



**HAL**  
open science

# Mouvement actif pour la localisation binaurale de sources sonores en robotique

Gabriel Bustamante

► **To cite this version:**

Gabriel Bustamante. Mouvement actif pour la localisation binaurale de sources sonores en robotique. Automatique. Université Paul Sabatier - Toulouse III, 2017. Français. NNT : 2017TOU30211 . tel-01681138v2

**HAL Id: tel-01681138**

**<https://laas.hal.science/tel-01681138v2>**

Submitted on 12 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Fédérale



Toulouse Midi-Pyrénées

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ FÉDÉRALE TOULOUSE MIDI-PYRÉNÉES

Délivré par :

*l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

---

---

Présentée et soutenue le *05/09/2017* par :

**GABRIEL BUSTAMANTE**

**Mouvement Actif pour la Localisation Binaurale  
de Sources Sonores en Robotique**

---

---

### JURY

LAURENT GIRIN	Professeur d'Université	GIPSA-lab
EMMANUEL VINCENT	Directeur de Recherche	INRIA Nancy
CHRISTINE EVERS	Research associate	Imperial College
YANNICK DEVILLE	Professeur d'Université	IRAP
RADU HORAUD	Directeur de Recherche	INRIA Grenoble
IVAN PETROVIC	Full Professor	University of Zagreb
ALEXANDER RAAKE	Full Professor	TU Ilmenau
PATRICK DANÈS	Professeur d'Université	LAAS-CNRS

---

**École doctorale et spécialité :**

*EDSYS : Automatique 4200046*

**Unité de Recherche :**

*Laboratoire d'analyse et d'architecture des systèmes du CNRS*

**Directeur de Thèse :**

*Patrick DANÈS*

**Rapporteurs :**

*Emmanuel VINCENT et Laurent GIRIN*



# Remerciements

Je souhaiterais dans un premier temps remercier Jean Arlat et Liviu Nicu, directeurs successifs du LAAS-CNRS durant mes années de doctorat, de m'avoir accueilli au sein de leur laboratoire.

Je tiens particulièrement à remercier Patrick Danès, professeur à l'Université Paul Sabatier et directeur de ma thèse, pour son accompagnement, ses conseils et son soutien. Merci de m'avoir écouté et encouragé dans les moments difficiles.

Je remercie ensuite Laurent Girin, professeur à l'Institut Polytechnique de Grenoble et Emmanuel Vincent, directeur de recherche à l'INRIA Nancy, d'avoir accepté de rapporter ma thèse. Merci aussi pour vos remarques constructives qui m'ont aidé à préparer la soutenance.

Je remercie également Christine Evers, chercheur au London Imperial College, Ivan Petrović professeur à l'Université de Zagreb, Alexander Raake professeur au TU Ilmenau, Radu Horaud, directeur de recherche à l'INRIA Grenoble, et enfin Yannick Deville, professeur à l'Université Paul Sabatier, d'avoir accepté d'examiner cette thèse.

Merci également aux membres de l'équipe RAP qui m'ont apporté leur aide durant ces années de thèse, et merci à mes collègues de bureau pour votre bonne humeur.

Enfin je souhaite remercier très chaleureusement ma famille et mes amis pour leurs encouragements tout au long de ma thèse. Plus particulièrement, merci Hélène pour ton soutien, tes encouragements quotidiens, ta patience et ton amour.



# Table des figures

1.1	Représentation planaire du problème . . . . .	16
1.2	Schéma de localisation en trois étapes . . . . .	20
1.3	Exemple de localisation à différents instants de simulation . . . . .	23
1.4	Plateforme robotique et vue schématique de son architecture logicielle	26
3.1	Espace admissible de mouvements . . . . .	58
3.2	Combinaison du mouvement et de l'observation . . . . .	59
3.3	Lignes de niveau et gradient . . . . .	63
3.4	Lignes de niveau de la fonction $F_1$ (modèle ITD) . . . . .	64
3.5	Lignes de niveau de la fonction $\tilde{F}_1$ (modèle ITD) . . . . .	65
3.6	Lignes de niveau des fonctions $F_1$ et $\tilde{F}_1$ (modèle ITD – situation particulière) . . . . .	66
3.7	Lignes de niveau de la fonction $F_1$ (modèle idéal) . . . . .	67
3.8	Lignes de niveau de la fonction $\tilde{F}_1$ (modèle idéal) . . . . .	68
3.9	Itérations de l'algorithme du gradient projeté . . . . .	74
3.10	Résultats de simulation . . . . .	82
3.11	Entropie de l'approximation moment-matched de la densité de filtrage en conditions réelles . . . . .	83
3.12	Expérience en conditions réelles (translation) . . . . .	83
3.13	Expérience en conditions réelles (mouvement circulaire) . . . . .	84
3.14	Expérience en conditions réelles (mouvement actif $\mathcal{P}_1$ ) . . . . .	84
4.1	Stratégies de mouvement actif avec différents horizons d'exploration $N = \{1, 2, 3, 5\}$ . . . . .	95
4.2	Distribution de l'entropie de la densité de filtrage pour un grand nombre d'expériences . . . . .	97
5.1	Localisation active d'une source avec un objectif de positionnement par rapport à une cible . . . . .	106



# Liste des Algorithmes

2.1	Évaluation de $F_1$ . . . . .	50
2.2	Évaluation de $F_i$ . . . . .	52
3.1	Gradient projeté pour la résolution de $\mathcal{P}_1$ . . . . .	78
4.1	Calcul des moments de $p(z_{k+1:k+i-1} z_{1:k})$ . . . . .	88
4.2	Exemple de programmation par différentiation automatique . . . . .	92
4.3	Gradient projeté pour la résolution de $\mathcal{P}_N$ . . . . .	94
A.1	Étape de prédiction de l'UKF . . . . .	114
A.2	Étape de mise à jour de l'UKF . . . . .	115



# Table des matières

<b>Introduction</b>	<b>1</b>
Perception et actions en robotique . . . . .	1
L’audition en robotique . . . . .	2
Organisation de la thèse . . . . .	3
<b>1 Localisation binaurale et mouvement actif</b>	<b>5</b>
1.1 Indices auditifs et modèles pour la localisation binaurale . . . . .	6
1.1.1 Indices binauraux et monauraux . . . . .	6
1.1.2 Modèle HRTF . . . . .	8
1.2 État de l’art de la localisation binaurale en robotique . . . . .	8
1.2.1 Localisation de sources à partir des indices binauraux . . . . .	9
1.2.2 Mouvements en boucle ouverte . . . . .	10
1.2.3 Génération de mouvements actifs . . . . .	12
1.3 Vers une localisation binaurale active en trois étapes . . . . .	14
1.3.1 Introduction . . . . .	15
A Notations . . . . .	15
B Modèle géométrique . . . . .	15
1.3.2 Estimation à court terme d’azimuts . . . . .	17
1.3.3 Localisation audio-motrice . . . . .	18
1.3.4 Position du problème de synthèse de mouvement actif . . . . .	20
1.4 Présentation des outils d’évaluation . . . . .	22
1.4.1 Simulateur binaural . . . . .	22
1.4.2 Localisation à partir d’un robot mobile . . . . .	24
<b>2 Outils théoriques pour la commande référencée information</b>	<b>27</b>
2.1 Définitions de l’information . . . . .	28
2.1.1 Estimation et mesure d’information . . . . .	28
2.1.2 Entropie et Information Mutuelle . . . . .	30
2.1.3 Entropie et filtrage bayésien . . . . .	32
2.2 Perception active en robotique . . . . .	35
2.2.1 Définition d’un modèle . . . . .	36
2.2.2 Choix du critère d’information . . . . .	37
2.2.3 Décisions à court terme contre décisions à long terme . . . . .	42
2.3 Fonction d’incertitude pour la localisation active de sources sonores . . . . .	45
2.3.1 Hypothèses du modèle d’exploration . . . . .	45
2.3.2 Définition d’une fonction d’incertitude . . . . .	47
A Récompense immédiate . . . . .	48
B Critère de décision à long terme . . . . .	50

<b>3</b>	<b>Génération de mouvement guidé par un critère de décision à court terme</b>	<b>55</b>
3.1	Définition du problème d'optimisation . . . . .	56
3.2	Intuitions sur le mouvement . . . . .	58
3.2.1	Mouvement guidé par les ITD . . . . .	58
3.2.2	Mouvement guidé par des mesures linéaires en azimuth . . . . .	61
3.3	Lignes de niveau . . . . .	62
3.3.1	Exploration guidée par les ITD . . . . .	63
3.3.2	Cas idéal de l'exploration guidée par des azimuths . . . . .	66
3.3.3	Discussions . . . . .	66
3.4	Gradient de la fonction de récompense . . . . .	69
3.4.1	Définition de $F_1$ . . . . .	70
3.4.2	Évaluation du gradient en un point . . . . .	71
3.5	Optimisation sous contraintes . . . . .	72
3.5.1	Conditions d'optimalité . . . . .	72
3.5.2	Résolution numérique . . . . .	74
3.6	Évaluation de la méthode . . . . .	77
3.6.1	Simulation . . . . .	77
3.6.2	Expériences en conditions réelles . . . . .	80
<b>4</b>	<b>Génération de mouvement avec critère de décision à long terme</b>	<b>85</b>
4.1	Position du problème . . . . .	85
4.1.1	Évaluation de $J_N$ . . . . .	86
4.1.2	Définition du gradient de $J_N$ . . . . .	87
4.2	Approches numériques . . . . .	89
4.2.1	Différentiation automatique . . . . .	89
4.2.2	Résolution de $\mathcal{P}_N$ par la méthode du gradient projeté . . . . .	93
4.3	Analyse de résultats de simulation . . . . .	93
4.3.1	Comparaison entre différents horizons d'exploration . . . . .	95
4.3.2	Évaluations statistiques . . . . .	96
<b>5</b>	<b>Prospectives pour la localisation binaurale à mouvement actif</b>	<b>99</b>
5.1	Prospectives court terme . . . . .	99
5.1.1	Comparaison exhaustive avec une simplification rencontrée dans la littérature . . . . .	99
5.1.2	Prise en compte du bruit de dynamique . . . . .	100
5.1.3	Entropie d'un mélange de gaussiennes . . . . .	101
5.2	Évolution vers une stratégie multiobjectifs . . . . .	102
	<b>Conclusion</b>	<b>107</b>
<b>A</b>	<b>Le filtre de Kalman unscented</b>	<b>111</b>
A.1	La transformée unscented . . . . .	111
A.2	Implémentation de l'UKF . . . . .	113

---

<b>B Optimisation statique</b>	<b>117</b>
B.1 Définitions . . . . .	117
B.2 Optimisation sans contrainte . . . . .	117
B.3 Optimisation sous contraintes . . . . .	118
<b>C Acronymes</b>	<b>121</b>
<b>Bibliographie</b>	<b>123</b>



# Introduction

L'idée qui consiste à créer des machines autonomes ou « robots » est très ancienne puisque des références à des automates créés par l'Homme se retrouvent dans des œuvres qui remontent à l'Antiquité. Aujourd'hui les robots continuent de fasciner et nourrissent certains ouvrages de science-fiction. Le terme « robotique », qui désigne la discipline vaste regroupant la conception, la réalisation et l'exploitation des robots, a été introduit dans les œuvres de l'écrivain Isaac Azimov<sup>1</sup>. La robotique est aujourd'hui une discipline scientifique reconnue, dont les applications touchent des domaines variés. Si les robots ont depuis longtemps été utilisés dans l'industrie sur les chaînes de montage, la robotique intéresse aujourd'hui les acteurs du transport (apparition de robots dans les gares, véhicules autonomes, etc.) et trouve également des applications dans le milieu médical (robots chirurgiens) ou militaire. Enfin, très récemment, le marché des « robots personnels » se développe également. Dans tous les cas, les robots doivent être en interaction avec leur environnement et/ou les humains.

## Perception et action en robotique

Contrairement à certaines croyances, à l'heure actuelle les robots sont encore largement inadaptés à notre environnement et sont loin d'être entièrement autonomes. En effet, les robots sont conçus pour effectuer des tâches dans un contexte spécifique et en environnement contrôlé. En dehors de ce contexte, la ou les tâches pour lesquelles ils ont été programmés deviennent impossibles à réaliser. La capacité de perception des robots est une condition nécessaire à l'amélioration de leur comportement dans un environnement non (ou partiellement) contrôlé, de leur autonomie ainsi que de leur capacité d'adaptation. Ainsi, par l'intermédiaire de capteurs, un robot doit pouvoir recueillir et traiter un maximum d'informations sur : (1) son état interne (position/vitesse des moteurs, odométrie, etc.) ; il s'agit alors de proprioception ; (2) son environnement (nature et position d'obstacles, nature et position de sources sonores, présence et position d'humains, etc.) ; il s'agit dans ce cas d'extéroception. Sur la base de ces informations, il doit, selon le contexte, être capable d'agir en conséquence. Les premières lois de commande basées sur des capteurs extéroceptifs (lasers, caméras, microphones, etc.) ont permis aux robots d'interagir avec l'environnement pour qu'ils puissent réaliser des tâches telles que le positionnement et le suivi de cible [Weiss *et al.* 1987, Espiau *et al.* 1992], ou encore la navigation réactive [Cherubini & Chaumette 2012]. La perception active est un cas particulier de la thématique de contrôle basé sur des capteurs extéroceptifs. Au lieu de stabiliser des quantités d'intérêt (par exemple la perception d'une cible dans

---

1. Le mot « robot » apparaît pour la première fois dans une pièce de théâtre de Karel Čapek au début du XX<sup>e</sup> siècle.

le plan image d'une caméra), l'objectif de la perception active est de commander le robot ou de configurer les paramètres de ses capteurs de manière à améliorer la qualité de la perception [Aloimonos *et al.* 1988].

De manière plus générale, le processus central à l'autonomie et à l'adaptabilité d'un robot est la boucle perception-décision-action. Certaines actions sont dites « délibérées » [Ingrand & Ghallab 2014], car établies à partir de processus décisionnels de haut-niveau. Les raisonnements sous-jacents impliquent une description symbolique de la situation du robot et un objectif de mission défini sur le long terme (entrer dans une pièce, transporter un objet de A vers B, le donner à un humain, etc.). D'autres actions, « câblées », (ou « réflexes », à l'instar des réflexes humains), sont directement déclenchées par des stimuli sensoriels, sans l'intervention d'un processus cognitif de haut niveau. Elles sont effectuées avec une intelligence bas-niveau, reposent sur les signaux perçus bruts et sont définies sur des échelles de temps plus courtes. On retrouve dans celles-ci les commandes basées sur des capteurs extéroceptifs mentionnées plus haut, à des fins de stabilisation ou d'exploration.

## L'audition en robotique

Parmi toutes nos facultés sensorielles, l'audition et la vision figurent très certainement parmi les plus essentielles. Même si la vision a sans doute été la modalité sensorielle la plus étudiée en robotique, elle peut ne pas suffire dans de nombreuses situations : conditions d'éclairage mauvaises ou changeantes, occultations, champ de perception des caméras limités, etc. Pour pallier ces problèmes, l'audition en robotique se développe [Okuno & Nakadai 2015]. Cette thématique ouvre de nouvelles perspectives, en particulier pour améliorer l'interaction entre les robots et les humains. Les premiers développements de l'audition en robotique ont pu bénéficier des travaux existants en analyse de scènes auditives ([Computational Auditory Scene Analysis \(CASA\)](#)) [Rosenthal & Okuno 1998] dont les objectifs sont d'analyser et de comprendre un mélange sonore pouvant être constitué de parole, de bruit, ou encore de musique, et ce dans différents environnements acoustiques. Les principales thématiques selon l'approche [CASA](#) sont la localisation de sources, la séparation de flux audio et l'ensemble des prétraitements qui vont ouvrir la voie à la reconnaissance automatique de parole [Okuno *et al.* 2004]. Les premières contributions se sont focalisées sur l'analyse de scènes auditives statiques (sources et récepteurs immobiles) par un robot immobile selon ce paradigme. Par la suite, des approches « actives » exploitant la mobilité des robots qui embarquent les capteurs audio, se sont développées [Cooke *et al.* 2007] dans le but d'améliorer les résultats.

Certaines approches en audition tirent parti de la redondance d'information apportée par une antenne de plusieurs microphones. Au contraire, l'approche binaurale étudiée dans cette thèse repose sur l'exploitation du flux audio perçu par deux microphones uniquement. Les approches binaurales peuvent par certains aspects être plus simples à mettre en œuvre et permettent de créer des liens avec les neurosciences en expliquant ou en imitant les processus de perception chez

---

l'humain. Des projets de recherche portant sur l'audition active binaurale, tels le projet franco-japonais ANR/JST BINAAHR (*BINaural Active Audition for Humanoid Robotics*) ou le projet européen FP7 FET Two!Ears, ont vu le jour ces dernières années.

## Organisation de la thèse

Dans cette thèse nous étudions le mouvement actif d'un capteur binaural pour la localisation d'une source sonore. Ce travail s'inscrit dans la thématique de la perception active et consiste à commander le mouvement du capteur binaural de manière à améliorer un processus de localisation. Il s'agit d'une action câblée (qui ne fait donc pas intervenir de processus de décision), qui fait écho au mouvement réflexe effectué par l'humain ou certains animaux, concomitamment à la perception d'un son. La localisation est ici un processus dynamique qui combine une information spatiale extraite du flux binaural avec le mouvement du capteur. Une stratégie de filtrage stochastique est utilisée et nous parlerons de localisation « audio-motrice ». Deux volets importants de la thèse peuvent être définis comme suit :

- La mise en relation de l'information apportée par un processus d'observation avec le déplacement du capteur binaural, étant donnée une connaissance *a priori* sur la position relative capteur-source.
- La génération de mouvement par la résolution d'un problème d'optimisation consistant à maximiser le critère d'information exprimé en fonction de la commande en translation et en rotation du capteur binaural.

Dans le chapitre 1, après un état de l'art de la localisation binaurale en robotique, nous présentons le problème de synthèse de mouvement actif. Le chapitre 2 répond au premier des points évoqués ci-dessus. Des concepts issus de la théorie de l'information sont introduits et nous étudions comment ils peuvent donner lieu à la définition d'un critère d'information au sein d'un processus de filtrage stochastique. D'abord un critère d'information basé sur le prochain pas d'estimation est établi. En fin de chapitre, une extension de ce critère à un horizon temporel plus large est présenté. Dans le chapitre 3, un problème d'optimisation sous contraintes est défini. Le critère d'information exprime un objectif moyen d'exploration à un pas dans le futur. Une solution au problème exploitant des approximations numériques du critère d'information et de son gradient basées sur la transformée *unscented* est développée. Dans le chapitre 4, le problème d'optimisation est étendu de façon à exprimer un objectif moyen d'exploration à  $N$  pas dans le futur. Le critère étant significativement plus complexe, le calcul de son gradient est effectué par un algorithme basé sur l'algèbre des nombres duaux. Enfin dans le chapitre 5 nous discutons des améliorations envisagées et des perspectives ouvertes par nos travaux.



# Localisation binaurale et mouvement actif

---

## Sommaire

<b>1.1 Indices auditifs et modèles pour la localisation binaurale</b>	<b>6</b>
1.1.1 Indices binauraux et monauraux	6
1.1.2 Modèle HRTF	8
<b>1.2 État de l’art de la localisation binaurale en robotique</b>	<b>8</b>
1.2.1 Localisation de sources à partir des indices binauraux	9
1.2.2 Mouvements en boucle ouverte	10
1.2.3 Génération de mouvements actifs	12
<b>1.3 Vers une localisation binaurale active en trois étapes</b>	<b>14</b>
1.3.1 Introduction	15
A Notations	15
B Modèle géométrique	15
1.3.2 Estimation à court terme d’azimuts	17
1.3.3 Localisation audio-motrice	18
1.3.4 Position du problème de synthèse de mouvement actif	20
<b>1.4 Présentation des outils d’évaluation</b>	<b>22</b>
1.4.1 Simulateur binaural	22
1.4.2 Localisation à partir d’un robot mobile	24

---

Les méthodes qui associent la perception audio des robots et leur mobilité se sont beaucoup développées à partir des années 2000, dans le but d’améliorer les processus de perception. En effet, l’audition est un processus qui peut être mal adapté à une vision statique du monde : lorsque les sources et les récepteurs sont immobiles, les modèles de perception audio ne sont représentatifs de la réalité que sur des instants très courts [Cooke *et al.* 2007]. Ainsi la mobilité intrinsèque des robots doit être mise à profit dans des processus « actifs » de façon à améliorer la perception audio, que ce soit à des fins de localisation, de séparation des sources, ou encore de reconnaissance de parole.

Précisons que le terme « actif » est souvent utilisé dans la littérature pour décrire de façon générale la combinaison de la perception et de l’action. Sauf indication contraire, dans ce manuscrit, l’action concernera le mouvement. Dans certains cas, la combinaison perception-action est effectuée « en boucle ouverte » : le mouvement du robot est déterminé par l’utilisateur ou par un processus indépendant de la

perception considérée. Une autre approche consiste à générer le mouvement du robot en boucle fermée sur la perception. Les commandes appliquées au robot peuvent alors être calculées à partir de cette seule information ou par sa combinaison avec un autre critère. Dans cette thèse, le terme de « mouvement actif » désigne l’approche en boucle fermée.

Dans ce premier chapitre, nous nous focalisons principalement sur la localisation binaurale de sources sonores, qui est un élément central de notre étude. Cependant, nous évoquons parfois des travaux qui se situent en dehors du paradigme binaural, en particulier dans la section 1.2.3. Après une présentation des principes de localisation binaurale en 1.1 et des développements effectués en robotique à ce sujet 1.2, nous présentons le problème de mouvement actif étudié dans cette thèse, dans la section 1.3. En fin de chapitre (section 1.4), nous présentons les outils d’évaluation qui ont été utilisés pour valider les approches développées.

## 1.1 Indices auditifs et modèles pour la localisation binaurale

La localisation binaurale consiste à déterminer l’origine spatiale de sources sonores à partir d’une analyse du flux audio perçu par un récepteur binaural. Chez l’humain ce récepteur correspond aux deux oreilles ; en robotique il s’agit de deux microphones qui peuvent être positionnés soit en champ libre (sans interférence acoustique entre la source sonore et chaque microphone), soit sur une tête sphérique ou anthropomorphe, éventuellement à l’intérieur de pavillons artificiels. Nous ferons régulièrement référence à l’axe interaural, qui désigne l’axe qui relie les deux oreilles ou les deux microphones. Une introduction au problème de localisation binaurale pour la robotique peut être trouvée dans [Argentieri *et al.* 2015]. Chez l’humain, la localisation repose en grande partie sur des indices binauraux (section 1.1.1) extraits du flux binaural (signaux perçus par les oreilles gauche et droite). Des indices monauraux, issus de l’analyse du signal perçu par chaque oreille séparément, participent également à la localisation. Tous ces indices peuvent être capturés par la relation spatio-fréquentielle qui unit le signal induit par la source en un point de référence (généralement le point milieu de la tête si celle-ci était absente) et le signal perçu en chaque oreille, au moyen de fonctions de transfert ([Head Related Transfert Function \(HRTF\)](#)) dont nous présentons le principe en 1.1.2.

Généralement, la position d’une source est décrite dans un système de coordonnées sphériques  $(O, r, \theta, \psi)$  dont l’origine  $O$  est le centre de la tête, avec  $r$  la distance de la source à l’origine,  $\theta$  l’azimut de la source par rapport à l’axe frontal et  $\psi$  l’élévation de la source par rapport au plan horizontal.

### 1.1.1 Indices binauraux et monauraux

La théorie *duplex*, introduite dans [Rayleigh 1907], énonce que la localisation horizontale chez l’humain repose essentiellement sur l’exploitation conjointe des in-

indices binauraux que sont la différence interaurale en phase ou en temps ([Interaural Phase Difference \(IPD\)](#) ou [Interaural Time Difference \(ITD\)](#)) et la différence interaurale en amplitude ([Interaural Level Difference \(ILD\)](#)). L'[ILD](#) désigne la différence d'intensité entre les signaux perçus par les oreilles gauche et droite. L'intensité du signal perçu en une oreille est plus importante lorsque la source est située du même côté, en raison de la diffusion de l'onde sur la tête avant d'atteindre la seconde oreille. Cet indice est très fortement dépendant de la fréquence du signal. Pour les signaux de fréquence relativement haute (supérieure à 750Hz), l'intensité des signaux perçus entre les deux oreilles est fortement modifiée par la géométrie de la tête. Au contraire, pour les signaux basse fréquence, la différence devient difficilement perceptible. En revanche, l'[ITD](#), qui représente le temps de retard séparant les perceptions entre les deux oreilles, semble moins sensible aux variations de fréquences. Néanmoins, lorsque la fréquence de la source excède un certain seuil (typiquement 1400Hz), l'[ITD](#) devient peu informatif en raison du repliement spatial. Ainsi chez l'humain, la localisation horizontale d'une source est effectuée par une combinaison « pondérée » de ces deux indices. Pour les sources de faible fréquence, la localisation est principalement effectuée par l'exploitation des [ITD](#), tandis que les [ILD](#) facilitent davantage la localisation de sources à haute fréquence.

Une source située dans le plan vertical symétrique de la tête (il s'agit en anatomie du plan sagittal médian) n'induit pas de différence entre les signaux perçus par les deux oreilles. Les indices relatifs à la localisation verticale sont monauraux. Bien que la perception de l'élévation d'une source soit perceptible chez l'humain (en grande partie grâce à la forme des pavillons des oreilles), les indices monauraux ne seront pas exploités dans cette thèse. Notre problème de localisation sera défini dans un plan horizontal contenant la source et les microphones ; nous y reviendrons dans la partie [1.3.1-B](#).

Tous les indices évoqués précédemment dépendent de la distance qui sépare le récepteur de la source ainsi que de l'angle d'incidence de l'onde acoustique. Cependant la détermination de la distance est une tâche bien plus ardue que l'estimation de l'angle d'incidence. L'humain est capable d'estimer partiellement la distance à une source en combinant les indices binauraux et en utilisant une information *a priori* sur l'environnement. Différents aspects de la perception de la distance chez l'humain sont abordés dans [[Zahorik et al. 2005](#)]. La perception de la distance est souvent surestimée pour les sources proches et sous-estimée pour les sources situées au-delà d'un mètre. Au contraire, la perception de l'azimut chez l'humain est relativement précise. Un individu peut généralement détecter des changements de  $10^\circ$  dans l'angle d'incidence horizontal d'une source et la précision peut être accrue à  $1^\circ$  lorsque la source est située dans la direction de plus grande acuité. Dans le plan transversal (horizontal) contenant les deux oreilles, il s'agit de la direction portée par l'axe normal à l'axe interaural et passant par le centre de la tête. Par la suite cette direction sera parfois appelée fovéa, par analogie à l'anatomie de l'oeil, où le terme fovéa désigne la zone de la rétine permettant la meilleure acuité visuelle.

### 1.1.2 Modèle HRTF

La **HRTF** est un modèle mathématique qui permet de caractériser le son perçu par un microphone, en prenant en compte l'ensemble des phénomènes de diffusion acoustique. C'est une relation spatio-fréquentielle entre le signal induit par la source en un point particulier, dit « de référence » (généralement le point milieu de la tête si celle-ci était absente), et le signal perçu par le microphone. Il s'agit donc d'une fonction de la position de la source (distance au point de référence, azimut et élévation dans un repère centré sur celui-ci) et de sa fréquence. L'antécédent temporel de la **HRTF** est la réponse impulsionnelle (**Head Related Impulse Response (HRIR)**). La fonction de transfert interaurale (**Interaural Transfert Function (ITF)**) résultant du rapport des deux **HRTF** gauche et droite, caractérise donc la propagation du son entre les deux microphones, laquelle dépend à la fois de la position de la source et des caractéristiques du récepteur binaural : microphones en champ libre, placés sur une tête sphérique ou anthropomorphe, placés dans des pavillons artificiels, présence d'un torse, etc. La phase et l'amplitude de l'**ITF** représentent respectivement l'**IPD** et l'**ILD** définis à chaque fréquence. Les **HRTF** dépendent fortement des caractéristiques physiques du récepteur mais également des propriétés du milieu acoustique. Deux méthodes d'obtention des **HRTF** sont possibles. Lorsque la forme du récepteur est simple (typiquement une tête sphérique ou ovale), les lois de propagation acoustique peuvent suffire à retranscrire son effet sur la perception en chaque microphone de manière analytique (voir par exemple [Duda & Martens 1998]). Dans le cas d'un mannequin anthropomorphe admettant une géométrie plus complexe, ce calcul devient impossible et les **HRTF** doivent être évaluées expérimentalement. Les réverbérations éventuelles du signal source n'étant pas prises en compte par le modèle, ces mesures doivent être effectuées en chambre anéchoïque.

## 1.2 État de l'art de la localisation binaurale en robotique

Les approches binaurales sont souvent considérées comme étant moins efficaces que celles utilisant une antenne de plusieurs microphones ; la redondance d'information est moins importante, ce qui rend la localisation moins précise. Les approches binaurales possèdent néanmoins de nombreux avantages et connaissent depuis plusieurs années un certain regain d'intérêt au sein de la communauté scientifique, notamment depuis l'apparition de processus actifs liant mouvement et perception. L'un des avantages majeurs des solutions binaurales est sans doute la simplicité de mise en œuvre (diminution des coûts et facilité d'utilisation grâce à des cartes d'acquisition stéréo). Notons cependant que le nombre réduit de microphones impose souvent des algorithmes plus complexes qu'avec des antennes de plus grande taille, ce qui peut donc engendrer un temps de calcul plus élevé. Par ailleurs, la perception binaurale en robotique permet de créer des liens intéressants avec les neurosciences, en tentant d'imiter et d'expliquer les processus de perception chez

l'humain ou chez les animaux. Ainsi l'approche binaurale est aussi une opportunité de mieux comprendre l'audition humaine.

### 1.2.1 Localisation de sources à partir des indices binauraux

En robotique, les microphones sont généralement disposés sur la tête du robot (tête sphérique ou anthropomorphe). Les indices binauraux (ILD et ITD) sont les indices les plus utilisés et leur extraction à partir du flux binaural est généralement effectuée en vue d'obtenir une information sur l'azimut de la source. Notons que dans le cas particulier où les microphones sont en champ libre, les ITD sont relativement indépendants de la fréquence et les ILD n'apportent que très peu d'information (la différence d'intensité entre les signaux gauche et droit tient essentiellement à la présence d'un élément diffusant entre les microphones), ce qui rend ces derniers indices difficilement exploitables. L'extraction de ces indices à partir du flux binaural peut être effectuée à partir de diverses méthodes (calculs dans les domaines temporels ou fréquentiels, méthodes bio-inspirées, etc.) dont une liste non exhaustive peut être trouvée dans [Youssef *et al.* 2012]. Dans certaines études [Lim & Duda 1994, Raspaud *et al.* 2010], les indices binauraux utilisés pour l'estimation de l'azimut sont obtenus à partir de données HRTF.

Cependant l'exploitation des indices binauraux ne donne qu'une information partielle sur la position de la source. Premièrement, ces indices ne fournissent pas suffisamment d'information pour estimer de manière satisfaisante la distance qui sépare le capteur de la source. Ensuite, les indices binauraux ne permettent pas d'estimer l'azimut sans ambiguïté. En effet, dans le cas d'une paire de microphones en champ libre, une même valeur d'ITD peut être associée à différents angles d'incidence de la source, formant un hyperboloïde à deux nappes (gauche et droite). Cette surface peut être assimilée à un cône pour des sources suffisamment éloignées. L'expression « cône de confusion » est utilisée dans la littérature. Les résultats de différentes études présentées dans [Shinn-Cunningham *et al.* 2000], montrent que ce cône se forme également lorsque les microphones sont placés sur une tête sphérique. Enfin, notons que cette surface d'incertitude en 3D est réduite à une ambiguïté avant-arrière lorsque la localisation est effectuée dans un plan horizontal.

Ce concept de cône de confusion est connu depuis longtemps et des études ont tenté de démontrer l'intérêt du mouvement de la tête pour la localisation chez l'humain. Dans [Wallach 1938], l'influence du mouvement de la tête est étudiée de façon à mieux caractériser ce cône de confusion. Dans [Thurlow & Runge 1967], la contribution du mouvement de la tête pour la localisation d'une source est considérée. Pour cette étude, des individus doivent localiser au mieux des sources sonores dans différentes conditions tout en effectuant divers mouvements de la tête. Les auteurs montrent que l'erreur de localisation horizontale est fortement diminuée lorsqu'un mouvement de rotation est effectué. La réduction d'erreur verticale n'est par contre pas significative.

Chez l'humain la tête n'est pas parfaitement sphérique, les oreilles ne sont pas diamétralement opposées et la tête n'est pas symétrique par rapport au plan trans-

versal contenant l'axe interaural. Par ailleurs, il est reconnu que le torse perturbe la diffusion de l'onde sonore aux basses fréquences [Algazi *et al.* 2001]. En conséquence, les indices binauraux ne sont pas parfaitement constants pour des sources placées sur un même cône de confusion [Shinn-Cunningham *et al.* 2000]. À partir de ce constat, les limitations apparaissant lors de l'estimation de la position d'une source à partir des indices binauraux pourraient être théoriquement levées en plaçant le capteur binaural sur un mannequin anthropomorphe (composé d'une tête et d'un torse). Ces mannequins sont conçus de manière à émuler les mêmes phénomènes de diffusion acoustique que chez l'humain. Cependant s'il a été démontré que les dissymétries du corps humain pouvaient participer à la localisation, de tels indices acoustiques encodent en réalité des relations extrêmement complexes entre la fréquence et la position de la source, rendant ces dissymétries difficilement exploitables.

Récemment, dans [Portello *et al.* 2013], deux microphones placés sur une tête sphérique sont exploités de façon à localiser horizontalement une source. Une fonction de « pseudo log-vraisemblance » de l'azimut est établie à partir d'une analyse temps-fréquence du flux binaural et de la fonction de transfert interaurale (ITF). L'ITF est naturellement définie à partir des HRTF gauche et droite, qui prennent en compte la diffusion acoustique sur la tête. Les auteurs montrent que la fonction de vraisemblance obtenue permet une estimation de l'azimut de meilleure qualité que dans les deux autres cas suivants : (1) flux binaural obtenu par des microphones en champ libre et exploitation d'une ITF décrivant une propagation en champ libre ; (2) flux binaural obtenu par des microphones effectivement disposés sur la tête, mais exploitation de HRTF et ITF décrivant une propagation en champ libre ou reposant sur l'approximation de Woodworth-Schlosberg (ILD nuls, et écriture simplifiée des ITD). Le cas (1) prouve que les ILD implicitement inclus dans la décomposition temps-fréquence du flux binaural relatif à la tête sphérique apportent une information significative sur l'origine spatiale de la source. Le cas (2) permet d'apprécier les erreurs commises lors de l'exploitation de modèles de propagation simplifiés. L'ambiguïté avant-arrière demeure naturellement présente. Cette méthode qui sera davantage détaillée en 1.3.2, a été étendue pour l'estimation des azimuts relatifs à plusieurs sources [Portello *et al.* 2014a].

### 1.2.2 Mouvements en boucle ouverte

En robotique, un moyen de lever l'ambiguïté avant-arrière et d'estimer la distance, est d'assimiler au cours du temps l'information spatiale provenant de l'analyse court-terme des indices binauraux et de la combiner avec les ordres moteurs du capteur binaural (qui induisent son déplacement selon une loi de dynamique *a priori* définie à l'avance) [Nakadai *et al.* 2000]. Il s'agit de processus de localisation audio-moteurs. Dans [Lu & Cooke 2010], les signaux gauche et droit sont d'abord traités par des filtres de type *gammatone*. Ensuite, les ITD sont estimés comme l'argument maximisant la somme de la corrélation croisée des sorties de ces filtres. La localisation repose sur un filtre à particules. L'étude est effectuée en simulation

et met en avant le fait que les performances de la localisation sont influencées par la nature des mouvements de la tête binaurale, tels que des mouvements aléatoires et les déplacements en direction de la source. Une approche différente est proposée dans [Portello *et al.* 2011] où, sur la base du temps de retard mesuré entre deux microphones en champ libre, un filtre de Kalman unscented multi-gaussien (**Multi-Gaussian Unscented Kalman Filter (MG-UKF)**) constitue une alternative au filtre particulaire pour la localisation d’une source unique. Le **MG-UKF** est pourvu d’une initialisation automatique et permet d’éviter une estimation trop optimiste de la covariance de la variable d’état représentant la position de la source. De par l’ambiguïté avant-arrière dans l’estimation de l’azimut, la densité de probabilité (ou **probability density function (pdf)**) de la position de la source relativement au capteur binaural est multimodale, et peut difficilement être assimilée à une loi gaussienne. Le mélange de gaussiennes permet une meilleure approximation. Dans [Portello *et al.* 2012], les auteurs étendent leur approche à la gestion des fausses mesures en implémentant un filtre multi-hypothèses à association de données probabiliste (**Multiple Hypothesis Probabilistic Data Association Filter (MH-PDAF)**). La méthode est inspirée du **Probabilistic Data Association Filter (PDAF)** de [Bar-Shalom *et al.* 2009], qui est modifié de façon à intégrer l’approche à mélange de gaussiennes de [Portello *et al.* 2011]. Le **MH-PDAF** est ensuite complété de façon à gérer l’intermittence de la source, par l’intégration d’un détecteur d’activité basé sur le test du rapport de vraisemblances généralisé (**Generalized Likelihood Ratio Test (GLRT)**). Typiquement, deux modèles  $M_1$  et  $M_2$  sont définis, qui représentent respectivement les cas où la source est active ou inactive. Le **GLRT** permet de détecter les transitions inter-modèles sans connaissance *a priori* relative à celles-ci. Une approche complémentaire est proposée dans [Nguyen *et al.* 2016] pour la localisation d’une source intermittente depuis un robot mobile muni d’une antenne de microphones. Les auteurs introduisent une extension d’un filtre de Kalman étendu multi-gaussien, appelée **Mixture Kalman Filter (MKF)**. Celle-ci permet d’estimer un vecteur d’état mixte discret-continu qui agrège l’activité de la source et sa position relativement au capteur, sur la base d’observations, également mixtes, issues de l’analyse court-terme du flux binaural par un algorithme de détection d’activité et un algorithme d’estimation d’azimut.

La fonction de pseudo log-vraisemblance de l’azimut évoquée en 1.2.1 et introduite dans [Portello *et al.* 2013], qui exploite directement les données **HRTF** et donc l’intégralité de la réponse acoustique de l’élément diffusant, a également été exploitée dans des méthodes de filtrage. Dans [Marković *et al.* 2013], cette fonction, souvent définie avec une résolution non maîtrisée (typiquement  $1^\circ$  ou  $5^\circ$  pour un ensemble tête-torse anthropomorphe, selon la résolution spatiale des relevés **HRTF**), est d’abord ajustée à des distributions circulaires, non normalisées, de von Mises ou Cauchy. L’intérêt d’introduire des lois circulaires est de prendre en compte naturellement le fait que l’argument de la pseudo-vraisemblance, *i.e.*, l’azimut, vit sur le cercle unité et non sur la droite réelle. Les deux modèles ont été comparés, et il a été montré que les distributions de von Mises permettent un ajustement plus pertinent. La vraisemblance originale et son expansion par une loi de von Mises ont ensuite été

intégrées dans une stratégie de localisation audio-motrice basée sur un filtre à particules. Les propriétés de ce filtre ont été testées par des simulations de Monte-Carlo, mettant en avant la pertinence de l’approche mais soulignant l’apparition ponctuelle d’inconsistance (estimé de la covariance *a posteriori* trop optimiste). Selon les auteurs, ce problème pourrait être réglé en augmentant le nombre de particules ou en ajustant l’initialisation du filtre. L’évaluation des vraisemblances des particules au moyen de l’expansion de la pseudo-vraisemblance par une loi de von Mises permet toutefois une meilleure estimation de la distance sans dégrader l’estimation de l’azimut. Enfin, la fonction de pseudo log-vraisemblance de l’azimut de [Portello *et al.* 2013] a également été approchée par un mélange de gaussiennes, de façon à être intégrée dans un schéma (modifié) de type MG-UKF [Portello 2013] [Portello *et al.* 2014b]. Nous reviendrons sur cette approche dans la partie 1.3.3.

Notons toutefois qu’en pratique, la qualité de la perception auditive par un robot, et donc la qualité des processus qui en dépendent, peut être altérée par le bruit qu’il génère, ou « ego-noise ». Une illustration de ce problème est donnée dans [Furukawa *et al.* 2013], où un *drone* multi-rotors équipé d’une antenne de microphones doit faire face au bruit stationnaire qu’il émet durant son vol, le tout en effectuant une tâche de localisation. La référence [Ince *et al.* 2009] étudie des solutions au problème de l’ego-noise et propose une méthode de soustraction spectrale pour réduire son impact. Les performances sont évaluées sur des expériences de reconnaissance automatique de la parole.

### 1.2.3 Génération de mouvements actifs

Chez l’humain, la localisation de sources sonores est améliorée par des mouvements réflexes exécutés naturellement et inconsciemment. En réalité, notre corps ne réagit pas seulement à un stimulus audio mais plutôt par rapport à ce que cette perception audio signifie dans une situation physique, émotionnelle et cognitive spécifique [Blauert 2017]. En robotique, relativement peu d’études portent sur la synthèse de mouvements permettant d’améliorer les processus de localisation audio-motrice. Plus largement, la commande de robots basée sur des capteurs extéroceptifs tels que des capteurs audio peut être abordée selon deux approches : les méthodes « sensor-based » (*sensor-based control*), et les approches « par retour d’état » (*situation-based control*) qui comportent une étape interne de reconstruction de la position et orientation 3D sur la base du flux sensorimoteur. Dans cette section nous présentons des études qui commandent le mouvement d’un capteur audio en boucle fermée sur les signaux qu’il capture. Nous évoquons dans un premier temps des travaux qui utilisent une approche « sensor-based », où la localisation de la source n’est pas l’objectif principal. Nous concentrons ensuite notre étude sur les approches par retour d’état qui mettent en jeu des processus de localisation et qui correspondent mieux aux particularités de notre problème.

Lorsque le problème de commande est exprimé dans l’espace des capteurs, le but est d’agir en boucle fermée sur la perception de façon que certaines primitives perceptuelles atteignent leurs valeurs de référence. Un « audio-servo » a ainsi été

proposé dans [Kumon *et al.* 2003], permettant de diriger l’orientation de la tête en direction de la source en se basant sur le signal audio brut perçu par deux microphones. Les mouvements sont limités à des rotations de la tête. Avec le même objectif, [Kumon & Noda 2011] proposent un système de pavillons déformables, par analogie à ceux des chats. Dans [Magassouba *et al.* 2015], le contrôle en translation et en rotation d’un capteur binaural dans un plan est étudié. Les auteurs utilisent des mesures d’ITD pour asservir les angles d’incidences d’une à trois sources, sur la base d’une méthode qui n’implique pas de localisation. Une validation expérimentale de cette approche est proposée dans [Magassouba *et al.* 2016], pour une ou deux sources mobiles. Une étude complète des commandes « audio-based » (analyse, synthèse, validation expérimentale) pour des primitives telles que l’ILD, l’ILD en conjonction avec l’énergie sonore absolue perçue, et l’ITD, est proposée dans [Magassouba 2016].

Les approches par retour d’état ont été davantage traitées dans la littérature. Comme notre application implique un processus de localisation, il nous paraît naturel de préférer ces méthodes. Dans certaines études, la notion d’entropie est utilisée. Il s’agit d’un critère d’information (que nous rappellerons dans la partie 2.1) qui représente l’incertitude associée à une variable aléatoire. Le problème de séparation de sources multiples est abordé dans [Sekiguchi *et al.* 2015]. Les auteurs proposent de diriger des robots équipés d’antennes de microphones vers les positions depuis lesquelles la tâche de séparation peut être effectuée de manière coopérative avec une efficacité maximale. Dans [Berglund & Sitte 2005], le problème de localisation binaurale est étudié à l’aide d’un robot mobile. Les auteurs montrent que l’estimation d’azimuts est plus précise quand la source se situe dans le plan sagittal médian ; c’est à dire quand la fovéa est orientée dans la direction de la source. À partir d’une analyse fréquentielle, le robot détermine la direction de la source et s’oriente dans sa direction grâce à des techniques d’apprentissage par renforcement. Plus récemment, un robot équipé de deux microphones placés dans des pavillons artificiels et d’une caméra, localise une source sonore en combinant acuité visuelle et auditive dans une méthode de filtrage à particules [Schult *et al.* 2015]. Cette méthode consiste à estimer récursivement la pdf *a posteriori* de la position relative capteur-source, qui constitue le vecteur d’état. Le mouvement appliqué au robot est celui qui maximise une fonction d’information calculée par une différence entre l’entropie de la pdf de l’état à l’instant courant et l’entropie de la pdf à l’instant suivant, après avoir appliqué une action et assimilé une nouvelle observation. Dans [Schymura *et al.* 2017], un algorithme d’exploration de Monte Carlo est utilisé pour diriger un robot vers une position spécifique en minimisant un critère de distance, tout en réduisant un critère d’entropie permettant de minimiser l’incertitude liée à la localisation de la source. En effet, pour un grand nombre d’applications, la localisation n’est pas le seul objectif. Les résultats montrent que la stratégie d’exploration pure guide le robot en direction de la source (nous verrons dans le chapitre 3 en quoi cela contribue à l’amélioration de la localisation). Néanmoins, dans ce cas, les performances de la localisation sont dégradées. Les auteurs précisent qu’en outre, le robot se rapproche alors trop de la position de la source, ce qui perturbe l’estimation des indices bi-

nauraux. Cela est dû à la [HRIR](#) utilisée, qui n'est pas adaptée au champ proche, ainsi qu'à d'importantes différences d'azimuts entre deux instants consécutifs.

Dans les études précédentes, les commandes sont établies de façon à maximiser un critère court terme ; c'est à dire que la récompense est attribuée immédiatement après l'application de la commande. Dans la littérature, de telles approches sont appelées stratégies « myopes » ou « gloutonnes ». Une autre méthode consiste à établir un critère à long terme, c'est à dire dépendant de l'état d'information du système après l'application d'une suite de  $N$  commandes. Nous qualifions ces approches de « stratégies à long terme » (voir section [2.2.3](#)). Dans [[Vincent et al. 2015](#)], le déplacement d'un robot équipé d'une antenne de microphones est planifié  $N$  pas à l'avance (avec  $N \geq 1$ ). L'objectif est de localiser une source sonore en minimisant un critère basé sur l'entropie d'une grille d'occupation, utilisée pour caractériser la probabilité de la position de la source. Pour tout instant  $k$ , l'entropie de la grille d'occupation à l'instant  $k + N$  dépend des futures observations et ne peut donc pas être évaluée de façon certaine. Le critère à minimiser est alors défini comme l'espérance de cette entropie par rapport aux futures observations. Afin de rendre possible le calcul du critère lorsque la valeur de l'horizon  $N$  augmente, l'espérance de l'entropie est calculée séparément pour chaque future position possible du robot (position et orientation possible de l'antenne de microphones). Les auteurs supposent donc que l'entropie associée à chaque position future du robot ne dépend pas de la trajectoire suivie pour atteindre cette position. Cette approximation permet d'aborder la minimisation du critère comme un problème de programmation dynamique qui est résolu par l'algorithme de Viterbi. Le même problème de localisation sur un horizon à long terme a été abordé très récemment dans [[Nguyen et al. 2017](#)] avec une approche différente. La *pdf a posteriori* relative à la position de la source, du robot et à l'activité de la source, est modélisée par un mélange de gaussiennes calculé récursivement au moyen d'un filtre [MKF](#). Ensuite, une méthode de recherche arborescente de Monte-Carlo ([Monte Carlo Tree Search \(MCTS\)](#)) est utilisée en vue d'obtenir la séquence de commandes qui minimise l'entropie de la *pdf a posteriori*. Même lorsque l'espace discret des commandes est de taille limitée, l'évaluation du critère pour toutes les séquences de commandes admissibles est généralement irréalisable. Cette méthode heuristique permet de n'évaluer le critère que pour certaines séquences de commandes. L'algorithme se termine quand le temps alloué à la recherche est épuisé et la séquence de commande qui conduit à l'entropie la plus basse est retenue. Les résultats de simulations montrent que la méthode permet une meilleure réduction d'entropie sur le long terme qu'une méthode gloutonne.

### 1.3 Vers une localisation binaurale active en trois étapes

Dans cette partie nous présentons des éléments de modélisation relatifs à la problématique de cette thèse. Le modèle géométrique liant une source à un repère

lié au capteur binaural est d'abord défini. Nous introduisons ensuite les solutions d'estimation de l'azimut et de localisation audio-motrice que nous adoptons. Enfin nous définissons le problème de mouvement actif étudié.

### 1.3.1 Introduction

#### A Notations

Certaines notations et conventions relatives au calcul des probabilités sont présentées ici. Soit  $\mathbf{x}$  une variable ou un vecteur aléatoire et  $x$  une réalisation de  $\mathbf{x}$ . Pour simplifier les notations, nous omettrons la mention de  $\mathbf{x}$  dans la définition de sa pdf  $p_{\mathbf{x}}$ , i.e., nous remplacerons  $p_{\mathbf{x}}(x)$  par  $p(x)$ .

