HAL
open science

Étude Statistique du Problème de la Trajectographie Passive

Benoit Landelle

To cite this version:

Benoit Landelle. Étude Statistique du Problème de la Trajectographie Passive. Mathématiques [math]. Université Paris Sud - Paris XI, 2009. Français. NNT : . tel-00386071v4

HAL Id: tel-00386071

https://theses.hal.science/tel-00386071v4

Submitted on 6 Jul 2009

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.

Numéro d'ordre : 9459

Université de Paris-Sud
U.F.R. Scientifique d'Orsay

THÈSE

présentée pour obtenir le grade de

DOCTEUR EN MATHÉMATIQUES
DE L'UNIVERSITÉ DE PARIS XI ORSAY

par

Benoît Landelle

Sujet :

ÉTUDE STATISTIQUE DU PROBLÈME DE LA TRAJECTOGRAPHIE PASSIVE

Soutenue le 30 juin 2009 après avis des rapporteurs :

Monsieur Fabien Campillo
Monsieur François Le Gland

devant la Commission d'examen composée de :

Monsieur Fabien Campillo	Rapporteur
Madame Élisabeth Gassiat	Directeur de Thèse
Monsieur François Le Gland	Rapporteur
Monsieur Pascal Massart	Président du Jury
Monsieur Éric Moulines	Examinateur
Monsieur Michel Prenat	Co-directeur de Thèse

À Aurélia, Maud, Victoire et Ferdinand

Remerciements

Je remercie chaleureusement mes directeurs de thèse, Élisabeth Gassiat et Michel Prenat, pour leur confiance et leur implication à mes côtés pendant ces quatre années de thèse. Je leur exprime toute ma reconnaissance pour leur disponibilité. Leur encadrement et leur manière d'appréhender chacun les mathématiques m'ont porté dans cette expérience passionnante de recherche. Je les remercie profondément pour l'aide qu'ils m'ont apportée tout au long de cette thèse.

Je suis très reconnaissant à Fabien Campillo et François Le Gland d'avoir accepté de rapporter cette thèse et de me faire l'honneur de participer au Jury. Je les remercie pour leur lecture attentive de ce document et les observations dont ils m'ont fait part qui ont contribué à améliorer ce travail.

Je remercie chaleureusement Éric Moulines pour la passionnante collaboration que nous avons eue et qui a joué un rôle décisif dans le bon déroulement de cette thèse. Ses multiples conseils, son dynamisme et sa bonne humeur ont été autant d'avancées dans l'élaboration de ce travail. Je suis très honoré de sa participation au Jury.

Je souhaite exprimer ma profonde gratitude à Pascal Massart qui m'a initié à la Statistique et m'a encouragé à tenter l'aventure de la recherche. Je le remercie de sa bienveillance et d'avoir accepté de présider le Jury.

J'ai reçu au sein de Thales Optronique un accueil très chaleureux et je souhaite remercier toutes celles et ceux qui m'ont accompagné dans ce périple : Laurence Gasnot, Claude Ramin, Gérard Berginc, Patrick Anstett, Pierre Nory, Gérard Gaillat, Michel Papuchon, Jean-Claude Fontanella, Christian Brisseau, Bruno Caneill, Alain Simon, Étienne Payot.

Je souhaite également exprimer ma gratitude à Emmanuel Trelat pour son éclairage en théorie du contrôle, Maurice Charbit, Sylvain Faure et Marc Lavielle pour leur aide sur les questions informatiques.

Je suis très reconnaissant à Valérie Lavigne pour sa gentillesse et sa disponibilité. Je la remercie grandement pour son aide précieuse en de multiples circonstances. J'adresse également mes remerciements à Pierre Pansu pour son engagement dans l'École Doctorale et tout ce qu'il fait pour chaque doctorant.

Je profite de cette occasion pour remercier Raphael Cerf qui me témoigna sa confiance et rendit possible mon retour aux mathématiques dans ce lieu incroyablement privilégié qu'est l'Université d'Orsay.

Ces quatre années de recherche ont été l'occasion de nombreuses rencontres et je souhaite remercier tous les doctorants qui ont fait de cette expérience une aventure humaine unique : Frédéric, Laurent T., Antoine L., Juliette, Aline, Adeline, Christine, Marianne, Amandine, Marie, Cathy, Camille, Bertrand Michel, Robin, Wilson, Nathalie, Cyril, Dominique, Pierre B., Ismaël C., Aurélien, Sophie D., Sylvain, Laurent M., Benoît S., Van Sang, Sébastien, Ramla et tous les autres.

Je souhaite remercier la famille et les amis qui sont toujours là quand tout va bien mais pas seulement : mes parents Jean-Marie et Geneviève, mes beaux-parents Jules et Marie-Lise, Laure, Rachel et Alain, les trois têtes brûlées Quentin, Marie et Paul, Hélène et Maurice, Daniel et Josianne, Pierre et Anne-Françoise, Y-lin et Pascal.

Enfin, je remercie mon épouse Aurélia pour sa patience et son soutien, toutes ces années durant. Je lui dédie ce travail ainsi qu'à nos trois rayons de soleil adorés Maud, Victoire et Ferdinand qui remplissent nos vies et nous comblent de bonheur.

Table des matières

I Introduction **3**

1 La trajectographie passive **5**
1.1 Présentation 5
1.2 Un problème inverse mal posé 6

2 L'observabilité **9**
2.1 Des dynamiques contraintes 9
2.2 Contributions 14

3 L'estimation **17**
3.1 Le phénomène de non-robustesse 17
3.2 Résultats obtenus dans cette thèse 19

4 Le filtrage **23**
4.1 Les Chaînes de Markov Cachées 23
4.2 Le Filtre Optimal 25
4.3 Principaux apports de cette thèse 27

II Observabilité **31**

5 Trajectoires paramétriques **33**
5.1 Introduction 33
5.2 Observabilité 34
5.3 Le mouvement rectiligne uniforme 41
5.4 Le mouvement circulaire uniforme 45
5.5 Conclusion 46
5.6 Preuves de la section 5.4 47

6 Trajectoires paramétriques par morceaux **51**
6.1 Introduction 51
6.2 Observabilité 51
6.3 Sections MRU et MCU 56
6.4 Extensions 59
6.5 Conclusion 60

7 Trajectoires régies par une équation différentielle **61**
7.1 Introduction 61
7.2 L'observabilité en théorie du contrôle 61

7.3 Observabilité à vitesse constante . . . 65
7.4 Vers un critère d'observabilité . . . 68
7.5 Perspectives . . . 71

III Estimation **73**

8 Trajectoires paramétriques **75**
8.1 Introduction . . . 75
8.2 Comportement asymptotique . . . 76
8.3 Test d'adéquation paramétrique . . . 79
8.4 Simulations . . . 80
8.5 Conclusion . . . 85
8.6 Annexe au test d'adéquation paramétrique . . . 86

9 Trajectoires paramétriques par morceaux **89**
9.1 Introduction . . . 89
9.2 Consistance . . . 89
9.3 Défaut d'observabilité . . . 91
9.4 Conclusion . . . 92

10 Robustesse en paramétrique **95**
10.1 Introduction . . . 96
10.2 Sensitivity to deterministic perturbation . . . 98
10.2.1 Perfect measurements . . . 98
10.2.2 Noisy measurements . . . 100
10.2.3 Continuous time problem . . . 101
10.2.4 Large sample results . . . 103
10.3 Sensitivity to stochastic perturbations . . . 104
10.3.1 Cyclostationary perturbation . . . 104
10.3.2 Stochastic perturbation on a parametric trajectory . . . 104
10.4 Simulations . . . 105
10.5 Conclusion . . . 110
10.6 Appendix . . . 110
10.7 Compléments . . . 113
10.8 Corrélations dans le cas sur-critique . . . 115
10.9 Un estimateur linéarisé . . . 116

11 Modèle de régression semiparamétrique avec bruit d'état **119**
11.1 Introduction . . . 120
11.2 Least squares estimation . . . 122
11.2.1 Consistency . . . 122
11.2.2 Asymptotic normality . . . 125
11.2.3 Application to BOT . . . 126
11.3 Likelihood and efficiency . . . 129
11.3.1 Application to BOT . . . 133
11.4 Further considerations . . . 133
11.5 Simulations . . . 135
11.6 Compléments . . . 142

IV Filtrage **145**

12 Filtrage de Kalman **147**
12.1 Introduction 147
12.2 Le modèle Linéaire Gaussien 147
12.3 Le filtre de Kalman étendu (EKF) 149
12.4 EKF avec bruit d'état 152
12.5 Conclusion 154

13 Stabilité asymptotique **155**
13.1 Introduction and notations 156
13.2 Main results 159
13.3 Nonlinear state-space models 166
13.3.1 Nonlinear state-space model with i.i.d. state noise 167
13.3.2 Nonlinear state-space model with dependent state noise 168
13.4 Proofs of Propositions 13.4 and 13.5 170
13.5 Proofs of Propositions 13.6 and 13.7 173

14 Filtre Particulaire **175**
14.1 Introduction 175
14.2 Stabilité asymptotique pour le BOT 176
14.3 Un pas vers le filtrage particulaire 180
14.4 Dynamique manœuvrante 181
14.5 Simulations 182
14.6 Perspectives 186

V Conclusion **187**

VI Annexe **191**

Présentation Générale

Cette thèse a pour objet l'étude statistique du problème de la trajectographie passive. Ce travail a été réalisé dans le cadre d'un contrat Cifre avec la société Thales Optronique[1] sur un sujet motivé par Michel Prenat, Ingénieur expert à Thales Optronique. Elle se décompose en 5 parties.

La première partie introduit le problème de la trajectographie passive en qualité de problème inverse mal posé. S'ensuit une présentation des trois thématiques centrales de ce travail : l'observabilité, l'estimation et le filtrage. Pour chacune, on expose un état de l'art suivi des contributions de cette thèse.

On s'intéresse, dans la deuxième partie, à la question de l'observabilité pour des trajectoires paramétriques puis paramétriques par morceaux et enfin, pour des trajectoires plus générales supposées à vitesse constante. On établit dans le cas paramétrique le lien qui unit la propriété d'observabilité à l'inversibilité de la matrice de Fisher. On présente des situations explicites d'observabilité et de non observabilité. Le cas particulier des trajectoires en mouvement rectiligne uniforme et en mouvement circulaire uniforme est considéré. Pour les trajectoires paramétriques par morceaux, des résultats généraux sont obtenus et sont appliqués à des trajectoires constituées de sections rectilignes et/ou circulaires. Dans le cas des trajectoires à vitesse constante, on présente les résultats existant en théorie du contrôle. Des expérimentations numériques suivent qui illustrent le caractère non observable des trajectoires à vitesse constante. Des critères additionnels sont envisagés pour garantir l'observabilité et sont expérimentés par simulations.

La troisième partie est consacrée à l'estimation. On décrit les propriétés asymptotiques de l'estimateur du maximum de vraisemblance pour des trajectoires paramétriques et paramétriques par morceaux. Le caractère non robuste de l'estimateur est décrit pour des scénarios réalistes où l'observabilité a lieu mais ne suffit pas pour fournir une estimation de qualité. On s'intéresse alors à la sensibilité de l'estimation quand le modèle d'état n'est pas totalement spécifié. On décrit le comportement de l'estimateur pour des perturbations d'état déterministes puis aléatoires. Dans ce deuxième cas, on établit, pour un bruit d'état isotrope, la consistance et la normalité asymptotique de l'estimateur des moindres carrés et on détermine une région de confiance quand la loi du bruit d'état est inconnue. On étudie également l'estimateur du maximum de vraisemblance et le caractère LAN du modèle statistique. Des pistes sont proposées pour traiter le cas d'un bruit d'état corrélé. Des simulations illustrent les différents résultats obtenus.

[1] Thales Optronique 2, avenue Gay-Lussac CS 90502 78995 - ELANCOURT CEDEX France

Dans la quatrième partie, on aborde la question de la trajectographie passive sous un point de vue différent des précédents en la considérant comme une chaîne de Markov cachée. On s'intéresse au problème du filtrage sous des aspects à la fois théoriques et algorithmiques. On rappelle le principe du filtre de Kalman et celui du filtre de Kalman étendu. Le caractère robuste du filtre étendu avec bruit d'état est illustré par des simulations numériques. Puis on établit pour des chaînes non ergodiques très générales la propriété de stabilité asymptotique du filtre optimal, propriété étroitement liée à la convergence des algorithmes particulaires. L'application au cas de la trajectographie passive est détaillée et des expérimentations numériques sont présentées pour des scénarios robustes et non robustes.

La cinquième et dernière partie propose une synthèse des résultats obtenus ainsi que des perspectives et des questions ouvertes pour des travaux futurs.

Première partie

Introduction

Chapitre 1

La trajectographie passive

1.1 Présentation

La mathématique se nourrit depuis toujours de problématiques issues du monde réel : phénomènes physiques, biologiques, économiques ou encore applications industrielles en aéronautique, énergétique, etc... La trajectographie passive est un de ces problèmes à la croisée de plusieurs disciplines : l'automatique, le traitement du signal et les mathématiques. Cette caractéristique en fait un sujet passionnant et difficile où les questions d'ordre mathématique ne se livrent pas d'emblée, leur formulation n'étant accessible qu'au prix d'une incontournable et nécessaire maturation.

La trajectographie passive consiste en la détermination de la trajectoire d'un mobile *cible* observé par un mobile *porteur*. Si les deux mobiles évoluent dans un même plan, les observations sont des mesures angulaires entre un axe fixe (le Nord magnétique) et l'axe de visée déterminé par la position du porteur et celle de la cible.

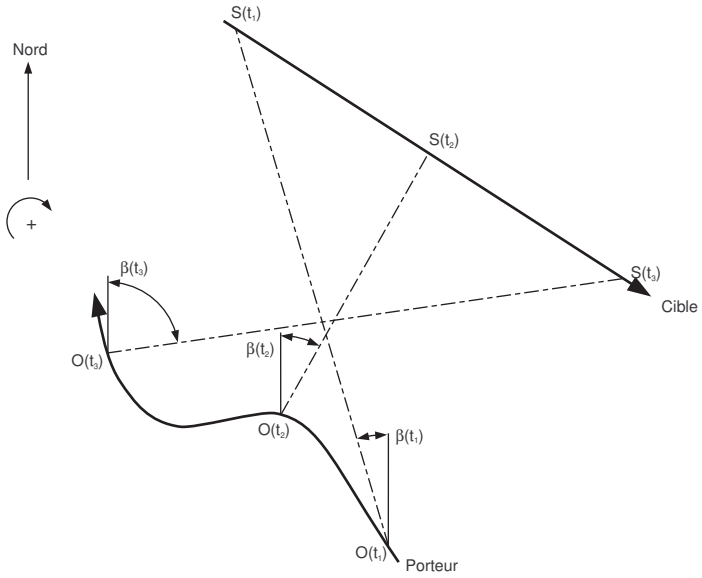


Fig. 1.1 – Mesures angulaires en trajectographie passive

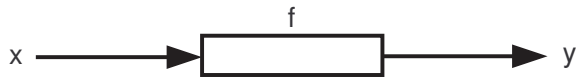
Dans sa version spatiale, les observations sont constituées de deux mesures d'angles qui correspondent, pour un repère de référence donné, aux deux composantes angulaires des coor-

données sphériques. Le trajectographie passive apparaît dans un contexte tactique présent dans différentes situations opérationnelles : un sous-marin porteur acquiert les mesures acoustiques des bruits émis par un sous-marin cible, un avion porteur réalise une suite de mesures optiques de la position d'un avion cible, etc ... Pour tous ces scénarios, l'observation réalisée est de nature *passive*, *id est* le porteur n'émet pas de signal pour acquérir l'observation. L'intérêt d'une mesure passive est évident : l'observateur reste discret vis-à-vis de la cible. Mais cette discrétion a un prix : l'observation est une mesure incomplète de la position de la cible, puisque la distance entre la cible et le porteur est inconnue.

1.2 Un problème inverse mal posé

Pourquoi et comment ce sujet séduit et résiste aux scientifiques qui s'y essayent depuis désormais une quarantaine d'année ? La raison est simple : la trajectographie passive appartient à la famille des *problèmes inverses mal posés*. La notion de problème *bien posé* a été initialement formulée par Hadamard [58] puis remaniée par Kirsch [72] dans sa forme actuelle. Soit f un opérateur qui modélise un appareil de mesure ou un périphérique électronique ou toute autre « boîte noire » qui reçoit une entrée et renvoie une sortie, x la quantité d'intérêt et y l'observation de x par le biais du périphérique, c'est-à-dire

$$y = f(x) \ . \tag{1}$$



Le problème inverse lié à l'équation (1), f étant connue, consiste, pour une observation y donnée, à déterminer l'entrée x correspondante.

Définition 1.1 (Problème bien posé)**.** *Soit* X *et* Y *des espaces vectoriels normés et* $f : \mathsf{X} \to \mathsf{Y}$ *une application continue (non nécessairement linéaire). Le problème* $y = f(x)$ *est bien posé au sens d'Hadamard*[1] *si les trois conditions suivantes sont réalisées :*

1. *pour tout* $y \in \mathsf{Y}$*, le problème* $y = f(x)$ *admet une solution ;*
2. *pour tout* $y \in \mathsf{Y}$*, il existe au plus une solution* $x \in \mathsf{X}$ *au problème* $y = f(x)$ *;*
3. *pour toute suite* $\{x_n\}_{n\geq 0} \in \mathsf{X}^{\mathbb{N}}$ *telle que* $\lim_{n\to\infty} f(x_n) = f(x)$*, alors* $\lim_{n\to\infty} x_n = x$*.*

Si l'une de ces conditions n'est pas satisfaite, le problème est dit mal-posé.

Dans toute la suite de ce travail, on se restreint désormais au cas de la trajectographie passive plane. Cette hypothèse permet une formulation mathématique simple donc plus maniable que la version spatiale sans pour autant dénaturer la question, l'essentiel de sa complexité étant préservée dans le cas plan.

Voyons en quoi la trajectographie passive s'inscrit dans la catégorie des problèmes inverses mal posés. On considère deux mobiles se déplaçant dans un même plan. Le porteur, encore appelé *plate-forme* ou *observateur*, suit une trajectoire connue[2]. Il observe la cible, aussi appelée

[1] ou plus simplement *bien posé*.
[2] C'est le cas en pratique grâce à un équipement de type centrale inertielle.

source, dont la trajectoire est inconnue. Soit $[0, T]$ l'intervalle de temps durant lequel le porteur effectue des mesures. Introduisons la fonction angle définie, pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, par

$$\mathrm{angle}(x, y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \mathrm{sgn}(x) \times \mathbb{1}_{y<0} & \text{si } x \neq 0 \text{ et } y \neq 0, \\ \frac{\pi}{2} \times \mathrm{sgn}(x) & \text{si } x \neq 0 \text{ et } y = 0, \\ \frac{\pi}{2} \times [1 - \mathrm{sgn}(y)] & \text{si } x = 0 \text{ et } y \neq 0, \end{cases}$$

où sgn représente la fonction signe. Pour un couple (x, y) de réels non simultanément nuls, la quantité $\mathrm{angle}(x, y)$ représente dans le plan l'angle à valeur dans $]-\pi, \pi]$ entre l'axe (Oy) et l'axe (OM) où M est le point de coordonnées (x, y).

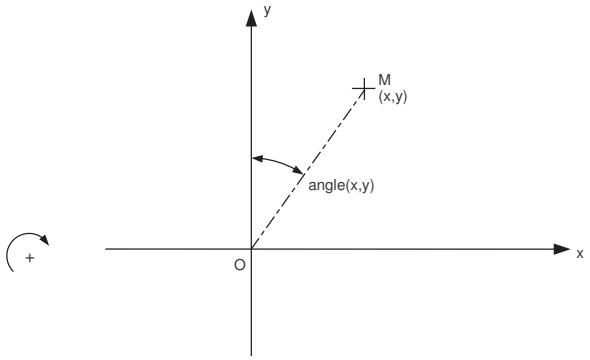


Fig. 1.2 – Fonction angle

À un instant t fixé, on note $O(t) = (O_1(t), O_2(t))$ la position du porteur[3] et $S(t) = (S_1(t), S_2(t))$ la position de la cible[4]. Soit Ψ la fonction définie pour tout $t \in [0, T]$ et pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ par

$$\Psi[(x_1, x_2), t] \stackrel{\text{def}}{=} \begin{cases} \mathrm{angle}[x_1 - O_1(t), x_2 - O_2(t)] & \text{si } (x, y) \neq O(t) \ , \\ 0 & \text{sinon} \ . \end{cases} \tag{2}$$

La mesure sans bruit collectée par le porteur à l'instant t est donnée par $\Psi[S(t), t]$. Introduisons alors la famille de trajectoires $\{S_\lambda\}_{\lambda>0}$ définies, pour tout $t \in [0, T]$, par

$$S_\lambda : t \longmapsto O(t) + \lambda(S - O)(t) \ .$$

Clairement, on a, pour tout réel $\lambda > 0$,

$$\forall t \in [0, T] \qquad \Psi[S_\lambda(t), t] = \Psi[S(t), t] \ .$$

[3]On choisit la lettre O comme initiale d'*observateur* en français et *observer* en anglais.
[4]On choisit la lettre S comme initiale de *source* en français et en anglais.

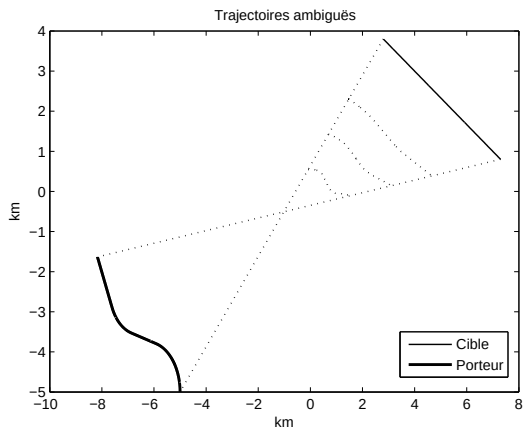


FIG. 1.3 – Trajectoires ambiguës indistingables de la cible

Autrement dit, toute trajectoire S_λ avec $\lambda > 0$ est conforme, sur le plan des mesures angulaires, avec la trajectoire cible S. Ainsi, notant Φ l'application qui à une trajectoire $S = \{S(t)\}_{t\in[0,T]}$ associe l'ensemble de ses mesures angulaires reçues par le porteur, c'est-à-dire

$$\Phi : S \longmapsto \{\Psi[S(t), t]\}_{t\in[0,T]} \ ,$$

le problème inverse associé à l'équation

$$Y = \Phi(S) \tag{3}$$

est mal posé puisque l'unicité requise par la deuxième condition de la définition 1.1 n'est pas satisfaite. On pourrait également montrer que la troisième condition n'est pas non plus vérifiée. Pour les problèmes linéaires mal posés, il existe des stratégies de résolution, la plus connues étant la technique de *régularisation de Tikhonov* [104], [105]. Mais un fossé sépare les problèmes linéaires de ceux non-linéaires et les techniques ne se transposent pas d'un contexte à l'autre.

Examinons ce qui rend le problème de la trajectographie passive difficile. L'expression (2) définissant Ψ est peu maniable. Dans un souci de simplification et sans pour autant altérer la complexité du problème, on fera dans toute la suite de ce travail, excepté[5] au chapitre 14, l'hypothèse que les trajectoires S et O sont telles que

(H_+) $\forall t \in [0, T] \qquad S_2(t) - O_2(t) > 0$

Sous cette hypothèse, la mesure d'angle à l'instant t est égale à

$$\Psi[S(t), t] = \arctan[S_1(t) - O_1(t)]/[S_2(t) - O_2(t)] \ .$$

Cette dernière expression permet de comprendre d'où vient le caractère mal posé. La fonction $\arctan$ réalise une bijection de $\mathbb{R}$ dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$ donc, la connaissance de $\Psi[S(t), t]$ équivaut à celle du quotient $[S_1(t) - O_1(t)]/[S_2(t) - O_2(t)]$. Résoudre le problème inverse associé à (3) revient à déterminer le numérateur et le dénominateur à partir de ce quotient, ce qui illustre sans détour la complexité de la trajectographie passive.

[5] On le rappellera opportunément.

Chapitre 2

L'observabilité

L'approche la plus naturelle pour répondre au problème mal posé de la trajectographie passive consiste à déterminer si l'on peut, par des restrictions sur l'ensemble des trajectoires de la cible, rendre le problème inverse bien posé. Deux questions s'imposent : quelles restrictions imposer et celles-ci étant choisies, permettent-elle effectivement de déterminer de manière unique une trajectoire à partir de ses mesures angulaires, *id est* la trajectoire cible devient-elle *observable* ?

L'observabilité n'est pas une question purement théorique. Les mesures considérées sont des mesures sans bruit, hypothèse non réaliste dans une situation opérationnelle. Mais cette étude est le préalable indispensable à l'estimation car le cadre asymptotique que l'on étudie permet, d'une certaine manière, de s'affranchir de l'aléa. Pour cette raison, la première partie de cette thèse est consacrée à l'observabilité.

Dans les chapitres 5, 6 et 7, on va considérer pour la cible des dynamiques que l'on pourrait qualifier de *rigides* et identifier des scénarios qui garantissent ou pas l'observabilité. On verra par la suite une approche différente où la dynamique est relâchée en autorisant des fluctuations aléatoires sur la trajectoire cible dans un cadre semiparamétrique puis dans le contexte des *chaînes de Markov cachées*.

2.1 Des dynamiques contraintes

La réponse à la question du choix des restrictions sur la trajectoire cible est en partie fournie par le caractère opérationnel du problème. Un avion ou un sous-marin se déplace essentiellement en ligne droite, à vitesse constante. C'est le mouvement rectiligne uniforme (MRU en abrégé) qui est une des dynamiques les plus étudiées dans la littérature sur le sujet. Le MRU s'inscrit dans le cadre plus général des modèles de trajectoires paramétriques, ceux où la trajectoire est complètement déterminée par la donnée d'un paramètre de dimension finie. Un modèle paramétrique est la donnée d'un ensemble Θ de paramètres admissibles, ensemble compact[1] de $\mathbb{R}^d$, et d'une application $\theta \mapsto \{S_\theta(t)\}_{t\in[0,1]}$. Étant donné un modèle paramétrique, on note, pour tout $t \in [0, T]$,

$$m(\theta, t) \stackrel{\text{def}}{=} \Psi[S_\theta(t), t] \ .$$

Pour une dynamique MRU, le paramètre de la trajectoire est $\theta = (x_0, y_0, v_x, v_y)$ où (x_0, y_0) désigne la position initiale du mobile et (v_x, v_y) les coordonnées de son vecteur vitesse. Pour tout $t \in [0, T]$, on a

$$S_\theta(t) = (x_0 + v_x t, y_0 + v_y t) \ .$$

[1] C'est le cas concrètement puisque les paramètres correspondent à des caractéristiques techniques du mobile cible comme la vitesse, l'accélération de giration, ... grandeurs qui sont bornées.

Par la suite, pour simplifier certaines écritures, on supposera que $T = 1$, ce qui ne change en rien la nature du problème.

Supposer une dynamique MRU ne suffit pas à rendre la trajectoire de la cible *observable*, c'est-à-dire complètement caractérisée par les données angulaires. Plus précisément, la cible de trajectoire S_{θ^*} est dite observable si, pour tout $\theta \in \Theta$,

$$m(\theta, \cdot) = m(\theta^*, \cdot) \Longrightarrow \theta = \theta^* \ .$$

Pour garantir l'observabilité d'un MRU, il faut que le porteur suive un certain type de trajectoire et notamment, qu'il ne soit pas lui-même en MRU. On voit donc apparaître une contrainte de *commande*, autrement dit le porteur doit agir pour rendre sa cible observable.

Le premier résultat théorique connu pour l'observabilité d'un MRU est celui de Aidala et al.[3]. Notant $\beta(t) = \Psi[S_{\theta^*}(t), t]$, on montre l'équivalence

$$\begin{aligned} & S_{\theta^*} \text{ observable} \\ & \qquad \Longleftrightarrow \ 2\dot{\beta}\,\dddot{\beta} - 3\ddot{\beta}^2 + 4\dot{\beta}^4 \neq 0 \ , \\ & \qquad \Longleftrightarrow \ \Big(\forall t \in [0,1] \quad \beta(t) \neq \arctan[\lambda_1 + \lambda_3 t]/[\lambda_2 + \lambda_4 t]\Big) \ , \end{aligned} \tag{1}$$

avec $(\lambda_i)_{1 \leq i \leq 4} \in \mathbb{R}^4$ tel que $\lambda_2 + \lambda_4 t > 0$ pour tout $t \in [0,1]$. C'est surtout la deuxième équivalence qui est utile puisqu'elle permet une discussion sur le choix de la trajectoire porteur.

On peut formuler des critères d'observabilité plus restreints pour des modèles paramétriques quelconques. Présentons en premier lieu les deux modèles statistiques d'observations bruitées que l'on considérera par la suite, le premier à temps discret décrit par l'équation

$$Y_k = m(\theta, t_k) + \sigma \varepsilon_k \ , \tag{2}$$

où $\{t_k\}_{k \geq 1}$ est une suite à valeurs dans $[0,1]$, σ un réel strictement positif et $\{\varepsilon_k\}_{k \geq 1}$ une suite de variables gaussiennes standards, et l'autre à temps continu régi par l'équation différentielle stochastique

$$\mathrm{d}Y_t = m(\theta, t)\mathrm{d}t + \varepsilon \mathrm{d}W_t \ , \tag{3}$$

où ε est un réel strictement positif et $(W_t)_{t \in [0,1]}$ est un processus de Wiener. Une grandeur fondamentale liée à ces modèles est la *matrice d'information de Fisher* définie, pour le modèle d'observations (2), par

$$I_n(\theta) \stackrel{\text{def}}{=} \sigma^{-2} \sum_{k=1}^{n} \nabla_\theta m(\theta, t_k) \nabla_\theta m(\theta, t_k)^{\mathrm{T}} \ , \tag{4}$$

où n désigne le nombre d'observations et, pour le modèle d'observations (3), par

$$I(\theta) \stackrel{\text{def}}{=} \int_0^1 \nabla_\theta m(\theta, t) \nabla_\theta m(\theta, t)^{\mathrm{T}} \, \mathrm{d}t \ . \tag{5}$$

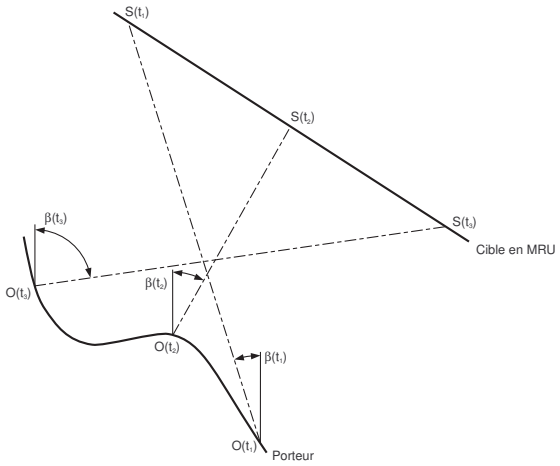


Fig. 2.1 – Cible en MRU

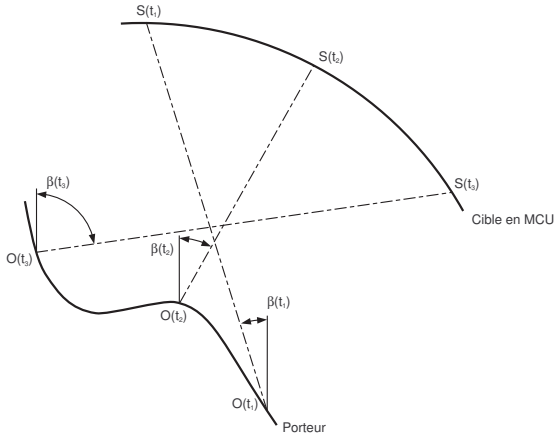


Fig. 2.2 – Cible en MCU

Jauffret [63] montre que l'inversibilité de la matrice de Fisher implique l'*observabilité locale* d'une trajectoire paramétrique. La trajectoire S_{θ^*} est dite localement observable s'il existe un voisinage V_{θ^*} de θ^* tel que, pour tout $\theta \in V_{\theta^*}$,

$$m(\theta, \cdot) = m(\theta^*, \cdot) \Longrightarrow \theta = \theta^* \ .$$

Contrairement au MRU où l'on dispose d'un résultat d'observabilité globale, le caractère local est le prix à payer pour un modèle paramétrique non spécifié. Cependant, ce lien entre observabilité et inversibilité pour la matrice de Fisher sert d'amorce à la fois pour comprendre des cas d'observabilité mais également des cas de non-observabilité. D'autres travaux conséquents sont consacrés à cette notion d'observabilité en trajectographie [51], [65], [102]. Les dynamiques considérées dans ces articles classent les modèles de trajectoire dans la catégorie paramétrique, même si la présentation des résultats ne met pas l'accent sur cet aspect paramétrique. Ceci est lié au choix du formalisme retenu, les auteurs privilégiant une présentation héritée de la théorie du contrôle. Les résultats obtenus identifient des situations d'observabilité pour des dynamiques plus générales que le MRU sous condition de commande du porteur, point essentiel en trajectographie passive.

L'extension naturelle aux modèles paramétriques est la catégorie des trajectoires *paramétriques par morceaux*, trajectoires qui sont des successions de trajectoires paramétriques élémentaires.

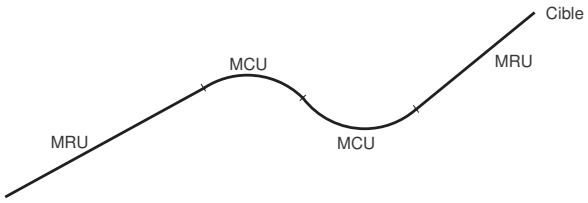


Fig. 2.3 – Trajectoire paramétrique par morceaux

On ne trouve que peu de travaux sur de tels modèles qui sont pourtant des modèles réalistes de trajectoires cibles. À titre d'exemples, mentionnons les trajectoires dites en *baïonnette* ou encore en *hippodrome* qui sont utilisées en situation opérationnelle.

Fig. 2.4 – Trajectoire en baïonnette

Fig. 2.5 – Trajectoire en hippodrome

Dans [106] et [64], les auteurs abordent le cas de trajectoires MRU par morceaux mais les instants de changement sont supposés connus. L'observabilité des trajectoires paramétriques par morceaux avec des instants de rupture inconnus est, à notre connaissance, une thématique quasiment inexistante de la littérature dédiée à la trajectographie passive. Plusieurs raisons expliquent sans doute cet insuccès. La détection des ruptures doit être faisable sous des hypothèses opérationnelles plausibles, autrement dit pour des trajectoires lisses (sans discontinuité des dérivées) et pas seulement en temps continu mais également pour des observations à temps discret. Ensuite, les algorithmes d'optimisation utilisés pour la détermination des paramètres de la trajectoire par morceaux s'avèrent insuffisamment performants et rendent l'estimation très difficile pour des observations bruitées. Mais surtout, les algorithmes particulaires connaissent un succès spectaculaire et mobilisent aujourd'hui la plupart des chercheurs travaillant sur la trajectoire passive, les autres voies étant *de facto* délaissées, du moins temporairement.

Dans le chapitre qui clôt l'étude d'observabilité, on s'intéresse à des trajectoires régies par une équation différentielle et plus précisément aux trajectoires à vitesse constante. Là encore, c'est le contexte opérationnel qui oriente le choix des dynamiques supposées. En situation, un avion ou un sous-marin, même s'il manœuvre, maintient dans la plupart des cas une vitesse constante. Ce modèle est très différent des modèles paramétriques ou paramétriques par morceaux. Pour ceux-ci, leur formulation est *quantitative* au sens où la donnée du paramètre ou de l'ensemble des paramètres fournit toute l'information pour reconstruire la trajectoire. Pour une trajectoire à vitesse constante, la description du modèle est *qualitative*, contenue dans l'équation différentielle qui régit la dynamique de la cible sans que celle-ci, même avec des conditions initiales, caractérise la trajectoire de la cible[2]. Il faut simultanément l'équation de la dynamique d'état et l'équation d'observation pour envisager de reconstruire la trajectoire cible. Présentons la formulation du problème de trajectographie passive dans le langage de la *théorie du contrôle* qui se prête à traduire la condition de vitesse constante. Pour une trajectoire $t \mapsto (x(t), y(t))$ telle que

$$(\mathrm{H}_+) \qquad \forall t \in [0,1] \qquad y(t) - O_2(t) > 0\,,$$

[2]Il y a une infinité de trajectoires à vitesse constante donnée partant d'un point initial donné.

on note $X(t)$ le vecteur d'état à l'instant t défini par

$$X(t) \stackrel{\text{def}}{=} \begin{pmatrix} x(t) & y(t) & \dot{x}(t) & \dot{y}(t) \end{pmatrix}^{\mathrm{T}} ,$$

et $Y(t)$ l'observation définie par

$$Y(t) \stackrel{\text{def}}{=} h[X(t), t] = \Psi[(x(t), y(t)), t] = \arctan[x(t) - O_1(t)]/[y(t) - O_2(t)] .$$

Le fait que la trajectoire soit à vitesse constante inconnue se traduit, pour tout $t \in [0, 1]$, par

$$\frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}^2 + \dot{y}^2) = 0 \iff \langle (\dot{x}, \dot{y}), (\ddot{x}, \ddot{y}) \rangle = 0 .$$

Le problème de trajectographie passive se modélise alors sous la forme d'un système différentiel partiellement observé décrit, pour tout $t \in [0, 1]$, par

$$\begin{cases} \dot{X}(t) = A(t)\, X(t) , \\ Y(t) = h[X(t), t] , \end{cases} \tag{6}$$

avec $A(t)$ la matrice définie par

$$A(t) \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\varrho(t) \\ 0 & 0 & \varrho(t) & 0 \end{pmatrix} ,$$

où $\varrho(t) = \dot{s}^{3/2}(t) c^{1/2}(t)$, s désignant l'abscisse curviligne de la trajectoire et c sa courbure. On définit la *résolvante* du système (6) notée $t \mapsto M(t)$ comme solution de

$$\begin{cases} \dot{M}(t) = A(t)\, M(t) , \\ M(0) = I_4 . \end{cases} \tag{7}$$

On dispose alors d'un critère d'observabilité locale pour le système (6), voir [103, chapter 6, theorem 26].

Proposition 2.1. *Le système* (6) *est localement observable en X si et seulement si la matrice*

$$G \stackrel{\text{def}}{=} \int_0^1 M^{-1^T}(t) \nabla_X h[X(t), t] \nabla_X h[X(t), t]^{\mathrm{T}} M^{-1}(t)\, \mathrm{d}t \tag{8}$$

est inversible.

Cependant, ce critère est inefficace pour une dynamique simple comme le MCU, conséquence de la description incomplète de l'équation différentielle $\dot{X}(t) = A(t)X(t)$ pour la caractéristique de vitesse constante. Une autre approche palliant aux insuffisances du modèle (6) nous entraîne vers un spécialité récente de la théorie contrôle qu'est la *platitude différentielle* qui ne sera qu'effleurée dans cette thèse, du fait de sa complexité et du fait de l'orientation privilégiée dans ce travail tournée vers l'inférence statistique et vers la modélisation stochastique.

2.2 Contributions

Dans le chapitre 5, on s'intéresse à des modèles paramétriques décrit par Θ un compact de $\mathbb{R}^{2p}$ et où la trajectoire S_θ est définie par

$$S_\theta = \left(\sum_{j=1}^{p} a_j e_j, \sum_{j=1}^{p} b_j e_j \right) \quad \text{avec} \quad \theta = (a_{1:p}, b_{1:p}) \in \Theta \ ,$$

où $(e_j)_{1 \leq j \leq p}$ est une famille libre de fonctions vérifiant la condition :

$$(\mathrm{H}_+) \qquad \forall (\theta, t) \in \Theta \times [0,1] \qquad S_{\theta,2}(t) - O_2(t) > 0.$$

Rappelons que cette hypothèse n'est pas réductrice de la difficulté du problème inverse mais permet de travailler avec une expression simple pour $\Psi[S_\theta(t), t]$. Pour cette catégorie de modèles, on démontre l'équivalence suivante :

Proposition 2.2.

$$S_{\theta^*} \textit{ observable} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta^*, \cdot) \right)_{1 \leq i \leq 2p} \textit{ libre} \iff I(\theta^*) \textit{ inversible} \ .$$

Ce résultat permet notamment de retrouver le critère de Aidala [3] pour une cible MRU et plus généralement pour toute cible à composantes polynomiales [51]. On en déduit également une classe de trajectoire porteur qui interdit l'observabilité de la cible.

Proposition 2.3. *Aucune des trajectoires porteur s'écrivant*

$$O = \left(\sum_{j=1}^{p} \alpha_j e_j, \sum_{j=1}^{p} \beta_j e_j \right) \quad \textit{avec} \quad (\alpha_j, \beta_j)_{1 \leq j \leq p} \in \mathbb{R}^{2p} \ ,$$

et vérifiant la condition (H_+) ne rend une trajectoire S_{θ^*} avec $\theta^* \in \Theta$ observable.

Ce résultat permet une interprétation intuitive de la notion d'observabilité : d'une certaine façon, pour que la cible devienne observable au porteur, celui-ci doit effectuer une trajectoire plus « complexe » que celle de la cible, ce qui lui permet d'extraire l'information des mesures angulaires.

Les dynamiques essentielles que sont le MRU et le mouvement circulaire uniforme (MCU) sont traitées en détail. Une trajectoire MCU est décrite par

$$S_\theta(t) = (x_0 + r\cos(\omega t + \alpha),\ y_0 + r\sin(\omega t + \alpha)) \ ,$$

avec $\theta = (x_0, y_0, r, \omega, \alpha) \in \Theta$ où Θ désigne un compact de $\mathbb{R}^5$. On s'intéresse au choix de la trajectoire porteur qui garantit l'observabilité de la cible. On obtient des résultats positifs d'observabilité[3] pour des dynamiques cibles en MRU et en MCU avec des trajectoires porteurs réalistes.

Proposition 2.4. – *Quand le porteur n'est pas immobile et suit une trajectoire MCU, le modèle statistique d'une cible dont la trajectoire est à composantes polynomiales vérifiant (H_+) est identifiable.*

[3]En fait, on a mieux puisque l'on obtient l'identifiabilité du modèle statistique associé.

– *Quand le porteur n'est pas immobile et suit une trajectoire à composantes polynomiales, le modèle statistique d'une cible en MCU est identifiable.*
– *Quand le porteur suit une trajectoire MCU, le modèle statistique d'une cible en MCU à vitesse angulaire différente de celle du porteur est identifiable.*

Bien que le modèle des trajectoires paramétriques par morceaux soit naturel, sa formulation rigoureuse s'avère lourde et complexe. Une des principales difficultés consiste à détecter les *instants de rupture* de la trajectoire cible, autrement dit les instants où la cible change de régime en passant d'un modèle paramétrique à un autre. Pour y parvenir, on suppose que ces instants sont le lieu d'une discontinuité pour la dérivée d'un certain ordre dans les composantes de la trajectoire. Cette hypothèse assez forte rend possible, pour des observations à temps continu, la détermination des instants de rupture. Une fois cette étape réalisée, la caractérisation des différentes séquences paramétriques s'obtient par application des résultats du premier chapitre.

Pour des trajectoires à vitesse constante, on propose des éléments de réponse à la question de l'observabilité. On montre par une méthode expérimentale très probante qu'une contrainte de vitesse constante ne suffit pas à caractériser une trajectoire avec des mesures angulaires connues. En utilisant l'algorithme d'Euler, on construit d'autres trajectoires que celle de la cible qui sont à vitesse constante et conformes à la cible au niveau des mesures angulaires. La figure 2.6 présente de telles trajectoires.

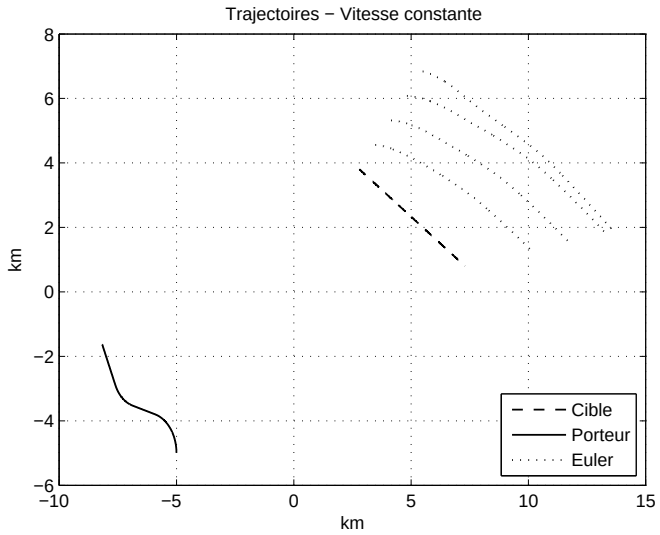


FIG. 2.6 – Trajectoires ambiguës à vitesse constante

$y(t_0)$ km	v m/s
4.56	378
5.32	432
6.08	486
6.84	486

FIG. 2.7 – Conditions initiales

Ces expérimentations permettent une réponse empirique négative à l'observabilité sous l'hypothèse de vitesse constante.

Des conditions additionnelles sont envisagées et illustrées là encore par des expérimentations, pour distinguer la trajectoire de la cible des autres trajectoires ambiguës à vitesse constante. Les orientations des vecteurs vitesses de la cible et du porteur ne présentent *a priori* pas de simultanéités de comportement. Au contraire, s'il y a des changements de régime simultanés, il y a tout lieu de croire que la trajectoire étudiée est ambiguë à la cible.

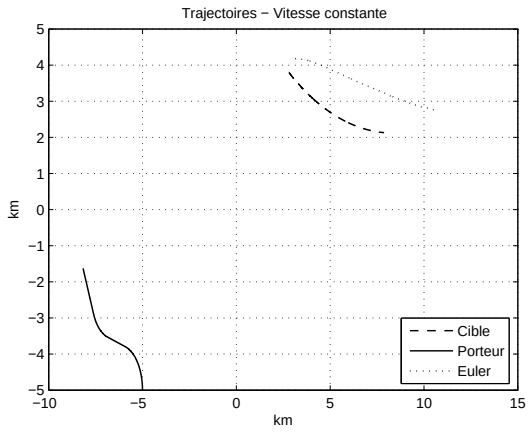


FIG. 2.8 – Trajectoire à vitesse constante

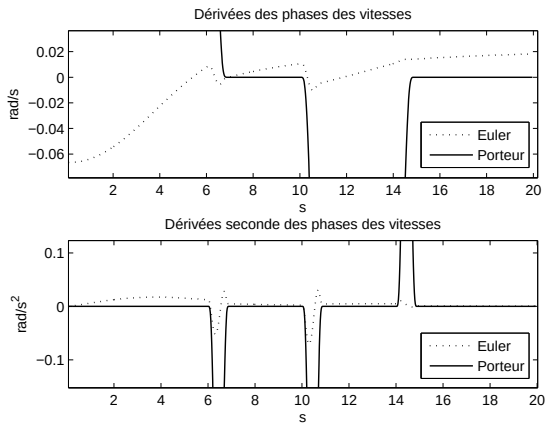


FIG. 2.9 – Dérivées première et seconde des phases des vitesses

Ces simulations ouvrent ainsi des perspectives vers un possible résultat théorique d'observabilité des trajectoires à vitesse constante.

Chapitre 3

L'estimation

Dans les contextes opérationnels qui sont ceux où intervient la trajectographie passive, les techniques d'estimation sont aussi appelées méthodes hors ligne (*offline* ou *batch* pour l'appellation anglo-saxone) par opposition aux méthodes en-ligne désignant les techniques de filtrage[1]. Ce qui justifie cette distinction, c'est que toutes les observations sont requises pour produire une estimation alors que les algorithmes de filtrage réalisent un calcul itératif lors de l'acquisition des observations. Ceci explique que les méthodes d'estimation soient peu développées dans la littérature spécialisée sur le BOT. Les moyens de calcul ayant considérablement progressé depuis les trois dernières décennies, il semblait nécessaire de revisiter ce « vieux » problème de trajectographie passive avec les outils informatiques actuels.

À la genèse de cette thèse, l'objectif était de pouvoir réaliser une trajectographie passive sur des cibles manœuvrantes. L'approche envisagée était celle de la *sélection de modèles* [86], [30], d'où la nécessité d'une étude préalable des modèles de trajectoires paramétriques et paramétriques par morceaux avec un nombre de ruptures connu. L'extension prévue était de basculer vers le cadre non-paramétrique avec des trajectoires paramétriques par morceaux ayant un nombre de ruptures inconnu.

Mais cette approche *a priori* raisonnable se heurte à un phénomène méconnu de la communauté scientifique travaillant sur le BOT : la non-robustesse de l'estimation paramétrique. Par exemple, pour une cible en MCU à faible courbure, si l'on suppose que le modèle paramétrique est celui du MRU, alors l'estimation peut fournir un résultat très éloigné de la trajectoire réelle. Ce phénomène combiné à d'autres difficultés algorithmiques va modifier la stratégie d'étude du BOT.

Les chapitres 8 et 9 sont dédiés à l'estimation de trajectoires paramétriques et paramétriques par morceaux. On se consacre ensuite à décrire, au chapitre 10, le phénomène de non-robustesse. Enfin, on présente au chapitre 11 une approche constructive pour répondre à la question de la robustesse avec un modèle semi-paramétrique.

3.1 Le phénomène de non-robustesse

La trajectographie passive pour un modèle paramétrique appartient à la catégorie des problèmes dits de *régression paramétrique* [52], [53], [61]. En effet, les observations à temps discret suivent l'équation

$$Y_k = m(\theta, t_k) + \sigma\varepsilon_k \ , \tag{1}$$

[1]Ces techniques sont présentées au chapitre 4.

où $m(\theta, t) = \Psi[S_\theta(t), t]$ avec $\{S_\theta\}_{\theta\in\Theta}$ un modèle de trajectoires paramétriques, $\{\varepsilon_k\}_{k\geq 1}$ une suite de variables gaussiennes standards et le modèle d'observations à temps continu décrit par l'équation différentielle stochastique

$$\mathrm{d}Y_t = m(\theta, t)\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t\ , \tag{2}$$

où $(W_t)_{t\in[0,1]}$ désigne un processus de Wiener. L'estimateur du maximum de vraisemblance est défini, respectivement pour ces deux modèles, par

$$\hat{\theta}_n \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \{Y_k - m(\theta, t_k)\}^2\ .$$

et

$$\hat{\theta}^\varepsilon \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} \left\{ \int_0^1 [m(\theta, t) - m(\theta^*, t)]^2\,\mathrm{d}t - 2\varepsilon \int_0^1 [m(\theta, t) - m(\theta^*, t)]\,\mathrm{d}W_t \right\}\ .$$

Sous l'hypothèse d'observabilité, on a les propriétés usuelles de l'estimateur du maximum de vraisemblance, à savoir consistance et normalité asymptotique.

L'apparente simplicité de la formulation de la trajectographie passive comme problème de régression paramétrique masque un phénomène inattendu lourd de conséquences : il s'agit de la non-robustesse de l'estimation vis-à-vis des écarts au modèle d'état. Présentons un cas d'école. Si la cible suit un MCU de courbure faible, on pourrait penser que l'estimation paramétrique d'un MRU fournit une trajectoire proche de la trajectoire réelle, du moins proche de la droite des moindres carrés de la trajectoire réelle. Considérons une situation idéale sans bruit de mesure pour des observations à temps continu. Soit $X : t \mapsto (X_1(t), X_2(t))$ la trajectoire de la cible[2]. On définit le critère des moindres carrés pour les mesures angulaires par

$$C(\theta, X) \stackrel{\text{def}}{=} \int_0^1 \{m(\theta, t) - \Psi[X(t), t]\}^2\ \mathrm{d}t\ ,$$

et la valeur du paramètre qui minimise ce critère par

$$\tilde{\theta}(X) \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} C(\theta, X)\ .$$

Sur la figure 3.1, on représente les trajectoires cible et porteur ainsi que la trajectoire $S_{\tilde{\theta}(X)}$.

[2]On opte pour la notation X dans l'étude de robustesse car on va considérer la trajectoire comme une *variable inconnue*.

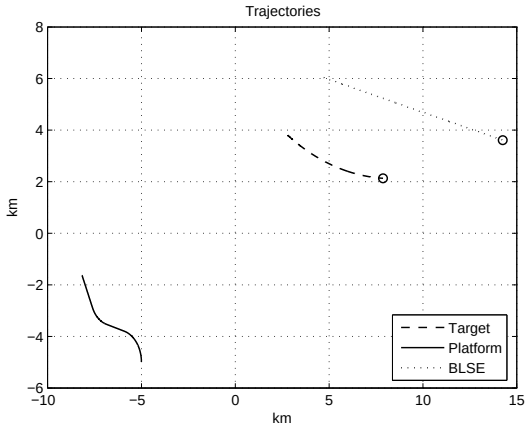


FIG. 3.1 – Sensibilité de l'estimateur des moindres carrés

On constate de manière flagrante la très haute sensibilité de $\tilde{\theta}(X)$ par rapport à X. L'étude détaillée de ce phénomène va permettre de mettre à jour des situations *surcritiques* et de quantifier cette sensibilité au modèle.

3.2 Résultats obtenus dans cette thèse

Pour décrire le phénomène de non-robustesse, on souhaite comprendre le comportement de l'application $X \mapsto \tilde{\theta}(X)$. En fait, on va se pencher sur une application définie un peu différement[3]. Notant $\mathrm{D}_{\theta,\theta}$ la différentielle seconde en la variable θ, on obtient

$$\mathrm{D}_{\theta,\theta}\, C(\theta, S_\theta) = 2I(\theta)\ .$$

Ainsi, pour une trajectoire S_{θ^*} telle que la matrice de Fisher correspondante soit inversible, on peut, par le *théorème des fonctions implicites*, définir localement une application $\bar{\theta}$ à valeurs dans Θ une partie de $\mathbb{R}^d$ telle que

$$\mathrm{D}_\theta\, C(\theta, X) = 0 \iff \theta = \bar{\theta}(X)\ .$$

Cette application est de classe $\mathscr{C}^1$ ce qui permet d'écrire, au voisinage de la trajectoire S_{θ^*}, le développement de Taylor-Young à l'ordre 1

$$\bar{\theta}(S_{\theta^*} + H) = \theta^* + \mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H) + \mathrm{o}(\|H\|_\infty)\ ,$$

et l'expression de la différentielle en S_{θ^*} est donnée par le résultat suivant :

Proposition 3.1.

$$\mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H) = I^{-1}(\theta^*) \int_0^1 \nabla_\theta m(\theta^*, t) \langle \nabla_x \Psi[S_{\theta^*}(t), t], H(t) \rangle_{\mathbb{R}^2}\, \mathrm{d}t\ .$$

On montre alors que, pour une perturbation périodique ou cyclostationnaire à fréquence élevée, la différentielle est de faible amplitude et l'estimation peu dégradée.

[3]La différence est mineure, comme cela est expliqué dans la section 10.2.1.

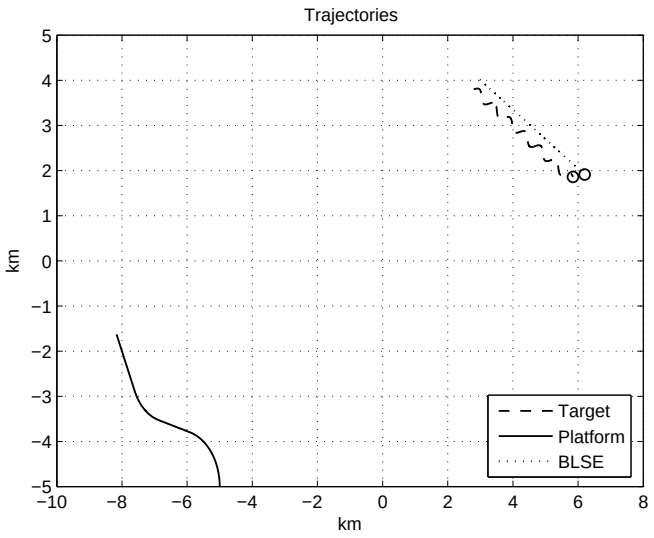


FIG. 3.2 – Perturbation périodique sur une trajectoire MRU

On montre également que le terme différentiel se décompose selon des directions et des intensités précises qui sont fonctions de la trajectoire porteur et du paramètre θ^*.

Théorème 3.1. *Soit S_{θ^*} une trajectoire paramétrique de matrice de Fisher associée inversible. Alors, on a, au voisinage de S_{θ^*}, le développement suivant*

$$\bar{\theta}(S_{\theta^*} + H) = \theta^* + \sum_{i=1}^{d} h_i D^{(i)}(\theta^*) + \mathrm{o}(\|H\|_\infty) \ ,$$

où les h_i sont fonctions de H telles que $\|h\|_{\mathbb{R}^d} \leq \|H\|_{\mathrm{L}^2}$, les vecteurs $\{D^{(i)}(\theta^*)\}_{1\leq i\leq d}$ sont orthogonaux de normes $\|D^{(i)}(\theta^*)\|_{\mathbb{R}^d} = \mu^{(i)}(\theta^*)$, $\mu^{(i)}(\theta^*)$ désignant la i-ème valeur propre d'un opérateur compact K_{θ^*} complètement déterminé par la donnée θ^* et la trajectoire porteur.

Ce résultat, outre qu'il décrit comment se dégrade l'estimation, permet de construire une perturbation *surcritique*, celle qui, pour une norme de perturbation donnée, génère une dégradation maximale de l'estimation. On peut alors mesurer l'ampleur du phénomène de non-robustesse en trajectographie passive. Ce résultat ouvre également des perspectives encourageantes pour des stratégies de commande optimale.

Dans la continuité de l'étude de robustesse, on propose un modèle de trajectographie passive avec un bruit d'état stochastique de loi inconnue. La cible est vue comme une trajectoire paramétrique à laquelle s'ajoute un bruit d'état, éventuellement dépendant, pour modéliser au mieux la trajectoire réelle d'un mobile cherchant à suivre une dynamique donnée (MRU par exemple). Les équations d'état et d'observation sont

$$\begin{cases} X_k = S_\theta(t_k) + \zeta_k \ , \\ Y_k = \Psi(X_k, t_k) + \varepsilon_k \ , \end{cases}$$

où $\theta \in \Theta \subset \mathbb{R}^d$ et $\{\varepsilon_k\}_{k\geq 0}$ est une suite de variables centrées i.i.d. de densité g et de variance σ^2. Dans un premier temps, on suppose les variables $\{\zeta_k\}_{k\geq 1}$ indépendantes et identiquement distribuées. On s'intéresse au comportement de l'estimateur des moindres carrés pour les mesures

angulaires défini par

$$\bar{\theta}_n \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \{Y_k - \Psi[S_\theta(t_k), t_k]\}^2 \ .$$

On fait alors une hypothèse clé qui va permettre d'obtenir les propriétés de consistance et de normalité asymptotique pour cet estimateur. On suppose que le bruit d'état vérifie la condition :

(N1) La distribution de ζ_1 est isotrope dans $\mathbb{R}^2$ et telle que, pour tout $t \in [0,1]$,

$$S_{\theta^*,2}(t) + \zeta_{1,2} - O_2(t) > 0 \quad \mathbb{P} - p.s.$$

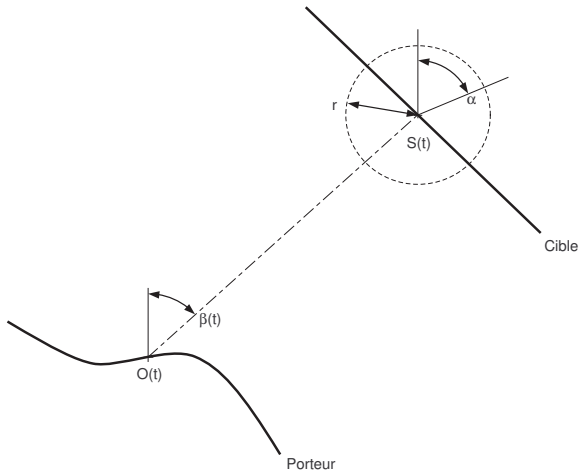


Fig. 3.3 – Bruit d'état isotrope

On définit, pour tout $\theta \in \Theta$,

$$\begin{aligned} I_R(\theta) &\stackrel{\text{def}}{=} \int_0^1 \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t \ , \\ I_\Psi(\theta) &\stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_\theta(t), t]\right\}^2 \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t \ . \end{aligned}$$

On a le résultat suivant :

Théorème 3.2. *Pour une trajectoire paramétrique S_{θ^*} observable, de matrice de Fisher inversible, l'estimateur $\bar{\theta}_n$ est consistant et vérifie*

$$\sqrt{n}\left(\bar{\theta}_n - \theta^*\right) = \\ I_R^{-1}(\theta^*) n^{-1/2} \sum_{k=1}^{n} \left\{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_{\theta^*}(t_k), t_k] + \varepsilon_k\right\} \nabla_\theta \Psi[S_{\theta^*}(t_k), t_k] + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

En particulier, $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ converge en loi, pour $n \to \infty$, vers $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ où

$$I_M^{-1}(\theta^*) = I_R^{-1}(\theta^*) \left[I_\Psi(\theta^*) + \sigma^2 I_R(\theta^*)\right] I_R^{-1}(\theta^*) \ .$$

Toutefois, le résultat de normalité asymptotique ne permet pas de construire une région de

confiance pour θ^*. En effet, la variance limite fait intervenir $I_\Psi(\theta^*)$, matrice dans laquelle apparaît la loi du bruit d'état qui est inconnue. Sous une hypothèse peu contraignante, on construit une région de confiance *conservative*, complètement déterminée par l'observation. On suppose que :

(N2) la trajectoire S_{θ^*} est telle que, pour tout $t \in [0,1]$,

$$\|S_{\theta^*}(t) - O(t)\| \geq R_{\min} \ ,$$

et on connaît une constante A^2 qui vérifie

$$\pi^2 \left(1 + \pi^{-2/3}\right)^3 \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{R_{\min}^2} \leq A^2 \ .$$

Théorème 3.3. *Soit S_{θ^*} une trajectoire paramétrique observable et de matrice de Fisher inversible. Si les conditions (N1) et (N2) sont satisfaites, alors, pour tout $\alpha > 0$, notant C_α une région de confiance de niveau $1-\alpha$ pour la loi normale standard dans $\mathbb{R}^d$, on a*

$$\liminf_{n\to\infty} \mathbb{P}_{\theta^*}\left[\sqrt{n}(A^2+\sigma^2)^{-1/2} I_R^{1/2}\left(\bar{\theta}_n\right)\left(\bar{\theta}_n - \theta^*\right) \in C_\alpha\right] \geq 1-\alpha \ .$$

Dans un deuxième temps, on étudie le comportement asymptotique de l'estimateur du maximum de vraisemblance. On note $\mathcal{F}$ l'ensemble des densités de probabilité f telles que pour tout $t \in [0,1]$ et tout $\theta \in \Theta$,

$$\int \Psi[S_\theta(t) + \zeta, t] f(\zeta)\,\mathrm{d}\zeta = \Psi[S_\theta(t), t] \ .$$

La log-vraisemblance normalisée est définie par

$$J_n(\theta, f) \stackrel{\text{def}}{=} n^{-1} \sum_{k=1}^{n} \log\left(\int g\left\{Y_k - \Psi[S_\theta(t_k) + u, t_k]\right\} f(u)\,\mathrm{d}u\right) .$$

et si la densité f^* est connue, l'estimateur du maximum de vraisemblance est donné par

$$\tilde{\theta}_n \stackrel{\text{def}}{=} \arg\max_{\theta\in\Theta} J_n(\theta, f^*) \ .$$

On démontre la consistance et la normalité asymptotique de $\tilde{\theta}_n$. On s'intéresse ensuite à l'optimalité des estimateurs dans le cas semiparamétrique quand la densité du bruit d'état f^* est inconnue et on démontre le caractère LAN du modèle statistique $(\mathbb{P}_{n,(\theta,f)})_{(\theta,f)\in\Theta\times\mathcal{F}}$.

Enfin, on décrit comment les différents résultats s'étendent au cas de variables dépendantes, sous condition sur les coefficients de mélange $\{\alpha_k\}_{k\geq 1}$ du bruit d'état [100].

Chapitre 4

Le filtrage

Les méthodes les plus prisées en poursuite (*tracking*) et notamment pour le BOT sont les méthodes dites de filtrage qui appartiennent à la catégorie des méthodes en ligne (*online methods*), par opposition aux méthodes *batch*[1].

On distingue deux familles parmi les techniques de filtrage utilisées en trajectographie passive : les méthodes apparentées au *filtre de Kalman* et les *méthodes de Monte Carlo séquentielles* (Sequential Monte Carlo Methods, SMC), encore appelées *filtres particulaires* (Particle filters), voir [6] ou [42] pour une synthèse récente. Ces techniques présentent un intérêt évident qui est l'adéquation à un traitement temps réel, point crucial pour les situations tactiques. Un autre aspect qui explique l'engouement pour ces méthodes est leur facilité d'implémentation.

Le filtre conçu par Kalman [70] ,[22], résout un cas très particulier de problème de filtrage, communément appelé le cas linéaire gaussien. Pour les cas non linéaires ou non gaussiens, un grand nombre de stratégies ont été développées qui sont des extensions ou qui utilisent le filtre de Kalman. Ces méthodes dérivées sont sujettes, de par leur conception, à une source d'erreur intrinsèque et ce quelque soit la puissance de calcul disponible.

Les filtres particulaires, dont le succès ne se dément pas depuis une quinzaine d'années, n'ont pas ce défaut majeur. Ces nouvelles générations de filtres adressent des modèles très généraux et leurs performances sont essentiellement liées à la puissance des calculateurs. Le coût de ces algorithmes est d'ailleurs le prix à payer pour s'affranchir du calcul approximatif des méthodes héritées de Kalman. Toutefois, vue la croissance remarquable des capacités des moyens de calculs, ce point ne constitue pas un handicap véritable pour l'usage de ces méthodes.

