Kalman filter* en fonction du nombre d'éoliennes sur le parc et du nombre de points de mesure.

pourraient être utilisés, par exemple des filtres de Kalman itérés ou des filtres à particules, les filtres proposés constituent d'ores et déjà une solution intéressante pour le suivi de l'état caché dans le problème de contrôle acoustique.

7.5 Problème général : apprentissage conjoint par un algorithme Espérance-Maximisation (EM)

7.5.1 Algorithme EM proposé

Après l'étude de sous-problèmes dans les parties précédentes, on propose dans cette partie une solution au problème général d'estimation simultanée des paramètres du modèle et de l'état

caché. On rappelle que les paramètres qui nous intéressent sont $\boldsymbol{\theta} = (\sigma_x, \sigma_a, \sigma_{r^\bullet}, \sigma_{r_1}, ..., \sigma_{r_J})$ qui sont les écarts-types des incertitudes sur la mesure et sur le modèle d'état. On utilise une extension d'un algorithme EM classique inspiré par [Mader et al., 2014].

L'étape E de l'algorithme consiste simplement à utiliser un filtre de Kalman non-linéaire sur l'ensemble de la série temporelle, comme présenté dans la partie précédente. La distribution de filtrage obtenue $Q(\mathbf{y}_{1:T}) = p(\mathbf{y}_{1:T}|\boldsymbol{\theta}, \mathbf{z}_{1:T})$ est supposée gaussienne et on peut alors montrer que la fonction $\mathcal{F}(Q, \theta)$ définie dans la partie 2.3.3 se calcule de la manière suivante :

$$\begin{aligned}\mathcal{F}(Q,\theta) = &-\frac{T}{2}\log(\det(\boldsymbol{\Sigma}_\mathbf{y})) - \frac{T}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_\mathbf{y}^{-1}\left[\sum_{t=1}^{T}\mathbb{E}(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})^T\right]\right)\\ &-\frac{T}{2}\log(\det(\boldsymbol{\Sigma}_\mathbf{z})) - \frac{T}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_\mathbf{z}^{-1}\left[\sum_{t=1}^{T}\mathbb{E}(\mathbf{z}_t - h(\mathbf{y}_t))(\mathbf{z}_t - h(\mathbf{y}_t))^T\right]\right)\\ &-\int_{\mathbf{y}_{1:T}} Q(\mathbf{y}_{1:T})\log Q(\mathbf{y}_{1:T})d\mathbf{y}_{1:T}.\end{aligned} \tag{7.53}$$

En dérivant $\mathcal{F}$ par rapport à $\boldsymbol{\Sigma}_\mathbf{y}$ et à $\boldsymbol{\Sigma}_\mathbf{y}$, on obtient les mises à jour suivantes pour les variances lors de l'étape M :

$$\widehat{\boldsymbol{\Sigma}}_\mathbf{y} = \frac{1}{T}\left[\sum_{t=1}^{T}\mathbb{E}(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})^T\right], \tag{7.54}$$

$$\widehat{\boldsymbol{\Sigma}}_\mathbf{z} = \frac{1}{T}\left[\sum_{t=1}^{T}\mathbb{E}(\mathbf{z}_t - h(\mathbf{y}_t))(\mathbf{z}_t - h(\mathbf{y}_t))^T\right]. \tag{7.55}$$

On obtient ensuite les mises à jour de $\sigma_x, \sigma_a, \sigma_{r^\bullet}$ et $(\sigma_{r_1}, ..., \sigma_{r_J})$ par les formules

$$\hat{\sigma}_x = \sqrt{\frac{1}{I}\sum_{k=1}^{I}\widehat{\boldsymbol{\Sigma}}_\mathbf{y}[k]}, \tag{7.56}$$

$$\hat{\sigma}_a = \sqrt{\frac{1}{I\times J}\sum_{k=I+1}^{I+IJ}\widehat{\boldsymbol{\Sigma}}_\mathbf{y}[k]}, \tag{7.57}$$

$$\hat{\sigma}_r^\bullet = \sqrt{\frac{1}{I\times J}\sum_{k=J+1}^{2J}\widehat{\boldsymbol{\Sigma}}_\mathbf{z}[k]} \tag{7.58}$$

et

$$\hat{\sigma}_{r_j} = \sqrt{\widehat{\boldsymbol{\Sigma}}_\mathbf{z}[J+j]}, \forall j \in \{1, ..., J\} \tag{7.59}$$

qui peuvent être obtenues par dérivation. La notation $\widehat{\boldsymbol{\Sigma}}_\mathbf{y}[k]$ désigne ici le k-ème élément diagonal de $\widehat{\boldsymbol{\Sigma}}_\mathbf{y}$.

Notons que comme $\mathbf{y}_t$ et $\mathbf{y}_{t-1}$ suivent des distributions gaussiennes, $(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})$ suit également une distribution gaussienne dont le moment d'ordre 2 non-centré $\mathbb{E}(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})(\mathbf{y}_t - \mathbf{F}\mathbf{y}_{t-1})^T$ peut se calculer facilement et de manière exacte à partir des moyennes et des variances/covariances de $\mathbf{y}_t$ et de $\mathbf{y}_{t-1}$.

Le terme $\sum_{t=1}^{T}\mathbb{E}(\mathbf{z}_t - h(\mathbf{y}_t))(\mathbf{y}_t - h(\mathbf{y}_t))^T$ est calculé en linéarisant la fonction h ou en échantillonnant (Monte-Carlo). On peut ici encore, linéariser cette fonction de différentes manières, en

utilisant un développement de Taylor (comme dans le filtre *extended Kalman*) ou une linéarisation *statistique* (comme dans les filtres *unscented Kalman* et *central difference Kalman*).

Pour finir, notons que la méthode des moments présentée dans la partie 7.3, bien qu'efficace dans le problème simplifié, n'est pas utilisée pour calculer la mise-à-jour de l'étape M. En effet, elle ne permet pas de calculer formellement la fonction $\mathcal{F}(Q, \theta)$. Cette approximation pourrait éventuellement être utilisée pour résoudre le problème général d'estimation avec un algorithme de gradient comme présenté dans les perspectives de la thèse dans le chapitre 9.

7.5.2 Évaluation expérimentale

Données

Pour évaluer l'algorithme EM, nous générons deux ensembles de test.

Le premier est généré de façon similaire à ce que nous avons fait pour générer le premier ensemble de test pour les méthodes de filtrage dans la partie 7.4.3 : nous utilisons les mêmes 19 heures de données de l'étude acoustique comprenant 5 points de mesure, acquises sur un parc éolien constitué de 6 éoliennes et nous générons les incertitudes par des tirages aléatoires gaussiens en faisant varier les valeurs des écarts-types des incertitudes de la même manière.

Le second ensemble de test est généré à partir des mêmes données d'étude mais avant de lancer la génération des données, nous soustrayons artificiellement 10 dB à l'ensemble des valeurs de niveau résiduel.

De cette manière nous obtenons deux ensembles de test qui présentent des valeurs terrain d'émergence différentes :

— Un premier ensemble où l'émergence moyenne sur le point le plus exposé est de 1,8 dBA.

— Un second ensemble où l'émergence moyenne sur le point le plus exposé est de 7.0 dBA.

À l'intérieur de chaque ensemble, les cas se différencient par les valeurs d'écart-type des incertitudes sur le modèle et la séparation de sources. Des paramètres différents sont donc appris pour chaque cas de test. Pour les deux ensembles, l'écart-type de l'incertitude de mesure du sonomètre $\mathbf{w}_{\mathbf{b}^*}$ est fixée à 1,5 dB, et les paramètres σ_{r_1}, σ_{r_2}, ..., σ_{r_J} à estimer sont fixés par les valeurs terrain des niveaux résiduels.

Filtre et approximation utilisées et paramétrage de l'algorithme EM

Pour l'étape E, nous utilisons le filtre *unscented*. Pour l'étape M, nous utilisons la méthode de linéarisation du filtre *unscented* afin de calculer la mise à jour (7.55). À l'initialisation, les valeurs des écart-types σ_x, σ_a et $\sigma_r^\bullet$ valent 3 dB. Pour chaque calcul, on effectue 1000 itérations de l'algorithme. Notons que nous mettons ainsi de côté la méthode la méthode des moments présentée dans la partie 7.3, qui ne peut pas être intégrée directement dans la méthode EM car elle ne permet pas une mise-à-jour formelle à l'étape M. L'utilisation de cette approximation constitue une perspective de la thèse présentée dans le chapitre 9.

Résultats

Le Tableau 7.3 résume l'erreur RMSE totale de l'estimation de chaque paramètre sur chaque ensemble. Le détail des résultats est présenté dans l'annexe B.

L'estimation des paramètres apparaît relativement efficace et génère une erreur RMSE de 0,5 dB au plus sauf pour l'estimation de σ_x et σ_a qui est mauvaise. Pourtant, ces valeurs sont

	σ_x	σ_a	$\sigma_r^\bullet$	σ_{r_1}	σ_{r_2}	σ_{r_3}	σ_{r_4}	σ_{r_5}
Ens. 1	1,74	1,97	0,09	0,10	0,54	0,00	0,01	0,05
Ens. 2	1,38	0,94	0,14	0,38	0,65	0,02	0,09	0,26

TABLEAU 7.3 – RMSE entre les écarts-types terrain et les écart-types estimés pour les deux ensembles de test.

calculées sans approximation particulière de manière formelle à partir de l'équation (7.54). L'estimation n'est pas performante et on peut donc penser qu'il s'agit donc d'un problème lié au filtrage trop peu précis et/ou de conditionnement du problème.

Afin de déterminer s'il s'agit d'un problème de conditionnement, nous proposons de reparamétrer les résultats. Si on considère non plus les paramètres σ_x sur l'émission et σ_a sur la propagation individuellement, mais le paramètre $\sigma_l = \sqrt{\sigma_x^2 + \sigma_a^2}$ qui est l'écart-type de l'incertitude sur les contributions particulières $l_{i,j}$, on obtient alors les résultats présentés dans le Tableau 7.4. Le détail des résultats est présenté dans l'annexe B. On observe alors que dans le cas où l'émergence est élevée (cas de l'ensemble 2), l'estimation des incertitudes donne des résultats plus satisfaisants avec une RMSE de l'estimateur de 0,5 dB sur l'ensemble de ces cas. En revanche, dans le cas où l'émergence est plus faible (cas de l'ensemble 1), l'estimation reste de mauvaise qualité. La reparamétrisation des résultats permet donc en partie de lever le problème d'estimation.

	σ_l
Ens. 1	2,44
Ens. 2	0,46

TABLEAU 7.4 – RMSE de l'estimation du paramètre σ_l pour les deux ensembles de test.

Bien entendu, cette reparamétrisation constitue une perte d'information dans la mesure où la connaissance individuelle des paramètres σ_x et σ_a permet plus de finesse dans le modèle. Cependant, la simple estimation de σ_l permet déjà de proposer une solution de contrôle plus adaptée et est utilisée dans l'adaptation stochastique de l'algorithme *branch and bound* au chapitre 8.

On observe cependant que la valeur de l'émergence conditionne cette fois la qualité de l'estimation. Une explication probable est l'effet de masquage acoustique représenté mathématiquement par la non-linéarité. En effet, on peut comprendre ce phénomène en considérant l'approximation suivante [Cook, 2011] :

$$b_{t,j} = 10\log_{10}(10^{\frac{l_{t,j}}{10}} + 10^{\frac{r_{t,j}}{10}}) \tag{7.60}$$

$$\simeq \max(l_{t,j}, r_{t,j}). \tag{7.61}$$

Bien que grossière, cette approximation met en avant le phénomène qui apparaît lors de l'estimation. Quand l'émergence est grande, on a $b_{t,j} \simeq l_{t,j}$ et l'estimation de σ_l fonctionne alors bien dans la mesure où l'incertitude qui apparaît sur $b_{t,j}$ est à peu près la même que celle sur $l_{t,j}$. *A contrario*, lorsque l'émergence est faible on a $b_{t,j} \simeq r_{t,j}$ et dans ce cas le modèle proposé ne permet plus d'estimer $l_{t,j}$. En effet, dans ce cas, $b^*_{t,j}$ et $r^{*\bullet}_{t,j}$ mesurent tous deux le niveau résiduel. Notons cependant que le système de séparation de sources estime en réalité le niveau résiduel et le niveau particulier et non pas seulement le niveau résiduel. Le modèle de filtrage proposé ne considère que l'estimation du niveau résiduel et les résultats seraient sûrement différents si on modélisait la dépendance de l'incertitude de la séparation de sources au rapport signal-à-bruit.

Néanmoins et pour conclure, les résultats montrent que l'approche est numériquement valide et que l'on pourra certainement espérer apprendre les paramètres d'incertitude du bruit particulier quand l'émergence sera suffisamment élevée, et au contraire du bruit résiduel quand elle sera faible, rendant l'algorithme EM proposé utilisable.