Dans cette thèse, nous travaillerons essentiellement avec des variables continues. Nous supposons que toutes les variables aléatoires continues considérées sont associées à une pdf. Nous considérerons fréquemment des distributions de probabilité gaussiennes réelles. Ainsi, un vecteur aléatoire  $\mathbf{x}$  gaussien admet une pdf de la forme

$$p(x) = |2\pi\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}, \quad (1.1)$$

où  $\mu$  et  $\Sigma$  sont respectivement la moyenne et la covariance de la distribution et  $|\cdot|$  le déterminant. La covariance est une matrice symétrique, définie positive. Nous ferons parfois référence à  $\mu$  et  $\Sigma$  comme étant les premiers moments de la distribution ( $\mu = \mathbb{E}[\mathbf{x}]$  et  $\Sigma = \mathbb{E}[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T]$ ). Nous utiliserons fréquemment la notation

$$p(x) = \mathcal{N}(x; \mu, \Sigma). \quad (1.2)$$

#### B Modèle géométrique

Considérons deux microphones  $R_1$  et  $R_2$  disposés sur un élément diffusant qui peut être une sphère ou une tête anthropomorphe. Ces microphones sont rattachés à un repère orthonormé  $\mathcal{F} = (O, \vec{x}_R, \vec{y}_R, \vec{z}_R)$ , avec  $\vec{R}_1\vec{O} = \vec{OR}_2$  (figure 1.1a). Le vecteur  $\vec{y}_R = \frac{\vec{R}_2\vec{R}_1}{\|\vec{R}_2\vec{R}_1\|}$  supporte l'axe interaural tandis que  $\vec{z}_R$  pointe vers l'avant de la tête. Une source sonore  $E$  est supposée fixe, ponctuelle et omnidirectionnelle. De plus, elle est placée suffisamment loin par rapport au capteur binaural, pour que les ondes sonores qui atteignent le capteur puissent être considérées comme planes. C'est l'hypothèse de champ lointain. Le capteur binaural étant mobile, désignons par  $\mathcal{F}_k = (O_k, \vec{x}_{Rk}, \vec{y}_{Rk}, \vec{z}_{Rk})$  son repère à l'instant discret  $k$ . Définissons le repère monde par le repère  $\mathcal{F}_k$  à l'état initial : le repère  $\mathcal{F}_0$  ( $\mathcal{F}_{k=0}$ ). Enfin, supposons que pour tout  $k$ , la source et le capteur binaural appartiennent au même plan  $(O_k, \vec{y}_{Rk}, \vec{z}_{Rk})$ . Les coordonnées de la source par rapport à  $\mathcal{F}_k$  sont alors notées  $x_k = (0, e_{y_k}, e_{z_k})^T$ . De même, la distance et l'azimut de la source par rapport au capteur sont respectivement définies par  $r_k = \sqrt{e_{y_k}^2 + e_{z_k}^2}$  et  $\theta_k = -\text{atan2}(e_{y_k}, e_{z_k})$ .

Pour tout  $x_k$ , la position relative de la source dans le repère  $\mathcal{F}_{k+1}$  est déterminée par la relation déterministe idéale  $x_{k+1} = f(x_k, u_k)$ .  $u_k$  est le vecteur de commande

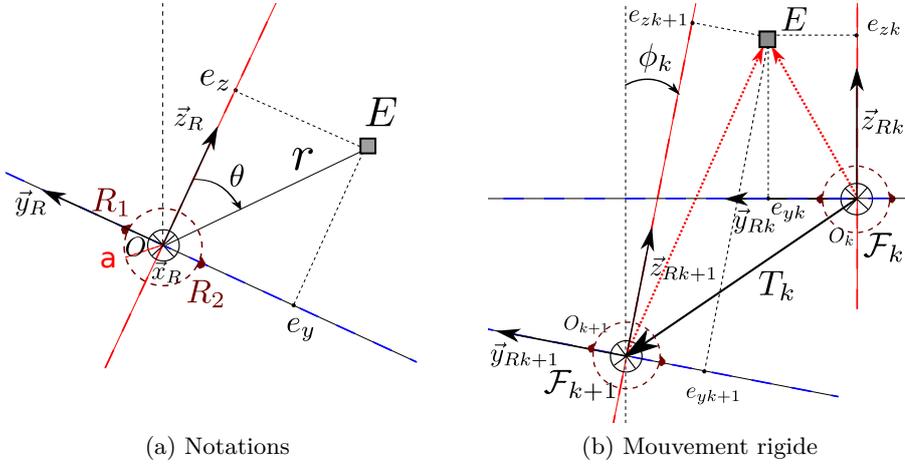


FIGURE 1.1 – Représentation planaire du problème – L’ensemble marron décrit le dispositif binaural avec les microphones  $R_1$ ,  $R_2$  et un élément diffusant ici représenté par une sphère de rayon  $a$ . La source sonore  $E$  est située sur le plan  $(O, \vec{y}_R, \vec{z}_R)$ . Notations (a) : Les ensembles  $\{e_y, e_z\}$  et  $\{r, \theta\}$  représentent respectivement les coordonnées cartésiennes et polaires de la source dans le repère lié au capteur binaural. Mouvement rigide du capteur (b) : La translation  $T_k$  suivie de la rotation  $\phi_k$  (exprimées dans le repère  $\mathcal{F}_k$ ) caractérisent le déplacement du capteur du repère  $\mathcal{F}_k$  à  $\mathcal{F}_{k+1}$ . Les coordonnées de la source sont représentées dans les deux repères.

décrivant la transformation rigide (figure 1.1b) définie par la translation

$$T_k = \begin{pmatrix} \overrightarrow{O_k O_{k+1}} \cdot \vec{x}_{Rk} \\ \overrightarrow{O_k O_{k+1}} \cdot \vec{y}_{Rk} \\ \overrightarrow{O_k O_{k+1}} \cdot \vec{z}_{Rk} \end{pmatrix} = \begin{pmatrix} T_{x_k} \\ T_{y_k} \\ T_{z_k} \end{pmatrix}, \quad (1.3)$$

dont les éléments sont les coordonnées du centre de  $\mathcal{F}_{k+1}$  exprimées dans  $\mathcal{F}_k$ , suivie de la rotation d’angle  $\phi_k \triangleq \widehat{(\vec{z}_{Rk}, \vec{z}_{Rk+1})}$  autour de l’axe  $\vec{x}_{Rk}$ . En définissant  $u_k = (T_k^T, \phi_k)^T = (T_{x_k}, T_{y_k}, T_{z_k}, \phi_k)^T$ , la relation entre  $x_k$  et  $x_{k+1}$  s’écrit alors

$$x_{k+1} = R^T(\phi_k)x_k - R^T(\phi_k)T_k, \quad (1.4)$$

avec  $R(\phi_k) \in SO(3)$ , la matrice de rotation dont les trois vecteurs colonnes sont les vecteurs de la base  $\mathcal{F}_{k+1}$  exprimés dans  $\mathcal{F}_k$ , i.e.,

$$R(\phi_k) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi_k) & -\sin(\phi_k) \\ 0 & \sin(\phi_k) & \cos(\phi_k) \end{pmatrix}. \quad (1.5)$$

*PREUVE.* Soit  $x_k = \overrightarrow{O_k \hat{E}}_{(\mathcal{F}_k)}$  le vecteur des coordonnées de la source exprimées dans le repère  $\mathcal{F}_k$ . L’égalité vectorielle  $\overrightarrow{O_k \hat{E}}_{(\mathcal{F}_k)} = \overrightarrow{O_k O_{k+1}}_{(\mathcal{F}_k)} + \overrightarrow{O_{k+1} \hat{E}}_{(\mathcal{F}_k)}$  peut

être transformée en

$$\overrightarrow{O_k \dot{E}_{(\mathcal{F}_k)}} = T_k + R(\phi_k) \overrightarrow{O_{k+1} \dot{E}_{(\mathcal{F}_{k+1})}}, \quad (1.6)$$

avec  $T_k = \overrightarrow{O_k O_{k+1}(\mathcal{F}_k)}$  le vecteur des coordonnées de la translation entre les deux repères  $\mathcal{F}_k$  et  $\mathcal{F}_{k+1}$ , et  $\overrightarrow{O_{k+1} \dot{E}_{(\mathcal{F}_k)}} = R(\phi_k) \overrightarrow{O_{k+1} \dot{E}_{(\mathcal{F}_{k+1})}}$  le vecteur des coordonnées de la source à l'instant  $k+1$ , exprimées dans le repère  $\mathcal{F}_k$ . Du fait que  $R(\phi_k)$  est une matrice orthogonale, la transformation entre les coordonnées  $x_k = \overrightarrow{O_k \dot{E}_{(\mathcal{F}_k)}}$  et les coordonnées  $x_{k+1} = \overrightarrow{O_{k+1} \dot{E}_{(\mathcal{F}_{k+1})}}$  s'écrit

$$x_{k+1} = R^T(\phi_k)x_k - R^T(\phi_k)T_k. \quad (1.7)$$

■

Soit  $\mathbf{x}_k$  le vecteur aléatoire associé à la position de la source à l'instant  $k$ . Le passage des états  $\mathbf{x}_k$  à  $\mathbf{x}_{k+1}$  peut alors être caractérisé par l'équation stochastique à temps discret

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, u_k) + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(0, Q_k), \quad (1.8)$$

avec  $\mathbf{w}_k$  le bruit de dynamique, supposé blanc, additif, gaussien, centré de matrice de covariance  $Q_k$ . La localisation étant réduite au plan  $(O_k, \overrightarrow{y_{Rk}}, \overrightarrow{z_{Rk}})$  parallèle au sol, nous considérerons le vecteur d'état réduit  $\mathbf{x}_k = (\mathbf{e}_{y_k}, \mathbf{e}_{z_k})^T$  et supposons que la première composante du bruit de dynamique 3D est identiquement nulle. De même, le vecteur de commande devient  $u_k = (T_{y_k}, T_{z_k}, \phi_k)^T$  et le vecteur de translation  $T_k$  ainsi que la matrice  $R(\phi)$  sont redéfinis par  $T_k = (T_{y_k}, T_{z_k})^T$  et

$$R(\phi_k) = \begin{pmatrix} \cos(\phi_k) & -\sin(\phi_k) \\ \sin(\phi_k) & \cos(\phi_k) \end{pmatrix}, \quad (1.9)$$

ce qui permet de définir l'équation de dynamique du capteur binaural dans le plan  $(O_k, \overrightarrow{y_{Rk}}, \overrightarrow{z_{Rk}})$ , par

$$\mathbf{x}_{k+1} = R^T(\phi_k)\mathbf{x}_k - R^T(\phi_k)T_k + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(0, Q_k), \quad (1.10)$$

### 1.3.2 Estimation à court terme d'azimuts

Notons dans un premier temps que l'expression « court terme » dans ce contexte, est employée pour caractériser la fenêtre temporelle courte sur laquelle l'estimation de l'azimut est effectuée. Considérons que les signaux  $\mathbf{r}_1, \mathbf{r}_2$  perçus par les microphones  $R_1, R_2$  (en champ libre, ou disposés sur une tête) soient modélisés par le processus stochastique

$$\begin{cases} \mathbf{r}_1(t) &= \mathbf{s}(t) + \mathbf{n}_1(t) \\ \mathbf{r}_2(t) &= (\mathbf{s} * h_\theta)(t) + \mathbf{n}_2(t). \end{cases} \quad (1.11)$$

Le signal  $\mathbf{s}$  représente la contribution de la source en  $R_1$  et  $\mathbf{n}_1, \mathbf{n}_2$  sont des bruits liés à la mesure.  $\{\mathbf{s}, \mathbf{n}_1, \mathbf{n}_2\}$  sont supposés réels, de moyenne nulle, conjointement gaussiens, à bande limitée, individuellement et conjointement stationnaires, et nous considérons que les bruits  $\mathbf{n}_1$  et  $\mathbf{n}_2$  sont statistiquement indépendants de  $\mathbf{s}$ . La réponse impulsionnelle interaurale ([Interaural Impulse Response \(IIR\)](#)), *i.e.*, l'antécédent temporel de la [ITF](#)) de  $R_1$  à  $R_2$  est notée  $h_\theta$ , et  $*$  désigne le produit de convolution. Notons que  $h_\theta$  ne dépend que de l'azimut car la source est supposée en champ lointain, la localisation se fait dans le plan horizontal et le modèle ne prend pas en compte les réverbérations. Le flux audio perçu par les microphones durant une courte période de temps  $\varepsilon$ , est associé à une itération  $k$  de sorte que cette période soit identifiée par un segment  $I_k = [kT_s - \varepsilon, kT_s]$ , avec  $T_s$  la période d'échantillonnage du signal. Cette période est choisie suffisamment courte pour que les conséquences du mouvement du capteur binaural sur le processus de perception soient négligeables (quelques dizaines de millisecondes).

La transformation de Fourier court terme est ensuite appliquée à ce segment de manière à obtenir ce que nous appelons la décomposition canal-temps-fréquence  $y_k$  du signal binaural. Cette décomposition  $y_k$  constitue une réalisation du vecteur d'observation  $\mathbf{y}_k$  à l'itération  $k$ . La pdf du vecteur d'observation  $\mathbf{y}_k$  est paramétrée par un vecteur déterministe inconnu  $\Theta_k = [\theta_k, \Psi_k^T]^T$ , avec  $\theta_k$  l'azimut de la source dans le repère  $\mathcal{F}_k$  et  $\Psi_k$  les paramètres spectraux de la source et du bruit (valeurs des densités spectrales de puissance de la source et du bruit, pour chaque indice fréquentiel appartenant à la plage de fréquences considérée). À partir de  $y_k$  et de la fonction interaurale en champ lointain, quotient des [HRTF](#) droite et gauche, la vraisemblance  $p(y_k|\Theta_k)$  peut être explicitée. Le problème présente une propriété de séparation qui permet de calculer l'estimé du maximum de vraisemblance  $\hat{\theta}_k$  en maximisant une fonction de « pseudo log-vraisemblance »  $L(\theta_k)$ , obtenue en remplaçant dans l'expression de  $p(y_k|\Theta_k)$  les paramètres spectraux  $\Psi_k$  par leurs valeurs les plus vraisemblables exprimées en fonction du paramètre spatial  $\theta_k$ . Une explication détaillée de la procédure peut être trouvée dans [[Portello et al. 2013](#), [Portello 2013](#)].

### 1.3.3 Localisation audio-motrice

Une stratégie de filtrage est mise en place afin qu'à tout instant  $k$ , sur la base des données observées  $y_{1:k} = y_1, \dots, y_k$  et des ordres moteur  $u_{0:k-1} = u_0, \dots, u_{k-1}$ , la pdf *a posteriori* (ou « état de croyance »)  $p(x_k|y_{1:k})$  du vecteur d'état caché  $\mathbf{x}_k$  conditionnellement aux observations  $y_{1:k}$ , puisse être établie. Comme les données observées ne renseignent que sur l'azimut de la source et ce de façon partielle à cause de l'ambiguïté avant-arrière (multimodalité de la fonction de pseudo log-vraisemblance  $L(\theta_k)$  développée par la méthode évoquée en [1.3.2](#)), l'objectif est d'obtenir une information plus complète sur la position de la source en supprimant cette ambiguïté et en recouvrant la distance capteur-source. La non-linéarité du modèle de dynamique et l'absence de modèle d'observation reliant explicitement une réalisation du vecteur d'état à une réalisation du vecteur d'observation, additionnées au fait que

la fonction de vraisemblance soit multimodale, font que la caractérisation de la *pdf a posteriori*  $p(x_k|y_{1:k})$  ne peut pas être effectuée au moyen de méthodes classiques de filtrage. Nos tentatives d'application des versions étendues (**Extended Kalman Filter (EKF)**) ou *unscented* (**Unscented Kalman Filter (UKF)**) du filtre de Kalman n'ont pas été satisfaisantes car elles conduisaient à des covariances *a posteriori* trop optimistes, et car la *pdf a posteriori* ne peut pas être décrite convenablement par une représentation gaussienne à cause de la multimodalité introduite par l'ambiguïté avant-arrière. Comme cela a été montré dans [Marković *et al.* 2013], ce problème peut être résolu par une méthode de filtrage particulière. Nous choisissons d'utiliser la stratégie de filtrage développée dans [Portello 2013], que nous introduisons brièvement ici.

La fonction  $L(\theta_k)$  n'a pas de forme analytique et ne donne que des valeurs numériques de la vraisemblance pour un ensemble discret d'azimuts. Afin de pouvoir assimiler l'information contenue dans cette fonction dans une stratégie de filtrage bayésien, une fonction de vraisemblance  $p(y_k|\theta_k)$  est construite comme un mélange de gaussiennes, de sorte que

$$p(y_k|\theta_k) = \sum_{j=1}^J \gamma_k^j \mathcal{N}(\theta_k; m_k^j, \phi_k^j), \quad (1.12)$$

avec  $\{\gamma_k^j, m_k^j, \phi_k^j\}_{j=1,\dots,J}$  les paramètres du mélange de gaussiennes, déduits des valeurs de  $L(\theta_k)$ . Cette approche permet d'éviter d'assimiler directement l'argument qui maximise la fonction de pseudo log-vraisemblance, ce qui n'aurait pas constitué une solution satisfaisante. En effet, cette fonction fait apparaître une ambiguïté avant-arrière (correspondant à deux modes distincts dans les valeurs de la fonction). Notons que cette ambiguïté persiste même quand les microphones sont placés dans des pavillons artificiels d'un mannequin anthropomorphe. En conséquence, sélectionner en tant que mesure l'azimut qui maximise la fonction de vraisemblance présente un risque non négligeable de sélectionner celui qui ne correspond pas à la vraie direction de la source. Ceci peut entraîner de fortes perturbations dans l'estimation avec une perte de consistance du filtre.

Nous utilisons un algorithme **MG-UKF** qui peut être vu comme un ensemble de filtres **UKF** non interactifs (une version de l'**UKF** est donnée en annexe A). À un instant  $k$  particulier et après avoir assimilé l'observation  $y_k$ , l'algorithme est formé de  $I_k$  filtres, chaque  $i^{\text{ème}}$  filtre gérant une distribution gaussienne  $\mathcal{N}(x_k; \hat{x}_{k|k}^i, P_{k|k}^i)$  d'espérance et de covariance  $\hat{x}_{k|k}^i$  et  $P_{k|k}^i$ . Tous les filtres évoluent de manière indépendante mais leur probabilité *a posteriori* (ou « poids ») est récursivement mise à jour par leur vraisemblance vis à vis des observations. La *pdf a posteriori*  $p(x_k|y_{1:k})$  est définie comme le mélange de gaussiennes

$$p(x_k|y_{1:k}) = \sum_{i=1}^{I_k} w_k^i \mathcal{N}(x_k; \hat{x}_{k|k}^i, P_{k|k}^i), \quad (1.13)$$

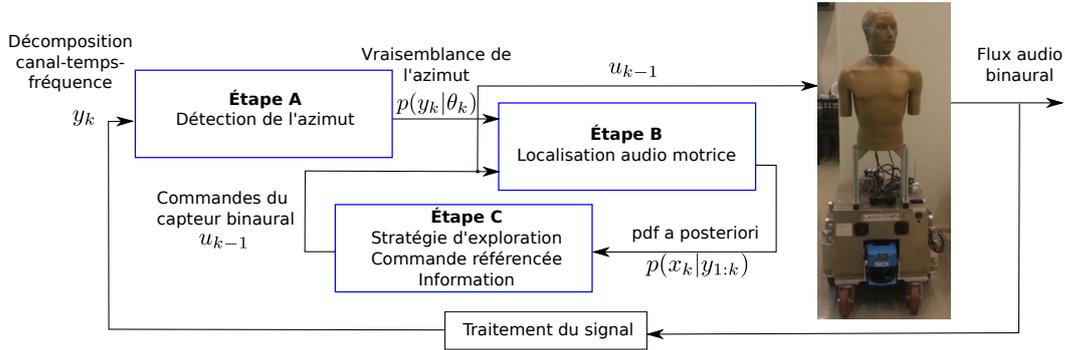


FIGURE 1.2 – Schéma de localisation en trois étapes.

avec  $I_k$  le nombre d'hypothèses et  $w_k^i$  la probabilité *a posteriori*, de chaque hypothèse du mélange.

### 1.3.4 Position du problème de synthèse de mouvement actif

Sur la base de la *pdf a posteriori*  $p(x_k | y_{1:k})$  de l'état  $\mathbf{x}_k$ , établie par la stratégie de localisation audio-motrice présenté en 1.3.3, nous recherchons les commandes à appliquer au capteur binaural qui conduisent à une réduction maximale de l'incertitude associée à la localisation. Le mouvement effectué par le capteur binaural, lorsque de telles commandes lui sont appliquées, correspond à la définition que nous donnons du mouvement actif : contrairement à des stratégies en boucles ouvertes où la commande à appliquer au capteur à chaque instant serait définie par l'utilisateur, la commande est ici établie en boucle fermée, sur la base de connaissances acquises sur l'état. Pour tout instant  $k$ , définissons la séquence des  $N$  futurs vecteurs de commande

$$\bar{u}_N = u_k, \dots, u_{k+N-1}, \quad (1.14)$$

qui correspondent aux mouvements finis appliqués au capteur entre les instants  $k$  et  $k+N$ . La *pdf*  $p(x_{k+N} | y_{1:k+N})$  correspond à la loi *a posteriori* à l'instant  $k+N$ , après l'application de ces  $N$  commandes et l'assimilation des futures variables observées  $y_{k+1:k+N} = y_{k+1}, \dots, y_{k+N}$ , c'est à dire les prochaines décompositions canal-temps-fréquence du signal perçu par les microphones. Ainsi la recherche de la séquence de vecteurs de commande  $\bar{u}_N$  qui conduit à la plus grande réduction de l'incertitude dans la localisation de la source, correspond par exemple à la minimisation d'un critère de coût lié à  $p(x_{k+N} | y_{1:k+N})$ . De manière équivalente, il peut aussi s'agir de la maximisation d'une fonction d'information qui serait liée à cette *pdf*. Comme un tel critère ne peut pas être évalué à l'instant  $k$  en l'absence des futures réalisations de  $y_{k+1:k+N}$ , une solution consiste à minimiser l'espérance sur les observations  $y_{k+1:k+N}$  de ce critère [Vlassis *et al.* 1999, Vincent *et al.* 2015, Nguyen *et al.* 2017]. L'absence d'un modèle d'observation analytique qui relie les variables d'observations  $y_{k+1}, \dots, y_{k+N}$  aux vecteurs d'état  $\mathbf{x}_{k+1}, \dots, \mathbf{x}_{k+N}$ , constitue ici une

difficulté majeure pour le calcul de cette espérance. Pour guider l'exploration, des variables d'observation alternatives  $z_{k+1:k+N} = z_{k+1}, \dots, z_{k+N}$  (qui peuvent donc différer de  $y_{k+1:k+N}$ ) sont alors définies. Dans l'approche que nous proposons dans cette thèse, nous définissons la variable d'observation  $\mathbf{z}_k$  (dont  $z_k$  est une réalisation à l'instant  $k$ ), reliée au vecteur d'état caché  $\mathbf{x}_k$  et à l'azimut qui lui est associé  $\boldsymbol{\theta}_k = -\text{atan2}(\mathbf{e}_{y_k}, \mathbf{e}_{z_k})$ , par un modèle d'observation scalaire de la forme suivante

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k = \bar{h}(\boldsymbol{\theta}_k) + \mathbf{v}_k, \quad \mathbf{z}_k \in \mathbb{R}, \quad \mathbf{v}_k \sim \mathcal{N}(0, R_k), \quad (1.15)$$

avec  $\mathbf{v}_k$  le bruit de mesure, supposé additif, gaussien centré et de variance  $R_k$ . Nous supposons que la dispersion du bruit de mesure peut varier dans le temps mais ne dépend pas de la valeur prise par l'état caché. Nous discuterons des implications d'une telle hypothèse dans la section 3.2. La variable d'observation  $\mathbf{z}_k$  est liée à l'azimut  $\boldsymbol{\theta}_k$  de la source mais reste indépendante de la distance capteur-source. Par exemple dans le cas d'une source en champ lointain,  $\bar{h}(\boldsymbol{\theta}_k)$  peut exprimer l'approximation des ITD de Woodworth-Schlosberg pour le cas où les microphones sont disposés sur une tête sphérique [Aaronson & Hartmann 2014].

L'assimilation des réalisations  $y_{1:k}$  des observations  $\mathbf{y}_{1:k}$  lorsqu'une séquence de vecteurs de commande  $u_{0:k-1}$  est appliquée conduit à la pdf *a posteriori*  $p(x_k|y_{1:k})$ . Pour alléger les notations dans l'étape d'exploration définie à partir de l'instant  $k$ , on désigne indifféremment cette loi par  $p(x_k|z_{1:k})$ . Ensuite, l'assimilation de  $z_{k+1:k+N}$  à la place de  $y_{k+1:k+N}$  à des fins d'exploration, combinée aux  $N$  prochaines entrées de commande  $\bar{u}_N = u_k, \dots, u_{k+N-1}$ , permet d'établir la pdf  $p(x_{k+N}|z_{1:k+N})$ . Dans les développements du chapitre 2, nous définissons un critère d'information lié à cette loi. La séquence de commande  $\bar{u}_N$  doit donc être calculée de façon à atteindre en moyenne, un maximum de ce critère à la fin des  $N$  pas (*i.e.* à l'instant  $k+N$ ).

L'ensemble du dispositif de localisation comporte donc trois étapes : (A) estimation à court terme de l'azimut de la source et calcul de la fonction de vraisemblance  $p(y_k|\boldsymbol{\theta}_k)$ ; (B) combinaison du résultat de (A) avec les commandes du capteur binaural dans une stratégie de filtrage stochastique, de façon à obtenir la pdf *a posteriori*  $p(x_k|y_{1:k})$ ; (C) partant de cette loi, stratégie d'exploration guidée par des observations  $z_{k+1:k+N}$  relatives à un modèle d'observation 1.15, permettant d'établir la séquence des vecteurs de commande  $\bar{u}_N$  qui maximise l'information sur l'état, *i.e.* qui minimise l'incertitude dans la localisation de la source. L'ensemble du processus est schématisé sur la figure 1.2.

Soit  $\bar{u}_N^*$  la séquence des  $N$  prochains vecteurs de commande du capteur binaural, qui conduit à la maximisation d'un critère d'information lié à la pdf  $p(x_{k+N}|z_{1:k+N})$ . Une stratégie de commande en boucle fermée à « horizon glissant », peut être adoptée. Cela consiste à appliquer le premier élément  $u_k^*$  de la séquence  $\bar{u}_N^*$  calculée. Ensuite une nouvelle observation  $y_{k+1}$  est disponible. Son assimilation mène à la pdf  $p(x_{k+1}|y_{1:k+1})$  de l'instant  $k+1$  et l'ensemble de la procédure peut être répété. Se pose alors la question du nombre de pas à prévoir dans la stratégie d'exploration. Intuitivement, un  $N$  suffisamment élevé apportera un bénéfice car la planification sera effectuée sur un horizon plus grand. Cela entraînera en retour un coût calculatoire

plus élevé. La recherche d'un compromis dans le choix de l'horizon d'exploration semble donc nécessaire. D'autant plus qu'intuitivement, une augmentation de  $N$  à partir d'un certain seuil, n'aura que peu d'effet sur la localisation.

## 1.4 Présentation des outils d'évaluation

L'objectif de cette thèse est de développer des approches de mouvement actif à court et long terme pour la localisation binaurale, mais aussi de valider les résultats obtenus par des simulations réalistes et des expériences sur un robot binaural réel. Ces travaux constituent une contribution au projet TWO!EARS. Il s'agit d'un programme de recherche européen achevé fin 2016 ([twoears.eu](http://twoears.eu)), dont l'objectif était de dresser un modèle computationnel complet de la perception auditive humaine, dans sa dimension active (*via* la mise en évidence de boucles de rétroaction multiples, au niveau réflexe mais également en provenance de niveaux cognitifs) tout en permettant l'incorporation de la multimodalité. Le projet avait pour but de développer une plateforme robotique auditive permettant d'implémenter ce modèle et de le confronter à des cas d'étude réalistes. J'ai eu la chance de bénéficier de cette plateforme matérielle et des logiciels développés dans ce projet. L'implémentation sur ce robot des algorithmes développés plus loin dans ce manuscrit, s'appuie sur de nombreuses bibliothèques codées par des collègues impliqués en tant qu'ingénieurs développeurs dans le projet, que je remercie vivement ici.

Ainsi les expériences ont pu être effectuées avec un mannequin ([Head And Torso Simulator \(HATS\)](#)) KEMAR (figure 1.4), manufacturé par G.R.A.S.<sup>®</sup> ([kemar.us](http://kemar.us)). Il est équipé de deux pavillons dans lesquels sont logés des microphones omnidirectionnels. Pour les aspects simulation que nous présentons dans la partie 1.4.1, nous supposons que la tête du mannequin est liée au torse et que celui-ci est mobile en translation et rotation. Dans une scène virtuelle, lorsque le [HATS](#) se déplace, le signal perçu par le capteur binaural est généré par rapport à la position relative capteur-source (en considérant que la pièce est anéchoïque). Pour les expériences en conditions réelles présentées dans la partie 1.4.2, le [HATS](#) est placé sur une base robotique mobile et la tête est munie d'un degré de liberté permettant une rotation autour de l'axe du cou. Enfin, en simulation et en conditions réelles, la stratégie complète de localisation présentée en 1.3.4 est appliquée.

### 1.4.1 Simulateur binaural

Afin de permettre l'évaluation réaliste d'un algorithme de localisation active binaurale, les signaux perçus par les microphones gauche et droit doivent être générés en tenant compte des positions du [HATS](#) et de la source, ainsi que des caractéristiques de l'environnement dans lequel ils se situent. Une difficulté particulière vient du fait que les signaux binauraux perçus dépendent de la position du mannequin qui elle-même peut évoluer en ligne en fonction de ces signaux. On considère un environnement anéchoïque. Sur la base du signal engendré par la source en un point de référence, les signaux binauraux perçus sont reconstitués par exploitation

des réponses impulsionnelles de la tête (HRIR) en les microphones gauche et droit du mannequin KEMAR. Les HRIR ont été identifiées expérimentalement dans un environnement anéchoïque. La base de données est disponible publiquement sur [www.twoears.eu](http://www.twoears.eu). La fonction de pseudo-vraisemblance de l'azimut de la source est calculée à chaque instant de la simulation à partir des signaux binauraux simulés. Cette information est ensuite combinée au vecteur de commande  $u_k$  du capteur binaural, au moyen du MG-UKF présenté précédemment. Selon les cas,  $u_k$  peut être une commande en boucle ouverte déterminée à l'avance (déplacement en translation, rotation, etc.), ou bien calculée au moyen d'une stratégie d'exploration en boucle fermée sur la perception faisant l'objet de la thèse. Un exemple de localisation pour le cas d'un mouvement circulaire du capteur binaural est présenté en figure 1.3.

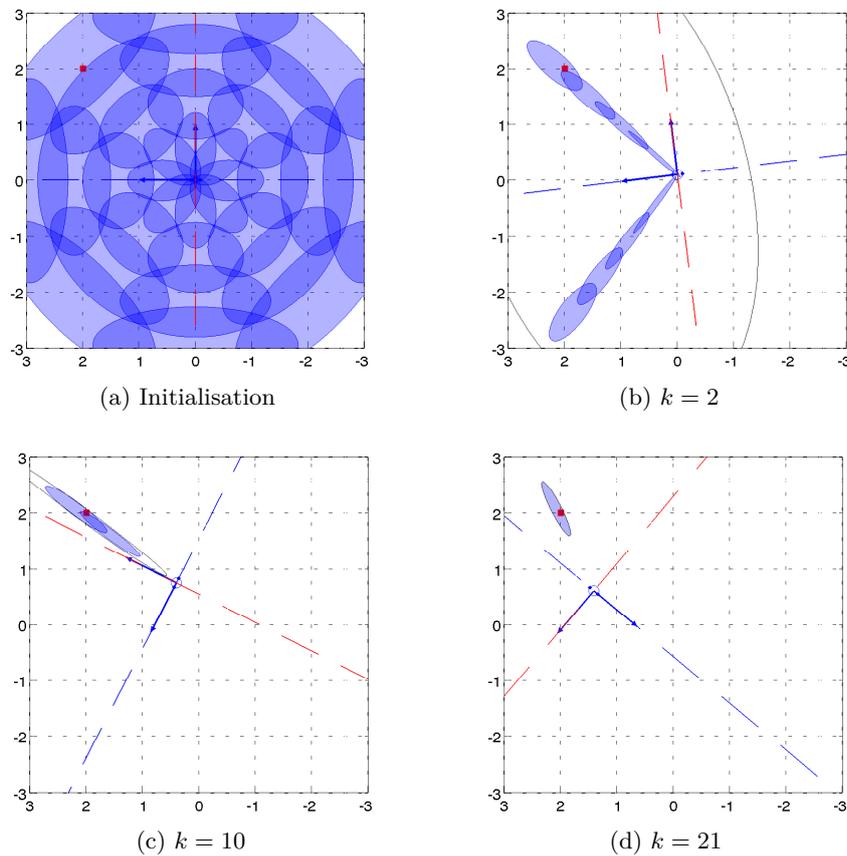


FIGURE 1.3 – Exemple de localisation à différents instants de simulation – Sont représentés : (1) le repère du capteur binaural (axe interaural en pointillés bleu et axe dans la direction de l'avant de la tête en rouge) ; (2) la position réelle de la source (carré rouge) ; (3) les ellipses de confiance à 99% de probabilité de chaque hypothèse constituant le mélange de gaussiennes.

### 1.4.2 Localisation à partir d'un robot mobile

Le **HATS KEMAR** a été monté sur un robot mobile non holonome<sup>1</sup> **NEOBOTIX MP-L655**. Cette base mobile peut avancer, reculer et tourner, de façon similaire à une voiture, i.e., en demeurant tangente à sa trajectoire. Pour les besoins expérimentaux, l'axe de la tête a été motorisé de façon à lui conférer un degré de liberté contrôlable en azimut. Couplée au mouvement de la base, la rotation de la tête permet de conférer au capteur binaural un mouvement omnidirectionnel de la façon suivante. Soit  $(v_y, v_z, w_x)$  les vitesses de translation et de rotation du repère binaural avec  $v_y, v_z$  les vitesses de translation selon les axes  $\vec{y}_R, \vec{z}_R$  et  $w_x$  la vitesse de rotation autour de  $\vec{x}_R$ , qui peuvent être déduites de  $u = (T_y, T_z, \phi)^T$ . Supposons que le **HATS** soit positionné sur la base mobile de telle manière que l'axe vertical de rotation de la tête soit : d'une part perpendiculaire au plan contenant l'axe des roues directrices de la base, d'autre part déporté d'une distance  $D$  de sorte que chaque point de l'axe soit équidistant des roues. En d'autres termes, nous supposons que le **HATS** n'est pas placé au dessus des roues directrices de la base mobile. Soient  $v_{base}$  la vitesse linéaire de la base tangente à sa trajectoire et  $w_{base}$  la vitesse de rotation autour de l'axe vertical de rotation de la base. La position et la vitesse angulaire de la tête **KEMAR** par rapport au torse et à la base, sont respectivement notées  $q$  et  $w_{tête}$ . Alors, lorsque  $D \neq 0$  (i.e., le **HATS** n'est pas placé au dessus des roues motrices de la base), il est montré qu'une combinaison des commandes de la base et de la rotation de la tête  $(v_{base}, w_{base}, w_{tête})$  peut être obtenue pour n'importe quelles vitesses de translation et rotation  $(v_y, v_z, w_x)$  imposées au capteur, *via* la relation

$$\begin{bmatrix} v_{base} \\ w_{base} \\ w_{tête} \end{bmatrix} = \begin{bmatrix} -\sin(q) & D \cos(q) & 0 \\ \cos(q) & D \sin(q) & 0 \\ 0 & -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} v_y \\ v_z \\ w_x \end{bmatrix}. \quad (1.16)$$

La démonstration de cette relation est disponible dans [Cadenat 1999]. Ainsi la combinaison de la rotation de la tête et du déplacement de la base permet de rendre le déplacement du capteur binaural omnidirectionnel. Dans la suite du manuscrit, nous considérerons toujours les commandes  $u_k$  du capteur binaural (translation et rotation du capteur binaural), en les confondant parfois par abus de langage avec les commandes réelles du robot.

L'architecture logicielle a été déployée sur l'intergiciel (middleware) **Robot Operating System (ROS)**, dédié à la robotique. **ROS** fournit des outils, des bibliothèques et des conventions qui permettent de simplifier et d'uniformiser le développement logiciel des robots. L'architecture est organisée en nœuds, processus indépendants qui peuvent communiquer entre eux par des sortes de tunnels appelés *topics*. Un nœud qui exporte des données « publie » sur un *topic* alors qu'un nœud qui reçoit des données en entrée « souscrit » à un *topic*. Les nœuds fournissent éventuellement des services qui permettent de les contrôler. Le développement des différents

1. Un robot ou un véhicule holonome a trois degrés de liberté : deux en translations et un en rotation.

nœuds a été effectué avec GenoM3 [Mallet *et al.* 2010]. L'intérêt réside dans le fait que GenoM3 permet une spécification des composants (services proposés, ports d'entrée-sortie, automates des activités, gestion des échecs, etc.) indépendante du middleware sous-jacent, et la génération automatique du code temps réel dans lequel est ensuite incorporé le contenu algorithmique. Cette séparation claire permet un code durable et l'élimination de nombreuses erreurs. Trois modules ont été développés avec GenoM3 :

- Le serveur audio Binaural. Il s'agit du composant logiciel qui récupère le flux audio perçu par le capteur binaural. La transformée de Fourier à court terme est appliquée sur des segments temporels du signal brut, de façon à obtenir la décomposition canal-temps-fréquence  $y_k$  du signal.
- Le schéma de localisation. Ce composant correspond au processus de localisation en 3 étapes que nous mettons en place. Il contient donc le calcul de la fonction de pseudo-vraisemblance de l'azimut à partir de  $y_k$ , la localisation audio-motrice qui combine l'estimation de l'azimut avec la commande du capteur binaural et une fonction qui calcule les commandes du capteur binaural. Ces commandes peuvent être déterminées à l'avance par l'utilisateur (translation, rotation, etc.) ou calculées au moyen de processus de mouvement actif qui font l'objet de nos recherches. Ces commandes correspondent au déplacement omnidirectionnel du capteur binaural décrit par l'équation de dynamique (1.4).
- Les commandes de rotation de la tête. Ce module permet de contrôler le moteur responsable de la rotation de la tête.

Le module qui contrôle les moteurs de la base n'a pas été développé avec GenoM3, s'agissant d'un nœud natif ROS. Un processus exécuté par un client MATLAB ou Tcl, permet d'organiser l'exécution de l'ensemble du programme. C'est-à-dire le démarrage ou l'arrêt des différents modules et des services de communications, la gestion et la sauvegarde des données, etc. L'architecture détaillée est donnée en figure 1.4.

Les expériences ont été conduites en intérieur dans une zone ouverte de  $15m \times 5m$  et de hauteur  $8m$ , délimitée par des murs en résine. Dans toutes les expériences, la source sonore a été placée sur un trépied de façon à se situer à la même hauteur que les microphones.

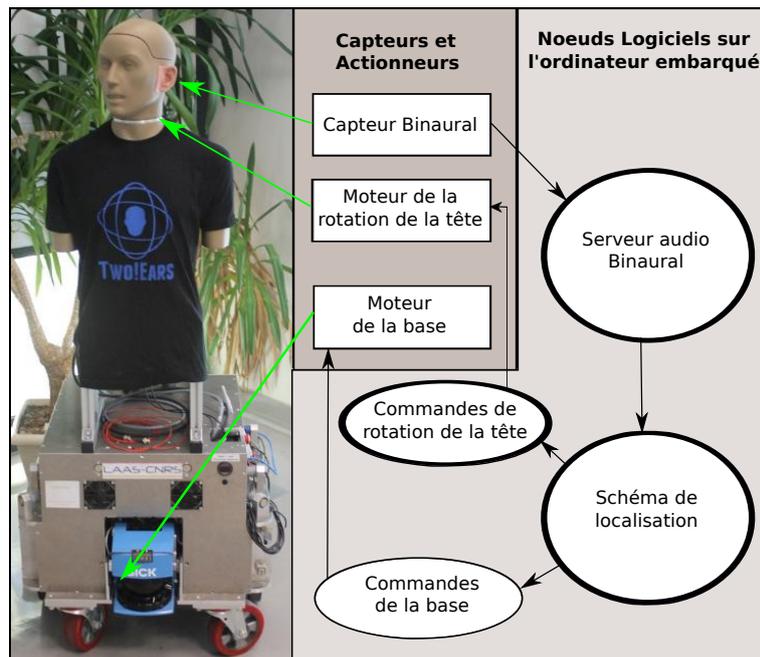


FIGURE 1.4 – Plateforme robotique et vue schématique de son architecture logicielle – À gauche, le mannequin KEMAR disposé sur la base mobile NEOBOTIX MP-L655. À droite, l'architecture logicielle. Les nœuds logiciels représentés en traits gras sont ceux développés avec GenoM3. Le module *Commandes de la base* est un nœud natif ROS. Le serveur audio binaural publie les données provenant du capteur binaural. Le nœud de localisation exécute la stratégie de localisation en 3 étapes et publie les commandes de locomotion de la base et de rotation de la tête.

# Outils théoriques pour la commande référencée information

---

## Sommaire

<b>2.1 Définitions de l'information</b>	<b>28</b>
2.1.1 Estimation et mesure d'information	28
2.1.2 Entropie et Information Mutuelle	30
2.1.3 Entropie et filtrage bayésien	32
<b>2.2 Perception active en robotique</b>	<b>35</b>
2.2.1 Définition d'un modèle	36
2.2.2 Choix du critère d'information	37
2.2.3 Décisions à court terme contre décisions à long terme	42
<b>2.3 Fonction d'incertitude pour la localisation active de sources sonores</b>	<b>45</b>
2.3.1 Hypothèses du modèle d'exploration	45
2.3.2 Définition d'une fonction d'incertitude	47
A Récompense immédiate	48
B Critère de décision à long terme	50

---

Afin de commander le capteur binaural pour améliorer le processus d'estimation de la position de la source, il est nécessaire de quantifier l'information associée à tout estimé. À tout instant du processus de localisation, la quantité d'information associée à un estimé futur dépend des prochaines commandes appliquées au capteur binaural et des futures observations. La commande du capteur binaural qui améliore le processus d'estimation peut donc être obtenue en maximisant l'espérance de cette quantité d'information sur les observations inconnues, sous réserve de définir un critère d'information et que celui-ci puisse être calculé. Plus généralement, les problèmes de commande optimale stochastique où l'objectif est formulé en terme de maximisation d'une fonction d'information, constituent le champ de la perception active.

Une première partie de ce chapitre présente différents concepts permettant de mesurer l'information contenue dans les estimateurs ou associée aux variables aléatoires. Nous étudions ensuite le lien entre ces mesures et le problème de filtrage. Nous présenterons ensuite le problème de perception active en robotique et comment

les différents moyens de mesurer l'information peuvent être utilisés pour formaliser des problèmes de commande. Dans ce type de problèmes, une fonction de récompense (ou d'incertitude), liée à l'information accumulée sur une variable d'intérêt est établie et la résolution consiste à déterminer la loi de commande qui maximise (ou minimise) cette fonction. Dans la dernière section de ce chapitre, nous proposerons un critère d'information qui constituera le cadre théorique des développements effectués dans le manuscrit, pour guider le mouvement d'un capteur binaural lors de la localisation audio-motrice.

## 2.1 Définitions de l'information

Dans cette section nous introduisons des outils permettant de mesurer l'information relative à un paramètre ou à une variable aléatoire. Nous présentons en section 2.1.1, l'information de Fisher qui permet de quantifier l'information qu'un vecteur aléatoire contient à propos d'un paramètre déterministe inconnu. Nous verrons également que cette définition peut être étendue au cas de l'estimation bayésienne, où les paramètres à estimer sont supposés être les réalisations de variables aléatoires de statistiques données. Nous présentons ensuite en section 2.1.2 la définition de l'entropie d'une variable aléatoire, qui est une mesure d'information introduite par Shannon [Shannon 1948]. Nous étudions enfin en 2.1.3 les relations entre l'entropie et le filtrage bayésien.

### 2.1.1 Estimation et mesure d'information

Considérons le problème d'estimation d'un vecteur de paramètres  $\theta$ . Soit  $\mathbf{z}$  une variable aléatoire qui dépend du vecteur de paramètres  $\theta$  et soit  $z$  une réalisation de  $\mathbf{z}$ . Le problème d'estimation de  $\theta$  consiste à construire un estimateur  $\hat{\theta}$ , fonction de  $\mathbf{z}$  de sorte que  $\hat{\theta} = g(\mathbf{z})$ . L'estimé  $\hat{\theta}$  de  $\theta$  est alors défini par  $\hat{\theta} = g(z)$ . Soit  $p(z|\theta)$  la fonction de vraisemblance de  $\theta$  définie comme la pdf de  $z$  conditionnellement à  $\theta$ . L'estimation de  $\theta$  peut être envisagée par deux approches :

- L'approche d'estimation classique, où aucune connaissance *a priori* sur  $\theta$  n'est disponible. En conséquence,  $\theta$  est déterministe inconnu. La méthode d'estimation par maximum de vraisemblance est une approche qui consiste à rechercher l'argument qui maximise  $p(z|\theta)$  avec

$$\hat{\theta} = \arg \max_{\theta} p(z|\theta), \tag{2.1}$$

où  $\hat{\theta}$  est l'estimé du maximum de vraisemblance.

- L'approche bayésienne, où une connaissance *a priori* sur la valeur du paramètre est disponible. Afin de prendre en compte cette connaissance, le vecteur de paramètres  $\theta$  que nous souhaitons estimer est vu comme une réalisation d'un vecteur aléatoire  $\boldsymbol{\theta}$ . La loi jointe  $p(\boldsymbol{\theta}, z)$  de  $\boldsymbol{\theta}$  et  $\mathbf{z}$  est donc disponible, et s'exprime comme le produit de la loi a priori  $p(\boldsymbol{\theta})$  et du modèle d'observation

$p(z|\theta)$ . Sur cette base, la loi *a posteriori*  $p(\theta|z)$  qui capture toute l'information sur  $\theta$  contenue dans la mesure  $\mathbf{z}$ , est déterminée par

$$p(\theta|z) = \frac{p(z|\theta)p(\theta)}{p(z)}, \quad (2.2)$$

avec  $p(z)$  une constante de normalisation qui ne dépend pas de  $\theta$ . Une approche d'estimation de  $\theta$  peut consister à établir l'estimé du maximum *a posteriori*  $\hat{\theta}$  défini par

$$\hat{\theta} = \arg \max_{\theta} p(z|\theta)p(\theta). \quad (2.3)$$

Dans les deux approches, l'estimateur  $\hat{\theta}$  est par construction un vecteur aléatoire dont la réalisation est  $\hat{\theta}$  dès lors que  $\mathbf{z}$  se réalise en  $z$ .

Avec le modèle d'estimation classique, la Borne de Cramér-Rao (**Cramér-Rao Lower Bound (CRLB)**) constitue une borne inférieure de la covariance de l'erreur d'estimation de tout estimateur non biaisé de  $\theta$  [Bar-Shalom & Li 1998]. La définition 2.1 peut ainsi être établie.

**Définition 2.1** (Borne Inférieure de Cramér-Rao et FIM). *La matrice de covariance de tout estimateur non biaisé  $\hat{\theta}$  d'un vecteur de paramètres déterministe inconnu  $\theta$  est bornée inférieurement par la **CRLB** de sorte que*

$$\mathbb{E}_{\mathbf{z}} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \geq J^{-1} = \text{CRLB}(\theta), \quad (2.4)$$

où la relation d'ordre est liée à la semi-définie positivité des matrices (pour  $A = \mathbb{E}_{\theta} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right]$  et  $B = J^{-1}$ ,  $A \geq B$  exprime la semi-définie positivité de  $A - B$ ).  $J$  est la Matrice d'Information de Fisher (**Fisher Information Matrix (FIM)**) calculée à partir de la fonction de vraisemblance  $p(z|\theta)$  du paramètre  $\theta$  par

$$J \triangleq \mathbb{E}_{\mathbf{z}} \left[ (\nabla_{\theta} \ln p(z|\theta))(\nabla_{\theta} \ln p(z|\theta))^T \right]^1 \quad (2.5)$$

$$= - \mathbb{E}_{\mathbf{z}} \left[ \nabla_{\theta} \nabla_{\theta}^T \ln p(z|\theta) \right]. \quad (2.6)$$

$\nabla_{\theta}$  est l'opérateur gradient et  $\nabla_{\theta} \nabla_{\theta}^T$  l'opérateur formant la matrice Hessienne.

La **FIM** est donc l'inverse de la **CRLB** pour un estimateur non biaisé. Dans les équations (2.4), (2.5) et (2.6) l'espérance est calculée par rapport à  $\mathbf{z}$ . La démonstration de l'équivalence des équations (2.5) et (2.6) est détaillée dans [Kay 1993].

Les définitions de la **CRLB** et de la **FIM** peuvent être étendues aux problèmes d'estimation traités par une approche bayésienne. Lorsque le paramètre  $\theta$  à estimer est aléatoire avec une **pdf**  $p(\theta)$  connue *a priori* et que la loi jointe  $p(\theta, z)$  est définie, la covariance de tout estimateur non biaisé de  $\theta$  est bornée inférieurement par

1. Afin de simplifier les notations, l'espérance d'une fonction  $f$  appliquée à une variable aléatoire  $\mathbf{x}$  sera indifféremment désignée par la notation classique  $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})]$  ou par la notation  $\mathbb{E}_{\mathbf{x}}[f(x)]$ , où la variable muette  $x$  exprime toute réalisation de  $\mathbf{x}$ .

l'inverse d'une matrice d'information  $J_T$  dans une expression similaire à (2.4) :

$$\mathbb{E}_{\mathbf{z},\theta} [(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] \geq J_T^{-1}, \quad (2.7)$$

avec

$$J_T \triangleq \mathbb{E}_{\mathbf{z},\theta} [(\nabla_{\theta} \ln p(z, \theta))(\nabla_{\theta} \ln p(z, \theta))^T] \quad (2.8)$$

$$= -\mathbb{E}_{\mathbf{z},\theta} [\nabla_{\theta} \nabla_{\theta}^T \ln p(z, \theta)]. \quad (2.9)$$

$J_T^{-1}$  constitue alors la Borne de Cramér-Rao *a posteriori* (**Posterior Cramér-Rao Bound (PCRB)**) [Tichavský *et al.* 1998]. Dans les équations (2.7), (2.8) et (2.9) l'espérance est effectuée sur  $\mathbf{z}$  et  $\theta$ .