4.1 Les Chaînes de Markov Cachées

Avant d'aborder le principe du filtrage, il convient de préciser le cadre mathématique qu'est celui des *Chaînes de Markov cachées* ou *Hidden Markov Model* (HMM). Une chaîne de Markov cachée est une suite de couples de variables aléatoires $\{(X_k, Y_k)\}_{k\geq 0}$ à valeurs dans des espaces probabilisés. La suite $\{X_k\}_{k\geq 0}$ est une chaîne de Markov qui n'est pas observée directement et la suite $\{Y_k\}_{k\geq 0}$ représente la séquence des observations de ces états cachés. On fait l'hypothèse que, conditionnellement aux états cachées, les observations sont indépendantes et que la loi de Y_k ne dépend que de l'état associé X_k, hypothèse dite de *canal sans mémoire*. L'hypothèse markovienne

[1] Voir page 17.

s'écrit

$$\mathbb{P}_\nu\left(X_0 \in \mathrm{d}x_0, \cdots, X_n \in \mathrm{d}x_n\right) = \underbrace{\mathbb{P}(X_0 \in \mathrm{d}x_0)}_{\nu(\mathrm{d}x_0)} \prod_{k=1}^{n} \underbrace{\mathbb{P}\left(X_k \in \mathrm{d}x_k | X_{k-1} = x_{k-1}\right)}_{Q_k(x_{k-1}, \mathrm{d}x_k)} ,$$

et l'hypothèse de canal sans mémoire

$$\begin{aligned}
&\mathbb{P}\left(Y_0 \in \mathrm{d}y_0, \ldots, Y_n \in \mathrm{d}y_n | X_0 = x_0, \ldots, X_n = x_n\right) \\
&\qquad = \prod_{k=0}^{n} \mathbb{P}\left(Y_k \in \mathrm{d}y_k | X_0 = x_0, \ldots, X_n = x_n\right) , \\
&\qquad = \prod_{k=0}^{n} \underbrace{\mathbb{P}\left(Y_k \in \mathrm{d}y_k | X_k = x_k\right)}_{g_k(x_k, y_k)\mu(\mathrm{d}y_k)} .
\end{aligned}$$

On définit la fonction de vraisemblance Ψ_k par

$$\Psi_k(x) \stackrel{\text{def}}{=} g_k(x, Y_k) .$$

On représente habituellement la chaîne de Markov cachée $\{(X_k, Y_k)\}_{k\geq 0}$ par le diagramme synthétique suivant :

$$\begin{tikzcd}
\arrow[r] & X_{k-1} \arrow[r] \arrow[d] & X_k \arrow[r] \arrow[d] & X_{k+1} \arrow[r] \arrow[d] & {} \\
& Y_{k-1} & Y_k & Y_{k+1} &
\end{tikzcd}$$

Pour illustrer ce type de dépendance, on peut décrire la chaîne de Markov cachée par les équations

$$\begin{cases} X_k = a_k(X_{k-1}, \zeta_k) , \\ Y_k = b_k(X_k, \varepsilon_k) , \end{cases}$$

avec $\{\zeta_k\}_{k\geq 0}$ et $\{\varepsilon_k\}_{k\geq 0}$ des suites mutuellement indépendantes de variables aléatoires i.i.d. également indépendantes de X_0.

Décrivons la formulation la plus couramment utilisée du BOT en tant que chaîne de Markov cachée. Les équations d'état et d'observation sont données par

$$\begin{cases} X_k = FX_{k-1} + \zeta_k , & (1) \\ Y_k = h(X_k, t_k) + \varepsilon_k , & (2) \end{cases}$$

avec X_k la matrice colonne $X_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, $F = \begin{pmatrix} I_2 & \Delta_T I_2 \\ 0 & I_2 \end{pmatrix}$ où Δ_T représente la période d'échantillonnage, $\{\zeta_k\}_{k\geq 0}$ définie par $\zeta_k = \beta G V_k$ où $\beta > 0$, $G = \begin{pmatrix} \frac{\Delta_T^2}{2} I_2 \\ \Delta_T I_2 \end{pmatrix}$ et $\{V_k\}_{k\geq 0}$ est une suite de variables indépendantes gaussiennes standards dans $\mathbb{R}^2$, $\{\varepsilon_k\}_{k\geq 0}$ une suite de variables i.i.d. de loi $\mathcal{N}(0, \sigma^2)$ indépendantes de $\{\zeta_k\}_{k\geq 0}$ et h l'application définie, pour $X = \begin{pmatrix} x & y & \dot{x} & \dot{y} \end{pmatrix}^{\mathrm{T}}$, par

$$h(X, t) \stackrel{\text{def}}{=} \Psi[(x, y), t] = \mathrm{angle}[x - O_1(t), y - O_2(t)] .$$

La dynamique de l'équation (1) correspond à un modèle d'accélération de type bruit blanc gaus-

sien avec l'hypothèse que ce bruit est constant sur un intervalle d'échantillonnage. La partie non bruitée de cette équation « $X_k = FX_{k-1}$ » correspond à une dynamique MRU. Ce modèle s'inscrit dans la catégorie des dynamiques dites non-manœuvrantes, voir [10], [11], [80]. Il existe des modèles plus sophistiqués où la dynamique peut basculer entre plusieurs modes et où la succession de ces modes constitue une chaîne de Markov. Ainsi, dans [17], [11], [7], on trouve des dynamiques manœuvrantes où la cible peut alterner des régimes MRU bruités et MCU bruités, avec une accélération de giration fixée. Ces modèles sont abondamment utilisés pour la définition des filtres particulaires.

La modélisation du BOT en chaîne de Markov cachée est une approche véritablement différente de celles rencontrées dans les méthodes *batch*. En effet, dans le cas des techniques d'estimation, la dynamique de la cible est décrite par un modèle paramétrique, paramétrique par morceaux ou éventuellement paramétrique bruité. Ces dynamiques sont relativement rigides et constituent un fort *a priori* sur la trajectoire de la cible. Le choix de la modélisation HMM donne de la souplesse à cette dynamique et permet d'envisager des trajectoires non conformes à des modèles rigides. Les ingénieurs qui emploient des filtres en trajectographie passive utilisent d'ailleurs les niveaux de bruit d'état et bruit d'observation comme des variables d'ajustement pour traduire la confiance qu'ils ont respectivement dans le modèle de dynamique et dans les observations. Toutefois, dès lors que la trajectoire cible est vue comme une réalisation aléatoire, l'objectif n'est plus l'estimation d'un paramètre mais l'acquisition de l'information maximale sur l'état courant de la cible sachant les observations acquises, ce qui amène à la notion de *filtre optimal*.

4.2 Le Filtre Optimal

Définition 4.1. *Soit $\{(X_k, Y_k)\}_{k\geq 0}$ une chaîne de Markov cachée à valeurs dans $\mathsf{X} \times \mathsf{Y}$. On définit le* filtre optimal*, appelé aussi* filtre bayésien optimal *comme la loi a posteriori de l'état courant X_n sachant toutes les observations acquises à l'instant n, c'est-à-dire $Y_{0:n} \stackrel{\text{def}}{=} \{Y_0, \ldots, Y_n\}$.*

On note $\phi_{\nu,n}[y_{0:n}]$ la loi de l'état caché X_n conditionnellement aux observations $y_{0:n} = \{y_0, \ldots, y_n\}$ donnée par

$$\begin{aligned} \phi_{\nu,n}[y_{0:n}](A) &= \mathbb{P}_\nu\left(X_n \in A | Y_0 = y_0, \ldots, Y_n = y_n\right) , \\ &= \frac{\nu\left[g_0(\cdot, y_0)Q_1 g_1(\cdot, y_1)Q_2 \ldots Q_n g_n(\cdot, y_n)\mathbb{1}_A\right]}{\nu\left[g_0(\cdot, y_0)Q_1 g_1(\cdot, y_1)Q_2 \ldots Q_n g_n(\cdot, y_n)\right]} . \end{aligned}$$

On notera également le filtre optimal $\pi_n \stackrel{\text{def}}{=} \phi_{\nu,n}[Y_{0:n}]$ en omettant volontairement la dépendance en ν lorsqu'il n'y a pas d'ambiguïté sur la loi initiale de l'état caché. Le filtre optimal s'interprète comme l'information optimale dont on dispose sur l'état courant de la chaîne cachée compte-tenu des observations effectuées.

La détermination du filtre optimal π_n pour une HMM est un problème difficile. La première démarche consiste à écrire les équations d'évolution de ce filtre. On utilisera les notations en usage pour les lois conditionnelles suivantes :

$$\pi_{k|k-1} = \mathscr{L}(X_k | Y_{0:k-1}) , \qquad \pi_k = \mathscr{L}(X_k | Y_{0:k}) .$$

Théorème 4.1. *La suite $\{\pi_k\}_{k\geq 0}$ vérifie les équations d'évolution, dites équations de* prédiction *et de*

correction, *décrites par le schéma*

$$\pi_{k-1} \xrightarrow[\textit{Prédiction}]{\mathscr{L}(X_k|X_{k-1})} \pi_{k|k-1} = \pi_{k-1}Q_k \xrightarrow[\textit{Correction}]{Y_k} \pi_k = \Psi_k \cdot \pi_{k|k-1}$$

où

$$\begin{aligned} \pi_{k-1}Q_k(\mathrm{d}x') &= \int \pi_{k-1}(\mathrm{d}x)Q_k(x,\mathrm{d}x') \ , \\ \Psi_k \cdot \pi_{k|k-1}(\mathrm{d}x) &= \frac{\Psi_k(x)\pi_{k-1}(\mathrm{d}x)}{\langle \pi_{k-1}, \Psi_k \rangle} \ , \quad \text{avec} \quad \langle \pi_{k|k-1}, \Psi_k \rangle = \int \Psi_k(x)\, \pi_{k|k-1}(\mathrm{d}x) \ . \end{aligned}$$

Il existe quelques cas simples pour lesquels on sait déterminer le filtre optimal :

- quand la chaîne cachée est à valeurs dans un espace d'état fini, problème résolu par l'algorithme *forward* de Baum [13], [14] ;
- dans le cas *linéaire gaussien*, quand l'équation d'état et l'équation d'observation sont linéaires avec des bruits gaussiens [70], [22].

$$\begin{cases} X_n = F_n X_{n-1} + c_n + G_n W_n \ , \\ Y_n = H_n X_n + d_n + V_n \ . \end{cases}$$

Dans le cas linéaire gaussien, l'espace d'état est non compact et la chaîne non nécessairement ergodique. Ceci va motiver la création d'un grand nombre de filtres dérivées du filtre de Kalman pour des cas non-linéaires et/ou non-gaussiens : EKF, UKF, GB(1), GB(2), IMM, ... [2], [69], [18], [47]. Ces techniques ont un point commun : les équations d'évolution des filtres sont calculées par des approximations [5]. Ainsi, dans le cas de l'EKF (Extended Kalman Filter), on réalise une approximation du type développement de Taylor au premier ordre ; pour l'UKF (Uscented Kalman Filter), une méthode de quadrature sert au calcul d'espérance. Il existe quelques rares cas où l'on a la convergence de ces filtres dérivés. Mais en général, le filtre dérivé n'approche pas le filtre optimal et ce comportement est intrisèque à la définition du filtre, ceci quelque soit la puissance des calculateurs qui réalisent les simulations.

Les filtres particulaires, génération d'algorithmes apparus il y a une quinzaine d'année, sont très différents dans leur conception. Le principe du filtrage particulaire consiste à approcher le filtre π_k par une somme pondérées de mesures de Dirac :

$$\pi_k \simeq \pi_k^N \stackrel{\text{def}}{=} \sum_{i=1}^N \omega_k^i \delta_{\xi_k^i} \ , \quad \text{avec} \quad \sum_{i=1}^N \omega_k^i = 1 \ .$$

Le filtre π_k^N est appelé *filtre particulaire* où N désigne le nombre de particules que sont les variables aléatoires $\{\xi_k^i\}_{1\leq i\leq N}$ et $\{\omega_k^i\}_{1\leq i\leq N}$ les poids de ces particules. L'idée phare des méthodes de Monte-Carlo séquentielles consiste à faire suivre au filtre particulaire les équations du filtre optimal. On représente alors les équations d'évolution du filtre particulaire par le schéma :

$$\pi_{k-1}^N \xrightarrow[\text{Mutation}]{} \pi_{k|k-1}^N \xrightarrow[\text{Sélection}]{} \pi_k^N$$

où les étapes de mutation et sélection sont :

– Mutation :

$$\pi^N_{k|k-1} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi^i_k}, \quad \xi^i_k \sim \pi^N_{k-1} Q_k$$

– Sélection :

$$\pi^N_k = \Psi_k \cdot \pi^N_{k|k-1} = \sum_{i=1}^{N} \frac{\Psi_k(\xi^i_k)}{\sum_{j=1}^{N} \Psi_k(\xi^j_k)} \delta_{\xi^i_k}$$

Ce schéma de filtre particulaire est un peu naïf mais décrit les deux étapes fondamentales de toutes les méthodes de Monte-Carlo séquentielles. Pour un aperçu récent et très complet de ces algorithmes, on peut consulter [45], [25], [42]. Si ces algorithmes ont littéralement envahi le champ d'investigation des scientifiques travaillant sur la poursuite, l'étendu des travaux théoriques où la convergence de ces filtres particulaires est prouvée reste assez restreinte. L'essentiel de ces résultats concernent les chaînes de Markov cachées ergodiques [36], [75], [113]. Les résultats établis par Oudjane et Rubenthaler [93] et Crisan et Heine [33] sont parmi les principales références qui traitent de filtres particulaires sur des chaînes non-ergodiques. Plus récemment, Handel [111], [112] obtient la stabilité asymptotique sans aucune condition sur le noyau de transition de la chaîne cachée mais avec certaines restrictions sur les lois initiales et pour des observations issues du modèle définissant le filtre. La convergence des filtres particulaires est étroitement liée à la propriété d'oubli du filtre optimal aussi appelée *stabilité asymptotique* [26], [93], [36], [37]. Cette propriété consiste, pour deux lois initiales ν et ν', à montrer que la distance en variation totale entre $\phi_{\nu,n}$ et $\phi_{\nu',n}$ est de limite nulle, *i.e.*

$$\lim_{n\to\infty} \left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} = 0 \quad \mathbb{P}_\star - p.s.$$

Un des enjeux de la stabilité asymptotique est son application à la convergence uniforme sur un horizon temporel infini des méthodes particulaires. La construction récursive de ces filtres requière en fait une convergence plus forte que celle décrite ci-avant. L'oubli à vitesse exponentielle est nécessaire pour montrer la convergence des filtres particulaires

$$\begin{aligned} \limsup_{n\to\infty} n^{-1} \log \left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} &< 0, \quad \mathbb{P}_\star - p.s. \,, \\ \limsup_{n\to\infty} n^{-1} \log \mathbb{E}_\star \Big[\left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} \Big] &< 0 \,. \end{aligned}$$

Une certaine prudence s'impose toutefois vis-à-vis des filtres particulaires pour le BOT. En effet, on pourrait, à la lecture de certains articles, penser que ces algorithmes révolutionnent les résultats obtenus en trajectographie passive. Mais les expérimentations menées au cours de cette thèse pour des trajectoires plausibles ne confortent pas cet optimisme, en particulier dans des situations où l'on observe des phénomènes de non-robustesse pour les méthodes batch. Enfin, soulignons que le filtre optimal peut ne pas être informatif. S'il est très étalé, l'information qu'il fournit n'est pas pertinente et, même si les filtres particulaires présentent un bon comportement asymptotique, ils donneront au mieux l'information contenue dans le filtre optimal, c'est-à-dire une information peu précise pour une décision opérationnelle.

4.3 Principaux apports de cette thèse

Le chapitre 12 présente l'EKF appliqué au problème de trajectographie passive. On y expose le comportement inattendu de ce filtre quand les équations d'évolution tiennent compte d'un bruit d'état.

Dans le chapitre 13, on étend les résultats de stabilité du filtre optimal pour des modèles de chaînes de Markov cachées non-ergodiques établis dans [24], [93] et [32]. On considère un modèle de chaîne cachée $\{X_k, Y_k\}_{k\geq 1}$ à valeurs dans $\mathsf{X} \times \mathsf{Y} = \mathbb{R}^{d_\mathsf{X}} \times \mathbb{R}^{d_\mathsf{Y}}$ décrit par les équations

$$\begin{cases} X_k = f(X_{k-1}) + \zeta_k \ , \\ Y_k = h(X_k) + \varepsilon_k \ . \end{cases} \tag{3}$$

Le bruit d'observations $\{\varepsilon_k\}_{k\geq 0}$ est une suite de variables aléatoires i.i.d de densité υ. On fait les hypothèses suivantes :

(E1) f est a-Lipschitzienne, i.e. $|f(x)-f(y)| \leq a|x-y|$ et h est uniformément continue et surjective telle que pour tout $y_1,\ y_2 \in \mathsf{Y}$ et $x_1,\ x_2 \in \mathsf{X}$ des antécédents de y_1 et y_2, il existent des constantes b_0 et b telles que
$$|x_1 - x_2| \leq b_0 + b|y_1 - y_2| \ .$$

(E2) La densité υ est strictement positive, continue et $\lim_{|u|\to\infty} \upsilon(u) = 0$.

On définit la notion de *fonction LD-set* qui, pour le système (3), est donnée par

$$y \longmapsto \mathsf{C}_\Delta(y) \stackrel{\text{def}}{=} \{x \in \mathsf{X} : |h(x) - y| \leq \Delta\} \ .$$

On définit également une suite de variables aléatoires pour tout $k \geq 1$ par

$$Z_k^\Delta \stackrel{\text{def}}{=} -\log \gamma^- \left[2(a+1)b_0 + (a+1)b\Delta + ab|\varepsilon_{k-1}| + |\zeta_k| + b|\varepsilon_k|\right] \ .$$

On obtient alors un résultat d'oubli à vitesse exponentielle presque sûrement.

Proposition 4.1. *Pour le modèle* (3) *sous les hypothèses (E1), (E2), avec un bruit d'état* $\{\zeta_k\}_{k\geq 0}$ *i.i.d. et indépendant de* $\{\varepsilon_k\}_{k\geq 0}$*, en supposant que pour tout* $\Delta > 0$*, on a*

$$\mathbb{E}|Z_1^\Delta| < \infty \ ,$$

alors, pour toute loi de probabilité ν_0*,* ν *et* ν' *sur* $(\mathsf{X}, \mathcal{X})$*, on a*

$$\limsup_{n\to\infty} n^{-1} \log \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} < 0, \quad \mathbb{P}_{\nu_0} - p.s.$$

Definition 4.1. *On définit la transformée de Legendre* $\psi_Z(\lambda)$ *de la variable aléatoire* Z *sur l'ensemble* $\{\lambda \geq 0 : \mathbb{E}[\mathrm{e}^{\lambda Z}] < \infty\}$ *par* $\psi_Z(\lambda) \stackrel{\text{def}}{=} \log \mathbb{E}[\mathrm{e}^{\lambda Z}]$.

On complète alors le résultat presque sûr par un résultat en espérance.

Proposition 4.2. *On considère le modèle* (3) *sous les hypothèses (E1), (E2), avec un bruit d'état* $\{\zeta_k\}_{k\geq 0}$ *i.i.d. et indépendant de* $\{\varepsilon_k\}_{k\geq 0}$*, et on suppose que pour tout* $\Delta > 0$*, il existe* $\alpha_0 > 0$ *tel que*

$$\psi_{Z_1^\Delta}(\alpha_0) > 0 \tag{4}$$

Alors, pour toute loi de probabilité ν *et* ν' *sur* $(\mathsf{X}, \mathcal{X})$ *et* $\Delta > 0$ *tels que*

$$\mathbb{E}\left\{\exp\left(\alpha_0[\log \nu g(\cdot, Y_0) Q \mathbb{1}_{\mathsf{C}(Y_1,\Delta)}]_-\right)\right\} < \infty \ , \quad \mathbb{E}\left\{\exp\left(\alpha_0[\log \nu' g(\cdot, Y_0) Q \mathbb{1}_{\mathsf{C}(Y_1,\Delta)})]_-\right)\right\} < \infty \ , \tag{5}$$

on a

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{E}\left[\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right] < 0 \ .$$

Ces résultats sont ensuite étendus au cas où le bruit d'état peut être dépendant de l'état courant :

(G) $\{\zeta_k\}_{k\geq 0}$ est une suite de variables aléatoires telle que, pour tout entier k, ζ_k est indépendant de ε_k et pour, tout $A \in \mathcal{X}$,

$$\mathbb{P}(\zeta_k \in A | X_{k-1} = x) = \int t(x, x') \mathbb{1}_A(x') \, \lambda^{\mathrm{Leb}}(\mathrm{d}x') \ .$$

De plus, il existe une densité de probabilité ψ et des constantes strictement positives $\mu^-, \ \mu^+$ telles que, pour tout $x, \ x' \in \mathsf{X}$,

$$\mu^- \psi(x') \leq t(x, x') \leq \mu^+ \psi(x') \ .$$

Ce type de dépendance permet notamment de couvrir le cas où le signal suit l'équation

$$X_k = f(X_{k-1}) + \sigma(X_{k-1}) \xi_k \ ,$$

avec $\{\xi_k\}_{k\geq 0}$ une suite i.i.d. ou encore le cas de certaines diffusions échantillonnées décrites par une EDS de la forme

$$\mathrm{d}X_t = \rho(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t \ ,$$

avec, pour ces deux équations, une condition d'ellipticité uniforme sur σ.

Enfin, dans le chapitre 14, on décrit un cadre de trajectographie pour lequel s'applique les résultats de stabilité du chapitre 13. On décrit également comment appliquer les résultats de convergence de filtre particulaire établis par Oudjane et al.[93] et on termine par des simulations de ces algorithmes sur des scénarios réalistes de trajectographie passive.

Deuxième partie

Observabilité

Chapitre 5

Trajectoires paramétriques

Sommaire

5.1 Introduction 33
5.2 Observabilité 34
5.3 Le mouvement rectiligne uniforme 41
5.4 Le mouvement circulaire uniforme 45
5.5 Conclusion 46
5.6 Preuves de la section 5.4 47

5.1 Introduction

La notion d'observabilité en trajectographie passive consiste à déterminer si une trajectoire cible est caractérisée par des mesures angulaires non bruitées. Autrement dit, peut-on reconstruire toute la trajectoire à partir de ces seules mesures ?

Sans restriction sur l'espace des trajectoires cibles considérées, il n'y a jamais observabilité. En effet, la trajectographie passive appartient à la catégorie des *problèmes inverses mal posés*. Pour un jeu de mesures angulaires donné, il existe une infinité de trajectoires conformes à ces mesures. On suppose que les observations sont réalisées sur l'intervalle de temps $[0,1]$. À un instant t fixé dans $[0,1]$, on note $O(t) = (O_1(t), O_2(t))$ la position du porteur et $S(t) = (S_1(t), S_2(t))$ la position de la cible. Soit la fonction angle définie, pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, par

$$\mathrm{angle}(x,y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \mathrm{sgn}(x) \times \mathbb{1}_{y<0} & \text{si } x \neq 0 \text{ et } y \neq 0, \\ \frac{\pi}{2} \times \mathrm{sgn}(x) & \text{si } x \neq 0 \text{ et } y = 0, \\ \frac{\pi}{2} \times [1 - \mathrm{sgn}(y)] & \text{si } x = 0 \text{ et } y \neq 0, \end{cases}$$

où sgn représente la fonction signe. On introduit la fonction Ψ définie pour tout $t \in [0,1]$ et pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ par

$$\Psi[(x_1,x_2),t] \stackrel{\text{def}}{=} \begin{cases} \mathrm{angle}[x - O_1(t), y - O_2(t)] & \text{si } (x,y) \neq O(t)\ , \\ 0 & \text{sinon}\ . \end{cases} \tag{1}$$

La mesure sans bruit collectée par le porteur à l'instant t est donnée par $\Psi[S(t),t]$[1]. Alors, toute

[1] Il s'agit d'un azimut, donc l'angle est mesuré relativement au nord magnétique d'où le sens inhabituel du rapport dans le terme $\arctan(x/y)$.

trajectoire de la forme

$$t \longmapsto O(t) + \lambda(S - O)(t) \ ,$$

avec par exemple λ un réel strictement positif est conforme, sur le plan des mesures angulaires, avec la trajectoire cible.

Pour qu'une trajectoire soit caractérisée par ses mesures angulaires, il faut donc imposer une restriction à l'ensemble des trajectoires admissibles. Les trajectoires *paramétriques*, c'est-à-dire qui sont complètement décrites par un paramètre comme l'est une trajectoire linéaire à vitesse constante, constitue le cadre le plus simple et le plus naturel pour amorcer l'étude de l'observabilité en trajectographie passive.

Dans ce chapitre, on s'intéresse, dans la section 5.2, à l'observabilité dans des modèles paramétriques généraux et aussi dans une famille de modèles paramétriques pour lesquels les composantes des trajectoires s'écrivent dans une base donnée. On présente ensuite, dans les sections 5.3 et 5.4, des résultats concernant les dynamiques les plus fréquentes en situation opérationnelle de trajectographie passive que sont le MRU et le MCU. Pour une meilleure lisibilité, les preuves de la section 5.4 sont rassemblées dans la section 5.6.

5.2 Observabilité

Donnons une définition formelle aux termes de trajectoire et de modèle paramétrique.

Définition 5.1. *Une trajectoire est une application continue de* $[0,1]$ *dans* $\mathbb{R}^2$.

Dans la réalité, la trajectoire d'un mobile opérationnel est plus lisse que simplement continue mais cette hypothèse suffit pour énoncer les résultats qui vont suivre.

Définition 5.2. *Un modèle de trajectoire paramétrique ou modèle paramétrique est la donnée d'un sous-ensemble compact* Θ *de* $\mathbb{R}^d$ *avec* d *un entier non nul et d'une application* $\theta \mapsto S_\theta$ *où* S_θ *est une trajectoire telle que, pour tout* t *fixé dans* $[0,1]$*, l'application* $\theta \mapsto S_\theta(t)$ *est de classe* $\mathscr{C}^2$ *avec* $t \mapsto \mathrm{D}_\theta S_\theta(t)$ *et* $t \mapsto \mathrm{D}_{\theta,\theta} S_\theta(t)$ *des applications continues sur* $[0,1]$*. On notera ce modèle paramétrique* $\{S_\theta\}_{\theta\in\Theta}$.

Supposons donnés une trajectoire porteur O et un modèle de trajectoire paramétrique $\{S_\theta\}_{\theta\in\Theta}$. On fait l'hypothèse suivante :

$$(\mathrm{H}_+) \qquad \forall(\theta,t) \in \Theta \times [0,1] \qquad S_{\theta,2}(t) - O_2(t) > 0.$$

Sous cette condition, on a, pour tout $\theta \in \Theta$ et $t \in [0,1]$,

$$\Psi[S_\theta(t), t] = \arctan[S_{\theta,1}(t) - O_1(t)]/[S_{\theta,2}(t) - O_2(t)] \ ,$$

et on note

$$m(\theta,t) \stackrel{\text{def}}{=} \Psi[S_\theta(t), t] \ .$$

On travaillera alternativement avec le modèle d'observations à temps discret

$$Y_k = m(\theta, t_k) + \sigma\varepsilon_k \ , \tag{2}$$

où $\{t_k\}_{1\le k\le n}$ est une suite à valeurs dans $[0,1]$, σ est un réel strictement positif et $\{\varepsilon_k\}_{1\le k\le n}$ est une séquence de variables gaussiennes standards, et avec le modèle à temps continu

$$\mathrm{d}Y_t = m(\theta,t)\mathrm{d}t + \varepsilon \mathrm{d}W_t \ , \tag{3}$$

où ε est un réel strictement positif et $(W_t)_{t\in[0,1]}$ est un processus de Wiener.

On définit l'application Φ sur l'ensemble des trajectoires par

$$\Phi : X \longmapsto (t \mapsto \Psi[X(t), t]) \ ,$$

autrement dit, $\Phi(X)$ est une application continue de $[0,1]$ dans $\mathbb{R}$ telle que $\Phi(X)(t) = \Psi[X(t), t]$. Par la suite, on note

$$\Phi(X)(t_{1:n}) \stackrel{\text{def}}{=} (\Phi(X)(t_1), \ldots, \Phi(X)(t_n)) \ .$$

Définition 5.3. *Pour le modèle d'observations à temps discret* (2)*, une trajectoire paramétrique S_{θ^*} est dite* observable *dans $\{S_\theta\}_{\theta\in\Theta}$ si, pour tout $\theta \in \Theta$,*

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \Longrightarrow \theta = \theta^* \ .$$

Pour le modèle d'observations à temps continu (3)*, une trajectoire paramétrique S_θ est dite* observable *dans $\{S_\theta\}_{\theta\in\Theta}$ si, pour tout $\theta \in \Theta$,*

$$\Phi(S_\theta) = \Phi(S_{\theta^*}) \Longrightarrow \theta = \theta^* \ .$$

S'il n'y a pas d'ambiguïté sur le modèle paramétrique, on dira simplement d'une trajectoire qu'elle est observable *si elle vérifie une des deux conditions mentionnées ci-avant.*

Ainsi, une trajectoire S_{θ^*} est observable si ses mesures angulaires non bruitées caractérisent de manière unique son paramètre et donc la trajectoire toute entière. Signalons que la notion d'observabilité est plus faible que l'identifiabilité du modèle paramétrique sous-jacent, notion habituellement utilisée en théorie de l'estimation. Détaillons ce point. On note indifféremment $(\mathbb{P}_\theta)_{\theta\in\Theta}$ le modèle statistique associé aux observations à temps discret (2) ou à temps continu (3).

Définition 5.4. *Le modèle statistique $(\mathbb{P}_\theta)_{\theta\in\Theta}$ est dit* identifiable *si l'application $\theta \mapsto \mathbb{P}_\theta$ est injective.*

On a, respectivement pour le modèle (2),

$$\mathbb{P}_\theta = \mathbb{P}_{\theta'} \iff \Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta'})(t_{1:n}) \ ,$$

et, pour le modèle (3)

$$\mathbb{P}_\theta = \mathbb{P}_{\theta'} \iff \Phi(S_\theta) = \Phi(S_{\theta'}) \quad \lambda^{\text{Leb}} - p.p.$$

Comme l'application $t \mapsto \Phi(S_\theta)(t)$ est continue, l'égalité presque partout dans le membre de droite de l'équivalence ci-dessus est en fait une égalité. Par suite, sous l'hypothèse d'identifiabilité, deux trajectoires paramétriques ayant mêmes séquences angulaires sont identiques. Pour une trajectoire S_{θ^*} observable, toute trajectoire paramétrique avec des mesures identiques à celles de S_{θ^*} est nécessairement S_{θ^*}. L'observabilité est donc relative à une trajectoire paramétrique donnée tandis que l'identifiabilité permet de distinguer n'importe quelle trajectoire paramétrique d'une autre. On verra que cette différence n'est pas anecdotique puisqu'il existe des scénarios pour lesquels une trajectoire S_{θ^*} est observable sans que le modèle statistique $(\mathbb{P}_\theta)_{\theta\in\Theta}$ soit identifiable, voir la proposition 5.9.

Précisons que la notion d'observabilité assure le caractère bien posé du problème de trajectographie passive.

Proposition 5.1. *Soit $\{S_\theta\}_{\theta\in\Theta}$ un modèle paramétrique et S_{θ^*} une trajectoire observable. Alors, notant*

$Y_{1:n} = \Phi(S_{\theta^*})(t_{1:n})$ et $Y = \Phi(S_{\theta^*})$, le problème inverse associé à

$$Y = \Phi(S_\theta) \quad ou \quad Y_{1:n} = \Phi(S_\theta)(t_{1:n}) \ ,$$

où θ à valeur dans Θ désigne la quantité d'intérêt est un problème inverse bien posé.

Démonstration. Traitons le cas à temps continu. Comme la trajectoire S_{θ^*} est observable, les deux premières conditions d'un problème bien posé sont clairement satisfaites. Supposons à présent qu'il existe une suite $\{\theta_n\}_{n\geq 0}$ à valeurs dans Θ telle que $\lim_{n\to\infty} \Phi(S_{\theta_n}) = \Phi(S_{\theta^*})$. Par compacité de Θ, la suite $\{\theta_n\}_{n\geq 0}$ admet une valeur d'adhérence $\theta' \in \Theta$. Alors, par continuité de Φ et de $\theta \mapsto S_\theta(t)$ à t fixé, il vient $\Phi(S_{\theta'}) = \Phi(S_{\theta^*})$ ce qui implique $\theta' = \theta^*$. Ensuite, si l'on considère une autre valeur d'adhérence $\tilde{\theta}$ de $\{\theta_n\}_{n\geq 0}$, on montre pour les mêmes raisons que $\tilde{\theta} = \theta^*$. La suite $\{\theta_n\}_{n\geq 0}$ admet donc θ^* pour unique valeur d'adhérence dans Θ compact et par suite $\lim_{n\to\infty} \theta_n = \theta^*$, ce qui clôt le caractère bien posé. □

La matrice d'information de Fisher, dans le cas du modèle d'observations (2), est définie par

$$I_n(\theta) \stackrel{\text{def}}{=} \sigma^{-2} \sum_{k=1}^{n} \nabla_\theta m(\theta, t_k) \nabla_\theta m(\theta, t_k)^{\mathrm{T}} \ , \tag{4}$$

et dans le cas du modèle d'observations (3) par

$$I(\theta) \stackrel{\text{def}}{=} \int_0^1 \nabla_\theta m(\theta, t) \nabla_\theta m(\theta, t)^{\mathrm{T}} \, \mathrm{d}t \ . \tag{5}$$

Ces matrices présentent une caractéristique qui va être abondamment exploitée pour discuter des cas d'observabilité et de non-observabilité.

Lemme 5.1. *les matrices de Fisher définies par les équations* (4) *et* (5) *sont des matrices de Gram.*

Démonstration. Pour tous $1 \leq i, \, j \leq d$, on a

$$I_n(\theta)_{i,j} = \sigma^{-2} \left\langle \frac{\partial m}{\partial \theta_i}(\theta, t_{1:n}), \frac{\partial m}{\partial \theta_j}(\theta, t_{1:n}) \right\rangle_{\mathbb{R}^n} \quad \text{et} \quad I(\theta)_{i,j} = \left\langle \frac{\partial m}{\partial \theta_i}(\theta, \cdot), \frac{\partial m}{\partial \theta_j}(\theta, \cdot) \right\rangle_{\mathrm{L}^2} \ .$$

□

Une conséquence immédiate de ce lemme est le critère d'inversibilité suivant :

Proposition 5.2. *On a les équivalences*

$$I_n(\theta) \text{ inversible} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta, t_{1:n}) \right)_{1\leq i\leq d} \text{ libre} \ , \tag{6}$$

et

$$I(\theta) \text{ inversible} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta, \cdot) \right)_{1\leq i\leq d} \text{ libre dans } \mathscr{C}([0,1]) \ . \tag{7}$$

L'introduction des matrices de Fisher dans cette partie dédiée à l'observabilité peut sembler gratuite. Il n'en est rien et on va montrer les liens existants entre ces notions. Le théorème de normalité asymptotique pour l'estimateur du maximum de vraisemblance est un premier élément de réponse puisque la variance, grandeur liée à la qualité de l'observabilité, est l'inverse de la matrice de Fisher. Définissons tout d'abord une notion d'observabilité restreinte dite *observabilité locale*.

Définition 5.5. *Pour le modèle d'observations à temps discret* (2)*, une trajectoire paramétrique S_{θ^*} est dite* localement observable *si il existe un voisinage $V_{\theta^*} \subset \Theta$ de θ^* tel que, pour tout $\theta \in V_{\theta^*}$,*

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \Longrightarrow \theta = \theta^* \ .$$

Pour le modèle d'observations à temps continu (3)*, une trajectoire paramétrique S_{θ^*} est dite* localement observable *si il existe un voisinage $V_{\theta^*} \subset \Theta$ de θ^* tel que, pour tout $\theta \in V_{\theta^*}$,*

$$\Phi(S_\theta) = \Phi(S_{\theta^*}) \Longrightarrow \theta = \theta' \ .$$

Rappelons un premier résultat établi dans [63].

Proposition 5.3. *Soit S_{θ^*} une trajectoire paramétrique. Dans le cas d'observations à temps discret, on a*

$$I_n(\theta^*) \text{ inversible} \Longrightarrow S_{\theta^*} \text{ localement observable} \ ,$$

et pour des observations à temps continu,

$$I(\theta^*) \text{ inversible} \Longrightarrow S_{\theta^*} \text{ localement observable} \ .$$

Démonstration. On propose deux schémas de preuves différents et plus courts que ceux décrits dans [63]. On se contente de détailler le cas à temps continu, la démonstration étant identique dans le cas discret. Posons

$$K_{\theta^*}(\theta) = \int_0^1 [m(\theta,t) - m(\theta^*,t)]^2 \, \mathrm{d}t \ .$$

Par dérivation sous l'intégrale, on obtient

$$\mathrm{D}_{\theta,\theta} \, K_{\theta^*}(\theta) = 2 \int_0^1 [m(\theta,t) - m(\theta^*,t)] \ \mathrm{D}_{\theta,\theta} \, m(\theta,t) \, \mathrm{d}t + 2I(\theta) \ .$$

En particulier, on a $\mathrm{D}_{\theta,\theta} \, K_{\theta^*}(\theta^*) = 2I(\theta^*)$. Or, la matrice de Fisher est, par définition, une matrice positive. Comme elle est supposée inversible, elle est donc définie positive. Alors, on sait par le résultat classique d'optimisation que θ^* est un minimum local strict de K_{θ^*}. Comme on a $K_{\theta^*}(\theta^*) = 0$, on en déduit le résultat attendu.

Variante pour la proposition 5.3. On peut également démontrer ce résultat sans passer par l'étude d'un extremum local de $K_{\theta^*}(\theta)$. D'après la formule de Taylor avec reste intégral, on a pour tout $t \in [0,1]$

$$m(\theta+h,t) = m(\theta,t) + \langle \nabla_\theta m(\theta,t), h \rangle_{\mathbb{R}^d} + \int_0^1 (1-u) \, \mathrm{D}_{\theta,\theta} \, m(\theta+uh,t)(h,h) \, \mathrm{d}u \ .$$

Ainsi, on en déduit que

$$|m(\theta+h,t) - m(\theta,t)| \geq \big| \, |\langle \nabla_\theta m(\theta,t), h \rangle_{\mathbb{R}^d}| - \mathrm{o}(\|h\|_{\mathbb{R}^d}) \big| \ ,$$

avec un $\mathrm{o}(\|h\|_{\mathbb{R}^d})$ uniforme en t. D'après le lemme 5.1, comme $I(\theta)$ est inversible, la famille des composantes de $\nabla_\theta m(\theta,\cdot)$ est libre et par suite, on a

$$\min_{(t,u) \in [0,1] \times \partial \bar{B}(0,1)} |\langle \nabla_\theta m(\theta,t), u \rangle_{\mathbb{R}^d}| > 0 \ .$$

Donc, il existe $\delta > 0$ tel que, pour tout $h \in B(0,\delta) \setminus \{0\}$, on a

$$|m(\theta + h, \cdot) - m(\theta, \cdot)| > 0 \ ,$$

ce qui prouve l'observabilité locale. □

Proposons un contre-exemple élémentaire à la réciproque de cette proposition avec des mobiles en mouvement[2]. Considérons le modèle de trajectoire paramétrique décrit par Θ un compact de $](-\frac{\pi}{2})^{1/3}, (\frac{\pi}{2})^{1/3}[$ contenant zéro et, pour $\theta \in \Theta$, $S_\theta(t) = (\theta^3 t, 1)$ pour tout $t \in [0,1]$. Soit la trajectoire porteur $O(t) = (t, 0)$. Il vient

$$m(\theta, \cdot) = m(\theta^*, \cdot) \iff \forall t \in [0,1] \quad \tan[\theta^3(t-1)] = \tan[\theta^{*3}(t-1)] \ ,$$

et clairement

$$m(\theta, \cdot) = m(\theta^*, \cdot) \iff \theta^3 = \theta^{*3} \iff \theta = \theta^* \ .$$

La trajectoire S_{θ^*} est donc observable. Or, on a

$$\frac{\partial m}{\partial \theta}(\theta, t) = 3\theta^2(t-1)\left\{1 + \tan^2[\theta^3(t-1)]\right\} \ ,$$

et pour $\theta^* = 0$, l'information de Fisher n'est pas inversible bien que la trajectoire soit observable.

On peut se demander s'il existe un cadre de modèles paramétriques dans lequel on peut étendre le résultat fourni par la proposition 5.3, en passant d'un caractère local à global et en rendant valide sa réciproque. C'est le cas pour des trajectoires dont les composantes s'écrivent dans une base de fonctions. Notons $(e_j)_{1 \le j \le p}$ une famille libre de fonctions de $\mathscr{C}([0,1])$.

Proposition 5.4. *Soit le modèle de trajectoire paramétrique vérifiant* (H_+) *décrit par* Θ *un compact de* $\mathbb{R}^{2p}$ *et*

$$S_\theta = \left(\sum_{j=1}^{p} a_j e_j, \sum_{j=1}^{p} b_j e_j\right) \quad \textit{avec} \quad \theta = (a_{1:p}, b_{1:p}) \in \Theta \ .$$

Pour le modèle d'observations à temps discret (2) *avec* $n \ge 2p$*, on a l'équivalence suivante*

$$S_{\theta^*} \textit{ observable} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta^*, t_{1:n})\right)_{1 \le i \le 2p} \textit{ libre} \ , \tag{8}$$

et par conséquent

$$S_{\theta^*} \textit{ observable} \iff I_n(\theta^*) \textit{ inversible} \ .$$

De même, pour le modèle d'observations à temps continu (3)*, on a*

$$S_{\theta^*} \textit{ observable} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta^*, \cdot)\right)_{1 \le i \le 2p} \textit{ libre} \ ,$$

et par suite

$$S_{\theta^*} \textit{ observable} \iff I(\theta^*) \textit{ inversible} \ .$$

[2]On pourrait faire encore plus simple avec un porteur immobile mais ce n'est pas l'esprit de la trajectographie passive.

Démonstration. On a

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \iff \sum_{j=1}^{p} a_j e_j(t_{1:n}) - \sum_{j=1}^{p} b_j(e_j \times \tan m(\theta^*, \cdot))(t_{1:n}) = (O_1 - O_2 \times \tan m(\theta^*, \cdot))(t_{1:n}) \ ,$$

et par suite

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \iff A_{\theta^*}(t_{1:n})(\theta - \theta^*) = 0 \ , \tag{9}$$

où $A_{\theta^*}(t_{1:n})$ désigne la matrice de $\mathbb{R}^{n\times 2p}$ dont les lignes sont constituées des vecteurs $A_{\theta^*}(t_k)$ pour k variant de 1 à n avec

$$A_{\theta^*}(t) = \begin{pmatrix} e_1(t) & \cdots & e_p(t) & e_1(t) \times \tan m(\theta^*, t) & \cdots & e_p(t) \times \tan m(\theta^*, t) \end{pmatrix} \ . \tag{10}$$

D'après le théorème du rang, on a

$$\operatorname{rg} A_{\theta^*}(t_{1:n}) + \dim \operatorname{Ker} A_{\theta^*}(t_{1:n}) = 2p \ . \tag{11}$$

Comme $\operatorname{rg} A_{\theta^*}(t_{1:n}) \leq n$, la condition $n \geq 2p$ est nécessaire à l'injectivité de l'application canoniquement associée à $A_{\theta^*}(t_{1:n}) \leq n$. Par suite, comme $n \geq 2p$, on en déduit que

$$\Big(\forall \theta \in \Theta \quad \Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \Longrightarrow \theta = \theta^*\Big) \iff \operatorname{rg} A_{\theta^*}(t_{1:n}) = 2p \ .$$

Par dérivation, on obtient

$$\nabla_\theta m(\theta, t) = [S_{\theta^*,2}(t) - O_2(t)] \, \|S_\theta(t) - O(t)\|^{-2} \times A_{\theta^*}(t) \ , \tag{12}$$

et par suite

$$\operatorname{rg} \left(\frac{\partial m}{\partial \theta_i}(\theta^*, t_{1:n}) \right)_{1\leq i\leq 2p} = \operatorname{rg} A_{\theta^*}(t_{1:n}) \ ,$$

ce qui prouve le résultat pour le cas discret. Dans le cas continu, on a

$$\Phi(X_\theta) = \Phi(X_{\theta^*}) \iff A_{\theta^*}(\cdot)(\theta - \theta^*) = 0 \ . \tag{13}$$

Donc, d'après la définition (10) de $A_{\theta^*}(t)$, on a

$$\Big(A_{\theta^*}(\cdot)(\theta - \theta^*) = 0 \Longrightarrow (\theta - \theta^*) = 0\Big) \iff \big(e_1, \cdots, e_p, e_1 \times \tan m(\theta^*, \cdot), \cdots, e_p \times \tan m(\theta^*, \cdot)\big) \text{ libre} \ .$$

D'après l'expression (12) du gradient de $\nabla_\theta m(\theta, t)$, on a

$$\big(e_1, \cdots, e_p, e_1 \times \tan m(\theta^*, \cdot), \cdots, e_p \times \tan m(\theta^*, \cdot)\big) \text{ libre} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta, \cdot) \right)_{1\leq i\leq 2p} \text{ libre} \ ,$$

ce qui clôt la preuve pour le cas continu. □

Typiquement, la famille de fonctions $(e_j)_{1\leq j\leq p}$ considérée sera une famille de polynômes, comme pour le MRU par exemple. Le résultat de la proposition 5.4 est utile car il permet de discuter de l'observabilité dans un certain type de modèle paramétrique avec des arguments simples

d'algèbre linéaire. Il permet également de décrire l'ensemble des trajectoires paramétriques ambiguës indistiguables de la trajectoire cible. On reprend la notation de $A_{\theta^*}(t)$ définie par l'équation (10) et celle de $A_{\theta^*}(t_{1:n})$, matrice de $\mathbb{R}^{n\times 2p}$ dont les lignes sont constituées des vecteurs $A_{\theta^*}(t_k)$.

Proposition 5.5. *Soit le modèle de trajectoire paramétrique vérifiant* (H_+) *décrit par* Θ *un compact de* $\mathbb{R}^{2p}$ *et*

$$S_\theta = \left(\sum_{j=1}^{p} a_j e_j, \sum_{j=1}^{p} b_j e_j\right) \quad \textit{avec} \quad \theta = (a_{1:p}, b_{1:p}) \in \Theta \ .$$

Pour des observations à temps discret (2)*, si* S_{θ^*} *n'est pas observable, alors l'ensemble des trajectoires paramétriques ayant les mêmes mesures angulaires que* S_{θ^*} *est décrit par l'espace affine*

$$\Sigma_{n,\theta^*} = \{S_\theta,\ \theta \in \theta^* + \operatorname{Ker} A_{\theta^*}(t_{1:n})\} \ .$$

Pour des observations à temps continu (3)*, l'ensemble des trajectoires conformes en mesures angulaires à une trajectoire non observable* S_{θ^*} *est décrit par l'espace affine*

$$\Sigma_{\theta^*} = \{S_\theta,\ \theta \in \theta^* + \operatorname{Ker} A_{\theta^*}\} \ .$$

Démonstration. C'est une conséquence directe de la preuve de la proposition 5.4. D'après (9), il vient

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \iff \theta \in \theta^* + \operatorname{Ker} A_{\theta^*}(t_{1:n}) \ ,$$

et d'après (13), on a

$$\Phi(S_\theta)(t_{1:n}) = \Phi(S_{\theta^*})(t_{1:n}) \iff \theta \in \theta^* + \operatorname{Ker} A_{\theta^*} \ .$$

La structure affine de Σ_{n,θ^*} en découle, par linéarité de l'application $\theta \mapsto S_\theta$.

Les arguments valent aussi pour le temps continu. □

Ainsi, étant donnée une trajectoire S_{θ^*} non observable, l'ensemble des trajectoires conformes en mesures angulaires à S_{θ^*} est infini. Plus précisément, d'après le théorème du rang, c'est un espace affine de dimension au plus égale à $2p$. On peut affiner cette majoration si l'on a plus d'information sur la famille $(e_j)_{1\le j\le p}$ et sur le choix de la séquence $t_{1:n}$. Par exemple, si $e_j(t) = t^j$ et $t_k = k/n$, on voit que pour tout $n \ge p$, on peut extraire de $A_{\theta^*}(t_{1:n})$ une matrice de Van der Monde de taille $p \times p$ et il vient

$$1 \le \dim \operatorname{Ker} A_{\theta^*}(t_{1:n}) \le 2p - p = p \ .$$

On peut signaler les travaux de Joannides et Le Gland [68] qui décrivent, dans un cadre bayésien, le comportement asymptotique de la loi *a posteriori* pour des modèles non-identifiables. Dans le cas d'observations à temps continu, notons μ la loi *a priori* sur Θ et μ^ε la loi *a posteriori*. D'après [68, Theorem 5], la loi μ^ε converge faiblement vers une distribution de probabilité aléatoire μ_{θ^*} qui est supportée par l'ensemble M_{θ^*} des paramètres indistinguables de la vraie valeur, c'est-à-dire $\theta^* + \operatorname{Ker} A_{\theta^*}$ dans le cas présent, et cette loi μ_{θ^*} possède une densité par rapport à la mesure de Lebesgue sur l'ensemble $\theta^* + \operatorname{Ker} A_{\theta^*}$.

En pratique, il est fondamental de savoir si l'on peut manœuvrer pour rendre la trajectoire cible observable. Il n'est pas aisé de répondre à cette question sans un choix explicite de la base $(e_j)_{1\le j\le p}$. Cependant, on peut détailler une classe de trajectoires porteurs qui ne rendent jamais la trajectoire cible observable et qui sont donc à proscrire.

Proposition 5.6. *Soit le modèle de trajectoire paramétrique décrit par Θ un compact de $\mathbb{R}^{2p}$ et*

$$S_\theta = \left(\sum_{j=1}^{p} a_j e_j, \sum_{j=1}^{p} b_j e_j \right) \quad \textit{avec} \quad \theta = (a_{1:p}, b_{1:p}) \in \Theta \ .$$

Aucune des trajectoires porteur s'écrivant

$$O = \left(\sum_{j=1}^{p} \alpha_j e_j, \sum_{j=1}^{p} \beta_j e_j \right) \quad \textit{avec} \quad (\alpha_k, \beta_k)_{1 \le k \le p} \in \mathbb{R}^{2p} \ ,$$

et vérifiant la condition (H_+) ne rend une trajectoire S_{θ^*} avec $\theta^* \in \Theta$ observable.

Démonstration. Écrivons la preuve dans le cas continu. D'après la proposition 5.4 et les équations (10) et (12), on a

$$S_{\theta^*} \text{ observable} \iff \left(\left(\sum_{j=1}^{n} (b_k - \beta_k)\, e_k\, e_i \right)_{1 \le i \le p}, \left(\sum_{j=1}^{n} (a_k - \alpha_k)\, e_k\, e_i \right)_{1 \le i \le p} \right) \text{ libre} \ .$$

Or, on a clairement

$$\sum_{i=1}^{p} (a_i - \alpha_i) \sum_{j=1}^{n} (b_j - \beta_j)\, e_j\, e_i - \sum_{i=1}^{p} (b_i - \beta_i) \sum_{j=1}^{n} (a_j - \alpha_j)\, e_j\, e_i = 0 \ ,$$

donc S_{θ^*} n'est pas observable. La preuve est identique dans le cas discret. □

Un cas particulier déjà connu de cette proposition est celui où cible et porteur sont en MRU [66]. Ce résultat complète aussi les travaux de Fogel et al.[51] où les auteurs montrent que, si la cible et le porteur ont des trajectoires à composantes polynomiales, alors la cible est observable si les composantes du porteur sont de degré supérieur à celles de la cible. On peut proposer une interprétation intuitive de ce résultat. Si la trajectoire porteur s'écrit dans la même base que la trajectoire cible, alors cible et porteur ont des trajectoires de complexité équivalente et les mesures angulaires sont insuffisantes pour extraire la trajectoire cible. Il faut donc, d'une certaine manière, que le porteur opère une manœuvre de complexité supérieure à celles effectuées par la cible pour qu'il acquière de l'information et puisse déterminer la trajectoire cible.

5.3 Le mouvement rectiligne uniforme

Dans cette section et la suivante, les preuves des différents énoncés sont présentées dans le cas d'observations en temps continu. On peut en déduire le cas des observations discrètes quand les arguments invoqués s'appuient uniquement sur des propriétés de liberté de famille de fonctions.

Proposition 5.7. *Soient $(f_1, \ldots, f_d)$ une famille libre de fonctions continues définies de $[0,1]$ dans $\mathbb{R}$ et $t_{1:n}$ la suite de points définie par $t_k = k/n$. Alors, pour n suffisamment grand, la famille de vecteurs $(f_1(t_{1:n}), \ldots, f_d(t_{1:n}))$ est une famille libre de $\mathbb{R}^n$.*

Démonstration. Considérons les matrices de Gram J_n et J définies, pour tous $1 \le i,\, j \le d$, par

$$(J_n)_{i,j} = \langle f_i(t_{1:n}), f_j(t_{1:n}) \rangle_{\mathbb{R}^n} \quad \text{et} \quad (J)_{i,j} = \langle f_i, f_j \rangle_{\mathrm{L}^2} \ .$$

On a clairement $\lim_{n\to\infty} n^{-1} J_n = J$. Par suite, s'agissant de matrices de Gram et la famille $(f_1, \ldots, f_d)$ étant libre, on a $\det J \neq 0$. Par continuité du déterminant, on conclut que, pour n assez grand, la matrice J_n est inversible, c'est-à-dire la famille de vecteurs $(f_1(t_{1:n}), \ldots, f_d(t_{1:n}))$ est libre. □

En fait, on sait, par des arguments d'algèbre linéaire, qu'il existe d points distincts $s_1, \ldots, s_d$ tels que la matrice de terme général $f_i(s_j)$ soit inversible. En revanche, on ne sait pas concrètement comment choisir de tels points. Pour une approche opérationnelle, le résultat de la proposition 5.7 est plus utile puisque, pour un nombre d'observations suffisant grand, le porteur a la garantie que la condition de liberté est satisfaite. On laisse au lecteur le soin d'appliquer cette proposition pour prolonger, dans le cas d'observations discrètes, les résultats démontrés ci-après.

Le modèle paramétriques le plus étudié dans la littérature du BOT est le mouvement rectiligne uniforme (MRU). L'ensemble des paramètres admissibles Θ est un compact de $\mathbb{R}^4$ et le paramètre d'une trajectoire est noté $\theta = (x_0, y_0, v_x, v_y)$ où (x_0, y_0) désigne la position initiale de la cible et (v_x, v_y) son vecteur vitesse. L'équation d'une trajectoire S_θ pour $\theta \in \Theta$ est donnée, pour tout $t \in [0, 1]$, par

$$S_\theta(t) = (S_{\theta,1}(t), S_{\theta,2}(t)) = (x_0 + v_x t,\ y_0 + v_y t)\ .$$

Retrouvons le critère d'observabilité établi par Aidala et al.[3] avec des arguments simples d'algèbre linéaire.

Proposition 5.8. *Soit le modèle de trajectoire paramétrique MRU caractérisé par Θ un compact de $\mathbb{R}^4$ tel que la condition (H_+) soit satisfaite. Alors, pour des observations à temps discret avec $n \geq 4$, on a*

$$S_{\theta^*} \text{ non observable} \iff \operatorname{rg} \begin{pmatrix} 1 & t_1 & \tan\beta(t_1) & t_1 \tan\beta(t_1) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & \tan\beta(t_n) & t_n \tan\beta(t_n) \end{pmatrix} < 4\ ,$$

et, pour des observations à temps continu,

$$S_{\theta^*} \text{ non observable} \iff (1, \mathrm{id}, \tan\beta(\cdot), \mathrm{id} \times \tan\beta(\cdot)) \text{ liée}\ .$$

Démonstration. D'après la proposition 5.4, on a en temps discret

$$S_{\theta^*} \text{ non observable} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta, t_{1:n})\right)_{1\leq i\leq 4} \text{ liée}\ ,$$

et en temps continu

$$S_{\theta^*} \text{ observable} \iff \left(\frac{\partial m}{\partial \theta_i}(\theta, \cdot)\right)_{1\leq i\leq 4} \text{ liée}\ .$$

Par dérivation, on obtient par

$$\nabla_\theta m(\theta, t) = [S_{\theta,2}(t) - O_2(t)]\, \|S_\theta(t) - O(t)\|^{-2} \times \begin{pmatrix} 1 & t & \tan m(\theta, t) & t \times \tan m(\theta, t) \end{pmatrix}\ ,$$

et le résultat suit. □

La démarche proposée ci-avant nous semble plus simple et plus naturelle que celle déployée par Aidala et al.[3] où le même critère est obtenu en établissant puis en résolvant une équation différentielle d'ordre 3.

D'après la proposition 5.6, on sait que si le porteur suit également un MRU, alors la trajectoire cible n'est pas observable.

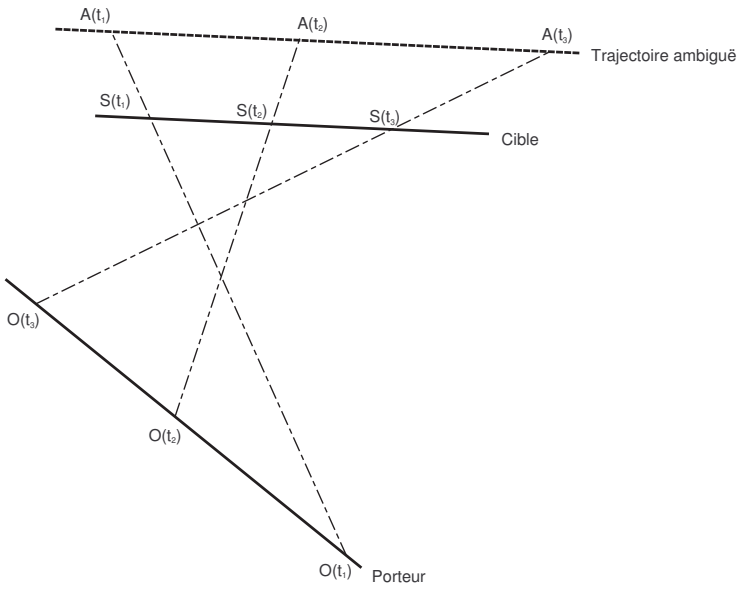


FIG. 5.1 – Cible non observable : cible et porteur en MRU

Concrètement, pour rendre un MRU observable, le porteur doit effectuer une manœuvre. Intuitivement, il paraît clair que plus la manœuvre est franche, plus le niveau d'observabilité sera bon. Toutefois, il se peut que la manœuvre du porteur ne permette pas de discerner toutes les trajectoires MRU.

Proposition 5.9. *On peut avoir simultanément une trajectoire S_{θ^*} observable et une autre trajectoire $S_{\theta'}$ non observable.*

Démonstration. On peut supposer, sans perte de généralité, que $(1,1,1,1)$ et $(0,1,1,1)$ sont éléments de Θ. Il suffit alors de considérer

$$S_{\theta^*}(t) = (t, 1+t) \ , \quad S_{\theta'}(t) = (1+t, 1+t) \ , \quad O(t) = (1+t+t^2, 1+t-t^2) \ .$$

□

On touche du doigt la frontière qui sépare l'observabilité de l'identifiabilité du modèle sous-jacent. Ce résultat montre que le modèle statistique associé au problème de régression paramétrique n'est, en général, pas identifiable et précise la différence entre la notion d'observabilité qui est relative à une valeur du paramètre et la notion d'identifiabilité qui couvre toutes les valeurs possibles du paramètre. Peut-on alors manœuvrer pour que le modèle associé aux trajectoires MRU soit identifiable ?

Pour cela, on va considérer un autre modèle paramétrique incontournable en trajectographie passive : le mouvement circulaire uniforme (MCU). C'est le mouvement que suit un mobile quand il effectue un virage avec une accélération de giration constante. Un mobile en MCU admet une trajectoire d'équation

$$X(t) = (x_0 + r\cos(\omega t + \alpha),\, y_0 + r\sin(\omega t + \alpha)) \ ,$$

pour tout $t \in [0,1]$. L'étude de ce modèle paramétrique du point de vue de la cible est présentée à la section 5.4. On s'intéresse ici à la situation où le porteur est en MCU.

Proposition 5.10. *Quand le porteur n'est pas immobile et suit une trajectoire MCU , le modèle statistique d'une cible dont la trajectoire est à composantes polynomiales vérifiant* (H_+) *est identifiable.*

Démonstration. Considérons deux trajectoires paramétriques S_θ et $S_{\theta'}$ avec les mêmes mesures angulaires. On note

$$S_\theta(t) = \left(\sum_{j=0}^{p} a_j t^j, \sum_{j=0}^{p} b_j t^j\right) \quad , \quad S_{\theta'}(t) = \left(\sum_{j=0}^{p} a'_j t^j, \sum_{j=0}^{p} b'_j t^j\right) \ ,$$

et $O(t) = (x_0 + r\cos(\omega t + \alpha),\ y_0 + r\sin(\omega t + \alpha))$. On peut supposer, sans perte de généralité, que $(x_0, y_0, r, \omega, \alpha) = (0, 0, 1, 1, 0)$. Pour tout $t \in [0,1]$, on a

$$m(\theta, t) = m(\theta', t) \iff$$
$$\left(\sum_{j=0}^{p} a_j t^j - \cos t\right)\left(\sum_{j=1}^{p} b'_j t^j - \sin t\right) = \left(\sum_{j=0}^{p} a'_j t^j - \cos t\right)\left(\sum_{j=1}^{p} b_j t^j - \sin t\right) \ .$$

En développant, il vient

$$m(\theta, t) = m(\theta', t) \iff$$
$$\sum_{j=0}^{p} a_j t^j \times \sum_{j=1}^{p} b'_j t^j - \sum_{j=0}^{p} a'_j t^j \times \sum_{j=1}^{p} b_j t^j + \sum_{j=0}^{p} (a'_j - a_j) t^j \sin t + \sum_{j=0}^{p} (b'_j - b_j) t^j \cos t = 0 \ .$$

On se convainc sans peine que la famille $\big((t^j)_{0 \le j \le 2p}, (t^j \cos t, t^j \sin t)_{0 \le j \le p}\big)$ est libre et il en résulte clairement que $\theta = \theta'$. □

Par conséquent, quand le porteur effectue une manœuvre à giration constante, il rend toute cible MRU observable.

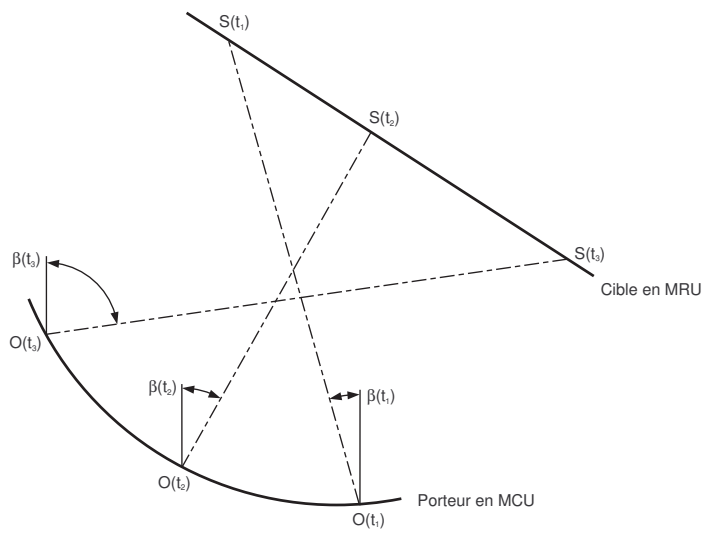


Fig. 5.2 – Cible observable : cible en MRU et porteur en MCU

Bien évidemment, il suffit que le porteur ait effectué un MCU durant la période d'observation et pas obligatoirement sur toute la période d'observation pour qu'une cible en MRU lui soit observable. Ainsi, un porteur dont la trajectoire serait la succession d'un MRU puis d'un MCU puis d'un MRU aurait l'observabilité d'une cible en MRU.

Précisons que, d'après la proposition 5.7, le résultat de la proposition 5.10 vaut pour le cas d'observations à temps discret avec un nombre d'observations suffisamment grand.

5.4 Le mouvement circulaire uniforme

Le mouvement circulaire uniforme (MCU) est l'autre modèle élémentaire fondamental de trajectoire paramétrique rencontré en trajectographie passive. L'équation d'une trajectoire est donnée par

$$S_\theta(t) = (x_0 + r\cos(\omega t + \alpha),\ y_0 + r\sin(\omega t + \alpha))\ ,$$

avec $\theta = (x_0, y_0, r, \omega, \alpha) \in \Theta$ où Θ désigne un compact de $\mathbb{R}^5$. À la différence du MRU, il n'existe pas de bases de fonctions appropriées à l'écriture du MCU. On ne peut donc pas invoquer les résultats des propositions 5.4, 5.5 et 5.6. Néanmoins, le porteur peut manœuvrer de sorte que le problème de régression paramétrique correspondant ait de bonnes propriétés.

Les preuves des résultats qui suivent sont fournies en section 5.6.

Proposition 5.11. *Pour un porteur non immobile dont la trajectoire est à composantes polynomiales, une cible en MCU admet une matrice de Fisher inversible.*

Signalons que le cas discret en découle, pour un nombre d'observations suffisamment grand.

D'après le résultat de la proposition 5.3, pour un porteur en MRU, une cible en MCU est localement observable. Ceci interdit clairement le cas du porteur stationnaire puisqu'il n'y a pas observabilité locale, comme l'illustre la figure 5.3.

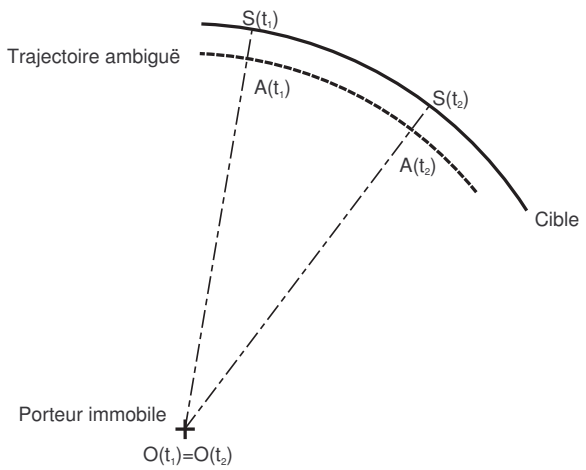


Fig. 5.3 – Non observabilité locale : cible en MCU et porteur immobile

En fait, on a beaucoup mieux que l'observabilité locale.

Proposition 5.12. *Quand le porteur suit une trajectoire non stationnaire et à composantes polynomiales, le modèle statistique d'une cible en MCU est identifiable.*

Signalons que le cas discret n'en découle pas car la preuve s'appuie sur un argument utilisant des séries entières qui ne semble pas se transposer pour des instants discrets.

On a vu en section 5.2 que, pour une cible dont la trajectoire est à composantes polynomiales, un porteur dont la trajectoire a les mêmes caractéristiques n'acquiert pas l'observabilité de la cible. Le comportement est assez différent dans le cas du MCU.

Proposition 5.13. *Pour un porteur avec une trajectoire MCU, une cible en MCU à vitesse angulaire différente de celle du porteur admet une matrice de Fisher inversible.*

Ainsi, d'après la proposition 5.3 pour une cible et un porteur en MCU avec des vitesses angulaires différentes, la cible est localement observable. Là encore, on dispose d'un résultat plus fort.

Proposition 5.14. *Quand le porteur suit une trajectoire MCU, le modèle statistique d'une cible en MCU à vitesse angulaire différente de celle du porteur est identifiable.*

D'après les propositions 5.10 et 5.12, une cible en MRU est observable par un porteur en MCU mais la réciproque est également vraie, à savoir, le porteur est observable par la cible, sous réserve de faire l'hypothèse d'une dynamique MCU. La question qui en découle est : le porteur peut-il observer la cible et préserver sa discrétion ?

Proposition 5.15. *L'observabilité n'est pas une notion symétrique : la cible peut être observable par le porteur sans que lui-même ne soit observable par la cible.*

Ainsi, le porteur peut préserver sa discrétion, d'une part du fait de la nature de la mesure qu'il réalise, mesure dite passive puisque le porteur n'émet pas de signal et d'autre part, en suivant une trajectoire non observable par la cible si celle-ci réalise des mesures semblables à celles du porteur.