8

Optimisation stochastique

8.1 Formulation du problème

On propose dans ce chapitre un algorithme pour la résolution du problème d'optimisation stochastique. Il s'agit d'un problème *chance-constrained programming* similaire à ceux présentés dans l'état de l'art au chapitre 2

$$\begin{cases} \max_{\mathbf{c}_t \in \mathcal{C}} \sum_{i=1}^{I} \mathbb{E}(p_{t,i} | c_{t,i}, o_t) \\ p(\mathbf{b}_t - \mathbf{r}_t < S_e \text{ ou } \mathbf{b}_t < S_s) > 1 - \varepsilon, \end{cases} \tag{8.1}$$

qui consiste à résoudre le problème d'optimisation sous une tolérance probabiliste ε à respecter. Comme décrit dans l'état de l'art, ce type de problème est difficile à résoudre dans un cadre général car on peut dans de nombreux cas observer une perte de convexité du problème (si on compare au problème déterministe) et une difficulté à vérifier l'admissibilité d'une solution (c'est-à-dire à vérifier qu'une solution potentielle ne viole pas la contrainte). De ce fait,on propose de considérer le problème non pas avec une loi jointe multivariée (*joint chance constraint*) mais avec les lois marginales (*single chance constraint*) :

$$\begin{cases} \max_{\mathbf{c}_t \in \mathcal{C}} \sum_{i=1}^{I} \mathbb{E}(p_{t,i} | c_{t,i}, \mathbf{v}_t) \\ \forall j \in \{1, ..., J\}, p(b_{t,j} - r_{t,j} < S_e \text{ ou } b_{t,j} < S_s) > 1 - \varepsilon. \end{cases} \tag{8.2}$$

Ce problème est plus simple à résoudre, mais également plus conservatif puisqu'on considère les lois marginales des contraintes. Dans le cadre du contrôle acoustique des éoliennes, on s'intéresse pour l'instant à cette formulation dans la mesure où les incertitudes sur les différents points de mesure sont décorrélées en pratique (distance élevée entre les points de mesure).

Pour résoudre ce problème, on propose de se ramener à un problème déterministe et d'adapter l'algorithme *branch and bound* pour la résolution. On présente dans la suite cette approche en deux temps :

1. **Caractérisation des solutions admissibles** : Dans le problème déterministe, on avait caractérisé l'admissibilité par une inégalité faisant intervenir une valeur maximale de bruit particulier $\mathbf{q}$ calculée à partir des seuils réglementaires S_s et S_e et du bruit résiduel à cet instant r_t. De manière analogue, on propose de caractériser l'admissibilité du problème stochastique par une valeur maximale du bruit particulier moyen en chaque point j comme une fonction déterministe de la moyenne $\widehat{m}_{r_{t,j}}$, des variances $\widehat{\sigma}^2_{l_{t,j}}$, $\widehat{\sigma}^2_{r_{t,j}}$, des seuils S_s, S_e et de la probabilité de non-respect tolérée ε. En résumé on cherche à déterminer la fonction $\mathfrak{q}$ telle que :
$$\widehat{m}_l^{\max} = \mathfrak{q}(\widehat{\sigma}_l, \widehat{\sigma}_r, \widehat{m}_r, \varepsilon, S_s, S_e). \tag{8.3}$$

2. **Adaptation de l'algorithme au problème stochastique** : Dans le problème déterministe, on construit l'arbre de recherche en calculant l'admissibilité des solutions à chaque nouveau branchement, et la majoration à partir d'un problème continu. De manière analogue, on propose de redéfinir une façon de construire l'arbre et la majoration pour le problème stochastique à partir de la caractérisation ci-dessus.

8.2 Caractérisation des solutions admissibles

8.2.1 Représentation graphique des solutions admissibles

Afin d'alléger les notations, on retire provisoirement l'indice j des variables et on ne s'intéresse qu'à un seul point de mesure. On s'intéresse ici à caractériser l'admissibilité d'une solution en un point par une méthode de Monte-Carlo. Pour simplifier le problème, on suppose que l_t et r_t sont décorrélées[23] et donc $\widehat{\sigma}_{l_t,r_t} = 0$.

On peut résumer graphiquement la condition d'admissibilité par la Figure 8.1.

Le bruit résiduel à un instant donné est représenté en abscisse, et le bruit particulier en ordonnée. On se place dans le cas d'un seuil d'émergence de 3 dBA, et avec un seuil de bruit ambiant maximum de 35 dBA. Sur cette figure, la contrainte q (le bruit particulier maximum) du problème déterministe est représentée par la ligne rouge. En effet, pour un niveau résiduel donné, on peut calculer q comme $q = 10\log_{10}(10^{\frac{\max(3+r,35)}{10}} - 10^{\frac{r}{10}})$. Les valeurs admissibles (respectant la contrainte acoustique) sont représentées par les valeurs situées dans l'aire en bleu dans la figure.

Cependant, dans le cadre du problème stochastique, r_t et l_t suivent tous deux des lois de probabilité gaussiennes qui sont supposées indépendantes. En effet, r_t suit une loi gaussienne dont la moyenne est donnée par la valeur à l'instant précédent $t-1$ et de variance calculée à l'aide de la loi d'évolution et/ou du filtrage. La loi de l_t est également approximativement une loi gaussienne dont les paramètres dépendent de la commande c_t et de la classe homogène o. Les courbes bleues foncées représentent la densité de probabilité de la loi jointe de $[l_t^T \ r_t^T]^T$. La probabilité que la contrainte acoustique soit respectée, est la probabilité que $[l_t^T \ r_t^T]^T$ se situe dans l'aire bleue qui est égale à l'intégrale prise sur l'aire bleue de la densité de probabilité.

8.2.2 Caractérisation par une méthode de Monte-Carlo

Une première méthode pour caractériser les solutions admissibles consiste à utiliser une méthode de Monte Carlo. Pour un n-uplet $(\widehat{\sigma}_l, \widehat{\sigma}_r, \widehat{m}_r, \varepsilon, S_s, S_e)$ donné on applique la méthode suivante :

1. On initialise $\widehat{m}_l$ par une valeur suffisamment grande afin que la contrainte ne soit pas admissible. Par exemple, on peut considérer à l'initialisation $\widehat{m}_l = \widehat{m}_l^{\text{theo}} + 4$ où $\widehat{m}_l^{\text{theo}}$ est la valeur maximale déterministe (image de $\widehat{m}_r$ par la courbe rouge dans la Figure 8.1)
2. On tire un grand nombre N d'échantillons selon la loi multivariée $\mathcal{N}\left(\begin{bmatrix}\widehat{m}_l \\ \widehat{m}_r\end{bmatrix}, \begin{bmatrix}\widehat{\sigma}_l^2 & 0 \\ 0 & \widehat{\sigma}_r^2\end{bmatrix}\right)$.
3. On vérifie s'il y a au moins $N(1-\varepsilon)$ tirages qui vérifient le critère déterministe.
4. Si ce n'est pas le cas, on diminue faiblement la valeur de $\widehat{m}_l$ et on recommence (0,1 dB par exemple).
   Si c'est le cas on obtient alors une approximation de la valeur de $\widehat{m}_l$ maximale admissible, que l'on note $\widehat{m}_l^{\max}$ et qui dépend de $\widehat{m}_r$, $\widehat{\sigma}_r$, $\widehat{\sigma}_l$, S_s, S_e et ε.

23. On peut néanmoins étendre cette analyse facilement au cas général où $\widehat{\sigma}_{l_t,r_t} \neq 0$.

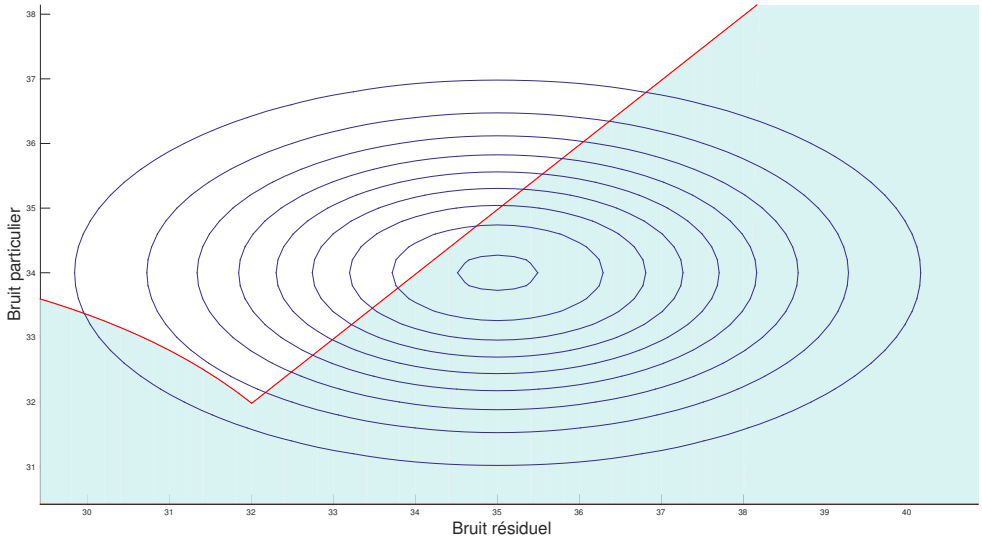


FIGURE 8.1 – Représentation de la loi jointe de $[l_t^T \ r_t^T]^T$ (courbes bleues foncées), de q (courbe rouge) et des solutions admissibles (aire bleue sous la courbe).

8.2.3 Caractérisation par une approximation basée sur la fonction de répartition d'une loi gaussienne

On présente maintenant une autre caractérisation plus simple, utilisant la fonction de répartition d'une loi gaussienne et en considérant deux cas : le cas où $\widehat{m}_{r_t} \gg 32$ et le cas où $\widehat{m}_{r_t} \ll 32$.

Approximation dans le cas où $\widehat{m}_{r_t} \gg 32$

Dans le cas où $\widehat{m}_{r_t} \gg 32$, seule la contrainte $b - r < S_e$ est considérée car on suppose que $b < S_s$ est vérifiée. On a alors :

$$\begin{aligned} b - r < S_e &\Leftrightarrow 10\log_{10}\left(10^{\frac{l}{10}} + 10^{\frac{r}{10}}\right) - 10^{\frac{r}{10}} < S_e \\ &\Leftrightarrow 10\log_{10}\left(\frac{10^{\frac{l}{10}}}{10^{\frac{r}{10}}} + 1\right) < S_e \\ &\Leftrightarrow \frac{10^{\frac{l}{10}}}{10^{\frac{r}{10}}} < 10^{\frac{S_e}{10}} - 1 \\ &\Leftrightarrow l - r < 10\log_{10}\left(10^{\frac{S_e}{10}} - 1\right) \\ &\Leftrightarrow l - r < \eta_1 \end{aligned} \tag{8.4}$$

où $\eta_1 = 10\log_{10}(10^{S_e/10} - 1)$ est une constante ne dépendant pas de l ni de r. On a ainsi

$$\begin{aligned} p(b - r < S_e) &= p(l - r < \eta_1) \\ &= p(l - r - \eta_1 < 0). \end{aligned} \tag{8.5}$$

On pose $\delta = l - r - \eta_1$. La variable δ est alors une variable aléatoire gaussienne qui vérifie

$$\delta \sim \mathcal{N}(\widehat{m}_l - \widehat{m}_r - \eta_1, \widehat{\sigma}_l^2 + \widehat{\sigma}_r^2). \tag{8.6}$$

En utilisant la formule de la fonction de répartition d'une gaussienne on en déduit

$$\begin{aligned} p(b - r < S_e) &= p(q < 0) \\ &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{-(\widehat{m}_l - \widehat{m}_r - \eta_1)}{\sqrt{2(\widehat{\sigma}_l^2 + \widehat{\sigma}_r^2)}}\right)\right]. \end{aligned} \tag{8.7}$$

On peut alors ré-écrire la condition d'admissibilité comme

$$\begin{aligned} p(b - r < S_e) \geq 1 - \varepsilon &\Leftrightarrow \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{-(\widehat{m}_l - \widehat{m}_r - \eta_1)}{\sqrt{2(\widehat{\sigma}_l^2 + \widehat{\sigma}_r^2)}}\right)\right] \geq 1 - \varepsilon \\ &\Leftrightarrow \operatorname{erf}\left(\frac{-(\widehat{m}_l - \widehat{m}_r - \eta_1)}{\sqrt{2(\widehat{\sigma}_l^2 + \widehat{\sigma}_r^2)}}\right) \geq 2\varepsilon - 1 \end{aligned} \tag{8.8}$$

où $\operatorname{erf}(\cdot)$ est la fonction d'erreur de Gauss.