### 2.1.2 Entropie et Information Mutuelle

Nous présentons dans cette section la mesure d'entropie et d'autres concepts qui s'y rapportent, tous issus des travaux sur l'information développés par Shannon. Une présentation plus complète de la notion d'entropie et de la théorie de l'information est fournie dans [Cover & Thomas 1991]. Historiquement l'entropie de Shannon concerne les variables aléatoires discrètes. Une extension de la définition première, nommée entropie « différentielle » et permettant d'étudier des variables aléatoires continues, a ensuite été introduite. Les variables que nous traitons dans cette thèse étant continues, nous présenterons directement la forme différentielle de l'entropie et nous omettrons le terme « différentiel », par souci de simplicité.

**Définition 2.2** (Entropie). *L'entropie  $H(\mathbf{x})$  d'une variable aléatoire continue  $\mathbf{x}$ , associée à la pdf  $p(x)$ , est définie par*

$$H(\mathbf{x}) = - \int p(x) \ln p(x) dx \quad (2.10)$$

$$= -\mathbb{E}_{\mathbf{x}} [\ln p(x)], \quad (2.11)$$

avec  $\ln$  le logarithme naturel de base  $e$ <sup>2</sup>.

L'entropie est une mesure d'incertitude d'une variable ou d'un vecteur aléatoire. Pour simplifier notre propos, nous désignerons parfois l'entropie d'une variable aléatoire  $\mathbf{x}$  par l'entropie de la pdf  $p(x)$  associée à cette variable. Lorsqu'un vecteur aléatoire  $\mathbf{x} \in \mathbb{R}^n$  admet une distribution gaussienne de loi  $\mathcal{N}(x; \mu, \Sigma)$ , il est montré dans [Cover & Thomas 1991] que l'entropie de  $\mathbf{x}$  s'écrit comme une fonction croissante monotone du logarithme du déterminant de la matrice de covariance  $\Sigma$  :

$$H(\mathbf{x}) = \frac{1}{2} \ln ((2\pi e)^n |\Sigma|). \quad (2.12)$$

---

2. L'entropie est dans ce cas exprimée en « nats ». Initialement, le logarithme en base 2 était utilisé et l'unité était le « bit ». Par la suite nous garderons le logarithme népérien et nous ne préciserons plus sa nature.

Ainsi l'entropie de  $\mathbf{x}$  gaussien est indépendante de l'espérance de  $\mathbf{x}$ . Dans ce qui suit, nous définissons également les entropies conditionnelles et jointes dans les définitions 2.3 et 2.4. Puis nous introduisons la notion d'information mutuelle (définition 2.6). Lorsque nous considérons un ensemble de variables aléatoires, nous supposons que la loi jointe de ces variables est définie.

**Définition 2.3** (Entropie conditionnelle). *Soient  $\mathbf{x}$  et  $y$  deux variables aléatoires continues de pdf jointe  $p(x, y)$  et soit  $p(x|y)$  la pdf de  $\mathbf{x}$  conditionnellement à l'événement  $y = y$ . L'entropie conditionnelle  $H_c(\mathbf{x}|y)$  est alors exprimée par*

$$H_c(\mathbf{x}|y) = - \int p(x, y) \ln p(x|y) dx dy \quad (2.13)$$

$$= -\mathbb{E}_{\mathbf{x}, y} [\ln p(x|y)]. \quad (2.14)$$

Notons que la définition (2.2) de l'entropie et l'entropie conditionnelle peuvent être liées par

$$\begin{aligned} H_c(\mathbf{x}|y) &= -\mathbb{E}_{\mathbf{x}, y} [\ln p(x|y)] \\ &= \mathbb{E}_y \left[ -\mathbb{E}_{\mathbf{x}|y} \ln p(x|y) \right] \\ &= \mathbb{E}_y [H(\mathbf{x}|y)], \end{aligned} \quad (2.15)$$

avec  $H(\mathbf{x}|y)$  l'entropie de la pdf conditionnelle  $p(x|y)$  conformément à la définition (2.2).

**Définition 2.4** (Entropie jointe). *L'entropie jointe d'une paire de variables aléatoires  $\{\mathbf{x}, y\}$  de pdf  $p(x, y)$ , est exprimée par*

$$H(\mathbf{x}, y) = - \int p(x, y) \ln p(x, y) dx dy \quad (2.16)$$

$$= -\mathbb{E}_{\mathbf{x}, y} \{\ln p(x, y)\}. \quad (2.17)$$

L'entropie jointe de deux variables aléatoires est égale à l'entropie de l'une de ces variables additionnée à l'entropie conditionnelle de l'autre de sorte que

$$H(\mathbf{x}, y) = H(\mathbf{x}) + H_c(y|\mathbf{x}) = H(y) + H_c(\mathbf{x}|y). \quad (2.18)$$

**Définition 2.5** (Divergence de Kullback Leibler). *Soit deux pdf  $f$  et  $g$ , la divergence de Kullback Leibler (ou entropie relative)  $D(f||g)$  est définie par*

$$D(f||g) = \int f \ln \frac{f}{g}. \quad (2.19)$$

$D(f||g) \geq 0$  et  $D(f||g) = 0$  si et seulement si  $f = g$ .

Bien qu'elle ne soit pas commutative, la divergence de Kullback Leibler est considérée comme une mesure de la distance entre deux distributions.

**Définition 2.6** (Information mutuelle). *L'information mutuelle  $I(\mathbf{x}, \mathbf{y})$  d'un couple de variables aléatoires continues  $\{\mathbf{x}, \mathbf{y}\}$  de pdf jointe  $p(x, y)$ , est exprimée par*

$$I(\mathbf{x}, \mathbf{y}) = \int p(x, y) \ln \frac{p(x, y)}{p(x)p(y)} dx dy \quad (2.20)$$

$$= \mathbb{E}_{\mathbf{x}, \mathbf{y}} \left[ \ln \frac{p(x, y)}{p(x)p(y)} \right]. \quad (2.21)$$

L'information mutuelle est une mesure de dépendance entre deux variables aléatoires. C'est une mesure de la quantité d'information qu'une variable contient sur une autre. L'information mutuelle  $I(\mathbf{x}, \mathbf{y})$  représente l'intersection de l'information de  $\mathbf{x}$  avec l'information de  $\mathbf{y}$ . C'est une quantité toujours positive, ou nulle si et seulement si les deux variables aléatoires considérées sont indépendantes. L'information mutuelle représente la réduction d'entropie d'une variable aléatoire lorsque la connaissance d'une autre variable non indépendante, est prise en compte. Cette relation correspond plus précisément à

$$I(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}) - H_c(\mathbf{x}|\mathbf{y}) = H(\mathbf{y}) - H_c(\mathbf{y}|\mathbf{x}). \quad (2.22)$$

La démonstration de (2.22) peut être trouvée dans [Cover & Thomas 1991]. Notons que  $I(\mathbf{x}, \mathbf{y}) = 0$ ,  $H(\mathbf{x}) = H_c(\mathbf{x}|\mathbf{y})$  et  $H(\mathbf{y}) = H_c(\mathbf{y}|\mathbf{x})$  si et seulement si les variables  $\mathbf{x}$  et  $\mathbf{y}$  sont indépendantes. L'information mutuelle peut également être définie au moyen de la divergence de Kullback Leibler par la relation

$$I(\mathbf{x}, \mathbf{y}) = D(p(x, y) || p(x)p(y)). \quad (2.23)$$

### 2.1.3 Entropie et filtrage bayésien

Considérons des processus aléatoires d'état et d'observation à temps discret, respectivement représentés à l'instant  $k$  par le vecteur aléatoire  $\mathbf{x}_k$  (caché) et le vecteur aléatoire  $\mathbf{z}_k$  (dont la réalisation est accessible). Notons  $p(x_0)$  la distribution initiale du vecteur d'état. Les modèles de dynamique et d'observation sont définis par

$$\begin{cases} \mathbf{x}_{k+1} &= f(\mathbf{x}_k, u_k) + \mathbf{w}_k \\ \mathbf{z}_k &= h(\mathbf{x}_k) + \mathbf{v}_k, \end{cases} \quad (2.24)$$

avec  $f$  et  $h$  des fonctions non linéaires. Les variables  $\mathbf{w}_k$  et  $\mathbf{v}_k$  représentent respectivement les bruits additifs de dynamique et d'observation qui sont blancs, mutuellement indépendants et indépendants de l'état initial. De (2.24) vont être déduits la loi de transition de l'état  $p(x_{k+1}|x_k)$  ainsi que le modèle d'observation  $p(z_k|x_k)$  qui est la fonction de vraisemblance de  $x_k$  par rapport au vecteur d'observation  $z_k$ . Notons que le processus d'état est markovien, c'est à dire qu'à tout instant  $k$ , la loi de l'état futur  $\mathbf{x}_{k+1}$  conditionnellement aux états présent et passés  $\{\mathbf{x}_l\}_{l \leq k}$  ne dépend que de l'état présent, *i.e.*,  $p(x_{k+1}|x_{0:k}) = p(x_{k+1}|x_k)$ . Soit  $p(x_k|z_{1:k})$  la

pdf de la variable d'état  $\mathbf{x}_k$  conditionnellement au fait que les observations  $\mathbf{z}_{1:k}$  se réalisent en la séquence de mesures  $z_{1:k}$ . Cette loi est alors qualifiée de pdf *a posteriori* ou de densité de filtrage et dépend implicitement des différentes commandes  $u_{0:k-1} = u_0, \dots, u_{k-1}$  appliquées au système depuis le début du processus jusqu'à l'instant considéré. Définissons également la loi de prédiction  $p(x_k|z_{1:k-1})$ , décrivant la pdf de l'état caché conditionnellement aux mesures jusqu'à l'instant  $k-1$ , avant l'assimilation de l'observation  $z_k$ .

Le filtrage bayésien consiste à calculer récursivement la densité de filtrage  $p(x_k|z_{1:k})$ , à partir des commandes appliquées au système et de l'assimilation des observations. Ce calcul peut être séparé en une étape de prédiction et de mise à jour. L'étape de prédiction consiste à calculer la loi  $p(x_k|z_{1:k-1})$  à partir de la densité de filtrage à l'instant  $k-1$  et de la loi de transition de l'état, au moyen de l'équation de Chapman-Kolmogorov

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}. \quad (2.25)$$

L'étape de mise à jour consiste à calculer la densité de filtrage  $p(x_k|z_{1:k})$  à partir de la loi de prédiction et de l'assimilation de l'observation  $z_k$ , au moyen de la règle de Bayes

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}. \quad (2.26)$$

Ci-dessus,  $p(z_k|z_{1:k-1})$  est une constante de normalisation car elle ne dépend pas des valeurs de la variable  $x_k$  de la loi de filtrage recherchée. La loi  $p(z_k|z_{1:k-1})$  de prédiction de la mesure à l'instant  $k$  conditionnellement aux observations passées  $z_{1:k-1}$ , dont l'évaluation en  $z_k$  constitue cette constante, satisfait

$$p(z_k|z_{1:k-1}) = \int p(z_k|x_k)p(x_k|z_{1:k-1})dx_k. \quad (2.27)$$

Définissons également  $p(x_k, z_k|z_{1:k-1})$  la pdf jointe de l'état et de l'observation, conditionnellement aux observations passées. Lorsque les vecteurs d'état  $\mathbf{x}_{0:k}$  et d'observation  $\mathbf{z}_{1:k}$  sont gaussiens dans leur ensemble et que les fonctions  $f$  et  $h$  du modèle (2.24) sont linéaires, les moments de la densité de filtrage peuvent être calculés récursivement en appliquant le filtre de Kalman. Dans ce cas,  $p(x_k|z_{1:k})$  est la loi normale  $\mathcal{N}(x; \hat{x}_{k|k}, P_{k|k})$  de moyenne  $\hat{x}_{k|k}$  et covariance  $P_{k|k}$ . Les filtres **EKF** et **UKF** sont des extensions du filtre de Kalman dans l'hypothèse où  $f$  et/ou  $g$  sont des fonctions non linéaires. Les équations standard du filtre de Kalman ne conviennent plus car la conservation de la propriété gaussienne ne peut pas être garantie lorsque les moments sont « propagés » au travers d'une fonction non linéaire. Nous présentons l'algorithme **UKF** en annexe A.

Soit  $H(\mathbf{x}_k|z_{1:k})$  l'entropie de la densité de filtrage  $p(x_k|z_{1:k})$  définie par

$$H(\mathbf{x}_k|z_{1:k}) = - \int p(x_k|z_{1:k}) \ln p(x_k|z_{1:k}) dx_k. \quad (2.28)$$

À tout instant  $k$ , comme le calcul des lois de filtrage des états futurs  $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \dots, \mathbf{x}_{k+N}$  dépend des futures observations  $z_{k+1}, z_{k+2}, \dots, z_{k+N}$ , avec  $N \in \mathbb{N}^*$ , nous calculons l'entropie moyenne sur ces futures observations conditionnellement aux observations passées. Pour  $N = 1$ , notons<sup>3</sup>  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[H(\mathbf{x}_{k+1}|z_{1:k+1})]$  l'espérance sur la prochaine mesure  $z_{k+1}$  conditionnellement aux mesures passées  $z_{1:k}$  de l'entropie de la loi de filtrage  $p(x_{k+1}|z_{1:k+1})$  à l'instant prochain  $k + 1$ . En appliquant la règle de Bayes définie en (2.26) à l'entropie moyenne  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[H(\mathbf{x}_{k+1}|z_{1:k+1})]$ , il est possible d'obtenir une équation de mise à jour de l'entropie

$$\begin{aligned} & \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[H(\mathbf{x}_{k+1}|z_{1:k+1})] \\ &= - \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} \left[ \int p(x_{k+1}|z_{1:k+1}) \ln p(x_{k+1}|z_{1:k+1}) dx_{k+1} \right] \\ &= - \int p(x_{k+1}, z_{k+1}|z_{1:k}) \ln p(x_{k+1}|z_{1:k}) dx_{k+1} dz_{k+1} \\ &\quad - \int p(x_{k+1}, z_{k+1}|z_{1:k}) \ln \frac{p(x_{k+1}, z_{k+1}|z_{1:k})}{p(z_{k+1}|z_{1:k})p(x_{k+1}|z_{1:k})} dx_{k+1} dz_{k+1} \\ &= H(\mathbf{x}_{k+1}|z_{1:k}) - I(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}|z_{1:k}). \end{aligned} \quad (2.29)$$

$H(\mathbf{x}_{k+1}|z_{1:k})$  est l'entropie de la pdf de prédiction  $p(x_{k+1}|z_{1:k})$  alors que  $I(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}|z_{1:k})$  est l'information mutuelle de  $p(x_{k+1}|z_{1:k})$  et  $p(z_{k+1}|z_{1:k})$ . Ainsi,  $I(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}|z_{1:k})$  représente la quantité d'information commune à  $p(x_{k+1}|z_{1:k})$  et  $p(z_{k+1}|z_{1:k})$  et constitue la réduction d'entropie, en moyenne, de la loi *a posteriori* suite à l'assimilation de l'observation future  $z_{k+1}$ . Comme l'information est d'autant plus importante que l'entropie est faible, l'information de la densité de filtrage augmente en moyenne de l'instant  $k$  à  $k + 1$  d'autant plus que  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[H(\mathbf{x}_{k+1}|z_{1:k+1})] < H(\mathbf{x}_k|z_{1:k})$ . Dans l'équation (2.29),  $H(\mathbf{x}_{k+1}|z_{1:k})$  est l'entropie de la loi de prédiction  $p(x_{k+1}|z_{1:k})$  après le déplacement du capteur et avant que l'observation  $z_{k+1}$  ne soit assimilée. Supposons que le modèle de dynamique exprime l'effet d'un mouvement rigide (défini par  $u_k$ ) et que l'entropie  $H(\mathbf{x}_{k+1}|z_{1:k})$  de  $p(x_{k+1}|z_{1:k})$  soit liée à l'entropie  $H(\mathbf{x}_k|z_{1:k})$  de la densité de filtrage  $p(x_k|z_{1:k})$  par la relation (qui sera vérifiée pour notre problème)

$$H(\mathbf{x}_{k+1}|z_{1:k}) \geq H(\mathbf{x}_k|z_{1:k}). \quad (2.30)$$

---

3. Notons que les observations  $z_{1:k}$  sont données alors que l'observation à l'instant  $k + 1$  est une variable aléatoire  $\mathbf{z}_{k+1}$ . Ainsi l'entropie de la loi de filtrage à l'instant  $k + 1$ , notée  $H(\mathbf{x}_{k+1}|z_{1:k+1})$  devrait s'écrire  $H(\mathbf{x}_{k+1}|\mathbf{z}_{k+1}, z_{1:k})$  car elle est fonction de  $\mathbf{z}_{k+1}$ . Conformément au fait que  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[f(\mathbf{z}_{k+1}, z_{1:k})]$  est notée  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[f(z_{1:k+1})]$ , l'espérance de  $H(\mathbf{x}_{k+1}|\mathbf{z}_{k+1}, z_{1:k})$  sur  $\mathbf{z}_{k+1}$  conditionnellement à  $z_{1:k}$  est désignée par  $\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}}[H(\mathbf{x}_{k+1}|z_{1:k+1})]$ .

La différence entre ces deux entropies représente donc l'information qui a été « perdue » au cours du mouvement (avant l'assimilation d'une observation, l'application d'une commande conduit à une augmentation de l'incertitude associée à l'état causée par le bruit de dynamique). En cas d'égalité, il n'y a pas de perte d'information suite au mouvement. Dans ce cas, l'entropie de la densité de filtrage diminue ou reste stable selon que la valeur de l'information mutuelle est strictement positive ou nulle. Nous reviendrons sur cette hypothèse dans la partie 2.3.1.

Dans le cas particulier où la densité de filtrage est gaussienne, à partir du résultat (2.12) et lorsque la loi *a posteriori* d'une variable aléatoire est connue, l'entropie de cette variable peut s'écrire comme une fonction de la matrice de covariance. À titre d'exemple, l'entropie de la densité de filtrage  $p(x_k|z_{1:k})$  peut s'écrire comme la fonction

$$H(\mathbf{x}_k|z_{1:k}) = \frac{1}{2} \ln \left( (2\pi e)^n |P_{k|k}| \right), \quad (2.31)$$

avec  $n$  la taille du vecteur d'état. Lorsque la loi de filtrage à un instant  $k+1$  est calculée récursivement au moyen d'un filtre de Kalman, la matrice de covariance de cette loi ne dépend pas de la réalisation  $z_{k+1}$  du processus d'observation  $\mathbf{z}_{k+1}$ . Ainsi,

$$\mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} H(\mathbf{x}_{k+1}|z_{1:k+1}) = H(\mathbf{x}_{k+1}|z_{1:k+1}), \quad (2.32)$$

et le résultat 2.29 correspond à la règle de mise à jour définie dans [Manyika 1993].

## 2.2 Perception active en robotique

Dans cette section, nous repositionnons le problème de mouvement actif pour la localisation binaurale au sein du contexte plus général de la perception active. En robotique ou pour toute application impliquant la commande de capteurs (tâches de surveillance ou de suivi de cibles par exemple), la perception active désigne le processus de décision en boucle fermée permettant de définir les commandes qui conduisent à la maximisation d'un critère, en partie ou en totalité défini par l'information délivrée par ces capteurs [Mihaylova *et al.* 2002]. En dehors de la robotique, le problème de perception active peut par exemple consister à sélectionner les paramètres de différents capteurs (sélection de la focale de caméras, des prises de vues, etc.), pour la reconnaissance d'objets ou de personnes [Denzler *et al.* 2003, Sommerlade & Reid 2008, Wenhardt *et al.* 2006]. En robotique, le but de la perception active est de commander le déplacement des capteurs embarqués sur un robot mobile, de façon à réaliser une tâche d'exploration combinée ou non avec d'autres objectifs dits de « stabilisation » (minimisation de l'énergie dépensée, positions cibles à atteindre ou à éviter, etc.). L'exploration est le problème de commande qui consiste à améliorer la connaissance de l'environnement du robot en maximisant un critère d'information. Lorsque ce critère est par exemple défini par l'entropie d'une variable aléatoire, il s'agit alors d'un problème de minimisation. Pour résoudre le problème

d'exploration, les techniques de localisation et de cartographie simultanées ([Simultaneous Localization And Mapping \(SLAM\)](#)) peuvent par exemple être étendues dans une version active. Le mouvement du robot est alors calculé de façon à améliorer la localisation du robot et des amers dans l'environnement [[Thrun et al. 2005](#)]. Dans certaines études, la maximisation d'un critère d'information conduit le robot à se diriger vers des zones inexplorées, ce qui peut être l'objectif recherché [[Julian et al. 2013](#)]. Dans notre application au contraire, la finalité réside uniquement dans la maximisation du critère d'information.

La perception active est une tâche complexe qui soulève un certain nombre de difficultés. Tout d'abord, les modèles des actions et des observations sont pour la plupart non linéaires. Ensuite, la solution du problème qui définit les actions à accomplir dépend souvent de la résolution de problèmes d'optimisation multiobjectif, le plus souvent sous contraintes. De plus, les processus d'état et d'observation sont de nature stochastique car ils sont soumis à des aléas (bruit de dynamique et d'observation, approximations faites pour la modélisation, etc.), qui doivent être pris en compte. Enfin, en robotique par exemple, la fréquence de la boucle perception-action doit être suffisamment élevée pour permettre de prendre en compte la dynamique d'évolution de l'environnement (l'apparition ou la disparition de sources dans une scène audio par exemple), ce qui impose des contraintes de temps possiblement sévères.

Nous étudions dans un premier temps la modélisation de ces problèmes en [2.2.1](#). Le choix du critère d'information pouvant être utilisé est ensuite discuté en [2.2.2](#). Enfin, nous comparons les approches court terme et long terme (que nous définissons), en [2.2.3](#).

### 2.2.1 Définition d'un modèle

Dans [[Mihaylova et al. 2002](#)], les auteurs présentent la perception active comme un problème de génération de trajectoires d'un système décrit par un modèle stochastique de dynamique et d'observation de la forme [\(2.24\)](#). Dans un certain nombre d'applications, les objectifs de réduction d'incertitude des variables d'état estimées sont combinés à d'autres objectifs inhérents à la mission (minimisation de l'énergie ou de la distance parcourue, minimisation de la distance à un objectif de position, etc.). Les auteurs définissent alors une fonction de valeur  $P$ , représentant la minimisation du coût (ou de façon équivalente la maximisation du gain) associé à chacun des objectifs de la mission, pour toute séquence d'actions (ou « politique »)  $\{a_1, \dots, a_n\}$ , de sorte que

$$P = \min_{a_1, \dots, a_n} \left[ \sum_{j=1} \alpha_j \mathcal{U}_j + \sum_{l=1} \beta_l \mathcal{C}_l \right]. \quad (2.33)$$

Les éléments relatifs à  $i$  et  $j$  doivent alors être associés aux objectifs d'exploration et aux objectifs de la mission évoqués plus haut. Les coefficients  $\alpha_j$  et  $\beta_l$  sont des éléments de pondération tandis que  $\mathcal{U}_j$  et  $\mathcal{C}_l$  sont les coûts de chaque objectif. Notons

que la solution du modèle (2.33) paramétré par les coefficients  $\alpha_j$  et  $\beta_l$  constitue une solution au sens de Pareto du problème d'optimisation multiobjectif

$$P' = \min_{a_1, \dots, a_n} [\mathcal{U}_1, \dots, \mathcal{U}_j, \mathcal{C}_l, \dots, \mathcal{C}_j]. \quad (2.34)$$

De façon à définir des actions réalisables, le problème de minimisation est souvent contraint, par exemple par un ensemble  $\mathcal{D}$  de sorte que

$$\forall i = 1, \dots, n, \quad a_i \in \mathcal{D}. \quad (2.35)$$

Notre application portant uniquement sur une tâche d'exploration, l'équation (2.33) peut être réduite à

$$P = \min_{a_1, \dots, a_n} \mathcal{U}, \quad (2.36)$$

avec  $a_1, \dots, a_n$  les actions, éventuellement contraintes, et  $\mathcal{U}$  un critère d'information qui reste à définir. Nous considérons que le critère ne peut pas tendre vers l'infini sur le domaine de contraintes considéré et que le problème admet ainsi un ou plusieurs minima finis. Nous étudierons la possibilité de prendre en charge un objectif supplémentaire (objectif de position à atteindre), dans la section 5.2.

### 2.2.2 Choix du critère d'information

Considérons un problème d'exploration où l'objectif est de commander un capteur (ou de définir ses paramètres), de façon à maximiser un critère d'information. Notons, que l'entropie est une mesure d'information qui caractérise l'incertitude d'une variable ou d'un vecteur aléatoire. En conséquence, l'information que porte l'entropie est maximale lorsque celle-ci est minimisée. Ainsi quand le problème de maximisation d'information consiste à minimiser une entropie, il s'agit plutôt d'un critère d'incertitude que d'un critère d'information. Dans cette partie nous présentons divers critères d'information ou d'incertitude.

La FIM et la CRLB ont été souvent utilisées pour résoudre certains types de problèmes de *Bearings-Only Tracking* (**Bearings-Only Tracking (BOT)**) en filtrage stochastique. Le BOT consiste à estimer la position d'une cible dans un repère lié à l'observateur, lorsque les observations n'apportent d'information que sur l'angle d'azimut formé par la cible et le Nord. Le problème du BOT actif consiste à rechercher la trajectoire optimale de l'observateur qui permet de maximiser une fonction de d'information. La FIM représentant l'information contenue dans les observations à propos du paramètre inconnu (azimut), l'erreur d'estimation peut être diminuée en établissant les commandes qui conduisent à maximiser le déterminant de la FIM [Oshman & Davidson 1999]. Dans [Helferty & Mudgett 1993], une comparaison de deux stratégies est effectuée pour un problème de BOT avec pour objectif la détermination de la position et de la vitesse de la cible. L'une de ces stratégies maximise le déterminant de la FIM et l'autre minimise un critère lié à la trace de la CRLB. Ce dernier est défini par la trace d'une somme pondérée des CRLB de

la position et de la vitesse de la source (décomposition de la **CRLB** du vecteur de paramètres total en deux matrices correspondant aux **CRLB** relatives à la position et à la vitesse). La pondération du critère sur la position et la vitesse est choisie de façon arbitraire selon l'objectif recherché. Il est montré que la trajectoire générée par la minimisation du critère lié à la **CRLB** pour un objectif sélectionné (position ou vitesse), engendre les meilleurs résultats sur cet objectif au détriment de l'autre. Ensuite, il est montré que la stratégie qui maximise le déterminant de la **FIM** constitue un compromis entre les deux critères. Pour éviter le calcul direct de la **FIM**, une stratégie peut être de calculer et de maximiser une borne inférieure [Liu 1988, Hammel *et al.* 1989, Le Cadre & Laurent-Michel 1999].

Bien que très similaire à notre application, des différences importantes avec le problème de localisation binaurale tel que nous le formulons sont mises en évidence dans [Portello 2013]. Notons par exemple que dans le **BOT**, le capteur est supposé ponctuel, l'azimut est défini par rapport à une référence (le Nord), et les observations renseignent toujours de la même façon sur l'azimut, quel que soit l'état courant de l'observateur. En effet dans le **BOT**, le système de coordonnées polaires est souvent utilisé, ce qui rend la fonction d'observation linéaire. Au contraire, dans notre problème de localisation binaurale, le capteur est constitué de deux microphones, l'azimut est défini relativement par rapport à la position courante du capteur (axe  $\vec{z}_R$ ) et l'incertitude spatiale liée à une estimation d'azimut dépend de la situation capteur-source (ce dernier point est illustré dans la section 3.2).

Pour la vision par ordinateur, la perception active (ou vision active dans ce contexte) consiste à accomplir des tâches de reconnaissance d'objets ou de personnes, de suivi de cibles ou encore de reconstruction 3D d'objets. Le problème peut alors être de déplacer de manière optimale une ou plusieurs caméras ou d'en modifier des paramètres tels que la distance focale, afin d'accomplir ce genre de tâches. Ces problèmes peuvent être résolus en maximisant un critère d'information construit à partir de l'entropie ou de l'information mutuelle, comme le montrent [Denzler & Brown 2002]. L'objectif est d'estimer la valeur de l'état  $\mathbf{x}_k$  à partir d'observations  $\mathbf{z}_k$  et de définir l'action  $a_k$  qui permet d'obtenir le maximum d'information sur  $\mathbf{x}_k$ . Une des spécificités du problème est que les actions  $a_k$  n'ont pas d'effet sur l'état  $\mathbf{x}_k$ , mais interviennent dans le processus d'observation  $p(z_k|x_k, a_k)$ . Les auteurs définissent l'information mutuelle  $I(\mathbf{x}_k, \mathbf{z}_k|a_k)$  de l'état et de l'observation conditionnellement à une action comme

$$I(\mathbf{x}_k, \mathbf{z}_k|a_k) = H(\mathbf{x}_k) - H_c(\mathbf{x}_k|\mathbf{z}_k, a_k), \quad (2.37)$$

avec  $H(\mathbf{x}_k)$  l'entropie de l'état *a priori* et  $H_c(\mathbf{x}_k|\mathbf{z}_k, a_k)$  l'entropie de l'état conditionnellement à la mesure obtenue pour l'application d'une action. L'action optimale  $a_k^*$  à appliquer à la caméra est définie comme l'argument maximum de cette information de sorte que

$$a_k^* = \arg \max_{a_k} I(\mathbf{x}_k; \mathbf{z}_k|a_k). \quad (2.38)$$

L'évaluation de l'information mutuelle est effectuée au moyen d'une méthode de Monte Carlo pour un problème de reconnaissance d'objet. Dans ce contexte, le problème d'estimation est statique car le vecteur d'état est constant dans le temps. Cependant l'approche proposée peut également permettre d'étudier des problèmes dynamiques, ce qui est proposé dans [Denzler *et al.* 2003] afin de résoudre un problème de suivi de cibles dans une scène 3D en réglant la distance focale de deux caméras. L'estimation de l'état est effectuée au moyen d'un filtre EKF. Pour sélectionner les actions, les auteurs définissent alors le critère d'information comme l'entropie conditionnelle  $H_c(\mathbf{x}_k|\mathbf{z}_k)$  et l'action optimale  $a_k^*$  à appliquer à la caméra est obtenue en recherchant l'argument minimum de cette entropie. Les auteurs utilisent ensuite les propriétés de l'entropie sur les vecteurs gaussiens pour exprimer le problème comme la minimisation de l'espérance, sur la future observation, d'une fonction du déterminant de la matrice de covariance de la pdf *a posteriori*. Le même formalisme est utilisé dans [Wenhardt *et al.* 2006] pour effectuer la reconstruction 3D d'un objet en utilisant les images d'une caméra mobile. L'objectif est de sélectionner la prochaine position de la caméra afin d'améliorer la reconstruction de l'objet. Une extension de ces travaux est proposée dans [Sommerlade & Reid 2008], où le zoom et l'orientation d'une caméra doivent être sélectionnés pour un problème de reconnaissance de personnes. Contrairement aux travaux précédemment cités, les auteurs recherchent un compromis dans leur critère de sélection de paramètres entre la reconnaissance et la possibilité d'acquérir de nouvelles cibles. La vision active semble donc être proche de notre problème dans le sens où l'objectif est d'établir une action de façon à maximiser un critère d'information. Notons cependant qu'une différence fondamentale avec notre application tient au fait que les actions interviennent dans le processus d'observation. Dans notre application, les actions sont les commandes du capteur binaural qui modifient l'état au travers d'une loi de dynamique *a priori*. Le processus d'observation au contraire est supposé indépendant des actions imposées au capteur. Par ailleurs la résolution de (2.38) est effectuée en brute force en évaluant l'information mutuelle pour toutes les actions possibles. Une telle approche n'est possible que si les variables de commande sont discrètes et prennent leurs valeurs dans un ensemble de faible dimension, ce qui n'est pas le cas dans notre application.

Des critères similaires utilisant l'entropie ou l'information mutuelle, ont été utilisés pour guider les modèles d'exploration en robotique. Dans [Burgard *et al.* 1997], le problème d'exploration pour la robotique est étudié avec comme objectif de déterminer les commandes du robot de manière à améliorer sa localisation dans une carte connue de l'environnement. Soit  $H_k$ , l'entropie de la pdf *a posteriori*<sup>4</sup> de la position du robot obtenue à l'instant  $k$  dans un processus de localisation markovien. La quantité  $\mathbb{E}_z [H_{k+1}]$  est l'espérance sur les observations possibles, après avoir exécuté une action  $a_k$ , de  $H_{k+1}$  (entropie de la loi *a posteriori* après avoir effectué une action  $a_k$  et assimilé une mesure  $z_k$ , réalisation de la variable d'observation  $\mathbf{z}_k$ ). Le

---

4. Dans ce contexte de processus de décision markovien, cette densité est appelée « état de croyance ».

problème est ensuite formulé de façon à déterminer l'action  $a_k^*$  qui minimise une somme pondérée du critère d'entropie  $\mathbb{E}_{\mathbf{z}} [H_{k+1}]$  et d'un critère de coût de l'action  $v(a_k)$ , de sorte que

$$a_k^* = \arg \min_{a_k} (\mathbb{E}_{\mathbf{z}} [H_{k+1}] + \alpha v(a_k)), \quad (2.39)$$

avec  $\alpha$  le coefficient de pondération qui détermine l'importance du coût de l'action par rapport au critère d'exploration. Une étude similaire est proposée dans [Fox *et al.* 1998], pour résoudre le même problème. Pour guider l'exploration, les auteurs définissent une fonction d'utilité des actions  $U_k(a_k)$ , comme la différence entre les entropies  $H_k$  et  $\mathbb{E}_{\mathbf{z}} [H_{k+1}]$ , *i.e.*,

$$U_k(a_k) = H_k - \mathbb{E}_{\mathbf{z}} [H_{k+1}]. \quad (2.40)$$

L'action  $a_k^*$  à appliquer au robot est ensuite choisie de façon à maximiser la différence entre la fonction d'utilité et une fonction de coût  $C_k(a_k)$ , dans un critère semblable à (2.39) par

$$a_k^* = \arg \max_{a_k} (U_k(a_k) - \beta C_k(a_k)), \quad (2.41)$$

avec  $\beta$  un coefficient de pondération. Nous retrouvons ici en (2.39) et (2.41) des critères similaires au modèle (2.33), où apparaît un terme correspondant à l'espérance du gain d'information et un autre détermine le coût des actions. À partir d'un tel modèle, un algorithme d'exploration de Monte Carlo est défini dans [Thrun *et al.* 2005]. Les intégrales qui interviennent dans les calculs des fonctions d'information et de coût sont calculées par un processus d'échantillonnage aléatoire. La complexité du calcul de la fonction d'utilité dépend du nombre de positions atteignables par le robot et du nombre d'observations.

Dans le but guider un robot pour localiser des cibles et cartographier l'environnement, [Vlassis *et al.* 1999] utilisent un critère d'information basé sur une loi de mise à jour de l'entropie qui est définie de la façon suivante. Soit  $\mathbf{x}$  un vecteur aléatoire qui caractérise la position du robot et  $p(x)$  la distribution *a priori*. Soit  $z$  une réalisation de la variable aléatoire  $\mathbf{z}$ , correspondant à une observation effectuée par le robot et caractérisée par la pdf conditionnelle  $p(z|x)$ . Les auteurs définissent une « fonction de risque »  $R$  comme étant l'espérance sur toutes les mesures possibles, de l'entropie  $H(\mathbf{x}|z)$  de la loi *a posteriori*  $p(x|y)$ , c'est à dire

$$\begin{aligned} R &= \mathbb{E}_{\mathbf{z}} [H(\mathbf{x}|z)], \\ &= H_c(\mathbf{z}|\mathbf{x}) + H(\mathbf{x}) - H(\mathbf{z}), \end{aligned} \quad (2.42)$$

avec  $H_c(\mathbf{z}|\mathbf{x})$ ,  $H(\mathbf{x})$  et  $H(\mathbf{z})$ , représentant respectivement l'entropie de  $\mathbf{z}$  conditionnellement à  $\mathbf{x}$  (entropie conditionnelle définie selon (2.13)), l'entropie de la position du robot et l'entropie de l'observation.  $R$  est l'entropie conditionnelle  $H_c(\mathbf{x}|\mathbf{z})$ . Ce résultat est similaire à la loi de mise à jour de l'entropie que nous avons définie

en (2.29). L'information mutuelle  $I(\mathbf{x}, \mathbf{z})$  peut être introduite en appliquant (2.22). L'entropie et l'information mutuelle peuvent aussi être utilisées de façon complémentaire. C'est le cas dans [Bourgault *et al.* 2002] où un robot doit générer une carte. Le problème consiste ici à maximiser l'information relative à cette carte en sélectionnant une action dans l'ensemble discrétisé des actions possibles, tout en localisant la position du robot. Une stratégie combinant un algorithme de SLAM et une grille d'occupation est utilisée. Alors que le SLAM utilise un filtre de type EKF afin d'estimer la position du robot et des amers dans la carte, la grille d'occupation discrétise l'environnement et associe à chaque case la probabilité *a posteriori* d'être occupée par le robot. Une fonction de coût différente est associée au SLAM et à la grille. Pour le SLAM, elle est déterminée à partir de l'entropie de la loi *a posteriori* supposée gaussienne, qui est donc une fonction de la matrice de covariance. La fonction de coût est définie comme la somme des aires des ellipses de confiance de la position des amers et du robot avec une formulation initialement proposée dans [Feder *et al.* 1999]. En parallèle, la deuxième fonction est définie de la façon suivante. À un instant  $k$ , en considérant pour chaque case  $i$  de la grille une pdf *a priori*  $P_i(x_i)$ , les auteurs définissent l'information mutuelle

$$I_i = H_i(\mathbf{x}_i) - \mathbb{E}_{\mathbf{z}} [H_i(\mathbf{x}_i|z_k)], \quad (2.43)$$

avec  $H_i(\mathbf{x}_i)$  l'entropie *a priori* de la loi  $P_i(x_i)$  et  $\mathbb{E}_{\mathbf{z}} [H_i(\mathbf{x}_i|z_k)]$  l'espérance sur l'observation future  $\mathbf{z}_k$ , de l'entropie de la loi  $p(x_i|z_k)$ , qui est la probabilité de l'état  $x_i$  conditionnellement à la seule mesure  $z_k$ . La fonction de coût est ensuite obtenue en sommant les  $I_i$  de chaque case, parmi toutes les cases qui se trouvent à portée de mesure. Enfin, la combinaison de ces deux fonctions de coût est effectuée par une combinaison linéaire convexe et le problème d'optimisation multiobjectif est résolu numériquement en discrétisant l'espace des actions pour un horizon court terme. La résolution du problème n'est pas optimale mais les auteurs soulignent que la combinaison de deux fonctions de coût est intéressante pour les problèmes d'exploration. En effet, la fonction de coût liée à la grille d'occupation attire le robot vers les zones inexplorées, tandis que la fonction liée au SLAM assure que la localisation du robot par rapport aux points d'intérêt de la carte reste suffisamment précise. Dans [Julian *et al.* 2013], il est également montré qu'une loi de commande qui maximise une fonction de récompense dérivée de l'information mutuelle d'une grille d'occupation et des futures mesures, attire le robot vers des zones inexplorées. Nous retrouverons un comportement similaire après application de notre stratégie d'exploration (voir chapitre 3).

Une autre approche de l'exploration consiste à faire coopérer plusieurs robots de façon à maximiser l'information apportée par les observations de chaque robot. Une approche privilégiée pour ce problème semble être la version décentralisée du filtre information [Manyika 1993]. Ce filtre est utilisé dans [Grocholsky *et al.* 2003] pour résoudre un problème de BOT actif multi-cibles. L'objectif est de calculer la trajectoire des robots de façon à maximiser une fonction d'information définie par rapport aux déterminants de la matrice d'information présente et future. Les commandes des

robots sont obtenues en résolvant numériquement un problème de commande optimale selon une méthode détaillée dans [Grocholsky 2006]. Dans [Stump *et al.* 2009] un problème similaire est étudié mais les robots ne disposent que de capteurs permettant de mesurer leur distance relative à des amers. La loi de commande est mise en place en suivant la direction du gradient d'une fonction définie par le déterminant de la matrice d'information du filtre information. Dans [Schwager *et al.* 2011] un problème où des robots doivent se déplacer de façon à rassembler un maximum d'information sur un environnement inconnu comportant des zones potentiellement hostiles, est envisagé. Le problème d'estimation est étudié avec une approche de filtrage bayésien. La loi de commande des robots est générée en suivant le gradient par rapport à la position de chaque robot, de l'information mutuelle d'une grille d'occupation et de la mesure. La contribution principale consiste en la formulation d'une expression analytique de ce gradient, dont nous donnons l'expression ici. Soient  $\mathbf{s}$  et  $\mathbf{y}$  deux variables aléatoires dont la loi jointe  $p_x(s, y)$  dépend d'un paramètre  $x \in \mathbb{R}^{2n}$  (représentant la position de  $n$  robots dans une carte à deux dimensions).  $\mathbf{s}$  est le vecteur d'état discret représentant l'état de l'environnement (présence ou non d'un danger), dont chaque composante est associée à une position et prend ses valeurs dans un ensemble  $S = \{0, 1\}$ , et  $\mathbf{y}$  de taille  $n$ , est le vecteur d'observation dont les composantes sont les mesures effectuées par chaque robot. Une expression du gradient de l'information mutuelle  $I_x(\mathbf{s}, \mathbf{y})$  est démontrée dans un cas général sans aucune restriction sur les pdf mises en jeu ni sur la dépendance entre le paramètre  $x$  et la pdf  $p_x(s, y)$ . Cependant, les auteurs précisent que l'évaluation de ce gradient est de complexité exponentielle par rapport au nombre de capteurs, de cases dans la grille d'occupation et de zones dangereuses à éviter, ce qui rend cette approche difficilement exploitable même pour un nombre réduit de robots et de cases dans la grille. Dans [Dames *et al.* 2012], les mêmes auteurs considèrent que les capteurs ont une portée limitée, de façon à réduire la complexité de calcul du gradient de l'information mutuelle. Des approches numériques telle que la méthode de Monte Carlo peuvent également être utilisées afin de calculer une approximation de ce gradient [Julian *et al.* 2011].

### 2.2.3 Décisions à court terme contre décisions à long terme

À tout instant discret  $k$ , le problème de perception active consiste donc à établir la séquence d'actions  $\{a_1, \dots, a_n\}$  à appliquer aux instants  $\{k + 1, \dots, k + n\}$ , qui maximise un critère d'information lié à la densité de filtrage à l'instant  $k + N$ . Lorsque le problème est posé en ces termes, il s'agit d'une planification sur plusieurs pas, que nous qualifions de stratégie à long terme. Une autre approche peut consister à établir l'action  $a_1$  qui maximise le critère d'information lié à l'instant  $k + 1$ . Dans ce cas il ne s'agit plus de planification mais d'une stratégie à court terme, où la récompense de l'action appliquée est immédiate. Dans cette section nous évoquons les avantages et les faiblesses de chaque approche.

**Récompense immédiate.** Les problèmes dont le critère d'information repose sur le prochain pas d'estimation uniquement, répondent à la question suivante : « quelle est la meilleure action à effectuer pour avoir la plus grande récompense après l'application de cette action et l'assimilation d'une observation ? » Dans une telle approche, parfois appelée gloutonne (*greedy*) dans la littérature, l'objectif est de gagner un maximum d'information dans l'immédiat.

Comme le critère n'est évalué que sur un seul pas d'estimation, les calculs peuvent dans certains cas rester assez simples, ce qui rend l'approche très adaptée pour certains types de problèmes embarqués où la rapidité de calcul est une contrainte importante. Cependant, la résolution du problème même à court terme, peut nécessiter une forte capacité de calcul lorsque la dimension des vecteurs d'état et d'observation augmente. Par exemple, l'algorithme d'exploration de Monte Carlo [Thrun *et al.* 2005] précédemment évoqué ne peut être résolu que lorsque la dimension du vecteur d'observation est limité. L'espérance sur les futures observations est calculée en effectuant un tirage aléatoire des mesures suivant leur loi de probabilité conditionnellement aux mesures jusqu'à l'instant présent. Ainsi la complexité de calcul peut facilement devenir très grande quand la taille de l'espace des mesures augmente. Cette approche est donc difficilement exploitable lorsque la commande doit être calculée en ligne, au cours de l'exploration. Les auteurs soulignent également que, de par cette complexité, cette méthode est généralement inadaptée aux problèmes de décision à long terme. Afin de rendre le problème soluble, des approximations sur le modèle des capteurs [Burgard *et al.* 1997, Julian *et al.* 2013] ou une réduction sévère du nombre d'actions [Bourgault *et al.* 2002] peut être effectuée.

Lorsque le processus d'estimation est réalisé au moyen d'un filtre gaussien (Kalman, EKF, UKF ou filtre information), la stratégie de décision court terme présente un autre avantage. En effet dans l'étape de mise à jour de ce type de filtres, la matrice de covariance *a posteriori* dépend de l'observation prédite mais est indépendante de la réalisation du processus d'observation. Ainsi, lorsque le critère d'information est construit par une fonction du déterminant de la matrice de covariance (entropie) ou par son inverse (FIM), celui-ci peut être évalué sans avoir à calculer une espérance sur la future observation.

Toutefois, une stratégie à récompense immédiate n'est pas nécessairement optimale sur plusieurs pas. En effet, appliquer  $N$  fois l'action optimale à un pas, calculée à partir d'un critère d'information court terme en chaque instant, ne garantit pas un résultat optimal à l'issue de ces  $N$  pas.

**Long terme.** Pour pallier ce problème, l'étude d'un critère d'information sur un horizon temporel plus lointain peut être envisagée.

Comme la densité de filtrage  $N$  pas dans le futur dépend des  $N$  futures observations, tout critère d'information construit à partir de cette densité doit en conséquence contenir un calcul d'espérance sur ces observations. Dans le cas du filtre de Kalman, contrairement à la situation court terme ( $N = 1$ ), la matrice de

covariance de la densité de filtrage à l’instant  $k + N$  (avec  $N > 1$ ) dépend implicitement des réalisations des observations qui précèdent cet instant<sup>5</sup>. Dans un certain nombre d’études [Forster *et al.* 2014, Deutsch *et al.* 2004, Leung *et al.* 2006], les auteurs font l’hypothèse que les observations futures sont égales à leurs prédictions afin de permettre l’évaluation du critère d’information. Dans le cadre du filtrage de Kalman, cela revient à annuler le terme d’innovation dans l’étape de mise à jour. Une autre approche peut consister à évaluer cette espérance au moyen d’une méthode de Monte Carlo [Ryan & Hedrick 2010, Nguyen *et al.* 2017]. Dans le développement du chapitre 4, nous montrerons qu’une approximation de cette espérance peut être calculée via la transformée unscented.

Pour utiliser un modèle de décision à long terme dans un problème d’exploration, une stratégie à horizon « glissant » est généralement mise en place. À un instant  $k$ , les actions  $\{a_k^*, a_{k+1}^*, \dots, a_{k+N-1}^*\}$  qui maximisent le critère d’information lorsqu’elles sont respectivement appliquées en chaque instant  $\{k, k + 1, \dots, k + N - 1\}$ , sont déterminées. Ensuite, seule l’action  $a_k$  est appliquée afin de placer le robot (ou plus généralement le capteur) dans une nouvelle configuration, de façon à pouvoir assimiler une nouvelle observation  $z_{k+1}$ . Cette procédure est ensuite répétée à l’instant suivant. Comparée à une approche à horizon fixe qui consisterait à calculer toutes les commandes jusqu’à un horizon lointain et de les appliquer en chaque instant, la méthode à horizon « glissant » semble la plus adaptée pour plusieurs raisons :

- Dans les problèmes d’exploration, l’instant final de l’expérience n’est pas nécessairement déterminé à l’avance. La tâche d’exploration peut être effectuée jusqu’à ce qu’un critère soit satisfait ou même être effectuée en continu (tâches de configuration et de contrôle actif de caméras pour la surveillance par exemple). Dans ce cas les problèmes à horizon fixe ne sont pas adaptés.
- Un horizon fixe au-delà d’une certaine valeur de  $N$  peut engendrer une trop grande complexité de calcul et rendre la résolution du problème impossible. Poser le problème sur un horizon glissant permet de limiter la taille de  $N$  sans imposer de restriction sur l’horizon final de l’exploration.
- Des observations relativement éloignées de leur prédiction peuvent apparaître et ne seraient pas prises en compte dans une stratégie à pas fixe, alors qu’une stratégie glissante incorpore la dernière observation disponible pour mettre à jour la politique d’action à appliquer.

La taille  $N$  de l’horizon de calcul doit être choisie avec attention et dépend de plusieurs considérations. Premièrement, cette décision peut dépendre de la nature du problème abordé. En effet, dans le cadre par exemple de l’exploration d’un environnement par un robot, nous pouvons imaginer qu’une valeur de  $N$  lointaine permettra au robot de se diriger dans des zones qu’il aurait évité avec un horizon de décision plus court. Dans d’autres situations, l’objectif peut être d’obtenir le

---

5. De la même façon que la covariance  $P_{k+1|k+1}$  de la loi  $p(x_{k+1}|z_{1:k+1})$  dépend implicitement des réalisations  $z_{1:k}$  du processus d’observation  $\mathbf{z}_{1:k}$ . Cette information est alors portée par la loi  $p(x_k|z_{1:k})$  dont les premiers moments  $\hat{x}_{k|k}$  et  $P_{k|k}$  sont connus à l’instant  $k$ .

maximum d'information dans un temps très court. Dans ce cas, un choix de  $N$  assez faible ( $N = 1, 2, 3, \dots$ ) sera plus adapté. Ensuite, notons que l'augmentation de  $N$  entraîne une augmentation de la complexité de calcul. La taille de  $N$  doit donc être limitée de sorte qu'une solution puisse être trouvée dans un temps raisonnable. Ce critère de temps dépend des problèmes étudiés, suivant que les actions sont établies hors ligne (l'horizon est alors fixe), ou en ligne, à chaque instant du processus. Un compromis entre une stratégie à récompense immédiate ( $N = 1$ ) et une stratégie long terme avec un horizon  $N$  lointain doit donc être trouvé.