5.5 Conclusion

Dans ce premier chapitre, on présente des résultats concrets d'observabilité pour des cibles avec des dynamiques réalistes comme le MRU ou le MCU. Les rapports reliant observabilité et inversibilité de la matrice de Fisher pour le modèle de régression sont étudiés et généralisés pour de large classes de trajectoire cible. Le cas pathologique où la cible n'est pas observable est détaillé et on décrit l'espace des trajectoires ambiguës. Enfin, des manœuvres réalistes rendant observables les dynamiques usuelles sont présentées et permettent dès lors une approche rigoureuse au problème d'estimation.

La question de l'observabilité n'est pas caractéristique de la trajectographie passive mono-porteur. En effet, le cas de deux mobiles porteurs n'est pas automatiquement un problème bien posé. Pour des observations discrètes, les temps d'observations pour chaque porteur sont notés respectivement

$$I_s = \{s_0 + k\Delta_s,\ k \in \mathbb{N}\}\ , \qquad I_t = \{t_0 + l\Delta_t,\ l \in \mathbb{N}\}\ .$$

Sans hypothèse restrictive sur la nature de la trajectoire et en supposant $I_s \cap I_t = \emptyset$, le problème s'inscrit, pour les mêmes raisons que celles du cas mono-porteur, dans la catégorie des problèmes inverses mal posés. Toutefois, sous une hypothèse de trajectoire paramétrique, l'estimation s'avère particulièrement robuste et ne souffre nullement des faiblesses des méthodes *batch* ou en-ligne. Si l'on veut que la fonction $\Psi(\cdot, t_k)$ soit inversible, autrement dit si on veut concrètement pouvoir faire de la triangulation, il faut $I_s \cap I_t \neq \emptyset$. On vérifie sans difficulté que cette condition est satisfaite si et seulement si le rapport des fréquences d'échantillonnage $\frac{\Delta_s}{\Delta_t}$ est un rationnel et si $s_0 - t_0$ est également dans un rapport rationnel avec une des deux fréquences d'échantillonnage.

5.6 Preuves de la section 5.4

Preuve de la proposition 5.11. Pour plus de souplesse, on notera

$$R_1(\theta,t) = x_0 + r\cos(\omega t+\alpha) - O_1(t)\ , \quad R_2(\theta,t) = y_0 + r\sin(\omega t+\alpha) - O_2(t)\ .$$

Ainsi, on a

$$\begin{aligned}\nabla_\theta m(\theta,t) = \|S_\theta(t) - O(t)\|^{-2} \times \\ \big(R_2(\theta,t),\quad -R_1(\theta,t),\quad \cos(\omega t+\alpha)R_2(\theta,t) - \sin(\omega t+\alpha)R_1(\theta,t),\quad -tc(\theta,t),\quad -c(\theta,t)\big)\end{aligned}$$

où

$$c(\theta,t) = r\sin(\omega t+\alpha)R_2(\theta,t) + r\cos(\omega t+\alpha)R_1(\theta,t)\ .$$

On en déduit que

$$\begin{aligned}&\left(\frac{\partial m}{\partial \theta_i}(\theta,\cdot)\right)_{1\le i\le 5} \text{ libre } \iff \\ &\quad\Big(\forall t\in[0,1]\quad R_2(\theta,t)\big(\lambda_1+\lambda_3\cos(\omega t+\alpha)+(\gamma_1+\gamma_2 t)\theta_3\sin(\omega t+\alpha)\big)+ \\ &\qquad R_1(\theta,t)\big(\lambda_2+\lambda_3\sin(\omega t+\alpha)+(\gamma_1+\gamma_2 t)\theta_3\cos(\omega t+\alpha)\big) = 0 \Longrightarrow (\lambda_i,\gamma_i)_i = 0\Big)\ .\end{aligned}$$

On peut désormais supposer, sans perte de généralité, que $\theta = (0,0,1,1,0)$. Le membre de droite de l'équivalence ci-dessus, pour un porteur à composantes polynomiales, peut s'écrire sous la forme synthétique suivante

$$[\sin t + Q(t)][\lambda_1 + \lambda_3\cos t + (\gamma_1+\gamma_2 t)\sin t] + [\cos t + P(t)][\lambda_2 + \lambda_3 \sin t + (\gamma_1+\gamma_2 t)\cos t] = 0\ ,$$

avec P et Q des polynômes de degré inférieur ou égal à p qui ne sont pas tous deux constants. De manière équivalente, on a

$$\begin{aligned}\gamma_1+\gamma_2 t + \lambda_1 Q(t) + \lambda_2 P(t) + \lambda_3\sin(2t) + \lambda_3[Q(t)\cos t + P(t)\sin t] + \\ \lambda_1\sin t + \lambda_2\cos t + (\gamma_1+\gamma_2 t)[Q(t)\sin t + P(t)\cos(t)] = 0\ .\end{aligned}$$

On se convainc sans peine que la famille de fonctions

$$\Big(\sin(2t), (t^k)_{0\le k\le p}, (t^k\cos t)_{0\le k\le p+1}, (t^k\sin t)_{0\le k\le p+1}\Big)$$

est libre. On en déduit d'abord que $\lambda_3 = 0$, puis

$$\lambda_1\sin t + \lambda_2\cos t + (\gamma_1+\gamma_2 t)[Q(t)\sin t + P(t)\cos(t)] = 0\ . \tag{14}$$

Comme P ou Q n'est pas contant, on a nécessairement $\gamma_1 = \gamma_2 = 0$ puis $\lambda_1 = \lambda_2 = 0$. □

Preuve de la proposition 5.12. Considérons deux trajectoires paramétriques S_θ et $S_{\theta'}$ avec les mêmes mesures angulaires. On note

$$\theta = (x_0, y_0, r, \omega, \alpha)\ , \quad \theta' = (x_0', y_0', r', \omega', \alpha')\ ,$$

et $O(t) = (P(t), Q(t))$. Pour tout $t \in [0,1]$, on a

$$\begin{aligned} m(\theta,t) = m(\theta',t) \iff & \\ & [x_0 + r\cos(\omega t+\alpha) - P(t)][y_0' + r'\sin(\omega' t+\alpha') - Q(t)] = \\ & \qquad [y_0 + r\sin(\omega t+\alpha) - Q(t)][x_0' + r'\cos(\omega' t+\alpha') - P(t)] \ . \end{aligned}$$

En développant, il vient

$$\begin{aligned} x_0y_0' - y_0x_0' + (x_0' - x_0)Q(t) + (y_0 - y_0')P(t) + & \\ r'\sin(\omega' t+\alpha')[x_0 - P(t)] + r\cos(\omega t+\alpha)[y_0 - Q(t)] & \\ -r\sin(\omega t+\alpha)[x_0' - P(t)] - r'\cos(\omega' t+\alpha')[y_0' - Q(t)] = 0 \ . & \end{aligned}$$

On se convainc aisément que les familles $\left(t^k\right)_{0\le k\le p}$ et $\left(t^k\cos(\omega t+\alpha), t^k\sin(\omega t+\alpha), t^k\cos(\omega' t+\alpha'), t^k\sin(\omega' t+\alpha')\right)_{0\le k\le p}$ sont linéairement indépendantes. On en déduit que

$$\begin{aligned} r'\sin(\omega' t+\alpha')[x_0 - P(t)] + r\cos(\omega t+\alpha)[y_0 - Q(t)] & \\ -r\sin(\omega t+\alpha)[x_0' - P(t)] - r'\cos(\omega' t+\alpha')[y_0' - Q(t)] = 0 \ . & \end{aligned} \tag{15}$$

Si $\omega \neq \omega'$, alors la famille $\left(t^k\cos(\omega t+\alpha), t^k\sin(\omega t+\alpha), t^k\cos(\omega' t+\alpha'), t^k\sin(\omega' t+\alpha')\right)_{0\le k\le p}$ est libre ce qui implique, pour tout $t \in [0,1]$,

$$P(t) = x_0 = x_0' \ , \quad Q(t) = y_0 = y_0' \ .$$

Comme le porteur est supposé non stationnaire, l'égalité ci-dessus n'a pas lieu et par suite, $\omega = \omega'$. On peut supposer, sans perte de généralité, que $\omega = 1$. Reformulons (15). On a pour tout $t \in [0,1]$

$$[r\sin(t+\alpha) - r'\sin(t+\alpha')]P(t) + [r'\cos(t+\alpha') - r\cos(t+\alpha)]Q(t) = c(t) \ ,$$

avec c une fonction 2π périodique. Comme les membres de cette égalité sont développables en série entière, on peut la prolonger à $\mathbb{R}$ tout entier. Pour $t = -\alpha' + 2\pi k$ avec k entier, on obtient

$$r\sin(\alpha-\alpha')P(-\alpha'+2\pi k) + [r' - r\cos(\alpha-\alpha')]Q(-\alpha'+2\pi k) = c(-\alpha') \ ,$$

Par conséquent, le polynôme $r\sin(\alpha-\alpha')P + [r' - r\cos(\alpha-\alpha')]Q$ est constant et égal à

$$r\sin(\alpha-\alpha')P + [r' - r\cos(\alpha-\alpha')]Q = c(-\alpha') \ .$$

Puis, en prenant $t = -\alpha + 2\pi k$ avec k entier, on obtient

$$r'\sin(\alpha-\alpha')P + [r'\cos(\alpha-\alpha') - r]Q = c(-\alpha) \ .$$

On a le système d'équations en P et Q suivant

$$\left\{ \begin{array}{rcl} r\sin(\alpha-\alpha')P + [r' - r\cos(\alpha-\alpha')]Q & = & c(-\alpha') \ , \\ r'\sin(\alpha-\alpha')P + [r'\cos(\alpha-\alpha') - r]Q & = & c(-\alpha) \ . \end{array} \right.$$

On sait, par hypothèse, que P et Q ne sont pas tous deux polynômes constants. Donc, le système

ne doit pas être un système de Cramer ou, de manière équivalente,

$$\begin{vmatrix} r\sin(\alpha-\alpha') & r'-r\cos(\alpha-\alpha') \\ r'\sin(\alpha-\alpha') & r'\cos(\alpha-\alpha')-r \end{vmatrix} = 0 \ .$$

Le calcul donne

$$\begin{aligned} \begin{vmatrix} r\sin(\alpha-\alpha') & r'-r\cos(\alpha-\alpha') \\ r'\sin(\alpha-\alpha') & r'\cos(\alpha-\alpha')-r \end{vmatrix} &= -r^2 - r'^2 + 2rr'\cos(\alpha-\alpha') \ , \\ &= \underbrace{-(r-r')^2}_{\leq 0} \ \underbrace{-2rr'[1-\cos(\alpha-\alpha')]}_{\leq 0} \ , \end{aligned}$$

et finalement

$$\begin{vmatrix} r\sin(\alpha-\alpha') & r'-r\cos(\alpha-\alpha') \\ r'\sin(\alpha-\alpha') & r'\cos(\alpha-\alpha')-r \end{vmatrix} = 0 \iff (r,\alpha) = (r',\alpha') \ .$$

En remplaçant dans (15), il vient clairement $(x_0, y_0) = (x'_0, y'_0)$ ce qui clôt l'observabilité du MCU. □

Preuve de la proposition 5.13. A l'instar de ce qui est fait dans la preuve de la proposition 5.11, on a

$$\begin{aligned} &\left(\frac{\partial m}{\partial \theta_i}(\theta,\cdot)\right)_{1\leq i\leq d} \text{ libre} \iff \\ &\quad \Big(\forall t\in[0,1] \quad R_2(\theta,t)[\lambda_1+\lambda_3\cos(\omega t+\alpha)+(\gamma_1+\gamma_2 t)r\sin(\omega t+\alpha)]+ \\ &\qquad R_1(\theta,t)[\lambda_2+\lambda_3\sin(\omega t+\alpha)+(\gamma_1+\gamma_2 t)r\cos(\omega t+\alpha)] = 0 \Longrightarrow (\lambda_i,\gamma_i)_i = 0\Big) \ . \end{aligned}$$

Le membre de droite de l'équivalence ci-dessus, pour un porteur en MCU, peut s'écrire sous la forme suivante

$$\begin{aligned} &[y_0 + r\sin(\omega t+\alpha) - b - r\sin(\eta t+\gamma)][\lambda_1+\lambda_3\cos(\omega t+\alpha)+(\gamma_1+\gamma_2 t)r\sin(\omega t+\alpha)] \\ &\quad + [x_0 + r\cos(\omega t+\alpha) - a - r\cos(\eta t+\gamma)][\lambda_2+\lambda_3\sin(\omega t+\alpha)+(\gamma_1+\gamma_2 t)r\cos(\omega t+\alpha)] = 0 \ . \end{aligned}$$

On développe et on obtient une combinaison linéaire des fonctions $\mathbb{1}$, $\cos(\omega t+\alpha)$, $\sin(\omega t+\alpha)$, $\cos(\omega t+\alpha)\sin(\eta t+\gamma)$, etc $\cdots$ Comme $\omega \neq \eta$, c'est une famille libre. En isolant le terme en $\sin(\eta t+\gamma)\sin(\omega t+\alpha)$ par exemple, il vient clairement $\gamma_1 = \gamma_2 = 0$. Ensuite, les termes en $\sin(\eta t+\gamma)$ et $\sin(\eta t+\gamma)\cos(\omega t+\alpha)$ donnent $\lambda_1 = \lambda_3 = 0$ et finalement $\lambda_2 = 0$. □

Preuve de la proposition 5.14. Supposons que deux trajectoires paramétriques S_θ et $S_{\theta'}$ ont même mesures angulaires. On note

$$\theta = (x_0, y_0, r, \omega, \alpha) \ , \quad \theta' = (x'_0, y'_0, r', \omega', \alpha') \ ,$$

et $O(t) = (a + r\cos(\eta t+\gamma),\ b + r\sin(\eta t+\gamma))$. Pour tout $t\in[0,1]$, on a

$$\begin{aligned} &m(\theta,t) = m(\theta',t) \iff \\ &\quad [x_0 + r\cos(\omega t+\alpha) - a - r\cos(\eta t+\gamma)][y'_0 + r'\sin(\omega' t+\alpha') - b - r\sin(\eta t+\gamma)] = \\ &\qquad [x'_0 + r'\cos(\omega' t+\alpha') - a - r\cos(\eta t+\gamma)][y_0 + r\sin(\omega t+\alpha) - b - r\sin(\eta t+\gamma)] \ . \end{aligned}$$

On développe et on obtient une combinaison linéaire des fonctions $\mathbb{1}$, $\cos(\omega t + \alpha)$, $\cos(\eta t + \gamma)$, $\cos(\omega t + \alpha)\sin(\omega' t + \alpha')$, etc ... Si on suppose $\omega \neq \omega'$, alors cette famille est libre ce qui impose $r = r' = 0$ qui est exclu par hypothèse. Donc $\omega = \omega'$. On obtient ensuite sans difficulté que $(x_0, y_0) = (x'_0, y'_0)$ puis un système de la forme

$$\begin{cases} r' \cos\alpha' - r\cos\alpha = 0 \ , \\ r' \sin\alpha' - r\sin\alpha = 0 \ . \end{cases}$$

Pour que $(r, r') = (0, 0)$ ne soit pas l'unique solution, il faut alors $\alpha = \alpha'$ puis $r = r'$. □

Preuve de la proposition 5.15. On peut supposer sans perte de généralité que $(1, 1, 0, 1) \in \Theta$. Considérons alors les trajectoires définies pour tout $t \in [0, 1]$ par

$$S(t) = (1 + t, t) \ , \quad O(t) = (t^2, t) \ .$$

Alors, d'après la proposition 5.4, la cible est observable. Supposons ensuite que la cible fasse pour le porteur l'hypothèse d'une trajectoire à composantes polynomiales de degré deux. Alors, d'après la proposition 5.9, le porteur n'est pas observable par la cible. □

Chapitre 6

Trajectoires paramétriques par morceaux

Sommaire

6.1 Introduction 51
6.2 Observabilité 51
6.3 Sections MRU et MCU 56
6.4 Extensions 59
6.5 Conclusion 60

6.1 Introduction

L'extension naturelle des modèles paramétriques décrits au chapitre 5 est le cas des trajectoires paramétriques par morceaux. Intuitivement, cette notion semble très simple puisqu'une telle trajectoire est une succession de sections paramétriques. Cependant, la formulation de ces modèles est plus lourde et sophistiquée qu'on ne pourrait l'envisager. La dénomination de *trajectoires paramétriques par morceaux* peut même induire en erreur sur la véritable nature du problème : quand le nombre de sections constituant la trajectoire est inconnu, le problème d'estimation associé est un problème d'estimation non-paramétrique.

Dans ce chapitre, on présente en section 6.2 le cadre des modèles *paramétriques par morceaux* ainsi que la notion d'observabilité pour ces modèles. On s'intéresse ensuite en section 6.3 aux modèles constitués de sections MRU ou MCU dans le cadre d'observations à temps continu. Des extensions sont envisagées en section 6.4 pour assouplir les contraintes formulées en section 6.2 mais aussi pour aborder le cas d'observations à temps discret.

6.2 Observabilité

Dans tout le chapitre, on travaille avec le modèle d'observations à temps continu. On considère une collection de modèles paramétriques $\{\mathcal{M}_i\}_{i\in I}$ où, pour $i \in I$, $\mathcal{M}_i$ est un modèle paramétrique, *i.e.* la donnée d'un ensemble $\Theta_i \subset \mathbb{R}^{d_i}$ et d'une application $\theta \mapsto S_{\mathcal{M}_i,\theta}$ que l'on note

$$\mathcal{M}_i : \begin{array}{ccc} \Theta_i & \longrightarrow & \mathscr{C}^\infty([0,1],\mathbb{R}^2) \\ \theta & \longmapsto & S_{\mathcal{M}_i,\theta} \end{array}$$

L'application $S_{\mathcal{M}_i,\theta}$ désigne donc la trajectoire du modèle paramétrique $\mathcal{M}_i$ de paramètre θ. Soulignons que la régularité est renforcée par rapport à la définition 5.1. Cette condition est nécessaire

car les changements de section d'une trajectoire par morceaux seront détectés comme des discontinuités de la trajectoire ou de ses dérivées. Rappelons que $O(t) = (O_1(t), O_2(t))$ désigne la position du porteur à l'instant t. Pour rester dans le domaine de différentiabilité de la fonction angle, on formule l'hypothèse suivante

$$(\mathrm{H}_*) \qquad \forall i \in I \quad \forall (\theta, t) \in \Theta_i \times [0,1] \qquad S_{\mathcal{M}_i,\theta}(t) - O(t) \notin \{(0,y) : y < 0\} \ .$$

Par la suite, on suppose donnée une collection de modèles paramétriques $\{\mathcal{M}_i\}_{i\in I}$ que l'on note $\mathcal{M}$. Pour une fonction f définie de $[0,1]$ dans $\mathbb{R}^n$ avec n entier non nul de classe $\mathscr{C}^\infty$ sauf éventuellement en un nombre fini de points, on définit

$$\mathcal{D}_f \stackrel{\text{def}}{=} \left\{ t \in [0,1] : \exists\, i \in \mathbb{N} \quad f^{(i)}(t^-) \neq f^{(i)}(t^+) \right\} \ ,$$

où $f^{(i)}$ désigne la dérivée d'ordre i de f ou le vecteur des dérivées d'ordre i si f est à valeurs vectorielles.

Définition 6.1. *Pour une trajectoire X de classe $\mathscr{C}^\infty$ sur $[0,1]$ sauf éventuellement en un nombre fini de points, on appellera* rupture *de X tout point $t \in \mathcal{D}_X$.*

On définit, pour tout entier $p \geq 1$,

$$\Delta_p \stackrel{\text{def}}{=} \left\{ (t_0, \cdots, t_p) \in [0,1]^{p+1},\ 0 = t_0 < t_1 < \cdots < t_p = 1 \right\} \ .$$

Définition 6.2. *Une trajectoire S est dite* paramétrique par morceaux *s'il existe un entier $p \geq 1$, une subdivision $\underline{\tau} = (\tau_0, \cdots, \tau_p) \in \Delta_p$ et φ est une application de $\mathbb{N}_p = \{1, \ldots, p\}$ dans I tels que*

$$S = \sum_{k=1}^{p} S_{\mathcal{M}_{\varphi(k)},\theta^{(k)}} \mathbb{1}_{[\tau_{k-1},\tau_k[} \quad \textit{avec} \quad \mathcal{D}_S = \underline{\tau} \ . \tag{1}$$

On note la trajectoire S_ξ avec $\xi = (\underline{\theta}, \underline{\tau}, \varphi)$ où

$$\underline{\theta} = \left(\theta^{(1)}, \cdots, \theta^{(p)}\right) \in \prod_{i=1}^{p} \Theta_{\varphi(i)} \ .$$

Pour un p fixé, on note Ξ_p l'ensemble des éléments caractérisant une trajectoire paramétrique à p morceaux, i.e.

$$\Xi_p \stackrel{\text{def}}{=} \left\{ \xi = (\underline{\theta}, \underline{\tau}, \varphi) : \varphi \in \mathcal{A}(\mathbb{N}_p, I),\ \underline{\tau} \in \Delta_p,\ \underline{\theta} \in \prod_{i=1}^{p} \Theta_{\varphi(i)} \textit{ tels que } \mathcal{D}_{S_\xi} = \underline{\tau} \right\} \ .$$

Cette définition est souple puisque les ruptures peuvent apparaître à n'importe quel ordre de dérivation. Elle permet de considérer à la fois des trajectoires MRU par morceaux mais aussi des trajectoires de classe $\mathscr{C}^1$ constituées de sections MRU et MCU. Surtout, cette définition est conforme à l'idée intuitive qu'on se fait d'un modèle dit par morceaux puisque les ruptures sont exactement les changements de régime. Cette définition interdit, par exemple, le découpage d'un MRU en plusieurs portions MRU successives. Par la suite, on utilisera les projections canoniques que l'on notera abusivement $\underline{\theta}(\cdot)$ et $\underline{\tau}(\cdot)$ définies, pour $\xi = (\underline{\theta}, \underline{\tau}, \varphi)$ par $\underline{\theta}(\xi) = \underline{\theta}$ et $\underline{\tau}(\xi) = \underline{\tau}$.

Avant d'introduire les définitions spécifiques au cadre paramétrique par morceaux, rappelons que l'observation non bruitée consiste en la mesure d'un azimut donné par

$$\Psi[(x_1, x_2), t] \stackrel{\text{def}}{=} \begin{cases} \text{angle}[x_1 - O_1(t), x_2 - O_2(t)] & \text{si } (x, y) \neq O(t) \ , \\ 0 & \text{sinon} \ . \end{cases} \tag{2}$$

où $O(t) = (O_1(t), O_2(t))$ désigne la position du porteur à l'instant t et on définit la fonction *mesure angulaire* par

$$\Phi : X \longmapsto (t \mapsto \Psi[X(t), t]) \ .$$

Les résultats d'observabilité d'une trajectoire paramétrique par morceaux vont naturellement s'appuyer sur l'observabilité des différentes sections paramétriques qui composent la trajectoire.

Définition 6.3. *La section paramétrique $S_{\mathcal{M}_{i^*},\theta^*}\mathbb{1}_{[a,b[}$ avec $i^* \in I$, $\theta^* \in \Theta_{i^*}$ est observable dans $\mathcal{M}$ si, pour tout $i \in I$, pour tout $\theta \in \Theta_i$,*

$$\Phi\left(S_{\mathcal{M}_i,\theta}\right)_{|[a,b[} = \Phi\left(S_{\mathcal{M}_{i^*},\theta^*}\right)_{|[a,b[} \Longrightarrow i = i^* \quad \text{et} \quad \theta = \theta^* \ .$$

Conformément au cas des trajectoires paramétriques, une trajectoire paramétriques par morceaux sera dite observable si elle est caractérisée par ses mesures angulaires.

Définition 6.4. *Soit $\Lambda \subset \bigcup_{p\geq 1} \Xi_p$ et $\xi^* \in \Lambda$. La trajectoire paramétrique par morceaux S_{ξ^*} est* observable dans Λ *si, pour tout $\xi \in \Lambda$,*

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \xi = \xi^* \ .$$

Soit Ξ l'ensemble défini par

$$\Xi \stackrel{\text{def}}{=} \left\{ \xi \in \bigcup_{p\geq 1} \Xi_p \ : \ \mathcal{D}_{\Phi(S_\xi)} \setminus \mathcal{D}_O = \underline{\tau}(\xi) \right\} \ .$$

La définition de Ξ suppose *de facto* que les instants de saut de la trajectoire paramétrique par morceaux sont observables puisque ce sont les seuls instants de rupture de $\Phi(S_\xi)$ privés de ceux correspondant à la trajectoire porteur. Comme les trajectoires de la cible et du porteur sont, d'une certaine manière, indépendantes, les ruptures de ces trajectoires sont *a priori* des ensembles disjoints. De plus, il est naturel d'imposer que les ruptures de la trajectoire cible se « lisent » dans les mesures angulaires ce qui légitime la définition de l'ensemble Ξ.

Il serait intéressant de pouvoir formuler une condition qui porte non sur la mesure angulaire $\Phi(S_\xi)$ mais sur la trajectoire cible S_ξ directement. La condition naturelle serait d'imposer

$$\mathcal{D}_{S_\xi} \cap \mathcal{D}_O = \emptyset \ , \tag{3}$$

c'est-à-dire les ruptures de la cible et du porteur ne sont pas simultanées. Cette hypothèse est naturelle dans la mesure où les décisions de manœuvre de la cible et du porteur sont indépendantes et que les changements de régime de ces mobiles n'ont aucune raison de coïncider. Peut-on alors détecter les ruptures de S_ξ *via* ses mesures angulaires ? Supposons la condition (3) vérifiée et considérons par exemple le cas où $\dot{S}_\xi(\tau_k^-) \neq \dot{S}_\xi(\tau_k^+)$. Il vient alors

$$\Phi(S_\xi)'(\tau_k^-) \neq \Phi(S_\xi)'(\tau_k^-) \iff \left\langle \nabla_x \Psi[S_\xi(\tau_k), \tau_k], \dot{S}_\xi(t_k^-) - \dot{S}_\xi(t_k^+) \right\rangle \neq 0 \ .$$

Cette dernière formulation montre que la condition $\Phi(S_\xi)'(\tau_k^-) \neq \Phi(S_\xi)'(\tau_k^-)$ n'est pas du tout restrictive. En effet, pour avoir

$$\left\langle \nabla_x \Psi[S_\xi(\tau_k), \tau_k], \dot{S}_\xi(t_k^-) - \dot{S}_\xi(t_k^+) \right\rangle = 0 \ ,$$

il faut imposer une contrainte très forte à la trajectoire porteur qui est complètement improbable dans une situation opérationnelle. On présente en figures 6.1 et 6.2 une simulation avec une cible en MRU par morceaux où le changement de régime intervient à l'instant $\tau_1 = 13\,\mathrm{s}$. On constate que $\Phi(S_\xi)'$ présente une discontinuité à l'instant $t = \tau_1$.

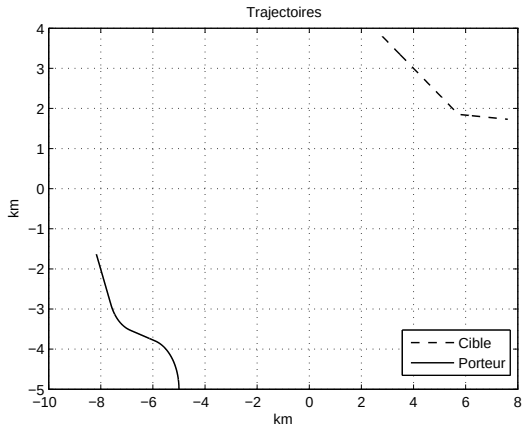


FIG. 6.1 – Cible S_ξ en MRU par morceaux

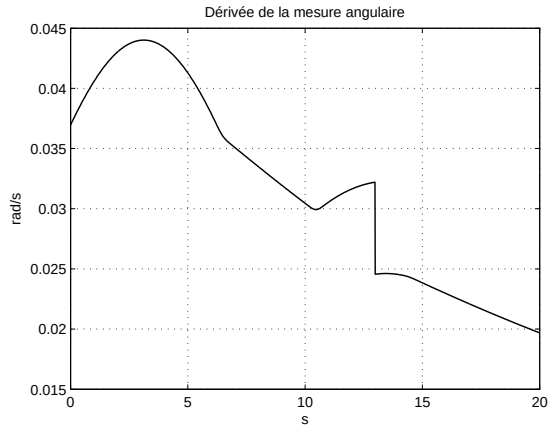


FIG. 6.2 – Tracé de $\Phi(S_\xi)'$

Décrivons également le cas plus réaliste où la trajectoire cible S_ξ est de classe $\mathscr{C}^1$ mais avec une discontinuité de la dérivée seconde $\ddot{S}_\xi(\tau_k^-) \neq \ddot{S}_\xi(\tau_k^+)$. Par dérivation, on a

$$\Phi(S_\xi)''(t) = \mathrm{D}_{x,x}\,\Psi[S_\xi(t), t]\left(\dot{S}_\xi(t), \dot{S}_\xi(t)\right) + \left\langle \nabla_x \Psi[S_\xi(t), t], \ddot{S}_\xi(t) \right\rangle \ ,$$

d'où il résulte

$$\Phi(S_\xi)''(\tau_k^-) \neq \Phi(S_\xi)''(\tau_k^-) \iff \left\langle \nabla_x \Psi[S_\xi(\tau_k), \tau_k], \ddot{S}_\xi(t_k^-) - \dot{S}_\xi(t_k^+) \right\rangle = 0 \ .$$

Là encore, cette condition n'est pas une hypothèse forte sur le porteur et sera toujours satisfaite dans une situation opérationnelle. Sur les figures 6.3 et 6.4, on présente le cas d'une cible avec une trajectoire de classe $\mathscr{C}^1$ constituée d'une section MCU suivie, à l'instant $\tau_1 = 9\,\mathrm{s}$ d'un MRU. On constate que $\Phi(S_\xi)''$ présente une discontinuité à l'instant $t = \tau_1$.

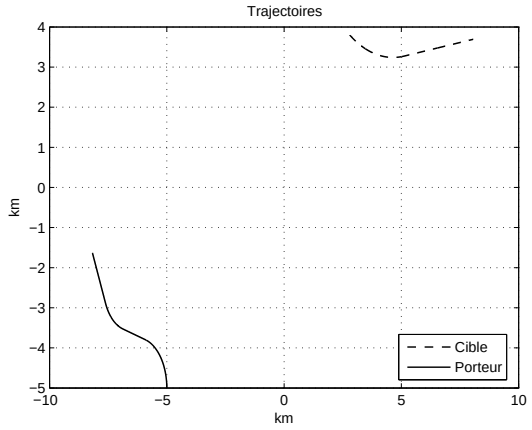


FIG. 6.3 – Cible S_ξ en MRU-MCU par morceaux

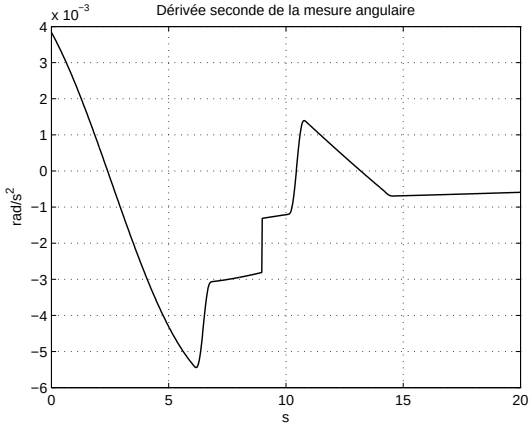


FIG. 6.4 – Tracé de $\Phi(S_\xi)''$

Proposition 6.1. *Soit $\mathcal{M}$ une collection de modèles paramétriques qui vérifie (H_*). Soit une trajectoire cible S_{ξ^*} paramétrique par morceaux avec $\xi^* \in \Xi$ et telle que toute section paramétrique est observable. Alors, la trajectoire cible est observable dans Ξ.*

Démonstration. Par définition de Ξ, on a $\mathcal{D}_{\Phi(S_\xi)} \setminus \mathcal{D}_O = \underline{\tau}(\xi)$ ce qui implique

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \underline{\tau}(\xi) = \underline{\tau}(\xi^*) \ .$$

Par suite, comme chaque section paramétrique est supposée observable, on obtient clairement l'observabilité de toute la trajectoire. □

Ce résultat est un peu naïf puisque chaque section paramétrique est supposée observable. La contrainte sur le porteur est forte et impose des manœuvres adéquates sur chaque intervalle de temps $[\tau_{k-1}, \tau_k[$, mais cette contrainte n'est pas irréaliste. On verra un résultat pratique conséquence de cette proposition à la section suivante.

Les hypothèses de régularité faites sur la trajectoire cible permettent d'annoncer des résultats plus fins. On va pour cela introduire la notion suivante :

Définition 6.5. *Soit $\mathcal{M}$ une collection de modèles paramétriques qui vérifie (H_*) et $\xi \in \bigcup_{p\geq 1} \Xi_p$ avec $\underline{\tau}(\xi) = \underline{\tau}$. Pour une section paramétrique $S_\xi \mathbb{1}_{[\tau_{k-1},\tau_k[}$, on appelle conditions aux bords d'ordre n la famille de vecteurs*

$$\left\{ S^{(j)}(\tau_{k-1}^+),\ S^{(j)}(\tau_k^-), \quad j \leq n \right\} \ .$$

Proposition 6.2. *Soit $\mathcal{M}$ une collection de modèle paramétriques, S_{ξ^*} une trajectoire continue paramétrique par morceaux avec $\xi^* \in \Xi$ et avec un nombre impair de sections. On suppose que la première section est observable puis une section sur deux. On suppose également que les modèles des sections paramétriques non observables sont connus. Alors, si les conditions aux bords d'ordre zéro déterminent toute section, la trajectoire est observable.*

Démonstration. Comme les modèles des sections qu'on ne suppose pas observables sont connus, on peut admettre, sans perte de généralité, que la collection $\mathcal{M}$ ne contient qu'un modèle. Par définition de Ξ, on a $\mathcal{D}_{\Phi(S_\xi)} \setminus \mathcal{D}_O = \underline{\tau}(\xi)$ et par suite

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \underline{\tau}(\xi) = \underline{\tau}(\xi^*) \ .$$

La première et la troisième sections sont supposées observables, donc

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \left(\theta^{(1)}, \theta^{(3)}\right) = \left(\theta^{(1)*}, \theta^{(3)*}\right) .$$

Par continuité de la trajectoire, on a

$$S_{\mathcal{M},\theta^{(1)}}(\tau_1^-) = S_{\mathcal{M},\theta^{(2)}}(\tau_1^+) \quad \text{et} \quad S_{\mathcal{M},\theta^{(2)}}(\tau_2^-) = S_{\mathcal{M},\theta^{(3)}}(\tau_2^+) .$$

Or, les conditions aux bords $S_{\mathcal{M},\theta^{(1)}}(\tau_1^+)$ et $S_{\mathcal{M},\theta^{(3)}}(\tau_2^-)$ déterminent le paramètre $\theta^{(2)}$ et comme $\left(\theta^{(1)}, \theta^{(3)}\right) = \left(\theta^{(1)*}, \theta^{(3)*}\right)$, il s'ensuit $\theta^{(2)} = \theta^{(2)*}$. On conclut en itérant le procédé. □

Ce résultat peut sembler opaque en première lecture. Mais son utilité est réelle pour des scénarios de trajectographie passive puisqu'on peut garantir l'observabilité d'une cible sans que toutes les sections soient observables, situation plausible si le porteur a des contraintes opérationnelles lourdes. On va présenter un certain nombre d'applications et de déclinaisons de ce résultat dans la section suivante.

Dans la même veine, on peut également annoncer des résultats d'observabilité pour des trajectoires de classe $\mathscr{C}^1$.

Proposition 6.3. *Soit $\mathcal{M}$ une collection de modèles paramétriques qui vérifie (H_*) et S_{ξ^*} une trajectoire de classe $\mathscr{C}^1$ paramétrique par morceaux avec $\xi^* \in \Xi$ et avec un nombre impair de sections. On suppose que la première section est observable puis une section sur deux. On suppose également que les modèles des sections paramétriques non observables sont connus. Alors, si les conditions aux bords d'ordre un déterminent toute section, la trajectoire est observable.*

Démonstration. La preuve est identique à celle de la proposition 6.2. Les conditions aux bords d'ordre un caractérisent les sections qu'on ne suppose pas observables. □

6.3 Sections MRU et MCU

En utilisant les résultats d'observabilité des trajectoires MRU et MCU établis au chapitre 5, on peut annoncer des résultats utiles au praticien pour l'observabilité des trajectoires constituées de sections MRU et MCU. Comme pour le cas paramétrique, on utilisera l'hypothèse suivante :

$$(\mathrm{H}_+) \qquad \forall i \in I \quad \forall(\theta, t) \in \Theta_i \times [0,1] \qquad S_{\mathcal{M}_i,\theta,2}(t) - O_2(t) > 0.$$

Proposition 6.4. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui ne contient que le modèle MRU, c'est-à-dire que l'on considère des trajectoires MRU par morceaux. Soit une trajectoire S_{ξ^*} MRU par morceaux avec $\xi^* \in \Xi$ et telle que pour tout k, il existe un intervalle $[a_k, b_k[\subset [\tau_{k-1}, \tau_k[$ où le porteur est en MCU. Alors, la cible est observable.*

Démonstration. C'est une conséquence des propositions 5.10 et 6.1. □

Il ne s'agit pas d'un résultat très élaboré mais les hypothèses sont tout à fait raisonnables pour une situation opérationnelle. En fait, le résultat tient même si la trajectoire n'est pas supposée continue.

Proposition 6.5. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui ne contient que le modèle MRU. Soit S_{ξ^*} une trajectoire continue MRU par morceaux avec $\xi^* \in \Xi$. On suppose qu'une section sur deux est observable. Si de plus $\beta(\cdot)$ est non constant durant la première et la dernière section, alors la trajectoire est observable.*

Démonstration. C'est essentiellement une conséquence de la proposition 6.2. Par définition de Ξ, il vient

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \underline{\tau}(\xi) = \underline{\tau}(\xi^*) \ .$$

Ensuite, détaillons le cas où la dernière section n'est pas observable[1]. Soit $p-1$ le nombre de rupture de S_ξ. L'avant-dernière section est observable donc

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \theta^{(p-1)} = \theta^{(p-1)*} \ .$$

Notons $\theta^{(p)} = (x_0, y_0, v_x, v_y)$ et $\theta^{(p)*} = (x_0^*, y_0^*, v_x^*, v_y^*)$. Par continuité des trajectoires, on obtient que

$$\begin{cases} x_0 + \tau_{p-1} v_x = x_0^* + \tau_{p-1} v_x^* \ , \\ y_0 + \tau_{p-1} v_y = y_0^* + \tau_{p-1} v_y^* \ . \end{cases} \tag{4}$$

Pour la dernière section, on a

$$\Phi\left(S_{\mathcal{M},\theta^{(p)}}\right)_{|[\tau_{p-1},\tau_p[} = \Phi\left(S_{\mathcal{M},\theta_p^*}\right)_{|[\tau_{p-1},\tau_p[} \iff$$
$$\forall t \in [\tau_{p-1}, \tau_p[\quad (x_0 - x_0^*) + \beta(t)(y_0 - y_0^*) + t(v_x - v_x^*) + t\beta(t)(v_y - v_y^*) = 0 \ .$$

Cette égalité et le système (4) impliquent

$$\forall t \in [\tau_{p-1}, \tau_p[\quad (v_x - v_x^*) + \beta(t)(v_y - v_y^*) = 0 \ ,$$

et la conclusion suit pour $\beta(\cdot)$ non constant. □

Ainsi, en tenant compte de la continuité de la trajectoire, on peut relâcher l'hypothèse d'observabilité des sections, c'est-à-dire assouplir les contraintes de manœuvre pour le porteur. Par exemple, il suffit que le porteur réalise des MCU une fois sur deux lors des changements de régime de la cible et que sa mesure $\beta(\cdot)$ soit non-constante[2] pour que la cible soit observable.

Proposition 6.6. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui contient le modèle MRU et le modèle MCU. Soit S_{ξ^*} avec $\xi^* \in \Xi$ une trajectoire continue MRU-MCU par morceaux telle que, pour tout k,*

$$\det\left(\dot{S}_\xi(\tau_k^-), \dot{S}_\xi(\tau_k^+)\right) = 0 \quad \text{et} \quad \left\langle \dot{S}_\xi(\tau_k^-), \dot{S}_\xi(\tau_k^+)\right\rangle_{\mathbb{R}^2} \geq 0 \ . \tag{5}$$

On suppose que le porteur effectue des MCU lors de la première section puis une fois sur deux et lors de la dernière section avec des vitesses angulaires différentes de la cible quand celle-ci est en MCU. Alors, la trajectoire est observable.

Démonstration. Le schéma de preuve est semblable à celui de la proposition 6.5. On a tout d'abord

$$\Phi(S_\xi) = \Phi(S_{\xi^*}) \Longrightarrow \underline{\tau}(\xi) = \underline{\tau}(\xi^*) \ .$$

Précisons ensuite en quoi la trajectoire porteur permet l'observabilité du modèle, puisque deux modèles cohabitent dans la collection. Il suffit d'écrire l'égalité des mesures angulaires pour une cible en MRU et une cible en MCU dont la vitesse angulaire est différente de celle du porteur pour aboutir à une impossibilité, par un argument de famille libre (à l'instar de ce qui a été fait

[1] C'est identique pour la première section.

[2] Cette condition n'est pas du tout restrictive en pratique. Au contraire, imposer $\beta(\cdot)$ constant serait totalement irréaliste.

dans les preuves du chapitre 5), c'est-à-dire

$$\Phi(S_\xi)_{|[\tau_{k-1},\tau_k[} = \Phi(S_{\xi^*})_{|[\tau_{k-1},\tau_k[} \Longrightarrow \varphi(k) = \varphi^*(k) \ .$$

Ainsi, le modèle est observable et le paramètre également, donc toute section où le porteur est en MCU est observable. Détaillons ensuite pourquoi les conditions aux bords déterminent effectivement les sections qu'on ne suppose pas observables. Soit $[\tau_{k-1}, \tau_k[$ une de ces sections. Par continuité de la trajectoire, les points aux instants τ_{k-1} et τ_k sont déterminés. La condition (5) indique le sens et l'orientation du vecteur vitesse à tout instant. Si la direction reliant les deux points aux instants τ_{k-1} et τ_k est la même, alors, il y a unique MRU qui passe par ces deux points. Sinon, il s'agit d'un MCU dont on connaît la tangente à l'instant τ_{k-1}. Or, il existe un unique MCU qui passe par deux points avec une tangente imposée pour l'un deux. Ainsi, les conditions aux bords déterminent les sections manquantes. □

La condition (5) permet de considérer des trajectoires MRU-MCU qui ne sont pas nécessairement à vitesse constante mais telles qu'il n'y a pas de discontinuité de l'orientation du vecteur vitesse. Ainsi, on peut avoir deux MRU qui se succèdent avec des vitesses différentes mais des sens et des orientations identiques. Ce résultat permet désormais de considérer des trajectoires réalistes, comme des *hippodromes*[3] ou des *baïonnettes*[4] ou plus simplement des MCU par morceaux.

On peut également annoncer des résultats d'observabilité pour des trajectoires de classe $\mathscr{C}^1$ comme par exemple les trajectoires dont les composantes sont des splines cubiques.

Définition 6.6. *Soit $0 = t_0 < t_1 < \ldots < t_{k-1} = 1$ une subdivision de $[0,1]$. On appelle spline cubique sur $[0,1]$ toute courbe $s \in \mathscr{C}^2([0,1])$ telle que, pour tout $i \in \mathbb{N}_{k-1}$, $s_{|[t_i,t_{i-1}]} \in \mathbb{R}_2[X]$.*

Proposition 6.7. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui ne contient que le modèle spline cubique. Soit S_{ξ^*} une trajectoire de classe $\mathscr{C}^1$ constituée de splines cubiques avec $\xi^* \in \Xi$. On suppose que la première section est observable et puis une section sur deux. Alors, la trajectoire est observable.*

Démonstration. Il suffit de vérifier que les conditions aux bords d'ordre un détermine une section spline cubique, c'est-à-dire dont les composantes sont des polynômes de degré inférieur ou égal à trois, ce qui immédiat puisque les conditions aux bords fournissent un système de Cramer en les inconnues. □

On pourrait, comme pour la proposition 6.5, proposer une condition suffisante sur $\beta(\cdot)$ pour éviter d'imposer l'observabilité de la première section mais son énoncé est moins simple et moins intuitif.

On peut affiner le résultat de la proposition 6.6 sous la régularité $\mathscr{C}^1$.

Proposition 6.8. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui contient le modèle MRU et le modèle MCU. Soit S_{ξ^*} avec $\xi^* \in \Xi$ une trajectoire de classe $\mathscr{C}^1$ MRU-MCU par morceaux. On suppose que le porteur effectue des MCU une fois sur deux par section avec des vitesses angulaires différentes de la cible quand celle-ci est en MCU. Alors, la trajectoire est observable.*

La régularité $\mathscr{C}^1$ impose la vitesse constante pour une telle trajectoire et il n'est plus indispensable que les sections aux extrémités soient observables. On ne détaille pas la preuve de ce résultat qui est identique à celle de la proposition 6.6.

[3] La cible se maintient dans un périmètre donné en suivant une trajectoire en forme d'hippodrome.

[4] La cible est en MRU, puis en MCU suivi d'un MCU de courbure inverse et à nouveau en MRU, d'où cette appellation. Cette trajectoire est utile quand on souhaite manœuvrer ponctuellement et maintenir un cap, typiquement en trajectographie.

D'autres résultats sont accessibles avec, par exemple, l'observabilité d'une section une fois sur trois.

Proposition 6.9. *Soit la collection $\mathcal{M}$ qui vérifie (H_+) et qui contient le modèle MCU. Soit S_{ξ^*} avec $\xi^* \in \Xi$ une trajectoire de classe $\mathscr{C}^1$ MCU par morceaux. On suppose que le porteur effectue des MCU une fois sur trois par section avec des vitesses angulaires différentes de la cible quand celle-ci est en MCU. Alors, la trajectoire est observable.*

La preuve est laissée au soin du lecteur. Le seul point non trivial concerne la détermination de deux séquences MCU avec un raccord $\mathscr{C}^1$ à partir des conditions aux bords d'ordre un. Un dessin permet de se convaincre qu'il existe un unique couple de MCU dont le raccord est $\mathscr{C}^1$ avec des vecteurs vitesses imposés aux bords.

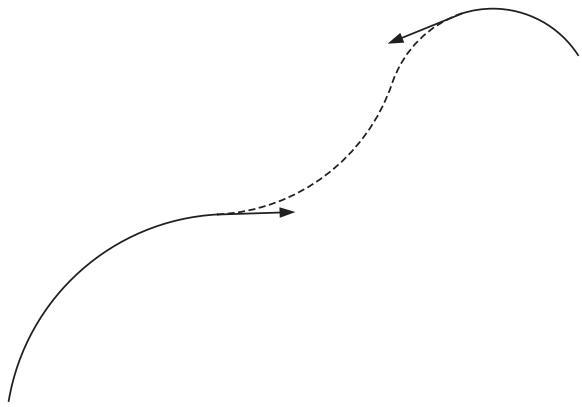


Fig. 6.5 – Trajectoire de classe $\mathscr{C}^1$ MCU par morceaux

On peut imaginer une multitude de variantes avec l'observabilité des sections au plus une fois sur deux ou au plus une fois sur trois ou encore avec de nouvelles contraintes opérationnelles.

6.4 Extensions

On pourrait souhaiter faire cohabiter des modèles emboîtés, par exemple que la collection $\mathcal{M}$ contienne le modèle polynomial de degré un (MRU) et le modèle polynomial de degré inférieur ou égal à deux. La définition de section observable n'aurait plus de sens car on pourrait coder toute section MRU dans deux modèles distincts. Une façon de contrer cette difficulté serait de tenir compte de la *complexité du modèle*.

Définition 6.7. *Une section paramétrique $S_{\mathcal{M}_i,\theta}$ avec $\theta \in \Theta_i \subset \mathbb{R}^{d_i}$ est dite de complexité d_i.*

Et la notion d'observabilité d'une section paramétrique pourrait devenir :

Définition 6.8. *La section paramétrique $S_{\mathcal{M}_{i^*},\theta^*|[a,b[}$ avec $i^* \in I$, $\theta^* \in \Theta_{i^*}$ est observable dans $\mathcal{M}$ à complexité minimale si, pour tout $i \in I$, pour tout $\theta \in \Theta$,*

$$\begin{cases} \Phi(S_{\mathcal{M}_i,\theta|[a,b[}) = \Phi(S_{\mathcal{M}_{i^*},\theta^*|[a,b[}) \ , \\ d_i \leq d_{i^*} \ , \end{cases} \implies i = i^* \quad \text{et} \quad \theta = \theta^* \ .$$

Dans la présentation de la section 6.2, les instants de rupture sont détectés parce qu'ils provoquent des discontinuités pour la dérivée d'un certain ordre de la trajectoire. Mais en pratique, les trajectoires sont de classe $\mathscr{C}^\infty$ et les changements de régimes correspondent à des dépassements de valeur seuil par exemple pour les dérivées secondes[5]. Typiquement, un changement de régime correspond à une situation où on observe $\|\ddot{S}_\xi\| > \Gamma_s$ avec Γ_s une valeur qui détermine quand la cible amorce une manœuvre. Il faudrait alors pouvoir adapter la notion d'observabilité pour de telles situations ce qui amènerait à l'introduction d'une notion moins binaire que l'observabilité.

Par ailleurs, dans le cas d'observations à temps discret, on ne peut plus travailler avec les dérivées de $\Phi(X_\xi)$ qui ne sont pas des grandeurs accessibles. Il faut détecter les instants de saut avec des conditions de non conformité aux modèles paramétriques de $\mathcal{M}$. Formalisons cette idée.

Définition 6.9. *Une section $S_{|[a,b[}$ est dite détectable dans $\mathcal{M}$ s'il existe $i \in I$ et $\theta \in \Theta_i$ tels que*

$$\Phi\left(S_{|[a,b[}\right)(t_{1:n}) = \Phi\left(S_{\mathcal{M}_i,\theta|[a,b[}\right)(t_{1:n}) \ .$$

Considérons le cas où la cible S_ξ est en MRU par morceaux. Supposons pour simplifier que $\underline{\tau} \in \{t_{1:n}\}$ qui désignent les instants d'observations. Le porteur réalise un MCU avant et après τ_1. Posons alors

$$s_1 = \sup\left\{s > 0 \text{ tel que } \ S_{\xi|[0,s[} \text{ est une section détectable dans } \mathcal{M}\right\} \ .$$

On a clairement $s_1 \geq \tau_1 (= t_k)$ et on peut montrer que pour $s_1 > \tau_1$ $(s_1 = t_{k+1})$, $S_{\xi|[0,s_1[}$ n'est pas détectable d'où $s_1 = \tau_1$. On peut ensuite définir

$$s_2 = \sup\left\{s > \tau_1 \text{ tel que } S_{\xi|[\tau_1,s[} \text{ est une section détectable}\right\} \ ,$$

et on itère le procédé jusqu'à obtenir la subdivision complète. Ainsi, τ est caractérisée par l'observation. Si on ne suppose pas que $\tau_k \in \{t_{1:n}\}$, on peut le déduire comme l'instant où les sections paramétriques obtenues avant et après se rejoignent.

Enfin, signalons qu'il doit être possible dans certains cas, de définir une matrice d'information comme dans le cas paramétrique et d'établir des liens entre observabilité et inversibilité de cette matrice. L'intérêt d'une telle étude reste très limité, compte-tenu des difficultés d'estimation évoquées au chapitre 9.

6.5 Conclusion

Dans ce chapitre, on dresse le formalisme approprié au modèle de trajectoires paramétriques pas morceaux. De nombreuses situations d'observabilité sont ensuite établies, d'abord pour des modèles généraux puis appliquées à des cas réalistes de trajectoires constituées de sections MRU et MCU. Les hypothèses identifiées sont plausibles dans des situations opérationnelles ce qui donne un intérêt pratique véritable à ces résultats d'observabilité. Toutefois, établir l'observabilité est une chose, être en mesure d'estimer une telle trajectoire en est une autre, comme cela est expliqué au chapitre 9. Des perspectives sont envisagées pour étendre les résultats dans le cadre d'observations à temps discret et de trajectoires lisses.

[5]Si un mobile passe d'un régime MRU à un régime MCU de manière "lisse", les variations d'accélération sont très fortes pendant la transition.

Chapitre 7

Trajectoires régies par une équation différentielle

Sommaire

7.1 Introduction 61
7.2 L'observabilité en théorie du contrôle 61
7.3 Observabilité à vitesse constante 65
7.4 Vers un critère d'observabilité 68
7.5 Perspectives 71

7.1 Introduction

On a traité précédemment l'observabilité de trajectoires paramétriques et paramétriques par morceaux. Malgré leurs complexités, ces modèles impliquent des dynamiques assez rigides qui imposent à la cible de suivre une trajectoire relativement contrainte. En pratique, les mobiles considérés en trajectographie passive comme les avions, les bateaux ou les sous-marins, ont des possibilités de manœuvre très supérieures aux seules dynamiques paramétriques ou paramétriques par morceaux[1]. Le cas d'une cible à vitesse constante englobe le cas du MRU, du MCU, du MRU-MCU à vitesse constante et laisse le champ libre pour des dynamiques beaucoup plus générales. La question de l'observabilité prend alors tout son sens. À notre connaissance, ce problème n'est pas abordé dans la littérature sur la trajectographie passive.

Ce chapitre contient des résultats expérimentaux. Dans la section 7.3, on décrit le procédé qui permet d'étudier les trajectoires ambiguës à la cible ayant des vitesses constantes. On montre ensuite des simulations qui répondent de manière empirique à la question de l'observabilité. La section 7.4 étudie la possibilité de définir un critère garantissant l'observabilité. Les perspectives de cette première étude sont présentées en section 7.5.

7.2 L'observabilité en théorie du contrôle

Le cas des dynamiques MRU et MCU est un cas particulier de dynamique à vitesse constante. Il existe en théorie du contrôle un formalisme et des résultats concernant des systèmes régis par des équations différentielles et partiellement observés. On rappelle que $O : t \mapsto (O_1(t), O_2(t))$

[1]Cela n'altère en rien l'intérêt de ces modèles qui sont fréquemment rencontrés en pratique. Par exemple, avion, bateau ou sous-marin sont très souvent en MRU.

désigne la trajectoire du porteur. Pour une trajectoire $t \mapsto (x(t), y(t))$ qui vérifie la condition

$$(\mathrm{H}_+) \qquad \forall t \in [0,1] \qquad y(t) - O_2(t) > 0 \ ,$$

on note $X(t)$ le vecteur d'état à l'instant t défini par

$$X(t) \stackrel{\text{def}}{=} \begin{pmatrix} x(t) & y(t) & \dot{x}(t) & \dot{y}(t) \end{pmatrix}^{\mathrm{T}} \ .$$

et $Y(t)$ l'observation

$$Y(t) \stackrel{\text{def}}{=} h[X(t), t] = \Psi[(x(t), y(t)), t] = \arctan[x(t) - O_1(t)]/[y(t) - O_2(t)] \ .$$

Faisons l'hypothèse que la trajectoire X est à vitesse constante inconnue, c'est-à-dire, pour tout $t \in [0,1]$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}^2 + \dot{y}^2) = 0 \iff \langle (\dot{x}, \dot{y}), (\ddot{x}, \ddot{y}) \rangle = 0 \ .$$

Alors, on peut formaliser le problème de trajectographie passive d'une cible à vitesse constante sans bruit de mesure sous la forme d'un système différentiel partiellement observé. Pour tout $t \in [0,1]$, on a

$$\begin{cases} \dot{X}(t) = A(t)\, X(t) \ , \\ Y(t) = h[X(t), t] \ , \end{cases} \tag{1}$$

avec $A(t)$ la matrice définie par

$$A(t) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\varrho(t) \\ 0 & 0 & \varrho(t) & 0 \end{pmatrix} \ ,$$

où $\varrho(t) = \dot{s}^{3/2}(t) c^{1/2}(t)$, s désignant l'abscisse curviligne de $t \mapsto (x(t), y(t))$ (de dérivée constante ici car la vitesse est constante) et c sa courbure[2]. Précisons que la matrice $A(t)$ est inconnue, la fonction ϱ étant elle-même inconnue.

Définition 7.1. *Le système* (1) *est localement observable en X si il existe un voisinage V_X de X tel que, pour toute trajectoire $Z \in V_X$,*

$$\forall t \in [0,1] \quad \begin{cases} \dot{Z}(t) = A(t)\, Z(t) \ , \\ h[X(t), t] = h[Z(t), t] \ , \end{cases} \implies X = Z \ .$$

On définit la résolvante du système (1) notée $t \mapsto M(t)$ comme solution de

$$\begin{cases} \dot{M}(t) = A(t)\, M(t) \ , \\ M(0) = I_4 \ . \end{cases} \tag{2}$$

On peut alors formuler le critère d'observabilité locale du système (1), voir [103, Chapter 6, Theorem 26].

[2]On rappelle que la courbure est définie, en coordonnées cartésiennes, par $c = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}$.

Proposition 7.1. *Le système* (1) *est localement observable en X si et seulement si la matrice*

$$G \stackrel{\text{def}}{=} \int_0^1 {M^{-1}}^{\mathrm{T}}(t)\nabla_X h[X(t),t]\nabla_X h[X(t),t]^{\mathrm{T}} M^{-1}(t)\,\mathrm{d}t \tag{3}$$

est inversible.

Il peut sembler, en première lecture, que ce critère permet de discuter l'observabilité locale de toute trajectoire à vitesse constante. En réalité, la matrice résolvante M est définie à partir de la matrice A elle-même inconnue. Il n'est donc pas du tout clair que la matrice G puisse avoir une expression déterminée à partir de la seule observation $\{Y(t)\}_{t\in[0,1]}$. Ce critère peut être réécrit de la manière suivante.

Proposition 7.2. *Notons $s(X)(t) = h[X(t),t]$. Le système* (1) *est localement observable en X si et seulement si la famille*

$$\left(\frac{\partial s(X)}{\partial x}, \frac{\partial s(X)}{\partial y}, u\frac{\partial s(X)}{\partial x} + v\frac{\partial s(X)}{\partial y}, -v\frac{\partial s(X)}{\partial x} + u\frac{\partial s(X)}{\partial y}\right)$$

est libre, les fonctions u et v étant définies par $u = -(a\cos\gamma + b\sin\gamma)$ et $v = a\sin\gamma - b\cos\gamma$ avec $a(t) = \int_0^t \cos\gamma$, $b(t) = \int_0^t \sin\gamma$ et $\gamma(t) = \int_0^t \varrho$.

Démonstration. On trouve pour solution du système (2) la matrice définie pour tout $t \in [0,1]$ par $M(t) = \begin{pmatrix} I_2 & B(t) \\ 0 & \mathrm{R}_{\gamma(t)} \end{pmatrix}$ où $\mathrm{R}_{\gamma(t)}$ est la matrice de rotation d'angle $\gamma(t)$ et $B(t) = \begin{pmatrix} a(t) & -b(t) \\ b(t) & a(t) \end{pmatrix}$. Son inverse est donnée par

$$M^{-1}(t) = \begin{pmatrix} I_2 & D(t) \\ 0 & \mathrm{R}_{-\gamma(t)} \end{pmatrix}, \quad \text{avec} \quad D(t) = -B(t)\mathrm{R}_{-\gamma(t)}\,.$$

Pour $t \in [0,1]$, le gradient $\nabla_X s(X)(t)$ s'écrit

$$\nabla_X s(X)(t) = \begin{pmatrix} \frac{\partial s(X)}{\partial x}(t) & \frac{\partial s(X)}{\partial y}(t) & 0 & 0 \end{pmatrix}^{\mathrm{T}},$$

et par suite

$$\nabla_X s(X)(t)^{\mathrm{T}} M^{-1}(t) = \begin{pmatrix} \frac{\partial s(X)}{\partial x}(t) & \frac{\partial s(X)}{\partial y}(t) & u(t)\frac{\partial s(X)}{\partial x}(t) + v(t)\frac{\partial s(X)}{\partial y}(t) & -v(t)\frac{\partial s(X)}{\partial x}(t) + u(t)\frac{\partial s(X)}{\partial y}(t) \end{pmatrix},$$

avec $u = -(a\cos\gamma + b\sin\gamma)$ et $v = a\sin\gamma - b\cos\gamma$. □

On voit l'intervention de la fonction ϱ, donnée manquante dans la formulation du système (1). Qu'obtient-on pour les dynamiques couramment utilisées que sont le MRU et le MCU ?

Proposition 7.3. *Étant donnée une trajectoire X MRU, on a*

$$G \text{ inversible} \iff (1, \tan s(X), \mathrm{id}, \mathrm{id}\times\tan s(X)) \text{ libre}\,.$$

Démonstration. On a $\varrho(t) = 0$ pour tout $t \in [0,1]$ et on en déduit

$$a(t) = t\,, \quad b(t) = 0\,, \quad u(t) = -t\,, \quad v(t) = 0\,.$$

Tous calculs effectués, on trouve

$$\nabla_X s(X)(t)^{\mathrm{T}} M^{-1}(t) \quad \text{colinéaire à} \quad \begin{pmatrix} 1 & -\tan s(X)(t) & -t & t \times \tan s(X)(t) \end{pmatrix}$$

et on conclut par le fait G est une matrice de Gram. □

Ainsi, dans le cas d'une trajectoire MRU, le critère de la proposition 7.1 est le critère usuel formulé au chapitre 5, proposition 5.8 (page 42) qui ne fait intervenir que l'observation de la trajectoire, donc précisément l'information à laquelle on a accès.

Proposition 7.4. *Pour une trajectoire X MCU, on a $\varrho(t) = \varrho_C \neq 0$ pour tout $t \in [0,1]$. Par suite,*

$$G \textit{ inversible} \iff \big(1, \tan s(X), \varrho_C^{-1}[\tan s(X) - \sin(\varrho_C\ \cdot)], \varrho_C^{-1}[1 - \tan s(X) \sin(\varrho_C\ \cdot)]\big) \textit{ libre}\ .$$

Démonstration. Le MCU admet une vitesse et une courbure constante non nulle ce qui implique que $\varrho(\cdot) = \varrho_C \neq 0$. Ensuite, on trouve, pour tout $t \in [0,1]$,

$$a(t) = \varrho_C^{-1} \sin(\varrho_C t)\ , \quad b(t) = \varrho_C^{-1}[1 - \cos(\varrho_C t)]\ , \quad u(t) = -\varrho_C^{-1} \sin(\varrho_C t)\ , \quad v(t) = -\varrho_C^{-1}\ .$$

Tous calculs effectués, on trouve que $\nabla_X s(X)(t)^{\mathrm{T}} M^{-1}(t)$ est colinéaire à

$$\begin{pmatrix} 1 & -\tan s(X)(t) & \varrho_C^{-1}[\tan s(X)(t) - \sin(\varrho_C t)] & \varrho_C^{-1}[1 - \tan s(X)(t) \sin(\varrho_C t)] \end{pmatrix}\ .$$

□

Contrairement au MRU, le critère de la proposition 7.1 n'est pas utilisable pour une trajectoire MCU puisqu'il fait intervenir la constante ϱ_C inconnue. On peut remarquer que si on connaissait cette constante, le problème perd un degré de liberté et on est alors ramené à un critère de liberté sur une famille de quatre fonctions, ce qui est cohérent avec le paramétrage d'un MCU.

Finalement, la formulation du problème de trajectographie par le système (1) est inappropriée puisqu'elle ne permet pas de retrouver le critère d'observabilité locale d'une dynamique simple comme le MCU. Ceci tient essentiellement à l'écriture explicite linéaire de la contrainte différentielle $\dot{X}(t) = A(t)\, X(t)$, écriture incomplète puisque la matrice $A(t)$ est inconnue. Pour avoir une formulation plus adéquate à la dynamique MCU, il faudrait écrire la contrainte différentielle sur la courbure et sur la vitesse

$$\frac{\mathrm{d}c}{\mathrm{d}t} = 0 \iff \frac{\mathrm{d}}{\mathrm{d}t}\,(\dot{x}\ddot{y} - \ddot{x}\dot{y}) = 0 \quad \text{et} \quad \dot{x}\ddot{x} + \dot{y}\ddot{y} = 0\ .$$

Mais la difficulté est toute autre car le système différentiel est implicite et non-linéaire. On peut l'écrire sous la forme synthétique suivante :

$$\begin{cases} f(\dot{X}, X) = 0\ , & (4) \\ Y = g(X)\ . & (5) \end{cases}$$

L'observabilité de la trajectoire, c'est-à-dire sa détermination à partir de l'observation Y équivaut à la *platitude* du système différentiel (4) avec comme sortie plate[3] (5), voir [48], [50] [49]. La question de la platitude différentielle est un problème difficile qui n'est pas abordé dans ce travail. Cependant, on peut préciser que la dynamique MCU est localement observable sous réserve de

[3] Le système différentiel $f(\dot{x}, x) = 0$ est plat avec comme sortie plate y si il existe une fonction ϕ et un entier p tel que $f(\dot{x}, x) = 0 \iff x = \phi(y, \dot{y}, \cdots, y^{(p)})$.

manœuvre, donc le système différentiel (4) pour la contrainte de courbure constante doit être plat. En revanche, on constate dans la section suivante que la contrainte de vitesse constante ne suffit pas à rendre la trajectoire observable et donc, le système différentiel afférent n'est pas plat.

7.3 Observabilité à vitesse constante

En supposant que la cible est à vitesse constante, sa trajectoire est-elle observable ? Autrement dit, parmi l'ensemble des trajectoires à vitesse constante, y a-t-il une unique trajectoire qui corresponde aux observations non bruitées ?

On ne répond pas catégoriquement à cette question, au sens où on ne prouve pas qu'une trajectoire à vitesse constante est ou n'est pas observable. En revanche, on présente des simulations suffisamment pertinentes pour se convaincre qu'une trajectoire cible à vitesse constante n'est pas observable.