Approximation dans le cas où $\widehat{m}_{r_t} \ll 32$

Au contraire, dans le cas où $\widehat{m}_{r_t} \ll 32$, on ne considère que la contrainte $b < S_s$. Pour ce faire, on propose de linéariser la contrainte autour de $\widehat{m}_{r_t}$ afin de se ramener à une approximation du même type. Tout d'abord, on a :

$$\begin{aligned} b < S_s &\Leftrightarrow 10\log_{10}\left(10^{\frac{l}{10}} + 10^{\frac{r}{10}}\right) < S_s \\ &\Leftrightarrow 10^{\frac{l}{10}} < 10^{\frac{S_s}{10}} - 10^{\frac{r}{10}} \\ &\Leftrightarrow l < f(r), \end{aligned} \tag{8.9}$$

où f est la fonction de r définie par

$$f(r) = 10\log_{10}\left(10^{\frac{S_s}{10}} - 10^{\frac{r}{10}}\right). \tag{8.10}$$

En calculant la dérivée de cette fonction

$$f'(r) = \frac{-10^{\frac{r}{10}}}{10^{\frac{S_s}{10}} - 10^{\frac{r}{10}}} \tag{8.11}$$

et en approximant le terme de droite dans (8.9) par un développement de Taylor à l'ordre 1 on obtient :

$$\begin{aligned} b < S_s &\Leftrightarrow l < 10\log_{10}\left(S_s - 10^{\frac{\widehat{m}_r}{10}}\right) + f'(\widehat{m}_r)(r - \widehat{m}_r) \\ &\Leftrightarrow l - f'(\widehat{m}_r)r - 10\log_{10}\left(S_s - 10^{\frac{\widehat{m}_r}{10}}\right) + f'(\widehat{m}_r)\widehat{m}_r < 0 \\ &\Leftrightarrow l - f'(\widehat{m}_r)r - \eta_2(\widehat{m}_r) < 0. \end{aligned} \tag{8.12}$$

On peut alors utiliser un développement similaire aux équations (8.5),(8.6),(8.7) et (8.8) obtenir finalement :

$$p(b - r < S_e) \geq 1 - \varepsilon \Leftrightarrow \operatorname{erf}\left(\frac{\widehat{m}_l - f'(\widehat{m}_r)\widehat{m}_r - \eta_2(\widehat{m}_r)}{\sqrt{2(\widehat{\sigma}_l^2 + f'(\widehat{m}_r)^2\widehat{\sigma}_r^2)}}\right) \geq 2\varepsilon - 1. \tag{8.13}$$

8.3 Évaluation et comparaison de la méthode de Monte-Carlo et de l'approximation proposée

On désire désormais expérimenter les deux méthodes en calculant la valeurs de $\widehat{m}_l^{\max}$ pour différentes valeurs de $\widehat{\sigma}_l$, $\widehat{\sigma}_r$, $\widehat{m}_r$ pour ε, S_s et S_e donnés. L'objectif est double. D'une part, on cherche à observer comment varie $\widehat{m}_l^{\max}$ afin de voir dans quelle mesure il serait possible d'apprendre une fonction $\mathfrak{q}$ définie par (8.3) et permettant de caractériser l'admissibilité d'une solution. D'autre part, on cherche à voir dans quelle mesure le calcul par Monte-Carlo pourrait être substitué par l'approximation qu'on a proposée dans la partie précédente.

8.3.1 Valeurs de tests

On a choisi en pratique de tester ces deux méthodes pour une valeur de $\varepsilon = 0,2$ (taux d'admissibilité de 80%) et avec les valeurs standard $S_s = 35$ et $S_e = 3$ pour les seuils à respecter. Concernant les valeurs de moyenne et de variance, on les fait varier de la façon suivante :

— On fait varier $\widehat{m}_r$ entre 10 dB et 50 dB par pas de 0,2 dB.
— On fait varier $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$ entre 0,5 et 6 dB par pas de 1 dB.

Pour chaque triplet $(\widehat{m}_r, \widehat{\sigma}_r, \widehat{\sigma}_l)$, on calcule $\widehat{m}_l^{\max}$ avec la méthode de Monte-Carlo mais également par l'approximation proposée. Notons que pour la méthode de Monte-Carlo nous effectuons 100 000 tirages pour chaque triplet.

Les valeurs obtenues par les deux méthodes sont ensuite comparées.

8.3.2 Résultats

À l'observation des résultats, on observe globalement trois cas :

— Lorsque $\widehat{m}_r$ est faible comme dans la Figure 8.2, on observe que $\widehat{\sigma}_r$ n'a pas d'influence sur la valeur de $\widehat{m}_l^{\max}$ qui ne dépend que de $\widehat{\sigma}_l$. Ce résultat est conforme à l'intuition que donne la Figure 8.1, où on peut voir que la courbe est quasiment horizontale lorsque $\widehat{m}_r$ est faible.
— Au contraire lorsque $\widehat{m}_r$ est grand comme dans la Figure 8.3, les valeurs $\widehat{\sigma}_l$ et $\widehat{\sigma}_r$ influent toutes deux de manière égale sur la valeur de $\widehat{m}_l^{\max}$.
— Pour finir, on observe que lorsque $\widehat{m}_r$ se situe autour de 32 dBA comme dans la Figure 8.4, une valeur élevée de $\widehat{\sigma}_r$ autorise une valeur de $\widehat{m}_l^{\max}$ plus élevée. Cela provient du fait que $\widehat{m}_r = 32$ est la valeur la plus conservative. Si on se situe dans un voisinage de cette valeur avec une probabilité élevée, on se trouvera alors dans un cas beaucoup moins conservatif.

La Figure 8.5 montre une comparaison entre les valeurs de $\widehat{m}_l^{\max}$ obtenues par Monte-Carlo, et par l'approximation proposée. La racine de l'erreur quadratique moyenne est de 0,7066 dB entre les valeurs des deux méthodes. Cependant, et comme en témoigne l'histogramme des erreurs dans la Figure 8.6, l'erreur entre les deux méthodes peut aller jusqu'à 5 dB de différence. Dans le cas où $m_r \simeq 32$, les valeurs de $\widehat{m}_l^{\max}$ obtenues par l'approximation proposée peuvent être bien plus petites que celles obtenus par la méthode de Monte-Carlo. Utiliser cette approximation dans ce cas pour l'optimisation stochastique engendrerait alors des pertes électriques trop fortes, une marge initialement élevée étant alors prise. Dans le cadre de la résolution du problème (8.2),on propose alors d'utiliser directement une méthode de Monte-Carlo afin que le critère d'admissibilité ne soit pas trop conservatif.

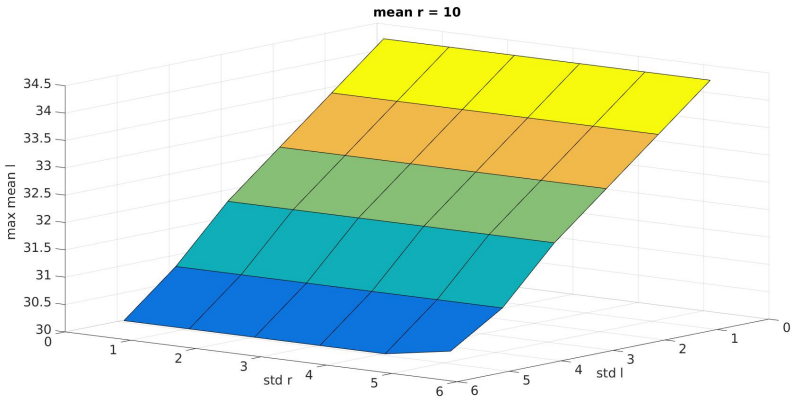


FIGURE 8.2 – Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 10$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$.

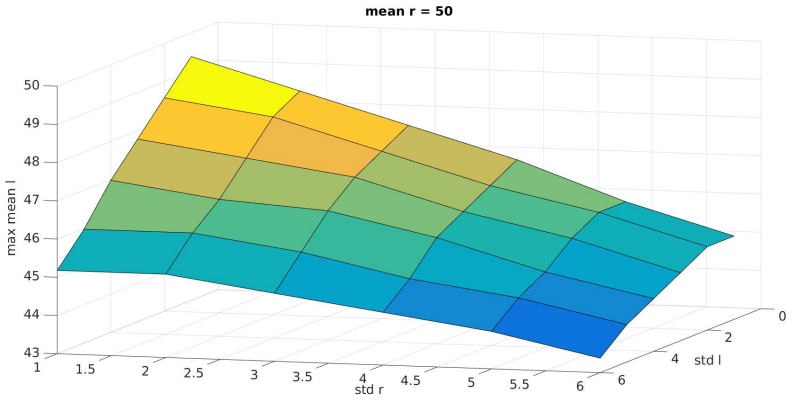


FIGURE 8.3 – Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 50$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$.

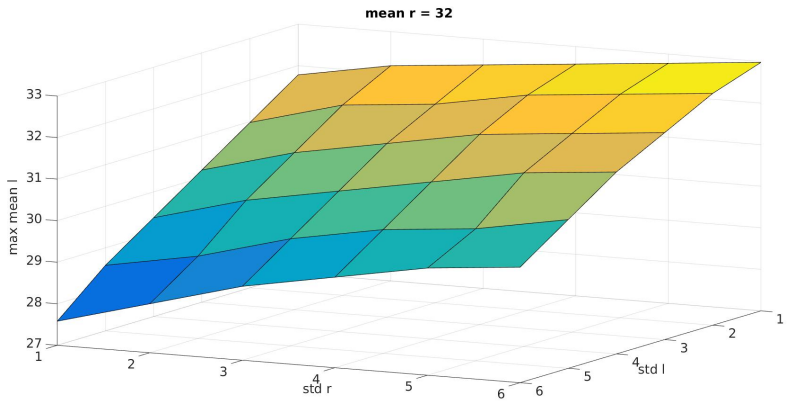


FIGURE 8.4 – Valeur de $\widehat{m}_l^{\max}$ obtenue par échantillonnage pour $\widehat{m}_r = 32$ et différentes valeurs de $\widehat{\sigma}_r$ et $\widehat{\sigma}_l$.

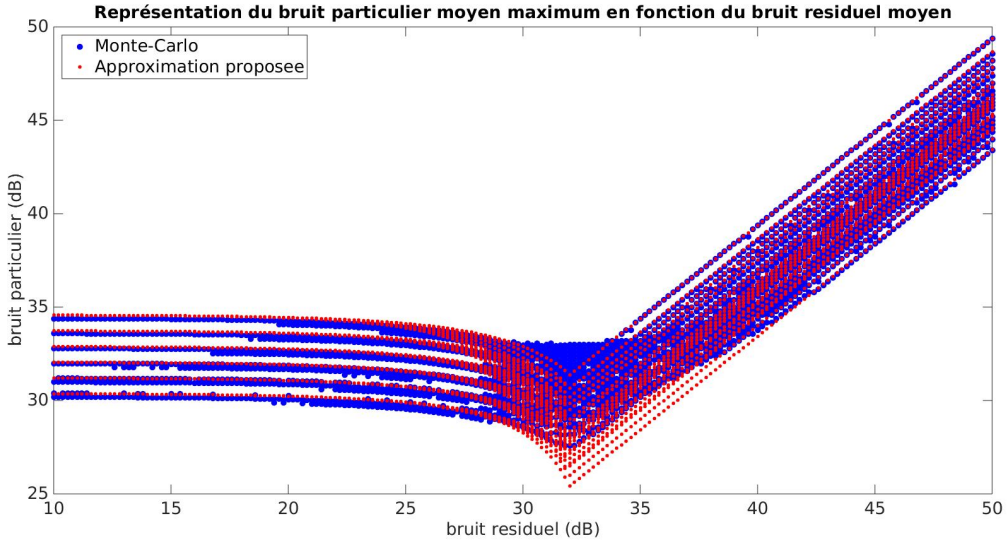


FIGURE 8.5 – Comparaison entre les valeurs de $\widehat{m}_l^{\max}$ obtenues par Monte-Carlo, et celles obtenues par l'approximation proposée en fonction de $\widehat{m}_r$.

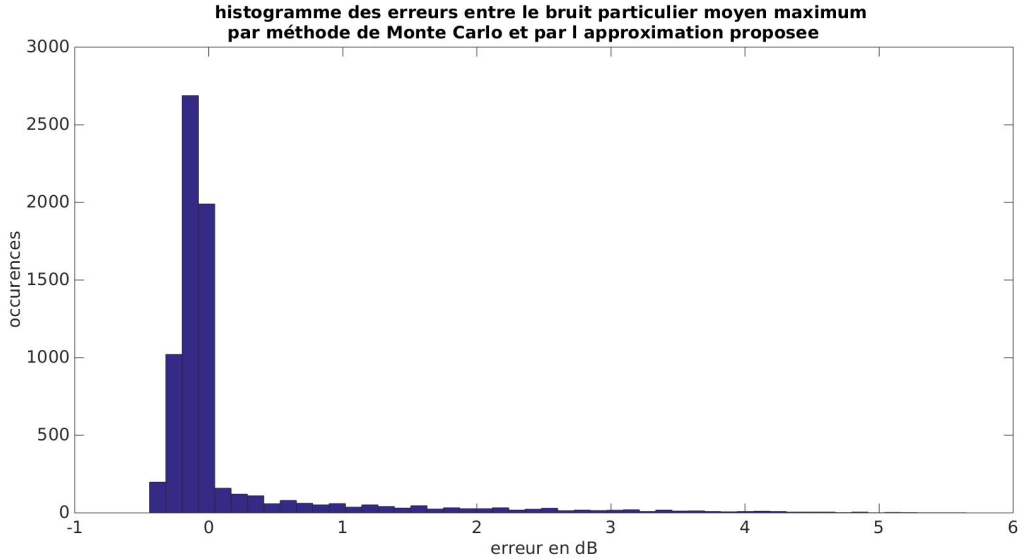


FIGURE 8.6 – Histogramme des erreurs entre les valeurs de $\widehat{m}_l^{\max}$ obtenues par Monte-Carlo, et celles obtenues par l'approximation proposée.