Dans notre application, le capteur binaural doit se déplacer de façon à améliorer le processus de localisation. Dans ce cas, l'objectif n'est pas d'obtenir l'information maximale au terme d'un horizon très grand mais plutôt d'obtenir le maximum d'information sur un temps relativement court. La valeur de  $N$  la plus adaptée au problème dans une stratégie glissante, ne correspond donc pas nécessairement à l'horizon le plus lointain pour lequel le problème peut être résolu. Nous illustrerons ce problème dans la partie 4.3 en comparant notre stratégie d'exploration pour différentes valeurs de  $N$ .

## 2.3 Fonction d'incertitude pour la localisation active de sources sonores

Dans cette partie, nous définissons en 2.3.1, en établissant certaines hypothèses, les limites du modèle d'exploration utilisé pour la synthèse du vecteur de commande du capteur binaural. Nous établissons ensuite en 2.3.2 une fonction d'incertitude originale liant le mouvement du capteur binaural à l'information disponible sur la position de la source. Notre fonction est construite sur la base de l'entropie de Shannon, ce qui nous permet d'exploiter les propriétés de l'entropie sur les vecteurs aléatoires gaussiens.

### 2.3.1 Hypothèses du modèle d'exploration

**Hypothèse gaussienne.** Comme nous l'avons évoqué en 1.3.4, les futures variables d'observation  $y_{k+1:k+N}$  encore inconnues à l'instant  $k$  sont remplacées par des variables alternatives  $z_{k+1:k+N}$ , qui sont des réalisations du processus d'observation  $\mathbf{z}_{k+1:k+N}$  relié au processus d'état caché par le modèle d'observation (1.15). Ainsi à chaque instant  $k$  du processus de localisation, l'exploration est initialisée par la pdf  $p(x_k|z_{1:k})$  que l'on suppose gaussienne, *i.e.*,

$$p(x_k|z_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}), \quad (2.44)$$

de moyenne  $\hat{x}_{k|k}$  et de covariance  $P_{k|k}$ . Nous faisons également une approximation gaussienne des lois de prédiction de l'état  $p(x_{k+1}|z_{1:k})$  et d'observation  $p(z_{k+1}|x_{k+1})$

de sorte que

$$p(x_{k+1}|z_{1:k}) \sim \mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k}), \quad (2.45)$$

$$p(z_{k+1}|x_{k+1}) \sim \mathcal{N}(z_{k+1}; h(x_{k+1}), R_{k+1}), \quad (2.46)$$

et nous définissons la pdf de la mesure prédite par la gaussienne

$$p(z_{k+1}|z_{1:k}) = \mathcal{N}(z_{k+1}; \hat{z}_{k+1|k}, S_{k+1}). \quad (2.47)$$

Nous supposons que les moments de (2.45) et (2.47) ainsi que ceux de (2.44) peuvent être calculés récursivement dans une stratégie de type UKF. L'hypothèse gaussienne facilite la définition d'une fonction de récompense en utilisant les résultats (2.31) de l'entropie sur les distributions gaussiennes.

Lorsque l'ensemble de la stratégie de localisation est considéré (voir figure 1.2), la pdf  $p(x_k|y_{1:k})$  forme un mélange de gaussiennes (voir par exemple la figure 1.3) et  $p(x_k|z_{1:k})$  peut alors être obtenue en retenant l'hypothèse de poids le plus fort au sein du mélange, ou bien en calculant l'approximation « moment-matched » (« moment-matched approximation » – approximation consistant à remplacer le mélange de gaussiennes par une distribution gaussienne dont les deux premiers moments sont ceux du mélange).

**Bruit de dynamique nul.** Nous considérons que le bruit de dynamique est suffisamment faible pour que les pertes d'information générées par le mouvement du capteur puissent être négligeables. Sous cette hypothèse le théorème 2.1 est vérifié.

**Theorème 2.1.** *Lorsque le bruit de dynamique est nul, l'entropie de la densité de filtrage à un instant  $k$  est égale à l'entropie de la loi de prédiction en  $k + 1$  après qu'un mouvement a été appliqué, i.e.,*

$$H(\mathbf{x}_k|z_{1:k}) = H(\mathbf{x}_{k+1}|z_{1:k}). \quad (2.48)$$

□

*PREUVE.* Lorsque les pdf  $p(x_k|z_{1:k})$  et  $p(x_{k+1}|z_{1:k})$  désignent respectivement les lois gaussiennes 2.44 et 2.45, les entropies associées ne dépendent que des matrices de covariance  $P_{k|k}$  et  $P_{k+1|k}$ , en effet,

$$H(\mathbf{x}_k|z_{1:k}) = \frac{1}{2} \ln \left[ (2\pi e)^n |P_{k|k}| \right], \quad (2.49)$$

et

$$H(\mathbf{x}_{k+1}|z_{1:k}) = \frac{1}{2} \ln \left[ (2\pi e)^n |P_{k+1|k}| \right]. \quad (2.50)$$

La covariance associée à l'état de prédiction peut être exprimée par la relation

$$P_{k+1|k} = R^T(\phi_k)P_{k|k}R(\phi_k) + Q_k, \quad (2.51)$$

avec  $R(\phi_k)$  la matrice de rotation définie en (1.10). La différence entre les deux entropies peut alors s'écrire

$$H(\mathbf{x}_{k+1}|z_{1:k}) - H(\mathbf{x}_k|z_{1:k}) = \frac{1}{2} \ln \frac{|R^T(\phi_k)P_{k|k}R(\phi_k) + Q_k|}{|P_{k|k}|} \quad (2.52)$$

Comme  $Q_k$  est supposée nulle,

$$|R^T(\phi_k)P_{k|k}R(\phi_k) + Q_k| = |R^T(\phi_k)P_{k|k}R(\phi_k)| = |P_{k|k}|, \quad (2.53)$$

donc,

$$H(\mathbf{x}_{k+1}|z_{1:k}) - H(\mathbf{x}_k|z_{1:k}) = 0. \quad (2.54)$$

■

Ainsi, un mouvement du capteur binaural avec bruit de dynamique nul peut soit apporter de l'information, *i.e.*, améliorer la localisation, soit laisser le système dans le même état d'information.

### 2.3.2 Définition d'une fonction d'incertitude

L'information disponible sur la position de la source à un instant  $k$  est entièrement décrite par l'entropie  $H(\mathbf{x}_k|z_{1:k})$ , exprimée en (2.28), de la loi de filtrage  $p(x_k|z_{1:k})$ . Une manière d'évaluer l'impact de l'application de la séquence des vecteurs de commande

$$\bar{u}_N = u_k, u_{k+1}, \dots, u_{k+N-1}, \quad (2.55)$$

sur les instants consécutifs à  $k$  jusqu'à  $k + N$ , est donc de calculer l'entropie  $H(\mathbf{x}_{k+N}|z_{1:k+N})$  de la loi de filtrage  $p(x_{k+N}|z_{1:k+N})$  à l'instant  $k + N$ . Comme cette entropie dépend des  $N$  futures observations  $z_{k+1:k+N}$  encore inconnues, elle ne peut pas être évaluée directement. Nous considérerons donc plutôt l'espérance de cette entropie sur ces mesures inconnues, conditionnellement aux mesures passées. Soit  $J_N$  la fonction d'information définie par

$$J_N = \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})]. \quad (2.56)$$

Nous proposons dans cette partie, à partir des considérations effectuées en 2.3.1 et des propriétés de l'entropie sur les vecteurs gaussiens, de développer une fonction d'information, approximation originale de (2.56). Cette fonction est dans un premier temps exprimée pour le cas de décision à court terme ( $N = 1$ ) puis son expression est

généralisée pour la décision à long terme ( $N \geq 2$ ). Notons que dans la suite de cette partie, l'information mutuelle  $I(\mathbf{x}_{k+i}, \mathbf{z}_{k+i} | z_{1:k+i-1})$  conditionnellement aux mesures précédentes  $z_{1:k+i-1}$  des vecteurs d'état  $\mathbf{x}_{k+i}$  et d'observation  $\mathbf{z}_{k+i}$  à l'instant  $k+i$ , sera notée  $I_{k+i}$ .

### A Récompense immédiate

Introduisons dans un premier temps les lemmes 2.1 et 2.2 qui nous permettent d'établir une expression de  $J_1$  (2.64) présentée dans le théorème 2.2.

**Lemme 2.1.** *L'espérance sur la future observation  $z_{k+1}$  conditionnée par les observations passées  $z_{1:k}$ , de l'entropie  $H(\mathbf{x}_{k+1} | z_{1:k+1})$  de la loi de filtrage  $p(x_{k+1} | z_{1:k+1})$ , peut être décomposée en*

$$\mathbb{E}_{\mathbf{z}_{k+1} | z_{1:k}} [H(\mathbf{x}_{k+1} | z_{1:k+1})] = H(\mathbf{x}_{k+1} | z_{1:k}) - I_{k+1}. \quad (2.57)$$

□

*PREUVE.* La preuve de (2.57) a été donnée en (2.29). ■

**Lemme 2.2.** *L'information mutuelle  $I_{k+1}$ , conditionnellement aux réalisations des mesures précédent l'instant  $k+1$ , de l'état  $\mathbf{x}_{k+1}$  et de l'observation  $\mathbf{z}_{k+1}$  peut être décomposée en*

$$I_{k+1} = K_1 + H(\mathbf{z}_{k+1} | z_{1:k}), \quad (2.58)$$

avec  $K_1$  une constante indépendante de la commande  $\bar{u}_1 = u_k$ . □

*PREUVE.* L'information mutuelle

$$I_{k+1} = \int p(x_{k+1}, z_{k+1} | z_{1:k}) \ln \frac{p(x_{k+1}, z_{k+1} | z_{1:k})}{p(z_{k+1} | z_{1:k})p(x_{k+1} | z_{1:k})} dx_{k+1} dz_{k+1} \quad (2.59)$$

peut être décomposée en  $I_{k+1} = A + B$  avec

$$A = \int p(x_{k+1}, z_{k+1} | z_{1:k}) \ln \frac{p(x_{k+1}, z_{k+1} | z_{1:k})}{p(x_{k+1} | z_{1:k})} dx_{k+1} dz_{k+1}, \quad (2.60)$$

et

$$B = - \int p(x_{k+1}, z_{k+1} | z_{1:k}) \ln p(z_{k+1} | z_{1:k}) dx_{k+1} dz_{k+1}. \quad (2.61)$$

En séparant la densité jointe  $p(x_{k+1}, z_{k+1}|z_{1:k})$  avec les règles de probabilité conditionnelles, la quantité  $A$  peut se réécrire :

$$\begin{aligned} A &= \int p(x_{k+1}, z_{k+1}|z_{1:k}) \ln \frac{p(x_{k+1}|z_{1:k})p(z_{k+1}|x_{k+1})}{p(x_{k+1}|z_{1:k})} dx_{k+1} dz_{k+1} \\ &= \int p(x_{k+1}|z_{1:k}) \left( \int p(z_{k+1}|x_{k+1}) \ln p(z_{k+1}|x_{k+1}) dz_{k+1} \right) dx_{k+1} \\ &= -\mathbb{E}_{\mathbf{x}_{k+1}|z_{1:k}} H(\mathbf{z}_{k+1}|x_{k+1}). \end{aligned} \quad (2.62)$$

Comme  $H(\mathbf{z}_{k+1}|x_{k+1})$  ne dépend que de la matrice de covariance du bruit de mesure  $R_{k+1}$ , la quantité  $A$  est donc constante par rapport à la commande  $\bar{u}_1 = u_k$  ; nous la renommons  $K_1$ . En utilisant l'indépendance de  $p(z_{k+1}|z_{1:k}) \ln p(z_{k+1}|z_{1:k})$  par rapport à la variable  $x_{k+1}$  la quantité  $B$  peut être développée en

$$\begin{aligned} B &= - \int p(z_{k+1}|z_{1:k}) \ln p(z_{k+1}|z_{1:k}) \left( \int p(x_{k+1}|z_{1:k+1}) dx_{k+1} \right) dz_{k+1} \\ &= - \int p(z_{k+1}|z_{1:k}) \ln p(z_{k+1}|z_{1:k}) dz_{k+1} \\ &= H(\mathbf{z}_{k+1}|z_{1:k}). \end{aligned} \quad (2.63)$$

■

**Theorème 2.2.** *Pour tout instant  $k$ , l'espérance sur la future observation  $\mathbf{z}_{k+1}$  conditionnellement aux réalisations des observations passées  $z_{1:k}$ , de l'entropie  $H(\mathbf{x}_{k+1}|z_{1:k+1})$  de la loi de filtrage  $p(x_{k+1}|z_{1:k+1})$  à l'instant  $k+1$ , notée  $J_1$ , peut être décomposée en*

$$\begin{aligned} J_1(\bar{u}_1) &= \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} [H(\mathbf{x}_{k+1}|z_{1:k+1})] \\ &= K_1' - H(\mathbf{z}_{k+1}|z_{1:k}), \\ &= K_1' - F_1(\bar{u}_1), \end{aligned} \quad (2.64)$$

Avec  $F_1$  (resp.  $K_1$ ) une fonction dépendant (resp. constante indépendante) du vecteur de commande  $\bar{u}_1$ . □

*PREUVE.* En appliquant le lemme 2.1 et en combinant le résultat avec le lemme 2.2, la fonction de récompense  $J_1$  peut s'écrire comme

$$J_1 = -K_1 - H(\mathbf{z}_{k+1}|z_{1:k}) + H(\mathbf{x}_{k+1}|z_{1:k}). \quad (2.65)$$

Notons que l'entropie  $H(\mathbf{x}_{k+1}|z_{1:k})$  associée à la pdf de prédiction est indépendante de la commande  $\bar{u}_1 = u_k$  car le bruit de dynamique est négligé (théorème 2.1). Ainsi cette quantité peut être ajoutée à  $(-K_1)$  sous la forme d'une constante  $K_1'$ . Ensuite, en écrivant  $H(\mathbf{z}_{k+1}|z_{1:k})$  comme une fonction  $F_1$  du vecteur de commandes  $\bar{u}_1$ , l'expression (2.64) est obtenue. L'évaluation de  $F_1$ , basée sur les calculs du filtre UKF (voir annexe A), est donnée dans l'algorithme 2.1. ■

**Algorithme 2.1 :** Évaluation de  $F_1$

$$[H_1] = F_1(u_k)$$

**Entrées :**

—  $u_k$

**Paramètres :**

- Moments de la loi de filtrage  $p(x_k|z_{1:k}) : \hat{x}_{k|k}, P_{k|k}$
- Modèles de dynamique et d'observation :  $f, h, Q_k, R_k$
- Paramètre de la transformée unscented :  $\gamma \in \mathbb{R}^{+*}$

**Sorties :**

— Entropie  $H(\mathbf{z}_{k+1}|z_{1:k}) : H_1$

Application de l'algorithme de prédiction (A.1) avec les paramètres  $f, h, Q_k, R_k, \gamma$

$$1 \left[ \sim, \sim, \hat{z}_{k+1|k}, S_{k+1|k}, \sim \right] = \text{UKF\_Pred} \left( \hat{x}_{k|k}, P_{k|k}, u_k \right);$$

Calcul de l'entropie  $F_1 = H(\mathbf{z}_{k+1}|z_{1:k})$

$$2 F_1 = \frac{1}{2} \ln \left( (2\pi e)^2 |S_{k+1|k}| \right);$$

## B Critère de décision à long terme

Nous cherchons à généraliser l'expression (2.64), pour le cas d'un critère de décision à long terme, c'est à dire pour un horizon d'exploration  $N \geq 2$ . Tout d'abord, nous étendons respectivement les lemmes 2.1 et 2.2 pour le cas d'un critère de décision à long terme (lemmes 2.3, 2.4). Une nouvelle expression de  $J_N$  généralisant le théorème 2.2 pour tout  $N$ , est ensuite définie par le théorème 2.3.

**Lemme 2.3.** *L'espérance, sur l'observation  $\mathbf{z}_{k+i}$  conditionnellement aux observations  $z_{1:k+i-1}$ , de l'entropie  $H(\mathbf{x}_{k+i}|z_{1:k+i})$  de la loi de filtrage  $p(x_{k+i}|z_{1:k+i})$ , peut être décomposée en*

$$\mathbb{E}_{\mathbf{z}_{k+i}|z_{1:k+i-1}} [H(\mathbf{x}_{k+i}|z_{1:k+i})] = H(\mathbf{x}_{k+i}|z_{1:k+i-1}) - I_{k+i}. \quad (2.66)$$

□

*PREUVE.* La relation (2.57) du lemme 2.1 peut être vérifiée pour tous instants consécutifs  $k+i-1$  et  $k+i$  par un développement similaire à (2.29). ■

**Lemme 2.4.** *L'information mutuelle  $I_{k+i}$  des vecteurs d'état  $\mathbf{x}_{k+i}$  et d'observation  $\mathbf{z}_{k+i}$  conditionnellement aux mesures  $z_{1:k+i-1}$  peut être décomposée en*

$$I_{k+i} = K_i + H(\mathbf{z}_{k+i}|z_{1:k+i-1}), \quad (2.67)$$

avec  $K_i$  une constante indépendante de la séquence de vecteurs de commande  $\bar{u}_i$ . □

*PREUVE.* Cette propriété peut être facilement démontrée en utilisant le lemme 2.2 traitant du cas particulier  $i = 1$ . L'information mutuelle  $I_{k+i}$  pour tout  $i > 0$ , peut être décomposée en  $I_{k+i} = A + B$  avec

$$A = -\mathbb{E}_{\mathbf{x}_{k+i}|z_{1:k+i-1}} [H(\mathbf{z}_{k+i}|x_{k+i})], \quad (2.68)$$

et

$$B = H(\mathbf{z}_{k+i}|z_{1:k+i-1}). \quad (2.69)$$

Comme  $H(\mathbf{z}_{k+i}|x_{k+i})$  ne dépend que de la matrice de covariance du bruit de mesure  $R_{k+i}$ , la quantité  $A$  est donc constante par rapport à la séquence de vecteurs de commande  $\bar{u}_i = (u_k, u_{k+1}, \dots, u_{k+i-1})^T$ ; nous la renommons  $K_i$ . ■

**Theorème 2.3.** *Pour tout instant  $k$ , l'espérance sur les futures observations  $\mathbf{z}_{k+1:k+N}$  conditionnellement aux réalisations des observations passées  $z_{1:k}$ , de l'entropie  $H(\mathbf{x}_{k+N}|z_{1:k+N})$  de la loi de filtrage  $p(x_{k+N}|z_{1:k+N})$  à l'instant  $k + N$ , notée  $J_N$ , peut être décomposée en*

$$\begin{aligned} J_N(\bar{u}_N) &= \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})] \\ &= K'_N - H(\mathbf{z}_{k+1}|z_{1:k}) - \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [H(\mathbf{z}_{k+i}|z_{1:k+i-1})] \\ &= K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})], \end{aligned} \quad (2.70)$$

où  $K'_N$  est une constante scalaire indépendante du vecteur de commandes  $\bar{u}_N$ . L'entropie  $H(\mathbf{z}_{k+i}|z_{1:k+i-1})$  dépend de la séquence des vecteurs de commande  $\bar{u}_i = u_k, \dots, u_{k+i-1}$  de l'instant  $k$  à  $k + i - 1$  et des futures observations  $z_{k+1:k+i-1}$ . Nous notons cette entropie comme une fonction de cette séquence de vecteurs de commande et de ces observations,  $F_i(\bar{u}_i, z_{k+1:k+i-1})$  (l'évaluation de  $F_i$  est donnée dans l'algorithme 2.2). La fonction  $F_1$  est celle définie dans le théorème 2.2. □

*PREUVE.* La démonstration est effectuée par récurrence.

**Initialisation :** Validons l'équation (2.70) pour  $N = 2$ . Considérons la séquence des vecteurs de commande  $\bar{u}_2 = u_k, u_{k+1}$ , l'équation suivante peut être établie

$$\begin{aligned} J_2(\bar{u}_2) &= \mathbb{E}_{\mathbf{z}_{k+1:k+2}|z_{1:k}} [H(\mathbf{x}_{k+2}|z_{1:k+2})] \\ &= \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} \left[ \mathbb{E}_{\mathbf{z}_{k+2}|z_{1:k+1}} [H(\mathbf{x}_{k+2}|z_{1:k+2})] \right]. \end{aligned} \quad (2.71)$$

En utilisant le résultat du lemme 2.3 pour  $i = 2$ ,  $J_2$  devient

$$J_2(\bar{u}_2) = \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} [H(\mathbf{x}_{k+2}|z_{1:k+1}) - I_{k+2}]. \quad (2.72)$$

En considérant qu'il n'y a pas de perte d'information durant le mouvement, l'égalité  $H(\mathbf{x}_{k+2}|z_{1:k+1}) = H(\mathbf{x}_{k+1}|z_{1:k+1})$  est vérifiée (théorème 2.1). Ensuite, en utili-

**Algorithme 2.2 :** Évaluation de  $F_i$

$$H_i = F_i(\bar{u}_i, z_{k+1:k+i-1})$$

**Entrées :**

- Vecteur de commandes :  $\bar{u}_i = (u_k, u_{k+1}, \dots, u_{k+i-1})^T$
- Observations futures :  $z_{k+1:k+i-1} = z_{k+1}, z_{k+2}, \dots, z_{k+i-1}$

**Paramètres :**

- Moments de la loi de filtrage  $p(x_k | z_{1:k}) : \hat{x}_{k|k}, P_{k|k}$
- Modèles de dynamique et d'observation :  $f, h, Q_k = \emptyset, R_k$
- Paramètre de la transformée unscented :  $\gamma$

**Sorties :**

- Entropie  $H(z_{k+i} | z_{1:k+i}) : H_i$

Application de l'algorithme UKF\_PRED(.) (A.1) avec les paramètres  $f, h, Q_k, R_k, \gamma$

$$1 \quad \left[ \hat{x}_{k+1|k}, P_{k+1|k}, \hat{z}_{k+1|k}, S_{k+1|k}, T_{k+1|k} \right] = \text{UKF\_PRED} \left( \hat{x}_{k|k}, P_{k|k}, u_k \right);$$

Boucle sur le nombre d'observations

2 **Pour**  $j = 1, \dots, i - 1$  **faire**

3     Mise à jour de l'état courant  $k + j$  avec l'algorithme UKF\_PRED(.) (A.2)

$$4 \quad \left[ \hat{x}_{k+j|k+j}, P_{k+j|k+j} \right] = \text{UKF\_MAJ} \left( \hat{x}_{k+j|k+j-1}, P_{k+j|k+j-1}, \hat{z}_{k+j|k+j-1}, S_{k+j|k+j-1}, T_{k+j|k+j-1}, z_{k+j} \right);$$

5     Application de l'algorithme UKF\_PRED(.) (A.1) avec les paramètres  $f, h, Q_k, R_k, \gamma$

$$5 \quad \left[ \hat{x}_{k+j+1|k+j}, P_{k+j+1|k+j}, \hat{z}_{k+j+1|k+j}, S_{k+j+1|k+j}, T_{k+j+1|k+j} \right] = \text{UKF\_PRED} \left( \hat{x}_{k+j|k+j}, P_{k+j|k+j}, u_{k+j} \right);$$

6 **fin**

$$7 \quad H_i = \frac{1}{2} \ln \left( (2\pi e)^2 |S_{k+j+1|k+j}| \right);$$

sant (2.64), l'équation (2.72) devient

$$J_2(\bar{u}_2) = K'_1 - F_1(\bar{u}_1) - \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} [I_{k+2}]. \quad (2.73)$$

À partir du lemme 2.4 pour  $i = 2$ , l'information mutuelle  $I_{k+2}$  peut être remplacée par  $K_2 + H(\mathbf{z}_{k+2}|z_{1:k+1})$ , avec  $K_2$  une constante et  $H(\mathbf{z}_{k+2}|z_{1:k+1})$  l'entropie de la loi de prédiction de la mesure  $p(z_{k+2}|z_{1:k+1})$ , que nous renommons  $F_2(\bar{u}_2, z_{k+1})$  (l'entropie  $H(\mathbf{z}_{k+2}|z_{1:k+1})$  dépend de  $z_{k+1}$  et implicitement de  $z_{1:k}$ , mais est indépendante de  $z_{k+2}$ ). En regroupant les constantes  $K'_1, K_2$  en  $K'_2$ , l'équation (2.70) est vérifiée pour  $N = 2$

$$J_2(\bar{u}_2) = K'_2 - F_1(\bar{u}_1) - \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} [F_2(\bar{u}_2, z_{k+1})]. \quad (2.74)$$

**Démonstration de la récurrence :** Supposons que pour un certain  $N$ , la relation (2.70) soit vérifiée. Démontrons qu'elle le soit également pour  $N + 1$ . En décomposant l'espérance  $\mathbb{E}_{\mathbf{z}_{k+1:k+N+1}|z_{1:k}} [\cdot]$  en  $\mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [\mathbb{E}_{\mathbf{z}_{k+N+1}|z_{1:k+N}} [\cdot]]$  et en appliquant le lemme 2.3 pour  $i = N + 1$  à  $J_{N+1}(\bar{u}_{N+1})$  nous obtenons

$$\begin{aligned} J_{N+1}(\bar{u}_{N+1}) &= \mathbb{E}_{\mathbf{z}_{k+1:k+N+1}|z_{1:k}} [H(\mathbf{x}_{k+N+1}|z_{1:k+N+1})] \\ &= \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N+1}|z_{1:k+N}) - I_{k+N+1}]. \end{aligned} \quad (2.75)$$

Comme  $H(\mathbf{x}_{k+N+1}|z_{1:k+N}) = H(\mathbf{x}_{k+N}|z_{1:k+N})$ , l'égalité suivante peut être déduite

$$\begin{aligned} J_{N+1}(\bar{u}_{N+1}) &= J_N(\bar{u}_N) - \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [I_{k+N+1}] \\ &= K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})] \\ &\quad - \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [I_{k+N+1}]. \end{aligned} \quad (2.77)$$

À partir du lemme 2.4 pour  $i = N + 1$ , l'information mutuelle  $I_{k+N+1}$  peut être remplacée par la somme  $K_{N+1} + H(\mathbf{z}_{k+N+1}|z_{1:k+N})$ , avec  $K_{N+1}$  une constante et  $H(\mathbf{z}_{k+N+1}|z_{1:k+N})$  l'entropie de la loi de prédiction de la mesure  $p(z_{k+N+1}|z_{1:k+N})$  renommée en  $F_{N+1}(\bar{u}_{N+1}, z_{k+1:k+N})$ . En regroupant les constantes en  $K'_{N+1}$ , l'équation 2.70 est retrouvée pour  $N + 1$ . ■



# Génération de mouvement guidé par un critère de décision à court terme

---

## Sommaire

<b>3.1</b>	<b>Définition du problème d'optimisation</b>	<b>56</b>
<b>3.2</b>	<b>Intuitions sur le mouvement</b>	<b>58</b>
3.2.1	Mouvement guidé par les ITD	58
3.2.2	Mouvement guidé par des mesures linéaires en azimuth	61
<b>3.3</b>	<b>Lignes de niveau</b>	<b>62</b>
3.3.1	Exploration guidée par les ITD	63
3.3.2	Cas idéal de l'exploration guidée par des azimuths	66
3.3.3	Discussions	66
<b>3.4</b>	<b>Gradient de la fonction de récompense</b>	<b>69</b>
3.4.1	Définition de $F_1$	70
3.4.2	Évaluation du gradient en un point	71
<b>3.5</b>	<b>Optimisation sous contraintes</b>	<b>72</b>
3.5.1	Conditions d'optimalité	72
3.5.2	Résolution numérique	74
<b>3.6</b>	<b>Évaluation de la méthode</b>	<b>77</b>
3.6.1	Simulation	77
3.6.2	Expériences en conditions réelles	80

---

Ce chapitre concerne la génération de mouvement pour la localisation audio-motrice d'une source guidée par un critère d'exploration court terme. Le comportement du capteur binaural est donc établi à partir de la minimisation de la fonction d'incertitude  $J_1$  définie par le Théorème 2.2. Ce problème simplifié fait l'objet d'un chapitre dédié car il présente le double avantage d'admettre une implémentation logicielle de complexité limitée et de permettre une compréhension intuitive de l'influence du mouvement sur la localisation.

La section 3.1 définit un problème d'optimisation admettant pour solution une politique de déplacement du capteur binaural. Suit une partie 3.2 quelque peu empirique, où nous montrons comment un déplacement du capteur binaural impacte la pdf *a posteriori* de la position relative de la source obtenue par assimilation d'une

observation consécutivement au déplacement. Nous appuyons notre raisonnement sur le modèle d'observation spécifique à la stratégie d'exploration, défini en 1.3.4. Nous donnons ensuite en 3.3 une représentation géométrique du problème d'optimisation par une analyse des lignes de niveau de la fonction de coût. Nous établissons ensuite en 3.4 le gradient de la fonction de coût et nous mettons en place en 3.5 la résolution numérique du problème d'optimisation. Enfin l'évaluation de la méthode en simulation et en conditions réelles est proposée dans la partie 3.6.

### 3.1 Définition du problème d'optimisation

Nous supposons tout au long du chapitre que les conditions d'exploration établies dans la partie 1.3.4 ainsi que les hypothèses relatives à l'exploration définies en 2.3.1 (bruit de dynamique nul et loi de filtrage  $p(x_k|z_{1:k})$  gaussienne), qui permettent d'établir la fonction de récompense, sont vérifiées. Ainsi l'étude se place dans le cas où la position de la source à tout instant  $k$  est caractérisée par la loi *a posteriori* gaussienne  $p(x_k|z_{1:k})$ , avec  $z_{1:k}$  la séquence de mesures, réalisation du processus d'observation  $\mathbf{z}_{1:k}$  lié au processus d'état caché  $\mathbf{x}_{0:k}$  selon le modèle

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k = \bar{h}(\boldsymbol{\theta}_k) + \mathbf{v}_k, \quad \mathbf{z}_k \in \mathbb{R}, \quad \mathbf{v}_k \sim \mathcal{N}(0, R_k). \quad (3.1)$$

En chaque instant  $k$ , le vecteur d'état  $\mathbf{x}_k$  décrit la position de la source dans le repère  $\mathcal{F}_k$  lié au capteur. Le processus d'observation  $\mathbf{z}_{1:k}$  est utilisé pour guider l'exploration et remplace les variables  $\mathbf{y}_{1:k}$  dont le lien formel avec le processus d'état caché  $\mathbf{x}_{0:k}$  ne peut être explicité (voir section 1.3.4). Dans la partie 2.3, nous avons établi une fonction d'incertitude  $J_1$  qui représente l'entropie moyennée sur la future observation  $\mathbf{z}_{k+1}$  conditionnellement aux observations passées  $z_{1:k}$ , de la densité *a posteriori*  $p(x_{k+1}|z_{1:k+1})$ . Cette fonction dépend de la commande  $\bar{u}_1 = u_k$  appliquée entre  $k$  et  $k + 1$ . Elle a été définie comme

$$J_1(\bar{u}_1) = K'_1 - F_1(\bar{u}_1) \quad (3.2)$$

(Théorème (2.2)). Après l'application d'une commande, l'information maximale est donc obtenue, en moyenne, lorsque la valeur minimale de  $J_1$  est atteinte. Soit  $\bar{u}_1 = \bar{u}_1^*$  la commande qui minimise  $J_1$ . Dans la réalité, le déplacement du robot ainsi que la rotation de la tête étant limités en vitesse et en accélération, la commande  $\bar{u}_1^*$  obtenue par une minimisation sans contrainte n'est pas nécessairement admissible. Entre deux instants consécutifs  $k$  et  $k + 1$ , ces limitations sont traduites en un mouvement fini maximum. Définissons donc les ensembles

$$\mathcal{T} = \{(T_y, T_z) \in \mathbb{R}^2 \mid T_y^2 + T_z^2 \leq r_{\max}^2\}, \quad (3.3)$$

et

$$\mathcal{R} = \{\phi \in \mathbb{R} \mid |\phi| \leq \phi_{\max}\}, \quad (3.4)$$

qui correspondent respectivement aux translations et rotations admissibles du capteur binaural (*i.e.*, les amplitudes maximales de translation et de rotation qui peuvent être effectuées entre deux instants consécutifs).  $r_{\max}$  représente la distance maximale en translation et  $\phi_{\max}$  la rotation maximale. L'ensemble des contraintes  $\mathcal{T} \times \mathcal{R}$  constitue un volume cylindrique représenté en figure 3.1. L'axe vertical du cylindre correspond à la rotation de la tête et toute coupe horizontale constitue l'ensemble des positions atteignables en translation, pour une rotation de la tête fixe. La rotation nulle est fixée à mi-hauteur du cylindre. Ainsi un problème d'optimisation sous contraintes  $\mathcal{P}_1$  est défini dans l'énoncé suivant.

**Énoncé 3.1** (Problème  $\mathcal{P}_1$ ). *Soit la densité de probabilité gaussienne  $p(x_k|z_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k})$  du vecteur d'état  $\mathbf{x}_k$  à un instant  $k$  du processus de localisation, conditionnellement au fait que le processus d'observation  $\mathbf{z}_{1:k}$  se réalise en la séquence de mesures  $z_{1:k}$ . Soit*

$$J_1(\bar{u}_1) = \mathbb{E}_{\mathbf{z}_{k+1}|z_{1:k}} [H(\mathbf{x}_{k+1}|z_{1:k+1})] \quad (3.5)$$

$$= K'_1 - F_1(\bar{u}_1), \quad (3.6)$$

la fonction de coût définie comme l'espérance sur la future observation  $\mathbf{z}_{k+1}$  conditionnellement aux observations passées  $z_{1:k}$ , de l'entropie de la densité  $p(x_{k+1}|z_{1:k+1})$  à l'instant  $k+1$ , après l'application de la commande  $\bar{u}_1$  (voir section 2.3.2). Le vecteur de commande optimal  $\bar{u}_1^* = u_k^* = (T_k^*, \phi_k^*)^T$  recherché minimise  $J_1$  sous les contraintes  $T_k^* \in \mathcal{T}$  et  $\phi_k^* \in \mathcal{R}$ , *i.e.*, est solution du problème

$$(\mathcal{P}_1) \begin{cases} \bar{u}_1^* = \arg \min_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} J_1(\bar{u}_1) \\ = \arg \max_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} F_1(\bar{u}_1), \end{cases} \quad (3.7)$$

avec  $F_1$  la fonction définie dans l'algorithme 2.1 par

$$F_1 = \frac{1}{2} \ln \left( (2\pi e)^2 |S_{k+1|k}| \right), \quad (3.8)$$

où  $S_{k+1|k}$  désigne la variance de la densité de probabilité unidimensionnelle de la mesure prédite  $p(z_{k+1}|z_{1:k}) = \mathcal{N}(z_{k+1}; \hat{z}_{k+1|k}, S_{k+1|k})$ .

La résolution de  $\mathcal{P}_1$  à tout instant  $k$  permet donc d'obtenir le vecteur de commande  $\bar{u}_1^*$  dont l'application conduit à la minimisation de l'entropie moyenne de la densité de filtrage *a posteriori* à l'instant  $k+1$ . Pour simplifier les notations et ne pas créer de confusion dans les indices, la dépendance de  $\mathcal{P}_1$ ,  $J_1$  et  $F_1$  vis à vis de l'indice temporel  $k$  ne sera généralement pas explicitée. Ainsi, dans les parties suivantes, et par souci de cohérence avec le chapitre 4, nous préférons également la notation  $\bar{u}_1$  (avec  $\bar{u}_1 = u_k$ ) pour désigner le vecteur de commande.

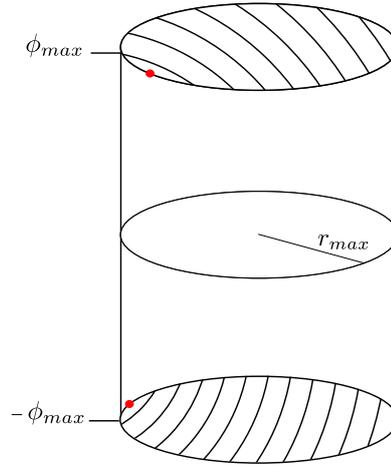


FIGURE 3.1 – Espace admissible de mouvements – Représentation du cylindre correspondant aux mouvements admissibles définis par l'ensemble  $\mathcal{T} \times \mathcal{R}$ . Des lignes de niveau de la fonction  $F_1$  ont été esquissées en fonction des translations, pour des valeurs fixes des rotations correspondant aux extrémités du cylindre ( $\phi = \pm\phi_{\max}$ ). Les marqueurs rouges désignent les valeurs maximales locales de  $F_1$  pour ces rotations.

## 3.2 Intuitions sur le mouvement

Dans cette partie, nous décrivons de manière intuitive la relation entre la position du capteur binaural et le processus de mise à jour de l'état. À partir d'une densité *a posteriori* gaussienne décrivant une connaissance de la position de la source, pour une situation relative capteur-source particulière, ainsi que d'un modèle de dynamique (1.10) et d'observation (3.1), nous décrivons comment le déplacement du capteur combiné à l'assimilation d'une mesure peut conduire à une réduction de l'incertitude sur la position de la source. Nous étudions deux configurations distinctes dans les parties 3.2.1 et 3.2.2 qui correspondent à un choix différent dans la fonction  $h$  (ou  $\bar{h}$ ) du modèle (3.1).

### 3.2.1 Mouvement guidé par les ITD

Considérons que la fonction  $\bar{h}$  du modèle d'observation (3.1) traduise une mesure d'ITD. Comme le capteur binaural est placé sur un élément diffusant, il peut s'agir de l'approximation des ITD de Woodworth-Schlosberg dont nous rappelons l'équation

$$\bar{h}(\theta) = \begin{cases} \frac{a}{c}(\sin(\theta) + \theta) & \theta \in [-\frac{\pi}{2}; \frac{\pi}{2}] \\ \frac{a}{c}(\sin(\theta) - \theta + \pi) & \theta \in ]\frac{\pi}{2}; \pi] \\ \frac{a}{c}(\sin(\theta) - \theta - \pi) & \theta \in ]-\pi; -\frac{\pi}{2}[ \end{cases}, \quad (3.9)$$

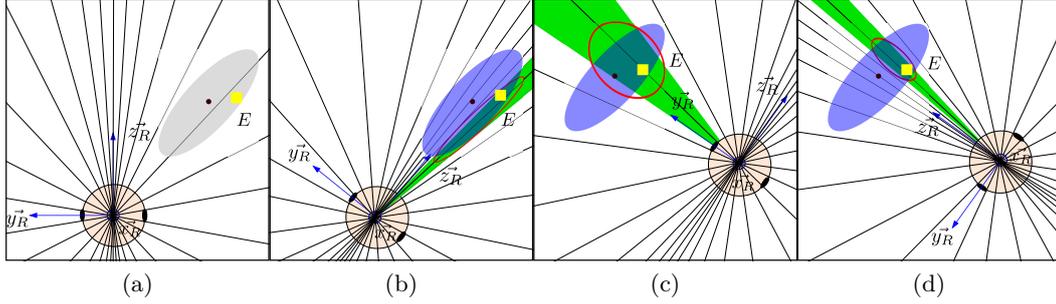


FIGURE 3.2 – Combinaison du mouvement et de l’observation – Les iso- $z$  sont représentées par les droites grises. Le repère  $\mathcal{F}_k$  bleu est attaché à la tête binaurale. Le carré jaune représente la position réelle de la source. En (a) l’ellipse grise est la zone de confiance associée à son estimation à l’instant  $k$ . En (b),(c) et (d), la situation (position et orientation) du capteur à l’instant suivant  $k+1$  est représentée. L’ellipse de confiance bleue est associée à la loi de prédiction du vecteur d’état (position relative capteur-source) après application de la commande. La zone de confiance associée à la densité de probabilité du vecteur d’état que l’on obtiendrait sur la base de la mesure en  $k+1$  seule est un cône (vert) de sommet le centre de la tête. L’ellipse associée à la densité de filtrage après incorporation de la mesure à l’instant  $k+1$  est esquissée en rouge.

avec  $\theta = -\text{atan2}(e_y, e_z)$  l’azimut de la source (voir figure 1.1a). Une hypothèse fondamentale du modèle d’observation (3.1) est que la dispersion du bruit de mesure (la variance  $R_k$ ) est supposée égale à une constante connue indépendante de la réalisation du vecteur aléatoire d’état  $\mathbf{x}_k$  exprimant la position relative capteur-source. Cependant le modèle étant non linéaire, des intervalles de taille identique autour de différentes valeurs de  $\bar{h}(\theta)$  (tels des intervalles de confiance relatifs à une probabilité donnée de réalisation du bruit de mesure additif défini en (3.1)) correspondent à des intervalles de tailles variables en azimuts. Notons que la fonction  $\bar{h}$  est impaire et  $2\pi$  périodique. Elle est quasi linéaire au voisinage de  $\theta = 0$  ou  $\theta = \pm\pi$  et la non-linéarité devient plus importante pour  $\theta \approx \pm\frac{\pi}{2}$ .

Pour une position relative capteur-source (cachée)  $x_k$ , la valeur de l’observation  $z_k$  peut être associée à une droite rigidement liée au repère  $\mathcal{F}_k$  et passant par le centre de la tête  $O_k$ , comme représenté schématiquement sur la figure 3.2a. La valeur de  $z_k$  est donc constante sur une même droite de l’espace des positions relatives capteur-source. Par la suite, un ensemble de telles droites indexées par des valeurs de  $z$  régulièrement espacées sera désigné par le terme « iso- $z$  ». La non-linéarité du modèle évoqué ci-dessus a pour conséquence une distribution non uniforme des iso- $z$  autour du capteur. Elles sont plus concentrées selon la direction  $\vec{z}_R$  qui correspond à l’avant et l’arrière de la tête (fovéa). Au contraire, l’espacement entre les iso- $z$  est plus grand à proximité de l’axe interaural  $\vec{y}_R$ .

Les figures 3.2b, 3.2c et 3.2d représentent l’évolution de l’incertitude sur la position relative capteur-source après application du vecteur de commande

$u_k = (T_k, \phi_k)^T$  au capteur et assimilation de la mesure  $z_{k+1}$  à l'issue du mouvement. Toutes les lois normales relatives au vecteur d'état (position relative, dans  $\mathbb{R}^2$ , de la source relativement au capteur) sont représentées par les ellipses de confiance correspondant à une probabilité donnée de réalisation de celui-ci : en bleu pour la densité de prédiction  $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k})$ , et en rouge pour la densité de filtrage  $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k+1}, P_{k+1|k+1})$ . Il convient de noter que si le bruit de dynamique est négligé pour guider l'exploration<sup>1</sup>, alors la densité de prédiction à l'instant  $k + 1$  est décrite par la même ellipse de confiance que celle représentant la loi de filtrage à l'instant  $k$ , mais « vue » depuis le capteur à l'issue de son déplacement.

Ensuite, la pdf de la position de la source dans le repère du capteur, conditionnellement au seul fait que la variable d'observation  $z_{k+1}$  se réalise en une mesure  $z_{k+1}$ , peut être décrite par un cône de confiance dont le sommet est le centre de la tête.<sup>2</sup> Ce cône constitue donc l'ensemble de confiance associé à la distribution de probabilité de la situation relative capteur-source déduite de la seule observation  $z_{k+1}$ . Son étendue (de part et d'autre de l'iso- $z$ <sup>3</sup>), est plus importante lorsque les iso- $z$  sont espacés, *e.g.*, le long de l'axe interaural.

Dans la partie 3.1, nous avons montré que la commande du capteur binaural est obtenue par la résolution du problème  $\mathcal{P}_1$ , *i.e.*, la maximisation de la fonction  $F_1$  (3.8). Le sens de variation de  $F_1$  est le même que  $|S_{k+1|k}|$ , avec  $S_{k+1|k}$  la variance de la loi de prédiction de la mesure  $\mathcal{N}(z_{k+1}; \hat{z}_{k+1|k}, S_{k+1|k})$ . Cette variance intervient dans l'étape de mise à jour du filtre, lorsque la loi de prédiction  $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k})$  et la densité issue de l'observation  $z_{k+1}$  seule, sont fusionnées en la densité de filtrage  $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k+1}, P_{k+1|k+1})$ . Intuitivement, la variance  $S_{k+1|k}$  de la densité de prédiction de la mesure à l'instant  $k + 1$  augmente avec le nombre d'iso- $z$  qui intersectent l'ellipse de confiance associée à la densité de prédiction de l'état  $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k})$ . En effet, l'intervalle de confiance correspondant à une certaine probabilité de réalisation de la loi de prédiction de la mesure, est intuitivement lié à la gamme des valeurs de  $z$  indiquant les iso- $z$  qui intersectent l'ellipse de confiance correspondant à la même probabilité de réalisation de la loi de prédiction du vecteur d'état. Dès lors, l'assimilation de l'observation  $z_{k+1}$  à l'issue du mouvement du capteur réduit d'autant plus l'incertitude dans la densité de filtrage que l'intersection du cône de confiance établi sur la base de  $z_{k+1}$  et de l'ellipse de confiance associé à la densité de prédiction est restreinte. Ces considérations permettent l'analyse qualitative de la loi de filtrage obtenue pour différents déplacements du capteur binaural.

1. Un bruit de dynamique non nul est pris en compte dans la stratégie de localisation audiomotrice (étape B).

2. Cette pdf est notée  $p(x_{k+1}|z_{k+1})$ . La densité de filtrage  $p(x_{k+1}|z_{1:k+1})$  (de moyenne  $\hat{x}_{k+1|k+1}$  et de covariance  $P_{k+1|k+1}$ ) est effectivement obtenue par la fusion de la loi de prédiction  $p(x_{k+1}|z_{1:k})$  (de moyenne  $\hat{x}_{k+1|k}$  et de covariance  $P_{k+1|k}$ ) et de la densité  $p(x_{k+1}|z_{k+1})$  (de moyenne  $\hat{x}_{k+1}$  et de covariance  $P_{k+1}$ ). Cette fusion est décrite par les équations du filtre information  $P_{k+1|k+1}^{-1} = P_{k+1|k}^{-1} + P_{k+1}^{-1}$  et  $P_{k+1|k+1}^{-1} \hat{x}_{k+1|k+1} = P_{k+1|k}^{-1} \hat{x}_{k+1|k} + P_{k+1}^{-1} \hat{x}_{k+1}$ .

3. Au vu de la non-linéarité de  $\hat{h}$ , l'axe de symétrie du cône de confiance ne correspond pas exactement à la valeur indexée par  $z_{k+1} = \hat{h}(\theta_{k+1})$ .

Dans un premier temps, à partir de la situation initiale décrite dans la figure 3.2a, le capteur binaural effectue une rotation de façon à aligner, à l'instant  $k + 1$ , l'axe  $\vec{z}_{R_{k+1}}$  de la fovéa avec le grand axe de l'ellipse associée à la densité de prédiction (figure 3.2b). Sur la figure 3.2c, une translation est appliquée au capteur de façon à placer, à l'instant  $k + 1$ , son centre  $O_{k+1}$  sur la droite portée par le petit axe de l'ellipse de prédiction. Une rotation dirige par la suite l'axe  $\vec{y}_{R_{k+1}}$  vers le centre de cette ellipse. La dernière stratégie, en figure 3.2d, dirige, à l'instant  $k + 1$ , l'axe de la fovéa auditive vers le petit axe de l'ellipse de prédiction en combinant également translation et rotation. Comme évoqué précédemment, le cône de confiance établi sur la base de l'observation  $z_{k+1}$  est large lorsque la source se trouve au voisinage de l'axe interaural, comme c'est le cas sur la figure 3.2c. Dans ce cas, un petit nombre d'iso- $z$  intersectent l'ellipse associée à la densité de prédiction en  $k + 1$ . Par conséquent, après l'assimilation de  $z_{k+1}$ , l'incertitude de la loi *a posteriori* ne diminue pas significativement. Au contraire, quand la fovéa est orientée vers la source, le cône de confiance est plus étroit et l'étape de mise à jour du filtre conduit à une réduction d'incertitude plus significative (figures 3.2b et 3.2d). Dans le cas de la figure 3.2d, la situation relative entre l'ellipse de prédiction en  $k + 1$  et le capteur binaural est encore plus favorable car il intersecte un nombre élevé d'iso- $z$ . Par les arguments heuristiques mentionnés précédemment, la variance  $S_{k+1|k}$  de la loi de prédiction de la mesure s'en trouve accrue. Notons enfin que quand le capteur se rapproche de la source, l'incertitude liée à la position de celle-ci sur la base de la seule observation  $z_{k+1}$  devient plus faible (la source est plus proche du sommet du cône de confiance). Ainsi, un plus grand nombre d'iso- $z$  intersectent l'ellipse de prédiction, ce qui entraîne également une augmentation de la variance  $S_{k+1|k}$ .

Finalement nous pouvons conclure qu'à tout instant  $k$ , la réduction d'entropie de la loi de filtrage à l'instant  $k + 1$  sera d'autant plus importante si les actions suivantes sont appliquées :

1. déplacer le capteur binaural de façon que l'axe  $\vec{z}_R$  (fovéa) pointe dans la direction de l'ellipse de confiance de la loi de prédiction, *i.e.*, de sorte que l'azimut mesuré à l'issue du déplacement soit voisin de  $\theta = 0$  ou  $\theta = \pm\pi$  ;
2. déplacer le capteur binaural de façon à placer le centre du repère  $O_{k+1}$  dans la direction du petit axe de l'ellipse.
3. réduire la distance entre le capteur et la source.