On suppose que la cible est telle que la condition (H_+) est satisfaite. Notons, pour tout $t \in [0,1]$,

$$\alpha(t) \stackrel{\text{def}}{=} \tan \Psi[S(t),t] = [S_1(t) - O_1(t)]/[S_2(t) - O_2(t)] \ .$$

La trajectoire cible $t \mapsto S(t)$ vérifie les équations

$$\begin{cases} \dot{x}^2 + \dot{y}^2 = v^2 \ , & (6) \\ x - O_1 = \alpha(y - O_2) \ . & (7) \end{cases}$$

On peut en déduire des équations différentielles ordinaire pour x et pour y. Ainsi, on a

$$x = \alpha(y - O_2) + O_1 \Longrightarrow \dot{x} = \dot{\alpha}(y - O_2) + \alpha(\dot{y} - \dot{O}_2) + \dot{O}_1$$

et la contrainte de vitesse constante donne

$$\dot{x}^2 + \dot{y}^2 = v^2 \iff [\alpha\dot{y} + r(y)]^2 + \dot{y}^2 = v^2 \ ,$$

ce qui équivaut à

$$(\alpha^2 + 1)\dot{y}^2 + 2\alpha r(y)\dot{y} + r(y)^2 - v^2 = 0 \ , \tag{8}$$

avec $r(y) = \dot{\alpha}(y - y_p) + \dot{O}_1 - \alpha\dot{O}_2$. On obtient donc une équation du second degré en $\dot{y}$. En pratique, on constate sur les simulations que, pour un jeu de trajectoires cible et porteur avec des dynamiques plausibles (MRU, MCU ou MRU-MCU par morceaux), le discriminant de (8) n'est jamais nul. On voit également que le choix du signe de la racine pour la résolution de (8) est fonction de la disposition des trajectoires. Par la suite, on admettra que le discriminant est positif et que l'on connaît le signe de la racine. Ainsi, on obtient

$$\dot{y} = F(y,v) \ , \quad \text{avec} \quad F(y,v) = (\alpha^2+1)^{-1}\{-\alpha r(y) + \varepsilon[\alpha^2 r^2(y) - (\alpha^2+1)(r^2(y) - v^2)]\}^{1/2} \ .$$

et $\varepsilon \in \{\pm 1\}$. On peut tout aussi bien privilégier la fonction x et on obtient de la même une équation différentielle pour x. On procède pour les simulations à la méthode de résolution par l'algorithme d'Euler, dont on sait la convergence sous les hypothèses adéquates.

Bien entendu, on pourrait tout-à-fait retenir des méthodes de résolution à des ordres plus élevés comme celles implémentées directement dans `Matlab`, voir par exemple `ODE23`, `ODE45`. Compte-tenu de la précision de l'échantillonnage $\max_{k\in\{1,\cdots,n\}} |t_{k+1} - t_k|$, on s'est contenté, pour

Algorithme 1 Algorithme d'Euler pour la reconstruction de trajectoires

Entrées: $y(t_0), v \in \mathbb{R}$
Sorties: la séquence des points $(x(t_{0:n}), y(t_{0:n}))$
  $x(t_0) \leftarrow \alpha(t_0)(y(t_0) - O_2(t_0)) + O_1(t_0)$
  pour $k = 1$ **à** n **faire**
    $y(t_k) \leftarrow y(t_{k-1}) + (t_k - t_{k-1})F(y(t_{k-1}), v)(t_k)$
    $x(t_k) \leftarrow \alpha(t_k)(y(t_k) - O_2(t_k)) + O_1(t_k)$
  fin pour
  Retourner $(x(t_{0:n}), y(t_{0:n}))$.

les simulations qui suivent, de la méthode d'Euler qui s'avère très suffisante pour l'usage qu'on en a.

Dans l'ensemble des simulations qui suivent, le porteur est à vitesse constante de $250\,\mathrm{m/s}$ et effectue des manœuvres de l'ordre de $50\,\mathrm{m/s^2}$. Le porteur collecte 1000 observations sans bruit de mesure sur une durée de $20\,\mathrm{s}$ à intervalle de temps réguliers.

Pour la première simulation, la cible suit un MRU à $270\,\mathrm{m/s}$. Sur la figure 7.1 sont représentées la trajectoire cible et la trajectoire porteur, mais également différentes trajectoires à vitesse constante et conformes à la cible pour les mesures angulaires. Les caractéristiques initiales de la cible sont $y(t_0) = 3.8\,\mathrm{km}$ et $v = 270\,\mathrm{m/s}$. les conditions initiales pour la résolution de l'équation différentielle (8) sont données dans le tableau 7.2.

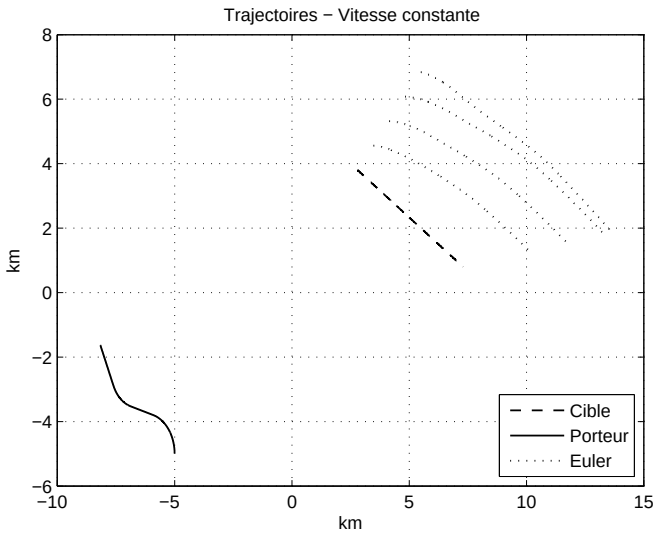


Fig. 7.1 – Trajectoires ambiguës à vitesse constante

$y(t_0)$ km	v m/s
4.56	378
5.32	432
6.08	486
6.84	486

Fig. 7.2 – Conditions initiales

Il n'y a pas, pour tout couple de conditions initiales $(y(t_0), v)$, une trajectoire à vitesse constante qui soit ambiguë à la cible. En effet, si on s'écarte de la trajectoire cible et si la vitesse choisie n'est pas suffisante ou au contraire si elle est trop élevée, la trajectoire pour cette vitesse ne pourra pas suivre les mesures angulaires imposées. C'est le cas, par exemple, pour le scénario de la première simulation si l'on prend $y(t_0) = 4.56\,\mathrm{km}$ et $v = 270\,\mathrm{m/s}$. Concrètement, on observe l'impossibilité de la trajectoire au cours de la résolution numérique avec l'apparition d'un discriminant négatif pour l'équation du second degré (8).

Pour la deuxième simulation, la cible suit un MCU à $270\,\mathrm{m/s}$ avec une accélération de $7\,\mathrm{m/s^2}$. On présente sur la figure 7.3 les trajectoires ambiguës à la cible avec les initialisations mentionnées

dans le tableau 7.4.

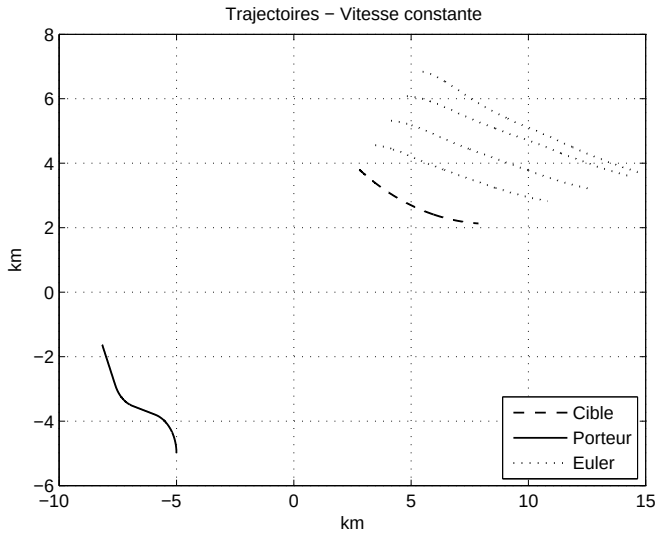


FIG. 7.3 – Trajectoires ambiguës à vitesse constante

$y(t_0)$ km	v m/s
4.56	378
5.32	432
6.08	486
6.84	486

FIG. 7.4 – Conditions initiales

Ces simulations illustrent le fait qu'une cible à vitesse constante n'est pas observable en trajectographie passive et répondent par là-même au problème de la platitude différentiel formulé en section 7.2.

Imposer une ordonnée initiale $y(t_0)$ ne suffit pas non plus à caractériser une trajectoire ambiguë à vitesse constante comme le montre la simulation présentée en figure 7.5.

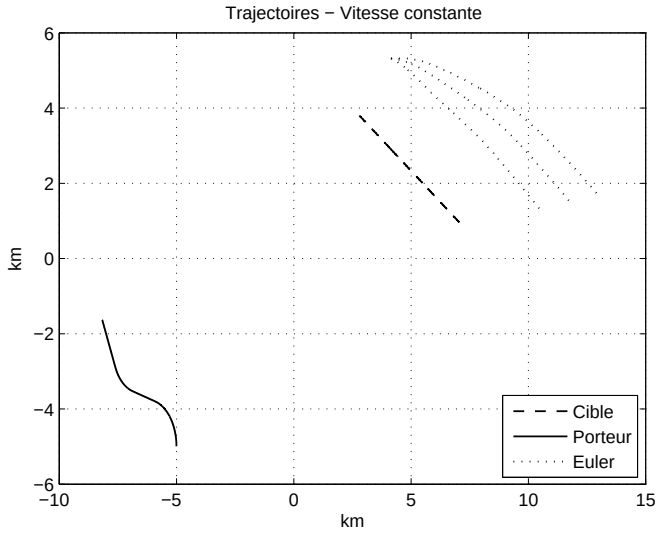


FIG. 7.5 – Trajectoires ambiguës à ordonnée initiale fixée

$y(t_0)$ km	v m/s
5.32	378
5.32	432
5.32	486

FIG. 7.6 – Conditions initiales

De même, imposer une vitesse initiale v ne suffit pas à caractériser une trajectoire ambiguë à vitesse constante comme le montre la simulation présentée en figure 7.7.

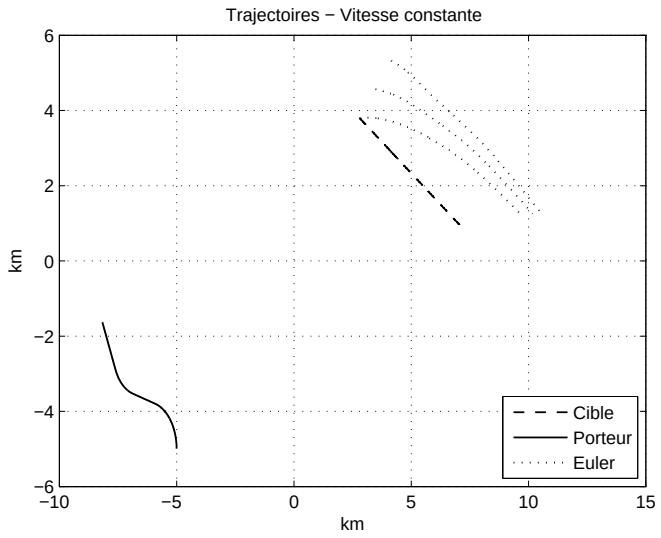


FIG. 7.7 – Trajectoires ambiguës à vitesse constante

$y(t_0)$ km	v m/s
3.8	378
4.56	378
5.32	378

FIG. 7.8 – Conditions initiales

7.4 Vers un critère d'observabilité

Malgré les résultats d'observabilité négatifs, la contrainte de vitesse constante semble être une condition appropriée pour séparer les trajectoires. Si l'on considère le cas très simple d'un porteur dont la trajectoire est composée de deux MRU, toute trajectoire construite par l'algorithme d'Euler et qui n'est pas la trajectoire cible va présenter une discontinuité d'orientation simultanément avec le porteur. Pour une trajectoire porteur de classe $\mathscr{C}^1$ avec une manœuvre franche, on observe sur les trajectoires ambiguës un changement de régime aux instants de manœuvre du porteur. C'est le cas de la simulation présentée en figure 7.9. Le porteur est en MRU puis fait une manœuvre avec une accélération très forte de $90\,\mathrm{m/s^2}$ et reprend un régime MRU. On observe alors nettement des changements dans l'allure des trajectoires ambiguës.

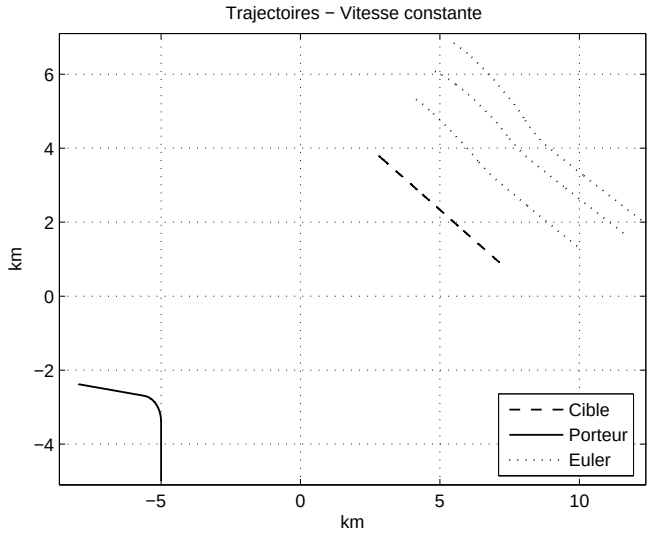


FIG. 7.9 – Trajectoires ambiguës à vitesse constante

$y(t_0)$ km	v m/s
5.32	351
6.08	405
6.84	432

FIG. 7.10 – Conditions initiales

D'une certaine manière, une trajectoire ambiguë est « corrélée » à la trajectoire porteur puisqu'elle présente des simultanéités de changement de régime, simultanéités qui n'ont pas lieu d'être pour la cible dont la trajectoire est décidée indépendamment de celle du porteur. Cette remarque a motivé à un grand nombre d'expérimentations numériques pour mesurer la corrélation des courbes ambiguës avec la trajectoire porteur.

Pour deux séquences $u_{1:n}$ et $v_{1:n}$ de $\mathbb{R}^n$, la fonction de covariance est définie par

$$\hat{\gamma}(u,v)(h) \stackrel{\text{def}}{=} \begin{cases} n^{-1}\sum_{k=1}^{n-|h|}(u_k-\bar{u}_n)(v_{k+|h|}-\bar{v}_n) & \text{si } h<0\ , \\ n^{-1}\sum_{k=1}^{n-|h|}(u_{k+|h|}-\bar{u}_n)(v_k-\bar{v}_n) & \text{si } h\geq 0\ , \end{cases}$$

et la fonction de corrélation par

$$\hat{\rho}(u,v)(h) \stackrel{\text{def}}{=} \frac{\hat{\gamma}(u,v)(h)}{\sqrt{\hat{\gamma}(u,u)(0)\hat{\gamma}(v,v)(0)}}\ .$$

Le critère de corrélation qui a intuitivement le plus de sens est $|\hat{\rho}(u,v)(0)|$ puisqu'on cherche à déceler des simultanéités dans les changements de régime. Mais d'autres critères ont également été testés comme $\max_{|h|\leq n}|\hat{\rho}(u,v)(h)|$ ou encore $\sum_{|h|\leq n}|\hat{\rho}(u,v)(h)|$. On écrit de manière générique $c(u,v)$ le critère utilisé. Pour une trajectoire $X=(x,y)$, on note $\rho(X,O)$ la grandeur qui définit une des stratégies de corrélation pour les grandeurs physiques des trajectoires, prise dans la liste

$$\Big\{c(x,O_1)+c(y,O_1),\ c(\dot{x},\dot{O}_1)+c(\dot{y},\dot{O}_2),\ c(\ddot{x},\ddot{O}_1)+c(\ddot{y},\ddot{O}_2),$$
$$c(\dot{x},\dot{O}_1)\times c(\dot{y},\dot{O}_2),\ c(\dot{x},\dot{x}_p)\times c(\dot{y},\dot{O}_2)+c(\ddot{x},\ddot{O}_1)\times c(\ddot{y},\ddot{O}_2),\cdots\Big\}\ ,$$

l'objectif étant de trouver la stratégie telle que,

$$S=\arg\min_{X\in\mathsf{A}}\rho(X,O)\ ,$$

où A désigne l'ensemble des trajectoires ambiguës à la cible et à vitesse constante. On a ainsi procédé à un grand nombre de simulations en testant les corrélations sur les positions, les vitesses, les accélérations sans réussir à départager, de manière significative pour différents types de dynamiques, la trajectoire cible des autres trajectoires ambiguës.

On a également comparé les corrélations des phases des vecteurs vitesses, c'est-à-dire les orientations de ces vecteurs. La phase de la vitesse semble être un critère adapté pour déceler les liens entre le porteur et une trajectoire ambiguë. En effet, lorsque le porteur manœuvre, la phase de sa vitesse varie fortement et on peut raisonnablement penser qu'il en est de même pour une trajectoire différente de la cible qui est contrainte de suivre les mêmes mesures angulaires. Pourtant, là encore, la mesure de corrélation sur la phase ne permet pas la distinction attendue. Sur la simulation présentée en figure 7.11, la cible est en MCU avec pour données initiales $y_0=3.8\,\mathrm{km}$ et $v=270\,\mathrm{m/s}$ et la trajectoire ambiguë a pour conditions initiales $y_0=4.18\,\mathrm{km}$ et $v=378\,\mathrm{m/s}$. On examine le critère de corrélation $\hat{\rho}(\phi,\phi_O)(0)$ où ϕ et ϕ_O désignent respectivement les phases des trajectoires X et O.

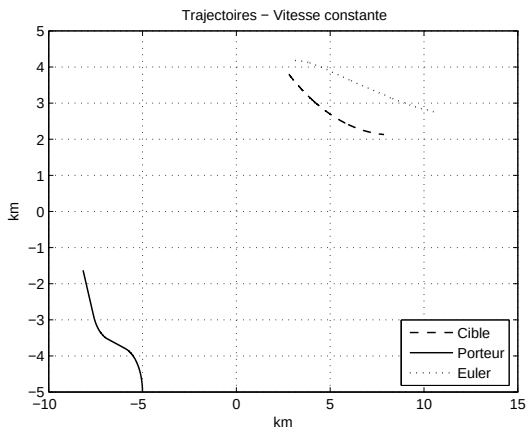


FIG. 7.11 – Trajectoire à vitesse constante

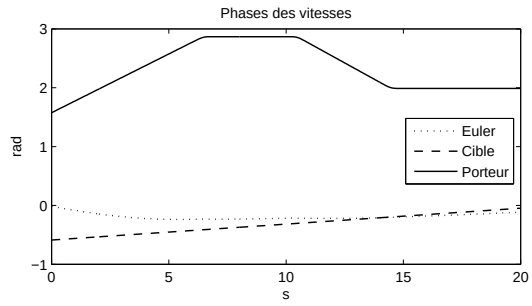


FIG. 7.12 – Phase des vitesses

On présente une autre simulation en figure 7.13 illustrant le même phénomène. La cible est en MRU-MCU avec pour données initiales $y_0 = 3.8\,\mathrm{km}$ et $v = 270\,\mathrm{m/s}$ et la trajectoire ambiguë a pour conditions initiales $y_0 = 1.9\,\mathrm{km}$ et $v = 404\,\mathrm{m/s}$.

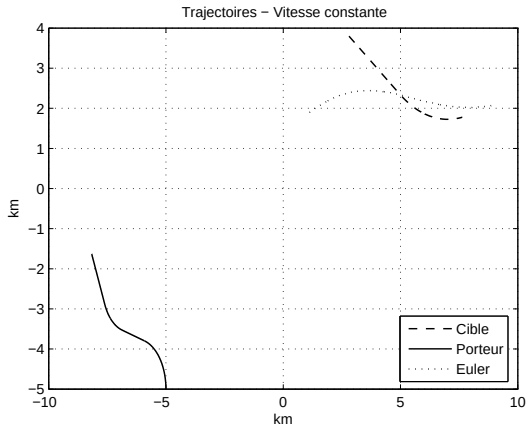


FIG. 7.13 – Trajectoire à vitesse constante

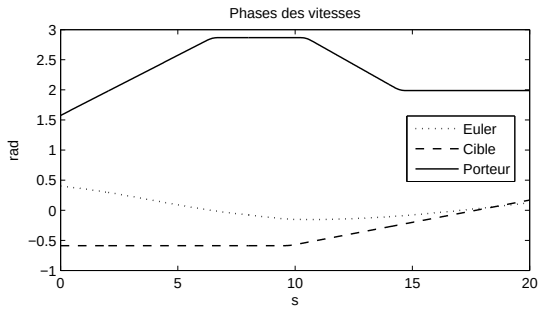


FIG. 7.14 – Phase des vitesses

Bien que l'examen des phases des vitesses semble, pour les raisons mentionnées précédemment, un bon critère, on constate un phénomène de lissage pour certaines trajectoires ambiguës qui rend indécelable les simultanéités de changement de régime. Comme on cherche à observer des variations franches, la dérivée des phases est sans doute un indicateur plus pertinent que les phases elles-mêmes pour détecter les simultanéités d'évolution. Cependant, la corrélation est une mesure encore insuffisante pour discerner la cible de trajectoires ambiguës.

En revanche, le critère $(X, O) \mapsto n^{-1} \sum_{k=1}^{n} \dot{\phi}^2(t_k) \dot{\phi_O}^2(t_k)$ proposé par Michel Prenat, semble prometteur et fournit de bons résultats pour des dynamiques de trajectoires assez générales. Si la fréquence d'échantillonnage est suffisamment fine, on pourrait également envisager un critère qui porte sur la dérivée seconde des phases, par exemple $(X, O) \mapsto n^{-1} \sum_{k=1}^{n} |\ddot{\phi}(t_k) \ddot{\phi_O}(t_k)|$, comme le suggèrent les simulations présentées en figure 7.15 et 7.17 avec le tracé des dérivées premières et secondes en figure 7.16 et 7.18.

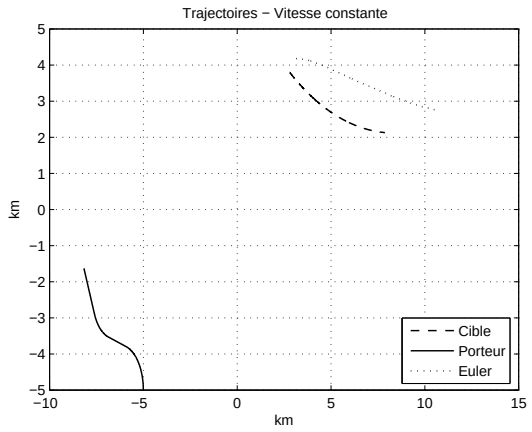


Fig. 7.15 – Trajectoire à vitesse constante

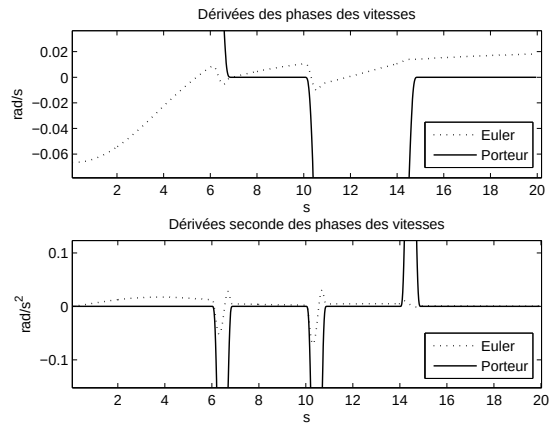


Fig. 7.16 – Dérivées première et seconde des phases des vitesses

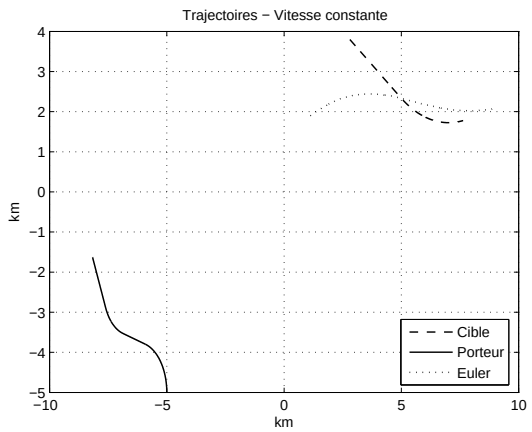


Fig. 7.17 – Trajectoire à vitesse constante

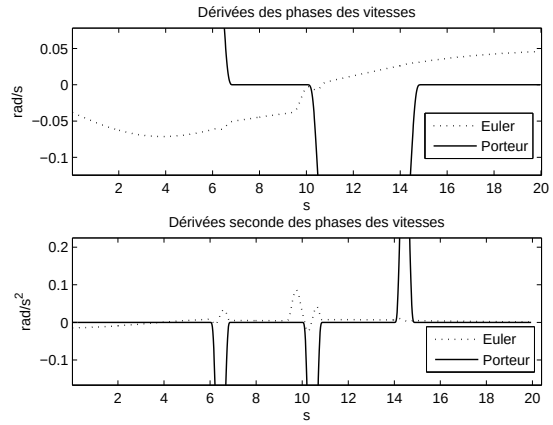


Fig. 7.18 – Dérivées première et seconde des phases des vitesses

7.5 Perspectives

Les expérimentations numériques confortent l'idée que les dérivées premières ou secondes des phases des vitesses sont une grandeur clé pour tester la simultanéité des changements entre une trajectoire conforme aux mesures angulaires et le porteur. Si ce fait se confirme, on pourrait construire un critère qui permette effectivement de discerner la trajectoire cible des autres trajectoires à vitesse constante qui lui sont ambiguës. Même si l'on atteint cet objectif, le problème est loin d'être clôt pour autant. Pour déterminer la trajectoire qui minimise un certain critère, il faudrait parcourir l'ensemble des trajectoires à vitesse constante qui fournissent les mêmes mesures angulaires que la cible. Or, chaque trajectoire est construite par le schéma d'intégration d'Euler. Il faudrait donc balayer toutes les trajectoires définies sur un ensemble donné de conditions initiales. On imagine aisément que le coût de calcul d'un tel algorithme est rédhibitoire. Dans le cas d'observations réelles bruitées, ce procédé s'inscrirait bien évidemment dans la catégorie des méthodes hors-ligne. De plus, un débruitage est nécessaire avant toute intégration ce qui alourdit un peu plus le procédé et dégrade la visibilité des simultanéités des changements

de régime. Enfin, on peut également souligner qu'en pratique, un mobile n'est jamais exactement à vitesse constante mais plutôt autour d'une valeur moyenne. Ainsi, une modélisation plus adéquate pourrait être une inclusion différentielle

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}^2+\dot{y}^2)\right| \leq \varepsilon \ ,$$

ou encore, écrire que le produit scalaire $\langle(\dot{x},\dot{y}),(\ddot{x},\ddot{y})\rangle$ est un bruit blanc gaussien

$$\langle(\dot{x},\dot{y}),(\ddot{x},\ddot{y})\rangle\,\mathrm{d}t = \mathrm{d}W_t \ ,$$

ce qui nous fait basculer vers des problèmes de complexité très supérieure.

Troisième partie

Estimation

Chapitre 8

Trajectoires paramétriques

Sommaire

8.1 Introduction 75
8.2 Comportement asymptotique 76
8.3 Test d'adéquation paramétrique 79
8.4 Simulations 80
8.5 Conclusion 85
8.6 Annexe au test d'adéquation paramétrique 86

8.1 Introduction

Lors de l'étude d'observabilité, on se place dans le cadre idéal de mesures parfaites, c'est-à-dire sans bruit de mesure. Le problème d'estimation adresse le cas où les observations sont des mesures bruitées, modélisation de la situation réelle par l'acquisition des données par le porteur. L'observabilité est le préalable à l'étude du problème d'estimation. On s'intéresse en effet au comportement asymptotique des estimateurs, quand le nombre d'observations croît pour des observations discrètes ou quand le niveau de bruit de mesure décroît pour des observations à temps continu. Le comportement asymptotique nous ramène, d'une certaine manière, au cadre idéal des mesures non bruitées en faisant s'évanouir l'aléa dans les équations.

On considère un modèle paramétrique $\{S_\theta\}_{\theta\in\Theta}$ qui satisfait la condition :

$$(\mathrm{H}_+) \qquad \forall(\theta,t)\in\Theta\times[0,1] \qquad S_{\theta,2}(t)-O_2(t)>0.$$

Comme dans les sections précédentes, on note $m(\theta,t)=\Psi[S_\theta(t),t]$ où l'on rappelle que Ψ est défini, pour tout $t\in[0,1]$, par $\Psi[(x_1,x_2),t]=\mathrm{angle}[x_1-O_1(t),x_2-O_2(t)]$, ce qui s'écrit, sous la condition (H_+),

$$\Psi[(x_1,x_2),t]=\arctan[x_1-O_1(t)]/[x_2-O_2(t)]\ .$$

Les observations à temps discret sont données par

$$Y_k=m(\theta,t_k)+\sigma\varepsilon_k\ , \tag{1}$$

avec $t_k=k/n$, $\sigma>0$ et $\{\varepsilon_k\}_{k\geq 0}$ est une séquence de variables gaussiennes standards et le modèle

d'observations à temps continu décrit, pour tout $t \in [0,1]$, par l'équation différentielle stochastique

$$\mathrm{d}Y_t = m(\theta, t)\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t\ , \tag{2}$$

avec $\varepsilon > 0$ et $(W_t)_{t\in[0,1]}$ est un processus de Wiener. On note indifféremment $(\mathbb{P}_\theta)_{\theta\in\Theta}$ le modèle statistique associé aux observations discrètes (1) ou à temps continu (2). Pour le modèle d'observations à temps discret, on définit

$$M_n(\theta) \stackrel{\text{def}}{=} n^{-1}\sum_{k=1}^{n}\{Y_k - m(\theta, t_k)\}^2\ ,$$

et l'estimateur du maximum de vraisemblance par

$$\hat{\theta}_n \stackrel{\text{def}}{=} \arg\max_{\theta\in\Theta} p_\theta(Y_1, \cdots, Y_n) = \arg\min_{\theta\in\Theta} M_n(\theta)\ .$$

Pour le modèle d'observations à temps continu, d'après le théorème de Girsanov [61], le rapport de log-vraisemblance $\log \frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\mathbb{P}_{\theta^*}}$ est l'opposé de la quantité définie par

$$M^\varepsilon(\theta) \stackrel{\text{def}}{=} \int_0^1 [m(\theta,t) - m(\theta^*,t)]^2\,\mathrm{d}t - 2\varepsilon\int_0^1 [m(\theta,t) - m(\theta^*,t)]\,\mathrm{d}W_t\ .$$

On souhaite définir l'estimateur du maximum de vraisemblance comme minimiseur de M^ε. Une précaution est à prendre car le paramètre θ^* est inconnu et on ne peut donc pas faire d'inférence statistique en utilisant M^ε. On contourne cette difficulté en introduisant la fonctionnelle stochastique

$$C^\varepsilon(\theta) \stackrel{\text{def}}{=} \int_0^1 m(\theta,t)^2\,\mathrm{d}t - 2\int_0^1 m(\theta,t)\,\mathrm{d}Y_t\ .$$

Cette quantité est bien accessible au statisticien. Rappelons en effet qu'observer dans le modèle (2) signifie que l'on a accès à toutes les intégrales stochastiques $\displaystyle\int_0^1 g(t)\,\mathrm{d}Y_t$ pour tout élément $g \in \mathrm{L}^2([0,1])$. On peut alors définir son minimiseur qui est un estimateur, comme fonction de l'observation et qui est précisément l'estimateur du maximum de vraisemblance

$$\hat{\theta}^\varepsilon \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} C^\varepsilon(\theta) = \arg\min_{\theta\in\Theta} M^\varepsilon(\theta)\ .$$

Dans ce chapitre, on étudie sous l'hypothèse d'observabilité, hypothèse usuelle en trajectographie passive, le comportement asymptotique de l'estimateur du maximum de vraisemblance pour les modèles d'observations (1) et (2). Les propriétés de consistance et de normalité asymptotique sont présentées en section 8.2. On aborde en section 8.3 la question de l'appartenance au modèle de régression paramétrique. La section 8.4 est dédiée aux simulations qui illustrent les résultats asymptotiques mais permettent également de dévoiler le caractère peu robuste de l'estimation.

8.2 Comportement asymptotique

Cette section est consacrée à l'étude des propriétés asymptotiques de l'estimateur du maximum de vraisemblance sous l'hypothèse d'observabilité. On est dans un cadre d'estimation bien maîtrisé puisqu'il s'agit d'un problème de régression paramétrique avec des bruits d'observations gaussiens i.i.d. [67], [52], [53], [61]. Les résultats qui suivent sont connus ; on en rappelle

les preuves pour assurer la cohérence de l'ensemble et aussi car les méthodes employées pour le modèle semiparamétrique décrit au chapitre 11 en sont inspirées. Signalons que le schéma de preuve de la proposition 8.4 diffère des schémas classiques [61] et introduit les idées qui permettront l'étude de robustesse présentée au chapitre 10.

On fait les hypothèses suivantes :

(E1) l'ensemble Θ compact et convexe[1] ;

(E2) L'application $(\theta, t) \mapsto m(\theta, t)$ est de classe $\mathscr{C}^2$.

Proposition 8.1. *Sous les hypothèses (E1) et (E2), étant donnée une trajectoire paramétrique S_{θ^*} observable, l'estimateur du maximum de vraisemblance pour le modèle d'observations à temps discret* (1) *est consistant, c'est-à-dire*

$$\hat{\theta}_n = \theta^* + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

Démonstration. On définit

$$M(\theta) \stackrel{\text{def}}{=} \int_0^1 \{m(\theta, t) - m(\theta^*, t)\}^2 \, \mathrm{d}t + \sigma^2 \ .$$

D'après l'inégalité de Bienaymé-Tchebycheff, on a clairement, pour tout $\theta \in \Theta$,

$$M_n(\theta) - M(\theta) = \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

Vérifions que cette convergence a lieu uniformément en $\theta \in \Theta$. Pour tous $\theta, \ \theta' \in \Theta$, on a

$$M_n(\theta) - M_n(\theta') =$$
$$n^{-1} \sum_{k=1}^n \{2m(\theta^*, t_k) - m(\theta, t_k) - m(\theta', t_k)\}\{m(\theta, t_k) - m(\theta', t_k)\} + 2n^{-1} \sum_{k=1}^n \varepsilon_k \{m(\theta, t_k) - m(\theta', t_k)\} \ ,$$

ce qui implique

$$\sup_{\substack{(\theta,\theta') \in \Theta^2 \\ \|\theta - \theta'\| \leq \delta}} |M_n(\theta) - M_n(\theta')| \leq \omega(\delta) \times 2n^{-1} \sum_{k=1}^n \left\{ |m(\theta^*, t_k)| + \sup_{(\theta,t) \in \Theta \times [0,1]} |m(\theta, t_k)| + |\varepsilon_k| \right\} \ ,$$

où ω désigne le module d'uniforme continuité de $(\theta, t) \mapsto m(\theta, t_k)$. On conclut en appliquant le théorème de convergence d'un processus [110, Theorem 18.14]. Comme l'hypothèse d'observabilité garantit l'unicité du minimum de M, d'après le théorème des M-estimateurs [110, Theorem 5.7], on en déduit la consistance pour $\hat{\theta}_n$. □

Proposition 8.2. *Sous les hypothèses (E1) et (E2), étant donné une trajectoire paramétrique S_{θ^*} observable, l'estimateur du maximum de vraisemblance pour le modèle d'observations à temps continu* (2) *est consistant, c'est-à-dire*

$$\hat{\theta}^\varepsilon = \theta^* + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

Démonstration. On a

$$\sup_{\theta \in \Theta} |M^\varepsilon(\theta) - M(\theta)| = 2\varepsilon \sup_{\theta \in \Theta} \left| \int_0^1 [m(\theta, t) - m(\theta^*, t)] \, \mathrm{d}W_t \right| \ ,$$

[1] C'est le cas en pratique.

ce qui implique

$$\lim_{\varepsilon\to 0}\sup_{\theta\in\Theta}|M^{\varepsilon}(\theta)-M(\theta)|=0\ ,\quad \mathbb{P}_{\theta^*}-p.s.$$

et on conclut par le théorème des M-estimateurs [110, Theorem 5.7]. □

On définit la matrice d'information de Fisher associée au modèle (1) par

$$I_n(\theta)\stackrel{\text{def}}{=}\sigma^{-2}\sum_{k=1}^{n}\nabla_\theta m(\theta,t_k)\nabla_\theta m(\theta,t_k)^{\mathrm{T}}\ , \tag{3}$$

et celle associée au modèle (2) par

$$I(\theta)\stackrel{\text{def}}{=}\int_0^1\nabla_\theta m(\theta,t)\nabla_\theta m(\theta,t)^{\mathrm{T}}\,\mathrm{d}t\ . \tag{4}$$

Proposition 8.3. *Sous les hypothèses (E1) et (E2), étant donnée une trajectoire paramétrique S_{θ^*} observable telle que $I_n(\theta^*)$ est inversible pour un certain n entier, on a*

$$\sqrt{n}\big(\hat\theta_n-\theta^*\big)=n^{-1}\sum_{k=1}^{n}\nabla_\theta m(\theta,t_k)\varepsilon_k+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ .$$

En particulier, on la convergence en loi

$$\sqrt{n}\big(\hat\theta_n-\theta^*\big)\rightsquigarrow\mathcal{N}\left(0,\sigma^2I^{-1}(\theta^*)\right)\ .$$

Démonstration. Il s'agit du schéma de preuve classique due à Wald [117] dont on ne fait qu'esquisser les grandes lignes. Le développement de Taylor avec reste intégral nous donne

$$\nabla_\theta M_n(\hat\theta_n)=0=\nabla_\theta M_n(\theta^*)+\int_0^1\mathrm{D}_{\theta,\theta}\,M_n(\theta^*+s(\hat\theta_n-\theta^*))\,\mathrm{d}s\,(\hat\theta_n-\theta^*)\ .$$

Le calcul direct montre que

$$\nabla_\theta M_n(\theta^*)=-2n^{-1}\sum_{k=1}^{n}\nabla_\theta m(\theta^*,t_k)\varepsilon_k\ .$$

Par des arguments de continuité uniforme et grâce à la consistance de $\hat\theta_n$, on vérifie que

$$\mathrm{D}_{\theta,\theta}\,M_n(\theta^*+s(\hat\theta_n-\theta^*))=\mathrm{D}_{\theta,\theta}\,M_n(\theta^*)+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ ,$$

et par suite, on obtient

$$\left[2I_n(\theta^*)+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\right]\sqrt{n}\big(\hat\theta_n-\theta^*\big)=2n^{-1/2}\sum_{k=1}^{n}\nabla_\theta m(\theta,t_k)\varepsilon_k+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ .$$

Précisons enfin que

$$I_n(\theta^*)\text{ inversible}\Longrightarrow I(\theta^*)\text{ inversible}\ .$$

En effet, on a sans aucune difficulté

$$\left(\frac{\partial m}{\partial\theta_i}(\theta^*,t_{1:n})\right)_{1\le i\le d}\text{ libre}\Longrightarrow\left(\frac{\partial m}{\partial\theta_i}(\theta^*,\cdot)\right)_{1\le i\le d}\text{ libre}\ .$$

La convergence en loi est une conséquence du théorème de Lindeberg-Feller. □

Proposition 8.4. *Sous les hypothèses (E1) et (E2), Étant donnée une trajectoire paramétrique S_{θ^*} observable telle que $I(\theta^*)$ est inversible, il existe $\delta > 0$ tel que, pour tout $\varepsilon \in]-\delta, \delta[$,*

$$\varepsilon^{-1}\big(\hat{\theta}^\varepsilon - \theta^*\big) = I^{-1}(\theta^*) \int_0^1 \nabla_\theta m(\theta^*, t)\, \mathrm{d}W_t + \mathrm{o}(1)\ , \quad \mathbb{P} - p.s.$$

En particulier, on la convergence en loi

$$\varepsilon^{-1}\big(\hat{\theta}^\varepsilon - \theta^*\big) \rightsquigarrow \mathcal{N}\left(0, I^{-1}(\theta^*)\right)\ .$$

Démonstration. Par dérivation, on trouve

$$\mathrm{D}_{\theta,\theta}\, M^\varepsilon(\theta) = 2I(\theta) + 2\int_0^1 [m(\theta,t) - m(\theta^*,t)]\, \mathrm{D}_{\theta,\theta}\, m(\theta,t)\, \mathrm{d}t - 2\varepsilon \int_0^1 \mathrm{D}_{\theta,\theta}\, m(\theta,t)\, \mathrm{d}W_t\ , \quad \mathbb{P} - p.s.$$

et en particulier

$$\mathrm{D}_{\theta,\theta}\, M^0(\theta^*) = 2I(\theta^*)\ , \quad \mathbb{P} - p.s.$$

Alors, d'après le théorème des fonctions implicites, presque sûrement il existe $\delta > 0$, un voisinage de θ^* noté V_{θ^*} et une application $\bar{\theta}^{(\cdot)} :]-\delta, \delta[\to V_{\theta^*}$ de classe $\mathscr{C}^1$ tels que, pour tout (θ, ε) dans $V_{\theta^*} \times]-\delta, \delta[$,

$$\mathrm{D}_\theta\, M^\varepsilon(\theta) = 0 \iff \theta = \bar{\theta}^\varepsilon\ ,$$

et sa différentielle est donnée par

$$\mathrm{D}_\varepsilon\, \bar{\theta}^\varepsilon = -\,\mathrm{D}_{\theta,\theta}\, M^\varepsilon\left(\bar{\theta}^\varepsilon\right) \mathrm{D}_{\varepsilon,\theta}\, M^\varepsilon(\bar{\theta}^\varepsilon)\ .$$

Par définition de $\hat{\theta}^\varepsilon$, on a $\mathrm{D}_\theta\, M^\varepsilon(\hat{\theta}^\varepsilon) = 0$ presque sûrement. La preuve de la proposition 8.2 donne en réalité un résultat plus fort que celui annoncé puisqu'on a la consistance forte de $\hat{\theta}^\varepsilon$, à savoir

$$\lim_{\varepsilon \to 0} \hat{\theta}^\varepsilon = \theta^*\ , \quad \mathbb{P} - p.s.$$

Ainsi, presque sûrement, pour ε assez petit, on a $\hat{\theta}^\varepsilon = \bar{\theta}^\varepsilon$. Le développement de Taylor-Young au premier ordre nous donne finalement

$$\hat{\theta}^\varepsilon = \theta^* + \varepsilon I(\theta^*)^{-1} \int_0^1 \nabla_\theta m(\theta^*, t)\, \mathrm{d}W_t + \mathrm{o}(\varepsilon)\ , \quad \mathbb{P} - p.s.$$

□

Si la cible est en MRU, l'hypothèse d'observabilité suffit à elle seule puisqu'elle garantit l'inversibilité de la matrice de Fisher, voir proposition 5.2. Pour une cible MCU, l'observabilité et l'inversibilité de $I_n(\theta^*)$ sont vérifiées pour un porteur dont la trajectoire est non immobile à composantes polynomiales, voir propositions 5.11 et5.12.

8.3 Test d'adéquation paramétrique

Sachant les propriétés asymptotiques de l'estimateur du maximum de vraisemblance, on peut souhaiter mettre en place un test pour s'assurer que le problème de régression est bien un

problème de régression paramétrique. Autrement dit, on veut tester l'hypothèse

$$\mathbf{H_0} : \quad « \Phi(S) \in \{\Phi(S_\theta),\ \theta \in \Theta\} » ,$$

contre une alternative

$$\mathbf{H_1} : \quad « \Phi(S) = \Phi(S_{\theta'}) + \rho_n g » ,$$

pour une certaine fonction g avec $\theta' \in \Theta$ et une suite $\{\rho_k\}_{k\geq 0}$ qui tend vers zéro. On rappelle que Φ désigne la fonction mesure angulaire et qu'elle est définie sur l'ensemble des trajectoires par

$$\Phi : X \longmapsto \{\Psi[X(t), t]\}_{t\in[0,T]} \ .$$

Avec les résultats d'observabilité du chapitre 5, on peut montrer que les procédures de test d'adéquation paramétrique établies par Horowitz et al.[59] ou encore par Guerre et al.[57] sont valides en trajectographie passive.

La vérification des hypothèses techniques de [59] est détaillée en section 8.6.

8.4 Simulations

On présente dans cette section les résultats de simulations numériques obtenus pour différents scénarios de trajectoires cible et porteur. Ces simulations ont été réalisées sous `Matlab`. Un des points délicats est la minimisation de la fonctionnelle non convexe M_n pour obtenir une valeur correcte de l'estimateur. On utilise la fonction `fminsearch` en fixant à 2000 les options `MaxIter` et `MaxFunEvals`.

Pour la première simulation, on considère une cible en MRU à vitesse constante de $270\,\mathrm{m/s}$. Le porteur est à vitesse constante à $250\,\mathrm{m/s}$ et effectue des manœuvres de l'ordre[2] de $50\,\mathrm{m/s^2}$. Le porteur collecte 1000 observations sur une durée de $20\,\mathrm{s}$ à intervalle de temps réguliers. La variance du bruit d'observation est de $0.01\,\mathrm{rad}$. Sur la figure 8.1 sont représentées les trajectoires cibles et porteur ainsi que la trajectoire MRU de paramètre $\hat{\theta}_n$ (Bearings-Least-Square Estimator, BLSE). On trace également l'ellipse de confiance de la position finale pour un niveau de confiance de 95%. On peut calculer cette ellipse car la position à l'instant final est une fonction linéaire du paramètre. Ceci se généralise d'ailleurs pour tous les modèles paramétriques pour lesquels la trajectoire s'écrit dans une base.

[2] Il s'agit d'accélérations relativement élevées.

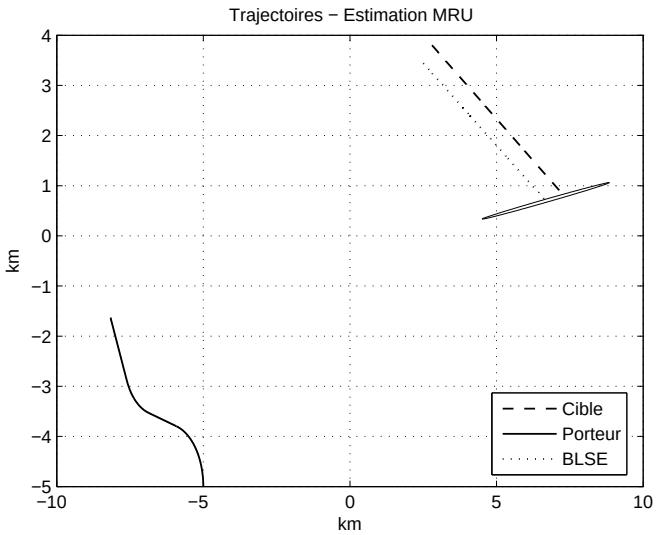


FIG. 8.1 – Estimation d'une trajectoire cible en MRU

Sur le tableau 8.1, on présente les intervalles de confiance à 95% pour les différentes coordonnées de θ^*. On note $|\mathrm{IC}(\cdot)|$ la taille de l'intervalle.

$\mathrm{IC}(\tilde{\theta}_{n,i})$		$\lvert\ \mathrm{IC}(\tilde{\theta}_{n,i})\rvert$
5.8863	8.7986	2.9123
0.5836	1.0725	0.4889
0.1932	0.2639	0.0706
-0.1777	-0.1167	0.0609

TAB. 8.1 – Intervalles de confiance à 95%

Les figures 8.2 et 8.3 sont des illustrations de la convergence en loi décrite par la proposition 8.3. On vérifie la normalité asymptotique de $\hat{\theta}_n$ sur les histogrammes des composantes de $\sqrt{n}(\hat{\theta}_n - \theta^*)$ superposés aux densités théoriques et les fonctions de répartition empirique superposées aux fonctions de répartition théoriques.

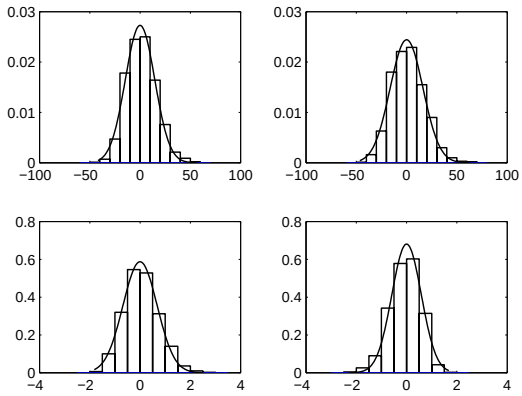


FIG. 8.2 – Histogrammes des composantes de $\sqrt{n}(\hat{\theta}_n - \theta^*)$

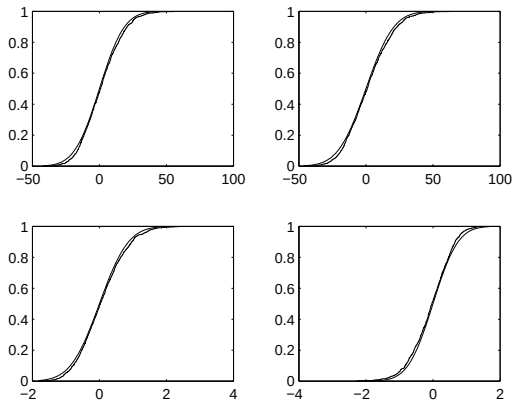


FIG. 8.3 – Fonctions de répartition empirique des composantes de $\sqrt{n}(\hat{\theta}_n - \theta^*)$

Pour la deuxième simulation, on considère une cible en MCU à vitesse constante de $270\,\mathrm{m/s}$ avec une accélération de giration de $7\,\mathrm{m/s^2}$. Le porteur a les mêmes caractéristiques que pour la première simulation. Sur la figure 8.4 sont représentées les trajectoires cibles et porteur ainsi que la trajectoire MCU de paramètre $\hat{\theta}_n$ (BLSE). Contrairement au cas du MRU, la position finale de la cible n'est pas une application linéaire du paramètre θ^*. Ceci est intrinsèque au paramétrage du MCU qui n'est pas linéaire en les paramètres. On présente le nuage des points pour la position finale obtenu par la méthode de Monte-Carlo.

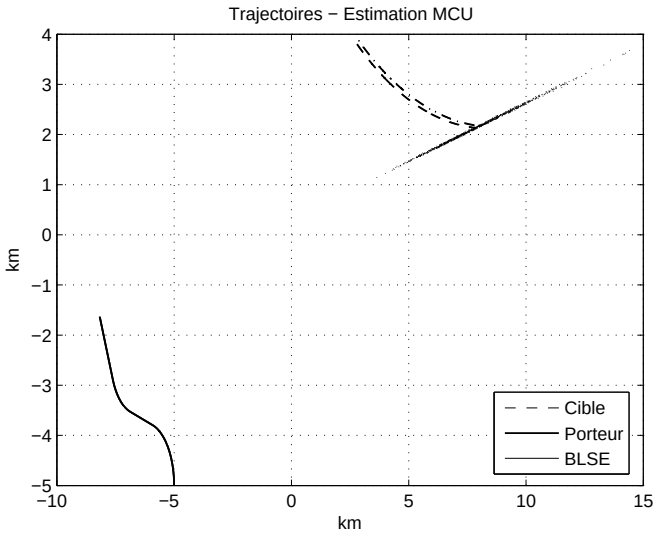


Fig. 8.4 – Estimation d'une trajectoire cible en CRU

Comme pour le MRU, on présente en figure 8.5 et 8.6 la normalité asymptotique de $\hat{\theta}_n$ avec histogrammes et fonctions de répartition empirique des composantes $\sqrt{n}(\hat{\theta}_n - \theta^*)$. Signalons que le comportement asymptotique est plus difficilement observable que dans le cas du MRU. Pour observer la normalité asymptotique de manière satisfaisante, le porteur collecte ici 2000 observations.

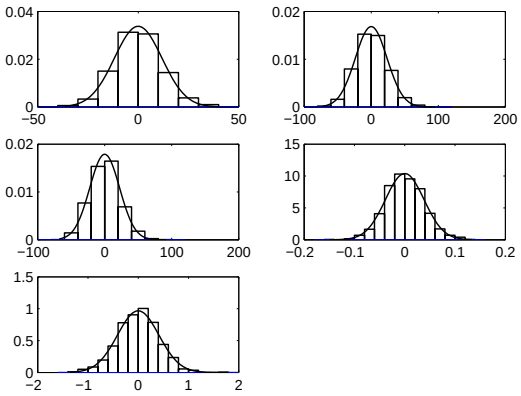


Fig. 8.5 – Histogrammes des composantes de $\sqrt{n}(\hat{\theta}_n - \theta^*)$

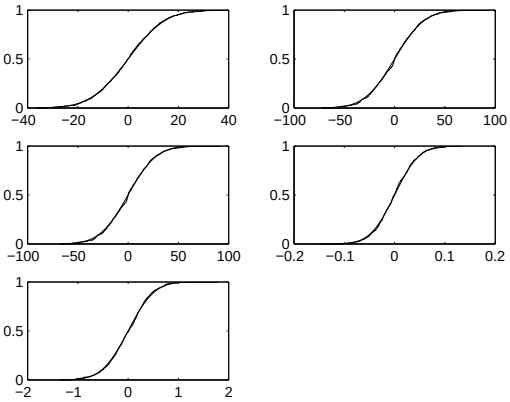


Fig. 8.6 – Fonctions de répartition empirique des composantes de $\sqrt{n}(\hat{\theta}_n - \theta^*)$

Pour la troisième simulation, la cible suit la même trajectoire MRU que pour la première

simulation et les caractéristiques de mesures sont également identiques. En revanche, le porteur effectue des manœuvres plus longues mais moins franches[3], de l'ordre de $30\,\mathrm{m/s^2}$. La conséquence est immédiate. L'ellipse de confiance est très largement étalée. On voit de manière flagrante combien la manœuvre du porteur est un élément clé dans la qualité de l'observation.

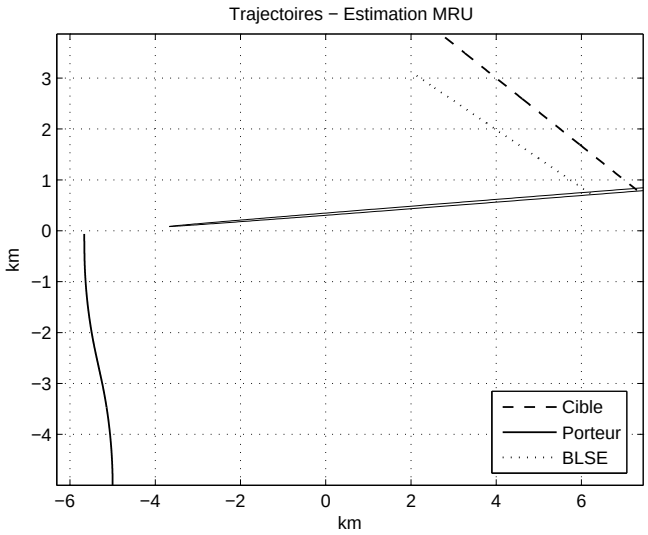


Fig. 8.7 – Faible niveau d'observabilité

IC($\tilde{\theta}_{n,i}$)		$\lvert$ IC($\tilde{\theta}_{n,i}$)$\rvert$
0.2598	13.5401	13.2803
0.3296	1.2048	0.8752
0.0777	0.3465	0.2688
-0.3441	0.0611	0.4052

Tab. 8.2 – Intervalles de confiance à 95%

Une manœuvre franche est donc nécessaire pour garantir un bon niveau d'observabilité et par là même une estimation convenable. Pour la quatrième simulation, le porteur est en MRU puis fait une manœuvre avec une accélération très forte[4] de $90\,\mathrm{m/s^2}$ et reprend un régime MRU.

[3]Comme on le voit sur la figure 8.7, le porteur manœuvre mais cherche à maintenir son cap. Cette trajectoire appelé *baïonnette* est un cas courant de manœuvre.

[4]Une telle accélération est possible en situation tactique mais pas réaliste pour une simple trajectographie.

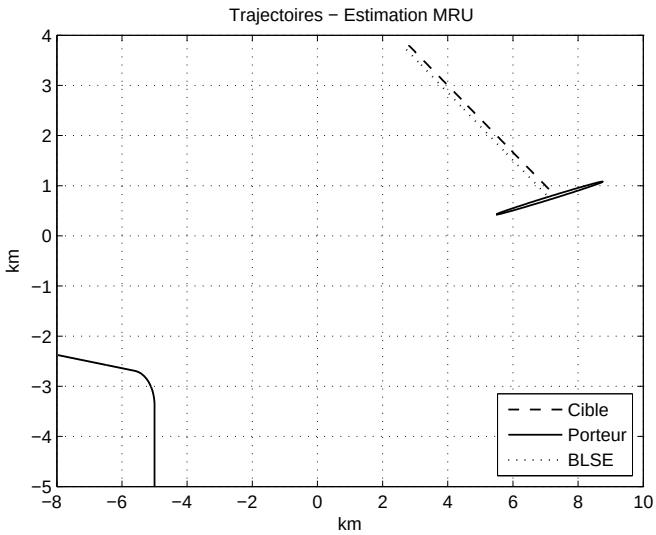


Fig. 8.8 – Manœuvre très forte

Comparée aux simulations 1 et 3, on voit d'après le tableau 8.3 que la manœuvre très franche de cette quatrième simulation améliore la pertinence de l'estimation.

IC($\tilde{\theta}_{n,i}$)		$\lvert$ IC($\tilde{\theta}_{n,i}$)$\rvert$
5.7536	7.9540	2.2005
0.4817	0.9248	0.4431
0.1780	0.2431	0.0650
-0.1618	-0.1290	0.0328

Tab. 8.3 – Intervalles de confiance à 95%

Ces simulations soulignent un des compromis de la trajectographie passive. Le porteur doit manœuvrer pour rendre la cible observable mais la manœuvre n'est pas quelconque car la trajectoire est avant tout une trajectoire opérationnelle. D'où la nécessité d'un arbitrage entre les contraintes opérationnelles et l'objectif de trajectographie.

Comme on peut le voir avec la troisième simulation, l'estimation est d'une grande sensibilité vis-à-vis de la trajectoire porteur. Mais ce n'est pas tout. L'estimation est également très sensible vis-à-vis du modèle paramétrique supposé. Si l'on procède à une estimation MRU sur une trajectoire cible qui est un MCU avec une courbure faible, on peut trouver un MRU très éloigné de la trajectoire réelle, ce qui choque l'intuition qui porterait à croire que l'estimation sera proche en norme euclidienne de la trajectoire réelle. La figure 8.9 illustre ce phénomène. Les positions finales de la cible et de la trajectoire estimée sont repérées par des disques.

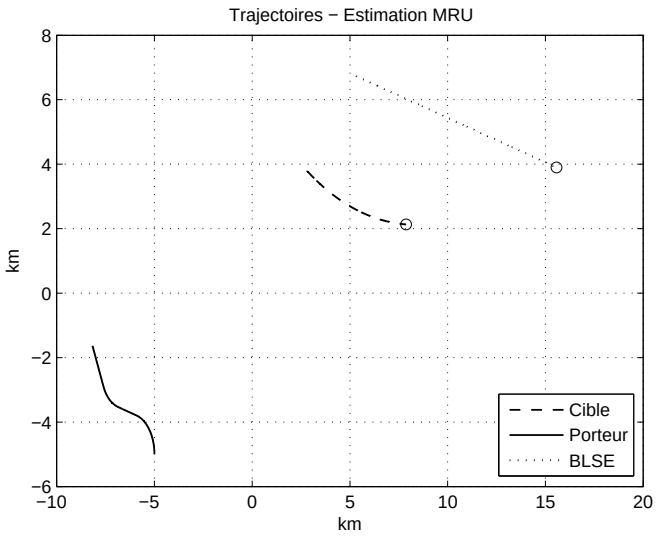


Fig. 8.9 – Sensibilité vis-à-vis du modèle paramétrique

Ce problème de *robustesse* est étudié en détail au chapitre 10 et met en lumière des situations de trajectographie pathologiques.

L'estimation dans les modèles paramétriques par morceaux décrite au chapitre 9 présente, comme on peut s'y attendre, des caractéristiques semblables.

8.5 Conclusion

Comme le prévoit la théorie asymptotique de la régression, l'observabilité d'une trajectoire paramétrique et l'inversibilité de la matrice de Fisher associée suffisent à garantir la consistance et la normalité asymptotique de l'estimateur du maximum de vraisemblance. Ces résultats théoriques sont à nuancer pour l'estimation en situation pratique. Observabilité et inversibilité de la matrice de Fisher sont des critères qualitatifs *binaires*, ils n'assurent pas la robustesse de l'estimation, comme le montrent les simulations de la section 8.4. Ceci s'avère très dommageable si, en plus, la trajectoire cible réelle n'est pas conforme au modèle d'état supposé, voir l'étude menée au chapitre 10. Ce comportement est le fait du caractère *mal posé* de la trajectographie passive. Pour une cible en MRU, le porteur peut manœuvrer et assurer l'observabilité mais avec une matrice d'information de Fisher presque singulière ce qui implique une estimation de qualité médiocre.

Ceci invite à considérer avec le plus grand intérêt les travaux sur la *commande optimale* pour l'observabilité de trajectoire paramétrique, voir [95], [83], [76], [34]. Cette thématique n'est pas abordée dans la présente thèse. En revanche le chapitre 10 ouvre des pistes pour de nouveaux critères utilisables en commande optimale.

Soulignons malgré tout que, pour des cas de manœuvres franches qui assurent l'identifiabilité du modèle statistique, l'estimation est de qualité et les hypothèses requises pour la mise en œuvre des tests d'adéquations au modèle paramétrique sont satisfaites, ce qui concerne une majorité de situations réelles en trajectographie passive.

8.6 Annexe au test d'adéquation paramétrique

Dans cette section, on s'intéresse à la vérification des hypothèses clés de l'article de Horowitz et al.[59] pour tester l'hypothèse

$$\mathbf{H_0}: \quad « \Phi(S) \in \{\Phi(S_\theta),\ \theta \in \Theta\} » ,$$

contre une alternative

$$\mathbf{H_1}: \quad « \Phi(S) = \Phi(S_{\theta'}) + \rho_n g » ,$$

On se place dans le cadre d'observations discrètes.

Proposition 1. *On suppose que les conditions (E1) et (E2) sont vérifiées. Pour un modèle de trajectoire $\{S_\theta\}_{\theta\in\Theta}$ tel que $\{\Phi(S_\theta)\}_{\theta\in\Theta}$ est identifiable et si toute trajectoire admet une matrice de Fisher inversible, alors les hypothèses 1 et 2 de [59] sont satisfaites.*

Preuve de la proposition 1. D'après (E2), les hypothèses de différentiabilité [5] de [59] sont satisfaites. D'après (E2) également, l'application $\theta \mapsto \left\| nI_n^{-1}(\theta) \right\|$ où $\| \ \|$ désigne la norme d'application linéaire est bien définie et continue. Par compacité de Θ, la quantité $\sup_{\theta\in\Theta} \left\| nI_n^{-1}(\theta) \right\|$ est donc minorée. Comme Θ est convexe, d'après la formule de Taylor-Lagrange avec reste intégral, on a pour tout $t \in [0,1]$

$$m(\theta+h,t) - m(\theta,t) = \langle \nabla_\theta m(\theta,t), h\rangle_{\mathbb{R}^d} + h\int_0^1 (1-u)\, \mathrm{D}_{\theta,\theta}\, m(\theta+uh,t)(h,h)\, \mathrm{d}t \ .$$

L'application $(t,\theta) \mapsto \mathrm{D}_{\theta,\theta}\, m(\theta,t)$ est continue sur le compact $[0,1]\times\Theta$ ce qui permet d'écrire

$$m(\theta+h,t) - m(\theta,t) = \langle \nabla_\theta m(\theta,t), h\rangle_{\mathbb{R}^d} + \mathrm{o}(\|h\|) \ ,$$

avec un $\mathrm{o}(\|h\|^2)$ uniforme en t. Par continuité de $(t,\theta) \mapsto \nabla_\theta m(\theta,t)$ sur le compact $[0,1]\times\Theta$, on a pour tout $h \in \mathbb{R}^d$ non nul

$$\begin{aligned}
\sum_{k=1}^n \|m(\theta+h,t_k) - m(\theta,t_k)\|^2 &= \|h\|^2 \sum_{k=1}^n \Big(\langle \nabla_\theta m(\theta,t_k), \|h\|^{-1}\, h\rangle_{\mathbb{R}^d} + \mathrm{o}(1) \Big)^2 \ , \\
&= \|h\|^2 \left(\sum_{k=1}^n \langle \nabla_\theta m(\theta,t_k), \|h\|^{-1}\, h\rangle^2_{\mathbb{R}^d} + \mathrm{o}(1) \right) \ .
\end{aligned}$$

L'inversibilité de la matrice de Fisher équivaut à la liberté de la famille $\left(\frac{\partial m}{\partial \theta_i}(\theta, t_{1:n})\right)_{1\le i\le d}$ qui implique la liberté de la famille de fonctions $\left(\frac{\partial m}{\partial \theta_i}(\cdot)\right)_{1\le i\le d}$. Pour tout u pris dans la boule unité de $\mathbb{R}^d$, d'après l'inégalité de Cauchy-Schwarz, on obtient

$$n^{-1} \sum_{k=1}^n \langle \nabla_\theta m(\theta,t_k), u\rangle^2_{\mathbb{R}^d} = \int_0^1 \langle \nabla_\theta m(\theta,t), u\rangle^2_{\mathbb{R}^d}\, \mathrm{d}t + \mathrm{o}(1) \ ,$$

avec un $\mathrm{o}(1)$ uniforme en u. La liberté de $\left(\frac{\partial m}{\partial \theta_i}(\cdot)\right)_{1\le i\le d}$ implique alors

$$\inf_{\|u\|=1,\, \theta\in\Theta} \int_0^1 \langle \nabla_\theta m(\theta,t), u\rangle^2_{\mathbb{R}^d}\, \mathrm{d}t > 0 \ ,$$

[5] Assumption 1,(i),(ii), [59]

et par suite

$$\sum_{k=1}^{n} \|m(\theta+h,t_k)-m(\theta,t_k)\|^2 \geq \text{cte} \times n \|h\|^2 \ .$$

Si la trajectoire est paramétrique, alors on a clairement $\sqrt{n}(\hat{\theta}_n - \theta^*) = \mathrm{O}_{\mathbb{P}}(1)$ et dans le cas contraire, il existe $\theta' \in \Theta$ tel que $\sqrt{n}(\hat{\theta}_n - \theta') = \mathrm{O}_{\mathbb{P}}(1)$ ce qui achève la vérification des hypothèses du modèle. □

Les hypothèses délicates dans [59] sont les hypothèses 1 et 2 qui concernent le modèle paramétrique. Les autres conditions sont des contraintes techniques mais ne présentent pas de difficultés particulières.

Cette proposition a d'intéressantes conséquences en pratique.

Proposition 2. *Pour un modèle paramétrique de cible MRU et un porteur qui suit un MCU non stationnaire, on peut appliquer le test d'adéquation à une famille paramétrique de [59].*

Démonstration. C'est une conséquence de la proposition 1, de la proposition 5.10 qui garantit l'identifiabilité et de la proposition 5.2 qui fournit l'inversibilité de toute matrice de Fisher sous ces hypothèses. □

Proposition 3. *Pour un modèle paramétrique de cible MCU et un porteur qui suit un MRU non stationnaire, on peut appliquer le test d'adéquation à une famille paramétrique de [59].*

Démonstration. C'est une conséquence de la proposition 1, de la proposition 5.11 qui garantit l'identifiabilité et de la proposition 5.12 qui établit l'inversibilité de toute matrice de Fisher sous ces conditions. □

Toutefois, la trajectographie n'en reste pas moins un problème relativement pathologique comme le souligne la proposition suivante.