8.4 Adaptation de l'algorithme branch and bound pour le problème stochastique

On veut adapter l'algorithme d'optimisation *branch and bound* déterministe pour l'optimisation du problème stochastique. On rappelle que, pour le problème déterministe, les deux éléments-clés étaient les suivants :

— **Admissibilité** : pour construire l'arbre de recherche, il était nécessaire dans l'algorithme déterministe de pouvoir vérifier l'admissibilité des nœuds fils possibles d'un nœud quel-

conque. Pour ce faire, on considérait la condition suivante : soient $\{1, ..., I_0\}$ les indices des éoliennes dont les modes de bridage sont déjà déterminés. Le nœud fils correspondant à l'éolienne $I_0 + 1$ et au mode c_{I_0+1} est admissible si et seulement si $\forall j \in \{1, ..., J\}$, $10^{\frac{\mathfrak{r}(c_{I_0+1})+a_{I_0+1,j}}{10}} + \sum_{i=1}^{I_0} 10^{\frac{l_{i,j}}{10}} \leq q_j$. De façon imagée, il s'agit de voir si en allumant l'éolienne $I_0 + 1$ avec le mode c_{I_0+1}, la contrainte est toujours respectée.

— **Majoration** : afin d'éviter de parcourir tout l'arbre, l'algorithme *branch and bound* repose sur l'utilisation d'une technique de majoration qui doit présenter les caractéristiques suivantes en chaque nœud :

1. Fournir un majorant de la production espérée $\mathbb{E}(p_{t,i}|c_{t,i}, o_t)$ pour toutes les solutions issues de l'arborescence du nœud considéré.
2. Être rapide à calculer et le plus proche possible d'une borne supérieure.

La première condition est une condition pour l'optimalité de l'algorithme, et la seconde une condition pour la rapidité de celui-ci. Dans le cadre déterministe, on avait proposé une majoration basée sur un problème de relaxation continue, où les bridages des éoliennes non déterminées d'un nœud donné sont considérés continus :

$$\begin{cases} \max_{(\tilde{x}_{I_0+1},...,\tilde{x}_I)\in[0,\tilde{\mathfrak{r}}_v(1)]^{(I-I_0)}} \sum_{i=I_0+1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{free}} \leq \mathbf{q} - \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{det}}. \end{cases} \tag{8.14}$$

Plaçons nous maintenant et pour simplifier dans le cadre du problème *single chance constraint* donné en (8.1) et construisons les deux critères nécessaires pour adapter l'algorithme.

On suppose désormais qu'on peut calculer $\widehat{m}_{l_j}^{\max}$ en chaque point de mesure comme une fonction notée $\mathfrak{q}$, de $\widehat{\sigma}_{l_j}$, $\widehat{\sigma}_{r_j}$, $\widehat{m}_{r_j}$, S_s, S_e et ε et on a donc $\widehat{m}_{l_j}^{\max} = \mathfrak{q}(\widehat{\sigma}_{l_j}, \widehat{\sigma}_{r_j}, \widehat{m}_{r_j}, S_s, S_e, \varepsilon)$. En pratique, la connaissance de cette fonction $\mathfrak{q}$ sera donnée par une approximation comme présentée dans la partie 8.2.3 ou par apprentissage d'une évalution de type Monte-Carlo comme présenté dans la partie 8.2.2.

D'autre part, on considère que $\widehat{\sigma}_{r_j}$, $\widehat{m}_{r_j}$, S_s, S_e et ε sont déterminés. Les valeurs de $\widehat{\sigma}_{r_j}$, $\widehat{m}_{r_j}$ sont données par l'étape de prédiction dans le filtre de Kalman et S_s, S_e, ε sont des valeurs fixées par l'utilisateur. En pratique, $\widehat{m}_{l_j}^{\max}$ ne dépend donc que de $\widehat{\sigma}_{l_j}$ et on note alors $\widehat{m}_{l_j}^{\max} = \mathfrak{q}(\widehat{\sigma}_{l_j})$. D'autre part, on note $\mathfrak{q}(\widehat{\mathbf{\Sigma}}_l) = [\mathfrak{q}(\widehat{\sigma}_{l_j})]_j$ et $\tilde{\mathfrak{q}}(\widehat{\mathbf{\Sigma}}_l) = \left[10^{\frac{\mathfrak{q}(\widehat{\sigma}_{l_j})}{10}}\right]_j$.

On peut alors remarquer que, par rapport au problème déterministe, on a une contrainte q qui n'est pas constante, mais qui dépend de la valeur de $\widehat{\sigma}_{l_j}$.

L'idée est alors de reformuler des conditions d'admissibilité et de majoration pour le problème stochastique de la manière suivante :

— **Admissibilité** : À chaque étape, on calcule $\widehat{\mathbf{m}}_{\mathbf{l}}$ et $\widehat{\mathbf{\Sigma}}_{\mathbf{l}}$ directement à partir de $\widehat{\mathbf{m}}_{\mathbf{x}}$, $\widehat{\mathbf{m}}_{\mathbf{a}}$, $\widehat{\sigma}_x$ et $\widehat{\sigma}_a$, par une méthode de linéarisation (développement de Taylor d'ordre 1 ou linéarisation statistique). On vérifie ensuite si $\forall j \in \{1, ..., J\}, \widehat{m}_{l_j} \leq \mathfrak{q}(\widehat{\sigma}_{l_j})$.

— **Majoration** : Pour l'étape de majoration, on utilise une relaxation continue similaire au problème déterministe, à la différence que le seuil $\tilde{\mathbf{q}}$ maximum de bruit admissible est ajusté en fonction de la variance $\widehat{\mathbf{\Sigma}}_l^{\text{det}}$ des éoliennes déjà déterminées :

$$\begin{cases} \max_{(\tilde{x}_{I_0+1},...,\tilde{x}_I)\in[0,\tilde{\mathfrak{r}}_v(1)]^{(I-I_0)}} \sum_{i=I_0+1}^{I} \mathfrak{f}(\tilde{x}_i) \\ \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{free}} \leq \tilde{\mathbf{q}}(\widehat{\mathbf{\Sigma}}_l^{\text{det}}) - \tilde{\mathbf{A}}^T \tilde{\mathbf{x}}_{\text{det}}. \end{cases} \tag{8.15}$$

En somme, il s'agit d'ajuster l'algorithme en prenant une marge adaptée et fonction de ε. Notons toutefois que cet algorithme constitue une heuristique dans la mesure où la condition d'admissibilité proposée est une approximation.

Cinquième partie

Conclusion et perspectives

9

Conclusion et perspectives

9.1 Travail effectué

Durant ce doctorat, on a construit un formalisme qui a permis de situer le problème dans la littérature scientifique et de proposer différents algorithmes issus et adaptés de l'état de l'art pour résoudre une partie des sous-problèmes spécifiques nécessaires au contrôle acoustique des parcs éoliens. Le formalisme, les algorithmes et les expériences présentées ont permis d'implémenter deux solutions de contrôle et de répondre à certaines des problématiques concernant la faisabilité technique d'un système commercial.

9.1.1 Solution déterministe

La solution déterministe s'appuie sur le découpage usuel des observations en classes homogènes, sur les données électriques et acoustiques et sur un algorithme d'optimisation permettant de résoudre le problème instantané consistant à optimiser la commande sans considérer d'éventuelles dépendances aux instants passés et futurs. En pratique, cette solution peut être mise en place en construisant un problème déterministe à partir des données acoustiques et météorologiques mesurées à un instant donné qui définissent alors les données du problème d'optimisation à résoudre. L'estimation par séparation de sources et la période de la journée définissent le niveau particulier maximal $\mathbf{q}$ de la contrainte, et les autres données du problème $\mathfrak{F}$, $\mathbf{A}_t$ et les valeurs possibles pour $\tilde{\mathbf{x}}_t$ sont définies à l'aide des données constructeurs, des données de propagation simulées, et de la vitesse et de la direction du vent mesurées à cet instant.

Résolution du problème instantané

En analysant le problème, on a d'abord montré que le problème instantané pouvait se formuler comme un problème d'optimisation convexe non-linéaire, dont les solutions sont à valeurs dans un ensemble discret et sous contraintes linéaires. En outre, la formulation du problème est plus générale que le problème initialement posé et il est en pratique possible de résoudre n'importe quel problème instantané de commande maximisant la production électrique, quelle que soit la contrainte de bruit particulier maximal admissible. Pour sa résolution, on a construit un algorithme *branch and bound*, initialisé par une heuristique gloutonne efficace, qui s'avère être déjà optimale dans 70% des cas étudiés. L'algorithme *branch and bound* permet ensuite de raffiner la solution. Pour majorer les nœuds de l'arbre, on résout pour chaque nœud un problème de relaxation continue construit en prolongeant la fonction objectif linéairement entre chaque

mode et à l'aide d'un algorithme de gradient adapté à la forme linéaire par morceaux. L'algorithme s'est avéré optimal dans la très grande majorité des cas (98%) et occasionnant une perte électrique relativement faible (2%). Il est de plus utilisable dans un contexte industriel, le temps de convergence n'excédant que très rarement les 2 minutes sur l'ensemble des cas testés.

Remarques sur le critère réglementaire

Concernant la solution déterministe, on a également étudié la prise en compte du critère basé sur des médianes temporelles de mesure de bruit éolien.

On a montré que ce critère, pertinent dans le cas d'un plan de bridage statique (bridage constant utilisé usuellement en ingénierie acoustique), perd sa pertinence dans le cas d'un plan de bridage dynamique. En effet, un des algorithmes testés permet de respecter le critère basé sur des médianes temporelles en jouant sur la définition de la médiane, tout en augmentant le niveau particulier moyen comparativement au bridage statique.

Ce fait permet de justifier l'utilisation du critère instantané qui est plus simple et s'inscrit toujours dans les normes françaises de mesure du bruit dans l'environnement. En effet, le critère basé sur des médianes a été à l'origine défini par l'agence de normalisation de manière pragmatique dans la mesure où les systèmes de mesure existants ne permettaient pas un suivi fin de l'évolution des niveaux sonores. Le nouveau système de contrôle proposé dans cette thèse résout cette limitation et permet de considérer un critère instantané plus respectueux des riverains.

9.1.2 Solution stochastique

La seconde solution repose sur une extension stochastique permettant de prendre en compte les incertitudes du problème. Le suivi et l'estimation des incertitudes se font avec un filtre de Kalman non-linéaire et un algorithme espérance-maximisation. Nous avons proposé un modèle de filtrage et avons découpé l'estimation des états cachés et des paramètres d'incertitude en trois travaux.

Un premier travail, semblable à l'étape M de l'algorithme EM et portant sur l'estimation des paramètres d'incertitude a été proposé. L'objectif a consisté à déterminer les paramètres d'incertitude en maximisant la fonction de log-vraisemblance sur une série de données et en supposant connu l'état caché du système à chaque instant. Deux approximations différentes de la fonction de log-vraisemblance ont été testées. Le problème est apparu être expérimentalement convexe, l'estimation de bonne qualité et présentant une erreur relative de l'ordre de 10%.

Ensuite, nous avons défini un modèle d'état pour le problème et différents filtres de Kalman ont été évalués expérimentalement. L'estimation de l'état caché par filtrage s'est avérée être en général de bonne qualité et le filtrage a permis de diminuer drastiquement l'incertitude sur la séparation de sources en la fusionnant avec l'incertitude du modèle. Parmi les filtre essayés, le filtre *unscented* a été retenu, s'avérant être plus le efficace et permettant d'obtenir une erreur RMSE moyenne d'environ 1,5 dB contre une erreur RMSE d'environ 3,5 dB pour l'estimation par séparation de source seule sans filtrage. En outre il s'est avéré être le plus consistant.

En revanche, l'estimation globale de l'état caché et des paramètres d'incertitude par un algorithme EM s'est montrée plus difficile. L'algorithme EM proposé fonctionne à ceci près qu'il n'arrive pas à estimer correctement les écarts-types sur l'émission et la propagation. En effet, le modèle n'arrive pas à « discerner » comment se répartit l'erreur entre l'émission et la propagation. Après reparamétrisation, on arrive tout de même à estimer l'écart-type σ_l du niveau particulier lorsque l'émergence est élevée sans toutefois y arriver finement lorsque celle-ci est faible. Ceci est dû à un phénomène de masquage du niveau particulier qui survient lorsque l'émergence est faible,

le niveau ambiant étant alors à peu près égal au niveau résiduel. Ces résultats sont à mettre en perspective avec un système de séparation réel où l'incertitude de séparation est certainement dépendante du rapport-signal à bruit. De ce fait, il serait donc nécessaire de modéliser cette dépendance afin de pouvoir conclure sur la précision de l'estimation dans un cas réel. Néanmoins les résultats de l'algorithme EM montrent que l'approche est numériquement valide et on pourra apprendre les paramètres d'incertitude du bruit particulier dans les périodes où l'émergence sera suffisamment élevée, et au contraire du bruit résiduel dans les périodes où elle sera faible.