La dernière règle heuristique citée, bien que peu intuitive compte tenu des résultats en audition prônant généralement une indépendance de la mesure d'ITD vis à vis de la distance capteur-source, sera discutée dans la partie 3.3.3.

### 3.2.2 Mouvement guidé par des mesures linéaires en azimut

Supposons que (3.1) soit remplacé par un modèle d'observation de la forme

$$\mathbf{z}_k = \boldsymbol{\theta}_k + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0, R_k). \quad (3.10)$$

Notons qu'un tel modèle où l'azimut est directement « contaminé » par un bruit de mesure de variance constante, n'est pas réaliste. En effet, l'incertitude associée à l'extraction d'azimut à partir du flux binaural est plus faible (*resp.* importante) lorsque la source est proche de l'axe de la fovéa (*resp.* de l'axe interaural). Nous imaginons ce modèle linéaire uniquement dans le but de mettre en avant d'éventuelles différences avec le modèle non linéaire précédent et ainsi pressentir quelles caractéristiques du mouvement actif peuvent être imputées aux non-linéarités<sup>4</sup>.

Dans ce contexte, les iso- $z$ , *i.e.*, les zones de l'espace d'état (positions relatives capteur-source) correspondant à des observations constantes, indexées par des valeurs régulièrement espacées, sont des droites radiales passant par le centre de la tête et espacées d'un angle constant. Du fait que le bruit de mesure demeure additif centré de variance constante  $R_k$ , les cônes de confiance (associés à la distribution de probabilité de la situation relative capteur-source établie sur la base de l'observation seule) ont la même taille, quel que soit l'azimut  $\theta_k$  de la source. En tenant compte de ces considérations, les règles heuristiques (1), (2) et (3) déclinées en fin de section précédente (section 3.2.1) sont réévaluées ci-dessous. Comme les droites iso- $z$  sont régulièrement espacées, l'assimilation, depuis une position fixe (translation nulle), d'une mesure dont la valeur non bruitée serait linéaire en l'azimut, apporte intuitivement la même quantité d'information quelle que soit l'orientation du capteur par rapport à la source. Dès lors, la règle (1) n'est plus valide. Au contraire, les règles (2) et (3) demeurent en vigueur. En effet, la nouvelle configuration des iso- $z$  n'a pas d'incidence sur (2) car cette règle ne dépend que de l'orientation de l'ellipse de confiance par rapport au capteur. Enfin concernant la règle (3), la réduction de la distance entre le capteur et la source permet effectivement d'augmenter le nombre d'intersections entre les droites iso- $z$  et l'ellipse de confiance.

### 3.3 Lignes de niveau

Pour compléter l'analyse effectuée en 3.2, nous caractérisons dans cette section des instances particulières du problème ( $\mathcal{P}_1$ ) paramétrées en sélectionnant des valeurs des moments  $\hat{x}_{k|k}$  et  $P_{k|k}$  de la loi de filtrage  $p(x_k|z_{1:k}) = \mathcal{N}(x_k|k; \hat{x}_{k|k}, P_{k|k})$  à l'instant  $k$ . La fonction  $F_1$  est évaluée en fonction de la translation  $T_k = (T_y, T_z)^T$  pour une rotation fixe  $\phi_k = \phi_0$ , de façon à tracer ses lignes de niveau dans le plan  $(O_k, \vec{y}_{Rk}, \vec{z}_{Rk})$ . Ainsi, pour tout point de coordonnées  $(T_y, T_z)$ , la projection du gradient de  $F_1$  dans le plan  $(O_k, \vec{y}_{Rk}, \vec{z}_{Rk})$  est orthogonale à la tangente de la ligne de niveau passant par ce point et indique la direction de plus forte pente de  $F_1$ . Un exemple est donné en figure 3.3. Le cercle autour du capteur binaural représente les limites du déplacement admissible en translation (il s'agit d'un plan de coupe perpendiculaire à l'axe de révolution du cylindre représenté en figure 3.1). Nous représenterons également les lignes de niveau de la restriction  $\tilde{F}_1$  de  $F_1$  à la surface du cylindre des mouvements admissibles. Les variables de cette fonction évoluent

4. L'analyse qui suit pourrait être faite avec le modèle  $\mathbf{z}_k = l\theta_k + \mathbf{v}_k$ ,  $l \in \mathbb{Z}^*$

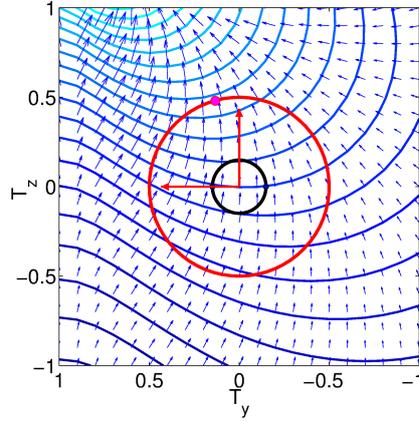


FIGURE 3.3 – Lignes de niveau et gradient – Les lignes de niveau ainsi que le gradient de  $F_1$  sont évalués en différents points du plan  $(O_k, T_y, T_z)$ . Aucune rotation du repère de la tête n'est appliquée. Le cercle rouge délimite les translations admissibles. Le point sur le cercle marque le maximum de la fonction sur le domaine admissible.

donc sur la surface de ce cylindre et  $\tilde{F}_1$  est définie de telle sorte que

$$\tilde{F}_1(\alpha, \phi) = [F_1 \circ g](\alpha, \phi), \quad (3.11)$$

avec

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}^3 \quad (3.12)$$

$$\begin{pmatrix} \alpha \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} r_{\max} \sin(\alpha) \\ r_{\max} \cos(\alpha) \\ \phi \end{pmatrix}.$$

Dans les figures 3.4, 3.5, 3.6, 3.7, et 3.8, les lignes de niveau des fonctions  $F_1$  et  $\tilde{F}_1$  ont été tracées en fonction du mouvement du capteur. Dans les figures relatives à la fonction  $F_1$ , le repère rouge est constitué des axes  $\vec{y}_{Rk}$  (horizontal) et  $\vec{z}_{Rk}$  (vertical) du repère  $\mathcal{F}_k$  lié au capteur binaural à l'instant  $k$ . Le cercle rouge délimite le domaine  $\mathcal{T}$  des translations  $T_k$  ( $T_y$  en abscisse,  $T_z$  en ordonnée) admissibles. Le repère bleu décrit l'orientation du repère  $\mathcal{F}_{k+1}$  obtenu par application successive de  $T_k$  et de la rotation d'angle  $\phi_0$  à  $\mathcal{F}_k$ . La loi de filtrage  $\mathcal{N}(x_{k|k}; \hat{x}_{k|k}, P_{k|k})$  à l'instant  $k$  est schématisée par l'ellipse de confiance bleue. Les couleurs des lignes de niveau deviennent plus « chaudes » lorsqu'elles correspondent à des valeurs croissantes. Le maximum de  $F_1$  sur  $\mathcal{T}$  et pour  $\phi$  fixé est mis en valeur.

### 3.3.1 Exploration guidée par les ITD

Considérons le cas où la fonction  $\bar{h}$  correspond à l'approximation des ITD de Woodworth-Schlosberg. Les iso- $z$  sont les « iso-ITD » décrits dans la section 3.2.1. Les lignes de niveau de  $F_1$  sont tracées dans la figure 3.4 en fonction des translations

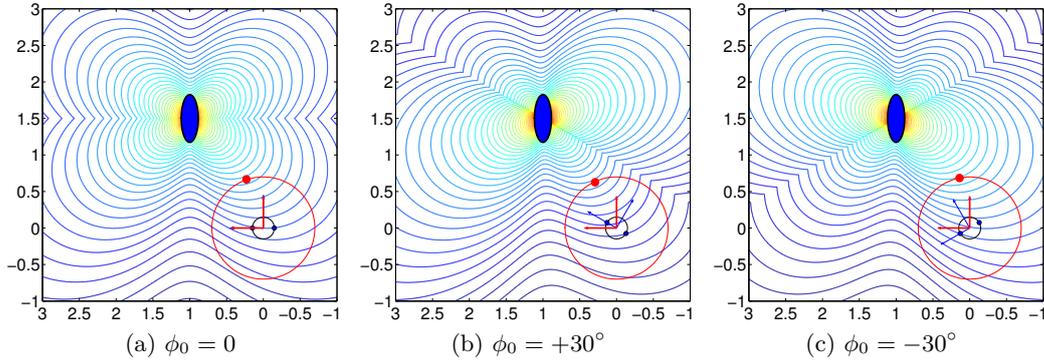


FIGURE 3.4 – Lignes de niveau du critère  $F_1(\bar{u}_1)$  par rapport à  $T_y$  (abscisses, en mètres) et  $T_z$  (ordonnées, en mètres) pour différentes valeurs de la rotation ultérieure  $\phi_0$ . L’exploration est guidée par les ITD, partant d’une loi de filtrage telle que  $\hat{x}_{k|k} = (1, 1.5)^T$  et  $P_{k|k} = \begin{pmatrix} 0.0033 & 0 \\ 0 & 0.0268 \end{pmatrix}$ .

$T_k = (T_y, T_z)^T$ , pour différentes rotations  $\phi = \phi_0$  de la tête. La loi de filtrage avant le déplacement du capteur satisfait  $\hat{x}_{k|k} = (1, 1.5)^T$  et  $P_{k|k} = \begin{pmatrix} 0.0033 & 0 \\ 0 & 0.0268 \end{pmatrix}$ .

La figure 3.4a correspond au cas sans rotation ( $\phi_0 = 0$ ). Nous observons que les lignes de niveau de  $F_1$  sont déformées et se rapprochent davantage de la position de la source lorsque la translation vaut  $T_k = (1, \cdot)^T$  ou  $T_k = (\cdot, 1.5)^T$ . Il s’agit de positions défavorables, où l’information acquise sera moins importante. Ainsi, en imaginant un déplacement en translation du capteur le long d’un cercle de centre  $\hat{x}_{k|k}$ , la valeur de la fonction  $F_1$  est plus faible lorsque le capteur se trouve à l’intersection de ce cercle et des axes  $T_k = (1, \cdot)^T$  ou  $T_k = (\cdot, 1.5)^T$ . Nous parlerons de directions ou d’axes de « déformations », qui sont ici parallèles aux axes  $\vec{y}_{Rk}$  et  $\vec{z}_{Rk}$ . Ces déformations sont cohérentes avec les règles heuristiques avancées en 3.2 : pour  $T_k = (1, \cdot)^T$ , des déformations apparaissent car le centre  $O_{k+1}$  se trouve sur la direction portée par le grand axe de l’ellipse ; les déformations au niveau de  $T_k = (\cdot, 1.5)$  correspondent au cas où l’axe interaural pointe vers l’ellipse. Notons que la valeur séparant deux lignes de niveau consécutives reste constante. En conséquence, lorsque le capteur se rapproche de la source et que les lignes deviennent plus proches les unes des autres, cela signifie que l’augmentation de l’information pour une même distance parcourue est plus importante. Néanmoins, rappelons que le modèle d’approximation des ITD de Woodworth-Schlosberg est valable sous l’hypothèse de champ lointain de la source. Par conséquent, les lignes de niveau situées trop près de l’ellipse de confiance doivent être considérées avec précaution. Nous discuterons de ce problème dans la partie 3.3.3. Les cas où la rotation du capteur binaural n’est pas nulle sont reportés dans les figures 3.4b ( $\phi_0 = +30^\circ$ ) et 3.4c ( $\phi_0 = -30^\circ$ ). Nous observons que la déformation des lignes due à la direction du grand axe de l’ellipse reste présente dans la direction de l’axe  $T_k = (1, \cdot)^T$ . Cependant la déformation imputée à l’axe interaural est tournée d’un angle équivalent à celui de la rotation appliquée au capteur binaural. La direction de déformation

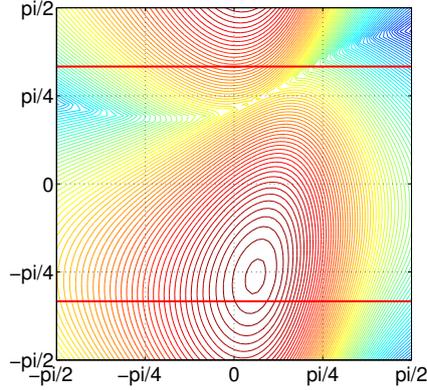


FIGURE 3.5 – Lignes de niveau de la fonction  $\tilde{F}_1(\alpha, \phi)$  par rapport à  $\alpha$  (abscisses en radians) et  $\phi$  (ordonnées en radians). Les axes rouges horizontaux représentent la rotation maximale de la tête fixée à  $\phi_{\max} = \frac{\pi}{3}$ . Les valeurs de  $\hat{x}_{k|k}$  et  $P_{k|k}$  sont les mêmes que dans la figure 3.4

reste donc parallèle à l'axe interaural  $\vec{y}_{Rk}$ . Nous observons que les solutions du problème  $\mathcal{P}_1$ , qui correspondent aux translations dans  $\mathcal{T}$  qui maximisent  $F_1$  pour une rotation  $\phi$  fixe, diffèrent. Il est ainsi plus intéressant d'appliquer une rotation  $\phi_0 = -30^\circ$  que  $\phi_0 = +30^\circ$ , car le point maximum obtenu par la rotation  $\phi_0 = -30^\circ$  est situé sur une ligne de niveau plus haute. Notons que dans les situations décrites, la solution du problème  $\mathcal{P}_1$  se retrouve sur la surface du cylindre, ce qui est cohérent avec la remarque de la section 3.2 selon laquelle l'information apportée par la mesure est plus importante lorsque le capteur se rapproche de la source. Les contours de la fonction  $\tilde{F}_1$  définie en (3.11) ont été reportés sur la figure 3.5 pour des mêmes valeurs de  $\hat{x}_{k|k}$  et  $P_{k|k}$ . Dans cet exemple, nous observons que le maximum de  $\tilde{F}_1$ , obtenu en  $(\alpha^*, \phi^*)$  est également le maximum de  $F_1$ , de sorte que  $\vec{u}_1^* = (T_y^*, T_y^*, \phi_k^*) = (r_{\max} \sin(\alpha^*), r_{\max} \cos(\alpha^*), \phi_k^*)$ .

Dans certaines situations, comme par exemple pour  $\hat{x}_{k|k} = (0, 1.5)^T$ , le problème  $\mathcal{P}_1$  peut admettre plusieurs solutions (voir figure 3.6). Cette situation a peu de chances de se réaliser dans la pratique car cela suppose un alignement parfait de l'axe  $\vec{z}_{Rk}$  avec l'axe principal de l'ellipse de confiance.

Dans les situations précédemment observées (figures 3.4 et 3.6), la solution du problème  $\mathcal{P}_1$  se situe sur la surface du cylindre et, au vu de ces observations ainsi que des conclusions de la partie 3.2, la position relative de la source par rapport au capteur  $\hat{x}_{k|k}$ , estimée avant le mouvement, semble correspondre au maximum de  $F_1$ . Ainsi, nous conjecturons que la solution de  $\mathcal{P}_1$  se trouve sur la surface du cylindre si et seulement si  $\hat{x}_{k|k}$  n'appartient pas à l'ensemble  $\mathcal{T}$ . En l'absence de forme analytique de la fonction  $F_1$ , il semble difficile d'apporter une démonstration formelle à cette conjecture. En conséquence, la méthode d'optimisation proposée dans la partie 3.5 doit permettre de rechercher le maximum de  $F_1$  sur l'ensemble du volume définissant la contrainte.

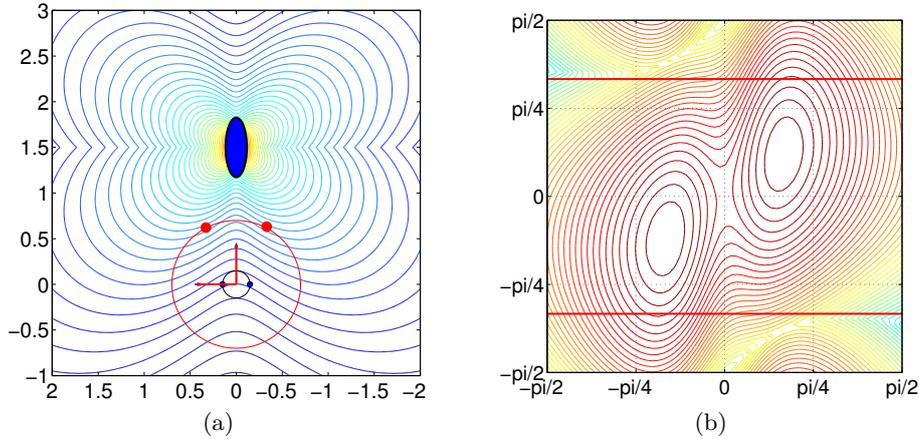


FIGURE 3.6 – Lignes de niveau des fonction  $F_1$  et  $\tilde{F}_1$  quand le problème  $\mathcal{P}_1$  admet deux solutions – (a) Lignes de niveau de  $F_1(\bar{u}_1)$  par rapport à  $T_y$  (abscisses, en mètres) et  $T_z$  (ordonnées, en mètres) avec  $\phi = 0$ . Les maxima admissible de  $F_1$  pour la rotation  $\phi = 0$  sont représentés par une marque sur le cercle de contrainte. (b) Lignes de niveau de  $\tilde{F}_1(\alpha, \phi)$  par rapport à  $\alpha$  (en abscisses) et  $\phi$  (en ordonnées). L’exploration est guidée par les ITD, partant d’une loi de filtrage telle que  $\hat{x}_{k|k} = (0, 1.5)^T$  et  $P_{k|k} = \begin{pmatrix} 0.0033 & 0 \\ 0 & 0.0268 \end{pmatrix}$ .

### 3.3.2 Cas idéal de l’exploration guidée par des azimuts

Considérons brièvement le cas abordé dans la section 3.2.2, où l’exploration est guidée par un modèle de la forme (3.10). L’analyse des lignes de niveau de la fonction  $F_1$  par rapport à  $(T_y, T_z)$  montre que celles-ci ne dépendent pas de la rotation  $\phi$  (figures 3.7a et 3.7b). Les lignes de niveau de  $\tilde{F}_1$  par rapport à  $(\alpha, \phi)$  sont alors verticales (figure 3.8). Notons cependant que les lignes de niveau dans les figures 3.7a et 3.7b sont toujours déformées dans la direction du grand axe de l’ellipse pour les mêmes raisons que celles évoquées dans la partie précédente.

Ces déformations disparaissent quand l’ellipse de confiance à 99% de probabilité de la loi  $p(x_k|z_{1:k})$  est circulaire (figure 3.7c). Les lignes de niveau sont alors des cercles concentriques. Dans ce dernier cas, la seule manière d’augmenter l’information apportée par la future mesure est de rapprocher le capteur de la source.

### 3.3.3 Discussions

Les règles heuristiques de la partie 3.2 et les observations effectuées sur les lignes de niveaux de  $F_1$  et  $\tilde{F}_1$  (sections 3.3.1) renseignent sur le mouvement du capteur binaural, résultant de la résolution du problème  $\mathcal{P}_1$ . Cependant la trajectoire du capteur binaural (*i.e.*, la succession de plusieurs translations et rotations) obtenue par la résolution à chaque instant du problème  $\mathcal{P}_1$ , ne sera satisfaisante que si la fonction  $\bar{h}$  qui caractérise la relation entre l’azimut de la source et les ITD, est suffisamment représentative de la réalité. En particulier, il faut s’assurer que certains phénomènes non modélisés dans la stratégie d’estimation court terme de

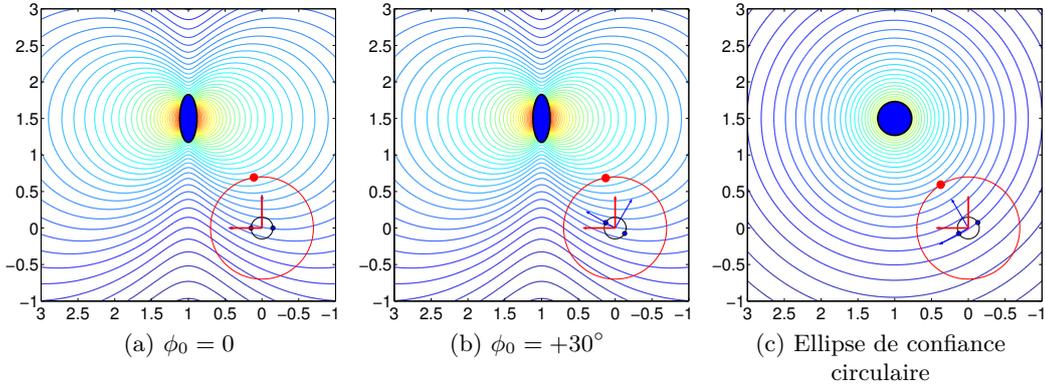


FIGURE 3.7 – Lignes de niveau de la fonction  $F_1(\bar{u}_1)$  par rapport à  $T_y$  (abscisses, en mètres) et  $T_z$  (ordonnées, en mètres), avec différentes valeurs de  $\phi_0$ . L'exploration est guidée par les mesures linéaires en azimuth, partant d'une loi de filtrage telle que  $\hat{x}_{k|k} = (1, 1.5)^T$  et  $P_{k|k} = \begin{pmatrix} 0.0033 & 0 \\ 0 & 0.0268 \end{pmatrix}$  (et  $P_{k|k} = \begin{pmatrix} 0.0134 & 0 \\ 0 & 0.0134 \end{pmatrix}$  en (c)). Le maximum de  $F_1$  sur le domaine  $\mathcal{T}$ , pour la rotation  $\phi = \phi_0$ , est représenté par une marque sur le cercle de contrainte.

l'azimut présentée en 1.3.2 (réverbérations, phénomènes négligés en champ proche, etc.), ne sont pas exacerbés par la trajectoire du capteur ce qui provoquerait une détérioration de l'estimation de l'azimut.

La politique de déplacement mise en place incite le capteur binaural à se rapprocher de la source. Comme nous l'avons vu, ce comportement améliore la qualité de la fusion du mouvement et des observations. Notons toutefois que ce résultat peut paraître contre-intuitif compte tenu des études réalisées sur l'audition binaurale humaine ou en robotique. En effet, les ITD sont relativement indépendants de la distance capteur-source, tant pour un modèle sphérique de la tête [Duda & Martens 1998] que pour un modèle basé sur un mannequin KEMAR [Brungart & Rabinowitz 1999]. Dans le modèle sur lequel repose notre stratégie d'estimation (voir la définition en 1.3.2), nous utilisons les HRIR établies expérimentalement sur un HATS KEMAR pour différentes distances entre le capteur et la source. Nous considérons donc notre modèle d'estimation de l'azimut relativement robuste à un changement de distance. Notons toutefois que lorsque la source et le capteur deviennent très proches, des effets non pris en compte par notre modèle peuvent apparaître. Dans notre application, les données HRIR que nous utilisons ne sont pas adaptées à une source en champ très proche et ne prennent pas en compte l'effet « parallaxe » [Brungart & Rabinowitz 1999]. Ce problème peut être expliqué de la façon suivante. Soit une source en champ lointain, placée de façon que l'onde sonore atteigne le microphone  $R_1$  en premier. L'angle que forme l'axe reliant le centre de la tête et la source avec l'axe  $\vec{z}_R$ , est quasiment identique à l'angle formé par l'axe allant du microphone  $R_1$  à la source, avec l'axe parallèle à  $\vec{z}_R$  passant par ce même microphone. En champ proche, si la source est située juste à côté de  $R_1$  la différence entre ces deux angles n'est plus négligeable. Ainsi, la straté-

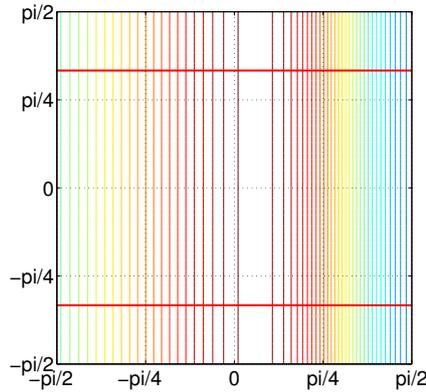


FIGURE 3.8 – Lignes de niveau de la fonction  $\tilde{F}_1(\alpha, \phi)$  – L’exploration est basée sur des mesures idéales d’azimuts. L’axe horizontal porte la variation selon la variable  $\alpha$  et l’axe vertical la variation selon  $\phi$ . Les axes rouges horizontaux représentent la rotation maximale de la tête. La position estimée de la source est  $\hat{x}_{k|k} = (1, 1.5)^T$ .

gie d’exploration dépendant d’un modèle d’observation en champ lointain, incite le capteur binaural à se déplacer vers la source sans restrictions, ce qui peut entraîner des problèmes dans l’estimation de l’azimut lorsque le capteur se retrouve trop proche de la source. Notre stratégie présente donc une faiblesse en ce point. De plus, notre dispositif expérimental n’est pas adapté à des tests en champ très proche (robot encombrant et haut parleurs posés sur des trépieds). Notons toutefois qu’en environnement réverbérant, le ratio entre l’énergie acoustique atteignant le capteur directement et l’énergie qui atteint le capteur via une ou plusieurs zones réverbérantes est une fonction monotone décroissante de la distance quadratique à la source [Zahorik *et al.* 2005]. Comme notre modèle ne prend pas en compte les réverbérations, le fait de se rapprocher de la source a un effet positif dans le sens où cela permet d’atténuer la perception des réverbérations qui pourraient venir « parasiter » l’estimation court terme d’azimut. Déplacer le capteur binaural de façon à le rapprocher de la source ne semble donc ne pas poser de problème tant que la distance entre les deux ne devient pas trop faible. Dans nos expériences en conditions réelles, nous avons choisi de diminuer le rayon du cylindre de contrainte lorsque la distance capteur-source estimée diminue (voir 3.6.2) afin de ne pas risquer une collision entre la source et le robot.

Ensuite nos intuitions sur le modèle d’observation et l’analyse des lignes de niveau nous montrent que la politique de déplacement incite le capteur binaural à orienter l’axe  $\vec{z}_R$  de la fovéa dans la direction de la source. Ce résultat est cohérent avec la localisation binaurale chez l’humain et en robotique où il est admis que la localisation de l’azimut d’une source est plus précise lorsque la source se situe dans l’axe de la fovéa. Par ailleurs, cette configuration permet également d’améliorer la qualité de la fonction de pseudo vraisemblance de l’azimut (1.3.2) établie à partir du flux audio [Portello 2013].

Enfin notre analyse effectuée en 3.2 nous montre que placer l'axe formé par le centre  $O_{k+1}$  et le premier moment  $\hat{x}_{k+1|k}$  de la loi  $p(x_{k+1}|z_{1:k})$ , dans la direction du petit axe de l'ellipse de confiance associée à la loi de prédiction, améliore également la fusion d'information. Dans la pratique, cela peut être difficilement exploitable. En effet, de façon empirique nous pouvons observer que les ellipses de confiance se forment généralement de façon que le grand axe soit aligné dans la direction définie par les centres de la tête et de l'ellipse (voir figure 3.2a par exemple). Ainsi, comme le mouvement du capteur binaural est contraint, le capteur ne peut pas se déplacer en une seule itération, au voisinage du petit axe de l'ellipse

### 3.4 Gradient de la fonction de récompense

Afin de déterminer le vecteur de commande  $\bar{u}_1^*$  solution de  $\mathcal{P}_1$ , une première approche naïve pourrait être d'évaluer la fonction de récompense  $J_1$  (3.2) pour un nombre fini de commandes et de sélectionner celle qui conduit à la plus petite valeur de  $J_1$ . En dehors de sa simplicité, une telle approche présente un inconvénient majeur : elle nécessite de déterminer un ensemble fini de commandes pour lesquelles la fonction  $J_1$  sera évaluée. Se pose alors le problème de la taille de cet ensemble et/ou de comment sélectionner les commandes à évaluer. Un nombre trop réduit de commandes entraînera une solution trop approximative alors qu'un nombre élevé aura pour conséquence une explosion du temps de calcul. Notons que si cette approche quelques peu triviale est à éviter pour la résolution du problème à court terme, elle devient complètement inadaptée aux problèmes à long terme que nous étudierons dans le chapitre 4.

Afin d'éviter une telle approche, le calcul du gradient de  $J_1$  par rapport au vecteur de commandes  $\bar{u}_1$  est crucial pour étudier les variations de l'information avec le mouvement du capteur et pour résoudre le problème d'optimisation. En outre, le gradient d'une fonction indique sa direction de plus forte pente et permet également d'établir la condition nécessaire d'optimalité d'un problème d'optimisation (voir annexe B), à condition qu'il puisse être exprimé sous forme analytique. Par ailleurs, de nombreuses méthodes numériques d'optimisation exploitent le gradient de la fonction à optimiser pour converger vers un optimum (voir section 3.5.2). En appliquant l'opérateur gradient

$$\nabla_{\bar{u}_1} = \begin{bmatrix} \frac{\partial}{\partial T_y} \\ \frac{\partial}{\partial T_z} \\ \frac{\partial}{\partial \phi} \end{bmatrix}, \quad (3.13)$$

à l'expression de  $J_1$  nous obtenons

$$\nabla_{\bar{u}_1} J_1(\bar{u}_1) = \nabla_{\bar{u}_1} K_1' - \nabla_{\bar{u}_1} F_1(\bar{u}_1), \quad (3.14)$$

$$= -\nabla_{\bar{u}_1} F_1(\bar{u}_1), \quad (3.15)$$

car la quantité  $K'_1$  dans l'expression de  $J_1$  est indépendante des variables de déplacement  $\bar{u}_1$ . Dans la suite de ce chapitre, l'étude portera sur l'évaluation de  $\nabla F_1$ . La fonction  $F_1$  ne pouvant pas être exprimée analytiquement, le gradient ne peut pas être déduit directement. Néanmoins comme  $F_1$  peut être calculée en un point particulier  $\bar{u}_1 = u$  à partir de l'algorithme (2.1), le gradient pourrait être évalué par une méthode numérique de type différence finie. Les méthodes à différences finies sont cependant connues pour provoquer des problèmes de précision numérique. Ainsi, la taille du pas de discrétisation doit être ajustée à la fonction à dériver. Cette méthode n'est donc pas satisfaisante pour notre application car les valeurs que prend la fonction  $F_1$  changent aux différents instant  $k$  du processus de localisation en fonction de la situation relative capteur-source et de l'estimation faite de la position de la source (le problème d'optimisation  $\mathcal{P}_1$  est différent en chaque instant  $k$ ).

Pour éviter cet écueil nous proposons une méthode permettant d'évaluer le gradient en un point à partir de la décomposition en série de Taylor de la composition des fonctions qui constituent  $F_1$ . Cette méthode permet l'évaluation rapide de  $\nabla F_1$  en un point particulier  $\bar{u}_1 = u$ .

### 3.4.1 Définition de $F_1$

Dans le but de définir le gradient  $\nabla F_1$  et en nous appuyant sur l'algorithme d'évaluation de  $F_1$  (algorithme 2.1) présenté en 2.3.2, définissons la fonction  $F_1$  comme la composition de plusieurs fonctions dont la première est une fonction du vecteur de commandes  $\bar{u}_1 = (T_y, T_z, \phi)^T$ . Considérons la loi gaussienne  $p(x_k | z_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k})$ . Sur la base de cette loi, une distribution unidimensionnelle  $\mathcal{N}(z_{k+1}; \hat{z}_{k+1|k}, S_{k+1|k})$  constituant une approximation gaussienne de la loi de prédiction de la mesure peut être établie au moyen de la transformée unscented, dont la définition est rappelée en annexe A.1.

Dans un premier temps, les *sigma-points*  $\{X_i^-\}$  de la loi *a posteriori*  $p(x_k | z_{1:k})$ , sont calculés à partir de sa moyenne  $\hat{x}_{k|k}$  et de la racine de Cholesky de la matrice de covariance  $P_{k|k} = L_{k|k} L_{k|k}^T$ , *i.e.*,

$$\left[ \{X_i^-\}, \{w_i^w\}, \{w_i^w\} \right] = \text{sp} \left( \hat{x}_{k|k}, L_{k|k}, \gamma \right), \quad (3.16)$$

avec  $\text{sp}(\cdot)$  la fonction définie en (A.3) et  $\gamma$  un paramètre de réglage.

À partir des *sigma-points*  $\{X_i^-\}$ , en exploitant l'équation de dynamique (1.10), qui dépend du vecteur de commande  $\bar{u}_1 = (T_y, T_z, \phi)^T$ , les *sigma-points*  $\{X_i^+\}$  de la loi de prédiction  $p(x_{k+1} | z_{1:k}) = \mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k})$  sont alors obtenus. Pour chaque *sigma-point*  $X_i^-$  nous remplaçons la fonction de l'équation de dynamique  $x_{k+1} = f(x_k, u_k)$  par une fonction de  $\bar{u}_1 = (T_y, T_z, \phi)$ , de sorte que

$$\forall i, X_i^+ = f_{X_i^-}(\bar{u}_1). \quad (3.17)$$

Les *sigma-points*  $\{X_i^-\}$  sont alors vus comme des paramètres de la fonction  $f_{X_i^-}$ , que nous pouvons réécrire en

$$f_{X_i^-} : \begin{pmatrix} T_y \\ T_z \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} f_{X_i^-[1]}(T_y, T_z, \phi) \\ f_{X_i^-[2]}(T_y, T_z, \phi) \end{pmatrix}, \quad (3.18)$$

avec  $X_i^-[1]$  et  $X_i^-[2]$  les composantes de chaque *sigma-point*  $X_i^-$ . Ensuite, l'ensemble des *sigma-points*  $\{Z_i^+\}$  de la pdf de la prédiction de la mesure  $p(z_{k+1}|z_{1:k}) = \mathcal{N}(z_{k+1}; \hat{z}_{k+1|k}, S_{k+1|k})$ , est obtenu à partir des  $\{X_i^+\}$  définis en (3.17), par

$$\forall i, Z_i^+(\bar{u}_1) = \bar{h} \left( -\text{atan2} \left( X_i^+[1](\bar{u}_1), X_i^+[2](\bar{u}_1) \right) \right), \quad (3.19)$$

avec  $X_i^+[1]$  et  $X_i^+[2]$  les composantes de  $X_i^+$ , et  $\bar{h}$  (3.9) l'équation de mesure (modèle de Woodworth-Schlosberg) utilisée dans le modèle d'observation (3.1), pour guider l'exploration. Finalement l'espérance  $\hat{z}_{k+1|k}$  et la variance  $S_{k+1|k}$  de  $p(z_{k+1}|z_{1:k})$  sont obtenues par

$$\hat{z}_{k+1|k}(\bar{u}_1) = \sum_i w_m^i Z_i^+(\bar{u}_1), \quad (3.20)$$

$$S_{k+1|k}(\bar{u}_1) = \sum_i w_c^i \left( Z_i^+(\bar{u}_1) - \hat{z}_{k+1|k}(\bar{u}_1) \right)^2, \quad (3.21)$$

avec  $w_m^i$  et  $w_c^i$  les poids de la transformée unscented. La fonction d'entropie  $H(\mathbf{z}_{k+1}|z_{1:k})$  que nous notons  $F_1$  est définie en appliquant la fonction

$$l(v) = \frac{1}{2} \log(\alpha |v|), \quad (3.22)$$

au déterminant de  $S_{k+1|k}$ , avec  $\alpha = (2\pi e)$ . Nous obtenons finalement

$$F_1(\bar{u}_1) = \frac{1}{2} \log \left( \alpha \left| S_{k+1|k}(\bar{u}_1) \right| \right), \quad (3.23)$$

avec  $\left| S_{k+1|k}(\bar{u}_1) \right| = S_{k+1|k}(\bar{u}_1)$  car  $S_{k+1|k}$  est scalaire.

### 3.4.2 Évaluation du gradient en un point

Définissons le développement de Taylor au premier ordre de la fonction  $f_{X_i^-}$  autour de  $\bar{u}_1 = u$  en appliquant un déplacement infinitésimal  $du = (dT_y, dT_z, d\phi)^T$ , de sorte que

$$f_{X_i^-}(u + du) = f_{X_i^-}(u) + Jf_{X_i^-}(u) du, \quad (3.24)$$

avec  $Jf_{X_i^-}(u)$  la matrice jacobienne de  $f_{X_i^-}$  évaluée au point  $u$ , définie par

$$Jf_{X_i^-} = \begin{pmatrix} \frac{\partial f_{X_i^- [1]}}{\partial T_y} & \frac{\partial f_{X_i^- [1]}}{\partial T_z} & \frac{\partial f_{X_i^- [1]}}{\partial \phi} \\ \frac{\partial f_{X_i^- [2]}}{\partial T_y} & \frac{\partial f_{X_i^- [2]}}{\partial T_z} & \frac{\partial f_{X_i^- [2]}}{\partial \phi} \end{pmatrix}. \quad (3.25)$$

Ce résultat est ensuite composé avec les développements de Taylor au premier ordre des fonctions  $\text{atan2}$  et  $\bar{h}$ ,

$$\text{atan2}(v, w) = \text{atan2}(v_0, w_0) + \nabla^T \text{atan2}(v_0, w_0) \begin{pmatrix} v - v_0 \\ w - w_0 \end{pmatrix} \quad (3.26)$$

$$\bar{h}(v) = \bar{h}(v_0) + \bar{h}'(v_0)(v - v_0). \quad (3.27)$$

Le résultat de la composition noté  $Z_i(u + du)$  est ensuite utilisé pour retrouver la moyenne  $\hat{z}(u + du)$  et la covariance  $S(u + du)$  à partir des équations (3.20) et (3.21). Enfin, la valeur  $S(u + du)$  est composée avec le développement de Taylor au premier ordre de la fonction  $l$  (3.22), *i.e.*,

$$l(v) = l(v_0) + l'(v_0)(v - v_0) \quad (3.28)$$

Ainsi le développement de Taylor au premier ordre de  $F_1(u + du)$  est obtenu

$$F_1(u + du) = F_1(u) + \nabla_{\bar{u}_1} F_1(u)^T du, \quad (3.29)$$

avec  $\nabla_{\bar{u}_1} F_1(u)$  le gradient de  $F_1$  évalué en  $\bar{u}_1 = u$ .

### 3.5 Optimisation sous contraintes

Dans cette partie nous recherchons des solutions au problème  $\mathcal{P}_1$ . Comme  $F_1$  ne peut être exprimé analytiquement la recherche d'une solution optimale de  $\mathcal{P}_1$  au moyen d'une analyse formelle soulève des problèmes que nous évoquons dans la partie 3.5.1. Nous proposons une approche numérique pour la résolution de  $\mathcal{P}_1$  dans la partie 3.5.2.

#### 3.5.1 Conditions d'optimalité

Les conditions d'optimalité des problèmes d'optimisation non contraints (conditions nécessaires du premier et second ordre : annulation du gradient et matrice hessienne symétrique définie positive; voir Annexe B.2), ne permettent pas de caractériser les valeurs optimales des problèmes sous contraintes.

Reformulons  $\mathcal{T} \times \mathcal{R}$  l'ensemble des variables admissibles du problème  $\mathcal{P}_1$ , en un ensemble  $\mathcal{C}$

$$\mathcal{C} = \left\{ x = (x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid g_i(x) \leq 0, \forall i = 1, \dots, 3 \right\}, \quad (3.30)$$

de sorte que

$$\bar{u}_1^* = \arg \max_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} F_1(\bar{u}_1) \quad (3.31)$$

$$= \arg \max_{\bar{u}_1 \in \mathcal{C}} F_1(\bar{u}_1). \quad (3.32)$$

Soient les fonctions  $g_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $g_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$  et  $g_3 : \mathbb{R}^3 \rightarrow \mathbb{R}$  définies par

$$g_1(x) = x_1^2 + x_2^2 - r_{\max}^2 \quad (3.33)$$

$$g_2(x) = x_3 - \phi_{\max} \quad (3.34)$$

$$g_3(x) = -x_3 - \phi_{\max}. \quad (3.35)$$

Soit  $u \in \mathcal{C}$  une variable de décision. Pour tout  $i = 1, \dots, 3$ , une contrainte  $g_i$  est dite saturée (*resp.* non saturée) si  $g_i(u) = 0$  (*resp.* si  $g_i(u) < 0$ ).

Une contrainte non saturée en  $u$  ne joue localement aucun rôle puisque tout point voisinage  $\mathcal{V}(u)$  de  $u$ , est admissible par rapport à cette contrainte. Au contraire, lorsqu'une contrainte est saturée en  $u$ ,  $\mathcal{V}(u)$  n'est pas nécessairement admissible. Soit  $E(u)$  l'ensemble des contraintes saturées simultanément lorsque  $\bar{u}_1 = u$ . Cet ensemble peut être égal à  $\{g_1\}$ ,  $\{g_2\}$ ,  $\{g_3\}$ ,  $\{g_1, g_2\}$ ,  $\{g_1, g_3\}$ . Les conditions de qualification des contraintes  $(QC)_u$  en un point  $u \in \mathcal{C}$ , sont exprimées dans le théorème (3.1).

**Theorème 3.1** (Condition de qualification des contraintes  $(QC)_u$ ). *Soit  $u \in \mathcal{C}$ , et soit  $E(u)$  l'ensemble des contraintes saturées simultanément en  $u$ . S'il n'existe pas de vecteur non nul  $\mu \in (\mathbb{R}^+)^n$  avec  $n = \text{Card}(E(u))$ , tel que  $\sum_{k=i}^n \mu_k \nabla g_k(u) = 0$  avec  $g_k$  le  $k^{\text{ème}}$  élément de  $E(u)$ , alors l'hypothèse  $(QC)_u$  de qualification des contraintes est vérifiée en  $u$ .  $\square$*

Pour le problème  $\mathcal{P}_1$ , ces conditions peuvent être facilement vérifiées en tout point  $u$  appartenant à la surface du cylindre. Les conditions de [Karush Kuhn Tucker \(KKT\)](#) (voir annexe B.3) permettent d'établir des conditions nécessaires d'optimalité pour les problèmes d'optimisation possédant des contraintes sous forme d'inégalités, combinées ou non avec des contraintes de type égalité. les conditions [KKT](#) appliquées à  $\mathcal{P}_1$  sont exprimées dans le théorème 3.2.

**Theorème 3.2.** *Si  $u^* \in \mathcal{C}$  maximise  $F_1$  sur  $\mathcal{C}$ , alors il existe un vecteur  $\mu = (\mu_0, \dots, \mu_3)^T$  non nul tel que*

$$\mu_0 \nabla F_1(u^*) + \sum_{i=1}^3 \mu_i \nabla g_i(u^*) = 0 \quad (3.36)$$

$$\mu_i \leq 0 \quad \forall i = 0, \dots, 3 \quad (3.37)$$

$$\mu_i g_i(u^*) = 0 \quad \forall i = 0, \dots, 3. \quad (3.38)$$

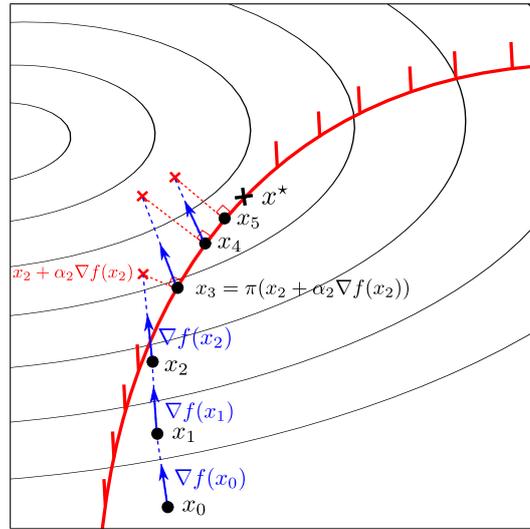


FIGURE 3.9 – Itérations de l'algorithme du gradient projeté – Le résultat des itérations  $j = 0, \dots, 5$  d'un algorithme de gradient projeté est représenté. La courbe rouge représente la contrainte et les segments verticaux désignent la zone non admissibles. Le maximum local de la fonction sur la contrainte est désigné par la marque  $x^*$ .

Comme la condition de qualification des contraintes  $(QC)_{u^*}$  sur  $u^*$  est vérifiée alors  $\mu_0 \neq 0$  et l'équation (3.36) peut être réécrite en

$$\nabla F_1(u^*) + \sum_{i=1}^3 \mu_i \nabla g_i(u^*) = 0. \quad (3.39)$$

□

L'équation (3.38) exprime que  $\mu_i = 0$  si la  $i^{\text{ème}}$  contrainte est non saturée ( $g_i < 0$ ). Par contre  $\mu_i$  peut être nul même en cas de saturation de  $g_i$ . Dans l'équation (3.37), la résolution de (3.39) pour toutes les configurations de saturation des contraintes possibles ( $\{g_1\}, \{g_2\}, \{g_3\}, \{g_1, g_2\}, \{g_1, g_3\}$ ), permet de mettre en avant les possibles solutions de  $\mathcal{P}_1$ . Cependant (3.39) ne peut pas être résolue de façon formelle car une forme analytique de  $F_1$  ne peut être exprimée. Par la suite nous résolvons donc  $\mathcal{P}_1$  par une méthode numérique.

### 3.5.2 Résolution numérique

La résolution du problème  $\mathcal{P}_1$  peut être effectuée numériquement au moyen de l'algorithme de gradient projeté. Cette méthode est une adaptation, pour l'optimisation sous contrainte, de l'algorithme de montée de gradient classique qui permet de résoudre les problèmes d'optimisation sans contrainte. Après avoir présenté les principes de ces deux algorithmes [Nocedal & Wright 2006], nous détaillons l'implémentation de l'algorithme du gradient projeté, pour résoudre le problème  $\mathcal{P}_1$ .

**Algorithme du gradient.** L'algorithme de montée de gradient permet de rechercher le maximum (ou le minimum, il s'agit alors d'un algorithme de descente de gradient) d'une fonction sans contrainte sur les variables de décision. Soit le problème d'optimisation sans contrainte défini par

$$x^* = \arg \max_x f(x), \quad (3.40)$$

avec  $x \in \mathbb{R}^n$ . L'algorithme de montée de gradient est une méthode itérative, qui partant d'une valeur initiale  $x_0$ , permet de se déplacer dans l'espace des variables de décision selon les directions données par le gradient  $\nabla f$  de  $f$  en chaque point  $x_i$ , de manière à atteindre un maximum. Les variables de décision  $x_j$  sont mises à jour de sorte que

$$x_{j+1} = x_j + \alpha_j d_j, \quad (3.41)$$

avec  $d_j = \nabla f(x_j)$  la direction de plus forte pente<sup>5</sup>. Le coefficient  $\alpha_j$  représente la distance parcourue le long de cette direction. Il peut être fixé à une valeur définie à l'instant initial, ou être calculé à chaque itération. Une valeur de  $\alpha_j$  trop importante induira un trop grand déplacement sur la fonction  $f$ , avec le risque de dépasser le point maximum. Au contraire, une valeur trop petite ne permettra pas de se déplacer suffisamment. Les conditions de Wolfe ou de Goldstein permettent de sélectionner un pas de descente adéquat [Nocedal & Wright 2006]. Afin de satisfaire de telles conditions, le pas  $\alpha_j$  peut être obtenu à chaque itération au moyen d'un algorithme de recherche linéaire.

La condition d'arrêt d'un algorithme de montée de gradient peut être définie de plusieurs manières. Il peut s'agir de tester les conditions nécessaires d'optimalité du premier ordre, c'est à dire vérifier si le gradient au point obtenu est nul (voir B.2). D'autres critères peuvent se baser sur l'observation de l'évolution de l'algorithme en définissant une valeur seuil  $\varepsilon$  à partir de laquelle l'algorithme doit être stoppé, par exemple,

$$\|x_{j+1} - x_j\| \leq \varepsilon, \quad (3.42)$$

ou

$$\|f(x_{j+1}) - f(x_j)\| \leq \varepsilon. \quad (3.43)$$

Enfin, une autre solution consiste à stopper l'algorithme quand un nombre maximum d'itérations défini a priori est dépassé.

---

5. La direction de plus forte pente n'est pas choisie systématiquement dans toutes les méthodes de montée ou de descente. Par exemple, dans l'algorithme de descente de Newton,  $d_j = -\nabla^2 f(x_j)^{-1} \nabla f(x_j)$ , avec  $\nabla^2 f(x_j)^{-1}$  l'inverse de la matrice Hessienne de  $f$ .

**Algorithme du gradient projeté.** Considérons maintenant un problème d'optimisation sous contraintes, exprimé par

$$x^* = \arg \max_{x \in \mathcal{D}} f(x), \quad (3.44)$$

avec  $\mathcal{D}$  le domaine admissible, supposé convexe, de la variable de décision  $x$ . Pour résoudre ce problème, l'équation (3.41) ne permet pas de déterminer l'argument qui maximise  $f$  sous la contrainte (si  $x_j \in \mathcal{D}$ , rien ne garantit que  $x_{j+1} = x_j + \alpha_j d_j \in \mathcal{D}$ ). Pour pallier ce problème, la valeur obtenue par l'application de (3.41) est projetée orthogonalement sur la contrainte. La projection orthogonale sur l'ensemble convexe des contraintes  $\mathcal{D}$  est définie par l'opérateur de projection

$$\pi_{\mathcal{D}}(x) \triangleq \arg \min_v \{\|x - v\|_2, v \in \mathcal{D}\}. \quad (3.45)$$

La mise à jour des valeurs est alors effectuée par la relation

$$x_{j+1} = \pi_{\mathcal{D}}(x_j + \alpha d_j), \quad (3.46)$$

qui remplace (3.41). En conséquence, les valeurs prises par l'algorithme « glissent » le long de la contrainte (voir figure 3.9) pour atteindre le point maximum recherché. Notons que le pas de descente est fixe. Un critère d'arrêt peut être mis en place en vérifiant les conditions KKT, en vérifiant les équations (3.42) et (3.43) ou en définissant un nombre maximal d'itérations.