Proposition 4. *Il existe un modèle paramétrique $\{S_\theta\}_{\theta\in\Theta}$ pour lequel une trajectoire S_{θ^*} est observable et pourtant, la condition 1 de [59] n'est pas vérifiée et on ne peut donc pas appliquer le test défini par Horowitz et Spokoiny pour tester l'adéquation au modèle paramétrique.*

Démonstration. D'après la proposition 5.9, on sait qu'il existe un modèle paramétrique $\{S_\theta\}_{\theta\in\Theta}$ tel que S_{θ^*} soit observable pour un certain $\theta^* \in \Theta$ et tel qu'il existe une infinité de trajectoires paramétriques non observables. Par conséquent, aucune des deux conditions d'identifiabilité formulées dans [59] n'est satisfaite. En effet, on a

$$\sup_{\theta\in\Theta} \left\| n I_n^{-1}(\theta) \right\| = \infty \ ,$$

et

$$\inf_{\substack{(\theta,\theta')\in\Theta^2 \\ \|\theta-\theta'\|\geq\delta}} |\Phi(X_\theta)(t_{1:n}) - \Phi(X_{\theta'})(t_{1:n})|^2 \geq 0 \ .$$

□

Pour résumer, même s'il existe des situations critiques comme celle décrite ci-avant, on peut, en associant les résultats d'observabilité établis au chapitre 5 avec les travaux décrits dans [59], envisager des scénarios où le porteur peut tester l'adéquation des mesures à un modèle paramétrique. Le problème de la trajectographie n'est pas pour autant devenu un problème bien posé. Pour que le test soit consistant, les hypothèses de [59] imposent que la suite $\{\rho_n\}_{n\geq 0}$ ne tend pas

plus vite vers zéro qu'une certaine vitesse. Autrement, dit une trajectoire qui serait ambiguë avec une trajectoire paramétrique, cas où la suite $\{\rho_n\}_{n\geq 0}$ est nulle, ne sera évidemment pas distinguée de la trajectoire paramétrique. On pourrait d'ailleurs opter pour un formalisme qui présente les résultats pour l'ensemble quotient des trajectoires par la relation d'équivalence "être ambiguë à".

Chapitre 9

Trajectoires paramétriques par morceaux

Sommaire

9.1 Introduction 89
9.2 Consistance 89
9.3 Défaut d'observabilité 91
9.4 Conclusion 92

9.1 Introduction

Les modèles de trajectoires paramétriques par morceaux sont séduisants parce qu'ils permettent de modéliser une très grande variété de trajectoires réalistes. De plus, ils sont adaptés à des trajectoires hors-modèle. Par exemple, toute trajectoire de classe $\mathscr{C}^1$ à vitesse constante peut être approchée de manière satisfaisante par des trajectoires de classe $\mathscr{C}^1$ MRU-MCU par morceaux. Si l'on considère des trajectoires dont le nombre de sections est fixé, on est alors dans un cadre paramétrique avec un paramètre fini-dimensionnel qui détermine la trajectoire. Le problème n'est pas simple pour autant. Si la trajectoire est constituée de sections paramétrique dont les modèles sont distincts, l'estimateur pour le modèle d'observation avec bruit gaussien peut exister dans un espace différent de celui de la vraie trajectoire, auquel cas on ne peut pas les comparer. Si l'on ne suppose plus le nombre de sections connus, alors le problème est non-paramétrique et requiert une approche type sélection de modèle avec une pénalité croissante en le nombre de sections. Cependant, cette approche n'a pas été concrétisée au cours de cette thèse, en particulier à cause du caractère pathologique de la fonction d'observation.

Dans la section 9.2, on montre la consistance de l'estimateur du maximum de vraisemblance puis on présente en section 9.3 le défault d'observabilité qui rend l'estimation des modèles par morceaux pratiquement inutilisables.

9.2 Consistance

On reprend le formalisme introduit au chapitre 6. On rappelle qu'un modèle paramétrique $\mathcal{M}_i$ est la donnée d'un ensemble $\Theta_i \subset \mathbb{R}^{d_i}$ et d'une application $\theta \mapsto S_{\mathcal{M}_i,\theta}$. On suppose donnée une collection de modèles paramétriques $\mathcal{M} = \{\mathcal{M}_i\}_{i\in I}$. On travaille avec le modèle d'observations à temps continu

$$\mathrm{d}Y_t = \Psi[S_\xi(t), t]\mathrm{d}t + \varepsilon \mathrm{d}W_t \ ,$$

où $(W_t)_{t\in[0,1]}$ désigne un processus de Wiener et

$$S_\xi = \sum_{k=1}^{p} S_{\mathcal{M}_{\varphi(k)},\theta_k} \mathbb{1}_{[\tau_{k-1},\tau_k[} \ ,$$

avec $\underline{\tau} = (\tau_0, \cdots, \tau_p) \in \Delta_p$, φ est une application de $\mathbb{N}_p = \{1, \ldots, p\}$ dans I et

$$\underline{\theta} = (\theta^{(1)}, \cdots, \theta^{(p)}) \in \Theta_{\varphi(1)} \times \cdots \times \Theta_{\varphi(p)} \ .$$

On rappelle que, pour tout entier $p \geq 1$,

$$\Delta_p \stackrel{\text{def}}{=} \left\{(t_0, \cdots, t_p) \in [0,1]^{p+1},\ 0 = t_0 < t_1 < \cdots < t_p = 1\right\} \ .$$

Pour un p fixé, on note Ξ_p l'ensemble des éléments caractérisant une trajectoire paramétrique par morceaux. Remarquons qu'il n'y a pas de métrique naturelle sur Ξ_p notamment à cause de $\underline{\theta}$ qui appartient au produit $\Theta_{\varphi(1)} \times \cdots \times \Theta_{\varphi(p)}$. Pour un autre choix de φ, le paramètre $\underline{\theta}$ peut vivre dans un ensemble produit différent. À présent, on souhaite estimer le paramètre $\xi^* \in \Xi_p$ de la trajectoire réelle. Les difficultés sont immédiates. L'estimateur du maximum de vraisemblance pour le modèle statistique $(\mathbb{P}_\xi)_{\xi\in\Xi_p}$ est défini[1] par

$$\tilde{\xi}^\varepsilon = \arg\min_{\xi\in\Xi_p} \int_0^1 \{\Phi(S_\xi)(t) - \Phi(S_{\xi^*})(t)\}^2 \ \mathrm{d}t - 2\varepsilon \int_0^1 \{\Phi(S_\xi)(t) - \Phi(S_{\xi^*})(t)\} \ \mathrm{d}W_t \ .$$

Si la collection contient plusieurs modèles, d'après la remarque précédente, il se peut que $\underline{\theta}(\tilde{\xi}^\varepsilon)$ ne soit pas dans le même espace que $\underline{\theta}^*$. Comme il n'y a pas de métrique raisonnable pour Ξ_p, on ne peut pas parler de consistance pour un tel estimateur. On pourrait envisager de montrer la consistance de la trajectoire de paramètre $\tilde{\xi}^\varepsilon$, c'est-à-dire, pour tout $\eta > 0$,

$$\lim_{\varepsilon\to 0} \mathbb{P}\left(\|S_{\tilde{\xi}^\varepsilon} - S_{\xi^*}\|_{\mathrm{L}^2} > \eta\right) = 0 \ .$$

Il doit alors être possible d'adapter la théorie classique des M-estimateurs avec ce nouveau critère.

Pour la suite, on va se contenter des cas où la collection $\mathcal{M}$ ne contient qu'un modèle ou alors on suppose la fonction φ fixée indépendamment de ξ, autrement dit on impose la séquence des modèle paramétriques de la trajectoire. Pour tout modèle $\mathcal{M}_i$ pris dans la collection $\mathcal{M}$, on suppose que l'application $(t,\theta) \mapsto S_{\mathcal{M}_i,\theta},\ [0,1] \times \Theta_i \to \mathbb{R}^2$ est continue. On note

$$\Lambda_p \stackrel{\text{def}}{=} \left\{\xi \in \Theta \times \bar{\Delta}_p : X_\xi \text{ continue}\right\} \ ,$$

avec $\Theta = {\Theta_i}^p$ si la collection ne contient qu'un modèle $\mathcal{M}_i$ et $\Theta = \prod_{k=1}^p \Theta_{\varphi(k)}$ sinon et $\bar{\Delta}_p$ l'adhérence de Δ_p dans $\mathbb{R}^{p+1}$. Précisons la nature topologique de Λ_p. On suppose pour simplifier que $\mathcal{M}$ ne contient qu'un modèle. L'application

$$\begin{array}{rcl} \Theta^p \times \bar{\Delta}_p & \to & (\mathbb{R}^2)^p \\ (\underline{\theta}, \underline{\tau}) & \mapsto & \left(S_{\mathcal{M},\theta_k}(\tau_k) - S_{\mathcal{M},\theta_{k-1}}(\tau_k)\right)_{1\leq k\leq p} \end{array}$$

est continue. Or, on a

$$\Lambda_p \stackrel{\text{def}}{=} \left\{\xi \in \Theta \times \bar{\Delta}_p : S_{\mathcal{M},\theta_k}(\tau_k) = S_{\mathcal{M},\theta_{k-1}}(\tau_k),\ \forall k \in \mathbb{N}_p\right\} \ ,$$

[1] Voir page 76 pour la définition précise de l'estimateur du maximum de vraisemblance.

donc Λ_p est un fermé dans le compact $\Theta \times \bar{\Delta}_p$, donc c'est un compact. Le modèle statistique considéré est désormais $(\mathbb{P}_\xi)_{\xi\in\Lambda_p}$.

Définition 9.1. *On définit l'estimateur du maximum de vraisemblance pour $(\mathbb{P}_\xi)_{\xi\in\Lambda_p}$ par*

$$\hat{\xi}^\varepsilon \stackrel{\text{def}}{=} \arg\min_{\xi\in\Lambda_p} \int_0^1 \{\Phi(S_\xi)(t) - \Phi(S_{\xi^*})(t)\}^2 \,\mathrm{d}t - 2\varepsilon \int_0^1 \{\Phi(S_\xi)(t) - \Phi(S_{\xi^*})(t)\} \,\mathrm{d}W_t \,.$$

Lemma 9.1. *L'application définie de Λ_p dans $\mathbb{R}^+$ par*

$$\xi \longmapsto \int_0^1 \{\Phi(S_\xi)(t) - \Phi(S_{\xi^*})(t)\}^2 \,\mathrm{d}t \tag{1}$$

est continue.

La preuve est une conséquence de la continuité de l'application $\xi \mapsto S_\xi$ qui se vérifie sans difficulté.

Proposition 9.1. *On suppose que $\mathcal{M}$ ne contient qu'un modèle ou que φ est fixé indépendamment de ξ. Soit une trajectoire S_{ξ^*} avec $\xi^* \in \Lambda_p$. Alors, si la trajectoire est observable dans Λ_p, c'est-à-dire*

$$\arg\min_{\xi\in\Lambda_p} \|\Phi(S_\xi) - \Phi(S_{\xi^*})\|_{\mathrm{L}^2} = \{\xi^*\} \,,$$

l'estimateur $\hat{\xi}^\varepsilon$ est consistant.

Il suffit d'appliquer le théorème de consistance des M-estimateurs [110, Theorem 5.7] en utilisant la continuité de l'application (1) et la compacité de Λ_p.

Sur un plan pratique, les conditions sur le porteur pour garantir l'observabilité d'une cible en MRU par morceaux (proposition 6.5) ou en MCU par morceaux (proposition 6.9) permettent l'observabilité dans Λ_p.

La consistance n'est qu'un pan du problème d'estimation. Il faudrait ensuite s'intéresser à la vitesse de convergence de l'estimateur. Il est probable que l'estimation des modèles paramétriques par morceaux s'inscrivent dans un cadre classique, voir [98], [91], [88]. Cette investigation n'a pas été menée plus loin à cause du comportement pathologique que l'on va décrire dans la section suivante.

9.3 Défaut d'observabilité

En réalité, les modèles de trajectoires paramétriques par morceaux sont très peu adaptés au problème de la trajectographie passive. Non pas qu'ils soient de mauvais candidats pour la description de trajectoires réelles, mais parce que le critère des mesures angulaires ne permet pas quasiment pas la distinction entre certains modèles de trajectoires distincts, et ce malgré l'hypothèse d'observabilité. Ce constat illustre la limite de la notion binaire d'observabilité. En pratique, il est nécessaire, pour que l'algorithme de minimisation atteigne la trajectoire réelle, que le critère à minimiser admette un minimum assez « franc ». On peut fabriquer, dans le cadre idéal des mesures sans bruit d'observation, quantité de scénarios avec des minima locaux très proches du minimum global et atteint en des trajectoires relativement éloignées de la trajectoire réelle.

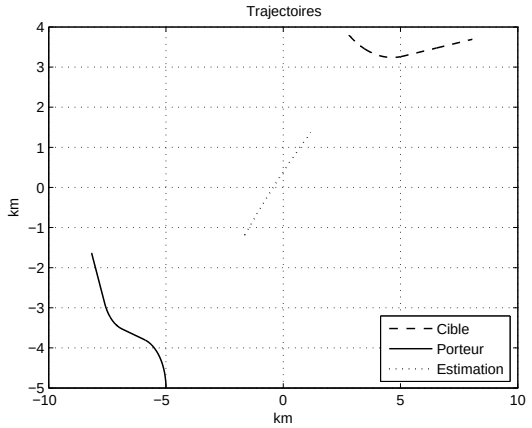


Fig. 9.1 – Estimation d'une trajectoire constituée d'un MCU suivi d'un MRU

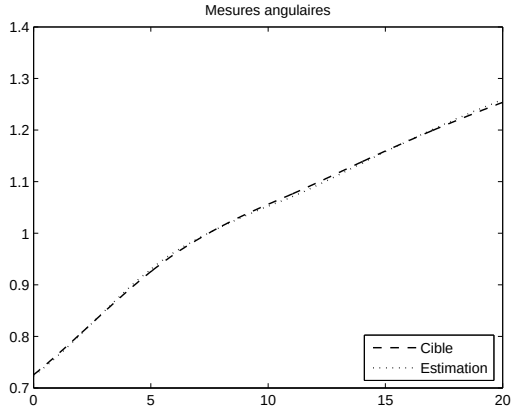


Fig. 9.2 – Écart des mesures angulaires

Sur la figure 9.1, on peut voir l'estimation d'une trajectoire MCU-MRU. Le tracé des mesures angulaires présenté en figure 9.2 montre à quel point les trajectoires sont proches sur le plan des mesures angulaires bien qu'elles soient très éloignées l'une de l'autre. On présente une situation similaire en figures 9.3 et 9.4 où la cible suit une trajectoire MRU-MCU. Les mesures angulaires entre trajectoires réelle et estimée sont quasiment confondues bien que les deux trajectoires soient très distinctes.

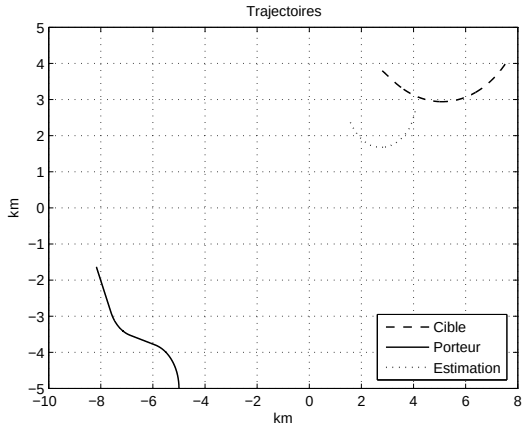


Fig. 9.3 – Estimation d'une trajectoire constituée d'un MCU suivi d'un MRU

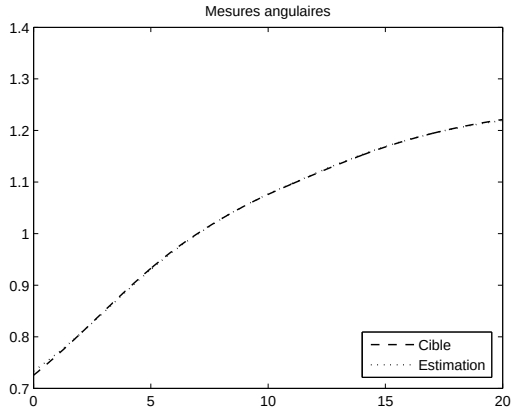


Fig. 9.4 – Écart des mesures angulaires

À cette situation s'ajoute la complexité algorithmiques de la minimisation. Par exemple, pour une trajectoire continues avec trois sections MRU, il faut estimer deux instants de rupture, quatre paramètres pour la première section MRU et deux pour chacune des sections suivantes ce qui signifie qu'il faut minimiser une fonction à dix paramètres. Le logiciel `Matlab`, très performant dans le cadre paramétrique, est d'un usage particulièrement laborieux pour de telles minimisations.

9.4 Conclusion

Le caractère mal posé de la trajectographie passive semble rendre impraticable l'estimation de trajectoires paramétrique par morceaux avec un nombre de sections fixé. Aussi, l'approche de

la sélection de modèles où on pénalise par rapport au nombre de sections n'a pas été étudiée plus avant. À cela s'ajoute un autre problème : l'estimation en trajectographie passive présente des instabilités très fortes : une trajectoire peut être proche en norme euclidienne d'un modèle paramétrique ou paramétriques par morceaux et l'estimation fournit un résultat très loin de la trajectoire réelle.

Chapitre 10

Robustesse en paramétrique

Sommaire

10.1 Introduction **96**
10.2 Sensitivity to deterministic perturbation **98**
10.2.1 Perfect measurements 98
10.2.2 Noisy measurements 100
10.2.3 Continuous time problem 101
10.2.4 Large sample results 103
10.3 Sensitivity to stochastic perturbations **104**
10.3.1 Cyclostationary perturbation 104
10.3.2 Stochastic perturbation on a parametric trajectory 104
10.4 Simulations **105**
10.5 Conclusion **110**
10.6 Appendix **110**
10.7 Compléments **113**
10.8 Corrélations dans le cas sur-critique **115**
10.9 Un estimateur linéarisé **116**

Le chapitre qui suit est, en partie, issu de l'article intitulée *Robustness Considerations for Bearings-only Tracking* publié en tant que *Proceedings of the 11th International Conference on Information Fusion, Cologne, Germany, 2008*.

Robustness Considerations for Bearings-only Tracking

ABSTRACT. This chapter deals with the analysis of robustness for bearings-only tracking. We focus on the case where the trajectory of the target follows some assumed parametric model with an unknown additive perturbation. We investigate the parametric estimator obtained using least squares. We prove that the difference between this estimator and the true parameter may be expanded along directions with intensities that may be very large due to the presence of the unknown perturbation. For a given parametric trajectory, we are able to exhibit the perturbation that leads to the highest intensities. We show through simulations that other methods, like Extended Kalman Filter or Particles Filter, present also the same lack of robustness.
Key words and phrases : Bearings-only tracking, Robustness, Estimation

10.1 Introduction

Bearings-only tracking (BOT) is the determination of the trajectory of a target solely from bearing measurements. This is a highly ill posed problem. At first, it requires the choice of a trajectory model such that, with a perfect observation, one is able to recover the trajectory. The need of ownship manoeuvre and its impact on target state observability have been explored extensively, see for example [11], [64], [65], [95], [107], [76], [63], [44]. Many algorithms have been proposed to track the target : Batch Least Squares (BLS) estimator [11], Extended Kalman Filter (EKF) [11], [101], Unscented Kalman Filter [101], Interacting Multiple Model (IMM) [11], [101], [9], and Particle Filters [101], [43]. All these strategies are designed with respect to particular classes of models. The performance of these algorithms have been well studied when they operate in the model they have been built for.

Even if the literature on BOT is very large, some basic questions still remain unsolved. In practice, the trajectory of the target could be very close to a known model but still not in the model. The question of whether the algorithm will give satisfying results in this context is a robustness question which, to our knowledge, has never been investigated .

In this paper, we focus on the behavior of the BLS estimator for a parametric trajectory model, such as e.g. uniform linear motion, when the real trajectory of the target is not exactly in the parametric model. It is commonly accepted that the BLS estimator presents good performances for the estimation of the parameter. We prove here that the BLS estimator can be highly sensitive to small perturbations and show, using simulations that other methods such as EKF or PF do not perform better than the BLS estimator.

Let us now introduce some notation and definitions. Let $O(t) = (O_1(t), O_2(t))$ be the position of the observer at time t, which is assumed to be known at the observation times $\{t_1, \ldots, t_n\}$. For all integer $k \in \{1, \ldots, n\}$ and for all $x = (x_1, x_2) \in \mathbb{R}^2$ such that $x_2 - O_2(t_k) > 0$, we define the angle measure

$$\Psi(x, t_k) = \arctan[x_1 - O_1(t_k)]/[x_2 - O_2(t_k)] \ .$$

Let $\theta \in \Theta$ be a finite dimensional parameter, where Θ stands for a compact subset of $\mathbb{R}^d$. We denote by S_θ a parametric trajectory which is completely described by the parameter θ. For example, if we assume a uniform linear motion, then we set

$$S_\theta(t) = (S_{\theta,1}(t), S_{\theta,2}(t)) = (\theta_1 + \theta_3 t,\ \theta_2 + \theta_4 t) \ .$$

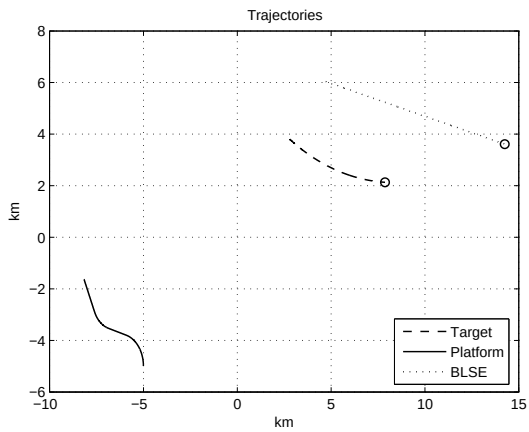


FIG. 10.1 – Sensitivity of BLS estimation

For a given parametric trajectory model satisfying

(H_+) $\quad S_{\theta,2}(t) - O_2(t) > 0 \quad$ for all t in [0,1] and for all θ in Θ ,

we define the associated noiseless measure by

$$m(\theta, t) = \Psi[S_\theta(t), t] \ .$$

If $X = \{X_1, \ldots, X_n\}$ is a sequence of positions of the target, the bearing least squares criterion is

$$C_n(\theta, X) = \sum_{k=1}^{n} [m(\theta, t_k) - \Psi(X_k, t_k)]^2 \ . \tag{1}$$

As Θ is a compact set, we know that there exists a value in Θ which minimizes the function $\theta \mapsto C_n(\theta, X)$. Let us define $\tilde{\theta}_n(X)$ as a minimizer of the criterion :

$$\tilde{\theta}_n(X) = \arg\min_{\theta \in \Theta} C_n(\theta, X) \ .$$

The parametric trajectory S_{θ^*} is said to be *observable* if the sequence $X = \{S_{\theta^*}(t_1), \ldots, S_{\theta^*}(t_n)\}$ uniquely determines the parameter θ^*. It then follows that the minimizer $\tilde{\theta}_n(X)$ is uniquely defined by $\tilde{\theta}_n(X) = \theta^*$.

A natural question is the following : what happens when we consider a general trajectory X close to an observable parametric trajectory ? Let us consider a case with a target following a uniform circular motion with speed of $270\,\mathrm{m/s}$ and with low curvature. The platform has a constant speed of $250\,\mathrm{m/s}$ and makes maneuvers with an acceleration of around $50\,\mathrm{m/s^2}$. The trajectory of the target is close in the Euclidian norm to a uniform linear trajectory with a maximum relative distance of $260\,\mathrm{m}$. Nevertheless, the simulations show that for such a situation, minimisation algorithms for (1) can give very bad results (see Figure 10.1). The real position and the estimate position at final time are flagged by a circle.

The paper is organized as follows. In section 10.2, we study the behavior of local minimizers of (1) when the real trajectory X is close to a parametric trajectory S_{θ^*}, see Theorem 10.1. This will allow to compute highly non robust situations. With noisy measurements, we prove that the BLS estimator is approximately Gaussian with mean $\tilde{\theta}_n(X)$, see Proposition 10.4. We prove similar results for continuous time observations, see Theorem 10.2, and for noisy sampled observations, see Theorem 10.3. In Section 10.3, we study the state space models, which may be interpreted

as adding a stochastic perturbation to a parametric trajectory. In such cases, the BLS estimator is asymptotically unbiased, see Propositions 10.7 and 10.12. In section 10.4, we present realistic simulations to exhibit various non-robust situations. We show that using EKF and PF does not help in such cases.

10.2 Sensitivity to deterministic perturbation

10.2.1 Perfect measurements

In the sequel, we shall denote the k-dim Euclidian space by $(\mathbb{R}^k, \langle\ ,\ \rangle_{\mathbb{R}^k})$. We consider the space of sequence of trajectories $(\mathbb{R}^2)^n$ endowed with the following norm :

$$\|X\|_\infty \stackrel{\text{def}}{=} \sup_{1\le k\le n} \|X_k\|_{\mathbb{R}^2} \ .$$

Let $I_n(\theta)$ be the matrix defined by

$$I_n(\theta) \stackrel{\text{def}}{=} \sum_{k=1}^{n} \nabla_\theta m(\theta, t_k) \nabla_\theta m(\theta, t_k)^{\mathrm{T}} \ ,$$

which is the usual Fisher information matrix for the classical parametric regression model

$$Y_k = m(\theta, t_k) + \varepsilon_k \ ,$$

where $\{\varepsilon_k\}_{k\ge 0}$ is a sequence of i.i.d standard Gaussian variables.

Let us define $C_n : \Theta \times (\mathbb{R}^2)^n \to \mathbb{R}$ by

$$C_n(\theta, X) \stackrel{\text{def}}{=} \sum_{k=1}^{n} [m(\theta, t_k) - \Psi(X_k, t_k)]^2 \ .$$

In the sequel, $\mathrm{D}_\theta\, C_n$ will denote the first differential of C_n in θ, $\mathrm{D}_{\theta,\theta}\, C_n$ the second differential of C_n in θ, $\mathrm{D}_{X,\theta}\, C_n$ the second differential of C_n in θ and X and $L(A, B)$ the set of continuous linear applications from A to B. It follows by direct calculations that

$$\mathrm{D}_{\theta,\theta}\, C_n(\theta^*, S_{\theta^*}) = 2 I_n(\theta^*) \ .$$

Recall that $(\mathbb{R}^d, \langle\ ,\ \rangle_{\mathbb{R}^d})$ and $((\mathbb{R}^2)^n, \|\ \|_\infty)$ are Banach spaces. Then, for $\theta^* \in \Theta$ such that $I_n(\theta^*)$ is non-singular, applying the implicit function theorem, there exist open subsets $V_{\theta^*} \subset \mathbb{R}^d$, $V_{S_{\theta^*}} \subset (\mathbb{R}^2)^n$ with $\theta^* \in V_{\theta^*}$, $S_{\theta^*} \in V_{S_{\theta^*}}$, and a continuously differentiable application $\bar\theta_n : V_{S_{\theta^*}} \to V_{\theta^*}$ such that, for all $(\theta, X) \in V_{\theta^*} \times V_{S_{\theta^*}}$,

$$\mathrm{D}_\theta\, C_n(\theta, X) = 0 \iff \theta = \bar\theta_n(X) \ .$$

Moreover, the differential $D_X \bar\theta_n(X) \in L((\mathbb{R}^2)^n, \mathbb{R}^d)$ is given by

$$\mathrm{D}_X\, \bar\theta_n(X)(H) = -\,\mathrm{D}_{\theta,\theta}\, C_n \left[\bar\theta_n(X), X\right]^{-1} \mathrm{D}_{X,\theta}\, C_n \left[\bar\theta_n(X), X\right](H) \ .$$

The need for non-singular Fisher information matrix is not surprising since we know [63] that it guarantees local observability for the parameter.

Proposition 10.1. *We can assume, by choosing a subset of $V_{S_{\theta^*}}$ that this set is connected. Then, for all $X \in V_{S_{\theta^*}}$*

$$\bar{\theta}_n(X) = \arg\min_{\theta \in V_{S_{\theta^*}}} C_n(\theta, X) \ .$$

In practice, a minimizing algorithm can reach other local minima than $\bar{\theta}_n(X)$. Also, the global $\arg\min$ can be different from $\bar{\theta}_n(X)$. But as we focus on robustness, the non robust behavior of $\bar{\theta}_n(X)$ implies an even worse behavior for $\tilde{\theta}_n(X)$. Anyway, for small perturbations, we observe in simulations that $\bar{\theta}_n(X) = \tilde{\theta}_n(X)$ almost ever.

It follows from the Taylor-Young formula that

$$\bar{\theta}_n(S_{\theta^*} + H) = \theta^* + \mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H) + \mathrm{o}(\|H\|_\infty) \ .$$

Our goal is now to understand the role played by the differential term $\mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H)$.

Proposition 10.2. *The differential $\mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H)$ is given by*

$$\mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H) = I_n^{-1}(\theta^*) \sum_{k=1}^{n} \nabla_\theta m(\theta^*, t_k) \langle \nabla_x \Psi[S_{\theta^*}(t_k), t_k], H_k \rangle_{\mathbb{R}^2}. \tag{2}$$

Since $I_n(\theta^*)$ is symmetric, there exists an orthonormal basis $\{e_n^{(1)}(\theta^*), \ldots, e_n^{(d)}(\theta^*)\}$ of eigenvectors with eigenvalues $\{\lambda_n^{(1)}(\theta^*), \ldots, \lambda_n^{(d)}(\theta^*)\}$. Let us define the operator $K_{n,\theta^*} \in L(\mathbb{R}^n)$ by

$$K_{n,\theta^*}(v) \stackrel{\text{def}}{=} \sum_{i=1}^{d} \langle \alpha_n^{(i)}(\theta^*), v \rangle_{\mathbb{R}^n}\, \alpha_n^{(i)}(\theta^*) \ ,$$

where the $\alpha_n^{(i)}(\theta^*) \in \mathbb{R}^n$ are defined by

$$\alpha_{n,k}^{(i)}(\theta^*) = \lambda_n^{(i)}(\theta^*)^{-1} \langle \nabla_\theta m(\theta^*, t_k), e_n^{(i)}(\theta^*) \rangle_{\mathbb{R}^d} \|\nabla_x \Psi[S_{\theta^*}(t_k), t_k]\|_{\mathbb{R}^2}.$$

Proposition 10.3. *The operator K_{n,θ^*} is a positive self-adjoint compact operator with d-dimensional range.*

Let us denote by $\{f_n^{(1)}(\theta^*), \ldots, f_n^{(d)}(\theta^*)\}$ the eigenvectors of the restriction of K_{n,θ^*} to its range $\mathcal{R}(K_{n,\theta^*})$ and let $U_\theta(t)$ be the normalized vector

$$U_\theta(t) = \|\nabla_x \Psi[S_\theta(t), t]\|^{-1}\, \nabla_x \Psi[S_\theta(t), t] \ .$$

Let us define for all $i \in \{1, \ldots, d\}$

$$D_n^{(i)}(\theta^*) \stackrel{\text{def}}{=} \sum_{j=1}^{d} \langle \alpha_n^{(j)}(\theta^*), f_n^{(i)}(\theta^*) \rangle_{\mathbb{R}^n}\, e_n^{(j)}(\theta^*) \ .$$

Notice that the vectors $\{D_n^{(1)}(\theta^*), \ldots, D_n^{(d)}(\theta^*)\}$ are orthogonal and that $\|D_n^{(i)}(\theta^*)\|_{\mathbb{R}^d} = \mu_n^{(i)}(\theta^*)$, $i = 1, \ldots, d$ are the non null eigenvalues of the operator K_{n,θ^*}.

Theorem 10.1. *We have*

$$\bar{\theta}_n(S_{\theta^*} + H) = \theta^* + \sum_{i=1}^{d} h_i D_n^{(i)}(\theta^*) + \mathrm{o}(\|H\|_\infty) \ ,$$

where the $h_i \in \mathbb{R}$ are defined by

$$h_i = \sum_{k=1}^{n} \langle H_k, U_{\theta^*}(t_k)\rangle_{\mathbb{R}^2} \, f_{n,k}^{(i)}(\theta^*) \ . \tag{3}$$

Writing $h = (h_1, \ldots, h_d)$, we also have $\|h\|_{\mathbb{R}^d} \leq \|\bar{H}\|_{\mathbb{R}^n}$ with $\bar{H} = (\|H_1\|_{\mathbb{R}^2}, \ldots, \|H_n\|_{\mathbb{R}^2})$.

The proofs of Propositions 10.1, 10.2, 10.3 and Theorem 10.1 are given in the Appendix.

The Gram-Schmidt algorithm allows us to compute the eigenvectors and eigenvalues required in the theorem. Thus, for a given parameter θ^* and an intensity perturbation $\|H\|_\infty$, it is possible to compute the perturbation H that leads to the worst BLS estimator $\bar{\theta}_n(S_{\theta^*} + H)$. Various realistic simulations are given in section 10.4.

10.2.2 Noisy measurements

We now consider the case where the platform receives noisy measurements. The sequence of observations is

$$Y_k = \Psi(X_k, t_k) + \sigma\varepsilon_k \ ,$$

where $\{\varepsilon_k\}_{k\geq 0}$ is a sequence of i.i.d standard Gaussian variables. Let us define the adjustment criterion by

$$C_n^\sigma(\theta, X) \stackrel{\text{def}}{=} \sum_{k=1}^{n} [m(\theta, t_k) - Y_k]^2 \ .$$

We assume that $I_n(\theta^*)$ is non-singular, so there exist open sets V_{θ^*}, $V_{S_{\theta^*}}$ and an application $\bar{\theta}_n$ as before. Fix some $X \in V_{S_{\theta^*}}$. It is clear that

$$\mathrm{D}_{\theta,\theta}\, C_n^\sigma(\theta, X) \to \mathrm{D}_{\theta,\theta}\, C_n(\theta, X) \quad \text{a.s.}$$

as $\sigma \to 0$. Then, applying the implicit function theorem, almost surely there exist $\delta > 0$ and an open set $V'_{\theta^*} \subset V_{\theta^*}$, and a continuously differentiable application $(-\delta, \delta) \to V_{\theta^*},\ \sigma \mapsto \bar{\theta}_\sigma(X)$ such that for all $|\sigma| < \delta$, $\theta \in V'_{\theta^*}$

$$\mathrm{D}_\theta\, C_n^\sigma(\theta, X) = 0 \iff \theta = \bar{\theta}_n^\sigma(X) \ .$$

The differential $D_\sigma \bar{\theta}_n^\sigma(X)$ is given by

$$\mathrm{D}_\sigma\, \bar{\theta}_n^\sigma(X) = -\,\mathrm{D}_{\theta,\theta}\, C_n^\sigma \left[\bar{\theta}_n^\sigma(X), X\right]^{-1} \mathrm{D}_{\sigma,\theta}\, C_n^\sigma \left[\bar{\theta}_n^\sigma(X), X\right] \ .$$

We can now describe the asymptotic behavior of $\bar{\theta}_n^\sigma(X)$.

Proposition 10.4. *Let θ^* be a parameter in Θ such that $I_n(\theta^*)$ is non-singular and let X be a trajectory in $V_{S_{\theta^*}}$. Then, almost surely, for σ small enough,*

$$\bar{\theta}_n^\sigma(X) - \bar{\theta}_n(X) = \sigma\, \mathrm{D}_{\theta,\theta}\, C_n \left[\bar{\theta}_n(X), X\right]^{-1} \sum_{k=1}^{n} \nabla_\theta m \left[\bar{\theta}_n(X), t_k\right] \varepsilon_k + \mathrm{o}(\sigma) \ .$$

Thus, for σ small enough, $\bar{\theta}_n^\sigma(X) - \bar{\theta}_n(X)$ is approximately distributed as $\mathcal{N}\left(0, \sigma^2 I_n(\bar{\theta}_n(X))^{-1}\right)$.

We can also describe the simultaneous asymptotic development in σ and H by considering the map $(\sigma, X) \mapsto \mathrm{D}_\theta\, C_n^\sigma(\theta, X)$. Indeed, we have $\mathrm{D}_{\theta,\theta}\, C_n^0(\theta^*, S_{\theta^*}) = I_n(\theta^*)$ and then, we apply the implicit function theorem at the point $(0, S_{\theta^*})$.

Proposition 10.5.

$$\bar{\theta}_n(X) - \theta^* = \mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H) + \sigma I_n^{-1}(\theta^*) \sum_{k=1}^{n} \nabla_\theta m(\theta^*, t_k)\varepsilon_k + \mathrm{o}(\sigma, \|H\|_\infty) \quad \text{a.s.}$$

The proofs of Proposition 10.4 and 10.5 are direct consequences of the implicit function theorem.

10.2.3 Continuous time problem

We shall prove similar results for continuous time observations. Let us define the space of trajectories $E = \mathscr{C}([0,1], \mathbb{R}^2)$ endowed with the uniform norm $\|f\| = \sup_{t\in[0,1]} \|f(t)\|_{\mathbb{R}^2}$. Consider, for $X \in E$, the stochastic differential equation

$$\mathrm{d}Y_t = \Psi[X(t), t]\mathrm{d}t + \varepsilon \mathrm{d}W_t \ ,$$

where $(W_t)_{t\in[0,1]}$ is a Wiener process. Let us denote by P_X the probability associated to this process. By Girsanov's theorem [61], the log likelihood ratio $\log \frac{\mathrm{d}\mathbb{P}_{S_\theta}}{\mathrm{d}\mathbb{P}_X}$ is proportional to

$$C^\varepsilon(\theta, X) \stackrel{\text{def}}{=} \int_0^1 \{m(\theta, t) - \Psi[X(t), t]\}^2 \ \mathrm{d}t - 2\varepsilon \int_0^1 \{m(\theta, t) - \Psi[X(t), t]\} \, \mathrm{d}W_t \ .$$

Let us define

$$C(\theta, X) \stackrel{\text{def}}{=} C^0(\theta, X) = \int_0^1 \{m(\theta, t) - \Psi[X(t), t]\}^2 \ \mathrm{d}t \ ,$$

and the Fisher information matrix $I(\theta)$ by

$$I(\theta) \stackrel{\text{def}}{=} \int_0^1 \nabla_\theta m(\theta, t) \nabla_\theta m(\theta, t)^{\mathrm{T}} \, \mathrm{d}t \ .$$

By direct calculations, it is clear that

$$\mathrm{D}_{\theta,\theta}\, C(\theta, S_\theta) = 2I(\theta) \ .$$

Recall that $(E, \| \ \|_\infty)$ is a Banach space. Then, for $\theta^* \in \Theta$ such that $I(\theta^*)$ is non-singular, applying the implicit function theorem, there exist open subsets $V_{\theta^*} \subset \mathbb{R}^d$, $V_{S_{\theta^*}} \subset E$ with $\theta^* \in V_{\theta^*}$, $S_{\theta^*} \in V_{S_{\theta^*}}$, and a continuously differentiable application $\bar{\theta} : V_{S_{\theta^*}} \to V_{\theta^*}$ such that, for all $(\theta, X) \in V_{\theta^*} \times V_{S_{\theta^*}}$

$$\mathrm{D}_\theta\, C(\theta, X) = 0 \iff \theta = \bar{\theta}(X) \ .$$

Recall that the differential $\mathrm{D}_X\, \bar{\theta}(S_{\theta^*})$ is a continuous linear operator from E to $\mathbb{R}^d$. We have :

Proposition 10.6.

$$\mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H) = I^{-1}(\theta^*) \int_0^1 \nabla_\theta m(\theta^*, t) \langle \nabla_x \Psi[S_{\theta^*}(t), t], H(t)\rangle_{\mathbb{R}^2} \, \mathrm{d}t \ . \tag{4}$$

The integral form of the expression given by (4) allows us to understand the robustness for a periodic function while the period increases.

Proposition 10.7. *Let H_T be a periodic perturbation with period T. Then*

$$\lim_{T\to 0} \mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H_T) = 0 \ .$$

This proposition illustrates the fact that a high cyclic perturbation does not affect the parametric estimation. A simulation with such a perturbation is given in Section 10.4.

Proof of Proposition 10.7. Let us write $H_T(t) = g(t/T)$ where g denotes a periodic function with period 1. Then the result follows by integration by parts. □

Recall that $\mathrm{L}^2([0,1],\mathbb{R}^k)$ denotes the set of functions f such that $\int_0^1 \|f(t)\|^2_{\mathbb{R}^k}\,\mathrm{d}t < \infty$ and $\mathrm{L}^2([0,1])$ denotes the set $\mathrm{L}^2([0,1],\mathbb{R})$. Let us define the operator $K_{\theta^*} : \mathrm{L}^2([0,1]) \to \mathrm{L}^2([0,1])$ by

$$K_{\theta^*}(v) \stackrel{\text{def}}{=} \sum_{i=1}^{d} \langle \alpha_i(\theta^*), v\rangle_{\mathrm{L}^2}\, \alpha_i(\theta^*) \ , \tag{5}$$

where the $\alpha_i(\theta^*) \in \mathrm{L}^2([0,1])$ are defined by

$$\alpha_i(\theta^*)(t) = \lambda^{(i)}(\theta^*)^{-1} \langle \nabla_\theta m(\theta^*,t), e^{(i)}(\theta^*)\rangle_{\mathbb{R}^d} \|\nabla_x \Psi[S_{\theta^*}(t),t]\|_{\mathbb{R}^2} \ .$$

Proposition 10.8. *The operator K_{θ^*} is a positive self-adjoint compact operator with d-dimensional range.*

Let us define for all $i \in \{1,\dots,d\}$

$$D^{(i)}(\theta^*) \stackrel{\text{def}}{=} \sum_{j=1}^{d} \langle \alpha^{(j)}(\theta^*), f^{(i)}(\theta^*)\rangle_{\mathrm{L}^2}\, e^{(j)}(\theta^*) \ .$$

Notice that $\{D^{(1)}(\theta^*),\dots,D^{(d)}(\theta^*)\}$ are orthogonal and that $\|D^{(i)}(\theta^*)\|_{\mathbb{R}^d} = \mu^{(i)}(\theta^*)$, $i = 1,\dots,d$ are the non null eigenvalues of the operator K_{θ^*}.

Theorem 10.2. *We have*

$$\bar{\theta}(S_{\theta^*} + H) = \theta^* + \sum_{i=1}^{d} h_i D^{(i)}(\theta^*) + \mathrm{o}(\|H\|_\infty) \ ,$$

where the $h_i \in \mathbb{R}$ are defined by

$$h_i = \Big\langle \langle H, U_{\theta^*}\rangle_{\mathbb{R}^2}, f^{(i)}(\theta^*)\Big\rangle_{\mathrm{L}^2} \ .$$

Writing $h = (h_1,\dots,h_d)$, we also have $\|h\|_{\mathbb{R}^d} \le \|H\|_{\mathrm{L}^2}$.

In the same way as in Section 10.2.2, we have

Proposition 10.9.

$$\bar{\theta}^\varepsilon(X) - \bar{\theta}(X) = \varepsilon\, \mathrm{D}_{\theta,\theta}\, C\left[\bar{\theta}(X), X\right]^{-1} \int_0^1 \nabla_\theta m\left[\bar{\theta}(X), t\right] \mathrm{d}W_t + \mathrm{o}(\varepsilon) \quad \text{a.s.}$$

Thus, for ε small enough, $\bar{\theta}^\varepsilon(X) - \bar{\theta}(X)$ is approximately distributed as $\mathcal{N}\left(0, \varepsilon^2 I^{-1}(\bar{\theta}(X))\right)$. **Remark :** Notice that for $X = S_{\theta^*}$, we retrieve the well-known development for classical parametric

regression

$$\bar{\theta}^{\varepsilon}(X) - \theta^* = \varepsilon I^{-1}(\theta^*) \int_0^1 \nabla_\theta m(\theta^*, t)\, \mathrm{d}W_t + \mathrm{o}(\varepsilon) \quad \text{a.s.}$$

By applying the implicit function theorem to the map $(\varepsilon, X) \mapsto C^{\varepsilon}(\theta, X)$ at the point $(0, S_\theta^*)$, we obtain

Proposition 10.10.

$$\bar{\theta}^{\varepsilon}(X) - \theta^* = \mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H) + \varepsilon I^{-1}(\theta^*) \int_0^1 \nabla_\theta m(\theta^*, t)\, \mathrm{d}W_t + \mathrm{o}(\varepsilon, \|H\|_\infty) \quad \text{a.s.}$$

The proofs of Propositions 10.6, 10.8, 10.9, 10.10 and Theorem 10.2 are the same as in Sections 10.2.1 and 10.2.2.

10.2.4 Large sample results

We present here a slightly different approach than in Section 10.2.2. We now consider an increasing sample size. The sequence of observations is

$$Y_k = \Psi[X(t_k), t_k] + \sigma \varepsilon_k \ ,$$

where $t_k = k/n$ and $\{\varepsilon_k\}_{k \geq 0}$ is a sequence of i.i.d standard Gaussian variables. Let us define the bearing adjustment criterion by

$$J_n(\theta, X) \stackrel{\text{def}}{=} n^{-1} \sum_{k=1}^{n} [m(\theta, t_k) - Y_k]^2 \ ,$$

and the minimizer in θ by

$$\hat{\theta}_n(X) \stackrel{\text{def}}{=} \arg\min_{\theta \in \Theta} J_n(\theta, X) \ .$$

As in section 10.2.3, we write

$$C(\theta, X) = \int_0^1 \{m(\theta, t) - \Psi[X(t), t]\}^2 \, \mathrm{d}t \ ,$$

and

$$\bar{\theta}(X) = \arg\min_{\theta \in \Theta} C(\theta, X) \ .$$

We have the following uniform convergence

$$\sup_{\theta \in \Theta} |J_n(\theta, X) - \sigma^2 - C(\theta, X)| = \mathrm{o}_{\mathbb{P}}(1) \ ,$$

and then apply the theory of M-estimators (see [109]) :

Theorem 10.3. *We have*

$$\sqrt{n}\big[\hat{\theta}_n(X) - \bar{\theta}(X)\big] = I_n^{-1}\left[\bar{\theta}(X)\right] n^{-1/2} \sum_{k=1}^{n} \varepsilon_k \nabla_\theta m\left[\bar{\theta}(X), t_k\right] + \mathrm{o}_{\mathbb{P}}(1) \ ,$$

so that, as $n \to \infty$,

$$\sqrt{n}\big[\hat{\theta}_n(X) - \bar{\theta}(X)\big] \rightsquigarrow \mathcal{N}\left(0, I^{-1}\left[\bar{\theta}(X)\right]\right) \ .$$

10.3 Sensitivity to stochastic perturbations

10.3.1 Cyclostationary perturbation

To extend the result of the Proposition 10.7, we can consider the case of cyclostationary perturbation. Recall that a $\mathbb{R}^k$-valued process $(Z_t)_{t\in I}$ is said to be cyclostationary or periodically correlated with period $T > 0$ if T is the smallest real such that for all $t, t' \in I$, $s \mapsto \mathbb{E}(Z_{t+s})$ and $s \mapsto \mathbb{E}(Z_{t+s} Z_{t'+s}^{\mathrm{T}})$ are periodic functions with period T.

Proposition 10.11. *We consider the continuous time model described in section 10.2.3. Let H_T be a continuous cyclostationary process with period T. Then*

$$\lim_{T\to 0} \mathbb{V}\left[\mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H_T)\right] = 0 \ .$$

Proof of Proposition 10.11. For convenience, we write $J(t) = \nabla_\theta m(\theta^*, t)\nabla_x \Psi[S_{\theta^*}(t), t]$. By Fubini's theorem, we have

$$\mathbb{V}\left(\int_0^1 \langle J(t), H(t)\rangle_{\mathbb{R}^2}\, \mathrm{d}t\right) = \mathbb{E}\left\{\int_0^1 \langle J(t), H(t) - \mathbb{E}[H(t)]\rangle_{\mathbb{R}^2}\ \mathrm{d}t\right\}^2 \ .$$

Applying Jensen's inequality and once again Fubini's theorem, we obtain

$$\mathbb{E}\left\{\int_0^1 \langle J(t), H(t) - \mathbb{E}[H(t)]\rangle_{\mathbb{R}^2}\, \mathrm{d}t\right\}^2 \leq \int_0^1 \|J(t)\|_{\mathbb{R}^2}^2\, \mathbb{E}\|H(t) - \mathbb{E}[H(t)]\|_{\mathbb{R}^2}^2\, \mathrm{d}t \ .$$

The application $t \mapsto \mathbb{E}\|H(t) - \mathbb{E}[H(t)]\|_{\mathbb{R}^d}^2$ is T-periodic and we conclude like in Proposition 10.7. □

10.3.2 Stochastic perturbation on a parametric trajectory

In this section, we consider the following model

$$\begin{cases} X_k = S_{\theta^*}(t_k) + b(h)\zeta_k \ , \\ Y_k = \Psi(X_k, t_k) + \sigma\varepsilon_k \ , \end{cases}$$

where $\{\zeta_k\}_{k\geq 0}$ are random variables in $\mathbb{R}^2$ independent of $\{\varepsilon_k\}_{k\geq 0}$ and b is a differentiable real-valued function with $b(0) = 0$. Let us define

$$M_n^\sigma(\theta, h) = \sum_{k=1}^n [m(\theta, t_k) - Y_k]^2 \ .$$

It follows from direct calculations that $\mathrm{D}_{\theta,\theta}\, M_n^0(\theta^*, 0) = I_n(\theta^*)$. Then, by the implicit function theorem, almost surely there exist $\eta > 0,\ \delta > 0$ and V_{θ^*} containing θ^* such that

$$\mathrm{D}_\theta\, M_n^\sigma(\theta) = 0 \iff \theta = \bar{\theta}_n^\sigma(h) \ .$$

By the same way arguments as in Section 10.2.2, we have

Proposition 10.12.

$$\bar{\theta}_n^\sigma(h) - \theta^* = \sigma I_n^{-1}(\theta^*) \sum_{k=1}^n \nabla_\theta m(\theta^*, t_k)\varepsilon_k$$
$$+ b'(0) h I_n^{-1}(\theta^*) \sum_{k=1}^n \nabla_\theta m(\theta^*, t_k) \langle \nabla_x \Psi[X_{\theta^*}(t_k), t_k], \zeta_k \rangle_{\mathbb{R}^2} + \mathrm{o}(\sigma, h) \quad \text{a.s.}$$

Hence, if $\{\zeta_k\}_{k\geq 0}$ are independent centered Gaussian variable, for h and σ small enough, $\bar{\theta}_n^\sigma(h) - \theta^*$ is approximately distributed as a centered Gaussian variable. This explains that sequences of independent Gaussian additive perturbations do not affect the robustness for BOT.

We could also fix $h > 0,\ \sigma > 0$ and study the asymptotic behavior of $\bar{\theta}_n^\sigma(h)$ as $n \to \infty$. If we assume that the random variables $\{\zeta_k\}_{k\geq 0}$ are identically distributed with a density f and satisfy some mixing property, then we have the convergence of $n^{-1} M_n^\sigma(\theta, h)$ to a function of θ and f. This approach is different from the previous one because there is no asymptotic on the level of perturbations. This question is studied in Chapter 11, see page 119.

10.4 Simulations

For all the simulations, the observation time is $20\,\mathrm{s}$ with video frequency. The platform has constant speed of $250\,\mathrm{m/s}$ and makes maneuvers with accelerations of approximately $50\,\mathrm{m/s^2}$. The assumed parametric model is the uniform linear motion. We point out that the problem of robustness occurs for other parametric models such as polynomial or uniform circular. The measure is corrupted by a Gaussian noise with level $\sigma = 5 \times 10^{-3}\,\mathrm{rad}$. To compare our results to those obtained by different methods, we consider two cases : in the first one, the target has a uniform circular motion with speed of $270\,\mathrm{m/s}$ and acceleration of $7\,\mathrm{m/s^2}$ (see Figures 10.2, 10.5, 10.6) ; in the second one, the target follows a uniform linear motion with a worst additive perturbation for a given intensity $\|H\|_\infty = 15\,\mathrm{m}$ (see Figures 10.3, 10.7, 10.8). We also present a case of a periodic perturbation which leads to good results for BLS estimation (see Figure 10.9).

The EKF algorithm designed for uniform linear motion has been implemented in modified polar coordinates [11]. The PF algorithm we use here is the SIR particle filter [101] designed for data issued from equation

$$\begin{cases} A_k = F A_{k-1} + \beta G V_k \ , \\ Y_k = h(A_k, t_k) + \sigma \varepsilon_k \ , \end{cases} \tag{6}$$

where the state vector is $A_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, $\{V_k\}_{k\geq 0}$ is a sequence of independent standard Gaussian vectors in $\mathbb{R}^2$, $\{\varepsilon_k\}_{k\geq 0}$ are independent standard real Gaussian variables independent of the sequence $\{V_k\}_{k\geq 0}$, $F = \begin{pmatrix} \Delta_T I_2 & I_2 \\ 0 & \Delta_T I_2 \end{pmatrix}$, where I_2 is the 2×2 matrix identity and $\Delta_T > 0$ is the sampling step, $G = \begin{pmatrix} \frac{\Delta_T^2}{2} I_2 \\ \Delta_T I_2 \end{pmatrix}$, and $h(A_k, t_k) = \Psi[(x_k, y_k), t_k]$. The initialisation for PF algorithm is uniform along the direction of first observation, which means that we have no *a priori* on initial distance.

The accelerations made by the platform are high enough to provide a good level of information for estimation or filtering, as it can be seen through confidence area or Cramer-Rao bound [101].

We choose θ^* as the parameter minimizing the euclidian norm for the distance between the target and a uniform linear motion. For the first simulation (case of uniform circular motion, see figure 10.2), we present the trajectory described by the BLS estimator (legend BLSE) and the trajectory corresponding to first order development with parameter $\theta^* + \mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H)$ (legend Taylor). We see that for such trajectory close to a uniform linear one, the first order development is a good approximation of the BLS estimator. It also illustrates that the term $\mathrm{D}_X\, \bar{\theta}(S_{\theta^*})(H)$ can be large for small perturbation on a linear motion.

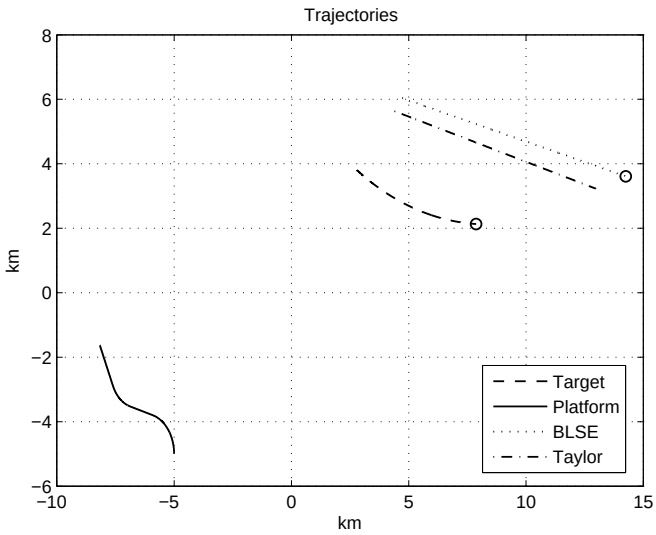


Fig. 10.2 – First order development

In the next simulation (see Figure 10.3), we present a case of the worst perturbation for a given intensity (much smaller than for simulations with circular motion). We see that, although the target is very close to a uniform linear motion, the BLS estimation is bad. For such pathologic cases, it is clear that the BLS estimation is a non-robust method and may lead to misleading conclusions. The non conformity of target can be more easily seen with trajectory speed (see Figure 10.4). These speed fluctuations are not large, but are sufficient to produce a very bad BLS estimation.

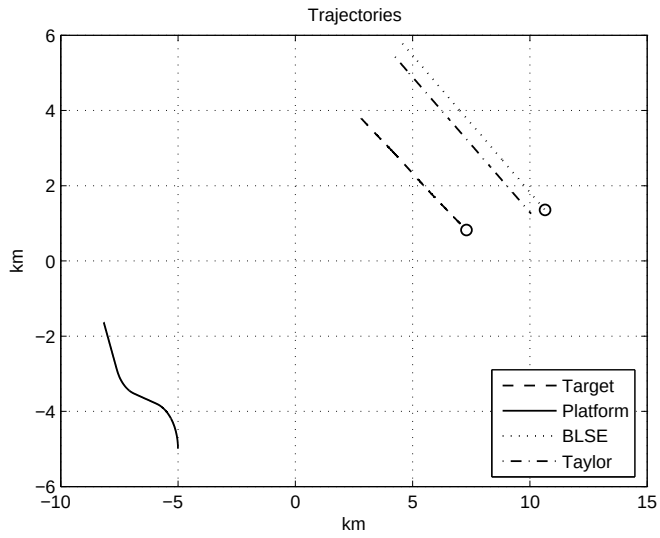


Fig. 10.3 – Worst case of estimation

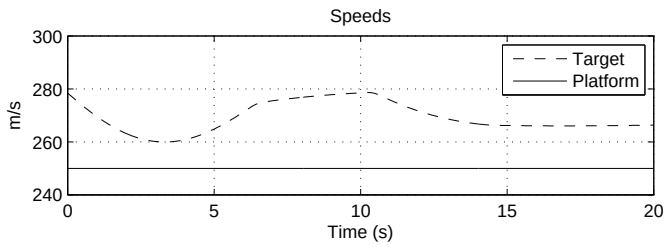


Fig. 10.4 – Speeds of objects in a worst case

A natural question is the following : What will produce other tracking algorithms for such cases ?

When the target has the uniform linear motion described above, the EKF algorithm presents first an instable behavior (see Figure 10.5) and then converges to the BLS estimation.

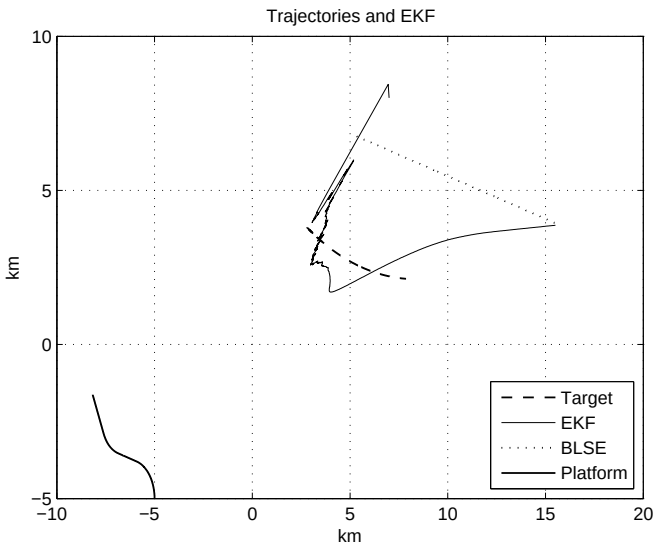


Fig. 10.5 – EKF for a circular motion

With particle filter, we present the particles at final time (see Figure 10.6). We see that the cloud of particles is not concentrated. With a more precise study, we see that the marginal densities of the filter are spread out and do not give relevant informations.

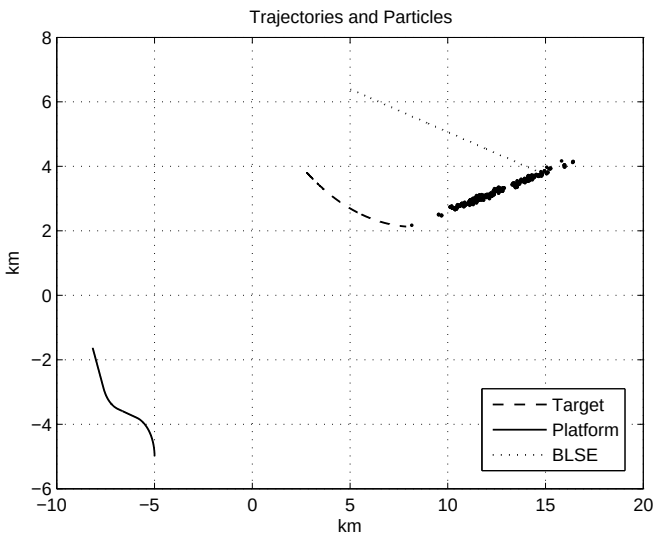


Fig. 10.6 – Particle Filter for a circular motion

When the target follows a parametric motion corrupted by worst perturbation, the EKF algorithm seems to behave like the BLS estimation. Bearings are very close to those generated by the BLS estimation. So the EKF designed for uniform linear motion presents the behavior we could expect and gives a result coherent with BLS estimation.

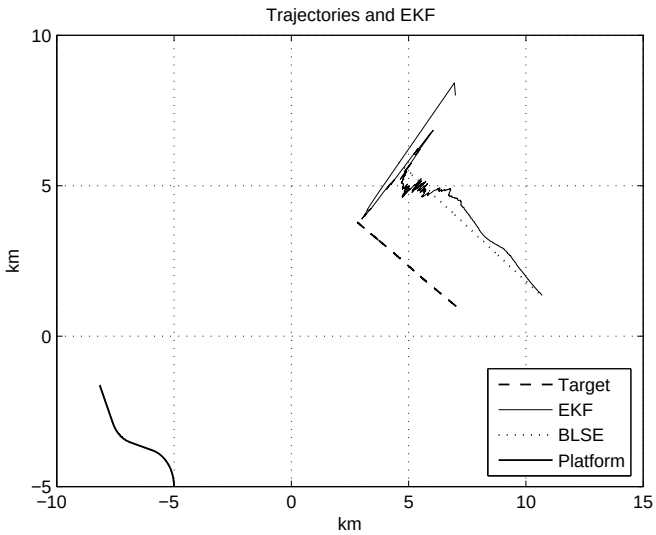


Fig. 10.7 – EKF for a worst perturbation

The particle filter has more flexibility. The marginal densities of the particle filter present a global maximum around the corresponding coordinate of $\bar{\theta}(X)$ and are quite spread out between this value and the value of the true trajectory. As the algorithm is designed for noisy uniform linear motion, it is coherent that the filter is quite concentrated around the BLS estimation for same reasons as EKF and the flexibility given by state noise in (6) allows the filter not to reject value between true trajectory and BLS estimation.

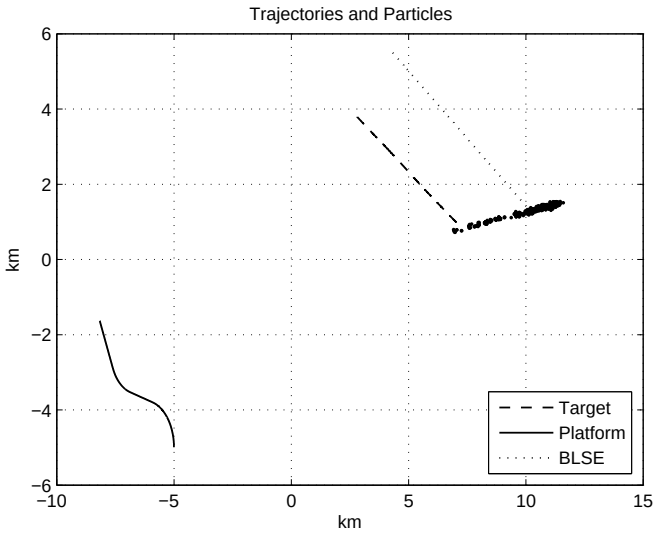


Fig. 10.8 – Particle Filter for a worst perturbation

The last simulation (see Figure 10.9) is an illustration of Proposition 10.7.

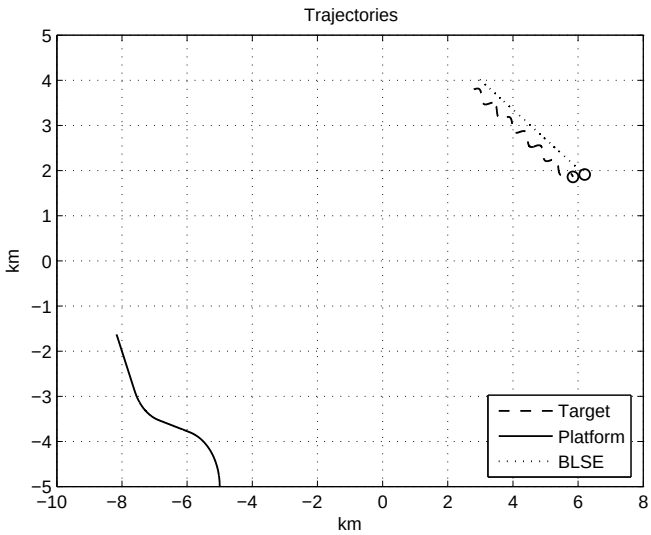


FIG. 10.9 – Effect of a periodic perturbation

10.5 Conclusion

This paper presented theoretical and practical considerations on robustness. We exhibited highly non robust situations where various classical methods are unable to recover the target's position. As a consequence, one has to be careful with results given by the usual algorithms used in BOT. We believe that the best way to avoid extreme sensitivity of the tracking to perturbations would be that the platform has a high maneuver availability. Future developments could study the hard and interesting problem of optimal maneuver for robustness.

10.6 Appendix

Proof of Proposition 10.1. Since $V_{S_{\theta^*}}$ is an open subset of $(\mathbb{R}^2)^n$, we can choose an open ball centered at S_{θ^*} and contained in $V_{S_{\theta^*}}$. This ball is clearly a connected set. So we assume that $V_{S_{\theta^*}}$ is connected. Let us consider the map $\varphi : V_{S_{\theta^*}} \to \mathbb{R}$ defined by $X \mapsto \det \mathrm{D}_{\theta,\theta}\, C_n\left(\bar{\theta}_n(X), X\right)$. This map is clearly continuous as composed from continuous applications. So its image is a connected set in $\mathbb{R}$. By the implicit function theorem, we know that the matrix $\mathrm{D}_{\theta,\theta}\, C_n\left(\bar{\theta}_n(X), X\right)$ is non-singular for all $X \in V_{S_{\theta^*}}$. Since $I_n(\theta^*)$ is a positive-definite matrix, then the range of φ is an interval of $(0, \infty)$ which proves that $\bar{\theta}_n(X)$ is a local minimum for $C_n(\theta, X)$. □

Proof of Proposition 10.2. Direct calculations yield

$$\mathrm{D}_{\theta,\theta}\, C_n(\theta, X) = 2\sum_{k=1}^{n} \mathrm{D}_{\theta,\theta}\, m(\theta, t_k)[m(\theta, t_k) - \Psi(X_k, t_k)] + 2\sum_{k=1}^{n} \nabla_\theta m(\theta, t_k) \nabla_\theta m(\theta, t_k)^{\mathrm{T}} ,$$

and for all $H \in (\mathbb{R}^2)^n$

$$\mathrm{D}_{X,\theta}\, C_n(\theta, X)(H) = -2\sum_{k=1}^{n} \nabla_\theta m(\theta, t_k) \langle \nabla_x \Psi(X_k, t_k), H_k \rangle_{\mathbb{R}^2}.$$

□

Proof of Proposition 10.3. It is immediate that K_{n,θ^*} is a self-adjoint operator. Clearly $\mathcal{R}(K_{n,\theta^*}) \subset \mathrm{span}(\alpha_n^{(1)}, \ldots, \alpha_n^{(d)})$. Let us check that this set is a linearly independent subset of $\mathbb{R}^n$. Let $\{c_1, \ldots, c_d\}$ be real numbers such that $\sum_{i=1}^d c_i \alpha_n^{(i)}(\theta^*) = 0$. By definition of $\alpha_n^{(i)}(\theta^*)$, it follows that

$$\sum_{i=1}^d c_i \alpha_{n,k}^{(i)}(\theta^*) = \|\nabla_x \Psi[S_\theta^*(t_k), t_k]\|_{\mathbb{R}^2} \times \left\langle \nabla_\theta m(\theta^*, t_k), \sum_{i=1}^d c_i \lambda_n^{(i)}(\theta^*)^{-1} e_n^{(i)}(\theta^*) \right\rangle_{\mathbb{R}^d} .$$

For convenience, let us write $u = \sum_{i=1}^d c_i \lambda_n^{(i)}(\theta^*)^{-1} e_n^{(i)}(\theta^*)$. Then $\sum_{i=1}^d c_i \alpha_n^{(i)}(\theta^*) = 0$ implies that $\langle \nabla_\theta m(\theta^*, t_k), u\rangle_{\mathbb{R}^d} = 0$ for all $k \in \{1, \ldots, n\}$. The matrix $I_n(\theta^*)$ is a Gram matrix. Since $I_n(\theta^*)$ is assumed to be non-singular, then the set of coordinates of $(\nabla_\theta m(\theta^*, t_1), \ldots, \nabla_\theta m(\theta^*, t_n))$ is linearly independent [96]. It gives that $u = 0$ and consequently that $c_1 = \ldots = c_d = 0$. It follows that $\mathcal{R}(K_{n,\theta^*}) = \mathrm{span}(\alpha_n^{(1)}, \ldots, \alpha_n^{(d)})$ is a d-dimensional space.