Pour finir, une adaptation simple de l'algorithme branch-and-bound déterministe a été présentée également afin de prendre en compte les incertitudes estimées et de résoudre le problème *single chance constraint programing* et d'assurer ainsi une certaine robustesse au contrôle. Tout d'abord, nous avons proposé de caractériser l'admissibilité des solutions en déterminant quelle est la valeur de niveau particulier maximale autorisée, non seulement par rapport aux seuils réglementaires S_s, S_e et au niveau résiduel $\mathbf{r}_t$ comme dans le cas déterministe, mais aussi par rapport au niveau de tolérance ε et aux écarts-types des incertitudes σ_r et σ_l. Pour ce faire, nous avons proposé une méthode de Monte-Carlo, ainsi qu'une approximation. Ensuite nous avons présenté une adaptation de l'algorithme *branch and bound* qui s'appuie sur cette caractérisation pour transformer le problème stochastique en un problème déterministe et le résoudre.

9.2 Perspectives concernant la définition et l'apprentissage des paramètres

Le contrôle acoustique des parcs éoliens étant un sujet totalement nouveau, les améliorations possibles sont vastes. Dans la suite, on présente des perspectives à court et long terme d'évolution du modèle proposé.

9.2.1 Redéfinition des classes homogènes par des méthodes de clustering

Le premier constat concernant le modèle est l'amélioration potentielle que pourrait apporter un découpage des variables environnementales plus fin et adapté. À l'heure actuelle, les classes homogènes sont définies intuitivement et suivant les orientations de la norme par les ingénieurs acousticiens lors de l'étude initiale des parcs (voir chapitre 1). Or rien ne garantit que cette méthode permet d'obtenir des classes homogènes où la distribution des incertitudes est concentrée. Une piste d'amélioration du modèle serait donc d'utiliser des méthodes d'apprentissage automatique afin de déterminer des classes homogènes plus adaptées aux données récupérée lors de l'étude acoustique du parc à contrôler. L'idée est de trouver une régression entre la vitesse de vent v, la direction du vent d et l'heure de la journée p et le niveau résiduel observé $\mathbf{r}$ et les valeurs d'atténuation $\mathbf{a}$. Pour ce faire, on pourrait s'inspirer par exemple de la méthode de [Deleforge et al., 2015] permettant de faire une régression paramétrique linéaire par morceaux entre des variables d'entrées et de sortie basé par le biais de mélanges de distributions gaussiennes.

9.2.2 Utilisation d'une méthode de gradient pour l'apprentissage des paramètres basée sur la méthode des moments proposée

La méthode des moments proposée dans la partie 7.3 est efficace lorsque le niveau résiduel est connu mais ne peut pas être utilisée directement dans l'algorithme EM proposé dans la partie 7.5 étant donné que l'intégrale de la log-vraisemblance sur l'état caché n'a pas à notre connaissance d'expression formelle. En revanche, cette méthode pourrait être utilisée dans le schéma général d'estimation des incertitudes en utilisant une méthode de gradient comme présentée dans l'état

de l'art du chapitre 2. En effet, on rappelle qu'il est possible de calculer à partir des données, le gradient de la log-vraisemblance en utilisant l'identité de Fisher :

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}_z(\boldsymbol{\theta}) = \int \nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta}) d\mathbf{y}_{1:T}. \tag{9.1}$$

Or le terme $\log p(\mathbf{z}_{1:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})$ est précisément la log-vraisemblance simplifiée où l'état caché $\mathbf{y}_{1:T}$ est supposé connu et qui est calculée par cette méthode des moments dans la partie 7.3. En calculant alors son gradient et en intégrant celui-ci numériquement par un échantillonnage de Monte-Carlo sur la distribution $p(\mathbf{y}_{1:T}|\mathbf{z}_{1:T}, \boldsymbol{\theta})$ obtenue par filtrage récursif bayésien, on obtient une nouvelle méthode d'estimation des paramètres du modèle d'incertitude $\boldsymbol{\theta}$. En effet, on peut ensuite appliquer la mise-à-jour des paramètres à chaque itération k l'aide de l'équation suivante :

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \lambda_k \nabla_{\boldsymbol{\theta}} \mathcal{L}_z(\boldsymbol{\theta}). \tag{9.2}$$

Ce faisant, on peut espérer une éventuelle amélioration par rapport à l'algorithme EM, dans la mesure où la méthode des moments s'est avérée très précise comparativement à la linéarisation simple lors de l'évaluation dans le problème simplifié. Cependant, il est nécessaire de choisir un pas λ_k et une méthode de Monte-Carlo adaptés.

9.2.3 Ajout d'une loi a priori sur les incertitudes et estimation des biais du modèle

Un aspect de l'estimation des incertitudes qui n'a pas été exploré est l'estimation des biais du modèle en vue d'une recalibration. En effet, les constructeurs prennent en général une marge de confiance concernant l'émission des éoliennes et la simulation numérique de la propagation souffre en général d'un biais. De plus, le biais peut évoluer lentement au cours du temps en ce qui concerne la propagation et l'émission. En effet, la propagation et l'émission dépendent de la température, de l'humidité, et de conditions de turbulence du vent. Leur influence peut être considérée comme faible sur une courte période de temps mais elles n'en demeurent pas moins une source d'incertitudes. De ce fait, une manière d'améliorer le modèle est de considérer le modèle avec des paramètres de biais suivant :

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{a}_t \\ \mathbf{r}_t \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathbf{x}} \mathbf{x}_{t-1} + \mathbf{G}_{\mathbf{x}} \mathbf{m}_{\mathbf{x}_t} + \boldsymbol{\mu}_{\mathbf{x}} + \mathbf{w}_{\mathbf{x}_t} \\ \mathbf{F}_{\mathbf{a}} \mathbf{a}_{t-1} + \mathbf{G}_{\mathbf{a}} \mathbf{m}_{\mathbf{a}_t} + \boldsymbol{\mu}_{\mathbf{a}} + \mathbf{w}_{\mathbf{a}_t} \\ \mathbf{F}_{\mathbf{r}} \mathbf{r}_{t-1} + \mathbf{G}_{\mathbf{r}} \mathbf{m}_{\mathbf{r}_t} + \boldsymbol{\mu}_{\mathbf{r}} + \mathbf{w}_{\mathbf{r}_t} \end{bmatrix}, \tag{9.3}$$

$$b^*_{t,j} = 10 \log_{10} \left(\sum_i^I 10^{\frac{a_{t,i,j}}{10}} 10^{\frac{x_{t,i}}{10}} + 10^{\frac{r_{t,j}}{10}} \right) + \mu_{b^*_j} + w_{b^*_{t,j}}, \tag{9.4}$$

$$r^{*\bullet}_{t,j} = r_{t,j} + \mu_{r^\bullet_j} + \mu_{b^*_j} + w_{r^\bullet_{t,j}} + w_{b^*_{t,j}}. \tag{9.5}$$

avec $\boldsymbol{\mu}_{\mathbf{x}}$, $\boldsymbol{\mu}_{\mathbf{a}}$, $\boldsymbol{\mu}_{\mathbf{r}}$, $\boldsymbol{\mu}_{\mathbf{b}^*}$ et $\boldsymbol{\mu}_{\mathbf{r}^\bullet}$ les vecteurs de biais ajoutés aux incertitudes du modèle. Le calcul des biais peut se faire en principe de façon similaire au chapitre 7 en utilisant un algorithme EM et un filtre de Kalman non-linéaire. Cependant, on ajoute au modèle de nouveaux degrés de liberté qui pourraient éventuellement rendre le problème non strictement convexe. Une manière de voir cela est de considérer une seule éolienne et un point de mesure : on observe par séparation de sources le niveau particulier l et on a l'égalité $\mu_l = \mu_x + \mu_a$, ce qui engendre une indétermination sur μ_x et μ_a. Pour remédier à cela, une manière de procéder serait de considérer une loi à *priori* sur l'ensemble des paramètres à estimer $\boldsymbol{\mu}_{\mathbf{x}}$, $\boldsymbol{\mu}_{\mathbf{a}}$, $\boldsymbol{\mu}_{\mathbf{r}}$, $\boldsymbol{\mu}_{\mathbf{b}^*}$, $\boldsymbol{\mu}_{\mathbf{r}^\bullet}$, $\sigma_{\mathbf{x}}$, $\sigma_{\mathbf{a}}$, $\sigma_{\mathbf{r}}$, $\sigma_{\mathbf{b}^*}$ et $\sigma_{\mathbf{r}^\bullet}$, idéalement une loi conjuguée, c'est-à-dire une loi conjointe Gaussian-inverse-Wishart sur l'ensemble des paramètres [Bishop, 2007, p.690].

9.2.4 Paramétrisation de l'incertitude de séparation de sources en fonction du rapport signal-à-bruit

Un des reproches que l'on peut adresser au modèle d'incertitude concerne la modélisation de l'incertitude sur la séparation de sources. Comme décrit dans la partie 7.2, l'incertitude sur la séparation de sources dépend en règle générale du rapport énergétique entre les différentes sources et donc, dans le contexte de l'éolien, de l'émergence. Une piste d'amélioration du modèle est d'utiliser une paramétrisation s'inspirant de Tran et al. [2015]. On construit le modèle en considérant que les paramètres de moyenne et de variance régissant l'incertitude $\mathbf{w}_{\mathbf{r}_t^\bullet}$ sont dépendants de l'émergence

$$\mathbf{w}_{\mathbf{r}_t^\bullet} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{r}_t^\bullet}(\mathbf{e}_t), \boldsymbol{\sigma}_{\mathbf{r}_t^\bullet}(\mathbf{e}_t)). \tag{9.6}$$

Dans ce cas, en lieu et place d'une simple estimation de paramètre, on obtient un problème d'estimation de l'incertitude conditionnellement à la valeur de l'émergence. Pour résoudre ce problème en pratique, une manière de procéder serait d'utiliser une méthode similaire à Tran et al. [2015] basée sur une régression sur une base de fonctions afin d'approcher les fonctions $\boldsymbol{\mu}_{\mathbf{r}_t^\bullet}(\mathbf{e}_t)$ et $\boldsymbol{\sigma}_{\mathbf{r}_t^\bullet}(\mathbf{e}_t)$.

9.2.5 Modèle de prédiction

Comme présenté dans le chapitre 7, des tendances apparaissent dans l'évolution du niveau résiduel. Notamment, il apparaît assez clairement une tendance journalière : des valeurs comparables s'observent d'un jour à l'autre aux même heures de la journée. Or, s'il est difficile de battre le prédicteur *persistance* à court terme comme cela a été présenté dans l'état de l'art au chapitre 2, il n'en demeure pas moins qu'une piste d'évolution serait d'utiliser un modèle ARIMA ou un modèle basé sur un réseau de neurones afin de prédire l'évolution du niveau résiduel statistiquement de façon plus fine.

En outre, un tel prédicteur pourrait être intégré dans le modèle de filtrage en incorporant dans le modèle d'état la nouvelle équation de prédiction et en modifiant légèrement les étapes de prédiction et de mise à jour du filtrage.

Par exemple, on peut étendre le filtre de Kalman proposé en remplaçant l'équation de prédiction par un modèle ARMA :

$$\mathbf{y}_t = \sum_{i=1}^{i_\alpha} \alpha_i \mathbf{y}_{t-i} + \sum_{i=0}^{i_\beta} \beta_i \mathbf{w}_{\mathbf{y}_{t-i}}, \tag{9.7}$$

$$\mathbf{z}_t = h(\mathbf{y}_t, \mathbf{w}_{\mathbf{z}_t}, \boldsymbol{\theta}_{\mathbf{z}}). \tag{9.8}$$

Dans ce cas, les étapes du filtrage de Kalman peuvent être adaptées. Concernant l'étape de prédiction, l'équation (9.7) du modèle ARMA permet de calculer la moyenne et la variance prédite, et ce calcul remplace alors les équations usuelles de prédiction de Kalman. Concernant l'étape de mise-à-jour, aucune modification n'est nécessaire ici. On peut utiliser directement les équations usuelles de mise à jour du filtre de Kalman car l'équation (9.8) est identique à celle utilisée dans le chapitre 7.

De manière générale, le filtre de Kalman non-linéaire peut être adapté avec d'autres modèles de prédiction du moment qu'il est possible de calculer une prédiction de la moyenne et de la variance de $\mathbf{y}_t$. On peut également imaginer des modèles plus complexes basés sur des réseaux de neurones et des filtres à particules, par exemple.

Notons cependant que l'estimation des paramètres d'incertitudes est alors plus difficile, la taille du problème et le nombre de paramètres augmentant. Une manière de procéder pour l'exemple ARMA-Kalman est par exemple d'augmenter l'état [Mader et al., 2014].