**Résolution de  $\mathcal{P}_1$ .** Afin de résoudre le problème  $\mathcal{P}_1$

$$\bar{u}_1^* = \arg \max_{\bar{u}_1 \in \mathcal{C}} F_1(\bar{u}_1), \quad (3.47)$$

au moyen de l'algorithme du gradient projeté, définissons tout d'abord la projection orthogonale d'un élément  $x \in \mathbb{R}^3$  sur l'ensemble  $\mathcal{C}$  par

$$\pi_{\mathcal{C}}(x) \triangleq \arg \min_v \{\|x - v\|_2, v \in \mathcal{C}\}. \quad (3.48)$$

Soit  $u = (T_y, T_z, \phi) \in \mathcal{C}$ . Notons que la fonction  $g_1$  est indépendante de  $\phi$  et que  $g_2$  ainsi que  $g_3$  sont indépendantes de  $T = (T_y, T_z)^T$ . Ainsi la projection d'un élément  $x \in \mathbb{R}^3$  sur l'ensemble  $\mathcal{C}$ , est équivalente aux projections successives  $\pi_{\mathcal{T}}(x)$  de  $x \in \mathbb{R}^2$  sur  $\mathcal{T}$  et  $\pi_{\mathcal{R}}(y)$  de  $y \in \mathbb{R}$  sur  $\mathcal{R}$  (l'ordre est sans importance). Ces projections sont définies par

$$\pi_{\mathcal{T}}(x) \triangleq \arg \min_v \{\|x - v\|_2, v \in \mathcal{T}\}, \quad (3.49)$$

et

$$\pi_{\mathcal{R}}(x) \triangleq \arg \min_v \{\|x - v\|_2, v \in \mathcal{R}\}. \quad (3.50)$$

L'algorithme du gradient projeté pour la résolution du problème  $\mathcal{P}_1$  est détaillé dans l'algorithme 3.1. Pour faciliter la projection de  $u = (T_y, T_z, \phi)^T$  sur  $\mathcal{C}$ , nous définissons  $v = (\rho, \theta, \phi)^T$  le vecteur de coordonnées cylindriques par la transformation

$$\begin{cases} \rho &= \sqrt{(T_y^2 + T_z^2)} \\ \theta &= \text{atan2}(T_y, T_z) \\ \phi &= \phi. \end{cases} \quad (3.51)$$

Définissons également la transformation inverse par

$$\begin{cases} T_y &= \rho \sin(\theta) \\ T_z &= \rho \cos(\theta) \\ \phi &= \phi. \end{cases} \quad (3.52)$$

Dans la solution que nous proposons, le pas  $\alpha$  est fixé heuristiquement et le critère d'arrêt est défini par une combinaison de l'équation (3.42) et de la limitation du nombre d'itérations.

## 3.6 Évaluation de la méthode

De façon à évaluer la stratégie d'exploration court terme développée, le schéma de localisation complet présenté en figure 1.2 a été évalué à la fois en simulation (3.6.1) et en conditions réelles à partir d'un robot binaural (3.6.2). Ce schéma de localisation comporte : (1) la détection court terme d'azimut ; (2) la localisation audio-motrice par la combinaison de l'observation issue de (1) avec le mouvement ; (3) la stratégie d'exploration permettant d'établir la commande du capteur en résolvant le problème  $\mathcal{P}_1$  à chaque instant  $k$  du processus. Les paramètres de la simulation ainsi que les conditions de l'expérience ont été décrits dans la section 1.4. Le mouvement du capteur binaural a été limité en translation et en rotation par  $r \leq r_{\max} = 0.1\text{m}$  et  $|\phi| \leq \phi_{\max} = 15^\circ$ . En simulation 3.6.1 comme en conditions réelles 3.6.2, plusieurs mouvements du capteur binaural ont été simulés : (I) la stratégie de mouvement actif par la résolution à chaque itération du problème  $\mathcal{P}_1$  ; (II) une translation dans la direction de l'axe interaural porté par le vecteur  $\vec{y}_R$  ; (III) un mouvement circulaire tel que l'avant de la tête (vecteur  $\vec{z}_R$ ) reste tangent à la trajectoire.

### 3.6.1 Simulation

Nous considérons un état initial de la simulation dans lequel une connaissance *a priori*  $p(x_0)$  gaussienne est disponible. À cet instant, la position estimée de la source dans le repère du robot  $\mathcal{F}_{k=0}$ , est définie par  $\hat{x}_0 = (1.5, 1.5)^T$  et  $P_0 = \begin{pmatrix} 0.0121 & 0.0074 \\ 0.0074 & 0.0231 \end{pmatrix}$ . La vraie position a été tirée aléatoirement dans cet état de croyance initial. Nous reportons respectivement dans les figures 3.10a et 3.10b la trajectoire du capteur

---

**Algorithme 3.1 :** Gradient projeté pour la résolution de  $\mathcal{P}_1$

---

**Entrées :**

- Fonction de récompense :  $F_1$
- Translation maximale admissible du capteur binaural :  $r_{\max}$
- Rotation maximale admissible du capteur binaural :  $\phi_{\max}$
- Opérateur de projection sur l'ensemble  $\mathcal{T} \times \mathcal{R}$  :  $\pi_{\mathcal{T} \times \mathcal{R}}(\cdot)$

**Paramètres :**

- Moments de l'état de croyance initial à l'instant  $k$  :  $\hat{x}_{k|k}, P_{k|k}$
- Pas de la descente :  $\alpha$
- Variable d'itération et itération maximale :  $i, M$
- Seuil pour l'arrêt de l'algorithme  $\varepsilon$

**Sorties :**

- $u^* = (T_y^*, T_z^*, \phi^*)^T \triangleq u_M = (T_{yM}, T_{zM}, \phi_M)^T$
- 

**Initialisation**

- 1  $u_0 = (T_{y,0}, T_{z,0}, \phi_0)^T$  ;
  - 2  $v_0 = (\rho_0, \theta_0, \phi_0)^T$  ;
  - 3  $u_{\text{prec}} = (\infty, \infty, \infty)^T$  ;
  - 4  $i = 0$  ;
- 

**Tant que  $i < M$  ET  $\|u_i - u_{\text{prec}}\|_2 \geq \varepsilon$  faire**

- 5 | Évaluation de  $d_i = \nabla F_1(u_i)$ , où  $F_1$  est défini sur la base de  $\hat{x}_{k|k}, P_{k|k}$  ;
  - 6 | Calcul de  $u_{i+1}$  avant projection :  $u_{i+1} = u_i + \alpha d_i$  ;
  - 7 | Transformation polaire de  $u_{i+1}$  :  $v_{i+1} = (\rho_{i+1}, \theta_{i+1}, \phi_{i+1})^T$  ;
  - 8 | **Projection** de  $u_{i+1}$  sur  $\mathcal{T} \times \mathcal{R}$  :  $u_{i+1} = \pi_{\mathcal{T} \times \mathcal{R}}(u_{i+1})$
  - 9 |     **Si**  $\rho_{i+1} > r_{\max}$  **alors** // Projection sur  $\mathcal{T}$
  - 10 |     |  $T_{y,i+1} = r_{\max} \sin(\theta_{i+1})$  ;
  - 11 |     |  $T_{z,i+1} = r_{\max} \cos(\theta_{i+1})$  ;
  - 12 |     **fin**
  - 13 |     **Si**  $|\phi_{i+1}| > \phi_{\max}$  **alors** // Projection sur  $\mathcal{R}$
  - 14 |     |  $\phi_{i+1} = \text{sign}(\phi_{i+1}) \phi_{\max}$  ;
  - 15 |     **fin**
  - 16 | **fin**
  - 17 |  $u_{\text{prec}} = u_i$  ;
  - 18 |  $u_{i+1} = (T_{y,i+1}, T_{z,i+1}, \phi_{i+1})^T$  ;
  - 19 |  $i = i + 1$  ;
  - 20 **fin**
-

binaural et l'entropie  $H(\mathbf{x}_k|y_{1:k})$  de la pdf *a posteriori*  $p(x_k|y_{1:k})$ , pour chaque stratégie. Afin d'obtenir des résultats comparables, les mouvements en boucle ouverte (stratégies (II) et (III)) ont les mêmes limitations en translation ( $r_{\max}$ ) et rotation ( $\phi_{\max}$ ) que le mouvement actif.

La stratégie de mouvement actif établie induit une translation du repère binaural et une rotation de la fovéa en direction de la position estimée de la source. De même, nous constatons que le capteur binaural est « attiré » par la source. Ce comportement est conforme aux analyses effectuées dans les parties 3.2 et 3.3.

Rappelons que comme la loi  $p(x_k|y_{1:k})$  est assimilée à une gaussienne, l'entropie  $H(\mathbf{x}_k|y_{1:k})$  est une fonction du déterminant de la matrice de covariance de cette loi et ne dépend pas de sa moyenne. De façon heuristique, cette entropie est donc liée à la « taille » (« l'aire ») de l'ellipse de confiance à 99% associée à  $p(x_k|y_{1:k})$ <sup>6</sup>. Ainsi, le fait que deux stratégies mènent à la même valeur d'entropie à un instant donné signifie que leurs ellipses de confiance respectives admettent des tailles similaires. Ceci peut par exemple être observé pour le mouvement circulaire et le mouvement actif à l'instant  $k = 16$  (figures 3.10d et 3.10c). Ensuite, nous constatons que l'entropie liée à la stratégie de mouvement actif reste inférieure à celle liée au mouvement de translation à tous les instants de la simulation. En revanche, entre les instants  $k = 12$  et  $k = 16$ , l'entropie obtenue pour le mouvement circulaire est la plus faible. Ceci ne contredit pas le fait qu'entre deux instants consécutifs la stratégie active établit la translation et la rotation du capteur qui mènent à la plus grande réduction d'entropie. En réalité, il s'agit d'un exemple intéressant qui montre qu'une séquence de  $N$  stratégies optimales à un pas ne constitue pas une stratégie optimale à  $N$  pas.

Dans les sections 3.2 et 3.3 nous avons mis en avant que l'augmentation de l'information était d'autant plus importante lorsque : (1) le capteur se rapproche de l'ellipse de confiance ; (2) l'axe de la fovéa est orientée dans la direction de l'ellipse de confiance ; (3) le centre de la tête est placé dans le prolongement du petit axe de l'ellipse. Ceci nous permet de comprendre l'augmentation de l'entropie à la fin des stratégies passives (II) et (III). Dans la situation décrite dans la figure 3.10d (où l'axe interaural en bleu est proche de l'ellipse de confiance), si le capteur continue dans une trajectoire circulaire tangente à l'axe  $\vec{z}_R$  sans bruit de dynamique, alors seulement une faible quantité d'information est apportée par l'observation. Dans la simulation, l'entropie augmente car la perte d'information due au bruit de dynamique ne peut pas être compensée par le peu d'information véhiculée par l'observation<sup>7</sup>. Un raisonnement similaire peut expliquer la faible augmentation de l'entropie dans la stratégie (II). Dans la configuration de la figure 3.10e, bien que l'axe de la fovéa en rouge soit proche de l'ellipse de confiance, le capteur binaural

6. On montre [Boyd et al. 1994] que le volume d'un ellipsoïde  $\mathcal{E} = \{x \in \mathbb{R}^{n_x}, (x - \bar{x})^T P^{-1} (x - \bar{x}) \leq \gamma^2\}$ , où le vecteur  $\bar{x}$ , la matrice définie positive  $P$  et le scalaire réel  $\gamma$  sont donnés, s'écrit  $\text{vol}(\mathcal{E}) = C(n_x) \sqrt{\det(\gamma^2 P)}$ , où  $C(n_x)$  désigne une constante ne dépendant que de la dimension  $n_x$  de  $x$ .

7. En simulation, un bruit de dynamique de covariance non nulle est pris en compte dans la stratégie de localisation audio-motrice. Le bruit de dynamique n'est négligé que pour guider l'exploration.

s'éloigne de l'ellipse de confiance, ce qui réduit la quantité d'information apportée par l'observation. En conséquence, l'entropie pour cette stratégie reste plus ou moins constante à partir de  $k = 16$  et nous observons même une augmentation à la toute fin de l'expérience lorsque le capteur s'éloigne significativement de l'ellipse. Au contraire pour la stratégie active (I), dans la situation de la figure 3.10c où l'axe de la fovéa en rouge est également proche de l'ellipse de confiance, le capteur binaural se rapproche de l'ellipse de confiance, ce qui conduit à une réduction de l'entropie, même lorsqu'un bruit affecte la dynamique.

### 3.6.2 Expériences en conditions réelles

Pour cette expérience, nous ne considérons aucune information *a priori* sur la position de la source, à l'instant initial du processus de localisation. La pdf *a posteriori*  $p(x_k|y_{1:k})$  définie comme un mélange de gaussiennes, est calculée récursivement dans la stratégie de filtrage MG-UKF. Afin de définir la loi gaussienne  $p(x_k|z_{1:k})$  qui guide l'exploration, la solution adoptée consiste à retenir l'hypothèse de poids la plus forte du mélange de gaussiennes décrivant la loi  $p(x_k|y_{1:k})$ . Une autre solution peut consister à utiliser l'approximation « moment-matched » de cette loi. Pour éviter que le robot ne se rapproche trop rapidement de la source, le coefficient  $r_{\max}$  définissant la contrainte  $\mathcal{T}$  est calculé à chaque instant en fonction de la distance estimée à la source.

Dans les figures 3.12, 3.13 et 3.14, les ellipses de confiance « à 99% de probabilité » de chaque hypothèse gaussienne de la loi  $p(x_k|y_{1:k})$  sont dessinées. Le repère du capteur binaural est représenté en bleu et les axes  $\vec{y}_R$  et  $\vec{z}_R$  sont respectivement prolongés par des lignes pointillées bleues et rouges. Enfin, la position de la source dans le repère monde (le repère  $\mathcal{F}_{k=0}$ ) est désignée par une marque rouge. À l'instant initial, le repère  $\mathcal{F}_{k=0}$  lié au capteur binaural est au centre de la figure avec l'axe  $\vec{z}_R$  orienté vers le haut.

Comme cela est prévisible, le mouvement de translation sans rotation ne permet pas de lever l'ambiguïté avant-arrière (figure 3.12). Le mouvement circulaire (figure 3.13) ainsi que la stratégie active (figure 3.14) permettent à la fois d'apporter une information sur la distance et de lever l'ambiguïté avant-arrière. Notons que comme en simulation, le capteur binaural dirigé par la stratégie active, s'oriente en direction de l'estimation de la source et s'en rapproche.

L'entropie de l'approximation « moment-matched » de la loi  $p(x_k|y_{1:k})$  a été évaluée à tous les instants de simulation des différentes stratégies (figure 3.11). Notons que lorsque les hypothèses du mélange de gaussiennes sont regroupées en un seul « groupe » (voir les figures 3.13 et 3.14 itérations {10, 20, 29}), cette entropie est très proche de  $H(\mathbf{x}_k|y_{1:k})$ . Au contraire, lorsque le mélange admet plusieurs groupes d'hypothèses (lorsque l'ambiguïté avant-arrière n'est pas levée, voir figure 3.12) l'entropie de l'approximation « moment-matched » est alors assez éloignée de l'entropie  $H(\mathbf{x}_k|y_{1:k})$ . Les résultats 3.11 présentent des similitudes avec les observations effectuées en simulation. La stratégie d'exploration développée conduit à une réduction d'entropie plus importante en moyenne que les autres stratégies. Notons cependant

---

que la courbe d'entropie correspondant à la stratégie circulaire passe ponctuellement sous la courbe de la stratégie active. Cela met en avant, comme nous l'avons évoqué en 3.6.1, le fait que l'application de  $N$  pas d'une stratégie optimale à 1 pas ne constitue pas une stratégie optimale à  $N$  pas.

Enfin, notons également que les expériences ont été menées dans un environnement ouvert et donc variable. Des bruits non stationnaires peuvent apparaître d'une expérience à l'autre. Par ailleurs, la situation initiale capteur-source peut également différer quelque peu. En conséquence, la comparaison entre les courbes d'entropie ne peut être effectuée que qualitativement.

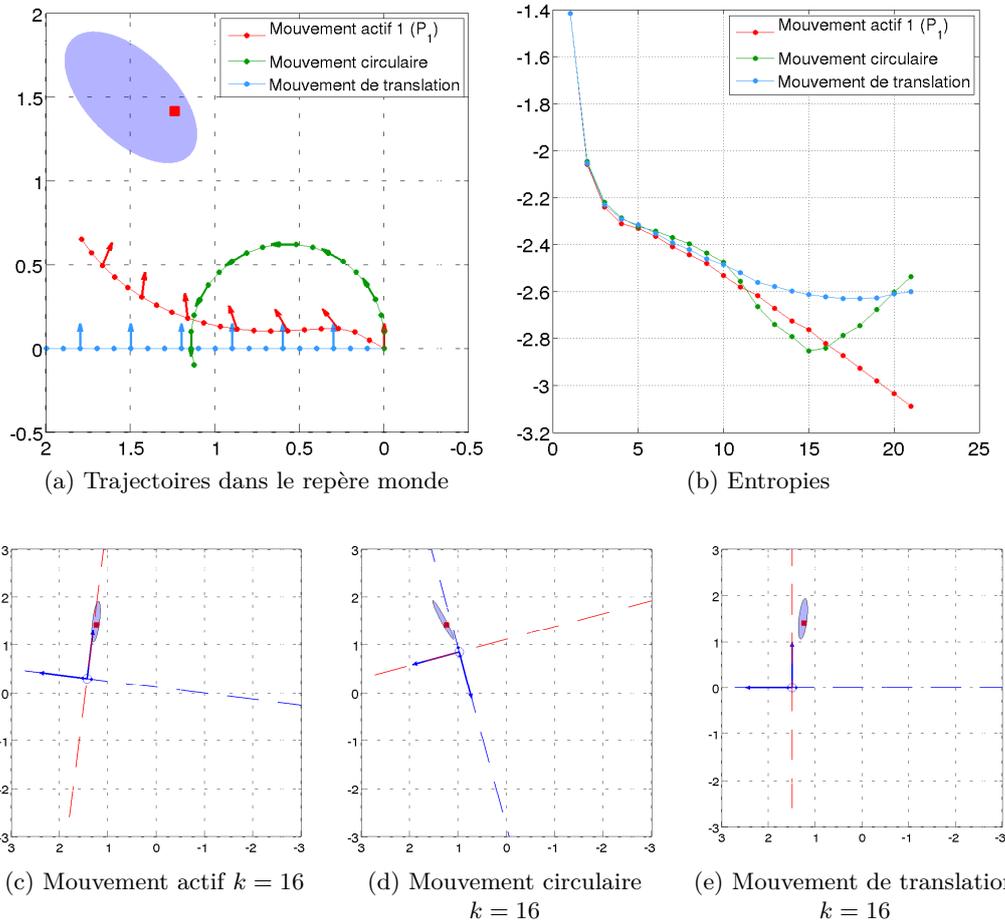


FIGURE 3.10 – Résultats de simulation – Localisation d’une source pour différents mouvements du capteur binaural : stratégie active, mouvement circulaire et mouvement de translation. (a) L’ellipse de confiance associée à l’estimation initiale de la position de la source est représentée. (b) Entropie (axe des ordonnées) pour chaque stratégie au cours du temps (abscisses). (c-d-e) Captures de l’état de la localisation à l’instant  $k = 16$ . L’ellipse de confiance est représentée en bleu. Le repère du capteur binaural est représenté en bleu et les axes  $\vec{y}_R$  et  $\vec{z}_R$  sont respectivement prolongés par des lignes pointillées bleues et rouges. Enfin, la position de la source dans le repère monde (le repère  $\mathcal{F}_{k=0}$ ) est désignée par une marque rouge.

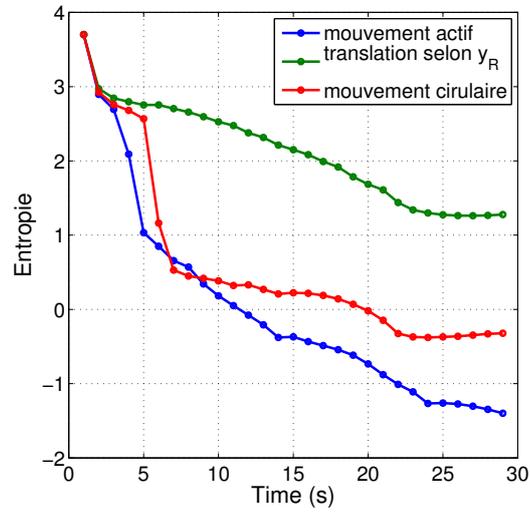


FIGURE 3.11 – Entropie de l’approximation moment-matched de la densité de filtrage en conditions réelles.

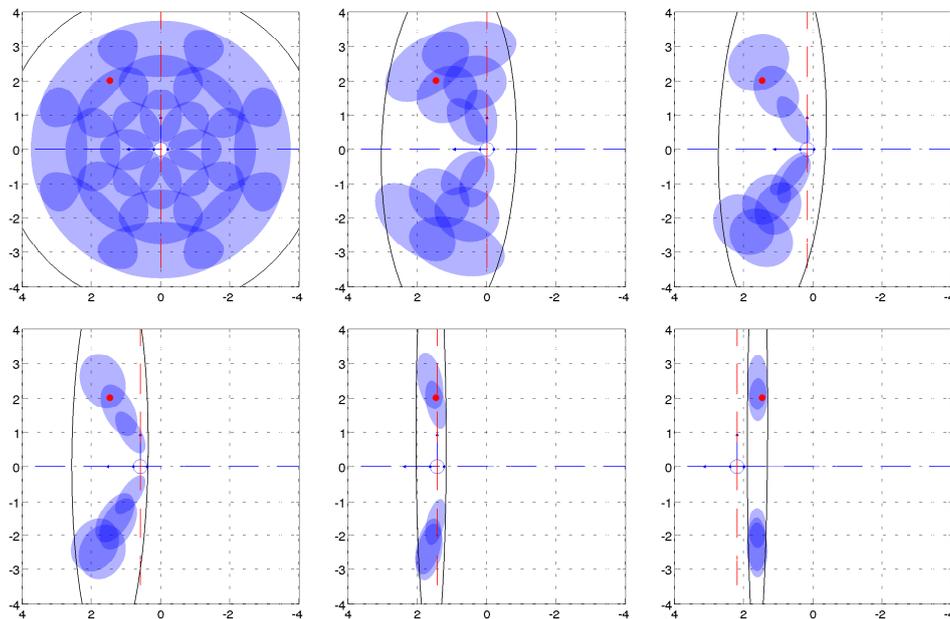


FIGURE 3.12 – Expérience en conditions réelles pour un déplacement en translation selon l’axe  $\vec{y}_R$  – Résultats de la localisation d’une source à partir du robot binaural dans le repère monde aux itérations  $\{1,2,5,10,20,29\}$  (sens de lecture de gauche à droite et de haut en bas).

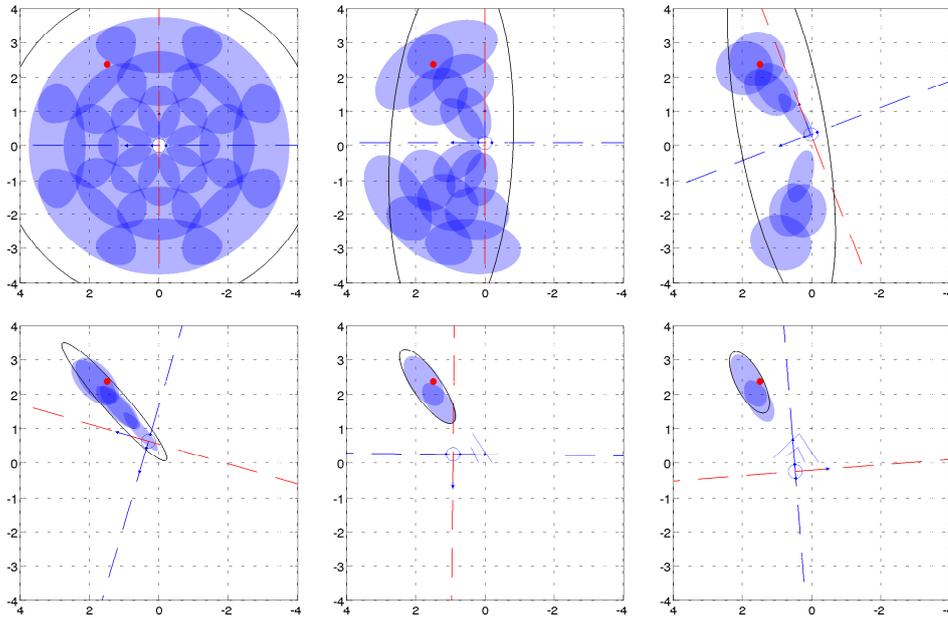


FIGURE 3.13 – Expérience en conditions réelles pour un mouvement circulaire de la tête binaurale – Résultats de la localisation d’une source à partir du robot binaural dans le repère monde aux itérations  $\{1,2,5,10,20,29\}$  (sens de lecture de gauche à droite et de haut en bas).

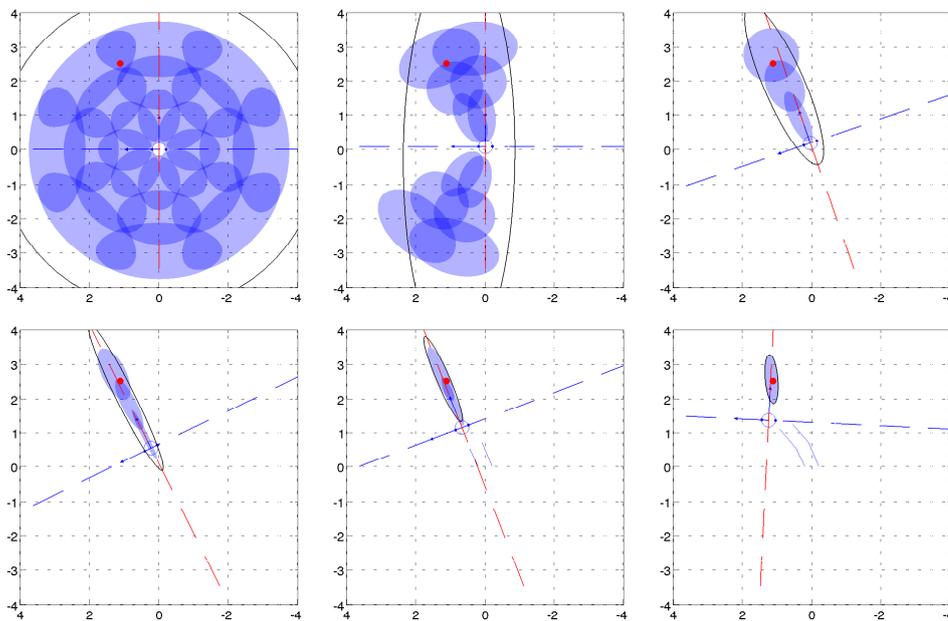


FIGURE 3.14 – Expérience en conditions réelles avec la stratégie de mouvement actif  $\mathcal{P}_1$  – Résultats de la localisation d’une source à partir du robot binaural dans le repère monde aux itérations  $\{1,2,5,10,20,29\}$  (sens de lecture de gauche à droite et de haut en bas).

# Génération de mouvement avec critère de décision à long terme

---

## Sommaire

---

<b>4.1</b>	<b>Position du problème</b>	<b>85</b>
4.1.1	Évaluation de $J_N$	86
4.1.2	Définition du gradient de $J_N$	87
<b>4.2</b>	<b>Approches numériques</b>	<b>89</b>
4.2.1	Différentiation automatique	89
4.2.2	Résolution de $\mathcal{P}_N$ par la méthode du gradient projeté	93
<b>4.3</b>	<b>Analyse de résultats de simulation</b>	<b>93</b>
4.3.1	Comparaison entre différents horizons d'exploration	95
4.3.2	Évaluations statistiques	96

---

Nous avons montré dans le chapitre précédent que la résolution du problème  $\mathcal{P}_1$ , à chaque pas du processus de localisation, permet une réduction significative de l'entropie de la densité de filtrage de la position relative de la source par rapport au capteur. Les résultats ont montré que cette stratégie court terme permet, à la fin de l'horizon temporel étudié, une plus grande réduction d'entropie que des stratégies passives sur ce même horizon. De plus, le vecteur de commande peut être calculé très rapidement, ce qui permet d'exploiter cette stratégie en temps réel. Cependant, l'approche court terme ne permet pas d'établir une politique de déplacement optimale sur le long terme. Dans ce chapitre nous traitons donc de la génération de mouvement avec un critère de décision défini sur un horizon fuyant de taille plus importante. Dans une première partie (4.1), un nouveau problème d'optimisation  $\mathcal{P}_N$  est défini. Nous proposons une méthode pour l'évaluation de la fonction d'incertitude  $J_N$  qui dépend des futures observations du système. Ensuite, le calcul du gradient de  $J_N$  et la résolution de  $\mathcal{P}_N$  sont proposés par des approches numériques en 4.2. Les résultats sont ensuite évalués en simulation (4.3).

## 4.1 Position du problème

Pour tout instant  $k$ , nous avons défini en 2.3.2 une fonction d'incertitude  $J_N$  comme l'espérance sur les futures observations  $z_{k+1:k+N}$  conditionnellement aux observations passées  $z_{1:k}$ , de  $H(\mathbf{x}_{k+N}|z_{1:k+N})$  (entropie de la loi de filtrage

$p(x_{k+N}|z_{k+N})$  à l'instant  $k + N$ , *i.e.*,

$$J_N = \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})]. \quad (4.1)$$

Cette fonction dépend de la séquence des vecteurs de commande  $\bar{u}_N = (u_k, \dots, u_{k+N-1})^T$  définie sur  $(\mathcal{T} \times \mathcal{R})^N$ . La recherche de la séquence optimale  $\bar{u}_N^*$  peut donc être définie par la résolution du problème  $\mathcal{P}_N$  défini dans l'énoncé (4.1).

**Énoncé 4.1** (Problème  $\mathcal{P}_N$ ). *Soit  $p(x_k|z_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k})$  la loi, gaussienne, du vecteur d'état  $\mathbf{x}_k$  à un instant  $k$  du processus de localisation, conditionnellement au fait que le processus d'observation  $\mathbf{z}_{1:k}$  se réalise en la séquence de mesures  $z_{1:k}$ . Soit*

$$J_N(\bar{u}_N) = K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})], \quad (4.2)$$

la fonction de coût définie dans le théorème (2.3), fonction de la séquence  $\bar{u}_N$  des vecteurs de commande. La séquence des vecteurs de commande  $\bar{u}_N^* = u_k^*, \dots, u_{k+N-1}^*$  qui minimise  $J_N$  sous les contraintes  $u_i \in \mathcal{T} \times \mathcal{R}$  avec  $i = k, k+1, \dots, k+N-1$ , est solution du problème

$$(\mathcal{P}_N) \begin{cases} \bar{u}_N^* &= \arg \min_{\bar{u}_N \in (\mathcal{T} \times \mathcal{R})^N} J_N(\bar{u}_N) \\ &= \arg \max_{\bar{u}_N \in (\mathcal{T} \times \mathcal{R})^N} F_1(\bar{u}_1) + \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})], \end{cases} \quad (4.3)$$

où chaque  $F_i$  est la fonction, définie dans l'algorithme 2.2, de la séquence des vecteurs de commande  $\bar{u}_i = u_k, \dots, u_{k+i-1}$  et des futures observations  $z_{k+1:k+i-1}$ .

La résolution de  $\mathcal{P}_N$  à tout instant  $k$  permet donc d'obtenir la séquence des vecteurs de commandes  $\bar{u}_N^*$  dont l'application conduit à la minimisation de l'entropie de la loi *a posteriori*  $p(x_{k+N}|z_{1:k+N})$  à l'instant  $k + N$ , moyennée sur les futures observations.

#### 4.1.1 Évaluation de $J_N$

Dans l'expression de  $J_N$  définie en (4.2), l'espérance de la fonction  $F_i$

$$\mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})] = \int p(z_{k+1:k+i-1}|z_{1:k}) F_i(\bar{u}_i, z_{k+1:k+i-1}) dz_{k+1:k+i-1}, \quad (4.4)$$

pour  $i \geq 2$ , ne peut pas être calculée formellement en l'absence de forme analytique de la fonction  $F_i$ . Une première approche consiste à admettre que les observations futures, inconnues, sont égales à leurs prédictions sur la base des observations passées et présentes, connues comme cela a été proposé dans [Deutsch *et al.* 2004, Leung

*et al.* 2006]. Nous préférons un calcul approché de (4.4) au moyen de la transformée unscented, dans la veine de l'approche proposée dans [Särkkä *et al.* 2014].

**Theorème 4.1** (Approximation d'une intégrale gaussienne par transformée unscented). *Soit  $x$  une réalisation du vecteur aléatoire gaussien  $\mathbf{x}$  de taille  $M$ , de pdf  $p(x) = \mathcal{N}(x; \hat{x}, P)$  et soit  $f$  une fonction de  $x$ . L'intégrale  $\int p(x)f(x)dx$  peut être approchée sur la base de l'évaluation de  $f$  en les  $2M + 1$  sigma-points  $\{X_i\}$ , par la somme pondérée*

$$\int p(x)f(x)dx \approx \sum_{i=0}^{2M} w_i f(X_i). \quad (4.5)$$

L'établissement des poids  $\{w_i\}$  de la transformée unscented et le calcul des sigma-points  $\{X_i\}$ , sont détaillés en Annexe A.1.  $\square$

Pour  $i \geq 2$ , la pdf  $p(z_{k+1:k+i-1}|z_{1:k})$  du vecteur aléatoire  $\mathbf{z}_{k+1:k+i-1}$  (de taille  $i-1$ ) conditionnellement aux observations passées  $z_{1:k}$ , peut être approchée par une distribution gaussienne

$$p(z_{k+1:k+i-1}|z_{1:k}) \approx \mathcal{N}(z_{k+1:k+i-1}; \hat{z}_{k+1:k+i-1|k}, C_{k+1:k+i-1|k}), \quad (4.6)$$

de moyenne  $\hat{z}_{k+1:k+i-1|k}$  et de covariance  $C_{k+1:k+i-1|k}$ . Les moments de cette distribution peuvent être calculés au moyen de la transformée unscented dans la procédure détaillée dans l'Algorithme (4.1). Ainsi en appliquant le théorème (4.1), l'intégrale (4.4) peut être approchée par

$$\int F_i(\bar{u}_i, z_{k+1:k+i-1})p(z_{k+1:k+i-1}|z_{1:k})dz_{k+1:k+i-1} \approx \sum_{j=0}^{2(i-1)} w_j F_i(\bar{u}_i, Z_j), \quad (4.7)$$

avec  $\{w_j\}$  les poids de la transformée unscented et  $2(i-1) + 1$  le nombre de sigma-points  $\{Z_j\}$  ( $Z_j \in \mathbb{R}^{(i-1)}$ ), calculés à partir de  $\hat{z}_{k+1:k+i-1|k}$  et  $C$ . La fonction d'information  $J_N$  (4.2) peut donc être réécrite en

$$J_N \approx K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \sum_{j=0}^{2(i-1)} w_j F_i(\bar{u}_i, Z_j). \quad (4.8)$$

Finalement, le problème  $\mathcal{P}_N$  (4.3) peut être redéfini avec la nouvelle expression de  $J_N$  par

$$(\mathcal{P}_N) \left\{ \begin{array}{l} \bar{u}_N^* = \arg \max_{\bar{u}_N \in (\mathcal{T} \times \mathcal{R})^N} F_1(\bar{u}_1) + \sum_{i=2}^N \sum_{j=0}^{2(i-1)} w_j F_i(\bar{u}_i, Z_j). \end{array} \right. \quad (4.9)$$

#### 4.1.2 Définition du gradient de $J_N$

Nous souhaitons établir le gradient de  $J_N$  de façon à étendre la méthode de résolution établie dans le chapitre 3 pour  $N = 1$ , à un horizon de  $N \geq 2$  pas dans

---

**Algorithme 4.1 :** Calcul des moments de  $p(z_{k+1:k+i-1}|z_{1:k})$

$$[\hat{z}_{k+1:k+i-1|k}, S_{k+1:k+i-1|k}] = Z\_Pred(\hat{x}_{k|k}, P_{k|k}, \bar{u}_i)$$


---

**Entrées :**

- $\hat{x}_{k|k}, P_{k|k}$
- Séquence des vecteurs de commande :  $\bar{u}_i = u_k^T, \dots, u_{k+i-1}^T$  avec  $u_k = (T_{y,k}, T_{z,k}, \phi_k)^T$

**Paramètres :**

- Nombre de vecteurs de commande :  $i \geq 2$
- Paramètre des *sigma-points* :  $\gamma$

**Sorties :**

- $\hat{z}_{k+1:k+i-1|k}, S_{k+1:k+i-1|k}$
- 

**Initialisation**

Concaténation des translations et rotations dans les vecteurs  $\bar{T}$  et  $\bar{\phi}$

- 1  $\bar{T} = (T_1^T, \dots, T_{i-1}^T)^T = ((T_{y,k}, T_{z,k})^T, \dots, (T_{y,k+i-1}, T_{z,k+i-1})^T)^T$
  - 2  $\bar{\phi} = (\phi_1, \dots, \phi_{i-1})^T = (\phi_k, \dots, \phi_{k+i-1})^T$
- 

Définition du vecteur  $\bar{x}$  regroupant les prédictions aux instants

$k+1, \dots, k+i-1$  de sorte que  $\bar{x} = (\hat{x}_{k+1|k}^T, \dots, \hat{x}_{k+i-1|k}^T)^T$

- 3  $\bar{x} = \bar{R}(\bar{\phi})\hat{x}_{k|k} - B\bar{T}$

avec  $\bar{R} = \begin{pmatrix} R^T(\phi_1) \\ \dots \\ R^T(\phi_1 + \dots + \phi_{i-1}) \end{pmatrix}$  de taille  $[2(i-1) \times 2]$  et

$$B = \begin{pmatrix} R^T(\phi_1) & \emptyset & \dots & \emptyset \\ R^T(\phi_1 + \phi_2) & R^T(\phi_2) & \dots & \emptyset \\ \vdots & \vdots & \ddots & \vdots \\ R^T(\phi_1 + \dots + \phi_{i-1}) & R^T(\phi_2 + \dots + \phi_{i-1}) & \dots & R^T(\phi_{i-1}) \end{pmatrix}$$

de taille  $[2(i-1) \times 2(i-1)]$

Calcul de  $\bar{P}$

- 4  $\bar{P} = \bar{R}P_{k|k}\bar{R}^T$

Décomposition de Cholesky de  $\bar{P}$

- 5  $\bar{P} = \bar{L}\bar{L}^T$  (Cette décomposition peut être facilement obtenue à partir de  $P_{k|k} = LL^T$ )
- 

Calcul des *sigma-points* de la loi  $p(x_{k+1:k+i-1}|z_{1:k})$

- 6  $[\{\bar{X}_j\}_{j=0, \dots, 2n_{\bar{x}}}, \{w_j^m\}, \{w_j^c\}] = \text{sp}(\bar{x}, \bar{L}, \gamma)$  avec  $n_{\bar{x}} = 2(i-1)$  la taille du vecteur  $\bar{x}$

Passage des *sigma-points* au travers de la fonction d'observation

(pour alléger la notation, la fonction  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$  définissant le modèle d'observation (1.15) page 21 est « étendue » à un argument  $\bar{x}$  de taille  $2(i-1)$  au sens où  $h(\bar{x})$  est constitué de l'empilement de l'application de  $h$  à chacun des sous-vecteurs successifs de dimension 2 qui constituent  $\bar{x}$ )

- 7 **Pour**  $j = 0, \dots, 2n_{\bar{x}}$  **faire**

- 8 |  $\bar{Z}_j = h(\bar{X}_j)$

- 9 **fin**
- 

Calcul des deux premiers moments de  $p(z_{k+1:k+i-1}|z_{1:k})$  avec (A.5) et (A.6)

- 10  $\hat{z}_{k+1:k+i-1|k} = \text{spMoy}(\{w_j^m\}, \{\bar{Z}_j\})$

- 11  $\hat{S}_{k+1:k+i-1|k} = \text{spCorr}(\{w_j^c\}, \{\bar{Z}_j\}, \hat{z}_{k+1:k+i-1|k}, \{\bar{Z}_j\}, \hat{z}_{k+1:k+i-1|k})$
-

le futur. Calculer le gradient de la fonction de récompense  $J_N$  par rapport à la séquence de vecteurs de commande  $\bar{u}_N$  est crucial pour caractériser les variations de l'information avec le mouvement et pour résoudre le problème d'optimisation  $\mathcal{P}_N$ . Avec la notation

$$\nabla_{\bar{u}_N} = (\nabla_{u_k}^T, \nabla_{u_{k+1}}^T, \dots, \nabla_{u_{k+N-1}}^T)^T, \quad (4.10)$$

le gradient de la fonction (4.8) s'écrit

$$\nabla_{\bar{u}_N} J_N = -\nabla_{\bar{u}_N} F_1(\bar{u}_1) - \sum_{i=2}^N \sum_{j=0}^{2(i-1)} w_j \nabla_{\bar{u}_N} F_i(\bar{u}_i, Z_j). \quad (4.11)$$

Notons que  $F_i(\bar{u}_i, Z_j)$  ne dépend pas des séquences de commandes postérieures à l'instant  $k+i-1$ , donc  $\nabla_{u_{k+i}} F_i = \dots = \nabla_{u_{k+N-1}} F_i = 0$  et  $\nabla_{\bar{u}_N} F_i$  ne dépend que de  $\{\nabla_{\bar{u}_l} F_l\}_{l \leq i}$ . Par la suite, nous concentrons notre étude sur le calcul de  $\nabla_{\bar{u}_i} F_i$  pour  $i = 2, \dots, N$ . La fonction  $F_i$  ne pouvant pas être écrite simplement sous la forme d'une équation, son gradient ne peut pas être évalué de façon formelle. Nous proposons dans la section suivante une approche numérique permettant d'évaluer ce gradient en toute séquence de vecteurs de commande.

## 4.2 Approches numériques

### 4.2.1 Différentiation automatique

Comme nous l'avons évoqué dans la partie 3.4, calculer le gradient de la fonction d'incertitude au moyen d'une méthode de type différences finies n'est pas souhaitable (variations du problème d'optimisation en fonction des instants du processus de localisation et difficulté d'établir le pas de dérivation conduisant à une bonne précision). Afin d'automatiser la procédure de dérivation de  $F_i$  pour tout  $i \geq 2$ , la méthode proposée dans la partie 3.4.2 est remplacée par une approche par différentiation automatique, dont nous présentons le principe dans cette partie.<sup>1</sup>

La différentiation automatique est une méthode numérique qui permet de transformer un programme calculant la valeur d'une fonction de façon qu'il calcule conjointement la valeur de la dérivée de cette fonction [Bartholomew-Biggs *et al.* 2000]. La différentiation automatique exploite le fait que tout programme informatique est composé de séquences d'opérations élémentaires  $+$ ,  $-$ ,  $\times$ ,  $/$ , ainsi qu'éventuellement de fonctions mathématiques connues (fonctions trigonométriques, logarithme, puissance, etc.). Ainsi, toute valeur intermédiaire  $v$  du programme peut être vue comme une fonction du vecteur d'entrée  $x$ , et écrite comme la composition de plusieurs fonctions simples ( $v = f_n \circ \dots \circ f_2 \circ f_1(x)$ ). La différentiation automatique calcule la dérivée de ces fonctions composées, par rapport au vecteur d'entrée  $x$ , en exploitant le théorème de dérivation des fonctions composées

1. Les deux procédures sont équivalentes pour le calcul de  $F_1$ , mais nous préférons dans ce cas l'utilisation de la méthode détaillée en 3.4, moins coûteuse en temps de calcul.

(*chain rule*). Par exemple, si un programme est composé de deux fonctions élémentaires  $f_1$  et  $f_2$  de telle sorte que la sortie soit  $v_2 = (f_2 \circ f_1)(x)$ , avec  $v_1 = f_1(x)$  une valeur intermédiaire (et  $v_2 = f_2(v_1)$ ), alors ce programme doit être transformé de façon à produire également

$$\frac{dv_2}{dx} = \frac{dv_2}{dv_1} \cdot \frac{dv_1}{dx}. \quad (4.12)$$

Les méthodes de différentiation automatique sont divisées en deux familles : les méthodes directes (*forward accumulation*) et les méthodes inverses (*reverse accumulation*). L'approche directe consiste à fixer la variable de départ pour laquelle la dérivation doit être effectuée, puis à évaluer le programme pour obtenir la dérivée de chaque fonction intermédiaire par rapport à cette variable. La méthode directe est relativement facile à implémenter car le programme d'origine est très peu modifié et le calcul des dérivées est effectué dans le sens d'exécution du programme (l'équation (4.12) est calculée de la droite vers la gauche). L'inconvénient de cette approche est que le programme doit être exécuté  $n$  fois (avec  $n$  la taille de  $x$ ) pour obtenir le gradient de la fonction (cf. l'exemple présenté dans l'algorithme 4.2). Avec les méthodes inverses, que nous n'aborderons pas ici, le programme d'origine doit être exécuté préalablement en stockant toutes les valeurs intermédiaires, puis le gradient est calculé en une fois dans le sens inverse du calcul précédent (l'équation (4.12) est calculée de la gauche vers la droite). Cette méthode nécessite cependant des modifications assez profondes du programme initial. Notre étude porte uniquement sur le mode direct, plus simple à implanter et plus flexible (une modification du programme initial peut facilement être prise en compte dans l'algorithme de différentiation automatique).

Un autre avantage de la différentiation automatique réside dans le fait que des outils permettant suivant les cas, un calcul via le mode direct et/ou inverse, ont été développés pour différents langages de programmation tels que le Fortran (le projet *GRESS* [Horwedel 1991] par exemple), le C++ (la liste des projets est assez longue, nous pouvons citer par exemple *CasADI* [Andersson et al. 2012], ou *OpenAD*) ou encore MATLAB (le projet *ADiMat* [Bischof et al. 2002] par exemple).

La différentiation automatique directe peut s'appuyer sur l'utilisation des nombres duaux. Un nombre dual  $z_d$ <sup>2</sup> est défini par deux éléments réels  $(z_a, z_b) \in \mathbb{R}^2$  par

$$z_d = z_a + \varepsilon z_b, \quad (4.13)$$

avec  $\varepsilon$  un élément nilpotent tel que  $\varepsilon^2 = 0$ . La multiplication de deux nombres duaux  $z_{1d} = z_{1a} + \varepsilon z_{1b}$  et  $z_{2d} = z_{2a} + \varepsilon z_{2b}$  conduit au nombre dual  $z_{3d} = z_{3a} + \varepsilon z_{3b}$

---

2. Les nombres duaux sont indicés de la lettre  $d$ .

qui vérifie

$$\begin{aligned} z_{3d} &= (z_{1a} + \varepsilon z_{1b})(z_{2a} + \varepsilon z_{2b}) \\ &= z_{1a}z_{2a} + \varepsilon(z_{1a}z_{2b} + z_{1b}z_{2a}) \\ &= z_{3a} + \varepsilon z_{3b}. \end{aligned} \tag{4.14}$$

Ainsi, les nombres duaux permettent de propager facilement, au sein d'un programme, les valeurs de dérivations au premier ordre.

Illustrons le principe de la différentiation automatique directe au travers d'un exemple. Soit  $f$  une fonction de  $x = (x_1, x_2)^T \in \mathbb{R}^2$ , à valeurs dans  $\mathbb{R}$ , définie par

$$\begin{aligned} f : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto x_1 \cos(x_2) + x_1^2. \end{aligned} \tag{4.15}$$

Supposons que la valeur de  $f$  pour  $x = \bar{x}$  puisse être évaluée par un programme appelé `f_eval` de sorte que

$$[f] = \text{f\_eval}(\bar{x}). \tag{4.16}$$

La différentiation automatique consiste donc à modifier le programme `f_eval` de façon à calculer les valeurs numériques du gradient  $\nabla f(\bar{x}_1, \bar{x}_2) = [\partial f / \partial x_1, \partial f / \partial x_2]_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)}^T$  de  $f$ , pour des valeurs particulières  $(\bar{x}_1, \bar{x}_2)^T$  de  $x = (x_1, x_2)^T$ . Ainsi un nouveau programme `f_eval_AD` est établi de sorte que

$$[f, \nabla f] = \text{f\_eval\_AD}(\bar{x}). \tag{4.17}$$

Le gradient est évalué par différentiation automatique par la méthode décrite dans l'algorithme 4.2. Comme il s'agit d'une approche directe, la fonction  $f$  est évaluée deux fois.

Dans la pratique, deux implémentations sont possibles pour la différentiation automatique à mode direct :

- La méthode par surcharge d'opérateurs et de fonctions : les opérateurs  $+$ ,  $-$ ,  $\times$ ,  $/$  ainsi que toutes les fonctions élémentaires utilisées dans le programme (fonction trigonométriques, logarithme, puissance, etc.) doivent être surchargés au sens de la programmation orientée objet, de manière à pouvoir être appliquées sur des nombres duaux. Ainsi les opérations constituant le programme initial peuvent être exécutées sans modification. C'est la méthode la plus simple à mettre en place et sans doute la plus élégante. En effet, comme il n'y a pas de modification du code, la méthode est robuste à toute évolution du programme initial. Si de nouvelles fonctions doivent être introduites, il suffit d'ajouter leur définition à la liste des fonctions surchargées. Un inconvénient de cette méthode est le temps d'exécution qui peut devenir très important suivant le langage de programmation et la taille du programme.