The operator is non-negative since, for all $v \in \mathbb{R}^n$, $\langle K_{n,\theta^*}(v), v\rangle_{\mathbb{R}^n} = \sum_{i=1}^d \langle \alpha_n^{(i)}, v\rangle_{\mathbb{R}^n}^2 \geq 0$. Applying the Cauchy-Schwarz inequality and the triangle inequality, we obtain that for all $v \in \mathbb{R}^n$

$$\|K_{n,\theta^*}(v)\|_{\mathbb{R}^n}^2 \leq \sum_{i=1}^n \|\alpha_n^{(i)}\|_{\mathbb{R}^n}^2 \; \|v\|_{\mathbb{R}^n}^2 .$$

The operator K_{n,θ^*} is bounded and takes value in a finite dimensional linear space. Then it is a compact operator [31]. □

Proof of Theorem 10.1. By definition of $\alpha_n^{(i)}(\theta^*)$ and $e_n^{(i)}(\theta^*)$, we have

$$\mathrm{D}_X \, \bar{\theta}_n(S_{\theta^*})(H) = \sum_{i=1}^d \left(\sum_{k=1}^n \alpha_{n,k}^{(i)}(\theta^*) \langle H_k, U_{\theta^*}(t_k)\rangle_{\mathbb{R}^2} \right) e_n^{(i)}(\theta^*) . \tag{7}$$

Let v_H be the vector of $\mathbb{R}^n$ defined by $v_{H,k} = \langle H_k, U_{\theta^*}(t_k)\rangle_{\mathbb{R}^2}$. Then by Pythagorean Theorem

$$\| \mathrm{D}_X \, \bar{\theta}_n(X)(H)\|_{\mathbb{R}^d}^2 = \sum_{i=1}^d \langle \alpha_i, v_H\rangle_{\mathbb{R}^n}^2, \tag{8}$$

which can also be interpreted as

$$\| \mathrm{D}_X \, \bar{\theta}_n(S_{\theta^*})(H)\|_{\mathbb{R}^d}^2 = \langle K_{n,\theta^*}(v_H), v_H\rangle_{\mathbb{R}^n} .$$

Since K_{n,θ^*} is a self-adjoint compact operator, we have the decomposition

$$\mathbb{R}^n = \mathcal{R}(K_{n,\theta^*}) \overset{\perp}{\oplus} \mathrm{Ker}\,(K_{n,\theta^*}) ,$$

and there exists an orthonormal basis of eigenvectors $(f_n^{(1)}(\theta^*), \ldots, f_n^{(d)}(\theta^*))$ of the restriction of K_{n,θ^*} to $\mathcal{R}(K_{n,\theta^*})$ with eigenvalues $0 < \mu_n^{(1)}(\theta^*) \leq \ldots \leq \mu_n^{(d)}(\theta^*)$. Then, by (7) and by definition of $\{h_i\}$ (3), we obtain

$$\mathrm{D}_X \, \bar{\theta}_n(S_\theta^*)(H) = \sum_{i=1}^d h_i D_n^{(i)}(\theta^*) .$$

Let us write $V = (V_1, \ldots, V_n) \in \mathbb{R}^n$ with $V_k = \langle H_k, U_{\theta^*}(t_k)\rangle_{\mathbb{R}^2}$. Then

$$\sum_{i=1}^{d} h_i^2 = \sum_{i=1}^{d} \langle V, f_n^{(i)}(\theta^*)\rangle_{\mathbb{R}^n}^2 = \|\Pi_{\mathcal{R}(K_{n,\theta^*})}(V)\|_{\mathbb{R}^n}^2 \ ,$$

where $\Pi_{\mathcal{R}(K_{n,\theta^*})}$ denotes the orthogonal projection on $\mathcal{R}(K_{n,\theta^*})$. This implies that $\sum_{i=1}^{d} h_i^2 \leq \|V\|_{\mathbb{R}^n}^2$ and the result follows. $\square$

Les sections qui suivent viennent détailler et compléter l'article *Robustness Considerations for Bearings-only Tracking*.

10.7 Compléments

Dans la section 10.2, on considère, dans le cas d'observations discrètes, la norme $\|X\|_\infty = \sup_{1\leq k\leq n} \|X_k\|_{\mathbb{R}^2}$ sur l'espace $(\mathbb{R}^2)^n$. Le choix de cette norme n'a pas beaucoup d'importance pour les observations discrètes puisqu'on est en dimension finie et que toutes les normes sont équivalentes. Il n'y a donc pas de difficulté afférente aux différentielles en la trajectoire.

Ce n'est pas aussi trivial pour les observations à temps continu. En temps continu, on utilise la norme $\|X\|_\infty = \sup_{t\in[0,1]} \|X(t)\|_{\mathbb{R}^2}$. L'ensemble des trajectoires E est un ouvert de $\mathscr{C}([0,1],\mathbb{R}^2)$ tel que, pour tout $X = (x,y) \in E$, on a $y(t) - O_2(t) > 0$ pour tout $t \in [0,1]$ et $\|X(t) - O(t)\|_{\mathbb{R}^2} \geq R_{\min}$[1]. Il faut que l'ensemble des trajectoires soit un ouvert d'un Banach, sans quoi il n'y a pas lieu d'y faire du calcul différentiel. Cette condition est effectivement satisfaite pour le choix de la norme décrite ci-avant. Il faut également le résultat suivant.

Lemme 10.1. *L'application $E \to \mathbb{R},\ X \mapsto \mathrm{D}_\theta\, C(\theta, X)$ est de classe $\mathscr{C}^1$ avec pour différentielle*

$$\mathrm{D}_{X,\theta}\, C(\theta, X)(H) = \int_0^1 \nabla_\theta m(\theta,t) \langle \nabla_x \Psi[X(t),t], H(t)\rangle_{\mathbb{R}^2}\, \mathrm{d}t\ .$$

Démonstration. D'après la formule de Taylor avec reste intégral, on a, pour tout $t \in [0,1]$,

$$\Psi[X(t)+H(t),t] = \Psi[X(t),t] + \langle \nabla_x \Psi[X(t),t], H(t)\rangle_{\mathbb{R}^2} + \int_0^1 (1-u)\, \mathrm{D}_{x,x}\, \Psi[X(t)+uH(t),t](H(t),H(t))\, \mathrm{d}u\ ,$$

Notons pour simplifier $R_1(t) = x(t) - O_1(t)$, $R_2(t) = y(t) - O_2(t)$ et $R(t) = \|X(t) - O(t)\|$. Le calcul nous donne

$$\mathrm{mat}\ \mathrm{D}_{x,x}\, \Psi[X(t),t] = R^{-4}(t) \begin{pmatrix} -2R_1(t)R_2(t) & R^2(t) - 2R_1^2(t) \\ R^2(t) - 2R_2^2(t) & -2R_1(t)R_2(t) \end{pmatrix}\ .$$

On a les majorations

$$\begin{aligned} |R^2(t) - 2R_1^2(t)| &\leq R^2(t)\ , \\ |R^2(t) - 2R_2^2(t)| &\leq R^2(t)\ , \\ |2R_1(t)R_2(t)| &\leq R^2(t)\ . \end{aligned}$$

Or, comme on a supposé $R(t) \geq R_{\min}$ pour tout $t \in [0,1]$, on en déduit que la norme de la Hessienne[2] $\|\mathrm{D}_{x,x}\, \Psi[X(t),t]\|_\infty$ est bornée uniformément en $t \in [0,1]$. En écrivant

$$\left| \int_0^1 (1-u)\, \mathrm{D}_{x,x}\, \Psi[X(t)+uH(t),t](H(t),H(t))\, \mathrm{d}u \right| \leq \sup_{t\in[0,1]} \|H(t)\|_{\mathbb{R}^2}^2 \int_0^1 \|\mathrm{D}_{x,x}\, \Psi[X(t),t]\|_\infty\, \mathrm{d}t\ ,$$

[1] En pratique, la trajectographie passive par capteur optique n'a de sens que sous cette hypothèse.
[2] On prend la norme d'application linéaire.

on conclut à la différentiabilité de $X \mapsto \Psi[X(t), t]$ pour tout $t \in [0, 1]$. La différentiabilité de l'application de $X \mapsto \mathrm{D}_\theta\, C(\theta, X)$ en est une conséquence quasi immédiate. □

Comme on l'a vu, l'inversibilité de la matrice de Fisher permet de définir, localement autour de la trajectoire paramétrique, le minimiseur des moindres carrés et d'en décrire le comportement. Mais l'inversibilité de la matrice de Fisher est avant tout une condition d'observabilité locale dans le cas général et globale pour certaines trajectoires, voir chapitre 5. Or, l'étude de robustesse montre que le terme différentiel qui contient l'essentiel de la dégradation de l'estimation fait intervenir directement dans son écriture (2) ou (4) la matrice d'information de Fisher et le gradient de la mesure. Ainsi, il semble qu'il existe un lien étroit entre la robustesse et la qualité d'observabilité, grandeur que l'on peut mesurer par exemple avec la trace ou le déterminant de la matrice de Fisher. Intuitivement, il semble naturel de penser que plus la qualité d'observabilité est élevée, meilleure est la robustesse de l'estimation paramétrique. Ce fait n'est pas établi pour l'instant, même si l'on dispose des majorations qui suivent.

Lemme 10.2. *Notons* $\|\mathrm{D}_X\, \bar{\theta}(S_\theta)\| = \sup_{\|H\|_\infty = 1} \|\mathrm{D}_X\, \bar{\theta}(S_\theta)(H)\|_{\mathbb{R}^d}$. *Pour des trajectoires paramétriques MRU ou MCU, on a*

$$\|\mathrm{D}_X\, \bar{\theta}(S_\theta)\| \leq \begin{cases} \operatorname{tr} I^{-2}(\theta) \times \operatorname{tr} \left(\displaystyle\int_0^1 I^2(\theta, t)\, \mathrm{d}t \right) , \\ \lambda_1^{-1}(\theta) \operatorname{tr} I(\theta) , \end{cases}$$

avec $I(\theta, t) = \nabla_\theta m(\theta, t) \nabla_\theta m(\theta, t)^{\mathrm{T}}$ *et* $\lambda_1(\theta)$ *la plus petite valeur propre de* $I(\theta)$.

Démonstration. La première majoration résulte de l'écriture de l'opérateur compact (5) $K_\theta(v) = \sum_{i=1}^d \langle \alpha_i(\theta), v \rangle_{\mathrm{L}^2}\, \alpha_i(\theta)$. Par suite, on a

$$\begin{aligned} \sup_{\|v\|_{\mathrm{L}^2}=1} \|K_\theta(v)\|_{\mathrm{L}^2} &\leq \sum_{i=1}^d \|\alpha_i(\theta)\|_{\mathrm{L}^2}^2 , \\ &\leq \sum_{i=1}^d \int_0^1 \lambda_i^{-2}(\theta)\, \|\nabla_\theta m(\theta, t)\|_{\mathbb{R}^d}^2\, \|\nabla_x \Psi[S_\theta(t), t]\|_{\mathbb{R}^2}^2\, \mathrm{d}t . \end{aligned}$$

Or, $\|\nabla_x \Psi[S_\theta(t), t]\|_{\mathbb{R}^2} \leq \|\nabla_\theta m(\theta, t)\|_{\mathbb{R}^d}$ pour un MRU ou un MCU. Ainsi, il vient

$$\|\mathrm{D}_X\, \bar{\theta}(S_\theta)\| \leq \operatorname{tr} I^{-2}(\theta) \int_0^1 \operatorname{tr} I^2(\theta, t)\, \mathrm{d}t ,$$

ce qui prouve la première inégalité. Pour la seconde, on part de l'écriture du terme différentiel (4)

$$\begin{aligned} \|\mathrm{D}_X\, \bar{\theta}(S_\theta)\| &\leq \|I^{-1}(\theta)\| \int_0^1 \|\nabla_\theta m(\theta, t)\|\, \|\nabla_x \Psi[S_{\theta^*}(t), t]\|\, \mathrm{d}t , \\ &\leq \|I^{-1}(\theta)\| \int_0^1 \left\|\nabla_\theta m(\theta, t) \nabla_\theta m(\theta, t)^{\mathrm{T}}\right\|\, \mathrm{d}t \end{aligned}$$

ce qui clôt la preuve. □

Enfin signalons que le problème de robustesse n'est pas endémique au MRU, même si pour des raisons techniques, les simulations de la section 10.4 ne concernent que des trajectoires paramétriques MRU. Ainsi, le MCU présente les mêmes défaillances de robustesse comme le montre la simulation présentée en figure 10.10 où la cible suit une trajectoire MRU-MCU par morceaux.

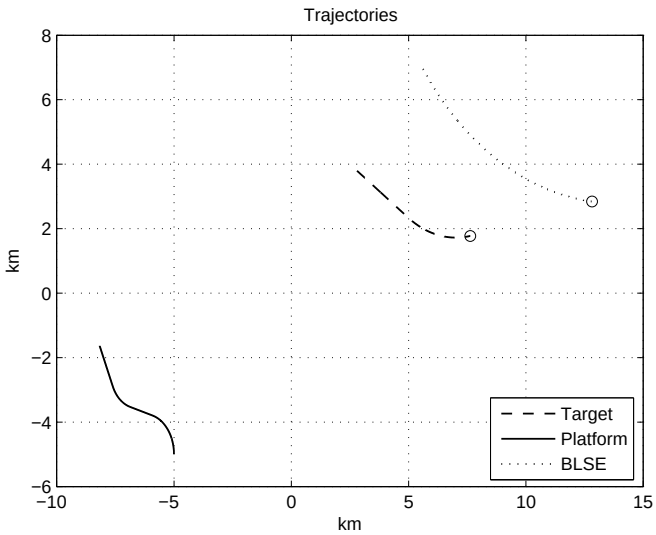


Fig. 10.10 – Sensibilité de l'estimation MCU

10.8 Corrélations dans le cas sur-critique

On a présenté dans la section 10.4 des simulations pour des cas *surcritiques*, ceux où l'estimation paramétrique est la moins robuste pour une intensité de perturbation $\|H\|_\infty$ donnée. On a donc pu voir que des fluctuations presque indécelables, puisqu'elles ne dégradent pratiquement pas le caractère linéaire d'une trajectoire MRU, voir figure 10.3, peuvent générer une estimation très fausse au regard de la trajectoire réelle. Il importe néanmoins de mentionner que de telles trajectoires ne sont pas plausibles. En effet, les trajectoires surcritiques sont calculées pour maximiser la perturbation et sont fonctions bien évidemment de la trajectoire porteur, comme cela apparaît clairement dans la preuve du théorème 10.1. Pour la simulation présentée en figure 10.3 où la cible suit une trajectoire surcritique, on représente le tracé des accélérations cible et porteur ainsi que les dérivées des phases des vitesses, voir figure 10.11. On observe clairement des simultanéités dans les variations de ces grandeurs entre la cible et le porteur, ce qui n'a pas lieu d'être puisque cible et porteur suivent, en pratique, des trajectoires décidées indépendamment[3].

Ainsi, ces cas de figures pathologiques ne seront pas observés dans des situations réelles de trajectographie passive. Pour autant, le problème de robustesse n'en est pas amoindri. En effet, on voit que pour des trajectoires MCU à courbure faible, l'estimation MRU peut être très mauvaise et ce cas de figure est tout à fait envisageable dans la réalité. L'intérêt des trajectoires surcritiques est essentiellement de fournir une borne supérieure pour la dégradation de l'estimation paramétrique. Si cette borne est faible, on sait alors que l'estimation est robuste. Par ailleurs, comme on le laisse entendre dans la section 10.5, cette borne peut être utilisé pour un travail autour de la commande optimale.

[3]Sauf bien sûr si le porteur commande pour améliorer son degré d'observabilité ou la robustesse de son estimation. Mais, dans ce cas, on peut imaginer qu'il y a un laps de temps entre les manœuvres de la cible et celles du porteur et non pas simultanéité.

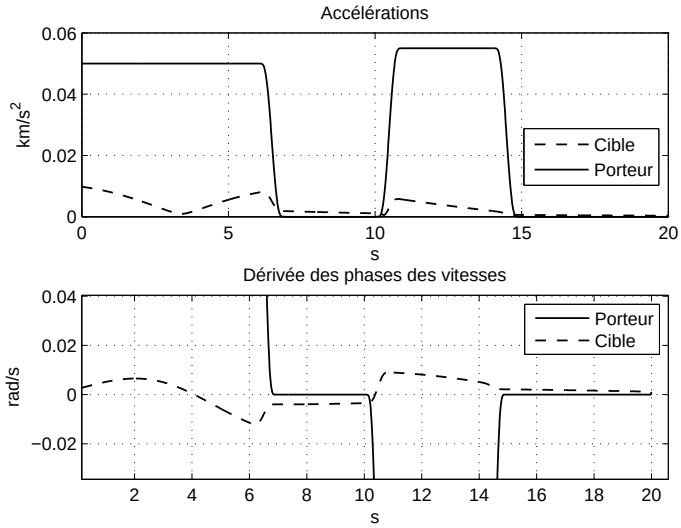


FIG. 10.11 – Simultanéités des changements de régime

10.9 Un estimateur linéarisé

En définissant les critères suivants,

$$\begin{aligned} L_n(\theta, X) &\stackrel{\text{def}}{=} \sum_{k=1}^{n} \{R_1(\theta, t_k)\cos\Psi[X(t_k), t_k] - R_2(\theta, t_k)\sin\Psi[X(t_k), t_k]\}^2 \ , \\ L_n^\sigma(\theta, X) &\stackrel{\text{def}}{=} \sum_{k=1}^{n} \{R_1(\theta, t_k)\cos Y_k - R_2(\theta, t_k)\sin Y_k\}^2 \ , \end{aligned}$$

où $R_1(\theta, t) = S_{\theta,1}(t) - O_1(t)$ et $R_2(t) = S_{\theta,2}(t) - O_2(t)$, on peut écrire, par les mêmes techniques que précédemment, l'étude asymptotique de l'estimateur

$$\check{\theta}_n^\sigma(X) \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} L_n^\sigma(\theta, X) \ ,$$

pour $X \to S_{\theta^*}$ et $\sigma \to 0$. Cet estimateur linéarisé est étudié dans la littérature de la trajectographie passive, en particulier le phénomène de biais observé quand le nombre d'observations $n \to \infty$, voir [4], [55], [39], [46].

Il y a plusieurs raisons qui motivent l'intérêt de ce nouvel estimateur. Tout d'abord, si la trajectoire S_{θ^*} est observable, on a

$$\theta^* = \arg\min_{\theta\in\Theta} L_n(\theta, S_{\theta^*}) \ ,$$

ce qui fait de $\check{\theta}_n^\sigma(X)$ un candidat intéressant dans un contexte de robustesse. Si le paramétrage est linéaire en le paramètre[4], alors le calcul de $\check{\theta}_n^\sigma(X)$ est explicite. Un autre fait plus remarquable est le lien qui existe entre $\check{\theta}_n(X)$, le minimiseur du critère linéarisé sans bruit de mesure, défini par

$$\check{\theta}_n(X) \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} L_n(\theta, X) \ ,$$

[4]Par exemple, pour des trajectoires à composantes polynomiales, donc en particulier le MRU.

et le minimiseur $\bar{\theta}_n(X)$. Par la suite, on note $R(\theta,t) = \|S_\theta(t) - O(t)\|_{\mathbb{R}^2}$.

Proposition 10.13. *On a les équivalences*

$$\mathrm{D}_{\theta,\theta}\, L_n(\theta, X_\theta) \text{ inversible} \iff I_n(\theta) \text{ inversible} ,$$

et

$$\mathrm{D}_{\theta,\theta}\, L(\theta, X_\theta) \text{ inversible} \iff I(\theta) \text{ inversible} .$$

Démonstration. Tous calculs effectués, on trouve

$$\mathrm{D}_{\theta,\theta}\, L_n(\theta, X_\theta) = 2\sum_{k=1}^{n} \eta(\theta,t_k)\eta(\theta,t_k)^{\mathrm{T}} ,$$

avec

$$\eta(\theta,t) = R(\theta,t)^{-1} \times \left(\tfrac{\partial R_1}{\partial \theta_i}(\theta,t)R_2(\theta,t) - \tfrac{\partial R_2}{\partial \theta_i}(\theta,t)R_1(\theta,t)\right)_{1\le i\le d} ,$$

et

$$\nabla_\theta m(\theta,t) = R(\theta,t)^{-2} \times \left(\tfrac{\partial R_1}{\partial \theta_i}(\theta,t)R_2(\theta,t) - \tfrac{\partial R_2}{\partial \theta_i}(\theta,t)R_1(\theta,t)\right)_{1\le i\le d} ,$$

d'où le résultat. □

Ainsi, on peut, sous la condition d'inversibilité de la matrice de Fisher $I_n(\theta^*)$, écrire le théorème des fonctions implicites et obtenir un développement asymptotique pour $\check{\theta}_n(X)$ et $\check{\theta}_n^\sigma(X)$, par un procédé identique à ce qui a été fait précédemment. Détaillons le comportement en l'absence de bruit.

Proposition 10.14. *Si la matrice $I_n(\theta^*)$ est inversible, alors il existe des voisinages de θ^* et S_{θ^*} et une application $\check{\theta}_n$ de classe $\mathscr{C}^1$ définie sur $V_{S_{\theta^*}}$ tel que, pour tout couple (θ, X) pris sur ces voisinages, on a*

$$\mathrm{D}_\theta\, L_n(\theta, S_\theta) = 0 \iff \theta = \check{\theta}_n(X) .$$

Le développement de Taylor-Young s'écrit

$$\check{\theta}_n(S_{\theta^*} + H) = \theta^* + \mathrm{D}_X\, \check{\theta}_n(S_{\theta^*})(H) + \mathrm{o}(\|H\|_\infty) ,$$

avec la différentielle donnée par la formule

$$\mathrm{D}_X\, \check{\theta}_n(S_{\theta^*})(H) = -\,\mathrm{D}_{\theta,\theta}\, L_n(\theta^*, S_{\theta^*})^{-1}\, \mathrm{D}_{X,\theta}\, L_n(\theta^*, S_{\theta^*})(H) ,$$

où

$$\mathrm{D}_{\theta,\theta}\, L_n(\theta, S_\theta) = 2\sum_{k=1}^{n} R(\theta,t_k)^2 \nabla_\theta m(\theta,t_k)\nabla_\theta m(\theta,t_k)^{\mathrm{T}} , \tag{9}$$

$$\mathrm{D}_{X,\theta}\, L_n(\theta, S_\theta)(H) = -2\sum_{k=1}^{n} R(\theta,t_k)^2 \nabla_\theta m(\theta,t_k)\langle \nabla_x \Psi[X_\theta(t_k), t_k], H_k\rangle_{\mathbb{R}^2} . \tag{10}$$

La preuve ne présente pas de difficulté particulière et est laissée au lecteur.

Dans le développement asymptotique de $\check{\theta}_n(S_{\theta^*} + H)$, on relève un point particulièrement intéressant. Si les trajectoires S_{θ^*} et O sont telles que $R(\theta^*, \cdot) \simeq \mathrm{cte}$[5], alors, vu les équations (9) et

[5]On peut écrire une asymptotique rigoureuse pour $R(\theta^*, \cdot) = R + h$, avec R une constante et $\|h\|_\infty \to 0$.

(10), on a $\mathrm{D}_X\, \check{\theta}_n(S_{\theta^*})(H) \simeq \mathrm{D}_X\, \bar{\theta}_n(S_{\theta^*})(H)$. Autrement dit, on va observer le même phénomène de non robustesse pour ce nouvel estimateur, comme l'illustre la simulation sur les figures 10.12 et 10.13. La trajectoire de paramètre $\check{\theta}_n(X)$ est désignée par la légende BL (Bearings Linearized). Les caractéristiques des trajectoires cible et porteur sont celles décrites dans le premier cas présenté en section 10.4.

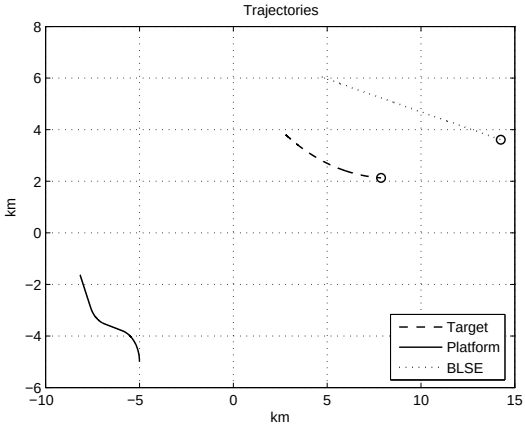


FIG. 10.12 – Estimateur $\bar{\theta}_n(X)$

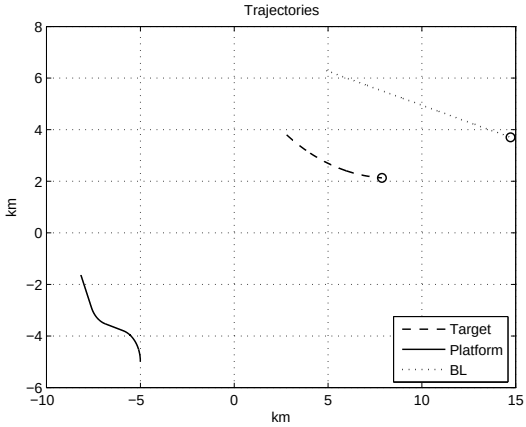


FIG. 10.13 – Estimateur $\check{\theta}_n(X)$

Précisons que supposer $R(\theta^*, \cdot) \simeq \mathrm{cte}$ a du sens en trajectographie passive par capteur optique puisque la distance cible-porteur est très élevée, relativement au déplacement des mobiles. Enfin, l'étude asymptotique pour $n \to \infty$ est également un sujet d'intérêt qui pourrait faire l'objet d'une étude future.

Chapitre 11

Modèle de régression semiparamétrique avec bruit d'état

Sommaire

11.1 Introduction **120**
11.2 Least squares estimation **122**
11.2.1 Consistency 122
11.2.2 Asymptotic normality 125
11.2.3 Application to BOT 126
11.3 Likelihood and efficiency **129**
11.3.1 Application to BOT 133
11.4 Further considerations **133**
11.5 Simulations **135**
11.6 Compléments **142**

Le chapitre qui suit est le contenu de l'article intitulé *Semiparametric regression estimation using noisy nonlinear non invertible functions of the observations*, co-écrit avec Élisabeth Gassiat et soumis en décembre 2008.

Semiparametric regression estimation using noisy nonlinear non invertible functions of the signal.

ABSTRACT. We investigate a semiparametric regression model where one gets noisy non linear non invertible functions of the signal. We focus on the application to bearings-only tracking. We first investigate the least squares estimator and prove its consistency and asymptotic normality under mild assumptions. We study the semiparametric likelihood process and prove local asymptotic normality of the model. This allows to define the efficient Fisher information as a lower bound for the asymptotic variance of regular estimators and to prove that the maximum likelihood estimator is regular and asymptotically efficient. Simulations are presented to illustrate our results.

Key words and phrases : Nonlinear regression, Semiparametric models, Bearings-only Tracking, Inverse models, Mixed Effects models.

11.1 Introduction

In bearings-only tracking (BOT), one gets information about the trajectory of a target only via bearing measurements obtained by a moving observer. This is a highly ill-posed problem which requires, so that one be able to propose solutions, the choice of a trajectory model. The literature on the subject is very large, and many algorithms have been proposed to track the target, see for instance [11], [43], [9], [101]. All these algorithms are designed for particular classes of models for the trajectory of the target. In Chapter 10, it is proven that the least squares estimator may be very sensitive to some small deterministic perturbations, in which case the algorithms are highly non robust. However, it has been also claimed in Chapter 10 that stochastic perturbations do not essentially alter the performances of the estimator. The aim of this chapter is to develop an estimation theory for a semiparametric model that applies to BOT. The model we study is the following :

$$\begin{cases} X_k = S_\theta(t_k) + \zeta_k \ , \\ Y_k = \Psi(X_k, t_k) + \varepsilon_k \ , \end{cases} \tag{1}$$

where $(t, \theta) \mapsto S_\theta(t)$ is a known map from $[0,1] \times \Theta$ to $\mathbb{R}^s$ and Θ is the parameter set (in general, a subset of a finite dimensional euclidian space), the map $(x,t) \mapsto \Psi(x,t)$ is a known function from $\mathbb{R}^s \times [0,1]$ to $\mathbb{R}$ and, in general, for a fixed t in [0,1], the map $x \mapsto \Psi(x,t)$ is non invertible, $\{t_k\}_{k\geq 0}$ is the known sequence of observation times in $[0,1]$, $\{\zeta_k\}_{k\geq 0}$ is a sequence of random variables taking values in $\mathbb{R}^s$, $\{\varepsilon_k\}_{k\geq 0}$ is a sequence of centered i.i.d. random variables with known density g with respect to Lebesgue measure on $\mathbb{R}$, with known variance σ^2 and independent of the sequence $\{\zeta_k\}_{k\geq 0}$. The process $\{X_k\}_{k\geq 0}$ is referred to as the *signal* process and is not observed while $\{Y_k\}_{k\geq 0}$ is the *observation* process. We aim at estimating θ using only the observations $\{Y_k\}_{k\geq 0}$.

In case of BOT, the signal $\{X_k\}_{k\geq 0}$ is the trajectory of the target, given by its euclidian coordinates ($s = 2$) at times $\{t_k\}_{k\geq 0}$, S_θ is the parametric trajectory the target is assumed to follow up to some parameter θ, for instance uniform linear motion, or a sequence of uniform linear and circular motions, $\{\zeta_k\}_{k\geq 0}$ is a noise sequence to take into account the fact that the model is only an idealization of the true trajectory and to allow stochastic departures of the trajectory model and $\{\varepsilon_k\}_{k\geq 0}$ is the observation noise. Since the observer is moving, if $O = \{O(t)\}_{t\in[0,1]}$ is its trajectory, the function $(x,t) \mapsto \Psi(x,t)$ is the angle, with respect to some fixed direction, of

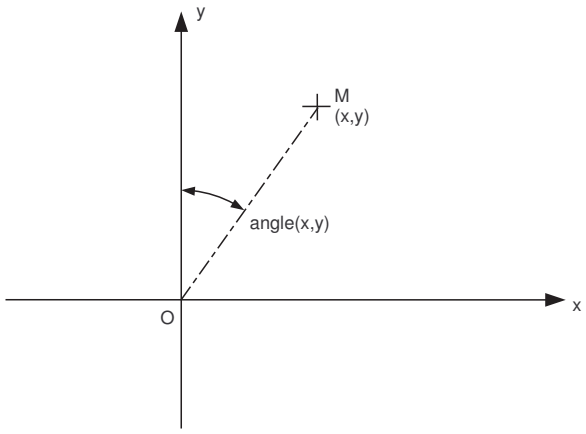


Fig. 11.1 – Angle function

$x - O(t)$ that is, for $x = (x_1, x_2)$ different from $O(t)$,

$$\Psi(x,t) = \text{angle}[x_1 - O_1(t), x_2 - O_2(t)] \ , \tag{2}$$

where the angle function is defined, see figure 11.1, by

$$\text{angle}(x,y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \text{sgn}(x) \times \mathbb{1}_{y<0} & \text{if } x \neq 0 \text{ and } y \neq 0, \\ \frac{\pi}{2} \times \text{sgn}(x) & \text{if } x \neq 0 \text{ and } y = 0, \\ \frac{\pi}{2} \times [1 - \text{sgn}(y)] & \text{if } x = 0 \text{ and } y \neq 0, \end{cases}$$

In such a case, for any z and fixed t, the set $\{x : \Psi(x,t) = z\}$ is infinite. Our aim here is to understand how it is possible to estimate the parameter θ in model (1), what are the limitations in the statistical performances, to propose estimation procedures, to build confidence regions for θ and to discuss their optimality under the weakest possible assumptions on the sequence $\{\zeta_k\}_{k\geq 0}$. Indeed, we would like to apply the results to BOT under realistic assumptions, for which it is not a strong assumption to assume that the observation noise $\{\varepsilon_k\}_{k\geq 0}$ consists of i.i.d. random variables with known distribution, but the trajectory noise $\{\zeta_k\}_{k\geq 0}$ may be quite complicated and unknown. To begin with, we will assume that the variables $\{\zeta_k\}_{k\geq 0}$ are i.i.d. with unknown distribution.

As such, the model may be viewed as a regression model with two variables, in which one of the variables is random, is not observed and follows itself a regression model. One could think that it looks like an inverse problem, or that the model may be understood as a state space model, or a mixed effects model, but in a nonstandard way, so that we have not been able to find results in the literature that apply to this setting. Throughout the chapter, observations $\{Y_k\}_{k\geq 0}$ are assumed to follow model (1) with true (unknown) parameter θ^* and the observation times are $t_k = k/n$ with $k \in \{1, \ldots, n\}$. All norms $\|\cdot\|$ are euclidian norms.

In Section 11.2, we consider least squares estimation and prove consistency and asymptotic normality in this setting, see Theorems 11.1 and 11.2. This allows to introduce basic considerations and set some assumptions. We prove that the results apply to BOT for linear observable trajectory models and when the trajectory noise has an isotropic distribution, see Theorem 11.3. Then, in Section 11.3 we study the likelihood process to set local asymptotic normality and efficiency in the parametric setting where the density of the noise $\{\zeta_k\}_{k\geq 0}$ is known, and define the efficient Fisher

information in the semiparametric setting where the density of the noise $\{\zeta_k\}_{k\geq 0}$ is unknown. This also gives an estimation criterion which may be used even if the trajectory noise is correlated. In Section 11.4, we propose strategies for semiparametric estimation and discuss possible extension of the results to possibly dependent trajectory noise $\{\zeta_k\}_{k\geq 0}$. Section 11.5 is devoted to simulations. In each section, particular attention is given to the application of the results to BOT.

11.2 Least squares estimation

In sections 11.2 and 11.3, we will use :

Assumption 11.1. *$\{\zeta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables.*

To be able to obtain a consistent estimator of θ, we require that, in the absence of noise (both observation noise and trajectory noise), the observation at all times is sufficient to retrieve the parameter. This is the observability assumption :

Assumption 11.2. *If $\theta \in \Theta$ is such that $\Psi[S_\theta(t), t] = \Psi[S_{\theta^*}(t), t]$ a.e. for all $t \in [0, 1]$, then $\theta = \theta^*$.*

If the observation noise is centered, in the absence of trajectory noise, the fact that only $\Psi[S_\theta(t), t]$ is observed with additive noise is not an obstacle to the estimation of θ under Assumption 11.2. But with trajectory noise, only the distribution of $\Psi[S_\theta(t) + \zeta_1, t]$ may be retrieved from noisy data. In case the marginal distribution of the ζ_k's is known, this may be enough, but in case it is unknown, one has to be aware of some link between the distribution of $\Psi[S_\theta(t) + \zeta_1, t]$ and θ. We thus introduce the following assumption, which will be proved to hold in some BOT situations.

Assumption 11.3. *For all t in $[0, 1]$, let us assume that*

$$\mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t]\} = \Psi[S_{\theta^*}(t), t] \ .$$

Let us now define the least squares criterion and the least squares estimator (LSE) by

$$M_n(\theta) \stackrel{\text{def}}{=} n^{-1} \sum_{k=1}^{n} \{Y_k - \Psi[S_\theta(t_k), t_k]\}^2 \ , \qquad \bar{\theta}_n \stackrel{\text{def}}{=} \arg\min_{\theta\in\Theta} M_n(\theta),$$

where $\arg\min_{\theta\in\Theta} M_n(\theta)$ is any minimizer of M_n.

11.2.1 Consistency

We assume that Θ is a compact subset of $\mathbb{R}^d$ and we will use :

Assumption 11.4. *The maps $t \mapsto \mathbb{E}\left\{\Psi[S_{\theta^*}(t) + \zeta_1, t]^2\right\}$ and $(t, \theta) \mapsto \Psi[S_\theta(t), t]$ define finite continuous functions on respectively $[0, 1]$ and $[0, 1] \times \Theta$. Moreover,*

$$\lim_{M\to\infty} \sup_{t\in[0,1]} \mathbb{E}\left(\Psi[S_{\theta^*}(t) + \zeta_1, t]^2 \mathbb{1}\left\{\Psi[S_{\theta^*}(t) + \zeta_1, t]^2 > M\right\}\right) = 0 \ .$$

Theorem 11.1. *Under assumptions 11.1 , 11.2, 11.3 and 11.4, $\bar{\theta}_n$ converges in probability to θ^* as n tends to infinity.*

The proof is a consequence of general results in M-estimation. We begin with a simple Lemma :

Lemma 1. *Under Assumption 11.1, if F is a real function on $\mathbb{R}^s \times [0,1]$ such that*

$$\sup_{t\in[0,1]} \mathbb{E}|F(\zeta_1,t)| < \infty \ , \qquad \lim_{M\to\infty} \sup_{t\in[0,1]} \mathbb{E}\{|F(\zeta_1,t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\} = 0 \ ,$$

and $\mathbb{E}[F(\zeta_1,\cdot)]$ is Riemann-integrable, then

$$n^{-1}\sum_{k=1}^{n} F(\zeta_k,t_k) = \int_0^1 \mathbb{E}[F(\zeta_1,t)]\,\mathrm{d}t + \mathrm{o}_{\mathbb{P}}(1) \ .$$

Proof. First of all, by the integrability assumption,

$$\lim_{n\to\infty} n^{-1}\sum_{k=1}^{n} \mathbb{E}[F(\zeta_k,t_k)] = \int_0^1 \mathbb{E}[F\left(\zeta_1,t\right)]\,\mathrm{d}t \ .$$

Then

$$\begin{aligned}
n^{-1}&\sum_{k=1}^{n}\{F(\zeta_k,t_k) - \mathbb{E}[F(\zeta_k,t_k)]\} \\
&= n^{-1}\sum_{k=1}^{n}\left\{F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|>M} - \mathbb{E}\left[F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|>M}\right]\right\} \\
&+ n^{-1}\sum_{k=1}^{n}\left\{F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\leq M} - \mathbb{E}\left[F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\leq M}\right]\right\} \ .
\end{aligned}$$

The variance of the second term is upper bounded by $2M^2/n$ so that the second term tends to 0 in probability as n tends to infinity and the absolute value of the first term has expectation upper bounded by $2\sup_{t\in[0,1]}\mathbb{E}\left[|F(\zeta_1,t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\right]$, which may be made smaller than any positive ϵ for big enough M, which proves the lemma. □

Proof of Theorem 11.1. Let us define

$$M(\theta) \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_\theta(t),t]\right\}^2 \,\mathrm{d}t + \sigma^2 \ .$$

Direct calculations yield

$$\begin{aligned}
M(\theta) &- M(\theta^*) \\
&= \int_0^1 \mathbb{E}\left(\left\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_\theta(t),t]\right\}^2 - \left\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_{\theta^*}(t),t]\right\}^2\right)\,\mathrm{d}t \ , \\
&= \int_0^1 \left\{\Psi[S_{\theta^*}(t),t] - \Psi[S_\theta(t),t]\right\} \times \left\{2\mathbb{E}\left(\Psi[S_{\theta^*}(t)+\zeta_1,t]\right) - \Psi[S_{\theta^*}(t),t] - \Psi[S_\theta(t),t]\right\}\,\mathrm{d}t.
\end{aligned}$$

By Assumption 11.3, it follows that

$$M(\theta) - M(\theta^*) = \int_0^1 \left\{\Psi[S_\theta(t),t] - \Psi[S_{\theta^*}(t),t]\right\}^2 \,\mathrm{d}t \ ,$$

so that $M(\theta)$ has a unique minimum at θ^* by Assumption 11.2. Also, under Assumption 11.4, the

map $\theta \mapsto M(\theta)$ is uniformly continuous from Θ to $\mathbb{R}$. Now, for any $\theta \in \Theta$,

$$M_n(\theta) = n^{-1}\sum_{k=1}^{n}\varepsilon_k^2 + 2n^{-1}\sum_{k=1}^{n}\varepsilon_k\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\} + n^{-1}\sum_{k=1}^{n}\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\}^2 \ . \quad (3)$$

It follows, by law of large numbers, that $n^{-1}\sum_{k=1}^{n}\varepsilon_k^2 = \sigma^2 + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1)$. The variance of the second expression in the right-hand side of (3) is given by

$$\mathbb{V}\left(n^{-1}\sum_{k=1}^{n}\varepsilon_k\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\}\right) = \sigma^2 n^{-2}\sum_{k=1}^{n}\mathbb{E}\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\}^2 \ ,$$

and converges to 0, so that, by Tchebychev's inequality,

$$2n^{-1}\sum_{k=1}^{n}\varepsilon_k\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\} = \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

Applying Lemma 1 yields

$$n^{-1}\sum_{k=1}^{n}\left\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]\right\}^2 = \int_0^1 \mathbb{E}\left\{\Psi[S_{\theta^*}(t)+\zeta_1,t]-\Psi[S_\theta(t),t]\right\}^2 \mathrm{d}t + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

Thus for any $\theta \in \Theta$, $M_n(\theta)$ converges in probability to $M(\theta)$. Using the compacity of Θ and the second part of Assumption 11.4, it is possible to strengthen this pointwise convergence to a uniform one :

$$\sup_{\theta\in\Theta}|M_n(\theta)-M(\theta)| = \mathrm{o}_{\mathbb{P}_{\theta^*}}(1). \quad (4)$$

Indeed, for any θ and θ' in Θ,

$$M_n(\theta)-M_n(\theta') = 2n^{-1}\sum_{k=1}^{n}\varepsilon_k\left\{\Psi[S_{\theta'}(t_k),t_k]-\Psi[S_\theta(t_k),t_k]\right\} + n^{-1}\sum_{k=1}^{n}\left\{2\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]-\Psi[S_\theta(t_k),t_k]-\Psi[S_{\theta'}(t_k),t_k]\right\}\left\{\Psi[S_{\theta'}(t_k),t_k]-\Psi[S_\theta(t_k),t_k]\right\} \ ,$$

so that, for any $\delta > 0$,

$$\sup_{\substack{(\theta,\theta')\in\Theta^2\\ \|\theta-\theta'\|\le\delta}}|M_n(\theta)-M_n(\theta')| \le \omega(\delta)\, n^{-1}\sum_{k=1}^{n}\left\{2|\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]|+2\sup_{(\theta,t)\in\Theta\times[0,1]}|\Psi[S_\theta(t),t]|+2|\varepsilon_k|\right\} \ ,$$

where $\omega(\cdot)$ is the uniform modulus of continuity of $(t,\theta)\mapsto\Psi[S_\theta(t),t]$. The right-hand side of the inequality converges in probability by Lemma 1 to a constant times $\omega(\delta)$, so that equation (4) follows from compacity of Θ. Theorem 11.1 now follows from [110, Theorem 5.7]. □

11.2.2 Asymptotic normality

Asymptotic normality of the least squares estimator will follow using usual arguments under further regularity assumptions.

Assumption 11.5. *There exists a neighborhood U of θ^* such that for all $t \in [0,1]$, the map $\theta \mapsto \Psi[S_\theta(t), t]$ possesses two derivatives on U that are continuous as functions of (θ, t) over $U \times [0,1]$.*

Let us define, for θ in U,

$$\begin{aligned} I_R(\theta) &\stackrel{\text{def}}{=} \int_0^1 \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t \ , \\ I_\Psi(\theta) &\stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left(\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_\theta(t), t]\}^2 \right) \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t \ . \end{aligned}$$

Then :

Theorem 11.2. *Under Assumptions 11.1 , 11.2, 11.3, 11.4 and 11.5, if $I_R(\theta^*)$ is non singular,*

$$\sqrt{n}\left(\bar{\theta}_n - \theta^*\right) = I_R^{-1}(\theta^*) n^{-1/2} \sum_{k=1}^n \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_{\theta^*}(t_k), t_k] + \varepsilon_k \right\} \nabla_\theta \Psi[S_{\theta^*}(t_k), t_k] + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ .$$

In particular, $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ converges in distribution, as $n \to \infty$, to $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ where

$$I_M^{-1}(\theta^*) = I_R^{-1}(\theta^*) \left[I_\Psi(\theta^*) + \sigma^2 I_R(\theta^*) \right] I_R^{-1}(\theta^*) \ .$$

Let us notice that, for a null sequence $\{\zeta_k\}_{k \geq 0}$, we retrieve the usual Fisher information matrix for the parametric regression model.

Proof. The proof follows Wald's arguments. On the set $\left\{\bar{\theta}_n \in U\right\}$ which can be assumed to be convex and which has probability tending to 1 according to Theorem 11.1,

$$\nabla_\theta M_n(\bar{\theta}_n) = 0 = \nabla_\theta M_n(\theta^*) + \int_0^1 \mathrm{D}_{\theta,\theta} \, M_n[\theta^* + s(\bar{\theta}_n - \theta^*)] \, \mathrm{d}s \, (\bar{\theta}_n - \theta^*) \ .$$

Direct calculations yield, for any $\theta \in U$,

$$\nabla_\theta M_n(\theta) = -2n^{-1} \sum_{k=1}^n \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k - \Psi[S_\theta(t_k), t_k] \right\} \nabla_\theta \Psi[S_\theta(t_k), t_k] \ ,$$

and

$$\begin{aligned} \mathrm{D}_{\theta,\theta} \, M_n(\theta) = 2n^{-1} \sum_{k=1}^n \nabla_\theta \Psi[S_\theta(t_k), t_k] \nabla_\theta \Psi[S_\theta(t_k), t_k]^{\mathrm{T}} \\ - 2n^{-1} \sum_{k=1}^n \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k - \Psi[S_\theta(t_k), t_k] \right\} \mathrm{D}_{\theta,\theta} \, \Psi[S_\theta(t_k), t_k] \ . \end{aligned} \tag{5}$$

Notice that, using Assumption 11.3, $\nabla_\theta M_n(\theta^*)$ is a centered random variable, and that, using Assumptions 11.4, 11.5, the variance of $\nabla_\theta M_n(\theta^*)$ converges to $4\left[I_\Psi(\theta^*) + \sigma^2 I_R(\theta^*)\right]$ as $n \to \infty$. Also using Assumptions 11.3, 11.4, 11.5, and applying Lemma 1, $\mathrm{D}_{\theta,\theta} \, M_n(\theta)$ converges in probability to $2I_R(\theta)$ as $n \to \infty$. Using Assumption 11.5, there exists an increasing function ω

satisfying $\lim_{\delta\to 0}\omega(\delta)=0$ such that, for all θ and θ' in U with $\|\theta-\theta'\|\leq\delta$,

$$\left\|\mathrm{D}_{\theta,\theta}\,M_n(\theta)-\mathrm{D}_{\theta,\theta}\,M_n(\theta')\right\|\leq\omega(\delta)\times n^{-1}\sum_{k=1}^{n}\left\{|\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]+\varepsilon_k|+2\right\}\ .$$

It follows that, on the set $\left\{\bar\theta_n\in U\right\}$, for all s in $[0,1]$,

$$\left\|\mathrm{D}_{\theta,\theta}\,M_n[\theta^*+s(\bar\theta_n-\theta^*)]-\mathrm{D}_{\theta,\theta}\,M_n(\theta^*)\right\|\leq\omega\left(\|\bar\theta_n-\theta^*\|\right)\times n^{-1}\sum_{k=1}^{n}\left\{|\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]+\varepsilon_k|+2\right\}\ .$$

By Lemma 1, $n^{-1}\sum_{k=1}^{n}|\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]+\varepsilon_k|=\mathrm{O}_{\mathbb{P}_{\theta^*}}(1)$ so that, using the consistency of $\bar\theta_n$, Lemma 1 and Assumption 11.5,

$$\int_0^1\mathrm{D}_{\theta,\theta}\,M_n[\theta^*+s(\bar\theta_n-\theta^*)]\,\mathrm{d}s=2I_R(\theta^*)+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ .$$

Finally, we obtain

$$\begin{aligned}&\left[I_R(\theta^*)+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\right]\sqrt{n}\left(\bar\theta_n-\theta^*\right)=\\&\qquad n^{-1/2}\sum_{k=1}^{n}\{\Psi[S_{\theta^*}(t_k)+\zeta_k,t_k]+\varepsilon_k-\Psi[S_{\theta^*}(t_k),t_k]\}\nabla_\theta\Psi[S_{\theta^*}(t_k),t_k]+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ .\end{aligned}$$

Using Assumption 11.5, the convergence in distribution to $\mathcal{N}\left(0,I_M^{-1}(\theta^*)\right)$ is a consequence of the Lindeberg-Feller Theorem and Slutzky's Lemma. □

Notice that, if $\hat I_M$ is a consistent estimator of $I_M(\theta^*)$, by Slutsky's Lemma, $n^{1/2}\hat I_M^{1/2}\left(\bar\theta_n-\theta^*\right)$ converges in distribution to the centered standard gaussian distribution in $\mathbb{R}^d$, which allows to build confidence regions with asymptotic known level. If the distribution of the trajectory noise $\{\zeta_k\}_{k\geq 0}$ is known, one may use $\hat I_M=I_M\left(\bar\theta_n\right)$. If the distribution of the noise is unknown, one could use bootstrap procedures to build confidence regions based on the empirical distribution of $\bar\theta_n$ using bootstrap replicates.

Another possibility occurs if one has a majoration

$$\mathbb{E}\left(\{\Psi[S_{\theta^*}(t)+\zeta_1,t]-\Psi[S_{\theta^*}(t),t]\}^2\right)\leq A^2, \tag{6}$$

where A denotes a known constant. Indeed, in such a case, $I_\Psi(\theta^*)$ is upper bounded (in the natural ordering of positive symetric matrices) by $A^2I_R(\theta^*)$, so that $I_M^{-1}(\theta^*)$ is upper bounded by $(A^2+\sigma^2)I_R^{-1}(\theta^*)$, and one may use $(A^2+\sigma^2)I_R^{-1}\left(\bar\theta_n\right)$ as variance matrix to obtain conservative confidence regions.

11.2.3 Application to BOT

To apply the results to BOT, one has to see whether Assumptions 11.1, 11.2, 11.3, 11.4 and 11.5 hold and if $I_R(\theta^*)$ is non singular. If $y>0$, the angle measure is given by

$$\mathrm{angle}(x,y)=\arctan(x/y)\ . \tag{7}$$

To ensure that the angle measure is given by (7), we will use the following hypothesis :

Assumption 11.6. *For all θ in Θ and t in $[0,1]$,*

$$S_{\theta,2}(t) - O_2(t) > 0 \ .$$

This hypothesis is usually made in BOT litterature. Under assumption 11.6, the observation is given by

$$\Psi(x,t) = \arctan[x_1 - O_1(t)]/[x_2 - O_2(t)] \ . \tag{8}$$

The bearing exact measurements of the non noisy possible trajectory stay inside an interval with length π. This may be seen as an assumption on the manoeuvres of the observer. Assumption 11.2 is the observability assumption which holds for models such as uniform linear motion if the observer does not move itself along uniform linear motion, or a sequence of uniform linear and circular motions, if the observer does not move along uniform linear motion or circular motion in the same time intervals as the target. Various observability properties are proved in Chapter 5. Assumptions 11.4 and 11.5 hold as soon as the trajectory model $S_\theta(t)$ is twice differentiable for all t in $[0,1]$ as a function of θ and the denominator in (8) may not be 0. The fact that $I_R(\theta^*)$ is non singular is equivalent to the observability assumption for linear models, see Proposition 5.4.

Let us recall such models. Let $(e_1,\ldots,e_p)$ be a family of continuous functions on $[0,1]$, $\theta = (a_1,\ldots,a_p,b_1,\ldots,b_p)$ in Θ a subset of $\mathbb{R}^{2p}$ and

$$S_\theta = \left(\sum_{j=1}^{n} a_j e_j, \sum_{j=1}^{p} b_j e_j\right) \ . \tag{9}$$

Thus under model (9), if the trajectory of the observer is such that Assumptions 11.2 and Assumption 11.6 hold, then Assumptions 11.4 and 11.5 hold and $I_R(\theta^*)$ is non singular.

What remains to be seen is whether Assumption 11.3 holds and it is the case under a simple assumption on the distribution of the trajectory noise :

Assumption 11.7. *ζ_1 has an isotropic distribution in $\mathbb{R}^2$ and a density with compact support such that, for all $t \in [0,1]$,*

$$S_{\theta^*,2}(t) + \zeta_{1,2} - O_2(t) > 0 \ .$$

We introduce some prior knowledge on the trajectory and on the variance of the trajectory noise to be able to obtain conservative confidence regions.

Assumption 11.8. *The trajectory $t \mapsto S_{\theta^*}(t)$ is such that, for all $t \in [0,1]$,*

$$\|S_{\theta^*}(t) - O(t)\| \geq R_{\min} \ ,$$

and a constant number A^2 such that

$$\pi^2 \left(1 + \pi^{-2/3}\right)^3 \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{R_{\min}^2} \leq A^2$$

is known.

This condition makes sense since in the context of passive tracking one usually assumes that the distance between target and observer is quite large. Let us denote by C_α a region with coverage $1-\alpha$ for the standard gaussian distribution in $\mathbb{R}^d$.

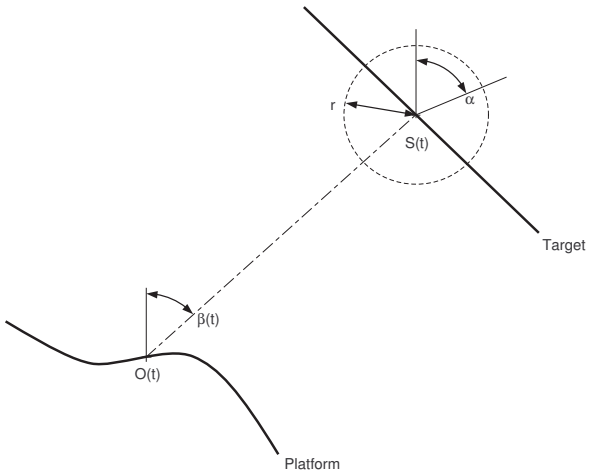


Fig. 11.2 – Isotropic noise

Theorem 11.3. *If the trajectory model $(t,\theta) \mapsto S_\theta(t)$ and the move of the observer are such that Assumptions 11.2, 11.4, 11.5 and 11.8 hold and $I_R(\theta^*)$ is non singular, or if the trajectory model is (9) satisfying Assumption 11.6 and Assumption 11.2 holds, if moreover Assumption 11.1 and 11.7 hold, then, for any $\alpha > 0$,*

$$\liminf_{n\to\infty} \mathbb{P}_{\theta^*}\left[\sqrt{n}(A^2+\sigma^2)^{-1/2} I_R^{1/2}\left(\bar\theta_n\right)\left(\bar\theta_n - \theta^*\right) \in C_\alpha\right] \geq 1-\alpha \ .$$

Proof. Under Assumption 11.7, let the density of ζ_1 be $F(\|\zeta_1\|)$. Recall that the trajectory of the observer is $\{O(t)\}_{t\in[0,1]}$. Let $\beta(t) = \Psi[S^*_\theta(t),t] = \arctan[S_{\theta^*,1}(t) - O_1(t)]/[S_{\theta^*,2}(t) - O_2(t)]$. It follows by trigonometric considerations, see figure 11.2, that

$$\begin{aligned}
&\mathbb{E}\{\Psi[S_{\theta^*}(t)+\zeta_1,t]\} \\
&\qquad = \iint_{\mathbb{R}\times(-\pi,\pi)} \arctan\left\{\frac{S_{\theta^*,1}(t) - O_1(t) + r\sin\alpha}{S_{\theta^*,2}(t) - O_2(t) + r\cos\alpha}\right\} F(r)\, r\,\mathrm{d}r\,\mathrm{d}\alpha \ , \\
&\qquad = \beta(t) + \iint_{\mathbb{R}\times(-\pi,\pi)} \arctan\left\{\frac{r\sin[\alpha-\beta(t)]}{\|S_{\theta^*}(t) - O(t)\| + r\cos[\alpha-\beta(t)]}\right\} F(r)\, r\,\mathrm{d}r\,\mathrm{d}\alpha \ .
\end{aligned}$$

Let

$$G_{\theta^*,t}(r,\alpha) = \arctan\left\{\frac{r\sin\alpha}{\|S_{\theta^*}(t) - O(t)\| + r\cos\alpha}\right\} \ .$$

Then,

$$\mathbb{E}\{\Psi[S_{\theta^*}(t)+\zeta_1,t]\} = \Psi[S_{\theta^*}(t),t] + \iint_{\mathbb{R}\times(-\pi,\pi)} G_{\theta^*,t}(r,\alpha) F(r)\, r\,\mathrm{d}r\,\mathrm{d}\alpha \ .$$

But for any $r>0$, for any $\alpha \in (-\pi,\pi)$, $G_{\theta^*,t}(r,-\alpha) = G_{\theta^*,t}(r,\alpha)$ so that

$$\mathbb{E}\{\Psi[S_{\theta^*}(t)+\zeta_1,t]\} = \Psi[S_{\theta^*}(t),t] \ .$$

Now,

$$\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_{\theta^*}(t),t] = \int_0^1 \nabla_x \Psi[S_{\theta^*}(t)+h\zeta_1,t]^{\mathrm{T}} \zeta_1\,\mathrm{d}h \ ,$$

and direct calculations provide $\|\nabla_x \Psi[x,t]\| = \|x - O(t)\|^{-1}$. Thus, for any $a \in]0,1[$,

$$\begin{aligned}\mathbb{E}\Big(\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_{\theta^*}(t),t]\}^2\Big) \\ &\leq \pi^2 \mathbb{P}\left(\|\zeta_1\| \geq a\|\Psi[S_{\theta^*}(t),t] - O(t)\|\right) + \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{(1-a)^2\|\Psi[S_{\theta^*}(t) - O(t)\|^2}\ , \\ &\leq \pi^2 \mathbb{P}\left(\|\zeta_1\| \geq aR_{\min}\right) + \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{(1-a)^2 R_{\min}^2}\ ,\end{aligned}$$

since $|\Psi(u) - \Psi(v)| \leq \pi$ for any real numbers u and v and by using the triangular inequality and Assumption 11.8. But Tchebychev's inequality leads to

$$\mathbb{E}\left(\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_{\theta^*}(t),t]\}^2\right) \leq \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{R_{\min}^2}\left(\frac{\pi^2}{a^2} + \frac{1}{(1-a)^2}\right)\ , \tag{10}$$

which is minimum for $a = \frac{1}{1+\pi^{-2/3}}$ leading to $\left(\frac{\pi^2}{a^2} + \frac{1}{(1-a)^2}\right) = \pi^2\left(1+\pi^{-2/3}\right)^3$ and

$$\mathbb{E}\left(\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_{\theta^*}(t),t]\}^2\right) \leq A^2\ .$$

To conclude one may apply the concluding remark of Section 11.2.2 to obtain asymptotic conservative confidence regions for θ. □

11.3 Likelihood and efficiency

Let $\mathcal{F}$ be the set of probability densities f on $\mathbb{R}^s$ such that for all t in $[0,1]$, for all θ in Θ,

$$\int_{\mathbb{R}^s} \Psi[S_\theta(t)+\zeta,t] f(\zeta)\,\mathrm{d}\zeta = \Psi[S_\theta(t),t]\ . \tag{11}$$

We will replace Assumptions 11.1 and 11.3 by :

Assumption 11.9. *$\{\zeta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables with density $f^* \in \mathcal{F}$.*

The normalized log-likelihood is the function on $\Theta \times \mathcal{F}$ given by

$$J_n(\theta,f) \stackrel{\text{def}}{=} n^{-1}\sum_{k=1}^n \log\left(\int_{\mathbb{R}^s} g\left\{Y_k - \Psi[S_\theta(t_k)+u,t_k]\right\}\ f(u)\,\mathrm{d}u\right). \tag{12}$$

Define

$$G\left[(\zeta,\varepsilon),t;\theta\right] \stackrel{\text{def}}{=} \log\left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t)+\zeta,t] + \varepsilon - \Psi[S_\theta(t)+u,t]\right\}\ f(u)\,\mathrm{d}u\right)\ .$$

As soon as for any (θ,f) in $\Theta\times\mathcal{F}$, it is possible to apply Lemma 1 to $G\left[(\cdot),\cdot;\theta\right]$ with the sequence $\{\zeta_k,\varepsilon_k\}_{k\geq 0}$, $J_n(\theta,f)$ converges in probability to

$$J(\theta,f) = \int_0^1\int_{\mathbb{R}^s}\int_{\mathbb{R}} \log\left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t)+\zeta,t] + \varepsilon - \Psi[S_\theta(t)+u,t]\right\}\ f(u)\,\mathrm{d}u\right) g(\varepsilon) f^*(\zeta)\,\mathrm{d}\varepsilon\,\mathrm{d}\zeta\,\mathrm{d}t. \tag{13}$$

Let

$$p_{(\theta,f)}(z,t) \stackrel{\text{def}}{=} \int_{\mathbb{R}^s} g\left\{z - \Psi[S_\theta(t)+u,t]\right\}\ f(u)\,\mathrm{d}u$$

be the density, for fixed t, of the random variable

$$Z \stackrel{\text{def}}{=} \Psi[S_\theta(t) + U, t] + V \ ,$$

where U is a random variable in $\mathbb{R}^s$ with density f independent of the real-valued random variable V with density g. Thus, $p_{(\theta^*, f^*)}(\cdot, t_k)$ is the probability density of Y_k. Then, the change of variable $z = \Psi[S_{\theta^*}(t) + \zeta, t] + \varepsilon$ in

$$\int_{\mathbb{R}} \log \left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t) + \zeta, t] + \varepsilon - \Psi[S_\theta(t) + u, t]\right\} f(u)\,\mathrm{d}u \right) g(\varepsilon)\,\mathrm{d}\varepsilon$$

leads to

$$J(\theta, f) = \iint_{\mathbb{R}\times(0,1)} p_{(\theta^*, f^*)}(z, t) \log p_{(\theta, f)}(z, t)\ \mathrm{d}z\,\mathrm{d}t \ .$$

Thus, for any (θ, f) in $\Theta \times \mathcal{F}$,

$$J(\theta^*, f^*) \geq J(\theta, f),$$

and $J(\theta^*, f^*) = J(\theta, f)$ if and only if $p_{(\theta,f)}(z,t) = p_{(\theta^*,f^*)}(z,t) \quad (t,z)$-a.e., that is the probability distribution of $Z = \Psi[S_\theta(t) + U, t] + V$ is the same as that of $\Psi[S_{\theta*}(t) + U^*, t] + V$, where U^* is a random variable in $\mathbb{R}^s$ with density f^* independent of the random variable V. But if $f \in \mathcal{F}$ and $f^* \in \mathcal{F}$, taking expectations leads to the fact that, t-a.e., $\Psi[S_\theta(t), t] = \Psi[S_{\theta*}(t), t]$, so that $\theta = \theta^*$ if Assumption 11.2 holds. In other words, $J(\theta, f)$ is maximum only for $\theta = \theta^*$.

Following the same lines as for the LSE, we may thus easily obtain that, if the probability density f^* is known, the parametric maximum likelihood estimator is consistent and asymptotically gaussian. Define the parametric maximum likelihood estimator as

$$\tilde{\theta}_n \stackrel{\text{def}}{=} \arg\max_{\theta\in\Theta} J_n(\theta, f^*) \ ,$$

where $\arg\max_{\theta\in\Theta} J_n(\theta, f^*)$ is any maximizer of $J_n(\cdot, f^*)$. If for any θ in Θ, there exists a small open ball containing θ such that Lemma 1 applies to $\sup_{\theta\in U} G[(\cdot), \cdot; \theta]$, it is possible, as in [110, Theorem 5.14], to strengthen the convergence of $J_n(\theta, f^*)$ to $J(\theta, f^*)$ in a uniforme one. The consistency of $\tilde{\theta}_n$ follows :

Theorem 11.4. *Under assumptions 11.2 and 11.9, if moreover Lemma 1 applies to* $\sup_{\theta\in U} G[(\cdot), \cdot; \theta]$, *then the estimator* $\tilde{\theta}_n$ *is consistent.*

We will use the notation Y_t for

$$Y_t \stackrel{\text{def}}{=} \Psi[S_{\theta^*}(t) + \zeta_1, t] + \varepsilon_1 \ ,$$

to simplify the writing of some integrals. We shall introduce the assumptions we need to prove the asymptotic distribution of $\tilde{\theta}_n$:

Assumption 11.10. *The following conditions are assumed :*

- *For all* (z, t) *in* $\mathbb{R} \times [0, 1]$, *the function* $\theta \mapsto p_{(\theta,f^*)}(z, t)$ *is twice continuously differentiable ;*
- *For any* θ *in* Θ, $t \mapsto \mathbb{E}\left\{\left\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t, t)\right\|^2\right\}$ *is finite and continuous ;*
- *There exists a neighborhood* U *of* θ^* *such that, for all* θ *in* U,
  - $t \mapsto \mathbb{E}\left\{\mathrm{D}_{\theta,\theta} \log p_{(\theta,f^*)}(Y_t, t)\right\}$ *is finite and continuous ;*

– *Lemma 1 applies to* $\log p_{(\theta,f)}(Y_t,t)$*, to* $\left\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t,t)\right\|^2$ *and to all components of* $\mathrm{D}_{\theta,\theta} \log p_{(\theta,f^*)}(Y_t,t)$.

Introduce the parametric Fisher information matrix :

$$I(\theta) \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{ \frac{\nabla_\theta p_{(\theta,f^*)}}{p_{(\theta,f^*)}}(Y_t,t) \frac{\nabla_\theta p_{(\theta,f^*)}}{p_{(\theta,f^*)}}^{\mathrm{T}}(Y_t,t) \right\} \mathrm{d}t \ .$$

Theorem 11.5. *Under assumptions 11.2, 11.9 and 11.10,* $\tilde{\theta}_n$ *converges in probability to* θ^* *as* n *tends to infinity.*
Moreover, if $I(\theta^*)$ *is non singular,*

$$\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big) = I^{-1}(\theta^*) \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}(Y_k,t_k) + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \ ,$$

and $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ *converges in distribution, as* $n \to \infty$*, to* $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$.