9.3 Perspectives concernant le calcul de la commande optimale

9.3.1 Utilisation d'un modèle acoustique par bandes de fréquences

Concernant le problème d'optimisation, il est possible de poser le problème plus finement en prenant en compte indépendamment chaque bande de fréquences. En effet, le problème d'optimisation posé dans le chapitre 4 suppose une certaine linéarité du problème d'optimisation résumé par l'équation $\tilde{l}_{i,j} = \tilde{a}_{i,j}\tilde{x}_i$. L'hypothèse sous-jacente est que la perte relative de puissance en décibel générée par un changement de mode de bridage est constante en fréquence. En réalité, l'atténuation diffère en fonction de la fréquence. Par exemple, considérons une éolienne et un unique point de mesure (on enlève alors les indices i et j). Supposons également que la variation d'un mode à un autre soit constante au sein de 2 bandes de fréquences, disons sur les bandes $f_1 = [50\text{ Hz } 1\text{ kHz}]$ et $f_2 = [1\text{ kHz } 20\text{ kHz}]$ par exemple mais différente d'une bande à l'autre. Dans ce cas le problème d'optimisation qui était posé jusqu'alors avec la contrainte

$$ax \leq q \tag{9.9}$$

devrait plutôt être réécrit avec la contrainte

$$a_{f_1}x_{f_1} + a_{f_2}x_{f_2} \leq q. \tag{9.10}$$

Par exemple, le passage du mode 1 au mode 2 pourrait très bien engendrer une perte d'1 dB pour x_{f_1} et de 2 dB pour x_{f_2} alors que le passage du mode 2 à 3 engendrerait lui une perte de 1 dB pour x_{f_1} et de 5 dB pour x_{f_2}. La prise en compte des différentes bandes de fréquence nécessite cependant un travail plus minutieux concernant les données constructeur et le modèle de propagation. D'autre part, elle augmente largement la complexité et la taille du problème d'optimisation.

9.3.2 Apprentissage de la contrainte dans le problème *chance constraint programing*

Dans le chapitre 8, on a proposé de caractériser l'admissibilité des contraintes, en recherchant la valeur de m_l maximale admissible en fonction de m_r, σ_r et σ_l, pour un seuil de tolérance ε donné. Pour ce faire, on a notamment proposé d'utiliser une méthode de Monte-Carlo, en faisant varier m_r, σ_r et σ_l et en recherchant la valeur m_l maximale itérativement. Cette analyse est prometteuse, mais nécessite un apprentissage. Une piste de recherche à ce sujet serait par exemple d'utiliser une méthode à noyau gaussien, comme la méthode de Nadaraya-Watson [Härdle, 2012], afin de permettre un calcul rapide, pour n'importe quelle valeur de m_r, σ_r, σ_l et ε, de la valeur maximale admissible m_l.

A

Formules Mathématiques

A.1 Formules mathématiques de passage d'une loi log-gaussienne à une loi gaussienne

Soit $\mathbf{y}$ vecteur gaussien de moyenne $\mathbf{m_y}$ et de covariance $\mathbf{\Sigma_y}$. On note $\tilde{\mathbf{y}} = \exp(\mathbf{y})$ le vecteur log-gaussien associé de moyenne $\mathbf{m_{\tilde{y}}}$ et de covariance $\mathbf{\Sigma_{\tilde{y}}}$. On a les formules de passage exactes suivantes [Gales, 1995, p. 47] :

$$m_{\tilde{y}_i} = \exp\left(m_{y_i} + \frac{\sigma^2_{y_i}}{2}\right) \tag{A.1}$$

et

$$\sigma_{\tilde{y}_i,\tilde{y}_j} = m_{\tilde{y}_i} m_{\tilde{y}_j}[\exp(\sigma_{y_i,y_j}) - 1]. \tag{A.2}$$

Inversement, mais cette fois de manière approchée, on a

$$m_{y_i} \simeq \log(m_{\tilde{y}_i}) - \frac{1}{2}\log\left(\frac{\sigma^2_{\tilde{y}_i}}{m^2_{\tilde{y}_i}} + 1\right) \tag{A.3}$$

et

$$\sigma_{y_i,y_j} \simeq \log\left(\frac{\sigma_{\tilde{y}_i,\tilde{y}_j}}{m_{\tilde{y}_i} m_{\tilde{y}_j}} + 1\right). \tag{A.4}$$

Notons que les formules que l'on utilise sont légèrement différentes car on considère $x_i = 10\log_{10}(\tilde{x}_i) = \frac{10}{\log(10)}\log(\frac{\tilde{x}_i}{p_0^2})$ (passage dB à linéaire et inversement). En posant donc $x_i = k \times y_i$ avec $k = \frac{10}{\log(10)}$ et $\tilde{x}_i = p_0^2 \times \tilde{y}_i$ on obtient

$$m_{\tilde{x}_i} = p_0^2 \exp\left(\frac{m_{x_i}}{k} + \frac{\sigma^2_{x_i}}{2k^2}\right), \tag{A.5}$$

$$\sigma_{\tilde{x}_i,\tilde{x}_j} = m_{\tilde{x}_i} m_{\tilde{x}_j}\left[\exp\left(\frac{\sigma_{x_i,x_j}}{k^2}\right) - 1\right], \tag{A.6}$$

$$m_{x_i} \simeq k \times \log\left(\frac{m_{\tilde{x}_i}}{p_0^2}\right) - \frac{1}{2}\log\left(\frac{\sigma^2_{\tilde{x}_i}}{m^2_{\tilde{x}_i}} + 1\right), \tag{A.7}$$

et

$$\sigma_{x_i,x_j} \simeq k^2 \times \log\left(\frac{\sigma_{\tilde{x}_i,\tilde{x}_j}}{m_{\tilde{x}_i} m_{\tilde{x}_j}} + 1\right). \tag{A.8}$$

B

Détails des résultats de l'algorithme EM

σ_x		σ_a		$\sigma_r^\bullet$		σ_{r_1}		σ_{r_2}		σ_{r_3}		σ_{r_4}		σ_{r_5}	
gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.
0,5	2,71	1,5	6,37	0,5	0,474	2,51	2,6	3,84	3,81	3,48	3,49	4,71	4,71	4,61	4,64
1,5	3,73	2,5	6,37		0,476		2,46		3,76		3,49		4,7		4,6
2,5	8,43	3,5	5,62		0,483		2,36		3,83		3,48		4,7		4,61
3,5	3,38	4,5	9,7		0,552		2,43		3,78		3,48		4,68		4,6
0,5	1,22	1,5	2,36	1,5	1,55		2,52		3,36		3,49		4,71		4,64
1,5	1,61	2,5	3,41		1,53		2,53		3,18		3,48		4,7		4,62
2,5	1,63	3,5	4,29		1,53		2,32		3,49		3,49		4,69		4,7
3,5	2,1	4,5	4,19		1,67		2,47		3,28		3,49		4,71		4,74
0,5	1,31	1,5	2,14	2,5	2,28		2,48		3,46		3,49		4,71		4,59
1,5	1,26	2,5	2,96		2,5		2,53		3,61		3,49		4,71		4,71
2,5	1,49	3,5	3,05		2,54		2,36		3,31		3,48		4,7		4,6
3,5	1,71	4,5	3,25		2,59		2,47		3,02		3,48		4,7		4,64
0,5	1,22	1,5	2,16	3,5	3,57		2,52		3,29		3,49		4,71		4,61
1,5	1,4	2,5	2,53		3,44		2,49		3,34		3,49		4,71		4,63
2,5	1,5	3,5	2,95		3,55		2,49		3,05		3,48		4,7		4,59
3,5	1,89	4,5	3,66		3,45		2,31		3,5		3,48		4,66		4,68
0,5	1,36	1,5	2,33	4,5	4,59		2,49		3,07		3,48		4,7		4,62
1,5	1,83	2,5	3,09		4,29		2,42		3,05		3,48		4,71		4,62
2,5	1,93	3,5	3,34		4,45		2,35		2,96		3,49		4,7		4,62
3,5	2,04	4,5	4,1		4,55		2,44		3,33		3,48		4,69		4,64

TABLEAU B.1 – Résultat de l'algorithme EM sur le premier ensemble de cas de test, contenant 20 cas différents où l'émergence moyenne sur le point le plus exposé est de 1,8 dBA. Les colonnes gt. et est. indiquent la vérité terrain du paramètre à estimer et la valeur obtenue à la fin de l'algorithme EM.

ensemble 2

σ_x		σ_a		$\sigma_r^\bullet$		σ_{r_1}		σ_{r_2}		σ_{r_3}		σ_{r_4}		σ_{r_5}	
gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.	gt.	est.
0,5	0,449	1,5	2	0,5	0,593	2,51	2,33	3,84	3,69	3,48	3,46	4,71	4,72	4,61	4,65
1,5	0,991	2,5	3,59		0,637		2,48		3,66		3,47		4,7		4,62
2,5	1,07	3,5	4,58		0,633		2,3		3,77		3,46		4,51		4,62
3,5	1,12	4,5	6,41		0,691		2,27		3,69		3,46		4,54		4,68
0,5	0,418	1,5	1,74	1,5	1,55		2,43		3,57		3,46		4,7		4,73
1,5	0,933	2,5	2,49		1,65		2,3		3,5		3,46		4,77		4,84
2,5	2,27	3,5	4,1		1,59		1,77		3,21		3,46		4,57		4,54
3,5	1,08	4,5	5,72		1,6		2,12		3,55		3,45		4,66		4,71
0,5	0,534	1,5	1,51	2,5	2,63		2,49		3,23		3,46		4,76		4,89
1,5	0,846	2,5	2,59		2,48		2,25		3,59		3,46		4,76		4,79
2,5	0,949	3,5	4,29		2,59		1,93		3,31		3,46		4,72		5,2
3,5	1,56	4,5	6,71		2,66		2,07		2,83		3,45		4,6		4,75
0,5	0,567	1,5	1,96	3,5	3,43		2,47		3,12		3,47		4,77		4,68
1,5	0,689	2,5	2,63		3,36		1,99		2,94		3,46		4,73		4,94
2,5	1,12	3,5	4,31		3,62		2		2,83		3,46		4,62		4,82
3,5	1,11	4,5	5,63		3,64		1,78		3,51		3,46		4,57		5,03
0,5	0,685	1,5	1,78	4,5	4,3		2,2		2,8		3,46		4,76		4,82
1,5	0,805	2,5	2,35		4,6		2,13		4,07		3,45		4,68		5
2,5	0,94	3,5	4,64		4,22		2,34		4		3,47		4,7		4,77
3,5	1,09	4,5	5,12		4,71		2,5		2,34		3,46		4,58		5,11

TABLEAU B.2 – Résultat de l'algorithme EM sur le second ensemble de cas de test, contenant 20 cas différents où l'émergence moyenne sur le point le plus exposé est de 1,8 dBA. Les colonnes gt. et est. indiquent la vérité terrain du paramètre à estimer et la valeur obtenue à la fin de l'algorithme EM.

ensemble 1		ensemble 2	
σ_l		σ_l	
gt.	est.	gt.	est.
1,58	6,92	1,58	2,05
2,92	7,38	2,92	3,72
4,3	10,1	4,3	4,7
5,7	10,3	5,7	6,51
1,58	2,66	1,58	1,79
2,92	3,77	2,92	2,66
4,3	4,59	4,3	4,69
5,7	4,69	5,7	5,83
1,58	2,51	1,58	1,6
2,92	3,22	2,92	2,73
4,3	3,4	4,3	4,39
5,7	3,67	5,7	6,89
1,58	2,48	1,58	2,04
2,92	2,89	2,92	2,72
4,3	3,3	4,3	4,45
5,7	4,12	5,7	5,74
1,58	2,7	1,58	1,91
2,92	3,59	2,92	2,48
4,3	3,86	4,3	4,73
5,7	4,58	5,7	5,23

TABLEAU B.3 – Résultat de l'algorithme EM sur les deux ensembles de cas de test où on a exprimé les résultats en fonction de σ_l, le paramètre d'incertitude des $l_{i,j}$, les contributions particulières. Les colonnes gt. et est. indiquent la vérité terrain du paramètre à estimer et la valeur obtenue à la fin de l'algorithme EM.

Bibliographie

Kumar Abhishek, Sven Leyffer, & Jeff Linderoth (2010). Filmint : An outer approximation-based solver for convex mixed-integer nonlinear programs. *INFORMS Journal on computing*, 22(4), 555–567.

Shabbir Ahmed & Alexander Shapiro (2008). Solving chance-constrained stochastic programs via sampling and integer programming. *Tutorials in Operations Research*, 10, 261–270.

MC Alexiadis, PS Dokopoulos, & HS Sahsamanoglou (1999). Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion*, 14(3), 836–842.

Christophe Andrieu, Nando De Freitas, & Arnaud Doucet (1999). Sequential mcmc for bayesian model selection. In *Higher-Order Statistics, 1999. Proceedings of the IEEE Signal Processing Workshop on* (pp. 130–134). : IEEE.

Christophe Andrieu, Arnaud Doucet, & Roman Holenstein (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 72(3), 269–342.

Christophe Andrieu, Arnaud Doucet, Sumeetpal S Singh, & Vladislav B Tadic (2004). Particle methods for change detection, system identification, and control. *Proceedings of the IEEE*, 92(3), 423–438.

Ioannis P Androulakis, Costas D Maranas, & Christodoulos A Floudas (1995). αbb : A global optimization method for general constrained nonconvex problems. *Journal of Global Optimization*, 7(4), 337–363.

M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, & Tim Clapp (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2), 174–188.

D Avitzour (1995). Stochastic simulation bayesian approach to multitarget tracking. *IEE Proceedings-Radar, Sonar and Navigation*, 142(2), 41–44.

Thomas Bayes (1763). An essay towards solving a problem in the doctrine of chances.

Bradley M Bell & Frederick W Cathey (1993). The iterated kalman filter update as a gauss-newton method. *IEEE Transactions on Automatic Control*, 38(2), 294–297.