---

**Algorithme 4.2 :** Exemple de programmation par différentiation automatique  $[f, \nabla f] = \text{f\_eval\_AD}(\bar{x}_1, \bar{x}_2)$  avec  $f(x_1, x_2) = x_1 \cos(x_2) + x_1^2$

---

**Entrées :**

— Valeurs de  $x$  :  $\bar{x}_1, \bar{x}_2$

**Paramètres :**

— Nombre de variables en entrée :  $n = 2$

**Sorties :**

— Évaluation de la fonction  $f$  pour  $\bar{x}_1, \bar{x}_2$  :  $f$

— Gradient de  $f$  par rapport à  $\bar{x}_1, \bar{x}_2$  :  $\nabla f$

---

**Pour**  $i = 1, \dots, n$  **faire**

1	Dérivation par rapport à la $i^{\text{ème}}$ variable $x_i$
2	$\frac{dx_i}{dx_i} = x'_i = 1$ ;
3	$\frac{dx_{j \neq i}}{dx_i} = x'_{j \neq i} = 0$ ;
	Initialisation des entrées sous forme duale
4	$x_{1d} = \bar{x}_1 + \varepsilon x'_1$ et $x_{2d} = \bar{x}_2 + \varepsilon x'_2$ ;
	Exécution du programme
5	$v_{1d} = \cos(x_{2d}) = \cos(\bar{x}_2) - \varepsilon x'_2 \sin(\bar{x}_2) = v_1 + \varepsilon v'_1$ ;
6	$v_{2d} = x_{1d} v_{1d} = \bar{x}_1 \times v_1 + \varepsilon(\bar{x}_1 v'_1 + \dot{x}_1 v_1) = v_2 + \varepsilon v'_2$ ;
7	$v_{3d} = x_{1d} x_{1d} = \bar{x}_1^2 + 2\varepsilon \bar{x}_1 x'_1 = v_3 + \varepsilon v'_3$ ;
8	$v_{4d} = v_{2d} + v_{3d} = v_4 + \varepsilon v'_4$ ;
	Sauvegarde de la dérivée partielle
9	$\frac{\partial f}{\partial x_i} = v'_4$ ;
10	$f = v_4$ ;
11	<b>fin</b>
12	$\nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)^T$ ;

---

- La méthode par transformation du code source : le code source du programme est modifié de façon à calculer explicitement les valeurs dérivées de chaque étape. Le code résultant est plus efficace (moins d'appels de fonctions) qu'avec la méthode par surcharge ce qui rend le programme plus rapide à l'exécution. Cependant cette méthode nécessite une bonne connaissance du programme initial et toute modification du code source initial doit être prise en compte dans un nouvel algorithme, ce qui la rend peu flexible.

#### 4.2.2 Résolution de $\mathcal{P}_N$ par la méthode du gradient projeté

Un programme de différentiation automatique a été développé afin de calculer le gradient  $\nabla_{\bar{u}_i} F_i(\bar{u}_i, Z_j)$  ( $\nabla_{\bar{u}_i}$  défini en (4.10)) avec  $\bar{u}_i$  une séquence des vecteurs de commande particulière, et  $Z_j$  le  $j^{\text{ème}}$  *sigma-point* déterminé à partir de la loi  $\mathcal{N}(z_{k+1:k+i-1}; \hat{z}_{k+1:k+i-1|k}, C)$ , dont le calcul est détaillé dans l'algorithme (4.1). L'évaluation de  $\nabla_{\bar{u}_N} J_N(\bar{u}_N)$ , défini dans l'équation (4.11), pour toute séquence  $\bar{u}_N$  peut alors être obtenue.

Un algorithme de gradient projeté est établi (par extension de l'algorithme (3.1)), de manière à résoudre le problème  $\mathcal{P}_N$  et à obtenir la séquence des vecteurs de commande  $\bar{u}_N^*$  qui minimise la fonction  $J_N$ . Si le problème  $\mathcal{P}_N$  est défini comme un problème de minimisation, l'algorithme du gradient projeté doit suivre la direction de descente  $d = -\nabla_{\bar{u}_N} J_N(\bar{u}_N)$ . Soit  $U \in (\mathbb{R}^3)^N$ , la superposition en un vecteur, de la séquence des vecteurs de commande  $\bar{u}_N = u_k, \dots, u_{k+N-1}$  de sorte que

$$U = (u_k^T, \dots, u_{k+N-1}^T)^T. \quad (4.18)$$

L'algorithme est initialisé à une valeur  $U^{(0)}$ , puis les valeurs  $U^{(i)} = (u_k^{(i)T}, \dots, u_{k+N-1}^{(i)T})$  sont calculées récursivement et les composantes  $u_j^{(i)}$  de  $U^{(i)}$  avec  $j = k, k+1, \dots, k+N-1$ , sont projetées sur l'ensemble  $\mathcal{T} \times \mathcal{R}$ . L'ensemble de la procédure est résumée dans l'algorithme (4.3).

### 4.3 Analyse de résultats de simulation

Nous avons évalué la stratégie d'exploration long terme en appliquant, comme dans le chapitre 3, le schéma complet de localisation présenté en figure 1.2. La configuration de la simulation est identique à celle de la partie 3.6.1. C'est à dire que nous considérons un état initial de la simulation dans lequel une connaissance *a priori*  $p(x_0)$  gaussienne est disponible, avec à cet instant la position estimée de la source dans le repère du robot  $\mathcal{F}_{k=0}$  caractérisée par  $\hat{x}_0$  et  $P_0$ . La vraie position a été tirée aléatoirement dans cet état de croyance initial. Notons que le temps d'exécution nécessaire à la résolution de  $\mathcal{P}_N$  augmente (tableau (4.1)) avec l'horizon d'exploration  $N^3$ . Pour cette raison, la taille  $N$  de l'horizon d'exploration en

---

3. Le calcul du gradient par l'algorithme de différentiation automatique développé sous MATLAB est très coûteux.

---

**Algorithme 4.3 :** Gradient projeté pour la résolution de  $\mathcal{P}_N$

---

**Entrées :**

- Fonction de coût :  $J_N$
- Translation maximale admissible du capteur binaural :  $r_{\max}$
- Rotation maximale admissible du capteur binaural :  $\phi_{\max}$
- Opérateur de projection sur l'ensemble  $\mathcal{T} \times \mathcal{R}$  :  $\pi_{\mathcal{T} \times \mathcal{R}}(\cdot)$

**Paramètres :**

- Moments de l'état de croyance initial à l'instant  $k$  :  $\hat{x}_{k|k}, P_{k|k}$
- Horizon de calcul :  $N$
- Variable d'itération et itération maximale :  $i, M$
- Seuil pour l'arrêt de l'algorithme et pas de la descente :  $\varepsilon, \alpha$

**Sorties :**

- Vecteur de commandes de taille  $3N$  :  $U^{(i)} = (u_k^{(i)T}, \dots, u_{k+N-1}^{(i)T})^T$
- 

**Initialisation**

- 1  $U^{(0)} = (u_k^{(0)T}, \dots, u_{k+N-1}^{(0)T})^T$   
 $= (T_{y,k}^{(0)}, T_{z,k}^{(0)}, \phi_k^{(0)}, \dots, T_{y,k+N-1}^{(0)}, T_{z,k+N-1}^{(0)}, \phi_{k+N-1}^{(0)})^T$ ;
  - 2  $U_{\text{prec}} = (\infty, \dots, \infty)^T$ ;
  - 3  $i = 0$ ;
- 

**Tant que  $i < M$  ET  $\|U^{(i)} - U_{\text{prec}}\|_2 \geq \varepsilon$  faire**

- 4 Évaluation de  $d_i = -\nabla J_N(U^{(i)})$ , où  $J_N$  est défini sur la base de  $\hat{x}_{k|k}, P_{k|k}$ ;
  - 5 Calcul de  $U^{(i+1)}$  avant projection :  $U^{(i+1)} = U^{(i)} + \alpha d_i$ ;
  - 6 **Pour  $j = k : k + N - 1$  faire // Boucle sur les éléments  $u_j^{(i+1)}$  de  $U^{(i+1)}$**
  - 7 Transformation polaire de  $u_j^{(i+1)}$  :  $v_j^{(i+1)} = (\rho_j^{(i+1)}, \theta_j^{(i+1)}, \phi_j^{(i+1)})^T$ ;
  - 8 **Projection** de  $u_j^{(i+1)}$  sur  $\mathcal{T} \times \mathcal{R}$  :  $u_j^{(i+1)} = \pi_{\mathcal{T} \times \mathcal{R}}(u_j^{(i+1)})$
  - 9     **Si**  $\rho_j^{(i+1)} > r_{\max}$  **alors** // Projection sur  $\mathcal{T}$
  - 10     |  $T_{y,j}^{(i+1)} = r_{\max} \sin(\theta_j^{(i+1)})$ ;
  - 11     |  $T_{z,j}^{(i+1)} = r_{\max} \cos(\theta_j^{(i+1)})$ ;
  - 12     **fin**
  - 13     **Si**  $|\phi_j^{(i+1)}| > \phi_{\max}$  **alors** // Projection sur  $\mathcal{R}$
  - 14     |  $\phi_j^{(i+1)} = \text{sign}(\phi_j^{(i+1)})\phi_{\max}$ ;
  - 15     **fin**
  - 16     **fin**
  - 17 **fin**
  - 18  $U_{\text{prec}} = U^{(i)}$ ;
  - 19  $U^{(i+1)} = (T_{y,k}^{(i+1)}, T_{z,k}^{(i+1)}, \phi_k^{(i+1)}, \dots, T_{y,k+N-1}^{(i+1)}, T_{z,k+N-1}^{(i+1)}, \phi_{k+N-1}^{(i+1)})^T$ ;
  - 20  $i = i + 1$ ;
  - 21 **fin**
-

$N$	Temps (s)
1	0,099
2	0,989
3	4,042
4	11,970
5	27,563
7	105,580
9	293,694

Tableau 4.1 – Temps d’exécution de l’algorithme de résolution de  $\mathcal{P}_N$  avec MATLAB pour différents horizons d’exploration  $N$ .

simulation doit être limitée. Dans la section 4.3.1 nous comparons les résultats de l’application de la stratégie d’exploration pour différents horizons  $N = \{1, 2, 3, 5\}$ . Ces résultats permettent de mettre en avant des différences de comportements du capteur en fonction de l’horizon choisi. Comme la stratégie d’exploration est calculée par rapport à un critère moyen sur les futures observations, et afin de pouvoir étudier l’impact de l’horizon d’exploration sur la réduction de l’entropie, nous effectuons une étude comparée des horizons d’exploration pour un grand nombre de simulations en section 4.3.2.

### 4.3.1 Comparaison entre différents horizons d’exploration

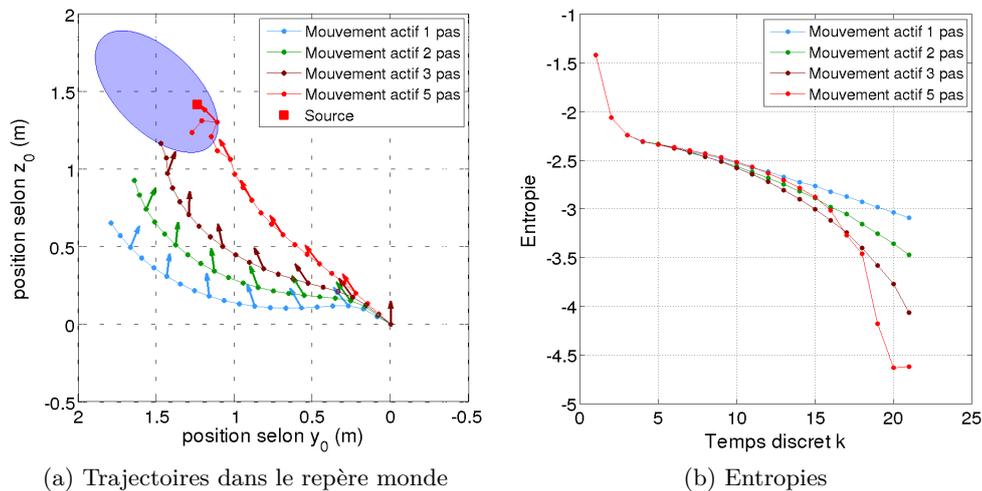


FIGURE 4.1 – Stratégies de mouvement actif avec différents horizons d’exploration  $N = \{1, 2, 3, 5\}$  – (a) L’ellipse de confiance associée à l’estimation initiale de la source est représentée en bleu. La position de la source dans le repère monde (le repère  $\mathcal{F}_{k=0}$ ) est désignée par une marque rouge. (b) Entropie (axe des ordonnées) pour chaque stratégie au cours du temps (abscisses).

La connaissance *a priori*  $p(x_0)$  est initialisée de sorte que  $\hat{x}_0 = (1.5, 1.5)^T$  et  $P_0 = \begin{pmatrix} 0.0121 & 0.0074 \\ 0.0074 & 0.0231 \end{pmatrix}$ . La stratégie d'exploration a été testée pour différents horizons (voir figure 4.1). Pour chaque scénario, le mouvement est calculé sur un horizon glissant : à chaque instant  $k$  de la simulation, la séquence des vecteurs de commande  $\bar{u}_N$  est calculée en résolvant  $\mathcal{P}_N$  avec  $N = \{1, 2, 3, 5\}$ ; puis le premier élément  $u_k$  de  $\bar{u}_N$  est appliqué; l'entropie  $H(\mathbf{x}_k|z_{1:k})$  de l'état de croyance est alors calculée et affichée pour chaque stratégie  $N = \{1, 2, 3, 5\}$ . Notons dans un premier temps que lorsque le capteur binaural se rapproche trop de la source (voir stratégie  $N = 5$  sur la figure 4.1b), la trajectoire du capteur binaural n'a plus beaucoup de sens (la stratégie d'exploration est guidée par un modèle d'observation champ lointain). Nous constatons que durant les premières secondes, les quatre stratégies conduisent à des valeurs d'entropie voisines bien que les positions relatives capteur-source soient différentes. Quand l'horizon de calcul augmente, le capteur se dirige vers la source de manière plus directe. La stratégie  $N = 5$  conduit le capteur binaural directement en direction de la source (4.1b) et offre les meilleurs résultats de localisation après  $k = 17$ . À l'instant final, l'entropie pour  $N = 5$  est la plus faible. Notons également qu'en fin de simulation, la différence en entropie entre les diverses stratégies sur la figure 4.1b est significative; l'entropie est d'autant plus faible que l'horizon  $N$  de la stratégie d'exploration est importante.

Nous remarquons donc que la stratégie ayant l'horizon le plus long ( $N = 5$ ) est la plus efficace sur le long terme mais les autres stratégies donnent de meilleurs résultats de localisation en début de simulation. La sélection de l'horizon d'exploration doit donc résulter d'un choix entre améliorer la localisation à court terme avec temps d'exécution faible ( $N$  faible) et améliorer la localisation à long terme au prix d'une augmentation importante du temps de calcul ( $N$  grand). Dans cet exemple, il semble que l'horizon  $N = 3$  constitue un choix judicieux au vu de ces considérations.

### 4.3.2 Évaluations statistiques

La connaissance *a priori*  $p(x_0)$  est initialisée de sorte que  $\hat{x}_0 = (1, 1)^T$  et  $P_0 = \begin{pmatrix} 0.0151 & 0.0117 \\ 0.0117 & 0.0151 \end{pmatrix}$ . Pour les horizons d'exploration  $N = \{1, 2, 7\}$ , l'entropie de la densité de filtrage  $p(x_k|z_{1:k})$  à chaque instant de localisation  $k = INIT, 1, \dots, 7$ , a été évaluée pour un jeu de 1000 expériences de Monte Carlo (figure 4.2). À l'instant  $k = INIT$  de chaque expérience, la position réelle de la source a été tirée aléatoirement dans la distribution initiale  $p(x_0)$ . À chaque instant  $k = INIT, 1, \dots, 6$ , les observations sont tirées aléatoirement selon le modèle d'observation (3.1), à partir de la vraie position relative de la source. Pour  $N = 7$ , le vecteur de commandes  $\bar{u}_{N=7}$  est calculé à l'instant initial et les commandes  $u_k, \dots, u_{k+6}$  qui le constituent sont appliquées à chaque instant de l'expérience. Pour  $N = \{1, 2\}$ , le vecteur de commandes  $\bar{u}_N$  est calculé à chaque instant  $k$  et seulement le premier élément est appliqué (stratégie à horizon glissant).

Pour chaque  $N$ , l'entropie résultant de l'application de la première commande  $u_0$  (commande permettant de passer de l'état initial  $k = INIT$  à  $k = 1$ ) est la même

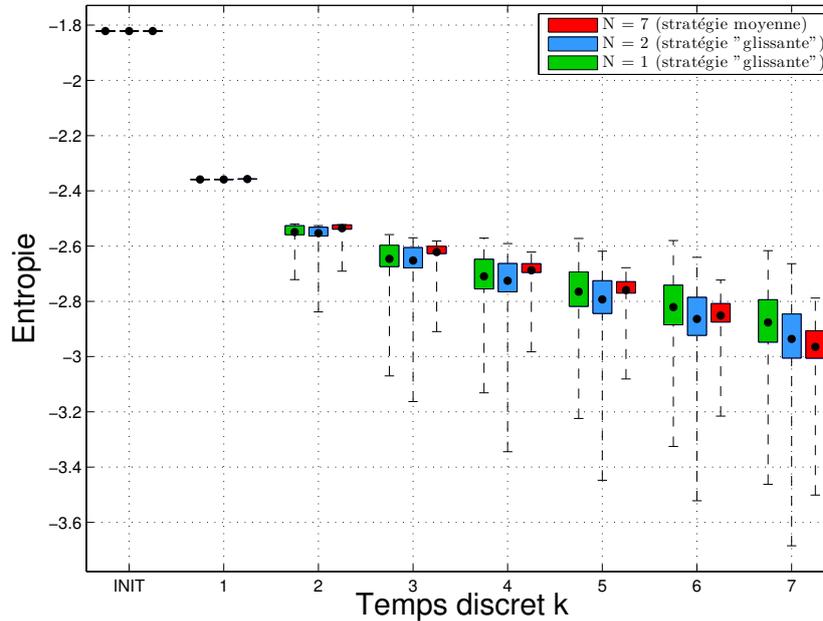


FIGURE 4.2 – Distribution de l’entropie de la densité de filtrage pour un grand nombre d’expériences – La localisation a été simulée 1000 fois pour chaque horizon d’exploration  $N = \{1, 2, 7\}$ . Les limites basses et hautes de chaque boîte représentent respectivement le premier et troisième quartiles des distributions de l’entropie ; soit 50% des données. Les extrémités des lignes discontinues constituent les valeurs minimales et maximales. Les moyennes empiriques sont représentées par des marques noires.

pour chaque expérience car elle ne dépend pas de l’assimilation de l’observation  $z_1$  (dans le filtrage de Kalman, le calcul de la covariance *a posteriori* ne dépend pas de l’assimilation de la vraie observation – voir algorithme A.2). En conséquence, il n’y a pas de dispersion de l’entropie à  $k = 1$ , celle-ci dépend néanmoins de la stratégie de mouvement actif. Bien que les entropies soient voisines en  $k = 1$  sur toutes les stratégies, les positions relatives capteur-source ne sont pas nécessairement identiques. La stratégie à court terme  $N = 1$  est très vite surpassée par le mouvement  $N = 2$ . Par ces simulations de Monte Carlo, nous retrouvons bien que la stratégie  $N = 7$  est optimale en moyenne, à l’issue de la simulation ; *i.e.*, elle conduit à l’entropie moyenne la plus basse.



# Prospectives pour la localisation binaurale à mouvement actif

---

## Sommaire

---

<b>5.1</b>	<b>Prospectives court terme</b> . . . . .	<b>99</b>
5.1.1	Comparaison exhaustive avec une simplification rencontrée dans la littérature . . . . .	99
5.1.2	Prise en compte du bruit de dynamique . . . . .	100
5.1.3	Entropie d'un mélange de gaussiennes . . . . .	101
<b>5.2</b>	<b>Évolution vers une stratégie multiobjectifs</b> . . . . .	<b>102</b>

---

Dans ce chapitre nous présentons des travaux plus prospectifs, permettant une évolution à court terme de notre approche (section 5.1) ou une extension consistant en l'étude d'un problème connexe (section 5.2).

## 5.1 Prospectives court terme

### 5.1.1 Comparaison exhaustive avec une simplification rencontrée dans la littérature

Dans le chapitre 2, nous avons défini la fonction d'incertitude  $J_N$  comme l'espérance sur les  $N$  futures observations conditionnellement aux observation passées, de l'entropie de la densité de filtrage à l'instant  $k + N$  de sorte que

$$J_N(\bar{u}_N) = \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})] \quad (5.1)$$

Dans la littérature, pour éviter le calcul de cette espérance, il est parfois considéré que les futures observations seront égales à leur prédiction, *i.e.*, le terme d'innovation du filtre de Kalman est nul. Par ailleurs, comme l'entropie de la loi *a posteriori* ne dépend que de la matrice de covariance de cette loi, le calcul de la fonction d'incertitude s'en trouve largement simplifié. Notons que dans le cas court terme ( $N = 1$ ), il n'y pas de différence avec notre approche puisque la matrice de covariance de la loi *a posteriori* est indépendante de la réalisation de la mesure (avec une approche Kalman uniquement). Cependant lorsque des instants plus lointains ( $N > 1$ ) sont considérés, la matrice de covariance  $P_{k+N|k+N}$  de la loi  $p(x_{k+N}|z_{1:k+N})$  dépend implicitement de la réalisation de toutes les observations précédant l'instant  $k + N$ .

Par l'application du théorème 2.3, sous les hypothèses définies en 2.3.1, nous avons établi que  $J_N$  pouvait s'écrire

$$J_N(\bar{u}_N) = K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \mathbb{E}_{z_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})], \quad (5.2)$$

avec  $K'_N$  un terme constant vis à vis de la séquence de vecteurs de commande  $\bar{u}_N$  et  $F_i$  une fonction de  $\bar{u}_i$  et des observations  $z_{k+1:k+i-1}$ . Dans la section 4.1.1, une approximation de l'espérance  $\mathbb{E}_{z_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})]$  a été obtenue au moyen de la transformée unscented. Cela nous a permis d'obtenir une fonction d'incertitude « moyennée » sur les futures observations.

Dans la suite de nos travaux, une comparaison de cette approche avec la simplification consistant à annuler l'innovation du filtre, pourra être considérée de manière à mieux qualifier l'importance du calcul de cette espérance.

### 5.1.2 Prise en compte du bruit de dynamique

Considérons que le bruit de dynamique  $w_k$ , supposé additif, gaussien, centré et de matrice de covariance  $Q_k$  non nulle, ne soit pas négligé pour la stratégie d'exploration. Nous montrons ici que lorsque ce bruit est de la forme  $\sigma^2 \mathbb{I}_2$ , les règles heuristiques avancées dans la partie 3.2 ainsi que notre stratégie d'exploration restent valides.

Soient  $\hat{x}_{k|k}$  et  $P_{k|k}$  les moments d'ordre un et deux de la loi de filtrage gaussienne  $p(x_k|z_{1:k})$  à un instant  $k$ . Soit  $\Lambda = \text{Diag}(\lambda_1, \lambda_2)$  la matrice diagonale constituée des valeurs propres  $\lambda_1, \lambda_2$  de  $P_{k|k}$ . La matrice de covariance  $P_{k|k}$  peut être diagonalisée de sorte que

$$V_{k|k}^T P_{k|k} V_{k|k} = \Lambda_{k|k}, \quad (5.3)$$

avec  $V_{k|k}$  la matrice orthogonale dont les colonnes sont les deux vecteurs propres de  $P_{k|k}$ . À partir du modèle de dynamique (1.10), lorsque la matrice de covariance  $Q_k$  du bruit de dynamique est nulle, la matrice de covariance de la densité de prédiction du vecteur d'état s'écrit

$$\begin{aligned} P_{k+1|k} &= R(\phi)^T P_{k|k} R(\phi) \\ &= R(\phi)^T V_{k|k} \Lambda_{k|k} V_{k|k}^T R(\phi) \\ &= V_{k+1|k}(\phi) \Lambda_{k|k} V_{k+1|k}^T(\phi). \end{aligned} \quad (5.4)$$

Lorsque la matrice  $Q_k$  est non nulle et de la forme  $\sigma^2 \mathbb{I}_2$ , cette relation est transformée en

$$\begin{aligned} P_{k+1|k} &= R(\phi)^T P_{k|k} R(\phi) + \sigma^2 \mathbb{I}_2 \\ &= R(\phi)^T V_{k|k} \Lambda_{k|k} V_{k|k}^T R(\phi) + \sigma^2 \mathbb{I}_2 \\ &= V_{k+1|k}(\phi) (\Lambda_{k|k} + \sigma^2 \mathbb{I}_2) V_{k+1|k}^T(\phi). \end{aligned} \quad (5.5)$$

Ainsi, lorsque le bruit de dynamique  $\sigma^2\mathbb{I}_2$  est pris en compte dans l'équation de dynamique, la matrice  $V_{k+1|k}(\phi)$  des vecteurs propres de  $P_{k+1|k}$ , qui supportent les axes de l'ellipse de confiance associée à la loi de prédiction  $p(x_{k+1}|z_{1:k})$  du vecteur d'état, n'est pas modifiée. L'orientation et la position de l'ellipse sont donc inchangées avec l'ajout de ce bruit. Seules les valeurs propres de  $P_{k+1|k}$  sont modifiées, ce qui a une incidence sur la taille de cette ellipse. En conséquence, l'entropie  $H(\mathbf{x}_{k+1}|z_{1:k})$  de la loi de prédiction augmente, mais reste indépendante du vecteur de commande  $u_k$ .

Ainsi, les calculs effectués en 2.3.2 permettant d'établir l'expression de la fonction d'incertitude  $J_1$ , restent inchangée. De même, les intuitions déclinées dans la partie 3.2 restent inchangées et les résultats obtenus pour un horizon court terme demeurent valides. Pour la stratégie à long terme, les résultats restent cohérents mais le théorème 2.3 développant l'expression récurrente de  $J_N$  est un peu modifié. En effet, comme la covariance du bruit de dynamique est non nulle, l'égalité entre l'entropie de la loi de filtrage à un instant  $k$  et l'entropie de la loi de prédiction suite au mouvement, énoncée dans le théorème 2.1, ne tient plus et doit être remplacée par

$$H(\mathbf{x}_k|z_{1:k}) \leq H(\mathbf{x}_{k+1}|z_{1:k}). \quad (5.6)$$

Ainsi, dans une approche par récurrence similaire à la démonstration du théorème 2.3, il peut être établi que

$$J_N = \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})] \quad (5.7)$$

$$\leq K'_N - F_1(\bar{u}_1) - \sum_{i=2}^N \mathbb{E}_{\mathbf{z}_{k+1:k+i-1}|z_{1:k}} [F_i(\bar{u}_i, z_{k+1:k+i-1})]. \quad (5.8)$$

De ce fait, l'expression (5.8) constitue, lorsqu'un bruit de dynamique de la forme  $\sigma^2\mathbb{I}_2$  est pris en compte, une borne supérieure de l'espérance sur les observations futures conditionnellement aux observations passées, de l'entropie  $H(\mathbf{x}_{k+N}|z_{1:k+N})$ . Dans ce cas, la stratégie d'exploration consiste en la minimisation de cette borne supérieure.

### 5.1.3 Entropie d'un mélange de gaussiennes

L'hypothèse de l'exploration depuis une connaissance *a priori* s'exprimant comme une loi gaussienne provient du fait que l'entropie d'un mélange de gaussiennes ne peut être calculé analytiquement. Cette hypothèse simplificatrice peut néanmoins être justifiée par le fait que, assez rapidement lors du processus de localisation, l'ambiguïté avant-arrière est levée. Dans ce cas, l'approximation « moment-matched » constitue alors une approximation raisonnable de la loi  $p(x_k|y_{1:k})$ . Cependant, pour les mouvements qui n'impliquent pas de rotation du capteur binaural, cette ambiguïté subsiste. Lorsque les modes du mélange de gaussiennes sont éloignés, l'erreur commise par l'approximation peut être assez importante (voir par exemple

le mouvement de translation en figure 3.12). Dans nos expériences en conditions réelles 3.6.2, l'exploration est guidée par l'hypothèse gaussienne de plus fort poids de la loi  $p(x_k|y_{1:k})$ . Cette solution n'est pas non plus idéale car le capteur binaural peut être guidé par une hypothèse fausse.

Une des perspectives à court terme de nos travaux consiste à faire évoluer notre fonction d'incertitude de façon à lever l'hypothèse gaussienne, trop restrictive. L'entropie d'une loi définie par un mélange de gaussiennes ne pouvant pas être calculée de façon formelle, des approximations doivent être utilisées. Dans [Huber *et al.* 2008], une expression analytique d'une borne supérieure de l'entropie d'une telle loi est établie. Soit  $\mathbf{x} \in \mathbb{R}^n$  un vecteur aléatoire dont la pdf  $p(x)$  est définie par le mélange de  $L$  hypothèses gaussiennes

$$p(x) = \sum_{i=1}^L w_i \mathcal{N}(x; \hat{x}_i, P_i), \quad (5.9)$$

avec  $\hat{x}_i$  et  $P_i$  les deux premiers moments de chaque hypothèse et  $w_i$  la probabilité (ou « poids ») de ces hypothèses. Notons  $H(\mathbf{x})$ , l'entropie de cette loi. Une borne supérieure de  $H(\mathbf{x})$ , notée  $H_{sup}(\mathbf{x})$  est définie par

$$H_{sup}(\mathbf{x}) = \sum_{i=1}^L w_i \cdot \left( -\ln w_i + \frac{1}{2} \ln ((2\pi e)^n |P_i|) \right). \quad (5.10)$$

Les auteurs affirment que  $H_{sup}(\mathbf{x})$  est significativement plus proche de la vraie entropie que la borne définie par l'entropie de l'approximation « moment-matched » de  $p(x)$ . De plus, la borne est exacte dans le cas gaussien ou lorsque le support partagé entre les hypothèses devient négligeable. Par ailleurs, le calcul de cette borne supérieure peut être simplifié en fusionnant certaines hypothèses, lorsque celles-ci décrivent un même mode (par exemple, dans la figure 3.12, les hypothèses qui décrivent chaque mode introduit par l'ambiguïté avant-arrière, pourraient être regroupées). Une fois le nombre d'hypothèses réduit, le calcul d'une borne supérieure simplifiée peut être réalisé à partir de 5.10.

Ainsi, en considérant que la loi  $p(x_k|z_{1:k})$  est un mélange de gaussiennes, nous pouvons définir la fonction d'incertitude  $\bar{J}_N$ , borne supérieure de l'espérance sur les futures observations conditionnellement aux observations passées, de la loi  $p(x_{k+N}|z_{1:k+N})$ , de sorte que

$$\bar{J}_N = \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H_{sup}(\mathbf{x}_{k+N}|z_{1:k+N})] \geq \mathbb{E}_{\mathbf{z}_{k+1:k+N}|z_{1:k}} [H(\mathbf{x}_{k+N}|z_{1:k+N})]. \quad (5.11)$$

Le problème d'optimisation pourrait alors consister à minimiser  $\bar{J}_N$ .

## 5.2 Évolution vers une stratégie multiobjectifs

Dans de nombreuses applications, le problème d'exploration ne constitue pas une fin en soi mais doit être combiné avec d'autres objectifs, qui peuvent être

de différentes natures (minimisation de l'énergie, position cible à atteindre, zones dangereuses à éviter, etc.). Dans cette section, nous étudions un problème similaire à celui posé dans [Schymura *et al.* 2017], où le mouvement du capteur binaural doit résoudre deux objectifs possiblement contradictoires : l'exploration qui consiste à améliorer le processus de localisation et un objectif de stabilisation consistant en la minimisation de la distance à une cible.

Soit  $S$  la position cible que le robot doit atteindre, exprimée dans le repère monde  $\mathcal{F}_{k=0}$ . Le vecteur aléatoire  $\mathbf{s}_k$  représente la position cette cible à un instant  $k$  dans le repère  $\mathcal{F}_k$  lié au capteur binaural. L'évolution de  $\mathbf{s}_k$  est donc décrite par le modèle de dynamique (1.10). Une loi de filtrage  $p(\mathbf{s}_k | o_{1:k})$  de  $\mathbf{s}_k$ , conditionnellement au fait qu'un processus d'observation  $o_{1:k}$  se réalise en la séquence de mesures  $o_{1:k}$ , pourrait être construite dans une méthode de filtrage stochastique indépendante du processus de localisation audio-motrice. Cependant, pour cette étude quelque peu prospective, nous considérons que la localisation de la cible est parfaite et que ainsi, sa position  $s_k$  dans le repère du robot est connue en chaque instant  $k$ .

Exprimons la distance du capteur à la cible à l'instant  $k + 1$ , après l'application d'un vecteur de commande  $\bar{u}_1$ , comme une fonction de ce vecteur, *i.e.*,

$$C(\bar{u}_1) = \|\mathbf{s}_{k+1}\|. \quad (5.12)$$

Le problème peut donc être exprimé comme la recherche du vecteur de commande  $\bar{u}_1^*$  minimisant  $C(\bar{u}_1)$ , *i.e.*, solution du problème d'optimisation

$$(\mathcal{Q}_1) \begin{cases} (\bar{u}_1) = \arg \min_{(T_y, T_z) \in \mathcal{T}} C(\bar{u}_1) \\ = \arg \min_{(T_y, T_z) \in \mathcal{T}} \|\mathbf{s}_{k+1}\|. \end{cases} \quad (5.13)$$

Le problème est formulé uniquement en terme de position et aucune orientation du capteur par rapport à la cible n'est imposée. Notons alors que le problème n'est contraint que sur le domaine  $\mathcal{T}$  car la norme de  $\mathbf{s}_{k+1}$  est indépendante de la rotation  $\phi_k$ . La minimisation simultanée de  $J_1(\bar{u}_1)$  et  $C(\bar{u}_1)$  dans un problème d'optimisation multiobjectif n'étant pas réalisable, le problème d'optimisation  $\mathcal{P}'_1$  est défini dans l'énoncé suivant.

**Énoncé 5.1** (Problème  $\mathcal{P}'_1$ ). *Soit la densité de probabilité gaussienne  $p(x_k | z_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k})$  du vecteur d'état  $\mathbf{x}_k$  à un instant  $k$  du processus de localisation, conditionnellement au fait que le processus d'observation  $\mathbf{z}_{1:k}$  se réalise en la séquence de mesures  $z_{1:k}$ . Soit  $J_1(\bar{u}_1)$  la fonction d'incertitude définie comme l'espérance sur la future observation  $\mathbf{z}_{k+1}$  conditionnellement aux observations passées  $z_{1:k}$ , de l'entropie de la densité  $p(x_{k+1} | z_{1:k+1})$  à l'instant  $k + 1$ , après l'application de la commande  $\bar{u}_1$  (voir section 2.3.2). Soit  $C(\bar{u}_1) = \|\mathbf{s}_{k+1}\|$  la distance du robot par rapport à la cible (dont la position est supposée connue), après l'application du vecteur de commande  $\bar{u}_1$ . Le vecteur de commande optimal*

$\bar{u}_1^* = u_k^* = (T_k^*, \phi_k^*)^T$  recherché minimise une fonction

$$P(\bar{u}_1) = \lambda J_1(\bar{u}_1) + (1 - \lambda)C(\bar{u}_1) \quad \lambda \in [0, 1], \quad (5.14)$$

sous les contraintes  $T_k^* \in \mathcal{T}$  et  $\phi_k^* \in \mathcal{R}$ , i.e., est solution du problème

$$(\mathcal{P}'_1) \begin{cases} \bar{u}_1^* = \arg \min_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} P(\bar{u}_1) \\ = \arg \min_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} \lambda J_1(\bar{u}_1) + (1 - \lambda)C(\bar{u}_1). \\ = \arg \max_{\bar{u}_1 \in \mathcal{T} \times \mathcal{R}} \lambda F_1(\bar{u}_1) + (\lambda - 1)C(\bar{u}_1). \end{cases} \quad (5.15)$$

$\lambda$  est un coefficient qui fixe « l'importance » donnée à chaque objectif.

Définissons  $\nabla P$ , le gradient de cette fonction par rapport au vecteur de commande  $\bar{u}_1$ , i.e.,

$$\nabla P = \lambda \nabla F_1(\bar{u}_1) + (\lambda - 1) \nabla C(\bar{u}_1) \quad \lambda \in [0, 1]. \quad (5.16)$$

Comme la rotation du capteur n'est dirigée que par l'objectif d'exploration, la composante de  $\nabla C(\bar{u}_1)$  correspondant à la dérivée de  $C$  par rapport à la rotation est nulle. Le problème  $\mathcal{P}'_1$  peut ensuite être résolu en modifiant l'algorithme du gradient projeté (3.1), en suivant la direction de montée de la fonction  $P$ , définie par son gradient  $\nabla P$ .

Des résultats préliminaires sont présentés dans la figure 5.1. La stratégie de déplacement alliant l'exploration et l'objectif de positionnement a été évaluée en simulation dans les mêmes conditions que celles décrites en 1.4.1. Le vecteur de commande du capteur binaural  $\bar{u}_1$  est calculé à chaque instant  $k$  en résolvant le problème  $\mathcal{P}'_1$ . À l'instar de la section 3.6.1, nous considérons une connaissance *a priori*  $p(x_0)$  gaussienne de la position de la source, avec ici,  $\hat{x}_0 = (1, 2)^T$  sa moyenne et  $P_0 = \begin{pmatrix} 0.0121 & 0.0074 \\ 0.0074 & 0.0231 \end{pmatrix}$  sa matrice de covariance. Deux scénarios (I) et (II), dans lesquels la cible est respectivement positionnée à  $S = (-1, 1)$  et  $S = (1, -1)$ , sont considérés. Pour chacun de ces scénarios, l'expérience a été menée pour différentes valeurs du coefficient  $\lambda = \{1, 0.98, 0.94, 0.9, 0\}$ . De telles valeurs très proches de  $\lambda = 1$  ont été choisies car les fonctions d'incertitude  $J_1$  et de coût  $C$  ne sont pas normalisées, si bien qu'une valeur intermédiaire  $\lambda = 0.5$  donne une plus grande importance à la minimisation de  $C$ . Dans le futur, une normalisation du critère devra être étudiée afin que la valeur  $\lambda = 0.5$ , confère la même importance à chaque objectif.

Nous observons que dans les deux scénarios, la stratégie d'exploration pure ( $\lambda = 1$ ), conduit à la plus faible entropie de la loi  $p(x_k | z_{1:k})$  en chaque instant de l'expérience. Lorsque la valeur de  $\lambda$  diminue, le capteur est de plus en plus attiré vers la cible et la réduction d'entropie par rapport à l'état initial est moins marquée. Notons que dans le cas où le mouvement est calculé en fonction de l'objectif de position seul ( $\lambda = 0$ ), l'entropie est maximale en chaque instant. Ensuite, comme la rotation n'est pas dirigée par l'objectif de position, le capteur est orienté dans

la direction estimée de la source même lorsque celui ci se dirige vers la cible. Le coefficient  $\lambda$  qui donne de l'importance à l'un des objectifs au détriment du second, doit être choisi par l'utilisateur avec attention. Nous voyons ici apparaître le dilemme d'exploration-exploitation, qui est au coeur du problème de perception en robotique pour les problèmes multiobjectifs. C'est à dire, comment décider entre :

- Renforcer la perception de l'environnement en exploitant l'information disponible. Il s'agit d'une tâche de stabilisation, consistant ici à satisfaire le critère de coût en dirigeant le robot vers la cible.
- Explorer les zones inconnues de façon à obtenir de nouvelles informations, *i.e.*, dans ce contexte, privilégier la stratégie d'exploration afin de maximiser l'information sur la position de la source.

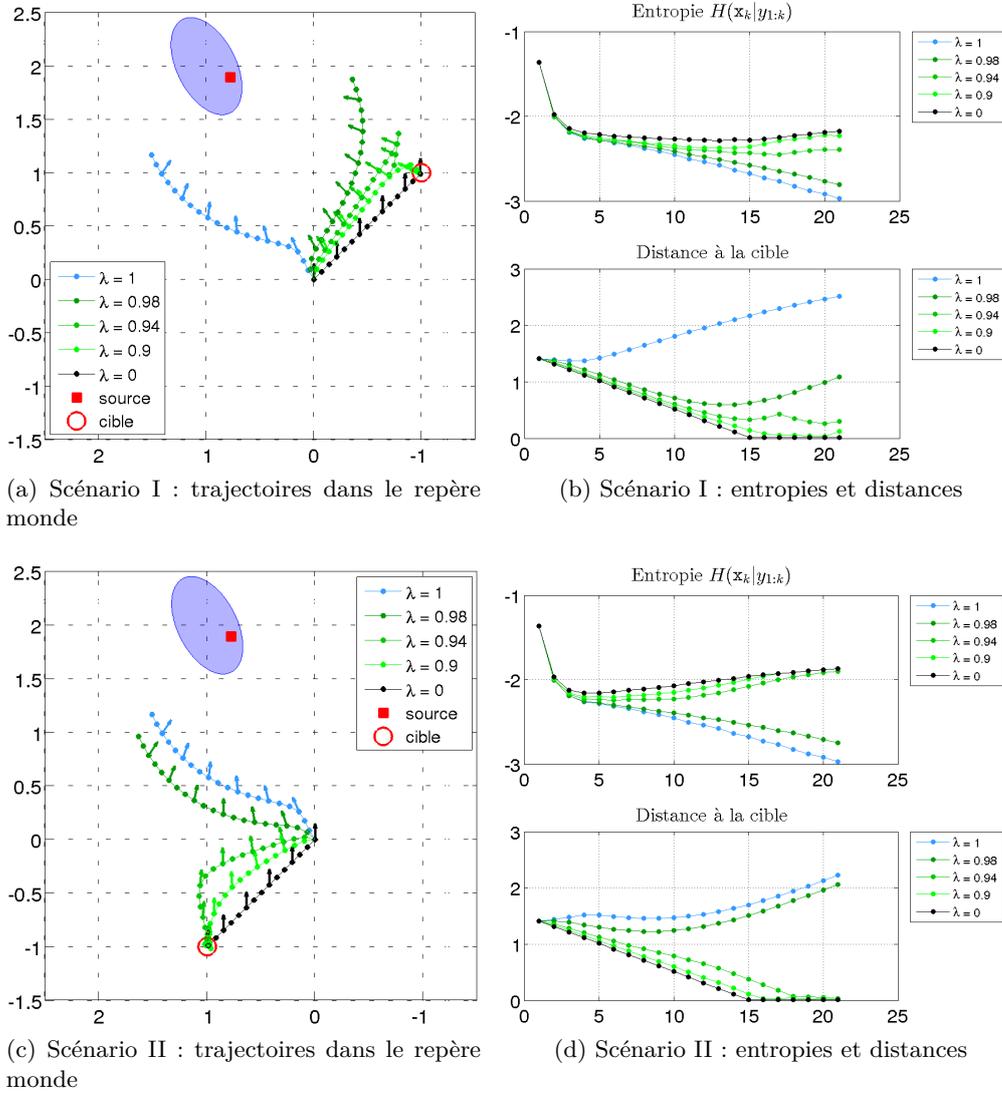


FIGURE 5.1 – Localisation active d’une source avec un objectif de positionnement par rapport à une cible – Le mouvement du capteur binaural est issu de la résolution de  $\mathcal{P}'_1$  en chaque instant. Différentes valeurs de  $\lambda = \{1, 0.98, 0.94, 0.9, 0\}$  sont envisagées dans chaque scénario (I,II). (a,c) L’ellipse de confiance est associée à l’estimé initial de la position relative de la source. Les positions de la source et de la cible dans le repère monde (le repère  $\mathcal{F}_{k=0}$ ) sont représentées. Les flèches indiquent la direction de l’axe frontal du capteur. (b,d) Entropies  $H(\mathbf{x}_k|z_{1:l_k})$  et distance à la cible (ordonnées) en chaque itération (abscisses).

# Conclusion

Dans cette thèse nous avons étudié la localisation binaurale d'une source sonore selon un schéma en trois phases : (A) estimation de primitives spatiales par une analyse court-terme du flux audio ; (B) localisation audio-motrice par l'assimilation de ces données et la combinaison avec les ordres moteurs du capteur au sein d'un schéma d'estimation stochastique ; (C) commande en boucle fermée du mouvement du capteur de façon à améliorer la qualité de la localisation. Les objectifs de cette thèse ont été de définir des stratégies de « mouvement actif » constituant la phase (C) et de les implémenter afin de les valider par des simulations réalistes et par des expérimentations sur un ensemble tête-torse anthropomorphe doté de perception binaurale et de mobilité.

À tout instant du processus de localisation, un critère d'information a été défini comme l'entropie de la loi de filtrage au terme d'un horizon glissant dans le futur, moyennée sur les futures observations (encore inconnues) conditionnellement aux observations passées. Le problème à ensuite été formulé comme la maximisation de ce critère d'information. L'horizon peut être constitué du prochain instant ou des  $N$  prochains instants, ce qui donne lieu à une stratégie court-terme ou long-terme, respectivement. Une évaluation de ce critère par utilisation de la transformée unscnted (permettant d'approximer l'espérance), ainsi que le calcul automatique du gradient de celle-ci par exploitation des nombres duaux, ont permis la détermination de la commande en boucle fermée sur l'audio à appliquer au capteur binaural.

**Prospectives.** Certaines perspectives ouvertes par nos travaux ont été évoquées dans le chapitre 5. Il s'agit essentiellement d'évolutions « court terme » de notre méthode qui permettraient d'améliorer les résultats, en particulier la section 5.1.3 où nous envisageons de lever l'hypothèse gaussienne, qui pourrait être, dans certaines situations, trop restrictive. Nous avons également présenté en 5.2 une extension de notre problème où le mouvement du capteur binaural est établi de façon à remplir plusieurs objectifs.

Une autre prospective aura pour objectif de redéfinir le modèle d'exploration pour la phase (C). Les variables d'observation issues de la phase (A) n'étant pas reliées aux vecteurs d'état par un modèle d'observation analytique, nous avons été contraints d'introduire un tel modèle (équation d'observation faisant intervenir les ITD de Woodworth-Schlosberg) afin d'explicitier ce lien. L'exploration n'est alors guidée que par ce modèle avec les restrictions qu'il impose (modèle d'ITD en champ lointain pour une tête sphérique), sans tenir compte d'autres indices tels que les ILD qui interviennent pourtant dans l'estimation de l'azimut. Une première évolution de notre modèle d'exploration pourra consister à inclure une expression des ILD dans le modèle d'observation. Ainsi, un nouveau vecteur d'observation  $\mathbf{z}_k$  pourra être défini, dont les composantes seront les mesures d'ITD et d'ILD. En dépit du fait que le modèle d'observation fera intervenir des fonctions mathématiques assez

complexes comme les polynômes de Legendre ou les fonctions de Hankel, le cadre développé dans cette thèse devrait permettre de synthétiser un mouvement actif sous ces hypothèses. Ensuite, une autre approche pourrait être de définir le problème d’exploration sans introduire de modèle d’observation analytique, en exploitant directement les observations issues de la phase (A). Se posera alors le problème d’exploration sur le long terme. Comment calculer l’espérance sur les observations futures alors qu’aucun modèle de leur réalisation en fonction d’un état futur possible n’est défini ?

D’autres travaux prospectifs pourront consister à faire évoluer l’étape (B) du processus de localisation et ainsi développer une stratégie de filtrage multi-cibles, de façon à assimiler des mesures d’azimuts provenant de différentes sources. Dans un premier temps, la stratégie appliquée dans l’étape (A), devrait être remplacée par une approche d’estimation de plusieurs azimuts, telle que proposée dans [Portello *et al.* 2014a]. Les principales difficultés du filtrage multi-cibles résident dans le fait que les mesures ne sont pas ordonnées (l’estimation court terme de l’azimut de plusieurs sources ne constitue qu’une information instantanée) et que le nombre de sources détectées (ainsi que le nombre de sources réellement actives) peut varier assez fréquemment. Une première approche de filtrage multi-cibles pourrait être d’appliquer une stratégie de filtrage mono-cible pour chaque source et d’assimiler dans chaque filtre, l’observation qui correspond le mieux à sa mesure prédite. Des solutions plus élaborées comme le [Joint Probabilistic Data Association Filter \(JPDAF\)](#), extension du filtre [PDAF](#) au cas multi-cibles, ou encore le [Reversible-Jump Markov Chain Monte Carlo \(RJ-MCMC\)](#) pourront être envisagées. Une approche plus adaptée au filtrage multi-cibles pourrait reposer sur le formalisme des ensembles finis aléatoires ([Random Finite Set \(RFS\)](#)). Ensuite, une adaptation de la phase (C) permettrait de développer une version active de la localisation multi-cibles et ainsi calculer la commande du capteur améliorant globalement la localisation de l’ensemble des sources, ou bien privilégiant la localisation d’une source au détriment d’une autre, dans une approche similaire à ce que nous avons présenté en 5.2.

Ensuite, une autre prospective consistera à développer une solution de [SLAM](#) audio binaural. Si les méthodes de [SLAM](#) sur la base de capteurs visuels ont reçu une attention particulière dans la littérature, relativement peu d’études traitent de la problématique [SLAM](#) dans le domaine de l’audition [[Martinson & Schultz 2009](#), [Sasaki \*et al.\* 2010](#), [Hu \*et al.\* 2011](#), [Evers \*et al.\* 2016b](#)], et se focalisent souvent sur des approches à base d’antennes. Les principales difficultés sont généralement celles soulevées par le filtrage multi-cibles, ainsi que la robustesse des algorithmes d’estimation court terme face aux bruits et réverbérations. À l’instar des approches par vision, des extensions actives du [SLAM](#) audio seront développées, dans la veine de [[Evers \*et al.\* 2016a](#)]. Ensuite seront étudiés les problèmes de localisation et de [SLAM](#) actif, sur la base d’une tête binaurale mobile et couplée à d’autres capteurs audio, éventuellement contrôlables, déployés dans l’environnement. Des méthodes de couplage entre l’audition et la vision pourront également être considérées.