The proof follows the same lines as that of Theorems 11.1 and 11.2 and is left to the reader. Notice that under the same assumptions, it is easy to prove that the parametric model is locally asymptotically normal in the sense of Le Cam (see [77]) so that if $I(\theta^*)$ is singular, there exists no regular estimator of θ which is $\sqrt{n}$-consistent. Thus if $I_R(\theta^*)$ is non singular and the assumptions in Theorem 11.2 hold, in which case the LSE is regular $\sqrt{n}$-consistent, then $I(\theta^*)$ is also non singular.

To investigate the optimality of possible estimators in the semiparametric situation, with f^* unknown but known to belong to $\mathcal{F}$, we use Le Cam's theory as developed for non i.i.d. observations by Mc Neney and Wellner [87]. Introduce the set $\mathcal{B}$ of integrable functions b on $\mathbb{R}^s$ such that :

- $\displaystyle\int_{\mathbb{R}^s} b(\zeta)\,\mathrm{d}\zeta = 0$ and there exists $\delta > 0$ such that $f^* + \delta b \geq 0$;
- for all t in $[0,1]$, for all θ in Θ, $\displaystyle\int_{\mathbb{R}^s} \Psi\left[S_\theta(t) + \zeta, t\right] b(\zeta)\,\mathrm{d}\zeta = 0$;
- $\displaystyle\int_0^1 \mathbb{E}\left\{ \left(\frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_t,t) \right)^2 \right\} \mathrm{d}t < \infty$.

Let $\mathcal{H} = \mathbb{R}^d \times \mathcal{B}$ be endowed with the inner product :

$$\big\langle (a,b),(a',b') \big\rangle_{\mathcal{H}} \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{ \left(\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_t,t) \cdot a + \frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_t,t) \right) \left(\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_t,t) \cdot a' + \frac{p_{(\theta^*,b')}}{p_{(\theta^*,f^*)}}(Y_t,t) \right) \right\} \mathrm{d}t \ .$$

We will need only local smoothness, so we introduce :

Assumption 11.11. *There exists a neighborhood* U *of* θ^* *such that for* θ *in* U *:*

- *for all* (z,t) *in* $\mathbb{R} \times [0,1]$*, the function* $\theta \mapsto p_{(\theta,f^*)}(z,t)$ *is twice continuously differentiable ;*
- *The maps* $t \mapsto \mathbb{E}\left\{\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t,t)\|^2\right\}$ *and* $t \mapsto \mathbb{E}\left\{\mathrm{D}_{\theta,\theta} \log p_{(\theta,f^*)}(Y_t,t)\right\}$ *are finite and continuous ;*

- *For any b in $\mathcal{B}$, for all (z,t) in $\mathbb{R}\times[0,1]$, the map $\theta\mapsto p_{(\theta,b)}(z,t)$ is continuously differentiable and $t\mapsto \mathbb{E}\left\{\left\|p_{(\theta^*,f^*)}[Y_t,t]^{-1}\cdot\nabla_\theta\int g\left\{Y_t-\Psi\left[S_\theta(t)+u,t\right]\right\}b(u)\,\mathrm{d}u\right\|\right\}$ is finite and continuous;*
- *Lemma 1 applies to $\left\|\nabla_\theta\log p_{(\theta,f^*)}[Y_t,t]\right\|^2$, all components of $\mathrm{D}_{\theta,\theta}\log p_{(\theta,f^*)}[Y_t,t]$ and to $\left\|p_{(\theta^*,f^*)}(Y_t,t)^{-1}\cdot\nabla_\theta p_{(\theta,b)}(Y_t,t)\right\|$ for θ in U.*

Let $\mathbb{P}_{n,(\theta,f)}$ be the distribution of $Y_1,\ldots,Y_n$ when the parameter is θ and the density of the trajectory noise is f. For (θ,f) in $\Theta\times\mathcal{F}$, let

$$\Lambda_n(\theta,f)\stackrel{\text{def}}{=}\log\left[\frac{\mathrm{d}\mathbb{P}_{n,(\theta,f)}(Y_1,\ldots,Y_n)}{\mathrm{d}\mathbb{P}_{n,(\theta^*,f^*)}(Y_1,\ldots,Y_n)}\right]=n[J_n(\theta,f)-J_n(\theta^*,f^*)]\ .$$

Proposition 11.1. *Assume that Assumption 11.11 holds. Then, the sequence $\left(\mathbb{P}_{n,(\theta,f)}\right)_{(\theta,f)\in\Theta\times\mathcal{F}}$ of statistical models is locally asymptotically normal with tangent space $\mathcal{H}$, that is, for (a,b) in $\mathcal{H}$,*

$$\Lambda_n\left(\theta^*+\frac{a}{\sqrt{n}},f^*+\frac{b}{\sqrt{n}}\right)=W_n\left(a,b\right)-\frac{1}{2}\|(a,b)\|^2_{\mathcal{H}}+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ ,$$

where

$$W_n\left(a,b\right)=\frac{1}{\sqrt{n}}\sum_{k=1}^n\left[\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_k,t_k)\cdot a+\frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_k,t_k)\right]\ ,$$

and for any finite subset $h_1,\ldots,h_q\in\mathcal{H}$, the random vector $(W_n(h_1),\ldots,W_n(h_q))$ converges in distribution to the centered Gaussian vector with covariance $\langle h_i,h_j\rangle_{\mathcal{H}}$.

Proof.

$$\begin{aligned}\Lambda_n\left(\theta^*+\frac{a}{\sqrt{n}},f^*+\frac{b}{\sqrt{n}}\right)&=\sum_{k=1}^n\log\left[1+\frac{p_{(\theta^*+\frac{a}{\sqrt{n}},f^*)}-p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}(Y_k,t_k)+\frac{1}{\sqrt{n}}\frac{p_{(\theta^*+\frac{a}{\sqrt{n}},b)}}{p_{(\theta^*,f^*)}}(Y_k,t_k)\right]\\&=W_n(a,b)-\frac{1}{2}\|(a,b)\|^2_{\mathcal{H}}+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\ ,\end{aligned}$$

by using Taylor expansion till second order of $u\mapsto\log(1+u)$, Taylor expansion till second order of $\theta\mapsto p_{(\theta,f^*)}(z,t)$ and Taylor expansion till first order of $\theta\mapsto p_{(\theta,b)}(z,t)$, which gives the first order term $W_n(a,b)$, and then applying Lemma 1 to the second order terms to get $\frac{1}{2}\|(a,b)\|^2_{\mathcal{H}}+\mathrm{o}_{\mathbb{P}_{\theta^*}}(1)$. The convergence of $(W_n(h))_{h\in\mathcal{H}}$ to the isonormal process on $\mathcal{H}$ comes from Lindeberg Theorem applied to finite dimensional marginals. □

The interest of Proposition 11.1 is that it gives indications on the limitations on the estimation of θ^* when f^* is unknown. Indeed, the *efficient Fisher information* I^* is given by

$$\inf_{b\in\mathcal{B}}\|(a,b)\|^2_{\mathcal{H}}=a^{\mathrm{T}}I^*a\ ,$$

and if I^* is non singular, any regular estimator $\widehat{\theta}_n$ that converges at speed $\sqrt{n}$ has asymptotic covariance Σ which is lower bounded (in the sense of positive definite matrices) by $(I^*)^{-1}$. In case $I_R(\theta^*)$ is non singular and the assumptions in Theorem 11.2 hold, one may deduce that I^* is non singular.

11.3.1 Application to BOT

As seen in Section 11.2.3, the set of isotropic densities with some compact support is a subset of $\mathcal{F}$. If g is twice differentiable, positive and upper bounded, if the trajectory model $\theta \mapsto S_\theta(t)$ is twice differentiable for all $t \in [0,1]$, then Assumptions 11.10 and 11.11 hold under almost any trajectory of the observer. Indeed, one may apply Lebesgue's Theorem to obtain derivatives of integrals, and use the fact that the function $z \mapsto \arctan z$ is infinitely differentiable, has vanishing derivatives at infinity, is bounded and has two bounded derivatives.

Moreover, as seen again in Section 11.2.3, if the trajectory model is (9) and satisfies Assumption 11.2, then $I_R(\theta^*)$ is non singular, so that the efficient Fisher information I^* is non singular, and all results of Section 11.3 apply.

11.4 Further considerations

It would be of great interest to have a more explicit general expression of I^* and of greater interest to exhibit an asymptotically regular and efficient estimator $\widehat{\theta}_n$. If one could approximate the profile likelihood $\sup_{f\in\mathcal{F}} J_n(\theta, f)$, one could hope that the maximizer $\widehat{\theta}_n$ of it be a good candidate.

Another possibility would be to use Bayesian estimators. Indeed, in the parametric context, the Bernstein-von Mises Theorem tells us that asymptotically, the posterior distribution of the parameter is gaussian, centered at the maximum likelihood estimator and with variance the inverse of Fisher information (see [110] for a nice presentation). Extensions to semiparametric situations are now available, see [27]. To obtain semiparametric Bernstein-von Mises Theorems, one has to verify assumptions relating the particular model and the choice of the non parametric prior. This could be the object of further work. Then, with an adequate choice of the prior on $\Theta \times \mathcal{F}$, taking advantage of MCMC computations, one could propose bayesian methods to estimate θ^* (mean posterior, maximum posterior, median posterior for example).

To extend the results of the preceding sections in the case where the trajectory noise is no longer a sequence of i.i.d. random variables, one needs to prove laws of large numbers and central limit theorems for empirical sums such as $n^{-1}\sum_{k=1}^n F(\varepsilon_k, t_k)$, we prove some below for stationary weakly dependent sequences $\{\zeta_k\}_{k\geq 0}$. In such a case, if $M(\theta)$ and $J(\theta, f^*)$ are still the limits of $M_n(\theta)$ and $J_n(\theta, f^*)$ respectively, then asymptotics for $\bar{\theta}_n$ and $\tilde{\theta}_n$ could be obtained. Here, $J_n(\theta, f^*)$ is no longer the normalized log-likelihood, rather the marginal normalized log-likelihood, but $J(\theta, f^*)$ is still a contrast function.

Since the convergence of the expectation relies on purely deterministic arguments (Rieman integrability), we focus on centered functions. We assume in this section that :

Assumption 11.12. *$\{\zeta_k\}_{k\geq 0}$ is a stationary sequence of random variables such that, for all t in $[0,1]$,*

$$\mathbb{E}[F(\zeta_1, t)] = 0 \ .$$

Denote by $\{\alpha_k\}_{k\geq 0}$ the strong mixing coefficients of the sequence $\{\zeta_k\}_{k\geq 0}$ defined as in [100], that is, for $k \geq 1$,

$$\alpha_k \stackrel{\text{def}}{=} 2 \sup_{\ell\in\mathbb{N},\, A\in\sigma(\zeta_i : i\leq \ell),\, B\in\sigma(\zeta_i : i\geq k+\ell)} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \ ,$$

and $\alpha_0 = 1/2$. Notice that they are also an upper bound for the strong mixing coefficients of the sequence $\{F(\zeta_k, t_k)\}_{k\geq 0}$ for any sequence $\{t_k\}_{k\geq 0}$ of real numbers in $[0,1]$.

Proposition 11.2. *Under Assumption 11.12, if α_k tends to 0 as $k \to \infty$, if $\sup_{t\in[0,1]} \mathbb{E}|F(\zeta_1,t)|$ is finite and $\sup_{t\in[0,1]} \mathbb{E}\{|F(\zeta_1,t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\}$ tends to 0 as $M \to \infty$, then, as $n \to \infty$,*

$$n^{-1}\sum_{k=1}^{n} F(\zeta_k,t_k) = \mathrm{o}_{\mathbb{P}}(1) \ .$$

Proof. Using Ibragimov's inequality [62], for any $M > 0$,

$$\begin{aligned}
\mathbb{V}\left(n^{-1}\sum_{k=1}^{n} F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\le M}\right) \\
&= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}\left[F(\zeta_i,t_i)\mathbb{1}_{|F(\zeta_i,t_i)|\le M}, F(\zeta_j,t_j)\mathbb{1}_{|F(\zeta_j,t_j)|\le M}\right] \ , \\
&\le 2M^2 n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{|i-j|} \ , \\
&\le 2M^2 n^{-1}\sum_{k=0}^{n-1} \alpha_k \ ,
\end{aligned}$$

which tends to 0 by Cesaro's lemma as $n \to \infty$. The end of the proof is similar to that of Lemma 1. □

Define now

$$\alpha^{-1}(u) \stackrel{\text{def}}{=} \inf\{k \in \mathbb{N} : \ \alpha_k \le u\} = \sum_{i\ge 0} \mathbb{1}_{u<\alpha_i} \ .$$

Define also, for any t in $[0,1]$,

$$Q_t(u) \stackrel{\text{def}}{=} \inf\{x \in \mathbb{R} \ : \ \mathbb{P}[|F(\zeta_1,t)| > x] \le u\} \ ,$$

and

$$Q(u) \stackrel{\text{def}}{=} \sup_{t\in[0,1]} Q_t(u) \ .$$

We shall assume that :

Assumption 11.13. $\displaystyle\int_0^1 \alpha^{-1}(u)\, Q^2(u)\, \mathrm{d}u < \infty$.

This condition is the same as the convergence of the series $\displaystyle\sum_{k\ge 0}\int_0^{\alpha_k} Q^2(u)\, \mathrm{d}u$.
Applying [100, Theorem 1.1], one gets, for any t in $[0,1]$ and $k \ge 0$,

$$|\mathrm{Cov}[F(\zeta_0,t), F(\zeta_k,t)]| \le 2\int_0^{\alpha_k} Q^2(u)\, \mathrm{d}u \ ,$$

so that if Assumption 11.13 holds, one may define

$$\gamma^2 \stackrel{\text{def}}{=} \int_0^1 \mathbb{V}[F(\zeta_0,t)]\, \mathrm{d}t + 2\sum_{k=1}^{+\infty}\int_0^1 \mathrm{Cov}[F(\zeta_0,t), F(\zeta_k,t)]\, \mathrm{d}t \ . \tag{14}$$

Now :

Proposition 11.3. *Under Assumptions 11.12 and 11.13, if $\sigma^2 > 0$ and if for any integer k, the real function $(t,u) \mapsto \mathrm{Cov}[F(\zeta_0,t), F(\zeta_k,u)]$ is continuous on $[0,1]^2$, then*

$$n^{-1/2} \sum_{k=1}^{n} F(\varepsilon_k, t_k) \rightsquigarrow \mathcal{N}(0, \gamma^2) ,$$

as $n \to \infty$.

Proof. Let $S_n = \sum_{k=1}^{n} F(\zeta_k, t_k)$. First of all, let us prove that $n^{-1}\mathbb{V}(S_n)$ converges to σ^2 as $n \to \infty$.

$$\begin{aligned} n^{-1}\mathbb{V}(S_n) &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}[F(\zeta_i, t_i), F(\zeta_j, t_j)] , \\ &= n^{-1} \sum_{k=1-n}^{n-1} \sum_{i=1\vee(1-k)}^{n\wedge(n-k)} \mathrm{Cov}[F(\zeta_0, t_i), F(\zeta_k, t_{i+k} t)] . \end{aligned}$$

For any $K \geq 1$, using again [100, Theorem 1.1],

$$\left| n^{-1} \sum_{K \leq |k| \leq n-1} \sum_{i=1\vee(1-k)}^{n\wedge(n-k)} \mathrm{Cov}[F(\zeta_0, t_i), F(\zeta_k, t_{i+k})] \right| \leq 2 \sum_{k \geq K} \int_0^{\alpha_k} Q^2(u) \, \mathrm{d}u ,$$

which is smaller than any positive ϵ for big enough K under Assumption 11.13. Now, for any fixed integer k,

$$\begin{aligned} &\left| n^{-1} \sum_{i=1\vee(1-k)}^{n\wedge(n-k)} \mathrm{Cov}[F(\zeta_0, t_i), F(\zeta_k, t_{i+k})] - \int_0^1 \mathrm{Cov}[F(\zeta_0, t), F(\zeta_k, t)] \, \mathrm{d}t \right| \\ &\qquad \leq \sup_{(t,u) \in [0,1]^2, \, |t-u| \leq k/n} |\mathrm{Cov}[F(\zeta_0, t), F(\zeta_k, t+u)] - \mathrm{Cov}[F(\zeta_0, t), F(\zeta_k, t)]| \\ &\qquad\quad + \frac{k}{n} \sup_{t \in [0,1]} |\mathrm{Cov}[F(\zeta_0, t), F(\zeta_k, t)]| , \end{aligned}$$

which goes to 0 as $n \to \infty$ under the continuity assumption. The convergence of $n^{-1}\mathbb{V}(S_n)$ to σ^2 follows. The end of the proof is a direct application of [99, Corollary 1]. □

11.5 Simulations

The simulations have been realized using `Matlab`. The minimisation is made with the function `searchmin` by setting to 2000 the options `MaxFunEvals` and `MaxIter`, so that the method reaches the minimum.

For all the simulations, the observation time is of $20\,\mathrm{s}$. The trajectory of the observer has a speed with constant norm $\left\| \frac{\mathrm{d}O(t)}{\mathrm{d}t} \right\|$ equal to $0.25\,\mathrm{km/s}$ and makes maneuvers with norm of acceleration $\left\| \frac{\mathrm{d}^2 O(t)}{\mathrm{d}t^2} \right\|$ of approximatively $50\,\mathrm{m/s^2}$. The trajectory is mainly composed of uniform linear motions and circular uniform motions. The different sequences of the trajectory of the platform are described in the following table. The null values of acceleration correspond to uniform linear motions and the others to uniform circular motion.

time interval (s)	$0-6$	$7-10$	$11-14$	$15-20$
norm of acceleration$(\mathrm{m/s^2})$	50	0	-55	0

The positive and negative values for norm of acceleration correspond respectively to anticlockwise and clockwise circular motion. The transition sequences between circular motion and linear motion which are the time intervals $[6,7]$, $[10,11]$, and $[14,15]$ are such that the whole trajectory is $\mathscr{C}^\infty$.

The assumed parametric model is a uniform linear motion with a speed of $0.27\,\mathrm{km/s}$. The parameter θ is defined by

$$\theta = (x_0, y_0, v_x, v_y)\ .$$

where (x_0, y_0) denotes the initial position and (v_x, v_y) the speed vector. The parametric trajectory is then defined by

$$S_\theta(t) = (x_0 + v_x t,\ y_0 + v_y t)\ .$$

The observation noise is a sequence of i.i.d centered Gaussian variables with variance $\sigma = 10^{-3}\,\mathrm{rad}$. The platform receives 2000 observations.

For the first simulation, we consider a sequence $\{\zeta_k\}_{k\geq 0}$ of i.i.d Gaussian centered random variables with variance $\sigma_X^2 \times I_2$ and $\sigma_X = 10\,\mathrm{m}$[1]. The figure 11.3 shows the trajectory of the platform with a realization of a trajectory of the target and the parametric trajectory with parameter $\bar{\theta}_n$ and also the confidence area with level of 95% for the position at final time. The figure 11.6 presents the same for the maximum likelihood estimator (MLE) $\tilde{\theta}_n$.

By using Monte-Carlo methods with 1000 experiments, histograms of the coordinates of $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ are presented on figure 11.4 with the marginal probability densities of the asymptotic law $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ in dotted line. The empirical cumulative distribution functions of the coordinates of $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ are presented on figure 11.5 juxtaposed to the marginal cumulative distributions of law $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$. These two figures illustrate the convergence in distribution given by Theorem 11.2, since the sequence $\{\zeta_k\}_{k\geq 0}$ is an i.i.d. sequence of isotropic random variables.

[1]The variance is small enough to make results of Theorems 11.2 and 11.5 valid even if the assumption 11.7 is not satisfied.

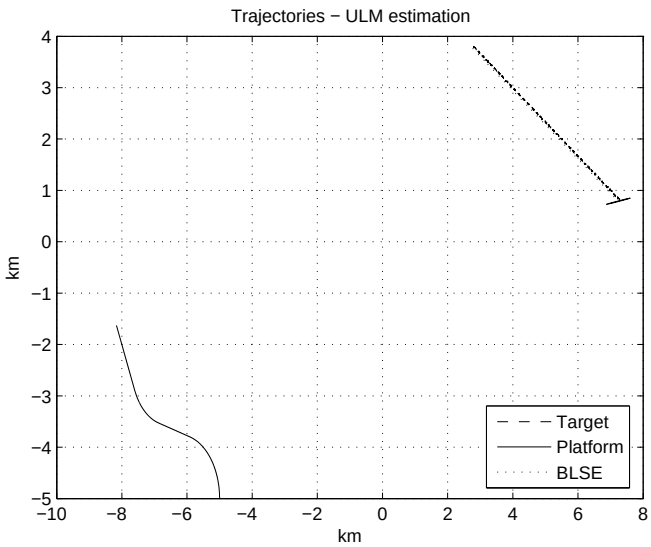


Fig. 11.3 – Trajectories with confidence area for BLSE at final position

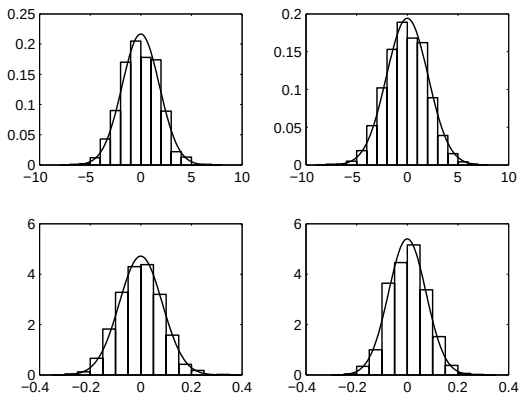


Fig. 11.4 – Histograms for BLSE with iid Gaussian isotropic sequence

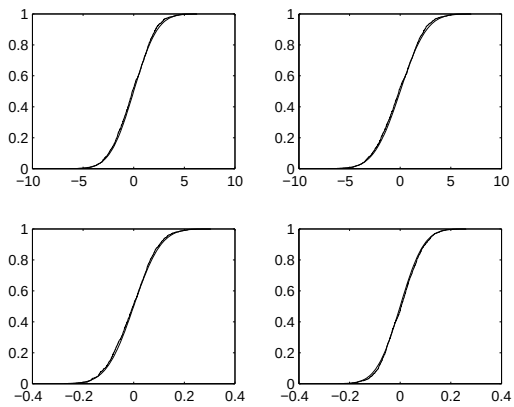


Fig. 11.5 – Cumulative distribution functions for BLSE with iid Gaussian isotropic sequence

The figure 11.7 present the histograms of the coordinates of $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ with the marginal probability densities of the asymptotic law $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$ in dotted line. Empirical cumulative distribution functions of the coordinates of $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ and marginal cumulative distributions of law $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$ are presented on figure 11.8. These two figures illustrate the convergence in distribution given by Theorem 11.5.

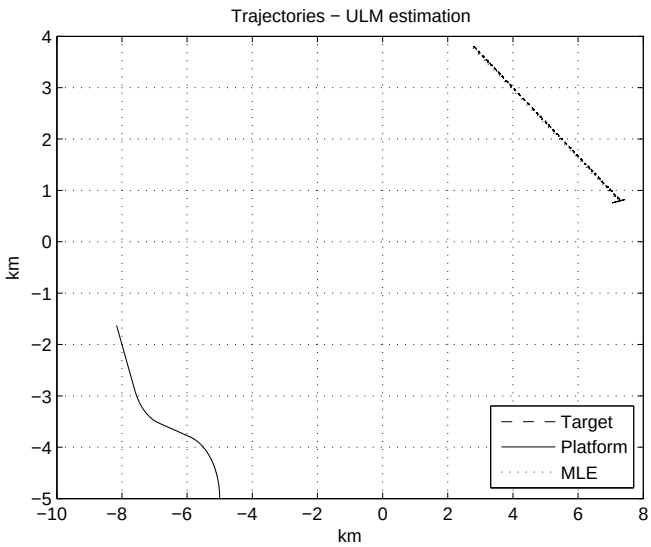


FIG. 11.6 – Trajectories with confidence area for MLE at final position

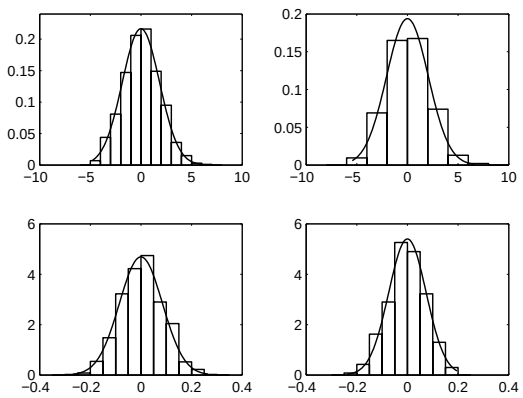


FIG. 11.7 – Histograms for MLE with iid Gaussian isotropic sequence

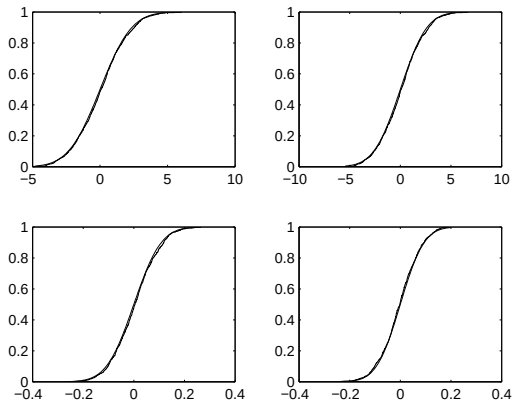


FIG. 11.8 – Cumulative distribution functions for MLE with iid Gaussian isotropic sequence

Confidence intervals for coordinates of θ^* with level of 95% are detailed in table 11.1 for $\bar{\theta}_n$ and in table 11.3 for $\tilde{\theta}_n$ and are respectively denoted by $\mathrm{IC}_1\left(\bar{\theta}_n\right)$ and $\mathrm{IC}_3\left(\tilde{\theta}_n\right)$. We also present in table 11.2 conservative confidence intervals denoted by $\mathrm{IC}_2\left(\bar{\theta}_n\right)$ built on the result provided by Theorem 11.3 with $R_{\min} = 6\,\mathrm{km}$. The choice of σ_X and $R_{\min}$ is a prior knowledge on the experiment and is made according to the knowledge of the tactical situation of BOT. Note that the majoration obtained in (10) shows that the accuracy of the conservative confidence intervals is proportional to the ratio $\mathbb{E}(\|\zeta_1\|^2)/R_{\min}^2$. This result is very interesting in practice since it shows that for high values of relative distance between target and observer and small values of state noise variance, conservative confidence intervals are of high accuracy.

For these simulations, one needs to calculate $I_\Psi\left(\bar{\theta}_n\right)$, $I_\Psi(\theta^*)$, $I\left(\tilde{\theta}_n\right)$ and $I(\theta^*)$ which involve expectations of functions of the r.v. ε_1 with law $\mathcal{N}(O, \sigma_X^2 \times I_2)$. All integrals of this type has been calculated using quadrature formula with 12 points. Abscissas and weight factors are given in

[1]. Let us detail the numerical values of $I_\Psi\left(\bar{\theta}_n\right)$ and $\sigma^2 \times I_R\left(\bar{\theta}_n\right)$ for one experiment used to build the estimators $\bar{\theta}_n$ and $\tilde{\theta}_n$. These numerical values illustrate that the contributions of state noise and observation noise are of the same level.

$$I_\Psi\left(\bar{\theta}_n\right) = 10^{-6} \times \begin{pmatrix} 0.0010 & -0.0014 & 0.0049 & -0.0094 \\ -0.0014 & 0.0024 & -0.0094 & 0.0220 \\ 0.0049 & -0.0094 & 0.0400 & -0.0950 \\ -0.0094 & 0.0220 & -0.0950 & 0.2709 \end{pmatrix},$$

$$\sigma^2 \times I_R\left(\bar{\theta}_n\right) = 10^{-6} \times \begin{pmatrix} 0.0015 & -0.0023 & 0.0082 & -0.0169 \\ -0.0023 & 0.0043 & -0.0169 & 0.0428 \\ 0.0082 & -0.0169 & 0.0728 & -0.1853 \\ -0.0169 & 0.0428 & -0.1853 & 0.5639 \end{pmatrix}.$$

Let us now precise the values of variance matrices. We have

$$I_M^{-1}\left(\bar{\theta}_n\right) = \begin{pmatrix} 3.4917 & 3.8949 & 0.1560 & -0.1399 \\ 3.8949 & 4.3496 & 0.1752 & -0.1561 \\ 0.1560 & 0.1752 & 0.0074 & -0.0062 \\ -0.1399 & -0.1561 & -0.0062 & 0.0056 \end{pmatrix},$$

and

$$I^{-1}\left(\tilde{\theta}_n\right) = \begin{pmatrix} 3.3918 & 3.7884 & 0.1526 & -0.1359 \\ 3.7884 & 4.2362 & 0.1715 & -0.1518 \\ 0.1526 & 0.1715 & 0.0072 & -0.0061 \\ -0.1359 & -0.1518 & -0.0061 & 0.0055 \end{pmatrix}.$$

The true parameter θ^* is

$$\theta^* = (2.8, 3.8, 0.225, -0.15),$$

and values of estimators $\bar{\theta}_n$ and $\tilde{\theta}_n$, used to calculate variance matrices, are

$$\bar{\theta}_n = (2.8753, 3.8841, 0.2284, -0.1530), \qquad \tilde{\theta}_n = (2.8067, 3.8077, 0.2253, -0.1502),$$

with initial coordinates $x_0,\ y_0$ given in km and $v_x,\ v_y$ given in $\mathrm{km/s}$ and the position at final time is $(7.3, 0.8)$.

$\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)\right\rvert$
7.3128	7.5747	0.2619
0.8017	0.8456	0.0439
0.2253	0.2316	0.0063
-0.1558	-0.1503	0.0055

Tab. 11.1 – Confidence intervals for BLSE at level 95%

$\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)\right\rvert$
6.0645	8.8230	2.7586
0.5917	1.0557	0.4640
0.1949	0.2619	0.0669
-0.1818	-0.1242	0.0576

Tab. 11.2 – Conservative confidence intervals for BLSE at level 95%

$\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)\right\rvert$
7.1842	7.4430	0.2588
0.7815	0.8249	0.0434
0.2222	0.2285	0.0063
-0.1529	-0.1475	0.0054

Tab. 11.3 – Confidence intervals for MLE at level 95%

It appears that the maximum likelihood estimator $\tilde{\theta}_n$ is a bit more accurate than $\bar{\theta}_n$. It is not surprising since the MLE is designed specifically for the model, and takes into account the state noise. Nevertheless, because of the high calculation cost for the MLE, the BLSE is in practice a very useful alternative.

For the second simulation, we consider the case of a sequence $\{\zeta_k\}_{k\geq 0}$ of i.i.d Gaussian centered random variables with variance $\sigma_X^2 \times \begin{pmatrix} 6^2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\sigma_X = 10\,\mathrm{m}$. It seems that the results given by Theorems 11.2 and 11.5 still hold, even though the sequence $\{\zeta_k\}_{k\geq 0}$ does not have an isotropic distribution, see Figures 11.9, 11.10, 11.11 and 11.12.

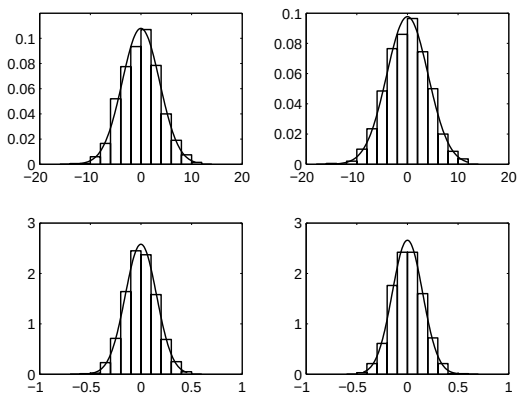


Fig. 11.9 – Histograms for BLSE with iid Gaussian non-isotropic sequence

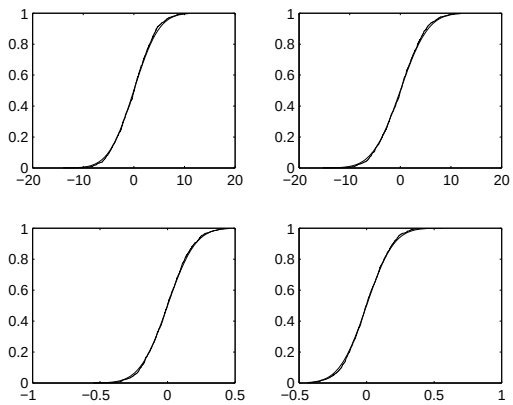


Fig. 11.10 – Cumulative distribution functions for BLSE with iid Gaussian non-isotropic sequence

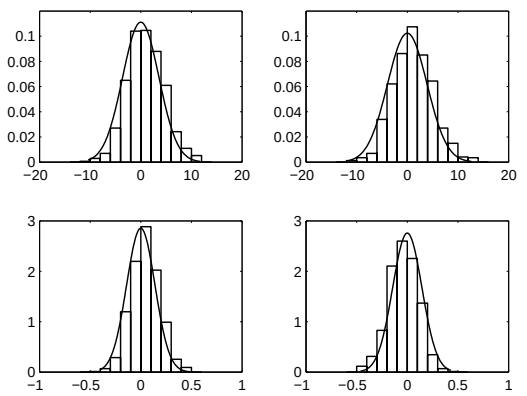


Fig. 11.11 – Histograms for MLE with iid Gaussian non-isotropic sequence

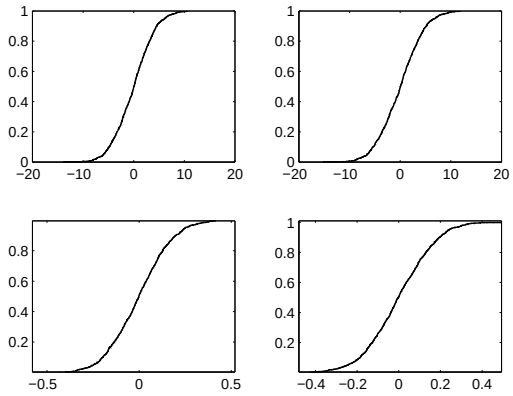


Fig. 11.12 – Cumulative distribution functions for MLE with iid Gaussian non-isotropic sequence

The estimators values are

$$\bar{\theta}_n = (2.8383, 3.8440, 0.2264, -0.1516)\ , \qquad \tilde{\theta}_n = (2.7984, 3.7999, 0.2253, -0.1499)\ .$$

The values of variance matrices for the two estimators are

$$I_M^{-1}\left(\bar{\theta}_n\right) = \begin{pmatrix} 15.4505 & 17.0122 & 0.6174 & -0.6253 \\ 17.0122 & 18.7661 & 0.6863 & -0.6889 \\ 0.6174 & 0.6863 & 0.0263 & -0.0250 \\ -0.6253 & -0.6889 & -0.0250 & 0.0253 \end{pmatrix}\ ,$$

and

$$I^{-1}\left(\tilde{\theta}_n\right) = \begin{pmatrix} 12.9538 & 14.0399 & 0.4766 & -0.5214 \\ 14.0399 & 15.2720 & 0.5262 & -0.5661 \\ 0.4766 & 0.5262 & 0.0197 & -0.0192 \\ -0.5214 & -0.5661 & -0.0192 & 0.0210 \end{pmatrix}\ .$$

The confidence intervals detailed in table 11.4 and table 11.6 show that the maximum likelihood estimator $\tilde{\theta}_n$ is significantly more accurate than the BLSE. Comparing to the first simulation where the difference is not so large, the higher accuracy of $\tilde{\theta}_n$ can be understood because of the higher level state noise in this simulation. Then, taking into account this state noise for estimating the parameter provides a significantly better result. The conservative intervals for $R_{\min} = 6\,\mathrm{km}$ described in table 11.5 are quite large compared to those obtained for the first simulation. This inaccuracy results directly from the large value of $\mathbb{E}\left(\|\zeta_1\|^2\right)$ chosen for the state noise.

$\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)\right\rvert$
7.1040	7.6275	0.5235
0.7698	0.8552	0.0854
0.2204	0.2323	0.0119
-0.1574	-0.1457	0.0117

Tab. 11.4 – Confidence intervals for BLSE at level 95%

$\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)\right\rvert$
1.5049	13.2266	11.7218
-0.1740	1.7990	1.9730
0.0842	0.3686	0.2844
-0.2740	-0.0291	0.2449

Tab. 11.5 – Conservative confidence intervals for BLSE at level 95%

$\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)\right\rvert$
7.0721	7.5366	0.4645
0.7643	0.8388	0.0746
0.2201	0.2305	0.0103
-0.1552	-0.1446	0.0107

Tab. 11.6 – Confidence intervals for MLE at level 95%

For the third and last simulation, the sequence $\{\zeta_k\}_{k\geq 0}$ is an AR(1) series such that, for all k integer,

$$\zeta_{k+1} = \Phi\zeta_k + \eta_k \ ,$$

where $\Phi = 0.6$ and $\{\eta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables with law $\mathcal{N}(0, \sigma_\eta^2)$ and $\sigma_\eta = 8\,\mathrm{m}$. Thus, the sequence of state noise $\{\zeta_k\}_{k\geq 0}$ is a dependent stationary sequence such that the mixing coefficient α_k tends exponentially fast to zero as k tends to infinity. Then, we observe the predicted behavior described by Proposition 11.3. Indeed, by drawing the densities and cumulative distribution functions of the centered Gaussian law with the empirical variance, we observe a very good adequacy to the Gaussian behavior, see figures 11.13 and 11.14.

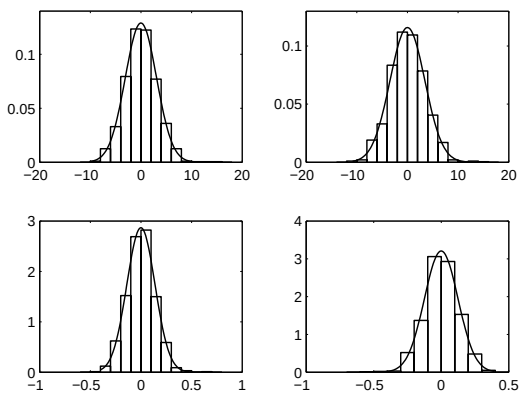


Fig. 11.13 – Histograms for AR(1) sequence, Gaussian adequacy

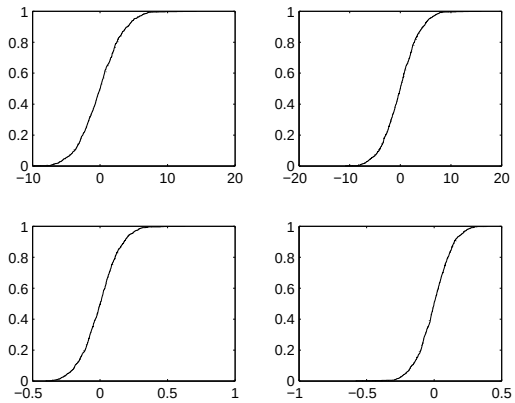


Fig. 11.14 – Cumulative distribution functions for AR(1) sequence, Gaussian adequacy

11.6 Compléments

L'étude présentée dans ce chapitre est focalisée sur le problème de la trajectographie passive par mesure d'angles dans le cas mono-porteur, problème éminemment mal posé. Il serait intéressant d'examiner le cas où la fonction $\Psi(\cdot, t)$ avec $t \in [0, 1]$ est une fonction inversible. Cette

condition est réalisée avec d'autres contextes concrets de trajectographie, comme le cas multi-porteur avec des mesures synchronisées, ou alors, la trajectographie par mesures d'angles et de distances relatives, mesures fournies respectivement par un capteur optique et une mesure radar. On pourrait ainsi construire un estimateur de la loi du bruit avec des techniques d'estimation non-paramétrique. Rappelons le modèle considéré qui suit les équations

$$\begin{cases} X_k = S_\theta(t_k) + \zeta_k \ , \\ Y_k = \Psi(X_k, t_k) + \varepsilon_k \ . \end{cases}$$

Pour un noyau K donné [108], si l'hypothèse 11.3 est satisfaite, l'estimateur

$$\hat{f}_n(x) \stackrel{\text{def}}{=} \frac{1}{nh} \sum_{k=1}^{n} K \left[\frac{\Psi(\cdot, t_k)^{-1}(Y_k) - S_{\bar{\theta}_n}(t_k) - x}{h} \right]$$

semble un bon candidat pour l'estimation de la densité de la loi du bruit et pourrait être utilisé par une technique de *plug-in* pour réaliser une meilleure estimation du paramètre θ^*.

Quatrième partie

Filtrage

Chapitre 12

Filtrage de Kalman

Sommaire

12.1 Introduction **147**
12.2 Le modèle Linéaire Gaussien **147**
12.3 Le filtre de Kalman étendu (EKF) **149**
12.4 EKF avec bruit d'état **152**
12.5 Conclusion **154**

12.1 Introduction

Le filtre inventé par Rudolf E. Kalman en 1960 permet de résoudre le problème de filtrage linéaire avec bruits additifs gaussiens. Ses travaux sont les premiers du genre à présenter la résolution du filtre optimal d'une chaîne de Markov cachée. Les articles de Kalman [70] et Bucy et al.[22] inspireront de nombreux travaux pour adresser des modèles de complexité très supérieure qui sont non-linéaires et/ou non gaussiens. Ainsi, de multiples stratégies dérivées de la technique de Kalman vont voir le jour et la trajectographie passive, de par ses caractéristiques défavorables avec une observation très non-linéaire, va constituer un terrain d'investigation privilégié pour ces recherches. Le plus simple et le plus couramment utilisé de ces filtres dérivés est l'EKF (Extended Kalman Filter). Il motive, aujourd'hui encore et malgré l'engouement pour les filtres particulaires, des recherches actives pour son application au BOT.

Dans ce chapitre, on rappelle en section 12.2 le filtre de Kalman pour le cas d'école linéaire gaussien. Puis, on présente en section 12.3 la définition de l'EKF et comment celui-ci est utilisé en trajectographie passive avec un jeu de coordonnées dites *polaires modifiées*. Dans la section 12.4, des simulations numériques montrent un phénomène inattendu qu'est le comportement robuste de l'EKF avec bruit d'état dans des situations où les méthodes batch sont inappropriées.

12.2 Le modèle Linéaire Gaussien

Le modèle linéaire gaussien est décrit par les équations

$$\begin{cases} X_k = F_k X_{k-1} + c_k + G_k W_k \ , \\ Y_k = H_k X_k + d_k + V_k \ . \end{cases}$$

On suppose que :

– les coefficients F_k, c_k, G_k, H_k et d_k sont déterministes et à valeurs respectivement dans $\mathbb{R}^{s\times s}$, $\mathbb{R}^s$, $\mathbb{R}^{s\times p}$, $\mathbb{R}^{q\times s}$ et $\mathbb{R}^q$;
– les séquences des bruits d'état et d'observation $\{W_k\}_{k\geq 0}$ et $\{V_k\}_{k\geq 0}$ sont des séquences de vecteurs gaussiens indépendants de variances respectives Q_k et R_k et également indépendants de X_0 ;
– la loi initiale de X_0 est la loi normale $\mathcal{N}(\bar{X}_0, Q_0)$.

Rappelons les notations introduites au chapitre 4 pour décrire l'équation d'évolution du filtre optimal

$$\pi_{k|k-1} \stackrel{\text{def}}{=} \mathscr{L}(X_k|Y_{0:k-1}) \ , \qquad \pi_k \stackrel{\text{def}}{=} \mathscr{L}(X_k|Y_{0:k}) \ .$$

On montre que la séquence $\{X_k, Y_k\}_{k\geq 1}$ est une suite de vecteurs gaussiens et les séquences de filtres $\{\pi_{k|k-1}\}_{k\geq 1}$ et $\{\pi_k\}_{k\geq 1}$ sont également des séquences de lois gaussiennes. On note $\hat{X}_{k|k-1} = \mathbb{E}(X_k|Y_{0:k-1})$, $P_{k|k-1} = \mathbb{V}(X_k|Y_{0:k-1})$, $\hat{X}_k = \mathbb{E}(X_k|Y_{0:k})$, $P_k = \mathbb{V}(X_k|Y_{0:k})$ et on a

$$\pi_{k|k-1} = \mathcal{N}\left(\hat{X}_{k|k-1}, P_{k|k-1}\right) \ , \qquad \pi_k = \mathcal{N}\left(\hat{X}_k, P_k\right) \ .$$

Ainsi, la détermination du filtre consiste au calcul de l'espérance $\hat{X}_k$ et de la matrice de covariance P_k. L'algorithme de Kalman-Bucy permet un calcul récursif de ces quantités. Les conditions initiales sont données par le système :

$$\begin{cases} \hat{X}_0 = \bar{X}_0 \ , \\ P_0 = Q_0 \ . \end{cases}$$

Les équations d'évolution du filtre sont décrites par
– les équations de prédiction :

$$\begin{cases} \hat{X}_{k|k-1} = F_k \hat{X}_k + c_k \ , \\ P_{k|k-1} = F_k P_{k-1} + G_k Q_k G_k^{\mathrm{T}} \ ; \end{cases}$$

– et les équations de correction

$$\begin{cases} \hat{X}_k = \hat{X}_{k|k-1} + K_k[Y_k - (H_k \hat{X}_{k|k-1} + d_k)] \ , \\ P_k = [I - K_k H_k] P_{k|k-1} \ , \\ K_k = P_{k|k-1} H_k^{\mathrm{T}} [H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k]^{-1} \ . \end{cases}$$

Plusieurs approches permettent d'établir ces équations d'évolution [5], [29]. Le modèle linéaire gaussien est un des rares cas solubles parmi les problème de filtrage. Le filtre de Kalman-Bucy s'étend au cas *conditionnellement gaussien* découvert par Lipster et al.[81]. Pour ce modèle, les coefficients dépendent des observations passées et la séquence $\{X_k, Y_k\}_{k\geq 1}$ n'est plus une suite gaussienne. Cependant, le filtre π_n est une loi gaussienne et l'algorithme de Kalman-Bucy produit la moyenne et la variance de cette loi. En relâchant l'hypothèse gaussienne sur les bruits, on montre que le filtre de Kalman-Bucy fournit l'estimée linéaire de variance minimale. Pour des situations plus générales avec un modèle non-linéaire et/ou non gaussien, le filtre de Kalman-Bucy n'est plus adapté et de nouvelles stratégies sont nécessaires.

Avant de clore cette section, penchons-nous sur la stabilité du filtre de Kalman-Bucy. Afin de tenir compte de la dépendance en la loi initiale, on opte pour la notation $\phi_{\nu,n}$ qui désigne le filtre optimal de loi initiale ν. Rappelons que la propriété de stabilité d'un filtre consiste, pour deux lois initiales ν et ν', à montrer que la distance en variation totale entre $\phi_{\nu,n}$ et $\phi_{\nu',n}$ est de limite

nulle, *i.e.*

$$\lim_{n\to\infty} \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} = 0 \quad \mathbb{P}_\star - a.s.$$

La stabilité du filtre de Kalman est un phénomène bien compris, voir [15], [38], [84], [85] et [82, sections 14.6, 16.2]. Toutefois, les méthodes mises en œuvre pour prouver la robustesse du filtre sont très spécifiques au cas linéaire et ne s'adaptent pas du tout aux problèmes de filtrage non-linéaire. De nombreuses autres stratégies ont été envisagées pour répondre à cette question difficile de stabilité du filtre optimal dans le cas non-linéaire, voir section 13.1. Mais il faut attendre les travaux récents de Kleptsyna et al.[73], Douc et al.[40], pour couvrir le cas d'un noyau gaussien pour des HMM ergodiques et ceux de Oudjane et al.[93], Crisan et al.[32], Handel [111], [112] ainsi que les résultats du chapitre 13 pour le cas non-ergodique.

12.3 Le filtre de Kalman étendu (EKF)

Le principe du filtre de Kalman étendu communément appelé EKF (Extended Kalman Filter) est très simple. On considère un système non-linéaire

$$\begin{cases} X_k = f_k(X_{k-1}) + g_k(X_{k-1})W_k \ , \\ Y_k = h_k(X_k) + V_k \ , \end{cases} \tag{1}$$

où $\{X_k\}_{k\geq 0}$, $\{Y_k\}_{k\geq 0}$, $\{W_k\}_{k\geq 0}$, $\{V_k\}_{k\geq 0}$ sont à valeurs respectivement dans $\mathbb{R}^s$, $\mathbb{R}^q$, $\mathbb{R}^p$ $\mathbb{R}^q$, les fonctions f_k, g_k et h_k sont déterministe définies sur $\mathbb{R}^s$ à valeurs dans $\mathbb{R}^s$, $\mathbb{R}^{s\times p}$, $\mathbb{R}^q$ et $\{W_k\}_{k\geq 0}$ et $\{V_k\}_{k\geq 0}$ sont des bruits blancs gaussiens de covariances respectives Q_k^W, Q_k^V indépendants entre eux et indépendants de la condition initiale X_0. On fait, à chaque itération, les approximations suivantes :

$$\begin{aligned} f_k(x) &\simeq f_k(\hat{X}_{k-1}) + \nabla f_k(\hat{X}_{k-1})(x - \hat{X}_{k-1}) \ , \\ g_k(x) &\simeq g_k(\hat{X}_{k-1}) \ , \\ h_k(x) &\simeq h_k(\hat{X}_{k|k-1}) + \nabla h_k(\hat{X}_{k|k-1})(x - \hat{X}_{k|k-1}) \end{aligned}$$

Le système (1) est remplacé par

$$\begin{cases} X_k = F_k(X_{k-1} - \hat{X}_{k-1}) + f_k + G_k W_k \ , \\ Y_k = H_k X_k + V_k \ , \end{cases}$$

où $F_k = \nabla f_k(\hat{X}_{k-1})$, $f_k = f_k(\hat{X}_{k-1})$, $G_k = g_k(\hat{X}_{k-1})$, $H_k = \nabla h_k(\hat{X}_{k|k-1})$ et $h_k = h_k(\hat{X}_{k|k-1})$. On applique ensuite le filtre de Kalman (classique) sur ces équations linéarisées.

On reprend la formulation du problème de trajectographie passive comme chaîne de Markov cachée décrite en page 24. On note l'état du système à l'instant t_k par $X_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, où (x_k, y_k) désigne la position de la cible dans le plan et $(\dot{x_k}, \dot{y_k})$ le vecteur vitesse. Les équations d'état et d'observation sont

$$\begin{cases} X_k = F X_{k-1} + \zeta_k \ , & (2) \\ Y_k = h(X_k, t_k) + \varepsilon_k \ , & (3) \end{cases}$$

avec X_k la matrice colonne $X_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, $F = \begin{pmatrix} I_2 & \Delta_T I_2 \\ 0 & I_2 \end{pmatrix}$ où Δ_T désigne la période

d'échantillonnage, $\{\zeta_k\}_{k\geq 0}$ définie par $\zeta_k = \beta G V_k$ où $\beta > 0$, $G = \begin{pmatrix} \frac{\Delta_T^2}{2} I_2 \\ \Delta_T I_2 \end{pmatrix}$ et $\{V_k\}_{k\geq 0}$ est une suite de variables indépendantes gaussiennes standards dans $\mathbb{R}^2$, $\{\varepsilon_k\}_{k\geq 0}$ une suite de variables i.i.d. de loi $\mathcal{N}(0, \sigma^2)$ indépendantes de $\{\zeta_k\}_{k\geq 0}$ et h l'application définie, pour $x = \begin{pmatrix} x_1 & x_2 & \dot{x_1} & \dot{x_2} \end{pmatrix}^{\mathrm{T}}$, par

$$h(x,t) \stackrel{\text{def}}{=} \Psi[(x_1, x_2), t] \ ,$$

où $\Psi[(x_1, x_2), t]$ est défini par $\Psi[(0,0), t] = 0$ et $\Psi[(x_1, x_2), t] = \mathrm{angle}[x_1 - O_1(t), x_2 - O_2(t)]$ pour $(x_1, x_2) \neq O(t)$ avec la fonction d'angle donnée par

$$\mathrm{angle}(x,y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \mathrm{sgn}(x) \times \mathbb{1}_{y<0} & \text{si } x \neq 0 \text{ et } y \neq 0, \\ \frac{\pi}{2} \times \mathrm{sgn}(x) & \text{si } x \neq 0 \text{ et } y = 0, \\ \frac{\pi}{2} \times [1 - \mathrm{sgn}(y)] & \text{si } x = 0 \text{ et } y \neq 0. \end{cases}$$

La partie non bruitée de l'équation (2) "$X_k = F X_{k-1}$" correspond à une dynamique MRU. Le niveau du bruit d'état choisi par les praticiens sert de variable d'ajustement pour traduire la confiance dans la dynamique supposée. Moins le bruit est élevé et plus on fait confiance à la dynamique d'état MRU, plus ce bruit est élevé et plus on autorise de variations à la dynamique de la cible.

L'implémentation du filtre EKF pour le BOT en coordonnées cartésiennes montre un comportement très instable. Un nouveau jeu de coordonnées baptisées *coordonnées polaires modifiées* ou MPC (Modified Polar Coordinates) élaborées par Aidala et al.[116] permettent de « transporter » la non-linéarité de l'équation d'observation vers l'équation d'état. Le vecteur d'état devient, par ce changement de coordonnées,

$$Z_k = \begin{pmatrix} \dot{\beta}_k & \dot{r_k}/r_k & \beta_k & 1/r_k \end{pmatrix}^{\mathrm{T}}$$

où $\beta_k = \Psi[(x_k, y_k), t_k]$ et $r_k = \|(x_k, y_k) - O(t_k)\|$. L'équation d'observation (3) est très simple dans ces nouvelles coordonnées puisqu'elle s'écrit

$$Y_k = H Z_k + \varepsilon_k \ ,$$

avec $H = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}$. Mais pour l'équation d'état, son écriture est beaucoup moins simple qu'en cartésiennes. Pour garder des notations légères, on utilisera le difféomorphisme φ_k qui permet de passer des coordonnées cartésiennes aux coordonnées polaires modifiées

$$Z_k = \varphi_k(X_k) \iff X_k = \varphi_k^{-1}(Z_k) \ .$$

On notera la dépendance en k pour ce changement de variables due au fait que les coordonnées polaires modifiées sont des coordonnées relatives. Le BOT écrit en MPC devient

$$\begin{cases} Z_k = \varphi_k \left[F \varphi_k^{-1}(Z_{k-1}) + \zeta_k \right] \ , \\ Y_k = H Z_k + \varepsilon_k \ , \end{cases}$$

et l'algorithme EKF est alors implémenté pour ce système d'équations.

Pour les simulations qui suivent, la durée d'observation est de $20\,\mathrm{s}$ et la période d'échantillonnage de $0.01\,\mathrm{s}$. Le porteur est à vitesse constante de $250\,\mathrm{m/s}$ et effectue des manœuvres de l'ordre

de $50\,\mathrm{m/s^2}$. On présente en figure 12.1 une réalisation du filtre EKF sur une cible MRU. Le bruit d'état du filtre qui est le paramètre β dans $\zeta_k = \beta G V_k$ est supposé nul. Le bruit d'observation est celui des observations réellement acquises à savoir $\sigma = 5 \times 10^{-3}\,\mathrm{rad}$.

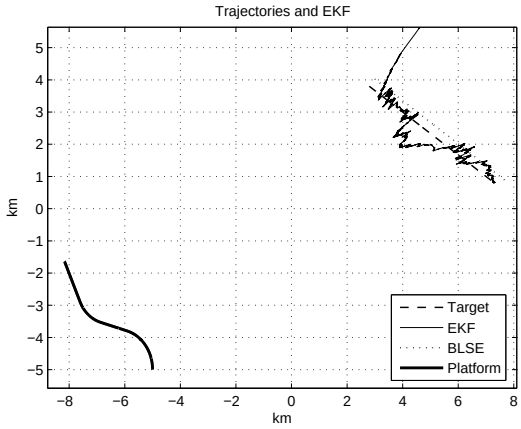


Fig. 12.1 – EKF sans bruit d'état, cible en MRU

Le comportement du filtre EKF est satisfaisant puisqu'on le voit converger vers la trajectoire réelle. Cette situation n'est pas marginale et peut être observée pour d'autres trajectoires MRU. Toutefois, à notre connaissance, les résultats théoriques de convergence de l'EKF [97], [19], [12], [74] ne s'appliquent au problème de la trajectographie passive. En effet, pour établir la convergence de l'EKF, on requiert des conditions d'injectivité sur la fonction d'observation, conditions nullement satisfaites pour le cas du BOT.

Mais on rencontre également des situations plus instables où l'EKF ne fait pas ce pour quoi il est conçu. Pour la simulation présentée en figure 12.2, les conditions sont identiques à la simulation précédente mais la cible suit un MRU dans une direction opposée (on repère sa position à l'instant final par un cercle). On constate que l'EKF ne parvient pas à converger vers la trajectoire réelle alors que l'estimation MRU (BLSE) fournit un résultat satisfaisant.

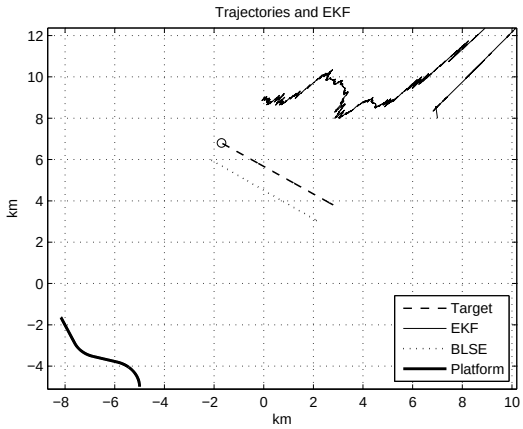


Fig. 12.2 – EKF sans bruit d'état, cible en MRU

Précisons toutefois que ce comportement défavorable est plus difficilement observable que celui décrit précédemment.

12.4 EKF avec bruit d'état

On va s'intéresser désormais au comportement de l'EKF dans des situations où l'estimation paramétrique présente un comportement non-robuste sur des trajectoires cibles qui ne sont pas conformes au modèle paramétrique supposé[1]. Si le filtre EKF est utilisé sans bruit d'état, on observe la plupart du temps le comportement attendu, à savoir que le filtre tend vers l'estimation MRU. En revanche, si le filtre est utilisé avec un bruit d'état, alors on constate que dans certains cas, le filtre présente un comportement plus robuste que la méthode batch.

Pour les simulations qui suivent, la durée d'observation est de $20\,\mathrm{s}$ et la période d'échantillonnage de $0.01\,\mathrm{s}$. Le porteur est à vitesse constante de $250\,\mathrm{m/s}$ et effectue des manœuvres de l'ordre de $50\,\mathrm{m/s^2}$. La cible est également à vitesse constante de $270\,\mathrm{m/s}$ et effectue, selon les cas, des manœuvres du même ordre que le porteur. Le bruit d'observation est celui des observations réellement acquises à savoir $\sigma = 5 \times 10^{-3}\,\mathrm{rad}$. Pour chaque simulation, on présente les résultats obtenus sans bruit d'état puis avec un bruit d'état de $1\,\mathrm{m/s^2}$.

Sur les figures 12.3 et 12.4, la cible suit une trajectoire constituée d'une section MCU puis d'une section MRU.

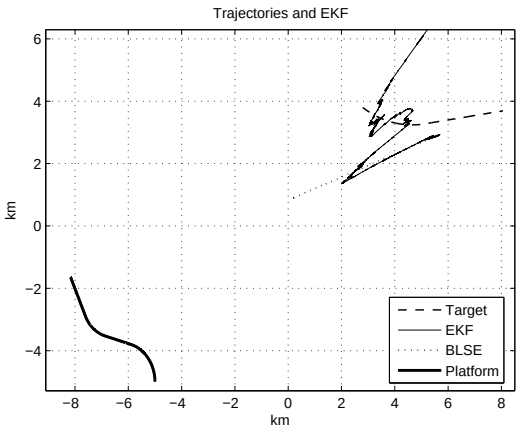


FIG. 12.3 – EKF sans bruit d'état

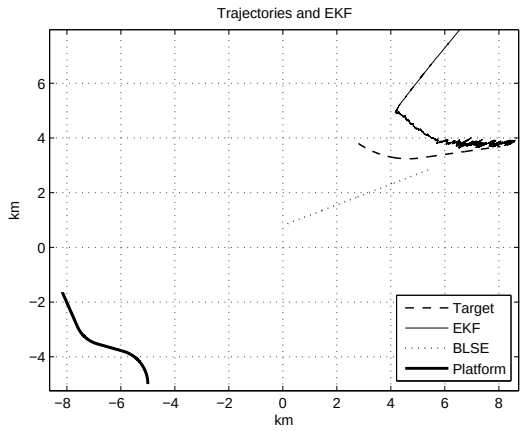


FIG. 12.4 – EKF avec bruit d'état

Sans bruit d'état, l'EKF se comporte comme l'estimation mais avec bruit d'état, il parvient à accrocher la trajectoire réelle.

Sur les figures 12.5 et 12.6, la trajectoire de la cible est polynomiale de degré 2, autrement dit la trajectoire est MRU avec une perturbation quadratique.

[1] Voir chapitre 10 pour l'étude détaillée de ce phénomène.

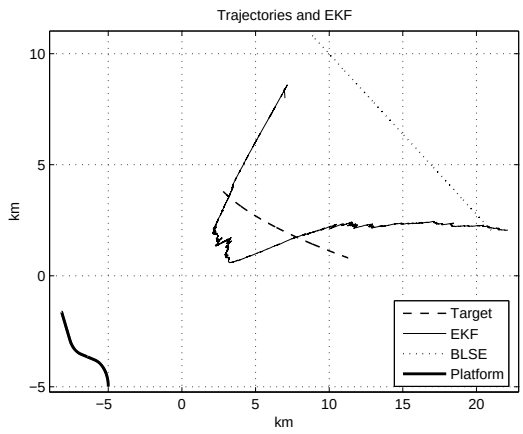


Fig. 12.5 – EKF sans bruit d'état

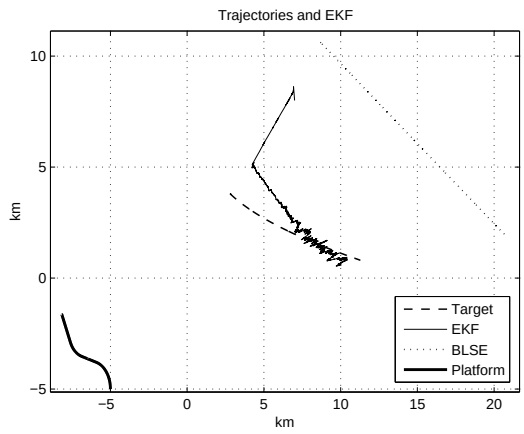


Fig. 12.6 – EKF avec bruit d'état

Le comportement de l'EKF est à peu près identique à celui de la simulation précédente.

Sur les figures 12.7 et 12.8, la trajectoire de la cible est un MCU de courbure faible.

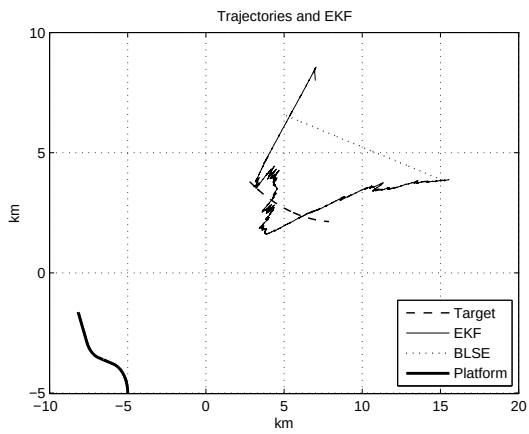


Fig. 12.7 – EKF sans bruit d'état

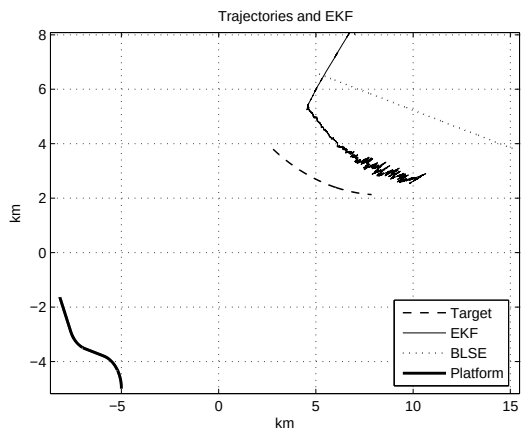


Fig. 12.8 – EKF avec bruit d'état

Pour cette simulation, l'EKF avec bruit d'état présente un meilleur comportement que l'estimation sans toutefois accrocher complètement la trajectoire réelle.

Cependant, l'injection d'un bruit d'état dans le filtre dégrade son comportement dans des situations où la dynamique réelle de la cible est conforme à celle supposée.

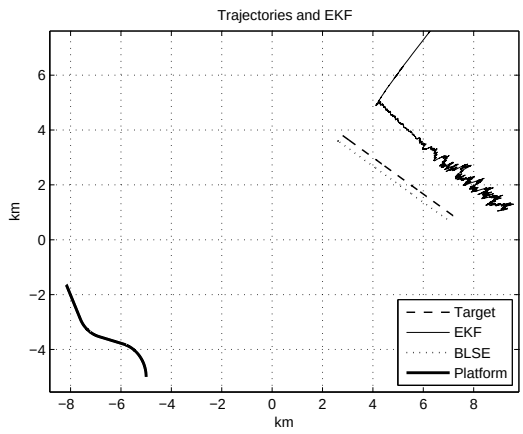


Fig. 12.9 – EKF avec bruit d'état, cible en MRU

Pour la simulation présentée en figure 12.9, la cible suit une trajectoire MRU qui a les mêmes caractéristiques que celles présentées à la figure 12.1. On constate que le filtre EKF n'est pas erratique mais il ne parvient pas à faire aussi bien que pour la simulation sans bruit d'état.

12.5 Conclusion

Le filtre EKF avec bruit d'état présente des caractéristiques intéressantes dans des situations où les méthodes batch mais aussi particulaires[2] ne permettent pas un pistage de la cible satisfaisant. L'injection du bruit d'état permet au filtre de s'affranchir, dans une certaine mesure, de la dynamique supposée et d'accorder une confiance plus forte aux observations. Le bruit d'état donne de la souplesse au filtre et lui confère ce comportement inattendu et favorable. Il faut toutefois rester prudent car le phénomène n'est pour l'instant pas compris et n'est pas observé dans toutes les situations. De plus, le bruit d'état peut, dans les cas d'estimations robustes par exemple, dégrader les résultats obtenus. Une étude plus approfondie menée par expérimentations numériques pourraient permettre de cataloguer des situations pour lesquelles on peut prédire le comportement de l'EKF.

[2]Voir section 10.4.