R Louis Bellaire, Edward W Kamen, & Serena M Zabin (1995). New nonlinear iterated filter with applications to target tracking. In *Signal and Data Processing of Small Targets 1995*, volume 2561 (pp. 240–252). : International Society for Optics and Photonics.

Adi Ben-Israel & Thomas NE Greville (2003). *Generalized inverses : theory and applications*, volume 15. Springer Science & Business Media.

James O Berger (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.

Niclas Bergman (1999). Recursive bayesian estimation. *Department of Electrical Engineering, Linköping University, Linköping Studies in Science and Technology. Doctoral dissertation*, 579, 11.

Dimitri P. Bertsekas (1976). *Dynamic Programming : Deterministic and Stochastic Models*. Academic Press.

Dimitri P Bertsekas (2014). *Constrained optimization and Lagrange multiplier methods*. Academic press.

Truman F Bewley (1994). *Advances in Econometrics : Volume 2 : Fifth World Congress*, volume 2. Cambridge University Press.

C Bishop (2007). *Pattern Recognition and Machine Learning*. Springer, New York, 2nd edition.

Karl Bolin, Gösta Bluhm, Gabriella Eriksson, & Mats E Nilsson (2011). Infrasound and low frequency noise from wind turbines : exposure and health effects. *Environmental research letters*, 6(3), 035103.

Pierre Bonami, Lorenz T Biegler, Andrew R Conn, Gérard Cornuéjols, Ignacio E Grossmann, Carl D Laird, Jon Lee, Andrea Lodi, François Margot, Nicolas Sawaya, & others (2008). An algorithmic framework for convex mixed integer nonlinear programs. *Discrete Optimization*, 5(2), 186–204.

Pierre Bonami, Mustafa Kilinç, & Jeff Linderoth (2012). Algorithms and software for convex mixed integer nonlinear programs. In *Mixed integer nonlinear programming* (pp. 1–39). Springer.

Pierre Bonami, Jon Lee, Sven Leyffer, & Andreas Wächter (2011). More branch-and-bound experiments in convex nonlinear integer programming. *Preprint ANL/MCS-P1949-0911, Argonne National Laboratory, Mathematics and Computer Science Division*.

J Frédéric Bonnans (1994). Local analysis of newton-type methods for variational inequalities and nonlinear programming. *Applied Mathematics and Optimization*, 29(2), 161–186.

Brian Borchers & John E. Mitchell (1994). An improved branch and bound algorithm for mixed integer nonlinear programs. *Computers & Operations Research*, 21(4), 359–367.

M Boutayeb, H Rafaralahy, & M Darouach (1997). Convergence analysis of the extended kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE transactions on automatic control*, 42(4), 581–586.

Stephen Boyd & Lieven Vandenberghe (2009). *Convex optimization*. Cambridge University Press.

Peter J Brockwell & Richard A Davis (2016). *Introduction to time series and forecasting*. springer.

Barbara G Brown, Richard W Katz, & Allan H Murphy (1984). Time series models to simulate and forecast wind speed and wind power. *Journal of climate and applied meteorology*, 23(8), 1184–1195.

Steven D Brown (1986). The kalman filter in analytical chemistry. *Analytica Chimica Acta*, 181, 1–26.

K Brusdal, Jean-Michel Brankart, G Halberstadt, Geir Evensen, Pierre Brasseur, PJ Van Leeuwen, E Dombrowsky, & Jacques Verron (2003). A demonstration of ensemble-based assimilation methods with a layered ogcm from the perspective of operational ocean forecasting systems. *Journal of Marine Systems*, 40, 253–289.

Samuel Burer & Adam N. Letchford (2012). Non-convex mixed-integer nonlinear programming : a survey. *Surveys in Operations Research and Management Science*, 17(2), 97–106.

Karen C Chou & Ross B Corotis (1981). Simulation of hourly wind speed and array wind power. *Solar Energy*, 26(3), 199–212.

Girish Chowdhary & Ravindra Jategaonkar (2010). Aerodynamic parameter estimation from flight data applying extended and unscented kalman filter. *Aerospace science and technology*, 14(2), 106–117.

Thomas F Coleman & Yuying Li (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization*, 6(2), 418–445.

John D Cook (2011). Basic properties of the soft maximum.

Alexandre Costa, Antonio Crespo, Jorge Navarro, Gil Lizcano, Henrik Madsen, & Everaldo Feitosa (2008). A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6), 1725–1744.

Dan Crisan & Arnaud Doucet (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on signal processing*, 50(3), 736–746.

Wade T Crow & Eric F Wood (2003). The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble kalman filtering : A case study based on estar measurements during sgp97. *Advances in Water Resources*, 26(2), 137–149.

Robert J Dakin (1965). A tree-search algorithm for mixed integer programming problems. *The computer journal*, 8(3), 250–255.

Claudia D'Ambrosio & Andrea Lodi (2011). Mixed integer nonlinear programming tools : a practical overview. *4OR*, 9(4), 329–349.

AR Daniel & AA Chen (1991). Stochastic simulation and forecasting of hourly average wind speed sequences in jamaica. *Solar energy*, 46(1), 1–11.

George B Dantzig, Alex Orden, Philip Wolfe, & others (1955). The generalized simplex method for minimizing a linear form under linear inequality restraints. *Pacific Journal of Mathematics*, 5(2), 183–195.

Antoine Deleforge, Florence Forbes, & Radu Horaud (2015). High-dimensional regression with gaussian mixtures and partially-latent response variables. *Statistics and Computing*, 25(5), 893–911.

V.S. Donchenko & N.F. Kirichenko (2003). Generalized inverse in control with constraints. *Cybernetics and Systems Analysis*, 39(6), 854–861.

Randal Douc & Olivier Cappé (2005). Comparison of resampling schemes for particle filtering. In *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on* (pp. 64–69). : IEEE.

Arnaud Doucet, Simon Godsill, & Christophe Andrieu (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3), 197–208.

Arnaud Doucet, Neil J Gordon, & Vikram Krishnamurthy (2001). Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on signal processing*, 49(3), 613–624.

Arnaud Doucet & Vladislav B Tadić (2003). Parameter estimation in general state-space models using particle methods. *Annals of the institute of Statistical Mathematics*, 55(2), 409–422.

Marco A Duran & Ignacio E Grossmann (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical programming*, 36(3), 307–339.

Jiancheng Fang & Xiaolin Gong (2010). Predictive iterated kalman filter for ins/gps integration and its application to sar motion compensation. *IEEE transactions on instrumentation and measurement*, 59(4), 909–915.

Ya Ping Fang, Kaiwen Meng, & Xiao Qi Yang (2012). Piecewise linear multicriteria programs : the continuous case and its discontinuous generalization. *Operations research*, 60(2), 398–409.

Paul Fearnhead (2002). Markov chain monte carlo, sufficient statistics, and particle filters. *Journal of Computational and Graphical Statistics*, 11(4), 848–862.

O Fégeant (1999). On the masking of wind turbine noise by ambient noise. In *EWEC-CONFERENCE-* (pp. 184–188).

Mark John Francis Gales (1995). *Model-based techniques for noise robust speech recognition.* PhD thesis, University of Cambridge Cambridge.

Arthur M Geoffrion (1972). Generalized benders decomposition. *Journal of optimization theory and applications*, 10(4), 237–260.

Zoubin Ghahramani & Geoffrey E Hinton (1996). *Parameter estimation for linear dynamical systems.* Technical report, Technical Report CRG-TR-96-2, University of Totronto, Dept. of Computer Science.

Zoubin Ghahramani & Sam T Roweis (1999). Learning nonlinear dynamical systems using an em algorithm. In *Advances in neural information processing systems* (pp. 431–437).

Esmaeil Ghahremani & Innocent Kamwa (2011). Online state estimation of a synchronous generator using unscented kalman filter from phasor measurements units. *IEEE Transactions on Energy Conversion*, 26(4), 1099–1108.

Gregor Giebel, Richard Brownsword, George Kariniotakis, Michael Denhard, & Caroline Draxl (2011). *The state-of-the-art in short-term prediction of wind power : A literature overview.* Technical report, ANEMOS. plus.

Donald Goldfarb & Leon Lapidus (1968). Conjugate gradient method for nonlinear programming problems with linear constraints. *Industrial & Engineering Chemistry Fundamentals*, 7(1), 142–151.

Alan A Goldstein (1964). Convex programming in hilbert space. *Bulletin of the American Mathematical Society*, 70(5), 709–710.

Neil J Gordon, David J Salmond, & Adrian FM Smith (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140 (pp. 107–113). : IET.

Yaqing Gu, Dean S Oliver, & others (2007). An iterative ensemble kalman filter for multiphase fluid flow data assimilation. *Spe Journal*, 12(04), 438–446.

Omprakash K Gupta & A Ravindran (1985). Branch and bound experiments in convex nonlinear integer programming. *Management science*, 31(12), 1533–1546.

Edward J Hannan & Jorma Rissanen (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, 69(1), 81–94.

Colin H Hansen & CH Sehrndt (2001). Fundamentals of acoustics. *Occupational Exposure to Noise : Evaluation, Prevention and Control. World Health Organization.*

Wolfgang Härdle (2012). *Smoothing techniques : with implementation in S.* Springer Science & Business Media.

Simon S Haykin & others (2001). *Kalman filtering and neural networks.* Wiley Online Library.

Tomoyuki Higuchi (1997). Monte carlo filter using the genetic algorithm operators. *Journal of Statistical Computation and Simulation*, 59(1), 1–23.

Dorit S. Hochbaum (1995). A nonlinear knapsack problem. *Operations Research Letters*, 17(3), 103–110.

David R Hunter & Kenneth Lange (2004). A tutorial on mm algorithms. *The American Statistician*, 58(1), 30–37.

Kazufumi Ito & Kaiqi Xiong (2000). Gaussian filters for nonlinear filtering problems. *IEEE transactions on automatic control*, 45(5), 910–927.

Simon J Julier (1998). A skewed approach to filtering. In *Proc. SPIE*, volume 3373 (pp. 271–282).

Simon J Julier (2002). The scaled unscented transformation. In *American Control Conference, 2002. Proceedings of the 2002*, volume 6 (pp. 4555–4559). : IEEE.

Simon J Julier & Jeffrey K Uhlmann (1997). New extension of the Kalman filter to nonlinear systems. In *Proc. AeroSense* (pp. 182–193).

Thomas Kailath, Ali H Sayed, & Babak Hassibi (2000). *Linear estimation*, volume 1. Prentice Hall.

Rudolph Emil Kalman & others (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1), 35–45.

Lalarukh Kamal & Yasmin Zahra Jafri (1997). Time series models to simulate and forecast hourly averaged wind speed in quetta, pakistan. *Solar Energy*, 61(1), 23–32.

Rambabu Kandepu, Bjarne Foss, & Lars Imsland (2008). Applying the unscented kalman filter for nonlinear state estimation. *Journal of process control*, 18(7-8), 753–768.

Nicholas Kantas, Arnaud Doucet, Sumeetpal Sindhu Singh, & Jan Marian Maciejowski (2009). An overview of sequential monte carlo methods for parameter estimation in general state-space models. *IFAC Proceedings Volumes*, 42(10), 774–785.

GN Kariniotakis, GS Stavrakakis, & EF Nogaret (1996). Wind power forecasting using advanced neural networks models. *IEEE transactions on Energy conversion*, 11(4), 762–767.

Rajesh G. Kavasseri & Krithika Seetharaman (2009). Day-ahead wind speed forecasting using f-arima models. *Renewable Energy*, 34(5), 1388–1393.

Christian L Keppenne & Michele M Rienecker (2003). Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble kalman filter. *Journal of Marine Systems*, 40, 363–380.

Padmanaban Kesavan & Paul I Barton (2000). Generalized branch-and-cut framework for mixed-integer nonlinear optimization problems. *Computers & Chemical Engineering*, 24(2-7), 1361–1366.

Genshiro Kitagawa (1996). Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of computational and graphical statistics*, 5(1), 1–25.

Genshiro Kitagawa (1998). A self-organizing state-space model. *Journal of the American Statistical Association*, (pp. 1203–1215).

Robert Kohn & Craig F Ansley (1986). Estimation, prediction, and interpolation for arima models with missing data. *Journal of the American Statistical Association*, 81(395), 751–761.

Kenneth Lange (2004). Elementary optimization. In *Optimization* (pp. 1–17). Springer.

Conny Larsson & Sven Israelsson (1991). Effects of meteorological conditions and source height on sound propagation near the ground. *Applied Acoustics*, 33(2), 109–121.

Conny Larsson & Olof Öhlund (2014). Wind turbine sound-metric and guidelines. In *43rd International Congress on Noise Control Engineering, Melbourne, Australia, November 16-19, 2014*.

Charles L Lawson & Richard J Hanson (1995). *Solving least squares problems*, volume 15. Siam.

Evgeny S Levitin & Boris T Polyak (1966). Constrained minimization methods. *USSR Computational mathematics and mathematical physics*, 6(5), 1–50.

Pu Li, Harvey Arellano-Garcia, & Günter Wozny (2008). Chance constrained programming approach to process optimization under uncertainty. *Computers & Chemical Engineering*, 32(1), 25–45.