Enfin, il sera opportun d’étudier un peu plus les connexions possibles entre l’approche que nous avons développée, en particulier les approximateurs à base de

---

transformée unscented, et les techniques plus classiques de décision sur un horizon fuyant telles que la commande optimale stochastique, la programmation neurodynamique [Scardovi 2005, Bertsekas & Tsitsiklis 1996] ou le formalisme des processus de décision markoviens partiellement observables (Partially Observable Markov Decision Process (POMDP)) [Ponzoni Carvalho Chanel 2013, Araya-Lopez 2013].



# Le filtre de Kalman unscented

---

Soit un processus tel que celui présenté en 2.1.3 avec un modèle de dynamique et d'observation de la forme

$$\begin{cases} \mathbf{x}_{k+1} = f(\mathbf{x}_k, u_k) + \mathbf{w}_k \\ \mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k, \end{cases} \quad (\text{A.1})$$

avec  $f, h$  deux fonctions non linéaires et  $\mathbf{w}_k, \mathbf{v}_k$  représentant respectivement les bruits de dynamique et d'observation qui sont blancs, mutuellement indépendants, indépendants de l'état initial et gaussiens centrés de covariances connues, *e.g.*,  $\mathbf{w}_k \sim \mathcal{N}(0, Q_k)$  et  $\mathbf{v}_k \sim \mathcal{N}(0, R_k)$ .

Le filtre de Kalman permet de calculer récursivement les deux premiers moments d'approximations gaussiennes de la densité de filtrage  $p(x_k|z_{1:k})$  ainsi que des pdf  $p(x_k|z_{1:k-1})$  et  $p(z_k|z_{1:k-1})$ . Dans le cas du filtre de Kalman classique, les modèles de dynamique et d'observation doivent être linéaires, pour que l'hypothèse gaussienne puisse être maintenue. Le filtre de Kalman unscented (**UKF**) est une extension du filtre classique et peut être appliqué à des modèles non linéaires.

L'**UKF** repose sur la transformée unscented que nous présentons en A.1 et constitue une alternative au filtre de Kalman étendu (**EKF**). Nous donnons ensuite l'algorithme complet d'une version de l'**UKF** dans la partie A.2. La présentation que nous faisons de la transformée unscented et de l'implémentation de l'**UKF** est inspirée de [Portello 2013].

## A.1 La transformée unscented

La transformée unscented, initialement proposée par [Julier & Uhlmann 2004] est une technique de linéarisation statistique qui permet d'estimer le résultat de l'application d'une transformation non linéaire sur une densité de probabilité. Un des avantages de la méthode est qu'elle se base sur un tirage déterministe de points de la distribution initiale (appelés *sigma-points*) afin de les propager au travers de la transformation non linéaire. Ainsi le procédé ne nécessite pas de calcul de dérivées ou de développement de Taylor.

Soit deux vecteurs aléatoires  $\mathbf{x}$  et  $\mathbf{y}$  de dimensions respectives  $n_x$  et  $n_y$ . Considérons que  $\mathbf{x}$  soit associée à une pdf gaussienne  $\mathcal{N}(x; \hat{x}, P_x)$ <sup>1</sup>. Soit  $f$  une fonction

---

1. La transformation unscented peut en fait être appliquée pour une distribution quelconque.

non linéaire quelconque liant  $\mathbf{x}$  et  $\mathbf{y}$  de sorte que

$$\mathbf{y} = f(\mathbf{x}). \quad (\text{A.2})$$

La transformée unscented permet de calculer une approximation gaussienne  $\mathcal{N}(\mathbf{y}; \hat{\mathbf{y}}, P_y)$ , de la densité  $p(\mathbf{y})$ . Des points particuliers  $\{X_i\}_{i=0, \dots, 2n_x}$  nommés *sigma-points* sont sélectionnés à partir de la loi de  $\mathbf{x}$ . Nous notons  $[\{X_i\}, \{w_i^m\}, \{w_i^c\}] = \text{sp}(\hat{\mathbf{x}}, L, \gamma)$  l'application permettant de calculer ces *sigma-points*, donnée par

$$(\text{sp}) \begin{cases} X_0 & = \hat{\mathbf{x}} \\ X_i & = \hat{\mathbf{x}} + (\sqrt{n_x + \gamma} L_x)_i \quad \forall i \in 1, \dots, n_x \\ X_{n_x+i} & = \hat{\mathbf{x}} - (\sqrt{n_x + \gamma} L_x)_i \quad \forall i \in 1, \dots, n_x, \end{cases} \quad (\text{A.3})$$

avec  $w_i^m$  et  $w_i^c$  les poids de la transformée définis par

$$\begin{aligned} w_0^m &= \frac{\gamma}{n_x + \gamma} & \text{et} & \quad w_i^m = \frac{\gamma}{(2n_x + \gamma)} & \text{pour} & \quad i = 1, \dots, 2n_x, \\ w_0^c &= \frac{\gamma}{n_x + \gamma} + n_x + \gamma & \text{et} & \quad w_i^c = \frac{\gamma}{(2n_x + \gamma)} & \text{pour} & \quad i = 1, \dots, 2n_x. \end{aligned}$$

Ci-dessus,  $\gamma = \alpha^2(n_x + \kappa) - n_x$  est un paramètre de réglage, fonction de  $\{\alpha, \kappa\}$  qui caractérise la dispersion des *sigma-points* autour de  $\hat{\mathbf{x}}$ , et  $L_x$  désigne la décomposition de Cholesky de  $P_x$  de sorte que  $P_x = L_x L_x^T$ . La notation  $(\sqrt{n_x + \gamma} L_x)_i$  désigne la  $i^{\text{ème}}$  colonne de cette matrice. La fonction  $f$  est ensuite appliquée sur les *sigma-points*  $\{X_i\}$

$$Y_i = f(X_i), \quad \text{pour } i = 0, \dots, 2n_x. \quad (\text{A.4})$$

La moyenne  $\hat{\mathbf{y}}$  ainsi que les covariances  $P_{yy}$  et  $P_{xy}$  respectivement définies par  $\mathbb{E}_y[(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})^T]$  et  $\mathbb{E}_{xy}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{y} - \hat{\mathbf{y}})^T]$ , sont ensuite approximées au moyen des applications  $\text{spMoy}$  et  $\text{spCorr}$  suivantes :

$$\hat{\mathbf{y}} = \text{spMoy}(\{w_i^m\}, \{Y_i\}) = \sum_{i=0}^{2n_x} w_i^m Y_i. \quad (\text{A.5})$$

et

$$P_{yy} = \text{spCorr}(\{w_i^c\}, \{Y_i\}, \hat{\mathbf{y}}, \{Y_i\}, \hat{\mathbf{y}}) = \sum_{i=0}^{2n_x} w_i^c (Y_i - \hat{\mathbf{y}})(Y_i - \hat{\mathbf{y}})^T, \quad (\text{A.6})$$

$$P_{xy} = \text{spCorr}(\{w_i^c\}, \{X_i\}, \hat{\mathbf{x}}, \{Y_i\}, \hat{\mathbf{y}}) = \sum_{i=0}^{2n_x} w_i^c (X_i - \hat{\mathbf{x}})(Y_i - \hat{\mathbf{y}})^T. \quad (\text{A.7})$$

## A.2 Implémentation de l'UKF

L'UKF peut être décomposé en une étape de prédiction et de mise à jour. La prédiction consiste en le calcul de  $\{\hat{x}_{k+1|k}, P_{k+1|k}\}$  et  $\{\hat{z}_{k+1|k}, S_{k+1|k}\}$  qui sont respectivement les premiers moments des pdf  $p(x_{k+1}|z_{1:k})$  et  $p(z_{k+1}|z_{1:k})$  suite à l'application du vecteur de commandes  $u_k$ . L'étape de mise à jour combine l'observation  $z_{k+1}$  avec le résultat de la prédiction de façon à obtenir les moments de la pdf *a posteriori*  $p(x_{k+1}|z_{1:k+1})$ . Les étapes de prédiction et de mise à jour du filtre sont respectivement décrites dans les algorithmes A.1 et A.2.

---

**Algorithme A.1 :** Étape de prédiction de l'UKF

$$\left[ \hat{x}_{k+1|k}, P_{k+1|k}, \hat{z}_{k+1|k}, S_{k+1|k}, T_{k+1|k} \right] = \text{UKF\_PRED} \left( \hat{x}_{k|k}, P_{k|k}, u_k \right)$$


---

**Entrées :**

—  $\hat{x}_{k|k}, P_{k|k}, u_k$

**Paramètres :**

—  $f, h, Q_k, R_k, \gamma$

**Sorties :**

—  $\hat{x}_{k+1|k}, P_{k+1|k}, \hat{z}_{k+1|k}, S_{k+1|k}, T_{k+1|k}$

---

Décomposition de Cholesky de  $P_{k|k}$

1  $P_{k|k} = L_{k|k} L_{k|k}^T$  ;

Calcul des *sigma-points* de  $x_{k|k}$

2  $\left[ X_{i,k|k}, w_i^m, w_i^c \right] = \text{sp}(\hat{x}_{k|k}, L_{k|k}, \gamma)$  ;

Application du modèle de dynamique

3 **Pour**  $i = 0, \dots, 2n_x$  **faire**

4 |  $\{X_{i,k+1|k}\} = f(X_{i,k|k}, u_k)$

5 **fin**

Calcul des deux premiers moments de  $p(x_{k+1}|z_{1:k})$

6  $\hat{x}_{k+1|k} = \text{spMoy}(\{w_i^m\}, \{X_{i,k+1|k}\})$  ;

7  $P_{k+1|k} = \text{spCorr}(\{w_i^c\}\{X_{i,k+1|k}\}, \hat{x}_{k+1|k}, \{X_{i,k+1|k}\}, \hat{x}_{k+1|k}) + Q_k$  ;

Nouveau tirage des *sigma-points* pour inclure le bruit de dynamique

8  $P_{k+1|k} = L_{k+1|k} L_{k+1|k}^T$  ;

9  $\left[ \{X_{i,k+1|k}\}, \{w_i^m\}, \{w_i^c\} \right] = \text{sp}(\hat{x}_{k+1|k}, L_{k+1|k}, \gamma)$  ;

Application du modèle d'observation

10 **Pour**  $i = 0, \dots, 2n_x$  **faire**

11 |  $\{Z_{i,k+1|k}\} = h(X_{i,k+1|k})$

12 **fin**

Calcul des deux premiers moments de  $p(z_{k+1}|z_{1:k})$

13  $\hat{z}_{k+1|k} = \text{spMoy}(\{w_i^m\}\{Z_{i,k+1|k}\})$  ;

14  $S_{k+1|k} = \text{spCorr}(\{w_i^c\}\{Z_{i,k+1|k}\}, \hat{z}_{k+1|k}, \{Z_{i,k+1|k}\}, \hat{z}_{k+1|k}) + R_k$  ;

Calcul de la covariance croisée de  $x_{k+1}$  et de  $z_{k+1}$  conditionnellement à  $z_{1:k}$

15  $T_{k+1|k} = \text{spCorr}(\{w_i^c\}\{X_{i,k+1|k}\}, \hat{x}_{k+1|k}, \{Z_{i,k+1|k}\}, \hat{z}_{k+1|k})$  ;

---

---

**Algorithme A.2 :** Étape de mise à jour de l'UKF

$$\left[ \hat{x}_{k+1|k+1}, P_{k+1|k+1} \right] = \text{UKF\_MAJ} \left( \hat{x}_{k+1|k}, P_{k+1|k}, \hat{z}_{k+1|k}, S_{k+1|k}, T_{k+1|k}, z_{k+1} \right)$$


---

**Entrées :**

- $\hat{x}_{k+1|k}, P_{k+1|k}, \hat{z}_{k+1|k}, S_{k+1|k}, T_{k+1|k}$
- observation  $z_{k+1}$

**Sorties :**

- $\hat{x}_{k+1|k+1}, P_{k+1|k+1}$
- 

Calcul des deux premiers moments de  $\mathbf{x}_{k+1|k+1}$

- 1  $K_k = T_{k+1|k} (S_{k+1|k})^{-1}$  ;
  - 2  $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_k (z_{k+1} - \hat{z}_{k+1|k})$  ;
  - 3  $P_{k+1|k+1} = P_{k+1|k} - K_k S_{k+1|k} K_k^T$  ;
-



# Optimisation statique

---

Après le rappel de certaines définitions fondamentales en B.1, nous présentons dans cette section les conditions d'optimalité pour des problèmes d'optimisation sans contraintes B.2 ou avec contraintes B.3. Une présentation plus complète des approches analytiques ainsi que des techniques numériques pour l'optimisation statique, peut être trouvée dans [Hiriart-Urruty 2008].

## B.1 Définitions

Dans cette section nous rappelons certaines définitions

**Définition B.1** (Minimum et maximum local). *Un point  $x^*$  est un minimum (resp. maximum) local de  $f$  sur  $\mathbb{R}^n$ , s'il existe un voisinage  $V(x^*)$  de  $x^*$ , tel que  $\forall x \in \mathbb{R}^n \cap V(x^*)$ ,  $f(x^*) \leq f(x)$  (resp.  $f(x^*) \geq f(x)$ ).*

**Définition B.2** (Minimum et maximum global). *Un point  $x^*$  est un minimum (resp. maximum) global de  $f$  sur  $\mathbb{R}^n$ , si  $\forall x \in \mathbb{R}^n$ ,  $f(x^*) \leq f(x)$  (resp.  $f(x^*) \geq f(x)$ ).*

Notons que lorsque les inégalités dans les définitions (B.1) et (B.2) sont strictes, le maximum ou le minimum est strict.

**Définition B.3** (Fonction convexe). *La fonction  $f : \mathbb{R}^n \mapsto \mathbb{R}$  est dite convexe si  $\forall (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$  et  $\forall t \in [0; 1]$  alors*

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2). \quad (\text{B.1})$$

## B.2 Optimisation sans contrainte

Soit  $f$  une fonction de  $x \in \mathbb{R}^n$ , à valeurs dans  $\mathbb{R}$ . Considérons le problème d'optimisation sans contrainte  $\mathcal{P}$  défini par

$$(\mathcal{P}) : \min_{x \in \mathbb{R}^n} f(x), \quad (\text{B.2})$$

**Théorème B.1** (Condition nécessaire du premier ordre). *Si  $x^*$  est un extremum local de  $f$  et si  $f$  est différentiable au voisinage de  $x^*$  et de dérivée continue, alors*

$$\nabla f(x^*) = 0. \quad (\text{B.3})$$

□

Cette condition n'est que nécessaire. En conséquence il existe des fonctions pour lesquelles cette condition est vérifiée en un point mais où ce point ne constitue pas un extremum de la fonction. C'est le cas notamment des fonctions possédant des points selles (la hessienne de la fonction en un tel point a des valeurs propres négatives et positives).

**Theorème B.2** (Condition nécessaire du second ordre). *Si  $x^*$  est un minimum (resp. maximum) local de  $f$  et  $\nabla^2 f$  est continue au voisinage de  $x^*$ , alors  $\nabla f(x^*) = 0$  et  $\nabla^2 f$  est semi-définie positive (resp. semi-définie négative).*  $\square$

**Theorème B.3** (Condition suffisante du second ordre). *Si  $\nabla f(x^*) = 0$  et  $\nabla^2 f$  est continue au voisinage de  $x^*$  et définie positive (resp. définie négative), alors  $x^*$  est un minimum (resp. maximum) local strict de  $f$ .*  $\square$

**Theorème B.4.** *Si la fonction  $f$  est convexe (resp. concave), alors tout minimum (resp. maximum) local de  $f$  est aussi un minimum (resp. maximum) global.*  $\square$

### B.3 Optimisation sous contraintes

Un problème d'optimisation peut avoir des contraintes de type égalité ou inégalité. Nous présentons dans un premier temps le cas particulier où toutes les contraintes sont de type égalité. Nous verrons ensuite le cas général où certaines contraintes sont des égalités et d'autres des inégalités.

**Contraintes égalitaires.** Soit  $f$  une fonction de  $x \in \mathbb{R}^n$  à valeurs dans  $\mathbb{R}$ . L'ensemble des contraintes  $\mathcal{C}$  du problème  $\mathcal{P}'$

$$(\mathcal{P}') : \min_{x \in \mathcal{C}} f(x). \quad (\text{B.4})$$

est exprimé par

$$\mathcal{C} = \{x \in \mathbb{R}^n \mid l_j(x) = 0, \forall j = 1, \dots, q\}, \quad (\text{B.5})$$

avec  $l_j$  les fonctions qui définissent les contraintes de type égalitaires. Tout point  $x \in \mathbb{R}^n \cap \mathcal{C}$  est un point admissible du problème  $\mathcal{P}'$ . Les conditions nécessaires d'optimalité de ce problème sont exprimées par les conditions de Lagrange définies dans le théorème (B.5).

**Theorème B.5** (Conditions de Lagrange). *Si  $x^* \in \mathbb{R}^n \cap \mathcal{C}$  est un minimum local de  $f$  sur  $\mathcal{C}$  et si les vecteurs  $\nabla l_j(x)$ ,  $j = 1, \dots, q$ , sont linéairement indépendants, alors il existe  $\lambda = (\lambda_1, \dots, \lambda_q)^T$  tel que*

$$\nabla f(x^*) + \sum_{j=1}^q \lambda_j \nabla l_j(x^*) = 0. \quad (\text{B.6})$$

Les coefficients  $\lambda$  sont appelés multiplicateurs de Lagrange.  $\square$

**Cas général.** Considérons le cas général où le problème  $\mathcal{P}'$  défini par

$$(\mathcal{P}') : \min_{x \in \mathcal{C}} f(x). \quad (\text{B.7})$$

est soumis à l'ensemble de contraintes  $\mathcal{C}$  comprenant des contraintes sous forme d'égalités et d'inégalités, *i.e.*,

$$\mathcal{C} = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \forall i = 1, \dots, p, \quad l_j(x) = 0, \forall j = 1, \dots, q\}, \quad (\text{B.8})$$

avec  $g_i$  et  $l_j$  les fonctions qui définissent respectivement les contraintes de type inégalitaires et égalitaires. Tout point  $x \in \mathbb{R}^n \cap \mathcal{C}$  est un point admissible du problème  $\mathcal{P}'$ .

**Theorème B.6** (Condition de qualification des contraintes  $(QC)_x$ ). *Si les vecteurs  $\nabla l_1(x), \dots, \nabla l_q(x)$  sont linéairement indépendants et si il n'existe pas de vecteur non nul  $\mu = (\mu_1, \dots, \mu_p)^T \in (\mathbb{R}^+)^p$  tel que  $\sum_{i=1}^p \mu_i \nabla g_i(x) = 0$ , alors l'hypothèse  $(QC)_x$  de qualification des contraintes est vérifiée.*  $\square$

**Theorème B.7** (Conditions de Karush-Kuhn-Tucker (KKT)). *Si  $x^* \in \mathbb{R}^n \cap \mathcal{C}$  est un minimum local de  $f$  sur  $\mathcal{C}$  et si l'hypothèse  $(QC)_x$  de qualification des contraintes est vérifiée, alors il existe  $(\mu_1, \dots, \mu_p)^T$  et  $(\lambda_1, \dots, \lambda_q)^T$  non tous nuls tels que*

$$\nabla F_1(u^*) + \sum_{i=1}^p \mu_i \nabla g_i(u^*) + \sum_{j=1}^q \lambda_j \nabla l_j(u^*) = 0 \quad (\text{B.9})$$

$$\mu_i \geq 0 \quad \forall i = 1, \dots, p \quad (\text{B.10})$$

$$\mu_i g_i(u^*) = 0 \quad \forall i = 1, \dots, p \quad (\text{B.11})$$

$\square$

Notons que pour un problème de maximisation, l'équation (B.10) est modifiée en

$$\mu_i \leq 0 \quad \forall i = 1, \dots, p. \quad (\text{B.12})$$



# Acronymes

---

- BOT** Bearings-Only Tracking. [37](#)
- CASA** Computational Auditory Scene Analysis. [2](#)
- CRLB** Cramér-Rao Lower Bound. [29](#)
- EKF** Extended Kalman Filter. [19](#)
- FIM** Fisher Information Matrix. [29](#)
- GLRT** Generalized Likelihood Ratio Test. [11](#)
- HATS** Head And Torso Simulator. [22](#)
- HRIR** Head Related Impulse Response. [8](#)
- HRTF** Head Related Transfert Function. [6](#)
- IIR** Interaural Impulse Response. [18](#)
- ILD** Interaural Level Difference. [7](#)
- IPD** Interaural Phase Difference. [7](#)
- ITD** Interaural Time Difference. [7](#)
- ITF** Interaural Transfert Function. [8](#)
- JPDAF** Joint Probabilistic Data Association Filter. [108](#)
- KKT** Karush Kuhn Tucker. [73](#)
- MCTS** Monte Carlo Tree Search. [14](#)
- MG-UKF** Multi-Gaussian Unscented Kalman Filter. [11](#)
- MH-PDAF** Multiple Hypothesis Probabilistic Data Association Filter. [11](#)
- MKF** Mixture Kalman Filter. [11](#)
- PCRB** Posterior Cramér-Rao Bound. [30](#)
- PDAF** Probabilistic Data Association Filter. [11](#)
- pdf** probability density function. [11](#)
- POMDP** Partially Observable Markov Decision Process. [109](#)
- RFS** Random Finite Set. [108](#)
- RJ-MCMC** Reversible-Jump Markov Chain Monte Carlo. [108](#)
- ROS** Robot Operating System. [24](#)
- SLAM** Simultaneous Localization And Mapping. [36](#)
- UKF** Unscented Kalman Filter. [19](#)



# Bibliographie

- [Aaronson & Hartmann 2014] N.L. Aaronson et W. Hartmann. *Testing, correcting, and extending the Woodworth model for interaural time difference*. The Journal of the Acoustical Society of America, vol. 135, pages 817–823, 2014. (Cité en page 21.)
- [Algazi *et al.* 2001] V. R. Algazi, C. Avendano et R. O. Duda. *Elevation localization and head-related transfer function analysis at low frequencies*. The Journal of the Acoustical Society of America, vol. 109, no. 3, pages 1110–1122, Mars 2001. (Cité en page 10.)
- [Aloimonos *et al.* 1988] J. Aloimonos, I. Weiss et A. Bandyopadhyay. *Active vision*. International Journal of Computer Vision, vol. 1, no. 4, pages 333–356, 1988. (Cité en page 2.)
- [Andersson *et al.* 2012] J. Andersson, J. Åkesson et M. Diehl. *CasADI : A Symbolic Package for Automatic Differentiation and Optimal Control*. In Recent Advances in Algorithmic Differentiation, volume 87 of *Lecture Notes in Computational Science and Engineering*, pages 297–307. Springer, Berlin, 2012. (Cité en page 90.)
- [Araya-Lopez 2013] M. Araya-Lopez. *Near-Optimal Algorithms for Sequential Information-Gathering Decision Problems*. PhD thesis, Université de Lorraine, 2013. (Cité en page 109.)
- [Argentieri *et al.* 2015] S. Argentieri, P. Danès et P. Souères. *A survey on sound source localization in robotics : From binaural to array processing methods*. Computer Speech and Language, vol. 34, no. 1, pages 87 – 112, 2015. (Cité en page 6.)
- [Bar-Shalom & Li 1998] Y. Bar-Shalom et X.-R. Li. *Estimation and tracking : Principles, techniques, and software*. Artech House, 1998. (Cité en page 29.)
- [Bar-Shalom *et al.* 2009] Y. Bar-Shalom, F. Daum et J. Huang. *The probabilistic data association filter*. IEEE Control Systems, vol. 29, no. 6, pages 82–100, Dec 2009. (Cité en page 11.)
- [Bartholomew-Biggs *et al.* 2000] M. Bartholomew-Biggs, S. Brown, B. Christianson et L. Dixon. *Automatic differentiation of algorithms*. Journal of Computational and Applied Mathematics, vol. 124, no. 1–2, pages 171 – 190, 2000. Numerical Analysis 2000. Vol. IV : Optimization and Nonlinear Equations. (Cité en page 89.)
- [Berglund & Sitte 2005] E. Berglund et J. Sitte. *Sound source localisation through active audition*. In 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 653–658, Aug 2005. (Cité en page 13.)
- [Bertsekas & Tsitsiklis 1996] D. Bertsekas et J. N. Tsitsiklis. *Neuro-dynamic programming*. Athena Scientific, 1996. (Cité en page 109.)

- [Bischof *et al.* 2002] C. H. Bischof, H. M. Bückner, B. Lang, A. Rasch et A. Vehreschild. *Combining Source Transformation and Operator Overloading Techniques to Compute Derivatives for MATLAB Programs*. In Proceedings of the Second IEEE International Workshop on Source Code Analysis and Manipulation (SCAM 2002), pages 65–72, Los Alamitos, CA, USA, 2002. (Cité en page 90.)
- [Blauert 2017] J. Blauert. *Reading the World with Two Ears*. In 2017 ICSV24 International Congress on Sound and Vibration, London, 2017. (Cité en page 12.)
- [Bourgault *et al.* 2002] F. Bourgault, A.A. Makarenko, S.B. Williams, B. Grocholsky et H.F. Durrant-Whyte. *Information based adaptive robotic exploration*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, (IROS'2002), Lausanne, Switzerland, 2002. (Cité en pages 41 et 43.)
- [Boyd *et al.* 1994] S. Boyd, L. El Ghaoui, E. Feron et V. Balakrishnan. *Linear matrix inequalities in system and control theory*. Society for Industrial and Applied Mathematics, 1994. (Cité en page 79.)
- [Brungart & Rabinowitz 1999] D. S. Brungart et W. M. Rabinowitz. *Auditory localization of nearby sources. Head-related transfer functions*. The Journal of the Acoustical Society of America, vol. 106, no. 3, pages 1465–1479, 1999. (Cité en page 67.)
- [Burgard *et al.* 1997] W. Burgard, D. Fox et S. Thrun. *Active mobile robot localization by entropy minimization*. In Proceedings Second EUROMICRO Workshop on Advanced Mobile Robots, pages 155–162, Oct 1997. (Cité en pages 39 et 43.)
- [Cadenat 1999] V. Cadenat. *Commande référencée multi-capteurs pour la navigation d'un robot mobile*. PhD thesis, Université Paul Sabatier - Toulouse III, 1999. (Cité en page 24.)
- [Cherubini & Chaumette 2012] A. Cherubini et F. Chaumette. *Visual navigation of a mobile robot with laser-based collision avoidance*. The International Journal of Robotics Research, vol. 32, no. 2, pages 189–205, Septembre 2012. (Cité en page 1.)
- [Cooke *et al.* 2007] M. Cooke, Y.C. Lu, Y. Lu et R. Horaud. *Active Hearing, Active Speaking*. In Int. Symp. on Auditory and Audiological Research (ISAAR'07), Marienlyst, Helsingør, Denmark, 2007. (Cité en pages 2 et 5.)
- [Cover & Thomas 1991] T. Cover et J. Thomas. *Elements of information theory*. Wiley, 1991. (Cité en pages 30 et 32.)
- [Dames *et al.* 2012] P. Dames, M. Schwager, V. Kumar et D. Rus. *A decentralized control policy for adaptive information gathering in hazardous environments*. In 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), pages 2807–2813, Dec 2012. (Cité en page 42.)
- [Denzler & Brown 2002] J. Denzler et C.M. Brown. *Information Theoretic Sensor Data Selection for Active Object Recognition and State Estimation*. IEEE

- Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 2, pages 145–157, 2002. (Cit  en page 38.)
- [Denzler *et al.* 2003] J. Denzler, M. Zobel et H. Niemann. *Information theoretic focal length selection for real-time active 3D object tracking*. In Ninth IEEE International Conference on Computer Vision, 2003. Proceedings, pages 400–407 vol.1, Octobre 2003. (Cit  en pages 35 et 39.)
- [Deutsch *et al.* 2004] B. Deutsch, M. Zobel, J. Denzler et H. Niemann. *Multi-step Entropy Based Sensor Control for Visual Object Tracking*. Pattern Recognition, pages 359–366, 2004. (Cit  en pages 44 et 86.)
- [Duda & Martens 1998] R. O. Duda et W. L. Martens. *Range dependence of the response of a spherical head model*. The Journal of the Acoustical Society of America, vol. 104, no. 5, pages 3048–3058, 1998. (Cit  en pages 8 et 67.)
- [Espiau *et al.* 1992] B. Espiau, F. Chaumette et P. Rives. *A new approach to visual servoing in robotics*. IEEE Transactions on Robotics and Automation, vol. 8, no. 3, pages 313–326, Jun 1992. (Cit  en page 1.)
- [Evers *et al.* 2016a] C. Evers, A. Moore et P. Naylor. *Towards Informative Path Planning for Acoustic SLAM*. In DAGA 2016, Aachen, Germany, 2016. (Cit  en page 108.)
- [Evers *et al.* 2016b] C. Evers, A. H. Moore et P. A. Naylor. *Acoustic simultaneous localization and mapping (A-SLAM) of a moving microphone array and its surrounding speakers*. In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 6–10, March 2016. (Cit  en page 108.)
- [Feder *et al.* 1999] H.J.S. Feder, J.J. Leonard et C.M. Smith. *Adaptive Mobile Robot Navigation and Mapping*. The International Journal of Robotics Research, vol. 18, no. 7, pages 650–668, 1999. (Cit  en page 41.)
- [Forster *et al.* 2014] Christian Forster, Matia Pizzoli et Davide Scaramuzza. *Appearance-based Active, Monocular, Dense Reconstruction for Micro Aerial Vehicles*. In Proceedings of Robotics : Science and Systems, Berkeley, USA, July 2014. (Cit  en page 44.)
- [Fox *et al.* 1998] D. Fox, W. Burgard et S. Thrun. *Active Markov localization for mobile robots*. Robotics and Autonomous Systems, vol. 25, no. 3, pages 195 – 207, 1998. (Cit  en page 40.)
- [Furukawa *et al.* 2013] K. Furukawa, K. Okutani, K. Nagira, T. Otsuka, K. Itoyama, K. Nakadai et H.G. Okuno. *Noise correlation matrix estimation for improving sound source localization by multirotor UAV*. In Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on, pages 3943–3948, Nov 2013. (Cit  en page 12.)
- [Grocholsky *et al.* 2003] B. Grocholsky, A. Makarenko et H. Durrant-Whyte. *Information-theoretic coordinated control of multiple sensor platforms*. In IEEE Int. Conf. on Robotics and Automation, (ICRA’03), Taipei, Taiwan, 2003. (Cit  en page 41.)

- [Grocholsky 2006] B. Grocholsky. *Information-theoretic control of multiple sensor platforms*. PhD thesis, University of Sydney, 2006. (Cité en page 42.)
- [Hammel *et al.* 1989] S. E. Hammel, P. T. Liu, E. J. Hilliard et K. F. Gong. *Optimal observer motion for localization with bearing measurements*. Computers & Mathematics with Applications, vol. 18, no. 1, pages 171–180, 1989. (Cité en page 38.)
- [Helferty & Mudgett 1993] J. P. Helferty et D. R. Mudgett. *Optimal observer trajectories for bearings only tracking by minimizing the trace of the Cramer-Rao lower bound*. In Decision and Control, 1993., Proceedings of the 32nd IEEE Conference on, pages 936–939 vol.1, Dec 1993. (Cité en page 37.)
- [Hiriart-Urruty 2008] J.B. Hiriart-Urruty. Les mathématiques du mieux faire. Editions ellipses, Paris, France., 2008. (Cité en page 117.)
- [Horwedel 1991] J. E. Horwedel. *GRESS, A Preprocessor for Sensitivity Studies of Fortran Programs*. In Andreas Griewank et George F. Corliss, éditeurs, Automatic Differentiation of Algorithms : Theory, Implementation, and Application, pages 243–250. SIAM, Philadelphia, PA, 1991. (Cité en page 90.)
- [Hu *et al.* 2011] J. Hu, C. Chan, C. Wang, M. Lee et C. Kuo. *Simultaneous Localization of a Mobile Robot and Multiple Sound Sources Using a Microphone Array*. Advanced Robotics, vol. 25, no. 1-2, pages 135–152, Janvier 2011. (Cité en page 108.)
- [Huber *et al.* 2008] M. F. Huber, T. Bailey, H. Durrant-Whyte et U. D. Hanebeck. *On entropy approximation for Gaussian mixture random vectors*. In 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, pages 181–188, Aug 2008. (Cité en page 102.)
- [Ince *et al.* 2009] G Ince, K Nakadai, T Rodemann, Y Hasegawa, H Tsujino et J Imura. *Ego noise suppression of a robot using template subtraction*. In Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on, pages 199–204, Octobre 2009. (Cité en page 12.)
- [Ingrand & Ghallab 2014] F. Ingrand et M. Ghallab. *Robotics and Artificial Intelligence : a Perspective on Deliberation Functions*. AI Communications, vol. 27, no. 1, pages 63–80, 2014. (Cité en page 2.)
- [Julian *et al.* 2011] B. J. Julian, M. Angermann, M. Schwager et D. Rus. *A scalable information theoretic approach to distributed robot coordination*. In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pages 5187–5194. IEEE, 2011. (Cité en page 42.)
- [Julian *et al.* 2013] B.J. Julian, S. Karaman et D. Rus. *On mutual information-based control of range sensing robots for mapping applications*. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 5156–5163, Novembre 2013. (Cité en pages 36, 41 et 43.)
- [Julier & Uhlmann 2004] S. J. Julier et J. K. Uhlmann. *Unscented filtering and nonlinear estimation*. Proceedings of the IEEE, vol. 92, no. 3, pages 401–422, Mar 2004. (Cité en page 111.)

- [Kay 1993] S. Kay. *Fundamentals of statistical signal processing : Estimation theory*. Prentice Hall PTR, New Jersey, 1993. (Cité en page 29.)
- [Kumon & Noda 2011] M. Kumon et Y. Noda. *Active soft pinnae for robots*. In IEEE/RSJ Intl. Conf. Intell. Robots and Systems (IROS'2011), pages 112–117, Sept 2011. (Cité en page 13.)
- [Kumon *et al.* 2003] M. Kumon, T. Sugawara, K. Miike, I. Mizumoto et Z. Iwai. *Adaptive audio servo for multirate robot systems*. In IEEE/RSJ Intl. Conf. Intell. Robots and Systems (IROS'2003), Octobre 2003. (Cité en page 13.)
- [Le Cadre & Laurent-Michel 1999] J.-P. Le Cadre et S. Laurent-Michel. *Optimizing the receiver maneuvers for bearings-only tracking*. *Automatica*, vol. 35, no. 4, pages 591–606, 1999. (Cité en page 38.)
- [Leung *et al.* 2006] C. Leung, S. Huang, N. Kwok et G. Dissanayake. *Planning under uncertainty using model predictive control for information gathering*. *Robotics and Autonomous Systems*, vol. 54, no. 11, pages 898–910, Novembre 2006. (Cité en pages 44 et 86.)
- [Lim & Duda 1994] C. Lim et R. O. Duda. *Estimating the azimuth and elevation of a sound source from the output of a cochlear model*. In *Signals, Systems and Computers*, 1994. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on, volume 1, pages 399–403. IEEE, 1994. (Cité en page 9.)
- [Liu 1988] P. T. Liu. *An optimum approach in target tracking with bearing measurements*. *Journal of Optimization Theory and Applications*, vol. 56, no. 2, pages 205–214, 1988. (Cité en page 38.)
- [Lu & Cooke 2010] Y.C. Lu et M. Cooke. *Motion strategies for binaural localisation of speech sources in azimuth and distance by artificial listeners*. *Speech Communication*, 2010. (Cité en page 10.)
- [Magassouba *et al.* 2015] A. Magassouba, N. Bertin et F. Chaumette. *Sound-based control with two microphones*. In *Intelligent Robots and Systems (IROS)*, 2015 IEEE/RSJ International Conference on, pages 5568–5573. IEEE, 2015. (Cité en page 13.)
- [Magassouba *et al.* 2016] A. Magassouba, N. Bertin et F. Chaumette. *First applications of sound-based control on a mobile robot equipped with two microphones*. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2557–2562, May 2016. (Cité en page 13.)
- [Magassouba 2016] A. Magassouba. *Aural servo : towards an alternative approach to sound localization for robot motion control*. PhD thesis, Université Rennes 1, 2016. (Cité en page 13.)
- [Mallet *et al.* 2010] A. Mallet, C Pasteur, M. Herrb, S. Lemaignan et F. Ingrand. *GenoM3 : Building middleware-independent robotic components*. In *IEEE Int. Conf. on Robotics and Automation, (ICRA'2010)*, Anchorage, Alaska, 2010. (Cité en page 25.)

- [Manyika 1993] J.M. Manyika. *An Information-Theoretic Approach to Data Fusion and Sensor Management*. PhD thesis, University of Oxford, 1993. (Cité en pages 35 et 41.)
- [Marković et al. 2013] I. Marković, A. Portello, P. Danès, I. Petrović et S. Argentiéri. *Active speaker localization with circular likelihoods and bootstrap filtering*. In IEEE/RSJ Intl. Conf. on Intell. Robots and Systems (IROS'2013), Tokyo, Japan, 2013. (Cité en pages 11 et 19.)
- [Martinson & Schultz 2009] E. Martinson et A. Schultz. *Discovery of sound sources by an autonomous mobile robot*. *Autonomous Robots*, vol. 27, pages 221–237, Octobre 2009. (Cité en page 108.)
- [Mihaylova et al. 2002] L. Mihaylova, T. Lefebvre, H. Bruyninckx, K. Gadeyne et J. De Schutter. *Active Sensing for Robotics - A Survey*. In in Proc. 5 th Int'l Conf. On Numerical Methods and Applications, pages 316–324, 2002. (Cité en pages 35 et 36.)
- [Nakadai et al. 2000] K. Nakadai, T. Lourens, H.G. Okuno et H. Kitano. *Active Audition for Humanoid*. In Nat. Conf. on Artificial Intelligence (AAAI'2000), Austin, TX, 2000. (Cité en page 10.)
- [Nguyen et al. 2016] Q. V. Nguyen, F. Colas, E. Vincent et F. Charpillet. *Localizing an intermittent and moving sound source using a mobile robot*. In 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 1986–1991, Oct 2016. (Cité en page 11.)
- [Nguyen et al. 2017] Q. V. Nguyen, F. Colas, E. Vincent et Charpillet F. *Long-Term Robot Motion Planning for Active Sound Source Localization with Monte Carlo Tree Search*. In Hands-free Speech Communication and Microphone Arrays (HSCMA), San Francisco, United States, Mars 2017. (Cité en pages 14, 20 et 44.)
- [Nocedal & Wright 2006] J. Nocedal et S.J. Wright. *Numerical optimization*. Springer series in operations research. Springer, New York, 2nd ed édition, 2006. OCLC : ocm68629100. (Cité en pages 74 et 75.)
- [Okuno & Nakadai 2015] H. G. Okuno et K. Nakadai. *Robot audition : Its rise and perspectives*. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'2015), pages 5610–5614, April 2015. (Cité en page 2.)
- [Okuno et al. 2004] H. G. Okuno, T. Ogata, K. Komatani et K. Nakadai. *Computational auditory scene analysis and its application to robot audition*. In International Conference on Informatics Research for Development of Knowledge Society Infrastructure, 2004. ICKS 2004., pages 73–80, March 2004. (Cité en page 2.)
- [Oshman & Davidson 1999] Y. Oshman et P. Davidson. *Optimization of Observer Trajectories for Bearings-Only Target Localization*. *IEEE Trans. on Aerospace and Electronic Systems*, vol. 35, no. 3, pages 892–902, 1999. (Cité en page 37.)

- [Ponzoni Carvalho Chanel 2013] C. Ponzoni Carvalho Chanel. *Planification de perception et de mission en environnement incertain : application à la détection et à la reconnaissance de cibles par un hélicoptère autonome*. PhD thesis, Institut Supérieur de l'Aéronautique et de l'Espace (ISAE), 2013. (Cité en page 109.)
- [Portello *et al.* 2011] A. Portello, P. Danès et S. Argentieri. *Acoustic models and Kalman filtering strategies for active binaural sound localization*. In IEEE/RSJ Intl. Conf. on Intell. Robots and Systems (IROS'2011), San Francisco, CA, 2011. (Cité en page 11.)
- [Portello *et al.* 2012] A. Portello, P. Danès et S. Argentieri. *Active Binaural Localization of Intermittent Moving Sources in the Presence of False Measurements*. In IEEE/RSJ Int. Conf. on Intell. Robots and Systems (IROS'2012), 2012. (Cité en page 11.)
- [Portello *et al.* 2013] A. Portello, P. Danès, S. Argentieri et S. Pleedel. *HRTF-Based Source Azimuth Estimation and Activity Detection from a Binaural Sensor*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'2013), Tokyo, Japan, 2013. (Cité en pages 10, 11, 12 et 18.)
- [Portello *et al.* 2014a] A. Portello, G. Bustamante, P. Danès et A. Mifsud. *Localization of Multiple Sources from a Binaural Head in a Known Noisy Environment*. In IEEE/RSJ Int. Conf. on Intell. Robots and Systems (IROS'2014), Chicago, IL, 2014. (Cité en pages 10 et 108.)
- [Portello *et al.* 2014b] A. Portello, G. Bustamante, P. Danès, J. Piat et J. Manhès. *Active Localization of an Intermittent Sound Source from a Moving Binaural Sensor*. In Forum Acustium (FA'2014), Krakow, Poland, 2014. (Cité en page 12.)
- [Portello 2013] A. Portello. *Localisation Binaurale Active de Sources Sonores en Robotique Humanoïde*. PhD thesis, Univ. Toulouse III Paul Sabatier, 2013. (Cité en pages 12, 18, 19, 38, 68 et 111.)
- [Raspaud *et al.* 2010] M. Raspaud, H. Viste et G. Evangelista. *Binaural Source Localization by Joint Estimation of ILD and ITD*. IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 1, pages 68–77, Janvier 2010. (Cité en page 9.)
- [Rayleigh 1907] L. Rayleigh. *On our perception of sound direction*. Philosophical Magazine Series, vol. 13, no. 74, pages 214–232, 1907. (Cité en page 6.)
- [Rosenthal & Okuno 1998] D. F. Rosenthal et H. G. Okuno. *Computational auditory scene analysis*. Lawrence Erlbaum Associates Publishers Computational auditory, Mahwah, NJ, US, 1998. (Cité en page 2.)
- [Ryan & Hedrick 2010] A. Ryan et J. K. Hedrick. *Particle filter based information-theoretic active sensing*. Robotics and Autonomous Systems, vol. 58, no. 5, pages 574–584, Mai 2010. (Cité en page 44.)
- [Särkkä *et al.* 2014] S. Särkkä, J. Hartikainen, L. Svensson et F. Sandblom. *Gaussian process quadratures in nonlinear sigma-point filtering and smoothing*.

- In 17th International Conference on Information Fusion (FUSION), pages 1–8, July 2014. (Cit  en page 87.)
- [Sasaki *et al.* 2010] Y. Sasaki, S. Thompson, M. Kaneyoshi et S. Kagami. *Map-generation and Identification of Multiple Sound Sources from Robot in Motion*. In IEEE/RSJ Intl. Conf. Intell. Robots and Systems (IROS’2010), pages 437–443, Taipei, Taiwan, 2010. (Cit  en page 108.)
- [Scardovi 2005] L. Scardovi. *Information Based Control for State and Parameter Estimation*. PhD thesis, Univ. of Genoa, 2005. (Cit  en page 109.)
- [Schult *et al.* 2015] N. Schult, T. Reineking, T. Kluss et C. Zetsche. *Information-Driven Active Audio-Visual Source Localization*. PLOS ONE, vol. 10, no. 9, pages 1–24, 09 2015. (Cit  en page 13.)
- [Schwager *et al.* 2011] M. Schwager, P. Dames, D. Rus et V. Kumar. *A multi-robot control policy for information gathering in the presence of unknown hazards*. In Proceedings of the International Symposium on Robotics Research (ISRR 11), 2011. (Cit  en page 42.)
- [Schymura *et al.* 2017] C. Schymura, J. Diego, R. Grajales et D. Kolossa. *Monte Carlo Exploration for Active Binaural Localization*. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP’2017), New Orleans, United States, Mars 2017. (Cit  en pages 13 et 103.)
- [Sekiguchi *et al.* 2015] K. Sekiguchi, Y. Bando, K. Itoyama et K. Yoshii. *Optimizing the layout of multiple mobile robots for cooperative sound source separation*. In Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, pages 5548–5554. IEEE, 2015. (Cit  en page 13.)
- [Shannon 1948] C. E. Shannon. *A Mathematical Theory of Communication*. The Bell System Technical Journal, vol. 27, pages 379–423, 1948. (Cit  en page 28.)
- [Shinn-Cunningham *et al.* 2000] B.G. Shinn-Cunningham, S. Santarelli et N. Kopco. *Tori of confusion : Binaural localization cues for sources within reach of a listener*. The Journal of the Acoustical Society of America, vol. 107, no. 3, pages 1627–1636, 2000. (Cit  en pages 9 et 10.)
- [Sommerlade & Reid 2008] E. Sommerlade et I. Reid. *Information-theoretic active scene exploration*. In IEEE Conf. on Computer Vision and Pattern Recognition, (CVPR’2008), Anchorage, Alaska, 2008. (Cit  en pages 35 et 39.)
- [Stump *et al.* 2009] E. Stump, V. Kumar, B. Grocholsky et P. M. Shiroma. *Control for Localization of Targets using Range-only Sensors*. The International Journal of Robotics Research, vol. 28, no. 6, pages 743–757, Juin 2009. (Cit  en page 42.)
- [Thrun *et al.* 2005] S. Thrun, W. Burgard et D. Fox. Probabilistic robotics. The MIT Press, 2005. (Cit  en pages 36, 40 et 43.)
- [Thurlow & Runge 1967] W R Thurlow et P S Runge. *Effect of induced head movements on localization of direction of sounds*. The Journal of the Acoustical Society of . . . , vol. 42, no. 2, pages 480–488, 1967. (Cit  en page 9.)

- [Tichavský *et al.* 1998] P. Tichavský, C.H. Muravchik et A. Nehorai. *Posterior Cramér-Rao Bounds for Discrete-Time Nonlinear Filtering*. IEEE. Trans. On Signal Processing, vol. 46, 1998. (Cité en page 30.)
- [Vincent *et al.* 2015] E. Vincent, A. Sini et F. Charpillat. *Audio source localization by optimal control of a mobile robot*. In IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'2015), Brisbane, Australia, 2015. (Cité en pages 14 et 20.)
- [Vlassis *et al.* 1999] N. Vlassis, Y. Motomura et B. Kröse. *An Information-Theoretic Localization Criterion for Robot Map Building*, 1999. (Cité en pages 20 et 40.)
- [Wallach 1938] H. Wallach. *On sound localization*. The Journal of the Acoustical Society of America, vol. 10, no. 1, page 83, 1938. (Cité en page 9.)
- [Weiss *et al.* 1987] LEEWEISS, ARTHUR SANDERSON et CHARLESP NEUMAN. *Dynamic sensor-based control of robots with visual feedback*. IEEE Journal on Robotics and Automation, vol. 3, no. 5, pages 404–417, 1987. (Cité en page 1.)
- [Wenhardt *et al.* 2006] S. Wenhardt, B. Deutsch, J. Hornegger, H. Niemann et J. Denzler. *An information theoretic approach for next best view planning in 3-d reconstruction*. In Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, volume 1, pages 103–106. IEEE, 2006. (Cité en pages 35 et 39.)
- [Youssef *et al.* 2012] K. Youssef, S. Argentieri et J.L. Zarader. *Towards a systematic study of binaural cues*. In Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, pages 1004–1009. IEEE, 2012. (Cité en page 9.)
- [Zahorik *et al.* 2005] P. Zahorik, D. S. Brungart et A. W. Bronkhorst. *Auditory distance perception in humans : A summary of past and present research*. Acta Acustica united with Acustica, vol. 91, no. 3, pages 409–420, 2005. (Cité en pages 7 et 68.)



---

**Résumé :** Ce travail s'inscrit dans le contexte de la localisation de source sonore depuis un capteur binaural (constitué de deux microphones placés sur un élément diffusant) doté de mobilité. Un schéma de localisation « active » en trois phases est considéré : (a) estimation de primitives spatiales par une analyse court-terme du flux audio ; (b) localisation audio-motrice par assimilation de ces données et combinaison avec les ordres moteurs du capteur au sein d'un schéma d'estimation stochastique ; (c) commande en boucle fermée du mouvement du capteur de façon à améliorer la qualité de la localisation. Les recherches portent sur la définition de stratégies de « mouvement actif » constituant la phase (c).

Le problème est formulé comme la maximisation d'un critère d'information défini à partir des lois de filtrage de la position relative capteur-source sur un horizon temporel glissant dans le futur (plus exactement de son espérance sur les observations qui seront assimilées sur cet horizon conditionnellement aux observations passées). Cet horizon peut être constitué du prochain instant ou des  $N$  prochains instants, ce qui donne lieu à une stratégie « one-step-ahead » ou «  $N$ -step-ahead », respectivement. Une approximation de ce critère par utilisation de la transformée « unscented » et le calcul automatique du gradient de celle-ci par exploitation des nombres duaux, permettent la détermination de la commande (en boucle fermée sur l'audio donc) à appliquer au capteur.

Les résultats ont été validés par des simulations réalistes, et, pour certains, par des expérimentations sur un ensemble tête-torse anthropomorphe doté de perception binaurale et de mobilité.

**Mots clés :** Audition en robotique, Localisation binaurale active, Théorie de l'information.

---