Pu Li, Moritz Wendt, & GüNter Wozny (2002). A probabilistically constrained model predictive controller. *Automatica*, 38(7), 1171–1176.

Pu Li, Moritz Wendt, & Günter Wozny (2004). Optimal production planning for chemical processes under uncertain market conditions. *Chemical engineering & technology*, 27(6), 641–651.

Jun S Liu & Rong Chen (1998). Sequential monte carlo methods for dynamic systems. *Journal of the American statistical association*, 93(443), 1032–1044.

Lennart Ljung (1979). Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1), 36–50.

Lennart Ljung (1998). System identification. In *Signal analysis and prediction* (pp. 163–173). Springer.

Wolfgang Mader, Yannick Linke, Malenka Mader, Linda Sommerlade, Jens Timmer, & Björn Schelter (2014). A numerically efficient implementation of the expectation maximization algorithm for state space models. *Applied Mathematics and Computation*, 241, 222–232.

Kamlesh Mathur, Harvey M. Salkin, & Susumu Morito (1983). A branch and search algorithm for a class of nonlinear knapsack problems. *Operations Research Letters*, 2(4), 155–160.

Nicholas Metropolis & Stanislaw Ulam (1949). The monte carlo method. *Journal of the American statistical association*, 44(247), 335–341.

Mohamed A Mohandes, Shafiqur Rehman, & Talal O Halawani (1998). A neural networks approach for wind speed prediction. *Renewable Energy*, 13(3), 345–354.

Jorge J Moré (1978). The levenberg-marquardt algorithm : implementation and theory. In *Numerical analysis* (pp. 105–116). Springer.

Hiroyuki Moriyama, Nobuo Yamashita, & Masao Fukushima (2003). The incremental gauss-newton algorithm with adaptive stepsize rule. *Computational optimization and Applications*, 26(2), 107–141.

L-J Natvik & Geir Evensen (2003). Assimilation of ocean colour data into a biochemical model of the north atlantic : Part 1. data assimilation experiments. *Journal of Marine Systems*, 40, 127–153.

Magnus NøRgaard, Niels K Poulsen, & Ole Ravn (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36(11), 1627–1638.

NF Norme (1996). S 31-010" acoustique-caractérisation et mesurage des bruits de l'environnement-méthodes particulières de mesurage.

Marcus D Odom & Ramesh Sharda (1990). A neural network model for bankruptcy prediction. In *Neural Networks, 1990., 1990 IJCNN International Joint Conference on* (pp. 163–168). : IEEE.

Manfred Padberg (2013). *Linear optimization and extensions*, volume 12. Springer Science & Business Media.

BK Pagnoncelli, Shapiro Ahmed, & Alexander Shapiro (2009). Sample average approximation method for chance constrained programming : theory and applications. *Journal of optimization theory and applications*, 142(2), 399–416.

K Paliwal & Anjan Basu (1987). A speech enhancement method based on kalman filtering. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87.*, volume 12 (pp. 177–180). : IEEE.

Eja Pedersen & Högskolan I Halmstad (2003). *Noise annoyance from wind turbines : a review.* Naturvårdsverket.

Ksenia Ponomareva, Paresh Date, & Zidong Wang (2010). A new unscented kalman filter with higher order moment-matching. *Proceedings of Mathematical Theory of Networks and Systems (MTNS 2010), Budapest.*

George Poyiadjis, Arnaud Doucet, & Sumeetpal S Singh (2011). Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. *Biometrika*, 98(1), 65–80.

Ignacio Quesada & Ignacio E Grossmann (1992). An lp/nlp based branch and bound algorithm for convex minlp optimization problems. *Computers & chemical engineering*, 16(10-11), 937–947.

Anthony L Rogers, James F Manwell, & Sally Wright (2006). Wind turbine acoustic noise. *Renewable Energy Research Laboratory, Amherst : University of Massachusetts.*

Donald B Rubin (1987). The calculation of posterior distributions by data augmentation : Comment : A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest : The sir algorithm. *Journal of the American Statistical Association*, 82(398), 543–546.

Hong S Ryoo & Nikolaos V Sahinidis (1996). A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8(2), 107–138.

Michel Sakarovitch (1984). *Optimisation combinatoire : Programmation discrète*, volume 2. Editions Hermann.

Jeremy Schild & Vincent Chavand (2015). State of the art and new perspectives for the development of noise regulation of wind farms. *Proc. 6th International Meeting on Wind Turbine Noise.*

Stanley F Schmidt (1981). Kalman filter : Its recognition and development for aerospace applications. *J. Guid. and Contr.*, 4(1), 4–7.

M. Schwartz & M. Milligan (2002). Statistical wind forecasting at the us national renewable energy laboratory. In *Proceedings of the First IEA Joint Action Symposium on Wind Forecasting Techniques, Norrköping, Sweden* (pp. 115–124B).

Eun-Hwan Shin & Naser El-Sheimy (2004). An unscented kalman filter for in-motion alignment of low-cost imus. In *Position Location and Navigation Symposium, 2004. PLANS 2004* (pp. 273–279). : IEEE.

Gabe Sibley, Gaurav S Sukhatme, & Larry Matthies (2006). The iterated sigma point Kalman filter with applications to long range stereo. In *Robotics : Science and Systems*, volume 8 (pp. 235–244).

Martin A Skoglund, Gustaf Hendeby, & Daniel Axehill (2015). Extended kalman filter modifications based on an optimization view point. In *Information Fusion (Fusion), 2015 18th International Conference on* (pp. 1856–1861). : IEEE.

AFM Smith & Mike West (1983). Monitoring renal transplants : an application of the multiprocess kalman filter. *Biometrics*, (pp. 867–878).

Edward MB Smith & Constantinos C Pantelides (1999). A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex minlps. *Computers & Chemical Engineering*, 23(4-5), 457–478.

William J Stewart (1994). *Introduction to the numerical solution of Markov chains*. Princeton University Press.

Zaiyong Tang, Chrys de Almeida, & Paul A Fishwick (1991). Time series forecasting using neural networks vs. box-jenkins methodology. *Simulation*, 57(5), 303–310.

Jacques Teghem (2012). La recherche opérationnelle tome 1 : Les méthodes d'optimisation.

Dung T Tran, Emmanuel Vincent, & Denis Jouvet (2015). Nonparametric uncertainty estimation and propagation for noise robust asr. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(11), 1835–1846.

Aditya Tulsyan, Biao Huang, R Bhushan Gopaluni, & J Fraser Forbes (2013). On simultaneous on-line state and parameter estimation in non-linear state-space models. *Journal of Process Control*, 23(4), 516–526.

Rudolph Van Der Merwe (2004). *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. PhD thesis, Oregon Health & Science University.

Rudolph Van Der Merwe & Eric A Wan (2001). The square-root unscented kalman filter for state and parameter-estimation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, volume 6 (pp. 3461–3464). : IEEE.

Ranjan Vepa (2011). Nonlinear, optimal control of a wind turbine generator. *IEEE Transactions on Energy Conversion*, 26(2), 468–478.

Emmanuel Vincent, Rémi Gribonval, & Cédric Févotte (2006). Performance measurement in blind audio source separation. *IEEE transactions on audio, speech, and language processing*, 14(4), 1462–1469.

Eric A Wan & Ronell Van Der Merwe (2000). The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000* (pp. 153–158). : Ieee.

Larry Wasserman (2013). *All of statistics : a concise course in statistical inference*. Springer Science & Business Media.

Robert WM Wedderburn (1974). Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika*, 61(3), 439–447.

Moritz Wendt, Pu Li, & Günter Wozny (2002). Nonlinear chance-constrained process optimization under uncertainty. *Industrial & engineering chemistry research*, 41(15), 3621–3629.

Tapio Westerlund & Frank Pettersson (1995). An extended cutting plane method for solving convex minlp problems. *Computers & Chemical Engineering*, 19, 131–136.

Margaret Wright (2005). The interior-point revolution in optimization : history, recent developments, and lasting consequences. *Bulletin of the American mathematical society*, 42(1), 39–56.

Meiliang Wu & Andrew W Smyth (2007). Application of the unscented kalman filter for real-time nonlinear structural system identification. *Structural Control and Health Monitoring*, 14(7), 971–990.

Lei Xu & Michael I Jordan (1996). On convergence properties of the em algorithm for gaussian mixtures. *Neural computation*, 8(1), 129–151.

Pifu Zhang, Jason Gu, Evangelos E Milios, & Peter Huynh (2005). Navigation with imu/gps/digital compass with unscented kalman filter. In *Mechatronics and Automation, 2005 IEEE International Conference*, volume 3 (pp. 1497–1502). : IEEE.

Résumé

Actuellement, la construction d'un parc éolien nécessite une étude acoustique qui doit assurer la tranquillité des habitants aux alentours et la conformité au regard de la réglementation en vigueur. Pour ce faire, des mesures acoustiques sont réalisées sur une période d'environ deux semaines. Durant ces mesures, des cycles de marche et d'arrêt des machines sont réalisés afin de mesurer la différence de niveau sonore entre le bruit ambiant (éoliennes en fonctionnement) et le bruit résiduel (éoliennes à l'arrêt). Un plan de bridage des machines est alors calculé et fourni à l'exploitant afin de l'implémenter dans le système de contrôle local des éoliennes (SCADA). Actuellement, ce plan dépend grossièrement des conditions météorologiques et des périodes de la journée, supposées corrélées aux conditions acoustiques. En pratique, cette manière de procéder engendre fréquemment des dépassements du critère réglementaire et/ou des pertes de production électrique. Ceci est dû aux conditions acoustiques qui évoluent sans cesse, à la fois pour le bruit particulier (bruit des éoliennes seules) qui dépend finement des conditions météorologiques, et pour le bruit résiduel qui dépend de toutes les autres sources de l'environnement et qui est fondamentalement de nature stochastique.

La thèse vise à proposer un algorithme de contrôle du parc éolien en temps réel basé sur un nouveau paradigme de contrôle. On y étudie la possibilité de contrôler un parc éolien à partir d'un système en boite noire d'estimation temps-réel du niveau résiduel et du niveau particulier par séparation de sources. Dans le manuscrit, on définit tout d'abord une formulation du problème dans le cadre du contrôle en identifiant les problématiques propres à ce sujet, une définition des variables du problème et en se rattachant à l'état de l'art du contrôle. Ensuite, on propose deux solutions complètes de contrôle et une évaluation expérimentale. La première est une solution déterministe, qui s'appuie sur un algorithme d'optimisation combinatoire sous contrainte, et qui s'inspire du contrôle actuel des parcs éoliens tout en tenant compte de l'estimation par séparation de sources, alors supposée exacte. On y propose en outre une étude de la capacité du système déterministe à satisfaire le critère réglementaire français qui est aujourd'hui calculé à l'aide de médianes temporelles des variables acoustiques. La seconde est une solution stochastique, qui est basée sur une représentation d'état des variables acoustiques et des incertitudes gaussiennes. Elle inclut un filtrage de Kalman non-linéaire, afin de fusionner l'incertitude sur le modèle acoustique et l'incertitude de séparation de sources, un algorithme espérance-maximisation afin de ré-estimer les incertitudes du problème qui varient d'un parc à un autre, et une adaptation robuste de l'algorithme combinatoire afin de prendre en compte les incertitudes estimées.

Mots-clés: éoliennes, commande optimale, filtrage de Kalman.

Abstract

Currently, acoustic studies are required to set wind farms up. They must ensure the tranquility of the inhabitants around the farms in accordance with current regulations. For this purpose, acoustic measurements are made during a couple of weeks. When measuring, the wind turbines are periodically stopped in order to evaluate the difference between ambient noise levels (with the turbines on) and residual noise levels (with the turbines off). A curtailment plan is then computed and sent to the wind farm owner in order to set it up in the local turbine control

system (SCADA). Currently, the curtailment plan roughly depends on the weather conditions and the time of the day which are allegedly correlated to the acoustic variables. In practice, it frequently leads to violations of the acoustic constraints or electrical power loss. This is because the acoustic conditions constantly and strongly evolve over time : the wind turbine noise level finely depends on the weather conditions and the residual noise level depends on all the other acoustic sources and has therefore a stochastic nature.

The goal of the thesis is to design a principled real-time control algorithm for wind farms. To do so, we investigate the use of a black-box source separation system that estimates the residual noise level and the wind turbine noise level. We first provide a theoretical formulation of the problem by accounting for specific practical issues, by defining the variables of the problem and by binding these issues to the state of the art. Then, we propose two complete control solutions and run an experimental evaluation. The first solution is a deterministic algorithm based on a constrained combinatorial optimization algorithm, which is inspired by the current approach for controlling wind farms while exploiting the source separation system. Moreover, we present a study of its ability to fulfill the French acoustic constraints that are computed as temporal medians of the acoustic variables. The second solution is stochastic and based on a state-space model defined by means of Gaussian uncertainties. It features a nonlinear Kalman filter in order to fuse the uncertainties of the model and of the source separation system, an Expectation-Maximization algorithm that computes the uncertainties for a specific farm, and a robust variant of the deterministic algorithm that takes the estimated uncertainties into account when computing the optimal command.

Keywords: wind turbines, optimal command, Kalman filtering.