Chapitre 13

Stabilité asymptotique

Sommaire

13.1 Introduction and notations **156**
13.2 Main results **159**
13.3 Nonlinear state-space models **166**
13.3.1 Nonlinear state-space model with i.i.d. state noise 167
13.3.2 Nonlinear state-space model with dependent state noise 168
13.4 Proofs of Propositions 13.4 and 13.5 . . . . . **170**
13.5 Proofs of Propositions 13.6 and 13.7 . . . . . **173**

Le chapitre qui suit est le contenu de l'article intitulé *Forgetting of the initial distribution for non-ergodic Hidden Markov Chains*, co-écrit avec Éric Moulines, Randal Douc et Élisabeth Gassiat et soumis en septembre 2008.

Forgetting of the initial distribution for non-ergodic Hidden Markov Chains

ABSTRACT. In this chapter, the forgetting of the initial distribution for a non-ergodic Hidden Markov Models (HMM) is studied. A new set of conditions is proposed to establish the forgetting property of the filter, which significantly extends all the existing results. Both a pathwise and mean convergence of the total variation distance of the filter started from two different initial distributions are considered. The results are illustrated using a generic non-ergodic state-space models for which both pathwise and mean exponential stability is established.

Key words and phrases :Non-linear filtering, forgetting of the initial distribution, non-ergodic Hidden Markov Chains, Feynman-Kac semigroup

13.1 Introduction and notations

There are many applications where the current state of a dynamical system need to be estimated from observations up to the current time. In this chapter, it is assumed that the underlying state process $\{X_k\}_{k\geq 0}$ (often referred to as the *signal* process) is a general state space discrete time Markov chain and the *observation process* $\{Y_k\}_{k\geq 0}$ is independent conditionally to the state sequence. More specifically, let X and Y be separable Polish spaces endowed with their Borel σ-fields $\mathcal{X}$ and $\mathcal{Y}$. We denote by Q the transition kernel on $(\mathsf{X}, \mathcal{X})$, μ a measure on $(\mathsf{Y}, \mathcal{Y})$ and a transition density g from $(\mathsf{X}, \mathcal{X})$ to $(\mathsf{Y}, \mathcal{Y})$. Consider the Markov transition kernel defined for any $(x, y) \in \mathsf{X} \times \mathsf{Y}$ and $C \in \mathcal{X} \otimes \mathcal{Y}$ by

$$T\left[(x,y), C\right] \stackrel{\text{def}}{=} \iint Q(x, \mathrm{d}x')\, g(x', y') \mathbb{1}_C(x', y') \mu(\mathrm{d}y') \ . \tag{1}$$

We consider $\{X_k, Y_k\}_{k\geq 1}$ the Markov chain with transition kernel T and initial distribution $C \mapsto \iint g(x,y)\mathbb{1}_C(x,y)\nu(dx)\mu(\mathrm{d}y)$, where ν is a probability measure on $(\mathsf{X}, \mathcal{X})$. With a slight abuse in the terminology, ν is referred to as the initial distribution of $\{X_k, Y_k\}_{k\geq 1}$ and we denote by $\mathbb{P}_\nu$ the distribution of this process over a suitably defined measurable space $(\Omega, \mathcal{F})$. We assume that the chain $\{X_k\}_{k\geq 0}$ is not observed. The distribution of the hidden state X_n conditionally on the observations $Y_{0:n} \stackrel{\text{def}}{=} [Y_0, \ldots, Y_n]$, denoted $\phi_{\nu,n}[Y_{0:n}]$, is referred to as the *filtering* distribution.

A typical question consists in finding conditions under which the filtering distribution is *stable*, *i.e.* that an appropriately chosen distance between the filtering distributions $\phi_{\nu,n}[Y_{0:n}]$ and $\phi_{\nu',n}[Y_{0:n}]$ for two different choices of the initial distribution ν and ν' vanishes as n goes to infinity. In this chapter, assuming that $\{Y_k\}_{k\geq 0}$ is a Y-valued stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P}_\star)$, our objective is to establish either pathwise or mean filter stability in the total variation distance

$$\limsup_{n\to\infty} \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} = 0 \ , \quad \mathbb{P}_\star - a.s. \ , \tag{2}$$

$$\limsup_{n\to\infty} \mathbb{E}_\star \left[\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right] = 0 \ , \tag{3}$$

where $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm. In contrast with most contributions on this subject, $\mathbb{P}_\star$ need not be equal to $\mathbb{P}_\nu$ which means that our results apply even if the filtering model is misspecified. Under more stringent conditions, we may strengthen (2) or (3) by specifying rates of convergence. Of particular importance are the exponential rates (or exponential stability), which

amounts to requiring that

$$\limsup_{n\to\infty} n^{-1} \log \left(\left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} \right) < 0 \ , \quad \mathbb{P}_\star - a.s. \ , \tag{4}$$

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{E}_\star \left[\left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} \right] < 0 \ . \tag{5}$$

As stressed by [32], the most important motivation for studying the stability of the filter is the time-uniform convergence of estimators of the filtering distribution. Since these estimators are most often defined recursively, the approximation error at a given time instant has an impact at all subsequent time instants. As shown in [35], the propagation of error can be considered as an incorrect initialization at the time when the error was made. If the rate of convergence of the filter is fast enough (e.g. if the filter is exponentially stable), then the effects of these local errors do not build up. Another important application of the stability is for the inference of the transition kernel Q or the likelihood g, when these quantities belong either to parametric or non-parametric family of distributions. As shown in [41], the convergence of the likelihood of the observation and the consistency of the maximum likelihood estimator rely on the stability of the filter for a mis-specified model of the observations (several examples of this type will be given later).

The stability of the filter in nonlinear state space models has attracted many research efforts; see for example the in-depth tutorial of [28]. The brief overview below is mainly intended to allow comparison of assumptions and results presented in this contribution with respect to those previously reported in the literature.

The filtering equation can be seen as a positive random non-linear operator acting on the space of probability measures; the stability can be investigated using tools from the theory of positive operators, namely the Birkhoff contraction inequality for the Hilbert projective metric (see [8], [78], [79]). The results obtained using this approach require strong *mixing* conditions for the transition kernels : there exist positive constants ϵ_- and ϵ_+ and a probability measure λ on $(\mathsf{X}, \mathcal{X})$ such that

$$\epsilon_- \lambda(f) \leq Q(x, f) \leq \epsilon_+ \lambda(f) \ , \quad \text{for any } x \in \mathsf{X}, A \in \mathcal{X}. \tag{6}$$

This condition in particular implies that the chain is uniformly geometrically ergodic. Under weak additional assumptions on the likelihood, (6) allows to establish both pathwise and mean exponential stability of the filter, with bounds which are uniform with respect to the observations $Y_{0:n}$.

In [78], the stability of the optimal filter is studied for a class of kernels referred to as *pseudo-mixing*. The definition of pseudo-mixing kernel is adapted to the case where the state space is $\mathsf{X} = \mathbb{R}^d$, equipped with the Borel sigma-field $\mathcal{X}$. A kernel Q on $(\mathsf{X}, \mathcal{X})$ is *pseudo-mixing* if for any compact set C with a diameter d large enough, there exist positive constants $\epsilon_-(d) > 0$ and $\epsilon_+(d) > 0$ and a measure λ_{C} (which may be chosen to be finite without loss of generality) such that

$$\epsilon_-(d) \lambda_{\mathsf{C}}(A) \leq Q(x, A) \leq \epsilon_+(d) \lambda_{\mathsf{C}}(A) \ , \quad \text{for any } x \in \mathsf{C}, \ A \in \mathcal{X}. \tag{7}$$

This condition is more general than (6), but still it is not satisfied in the linear Gaussian case (see [78, Example 4.3]).

A significant improvement has been achieved by [73], who considered the filtering problem of a signal $\{X_k\}_{k\geq 0}$ taking values in $\mathsf{X} = \mathbb{R}^d$ filtered from observations $\{Y_k\}_{k\geq 0}$ in $\mathsf{Y} = \mathbb{R}^\ell$,

$$X_{k+1} = f(X_k) + \sigma(X_k)\zeta_k \ , \tag{8}$$

$$Y_k = h(X_k) + \beta \varepsilon_k \ . \tag{9}$$

Here $\{(\zeta_k, \varepsilon_k)\}_{k\geq 0}$ is a i.i.d. sequence of random vectors in $\mathbb{R}^{d+\ell}$ with density $q_\zeta(x)q_\varepsilon(y)$, $f(\cdot)$ is a d-dimensional vector function, $\sigma(\cdot)$ a $d \times d$-matrix function, $h(\cdot)$ is a ℓ-dimensional vector-function and $\beta > 0$. The authors established both pathwise (2) and mean (3) stability of the filter under appropriate conditions on the functions f, h and σ and on the signal and measurement noise $\{(\zeta_k, \varepsilon_k)\}_{k\geq 0}$. These conditions cover (with some restrictions) the linear gaussian state space model. Note however that these results hold only if $\mathbb{P}_\star = \mathbb{P}_\nu$ and $\nu' \ll \nu$. These results were later extended in [40]. Both pathwise and mean stability are established for initial distributions ν and ν' that are not necessarily comparable and a distribution $\mathbb{P}_\star$ which might be different from $\mathbb{P}_\nu$. The results hold under weaker conditions than those mentioned above; in particular, these results cover the linear Gaussian state-space model without restriction on the measurement and noise variance.

The works mentioned above mainly are obtained under the assumption that the signal process is ergodic. Results for non-ergodic signals in the linear Gaussian case have been obtained in [92]. Non-linear non-ergodic state-space models have been considered much less frequently in the literature. These extensions are important because many models in engineering or econometrics are non-ergodic (see [43] and [101] and the references therein). In [24], the model (8)-(9) is considered : f is assumed to be Lipshitz and $h(x) = x$, for all $x \in \mathsf{X} = \mathbb{R}^d$. The pathwise exponential stability is established [24, Theorem 2.1] under the assumptions that the state and the observation noise are both gaussian and that the variance of the observation noise is small enough. More general distributions for the state and the observation noises are considered in [24, Theorem 2.4], but still the exponential stability is obtained only under the condition that the scale of the observation noise is small enough.

These results were later extended in [93] to allow more general functions h. The authors establish stability in the mean of the filter, under conditions essentially stating that the tails of the observation noise $\{\varepsilon_k\}_{k\geq 0}$ are sufficiently light compared to the tails of the signal noise $\{\zeta_k\}_{k\geq 0}$. These results are derived under the additional assumption that the two initial conditions ν and ν' are comparable (*i.e.* $\nu \ll \nu'$ and $\nu' \ll \nu$) and that the distribution of the observation $\mathbb{P}_\star = \mathbb{P}_\nu$. Similar conditions have been studied in [32], which established the pathwise stability, again under $\mathbb{P}_\nu$. The conditions in these two publications are not equivalent; in particular [32] assume that $\sigma \equiv 1$ in (8) and that the signal and observation noises are i.i.d. whereas [93] allow a form of weak dependence in the signal noise (see Section 13.4 for further discussion).

In a related work, [60] have considered the stability of the filter for denumerable Markov chains. In this work, the observation equation (9) holds with $\sigma \equiv 1$ and $\{X_k\}_{k\geq 0}$ is a finite or denumerable Markov chain. The authors establish exponential pathwise stability when the noise variance is sufficiently small and h is one-to-one. Here again, $\nu \ll \nu'$ and the distribution of the observation process is $\mathbb{P}_\star = \mathbb{P}_\nu$.

A significant weakening of these assumptions has been achieved in [111] and [112]. These contributions establish the stability of the filter (in bounded Lipshitz norm) for an observation model (9) under the conditions that h possesses a uniformly continuous inverse and the noise $\{\varepsilon_k\}_{k\geq 0}$ has a density with respect to the Lebesgue measure whose Fourier transform vanishes nowhere but without imposing any assumption on the transition kernel Q of the signal. Stability in total variation distance can be obtained under the uniform strong Feller assumption, *i.e.* that $x \mapsto Q(x, \cdot)$ is uniformly continuous for the total variation distance on the space of probability measures. The pathwise and the mean filter stability are obtained under the conditions that the initial distributions of the process ν and ν' satisfy $\nu \ll \nu'$ and the distribution of the observation process is $\mathbb{P}_\star = \mathbb{P}_\nu$.

In this contribution, we propose a new set of conditions to establish pathwise and mean filter stability under possible model mis-specification. We assume an observation model that can

be more general than (9) and do not assume that $\nu \ll \nu'$; in addition, the distribution of the observation process $\mathbb{P}_\star$ is not constrained to be $\mathbb{P}_\nu$ and may, on the contrary, be fairly general. Compared to the very weak conditions introduced in [111] and [112], the price to pay are stronger conditions on the transition kernel Q, which are reminiscent from the Local Doeblin condition introduced in [40].

The chapter is organized as follows. In section 13.2, the assumptions are introduced and the main results are stated. In Theorem 13.1, the pathwise stability of the filter (2) is established and an explicit bound of the deviation is given. In Theorem 13.2, the average stability of the filter (3) is established together with a computable bound. In section 13.3, different nonlinear state-space models are considered. For these models, we provide conditions upon which the exponential pathwise and mean stability hold. Several technical Lemmas required to study the examples are given in Sections 13.4 and 13.5.

13.2 Main results

Our results require the existence of a set-valued function, referred to as *Local Doeblin (LD) set function*, which extends the so-called LD-sets introduced in [115] and later exploited in [73]. The difference between LD-sets of [115] and LD-set functions lies in the dependence on the successive observations.

Definition 13.1 (LD-set function). *A set-valued function* $\mathsf{C} : y \longmapsto \mathsf{C}(y)$ *from* Y *to* $\mathcal{X}$ *is called a* Local Doeblin set function (LD-set function) *if there exist a measurable function* $(y, y') \longmapsto \left(\varepsilon^-_{\mathsf{C}}(y, y'), \varepsilon^+_{\mathsf{C}}(y, y')\right)$ *from* $\mathsf{Y} \times \mathsf{Y}$ *to* $(0, \infty)^2$ *and a transition kernel* $\lambda : \mathsf{Y} \times \mathsf{Y} \times \mathcal{X} \to [0, 1]$ [1] *such that, for all* $x \in \mathsf{C}(y)$ *and* $A \in \mathcal{X}$,

$$\varepsilon^-_{\mathsf{C}}(y, y')\lambda\left(y, y'; A \cap \mathsf{C}(y')\right) \leq Q[x, A \cap \mathsf{C}(y')] \leq \varepsilon^+_{\mathsf{C}}(y, y')\lambda\left(y, y'; A \cap \mathsf{C}(y')\right) . \tag{10}$$

Consider the following assumptions on the likelihood of the observations :

H 1. g is continuous and positive.

This excludes the case of additive noise with bounded support; see for example [23]. Stability of the filter may hold in such context, but the fact that the likelihood might vanish creates additional technical difficulties which will obscure the main points of the paper. In particular, under this assumption, for any distribution ν on $(\mathsf{X}, \mathcal{X})$, $n \geq 0$ and sequence $y_{0:n} \in \mathsf{Y}^{n+1}$,

$$\mathbb{E}^Q_\nu\left[\prod_{k=0}^n g(X_k, y_k)\right] \stackrel{\text{def}}{=} \int \cdots \int \nu(\mathrm{d}x_0) \prod_{k=1}^n Q(x_{k-1}, \mathrm{d}x_k) \prod_{k=0}^n g(x_k, y_k) > 0 . \tag{11}$$

The filtering distribution can thus be expressed, for $A \in \mathcal{X}$,

$$\phi_{\nu,n}[y_{0:n}] \stackrel{\text{def}}{=} \frac{\int \cdots \int \nu(\mathrm{d}x_0) \prod_{k=1}^n Q(x_{k-1}, \mathrm{d}x_k) \prod_{k=0}^n g(x_k, y_k) \mathbb{1}_A(x_n)}{\int \cdots \int \nu(dx_0) \prod_{k=1}^n Q(x_{k-1}, \mathrm{d}x_k) \prod_{k=0}^n g(x_k, y_k)} . \tag{12}$$

The continuity can also be relaxed, also at the expense of some minor technical adaptations. The main idea of the proof is that the states belong very often to the LD-sets. Every time the state is in a LD set and jumps to another LD set, the forgetting mechanism comes into play. From now on, for all $(x, x') \in \mathsf{X}^2$, denote by $\bar{x} = (x, x')$ and by $\bar{g}$ the product $\bar{g}(\bar{x}, y) = g(x, y)g(x', y)$.

[1] for any $(y, y') \in \mathsf{Y} \to \mathsf{Y}$, $\lambda(y, y'; \cdot)$ is a σ-finite measure on $(\mathsf{X}, \mathcal{X})$ and for any $A \in \mathcal{X}$, the function $(y, y') \mapsto \lambda(y, y'; A)$ is measurable from $(\mathsf{Y} \times \mathsf{Y}, \mathcal{Y} \otimes \mathcal{Y})$ for $[0, 1]$ equipped with its Borel σ-field

Similarly, for all $A \in \mathcal{X}$, denote $\bar{A} = A \times A$, for all LD-set function C, $\bar{\mathsf{C}}$ the set-valued function $\bar{\mathsf{C}}(y) = \mathsf{C}(y) \times \mathsf{C}(y)$. For all $(x, x') \in \mathsf{X}^2$, and A, $B \in \mathcal{X}$, set $\bar{Q}(x, x', A \times B) = Q(x, A)Q(x', B)$. Finally, for ν,ν' two probability distributions on $(\mathsf{X}, \mathcal{X})$, we denote by $\mathbb{E}_\nu^Q$ and $\mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}$ the expectation with respect to the distribution of a Markov chain on X (resp. on $\mathsf{X} \times \mathsf{X}$) with initial distribution ν (resp. $\nu \otimes \nu'$) and transition kernel Q (resp. $\bar{Q}$). Then, under the stated assumptions, for any $A \in \mathcal{X}$, any ν and ν' two probability distributions on $(\mathsf{X}, \mathcal{X})$, any integer n and any sequence $y_{0:n} \in \mathsf{Y}^{n+1}$, the difference $\phi_{\nu,n}[y_{0:n}](A) - \phi_{\nu',n}[y_{0:n}](A)$ may be expressed as

$$\begin{aligned}
&\phi_{\nu,n}[y_{0:n}](A) - \phi_{\nu',n}[y_{0:n}](A) \\
&\quad = \frac{\mathbb{E}_\nu^Q\left[\prod_{i=0}^n g(X_i, y_i)\mathbb{1}_A(X_n)\right]}{\mathbb{E}_\nu^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]} - \frac{\mathbb{E}_{\nu'}^Q\left[\prod_{i=0}^n g(X_i, y_i)\mathbb{1}_A(X_n)\right]}{\mathbb{E}_{\nu'}^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]} , \\
&\quad = \frac{\mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i)\mathbb{1}_A(X_n)\right] - \mathbb{E}_{\nu'\otimes\nu}^{\bar{Q}}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i)\mathbb{1}_A(X_n)\right]}{\mathbb{E}_\nu^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]\mathbb{E}_{\nu'}^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]} , \\
&\quad = \frac{\mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i)\{\mathbb{1}_A(X_n) - \mathbb{1}_A(X'_n)\}\right]}{\mathbb{E}_\nu^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]\mathbb{E}_{\nu'}^Q\left[\prod_{i=0}^n g(X_i, y_i)\right]} .
\end{aligned} \tag{13}$$

We compute bounds for the numerator and the denominator of the previous expression. Such bounds are given in the two following Propositions. For an LD-set function C denote :

$$\rho_{\mathsf{C}}(y, y') \stackrel{\text{def}}{=} 1 - (\varepsilon_{\mathsf{C}}^- / \varepsilon_{\mathsf{C}}^+)^2(y, y') . \tag{14}$$

For any integer n and any sequence $\{y_i\}_{i=0}^n$ in Y, let us define

$$\Delta_n(\nu, \nu', y_{0:n}) = \sup_{A\in\mathcal{X}} \left| \mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i)\mathbb{1}_A(X_n)\right] - \mathbb{E}_{\nu'\otimes\nu}^{\bar{Q}}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i)\mathbb{1}_A(X_n)\right] \right| . \tag{15}$$

Proposition 13.1. *Let C be an LD-set function and ν and ν' be two probability measures on $(\mathsf{X}, \mathcal{X})$. Then,*

$$\Delta_n(\nu, \nu', y_{0:n}) \le \mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}\left\{\bar{g}(\bar{X}_0, y_0)\prod_{i=1}^n \bar{g}(\bar{X}_i, y_i)\rho_{\mathsf{C}}^{\delta_i}(y_{i-1}, y_i)\right\} ,$$

where $\delta_i = \mathbb{1}_{\bar{\mathsf{C}}(y_{i-1})\times\bar{\mathsf{C}}(y_i)}(\bar{X}_{i-1}, \bar{X}_i)$.

Proof. For convenience, we write $\mathsf{C}_i = \mathsf{C}(y_i)$, $\varepsilon_i^- = \varepsilon_{\mathsf{C}}^-(y_{i-1}, y_i)$, $\varepsilon_i^+ = \varepsilon_{\mathsf{C}}^+(y_{i-1}, y_i)$, $g_i(x) = g(x, y_i)$, $\lambda_i(\cdot) = \lambda(y_{i-1}, y_i; \cdot)$ and $\rho_i = 1 - (\varepsilon_i^- / \varepsilon_i^+)^2$. Let us define $\bar{\lambda}_i \stackrel{\text{def}}{=} \lambda_i \otimes \lambda_i$. Since C is an LD-set function, for all $i = 1, \ldots, n$, $\bar{x} \in \bar{\mathsf{C}}_{i-1}$, and $\bar{f}$ a non-negative function on $\mathsf{X} \times \mathsf{X}$,

$$(\varepsilon_i^-)^2\bar{\lambda}_i(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) \le \bar{Q}(\bar{x}, \mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) \le (\varepsilon_i^+)^2\bar{\lambda}_i(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) . \tag{16}$$

Define the sequence of unnormalized kernels $\bar{Q}_i^0$ and $\bar{Q}_i^1$ as follows : for all $\bar{x} \in \mathsf{X}^2$, and $\bar{f}$ a non-negative function on $\mathsf{X} \times \mathsf{X}$,

$$\begin{aligned}
\bar{Q}_i^0(\bar{x}, \bar{f}) &= (\varepsilon_i^-)^2\mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{\lambda}_i(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) , \\
\bar{Q}_i^1(\bar{x}, \bar{f}) &= \bar{Q}(\bar{x}, \bar{f}) - (\varepsilon_i^-)^2\mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{\lambda}_i(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) .
\end{aligned}$$

It follows from (16) that, for all $\bar{x}$ in $\bar{\mathsf{C}}_{i-1}$, $0 \le \bar{Q}_i^1(\bar{x}, \mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) \le \rho_i\bar{Q}(\bar{x}, \mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f})$ which implies that, for

all $\bar{x} \in \mathsf{X}^2$,

$$\begin{aligned}\bar{Q}_i^1(\bar{x},\bar{f}) &= \mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{Q}_i^1(\bar{x},\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) + \mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{Q}_i^1(\bar{x},\mathbb{1}_{\bar{\mathsf{C}}_i^c}\bar{f}) + \mathbb{1}_{\bar{\mathsf{C}}_{i-1}^c}(\bar{x})\bar{Q}_i^1(\bar{x},\bar{f}) \ , \\ &\leq \rho_i \mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{Q}(\bar{x},\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{f}) + \mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\bar{Q}_i^1(\bar{x},\mathbb{1}_{\bar{\mathsf{C}}_i^c}\bar{f}) + \mathbb{1}_{\bar{\mathsf{C}}_{i-1}^c}(\bar{x})\bar{Q}_i^1(\bar{x},\bar{f}) \ , \\ &\leq \bar{Q}\left(\bar{x}, \rho_i^{\mathbb{1}_{\bar{\mathsf{C}}_{i-1}}(\bar{x})\mathbb{1}_{\bar{\mathsf{C}}_i}}\bar{f}\right) \ .\end{aligned} \tag{17}$$

We write $\Delta_n\big(\nu,\nu',y_{0:n}\big) = \sup_{A\in\mathcal{X}}|\Delta_n(A)|$, where

$$\Delta_n(A) \stackrel{\text{def}}{=} \nu\otimes\nu'\big(\bar{g}_0\bar{Q}\bar{g}_1\ldots\bar{Q}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) - \nu'\otimes\nu\big(\bar{g}_0\bar{Q}\bar{g}_1\ldots\bar{Q}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) \ .$$

We decompose $\Delta_n(A)$ into $\Delta_n(A) = \sum_{t_{1:n}\in\{0,1\}^n}\Delta_n(A,t_{1:n})$, where

$$\Delta_n(A,t_{1:n}) \stackrel{\text{def}}{=} \nu\otimes\nu'\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) - \nu'\otimes\nu\big(\bar{g}_1\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) \ .$$

Note that, for any $t_{1:n}\in\{0,1\}^n$ and any sets $A,B\in\mathcal{X}$,

$$\nu\otimes\nu'\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times B}\big) = \nu'\otimes\nu\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{B\times A}\big) \ .$$

If there is an index $i\in\{1,\ldots,n\}$ such that $t_i=0$, then

$$\begin{aligned}&\nu\otimes\nu'\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) \\ &\quad= \nu\otimes\nu'\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_{i-1}^{t_{i-1}}\bar{g}_{i-1}\mathbb{1}_{\bar{\mathsf{C}}_{i-1}}\big)\times(\varepsilon_i^-)^2\bar{\lambda}_i\big(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{g}_i\bar{Q}_{i+1}^{t_{i+1}}\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) \ , \\ &\quad= \nu'\otimes\nu\big(\bar{g}_0\bar{Q}_1^{t_1}\bar{g}_1\ldots\bar{Q}_{i-1}^{t_{i-1}}\bar{g}_{i-1}\mathbb{1}_{\bar{\mathsf{C}}_{i-1}}\big)\times(\varepsilon_i^-)^2\bar{\lambda}_i\big(\mathbb{1}_{\bar{\mathsf{C}}_i}\bar{g}_i\bar{Q}_{i+1}^{t_{i+1}}\ldots\bar{Q}_n^{t_n}\bar{g}_n\mathbb{1}_{A\times\mathsf{X}}\big) \ .\end{aligned}$$

Thus, $\Delta_n(A,t_{1:n})=0$ except if for all $i\in\{1,\ldots,n\}$, $t_i=1$, and we obtain

$$\Delta_n(A) = \nu\otimes\nu'\Big[\bar{g}_0\bar{Q}_1^1\bar{g}_1\ldots\bar{Q}_n^1\bar{g}_n\big(\mathbb{1}_{A\times\mathsf{X}} - \mathbb{1}_{\mathsf{X}\times A}\big)\Big] \ .$$

It then follows using (17)

$$\Delta_n\big(\nu,\nu',y_{0:n}\big) \leq \nu\otimes\nu'(\bar{g}_0\bar{Q}_1^1\bar{g}_1\ldots\bar{Q}_n^1\bar{g}_n) \leq \mathbb{E}_{\nu\otimes\nu'}^{\bar{Q}}\left[\bar{g}(\bar{X}_0,y_0)\prod_{i=1}^n\bar{g}(\bar{X}_i,y_i)\rho_i^{\delta_i}\right] \ ,$$

with $\delta_i = \mathbb{1}_{\bar{\mathsf{C}}_{i-1}\times\bar{\mathsf{C}}_i}(\bar{X}_{i-1},\bar{X}_i)$. □

We now compute a bound for the denominator. For a given LD-set function C , we set

$$\Phi_{\nu,\mathsf{C}}(y,y') \stackrel{\text{def}}{=} \mathbb{E}_\nu^Q\left[g(\cdot,y_0)g(\cdot,y')\mathbb{1}_{\mathsf{C}(y_1)}(\cdot)\right] = \nu\left[g(\cdot,y)Qg(\cdot,y')\mathbb{1}_{\mathsf{C}(y')}(\cdot)\right] \ , \tag{18}$$

$$\Psi_{\mathsf{C}}(y,y') \stackrel{\text{def}}{=} \lambda\left(y,y';g(\cdot,y')\mathbb{1}_{\mathsf{C}(y')}\right) \ . \tag{19}$$

Proposition 13.2. *Let C be an LD-set function and $\{y_i\}_{i=0}^n$ a sequence in Y. We have for all $n\in\mathbb{N}$*

$$\mathbb{E}_\nu^Q\left[\prod_{i=0}^n g(X_i,y_i)\right] \geq \Phi_{\nu,\mathsf{C}}(y_0,y_1)\prod_{i=2}^n\left(\varepsilon_{\mathsf{C}}^-(y_{i-1},y_i)\Psi_{\mathsf{C}}(y_{i-1},y_i)\right) \ . \tag{20}$$

Proof of Proposition 13.2. Since C is an LD-set function, there exist some applications $\varepsilon_{\mathsf{C}}^-$, $\varepsilon_{\mathsf{C}}^+$ such

that, for all $i = 1, \ldots, n$, for all $x \in \mathsf{C}(y_{i-1})$ and for all $A \in \mathcal{X}$ with $A \subset \mathsf{C}(y_i)$,

$$\varepsilon_{\mathsf{C}}^{-}(y_{i-1}, y_i)\lambda(y_{i-1}, y_i; A) \leq Q(x, A) \leq \varepsilon_{\mathsf{C}}^{+}(y_{i-1}, y_i)\lambda(y_{i-1}, y_i; A) \ . \tag{21}$$

Obviously,

$$\mathbb{E}_{\nu}^{Q}\left[\prod_{i=0}^{n} g(X_i, y_i)\right] \geq \mathbb{E}_{\nu}^{Q}\left[g(X_0, y_0)\prod_{i=1}^{n} g(X_i, y_i)\mathbb{1}_{\mathsf{C}(y_i)}(X_i)\right] \ .$$

Then, the right-hand side of this expression may be bounded using (21) by

$$\begin{aligned}
&\mathbb{E}_{\nu}^{Q}\left[g(X_0, y_0)\prod_{i=1}^{n} g(X_i, y_i)\mathbb{1}_{\mathsf{C}(y_i)}(X_i)\right] \\
&\quad = \mathbb{E}_{\nu}^{Q}\left[g(X_0, y_0)g(X_1, y_1)\mathbb{1}_{\mathsf{C}(y_1)}(X_1)\prod_{i=2}^{n} g(X_i, y_i)\mathbb{1}_{\mathsf{C}(y_{i-1})\times\mathsf{C}(y_i)}(X_{i-1}, X_i)\right] \ , \\
&\quad \geq \nu\left[g(\cdot, y_0)Qg(\cdot, y_1)\mathbb{1}_{\mathsf{C}(y_1)}(\cdot)\right]\prod_{i=2}^{n}\varepsilon_{\mathsf{C}}^{-}(y_{i-1}, y_i)\lambda\left(y_{i-1}, y_i; g(\cdot, y_i)\mathbb{1}_{\mathsf{C}(y_i)}\right) \ .
\end{aligned}$$

□

Under (H1), $\Phi_{\nu,\mathsf{C}}(y_0, y_1) > 0$ for any initial distribution ν such that $\nu Q[\mathsf{C}(y_1)] > 0$. In the examples considered in Section 13.3, this condition is satisfied for the choices of Local Doeblin sets by any initial distributions. Following the same lines as above, it is easily seen that the lower bound (20) can be more generally written as

$$\mathbb{E}_{\nu}^{Q}\left[\prod_{i=0}^{n} g(X_i, y_i)\right] \geq \mathbb{E}_{\nu}^{Q}\left[\prod_{i=0}^{k} g(\cdot, y_k)\mathbb{1}_{\mathsf{C}(y_k)}(\cdot)\right] \times \prod_{i=k+1}^{n}\left(\varepsilon_{\mathsf{C}}^{-}(y_{i-1}, y_i)\Psi_{\mathsf{C}}(y_{i-1}, y_i)\right) \ .$$

This lower bound is positive as soon as $\nu Q^k[\mathsf{C}(y_k)] > 0$. The statements of the results below can be directly extended to handle this more general condition.

By combining these two Propositions, we obtain an explicit bound for the total variation distance $\left\|\phi_{\nu,n}[y_{0:n}] - \phi_{\nu',n}[y_{0:n}]\right\|_{\mathrm{TV}}$. For a set $A \in \mathcal{X}$ and an observation $y \in \mathsf{Y}$, the supremum of the likelihood over A is denoted

$$\Upsilon_A(y) \stackrel{\text{def}}{=} \sup_{x\in A} g(x, y) \ . \tag{22}$$

Consider the following assumption :

H 2. For any $\eta \in (0, 1)$, there exists an LD-set function C_η such that $y \mapsto \Upsilon_{\mathsf{C}_\eta^c(y)}(y)$ is measurable and for all $y \in \mathsf{Y}$,

$$\Upsilon_{\mathsf{C}_\eta^c(y)}(y) \leq \eta\Upsilon_{\mathsf{X}}(y) \ . \tag{23}$$

When $\mathsf{X} = \mathbb{R}^d$, this assumption is typically satisfied when, for any given y, the likelihood goes to zero as the state $|x|$ goes to infinity : $\lim_{|x|\to\infty} g(x, y) = 0$. This condition is satisfied in many models of practical interest, and roughly implies that the observation effectively provides information on the state range of values.

It is worthwhile to note that the bound we obtain is valid for any sequence $y_{0:n}$ and any initial distributions ν and ν'. Under assumption (H2), for any $\eta \in (0, 1)$ there exists a LD-set function

C_η satisfying (23). For any $\alpha \in (0,1)$ and a sequence $y_{0:n} = \{y_i\}_{i=0}^n$ in Y, define

$$\Lambda_\eta(y_{0:n}, \alpha) \stackrel{\text{def}}{=} \max\left\{ \prod_{k=1}^n \rho_\eta^{\delta_k}(y_{k-1}, y_k) \ : \{\delta_k\}_{k=1}^n \in \{0,1\}^n \, , \sum_{k=1}^n \delta_k \geq \alpha n \right\} \, , \tag{24}$$

where ρ_η is a shorthand notation for ρ_{C_η} (see (14)).

Proposition 13.3. *Assume (H1)-(H2). Let C be an LD-set function. Let α be some number in $(0,1)$, ν and ν' some probability measures on $(\mathsf{X}, \mathcal{X})$ and $\{y_i\}_{i=0}^n$ a sequence in Y. Then, for any $\eta > 0$,*

$$\begin{aligned} \left\| \phi_{\nu,n}[y_{0:n}] - \phi_{\nu',n}[y_{0:n}] \right\|_{\mathrm{TV}} &\leq 2\Lambda_\eta(y_{0:n}, \alpha) + \\ &2\eta^{a_n} \prod_{i=0}^n \Upsilon_{\mathsf{X}}^2(y_i) \prod_{i=2}^n \left(\varepsilon_{\mathsf{C}}^-(y_{i-1}, y_i) \Psi_{\mathsf{C}}(y_{i-1}, y_i) \right)^{-2} \Phi_{\nu,\mathsf{C}}^{-1}(y_0, y_1) \Phi_{\nu',\mathsf{C}}^{-1}(y_0, y_1) \, , \end{aligned} \tag{25}$$

with $a_n \stackrel{\text{def}}{=} \frac{(1-\alpha)n}{2} - \frac{1}{2}$.

Proof. Eq. (13) implies

$$\left\| \phi_{\nu,n}[y_{0:n}] - \phi_{\nu',n}[y_{0:n}] \right\|_{\mathrm{TV}} = \frac{2\Delta_n(\nu, \nu', y_{0:n})}{\mathbb{E}_\nu^Q \left[\prod_{i=0}^n g(X_i, y_i) \right] \mathbb{E}_{\nu'}^Q \left[\prod_{i=0}^n g(X_i, y_i) \right]} \, ,$$

where $\Delta_n(\nu, \nu', y_{0:n})$ is defined by (15). We stress that we will use two different LD-set functions, C_η for the numerator and C for the denominator. Set

$$N_{\eta,n} \stackrel{\text{def}}{=} \sum_{i=1}^n \mathbb{1}\{(\bar{X}_{i-1}, \bar{X}_i) \in \bar{\mathsf{C}}_\eta(y_{i-1}) \times \bar{\mathsf{C}}_\eta(y_i)\} \, .$$

Using Proposition 13.1, we obtain

$$\begin{aligned} \Delta_n(\nu, \nu', y_{0:n}) \leq \mathbb{E}_{\nu \otimes \nu'}^{\bar{Q}} &\left[\bar{g}(\bar{X}_0, y_0) \prod_{i=1}^n \bar{g}(\bar{X}_i, y_i) \rho_\eta^{\delta_i}(y_{i-1}, y_i) \mathbb{1}\{N_{\eta,n} \geq \alpha n\} \right] \\ &+ \mathbb{E}_{\nu \otimes \nu'}^{\bar{Q}} \left[\bar{g}(\bar{X}_0, y_0) \prod_{i=1}^n \bar{g}(\bar{X}_i, y_i) \rho_\eta^{\delta_i}(y_{i-1}, y_i) \mathbb{1}\{N_{\eta,n} < \alpha n\} \right] \, , \end{aligned} \tag{26}$$

with $\delta_i = \mathbb{1}_{\bar{\mathsf{C}}_\eta(y_{i-1}) \times \bar{\mathsf{C}}_\eta(y_i)}(\bar{X}_{i-1}, \bar{X}_i)$. The first term in the right-hand side expression of (26) satisfies

$$\mathbb{E}_{\nu \otimes \nu'}^{\bar{Q}} \left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i) \prod_{i=1}^n \rho_\eta^{\delta_i}(y_{i-1}, y_i) \mathbb{1}\{N_{\eta,n} \geq \alpha n\} \right] \leq \mathbb{E}_{\nu \otimes \nu'}^{\bar{Q}} \left[\prod_{i=0}^n \bar{g}(\bar{X}_i, y_i) \right] \Lambda_\eta(y_{0:n}, \alpha) \, .$$

Consider now the second term of the rhs of (26). Let $M_{\eta,n} \stackrel{\text{def}}{=} \sum_{i=0}^{n-1} \mathbb{1}_{\bar{\mathsf{C}}_\eta^c(y_i)}(\bar{X}_i)$. For any sequence $\{u_j\}$, such that $u_j \in \{0,1\}$ for $j \in \{0, \ldots, n\}$,

$$n \geq \sum_{i=0}^{n-1} u_i \vee u_{i+1} = \sum_{i=0}^{n-1} (u_i + u_{i+1} - u_i u_{i+1}) \geq 2 \sum_{i=0}^{n-1} u_i - 1 - \sum_{i=0}^{n-1} u_i u_{i+1} \, ,$$

which implies that $\sum_{i=0}^{n-1} u_i \leq (n+1)/2 + (1/2) \sum_{i=1}^n u_{i-1} u_i$. Using this inequality with $u_i =$

$\mathbb{1}\{\bar{X}_i \in \bar{\mathsf{C}}_\eta(y_i)\}$ for $i \in \{0,\dots,n\}$ shows that $N_{\eta,n} < \alpha n$ implies that $M_{\eta,n} \geq a_n$. Then,

$$\mathbb{E}^{\bar{Q}}_{\nu\otimes\nu'}\left[\bar{g}(\bar{X}_0,y_0)\prod_{i=1}^n \bar{g}(\bar{X}_i,y_i)\rho_\eta^{\delta_i}(y_{i-1},y_i)\mathbb{1}\{N_{\eta,n}<\alpha n\}\right] \leq \mathbb{E}^{\bar{Q}}_{\nu\otimes\nu'}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i,y_i)\mathbb{1}\{M_{\eta,n}\geq a_n\}\right] .$$

By splitting this last product, we obtain

$$\prod_{i=0}^n \bar{g}(\bar{X}_i,y_i)\mathbb{1}\{M_{\eta,n}\geq a_n\} = \prod\nolimits_1 \bar{g}(\bar{X}_i,y_i) \times \prod\nolimits_2 \bar{g}(\bar{X}_i,y_i)\mathbb{1}\{M_{\eta,n}\geq a_n\} \leq \eta^{a_n}\prod_{i=0}^n \Upsilon^2_{\mathsf{X}}(y_i) ,$$

where $\prod_1$ is the product over the indices $i \in \{0,\dots,n\}$ such that $\bar{X}_i \in \bar{\mathsf{C}}^{\mathrm{c}}_\eta(y_i)$ and $\prod_2$ is the product on the remaining indices. This implies that

$$\mathbb{E}^{\bar{Q}}_{\nu\otimes\nu'}\left[\prod_{i=0}^n \bar{g}(\bar{X}_i,y_i)\mathbb{1}\{M_{\eta,n}\geq a_n\}\right] \leq \eta^{a_n}\prod_{i=0}^n \Upsilon^2_{\mathsf{X}}(y_i) .$$

By combining the above relations and Proposition 13.2, the result follows. □

The last step consists in finding conditions upon which the bound in the right hand side of (25) is small. This bound depends explicitly on the observations Y's; it is therefore not difficult to state general conditions upon which this quantity is small. Let $\{Y_k\}_{k\geq 0}$ be a stochastic process with probability distribution $\mathbb{P}_\star$ in $(\mathsf{Y},\mathcal{Y})$. We first formulate an almost sure convergence on the total variation distance of the filter initialized with two different probability measures ν and ν' and then later establish a convergence of the expectation. To complete this program, a last condition measurability condition is required :

H 3. For any $\eta > 0$, the functions $(y,y') \mapsto \Phi_{\nu,\mathsf{C}_\eta}(y,y')$ and $(y,y') \mapsto \Psi_{\mathsf{C}_\eta}(y,y')$ are measurable.

Theorem 13.1. *Assume (H1)-(H2)-(H3). Assume moreover that for some η_0, there exists some LD-set function C_{η_0} such that*

$$\limsup_{n\to\infty}\left[-n^{-1}\sum_{k=2}^n \log \varepsilon^-_{\mathsf{C}_{\eta_0}}(Y_{k-1},Y_k)\right] < \infty , \qquad \mathbb{P}_\star - a.s. \tag{27}$$

$$\limsup_{n\to\infty}\left[n^{-1}\sum_{k=0}^n \log \Upsilon_{\mathsf{X}}(Y_k)\right] < \infty , \qquad \mathbb{P}_\star - a.s. \tag{28}$$

$$\limsup_{n\to\infty}\left[-n^{-1}\sum_{k=2}^n \log \Psi_{\mathsf{C}_{\eta_0}}(Y_{k-1},Y_k)\right] < \infty . \qquad \mathbb{P}_\star - a.s. \tag{29}$$

Assume in addition that there exists $\alpha \in (0,1)$ such that for all $\eta > 0$,

$$\limsup_{n\to\infty} n^{-1}\log \Lambda_\eta(Y_{0:n},\alpha) < 0, \qquad \mathbb{P}_\star - a.s. \tag{30}$$

Then, for any initial probability distributions ν and ν' on $(\mathsf{X},\mathcal{X})$ such that

$$\Phi_{\nu,\mathsf{C}_{\eta_0}}(Y_0,Y_1) < \infty \quad \textit{and} \quad \Phi_{\nu',\mathsf{C}_{\eta_0}}(Y_0,Y_1) < \infty , \quad \mathbb{P}_\star\textit{-p.s.}$$

we have

$$\limsup_{n\to\infty} n^{-1}\log \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} < 0 , \quad \mathbb{P}_\star - a.s. .$$

Proof. We apply (25) with $\mathsf{C} = \mathsf{C}_{\eta_0}$. Note that for any positive sequences $\{u_n\}$ and $\{v_n\}$,

$$\limsup_{n\to\infty} n^{-1}\log(u_n+v_n) \le \sup\left(\limsup n^{-1}\log u_n, \limsup n^{-1}\log v_n\right) .$$

Under the stated assumptions, there exists some constant $0 < M < \infty$ such that for any $\eta > 0$,

$$\begin{aligned}\limsup n^{-1}\log \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} \\ \le \sup\left(\limsup_{n\to\infty} n^{-1}\log\Lambda_\eta(Y_{0:n},\alpha), \frac{1-\alpha}{2}\log(\eta) + M\right) , \quad \mathbb{P}_\star - a.s.\end{aligned}$$

The proof is concluded by choosing η small enough so that $\log(\eta)(1-\alpha)/2 + M < 0$. □

Compared to [40, Theorem 1] in the ergodic case , the conditions (27) and (30) are specific to the non-ergodic case, since they involve the functions $\varepsilon^-_{\mathsf{C}}$ and $\varepsilon^+_{\mathsf{C}}$. In the ergodic case, these functions are constant and assumptions (27) and (30) are trivially satisfied.

Theorem 13.2. *Assume (H1)-(H2)-(H3). Then, for any $\eta_0 > 0$, $M_i > 0$, $i = 0,\dots,3$, $\delta > 0$ and $\alpha \in (0,1)$, there exist constants $\eta > 0$ and $\beta \in (0,1)$ such that, for all $n \in \mathbb{N}$,*

$$\mathbb{E}_\star\left[\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right] \le 2\left(\beta^n + r_0(\nu,n) + r_0(\nu',n) + \sum_{i=1}^4 r_i(n)\right) \tag{31}$$

where the sequences in the right-hand side of (31) *are defined by*

$$r_0(\nu,n) \stackrel{\text{def}}{=} \mathbb{P}_\star\left(-\log\Phi_{\nu,\mathsf{C}_{\eta_0}}(Y_0,Y_1) \ge M_0 n\right), \tag{32}$$

$$r_1(n) \stackrel{\text{def}}{=} \mathbb{P}_\star\left(-\sum_{k=2}^n \log\varepsilon^-_{\mathsf{C}_{\eta_0}}(Y_{k-1},Y_k) \ge M_1 n\right), \tag{33}$$

$$r_2(n) \stackrel{\text{def}}{=} \mathbb{P}_\star\left(\sum_{k=0}^n \log\Upsilon_{\mathsf{X}}(Y_k) \ge M_2 n\right), \tag{34}$$

$$r_3(n) \stackrel{\text{def}}{=} \mathbb{P}_\star\left(-\sum_{k=2}^n \log\Psi_{\mathsf{C}_{\eta_0}}(Y_{k-1},Y_k) \ge M_3 n\right), \tag{35}$$

$$r_4(n) \stackrel{\text{def}}{=} \mathbb{P}_\star\left(\log\Lambda_\eta(Y_{0:n},\alpha) \ge -\delta n\right). \tag{36}$$

Démonstration. For any $\alpha \in (0,1)$ and $\gamma \in (0,1)$, we can choose η small enough and such that for all $n \ge 0$, $\eta^{a_n}\mathrm{e}^{2n\sum_{i=0}^3 M_i} \le \gamma^n$ where $a_n = n(1-\alpha)/2 - 1/2$. Denote by Ω_n the event

$$\begin{aligned}\Omega_n = \Big\{ &-\log\Phi_{\nu,\mathsf{C}_{\eta_0}}(Y_0,Y_1) < M_0 n,\ -\log\Phi_{\nu',\mathsf{C}_{\eta_0}}(Y_0,Y_1) < M_0 n, \\ &-\sum_{i=2}^n \log\varepsilon^-_{\mathsf{C}_{\eta_0}}(Y_{i-1},Y_i) < M_1 n, \sum_{i=0}^n \log\Upsilon_{\mathsf{X}}(Y_i) < M_2 n, \\ &-\sum_{i=2}^n \log\Psi_{\mathsf{C}_{\eta_0}}(Y_{i-1},Y_i) < M_3 n,\ \log\Lambda_\eta(Y_{0:n},\alpha) < -\delta n\Big\}.\end{aligned}$$

Under the stated assumptions, $\mathbb{P}_\star(\Omega_n^{\mathrm{c}}) \le r_0(\nu,n) + r_0(\nu',n) + \sum_{i=1}^4 r_i(n)$. On the event Ω_n, we

have

$$\Phi^{-1}_{\nu,\mathsf{C}_{\eta_0}}(Y_0,Y_1)\Phi^{-1}_{\nu',\mathsf{C}_{\eta_0}}(Y_0,Y_1)\prod_{i=2}^{n}\left[\varepsilon^{-}_{\mathsf{C}_{\eta_0}}(Y_{i-1},Y_i)\Psi_{\mathsf{C}_{\eta_0}}(Y_{i-1},Y_i)\right]^{-2}\prod_{i=0}^{n}\Upsilon^2_{\mathsf{X}}(Y_i)\leq \mathrm{e}^{2n\sum_{i=0}^{3}M_i}.$$

Then, by Proposition 13.3, on the event Ω_n, we have $\left\|\phi_{\nu,n}[Y_{0:n}]-\phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\leq 2\beta^n$ where $\beta=\max(\gamma,e^{-\delta})$. Since

$$\mathbb{E}_\star\left[\left\|\phi_{\nu,n}[Y_{0:n}]-\phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right]\leq \mathbb{E}_\star\left[\left\|\phi_{\nu,n}[Y_{0:n}]-\phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\mathbb{1}_{\Omega_n}\right]+2\mathbb{P}_\star(\Omega_n^{\mathrm{c}})\ ,$$

the result follows. □

Theorem 13.2 does not provide directly a rate of convergence. Indeed, only the first term of the right-hand side of equation (31) gives a geometric rate. In Section 13.3, for specific models, explicit bounds of the other terms will be obtained with geometric rates.

13.3 Nonlinear state-space models

Let $\mathsf{X}=\mathbb{R}^{d_\mathsf{X}}$, $\mathsf{Y}=\mathbb{R}^{d_\mathsf{Y}}$, $\mathsf{Z}=\mathbb{R}^{\ell}$ with $d_\mathsf{Y}\leq d_\mathsf{X}$, endowed with the Borel σ-algebra $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. We consider the non-linear state-space model :

$$\begin{cases} X_k=f_*(X_{k-1})+\tau^*(X_{k-1},\zeta_k)\ , \quad X_0\sim\nu_0\\ Y_k=h_*(X_k)+\varepsilon_k\ , \end{cases} \tag{37}$$

where $f_*:\mathsf{X}\to\mathsf{X}$, $h_*:\mathsf{X}\to\mathsf{Y}$ $\tau^*:\mathsf{X}\times\mathsf{Z}\to\mathsf{X}$ are some functions, and ν_0 the initial distribution of X_0. In the sequel, it is assumed that $\{\zeta_k\}_{k\geq 0}$ and $\{\varepsilon_k\}_{k\geq 0}$ are independent and $\{\varepsilon_k\}_{k\geq 0}$ is an i.i.d sequence.

For $a>0$, define the set Lip_a of a-Lipshitz functions, *i.e.* $g\in\mathrm{Lip}_a$ if for all $x,x'\in\mathsf{X}^2$, $|g(x)-g(x')|\leq a|x-x'|$. For $b_0,b>0$, consider the set $\mathrm{S}_{b_0,b}$ the set of functions $g:\mathsf{X}\to\mathsf{Y}$ which are surjective and which satisfy, for all $x,x'\in\mathsf{X}^2$, $|x-x'|\leq b_0+b|g(x)-g(x')|$.

For a function $h:\mathsf{X}\to\mathsf{Y}$, a transition density kernel $t:\mathsf{X}\times\mathsf{X}\to\mathbb{R}$ and a probability density υ on Y, we compute the filtering distribution (12) $\phi_{\nu,n}[Y_{0:n}]$ using

$$Q(x,A)=\int_A t\left(x,x'-f(x)\right)\lambda^{\mathrm{Leb}}(\mathrm{d}x')\ ,\quad x\in\mathsf{X}\ ,A\in\mathcal{X}\ , \tag{38}$$

$$g(x,y)=\upsilon\left(y-h(x)\right)\ . \tag{39}$$

Consider the following assumptions :

M 1. There exist constants $a,b_0,b>0$ such that f is Lip_a and $h\in\mathrm{S}_{b_0,b}$.

M 2. The density υ is positive, continuous, and $\lim_{|u|\to\infty}\upsilon(u)=0$.

Notice that f^* and f are not necessarily contracting so that the both the observations and the filtering model used to construct $\phi_{\nu,n}[Y_{0:n}]$ are (possibly) non-ergodic. The assumption (M1) has been first considered in [93]. A function h satisfying (M1) can be viewed as a perturbation of a bijective function whose inverse is b-Lipschtiz. The rationale for considering such assumption is the following. For $y_1,\ y_2\in\mathsf{Y}$, the maximal distance between any two elements in the preimages $h^{-1}(\{y_1\})$ and $h^{-1}(\{y_2\})$ is controlled by $|y_1-y_2|$. The assumption (M2) is satisfied, for example, by Gaussian densities.

13.3.1 Nonlinear state-space model with i.i.d. state noise

In this Section, we assume that the observations and the filtering model are matched, *i.e.* $f^* = f$ and $h^* = h$. In addition, we consider the following assumption :

E 1. The state noise $\{\zeta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables with a positive and continuous density γ with respect to the Lebesgue measure λ^{Leb} on X. In addition, the transition kernel t in (38) is taken to be $t(x, x') = \gamma(x')$, for all $(x, x') \in \mathsf{X}^2$.

E 2. The function $\tau^*(x, \zeta)$ in (37) is given by : $\tau^*(x, \zeta) = \zeta$.

Under this assumption, for any $A \in \mathcal{X}$, the transition kernel Q may be expressed as

$$Q(x, A) = \int_A \gamma[x' - f(x)]\, \lambda^{\mathrm{Leb}}(\mathrm{d}x') \; . \tag{40}$$

Define by $D : \mathsf{Y} \times \mathsf{Y} \to \mathbb{R}$ the function

$$D(y, y') \stackrel{\mathrm{def}}{=} \sup \left\{ |f(z) - z'| \; : z \in h^{-1}(\{y\}), z' \in h^{-1}(\{y'\}) \right\} \; . \tag{41}$$

For any $r > 0$, we consider the minimum and the maximum of the state noise density over a ball of radius r :

$$\gamma^-(r) \stackrel{\mathrm{def}}{=} \inf_{|s|\leq r} \gamma(s) \; , \qquad \gamma^+(r) \stackrel{\mathrm{def}}{=} \sup_{|s|\leq r} \gamma(s) \; , \tag{42}$$

Lemma 13.1. . *Assume (M1)-(M2)-(E1). Then, for any $\Delta \in (0, \infty)$, the set valued function $\mathsf{C}_\Delta : \mathsf{Y} \to \mathcal{X}$, defined by*

$$y \longmapsto \mathsf{C}_\Delta(y) \stackrel{\mathrm{def}}{=} \{x \in \mathsf{X} : |h(x) - y| \leq \Delta\} \; . \tag{43}$$

is a LD-set function : for all $A \in \mathcal{X}$ and $x \in \mathsf{C}_\Delta(y)$; more precisely,

$$\varepsilon^-_\Delta(y, y')\lambda^{\mathrm{Leb}}[A \cap \mathsf{C}_\Delta(y')] \leq Q[x, A \cap \mathsf{C}_\Delta(y')] \leq \varepsilon^+_\Delta(y, y')\lambda^{\mathrm{Leb}}[A \cap \mathsf{C}_\Delta(y')] \; , \tag{44}$$

where, setting $c = (a + 1)b_0$ and $d = (a + 1)b$,

$$\begin{aligned} \varepsilon^-_\Delta(y, y') &\stackrel{\mathrm{def}}{=} \gamma^-[c + d\Delta + D(y, y')] \; , \\ \varepsilon^+_\Delta(y, y') &\stackrel{\mathrm{def}}{=} \gamma^+[c + d\Delta + D(y, y')] \; . \end{aligned} \tag{45}$$

Let Z^Δ_k be defined as :

$$Z^\Delta_k \stackrel{\mathrm{def}}{=} -\log \gamma^- \left[2c + d\Delta + ab|\varepsilon_{k-1}| + |\zeta_k| + b|\varepsilon_k|\right] \; . \tag{46}$$

Proposition 13.4. *Assume (M1), (M2), (E1) and (E2). Assume in addition that for all $\Delta > 0$,*

$$\mathbb{E}|Z^\Delta_1| < \infty \; . \tag{47}$$

Then, for any probability distributions ν_0, ν, and ν' on $(\mathsf{X}, \mathcal{X})$, we have

$$\limsup_{n\to\infty} n^{-1} \log \left\| \phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}] \right\|_{\mathrm{TV}} < 0, \quad \mathbb{P}_{\nu_0} - a.s.$$

The condition (47) is not very restrictive. For example, assume that $\{\zeta_k\}_{k\geq 0}$ and $\{\varepsilon_k\}_{k\geq 0}$ are sequences of Gaussian random variables. It follows, that $\gamma^-(r) = \gamma(r)$ for all $r \geq 0$. The condition (47) holds if $\mathbb{E}(|\varepsilon_1|^2) < \infty$ and $\mathbb{E}(|\zeta_1|^2) < \infty$ which are trivially satisfied. This result also extends

[32, Theorem 1.1] ; these authors assume that the densities υ and γ are upper and lower bounded, *i.e.* that there exist positive constants m_υ, M_υ, α_υ, β_υ, and m_γ, M_γ, α_γ and β_γ such that

$$m_\upsilon \exp(-\alpha_\upsilon |x|^{\beta_\upsilon}) \leq \upsilon(x) \leq M_\upsilon \exp(-\alpha_\upsilon |x|^{\beta_\upsilon}) \ , \tag{48}$$

$$m_\gamma \exp(-\alpha_\gamma |x|^{\beta_\gamma}) \leq \gamma(x) \leq M_\gamma \exp(-\alpha_\gamma |x|^{\beta_\gamma}) \tag{49}$$

together with (M1) and a condition which is slightly more restrictive than (M2). Under these assumptions, we may set $\gamma_-(r) = m_\upsilon \exp(-\alpha_\upsilon r^{\beta_\upsilon})$ for $r > 0$ and the condition (47) simply reads,

$$\mathbb{E}\left[|\epsilon_1|^{\beta_\upsilon}\right] < \infty \quad \text{and} \quad \mathbb{E}\left[|\zeta_1|^{\beta_\upsilon}\right] < \infty \ ,$$

which is of course satisfied under (48) and (49), without any conditions on the constants α_υ, β_υ, α_γ, and β_γ. The stability of the filter holds without requiring that the tails of the observation noise be *light* compared to the tails of the signal noise (this type of conditions is prevalent in many works on this topic since [24]).

With more stringent conditions on initial distributions, the convergence of the expected value of the total variation distance $\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}$ may be shown to be geometric. Define the log-moment generating function $\psi_Z(\alpha)$ of the random variable Z is defined by $\psi_Z(\alpha) \stackrel{\text{def}}{=} \log \mathbb{E}[\mathrm{e}^{\alpha Z}]$.

Proposition 13.5. *Assume that (M1), (M2), (E1), and (E2) hold. Then, for all $\Delta > 0$, there exists $\alpha_0 > 0$ such that*

$$\psi_{Z_1^\Delta}(\alpha_0) > 0 \ . \tag{50}$$

Let C be the LD-set function defined by (43). *Then, for any ν_0, ν and ν' probability measures on $(\mathsf{X}, \mathcal{X})$ and $\Delta > 0$ such that,*

$$\mathbb{E}_{\nu_0}\left\{\exp\left(\alpha_0 \left[\log \Phi_{\nu,\mathsf{C}}(Y_0, Y_1)\right]_-\right)\right\} < \infty \ , \mathbb{E}_{\nu_0}\left\{\exp\left(\alpha_0 \left[\log \Phi_{\nu',\mathsf{C}}(Y_0, Y_1)\right]_-\right)\right\} < \infty \ , \tag{51}$$

we have

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{E}_{\nu_0}\left[\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right] < 0 \ .$$

Assume that $\{\zeta_k\}_{k\geq 0}$ and $\{\varepsilon_k\}_{k\geq 0}$ are sequences of Gaussian random variables. The condition $\mathbb{E}[\mathrm{e}^{\alpha_0 \, Z_1^\Delta}] < \infty$ is equivalent to

$$\int_{\mathbb{R}^{d_\mathsf{X}+2d_\mathsf{Y}}} \exp\left[(\alpha_0 - \varsigma)|u|^2\right] \, \mathrm{d}u < \infty \ ,$$

where ς denotes some positive constant. Therefore, for $\alpha_0 > 0$ small enough, the condition (50) is satisfied. Proofs of Lemma 13.1, Propositions 13.4 and 13.5 are postponed to Section 13.4.

13.3.2 Nonlinear state-space model with dependent state noise

We now consider the case where the state noise $\{\zeta_k\}_{k\geq 0}$ can depend on previous states. This model has been introduced in [93, Section 3] and is important because it covers the case of partially observed discretely sampled diffusions, as well as partially observed stochastic volatility models [24, Section 2].

G 1. There exist ψ a positive continuous probability density function and constants $\mu_- > 0$ and $\mu_+ > 0$ such that, and for all $x, x' \in \mathsf{X}^2$,

$$\mu_- \psi(x') \leq t(x, x') \leq \mu_+ \psi(x') \ .$$

where t is the density transition kernel defined in (38).

This example also illustrates that the forgetting property is kept even when the distributions of the observations differ from the model.

G 2.
1. For some positive constants a^*, b_0^* and b^*, $f^* \in \mathrm{Lip}_{a^*}$ and $h^* \in \mathrm{S}_{b_0^*,b^*}$, where f^* and h^* are defined in (37).
2. There exists a function $\tau_+ : \mathsf{Z} \to \mathbb{R}$ such that for all $x \in \mathsf{X}$ and $\zeta \in \mathsf{Z}$, $|\tau^*(x,\zeta)| \le \tau_+(\zeta)$.
3. f^* and h^* are such that $\|f - f^*\|_\infty < \infty$ and $\|h - h^*\|_\infty < \infty$.

A first example of state equation satisfying (G1) is considered in [24]. A signal takes its values in X and follows the equation

$$X_k = f(X_{k-1}) + \sigma(X_{k-1})\xi_k \ , \tag{52}$$

where $\{\xi_k\}_{k\ge 0}$ is a sequence of i.i.d random variables and where $\sigma : \mathsf{X} \to \mathbb{R}^{d_\mathsf{X}\times d_\mathsf{X}}$ is a measurable function that satisfies, for all $x,\ u \in \mathsf{X}$, the following uniform ellipticity condition :

$$\sigma^-|u|^2 \le |\sigma^*(x)u|^2 \le \sigma^+|u|^2 \ , \tag{53}$$

where σ^-, σ^+ are positive constants. Another important example where assumption (G1) is satisfied is the case of certain discretely sampled diffusions. Let $(X_t)_{t\ge 0}$ be the unique solution of the following stochastic differential equation

$$\mathrm{d}X_t = \rho(X_t)dt + \sigma(X_t)\mathrm{d}B_t \ ,$$

where B is the d_X-dimensional Brownian motion and the functions $\rho : \mathbb{R}^{d_\mathsf{X}} \to \mathbb{R}^{d_\mathsf{X}}$ and $\sigma : \mathbb{R}^{d_\mathsf{X}} \to \mathbb{R}^{d_\mathsf{X}\times d_\mathsf{X}}$ are respectively of class C^1 and C^3 . Then, the sequence $\{X_k\}_{k\ge 0}$ satisfies assumption (G2) if the function σ is uniformly elliptic (condition (53)) ; The assumptions (M1), (M2) and (G2) are similar to those made in [24] and [93]. This allows to establish the forgetting of the initial condition with probability one without restriction on the signal-to-noise ratio and for sequences of observations which are not necessarily distributed according to the model used to compute the filtering distribution.

For the same reasons as above, we consider the set-valued function C_Δ defined in (43). Denote

$$q^-(r) \stackrel{\mathrm{def}}{=} \mu_- \times \inf_{|v|\le r} \psi(v) \ , \qquad q^+(r) \stackrel{\mathrm{def}}{=} \mu_+ \times \sup_{|v|\le r} \psi(v) \ , \tag{54}$$

Lemma 13.2. *Assume (M1)-(M2)-(G1). Then, for all $A \in \mathcal{X}$, $y \in \mathsf{Y}$, and $x \in \mathsf{C}_\Delta(y)$,*

$$\varepsilon_\Delta^-(y,y')\lambda^{\mathrm{Leb}}[A \cap \mathsf{C}_\Delta(y')] \le Q[x, A \cap \mathsf{C}_\Delta(y')] \le \varepsilon_\Delta^+(y,y')\lambda^{\mathrm{Leb}}[A \cap \mathsf{C}_\Delta(y')] \ , \tag{55}$$

where

$$\varepsilon_\Delta^-(y,y') \stackrel{\mathrm{def}}{=} q^-[c + d\Delta + D(y,y')] \ , \qquad \varepsilon_\Delta^+(y,y') \stackrel{\mathrm{def}}{=} q^+[c + d\Delta + D(y,y')] \ .$$

The proof is similar to the proof of Lemma 13.1 and is omitted for brevity.

Lemma 13.3. *Under (G1)-(G2), for all integer $k \ge 1$,*

$$D(Y_{k-1}, Y_k) \le \kappa + (a^*+1)b_0^* + a^*b^*|\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+(\zeta_k) \ ,$$

where

$$\kappa \stackrel{\mathrm{def}}{=} \|f - f^*\|_\infty + (b_0 + b\,\|h^* - h\|_\infty)(1 + a^*)$$

Now, define for all $\Delta > 0$

$$V_+^{*\Delta} \stackrel{\text{def}}{=} -\log q^- \left[c + d\Delta + \kappa + (1+a^*)b_0^* + a^*b^*|\varepsilon_0| + b^*|\varepsilon_1| + \tau_+(\zeta_0)\right] \ . \tag{56}$$

Proposition 13.6. *Assume (M1), (M2),(G1), and (G2). Assume in addition that for all $\Delta > 0$,*

$$\mathbb{E}\left(|V_+^{*\Delta}|\right) < \infty \ . \tag{57}$$

Then, for any initial probability distributions ν and ν' on $(\mathsf{X}, \mathcal{X})$, we have

$$\limsup_{n\to\infty} n^{-1} \log \left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}} < 0, \qquad \mathbb{P}_\star - a.s. \ ,$$

where $\mathbb{P}_\star$ is the distribution of the process specified by (37).

Let $\{Z_k^{*\Delta}\}_{k\geq 0}$ be the sequence defined for all $\Delta > 0$ and for all integer $k \geq 1$ by

$$Z_k^{*\Delta} \stackrel{\text{def}}{=} -\log q^- \left[c + d\Delta + \kappa + (a^*+1)b_0^* + a^*b^*|\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+^*(A_k^*)\right] \ .$$

Proposition 13.7. *Assume (M1), (M2), (G1), and (G2). Assume in addition that for all $\Delta > 0$, there exists a neighborhood $\alpha_0 > 0$ such that $\Psi_{Z_1^{*\Delta}}(\alpha_0) < \infty$. Let C_Δ be the LD-set function defined by* (43). *Then, for ν and ν' two probability measures on $(\mathsf{X}, \mathcal{X})$ and $\Delta > 0$ such that,*

$$\mathbb{E}_\star\left\{\exp\left(\alpha_0[\log \Phi_{\nu,\mathsf{C}_\Delta}(Y_0,Y_1)]_-\right)\right\} < \infty \ , \mathbb{E}_\star\left\{\exp\left(\alpha_0[\log \Phi_{\nu',\mathsf{C}_\Delta}(Y_0,Y_1)]_-\right)\right\} < \infty \ , \tag{58}$$

we have

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{E}_\star\left[\left\|\phi_{\nu,n}[Y_{0:n}] - \phi_{\nu',n}[Y_{0:n}]\right\|_{\mathrm{TV}}\right] < 0 \ .$$

Proofs of Propositions 13.6 and 13.7 are given in Section 13.5.

13.4 Proofs of Propositions 13.4 and 13.5

Proof of Lemma 13.1. Under assumption (M1), for any z in $h^{-1}(\{y\})$, and $x \in \mathsf{C}_\Delta(y)$,

$$|x - z| \leq b_0 + b\Delta \ . \tag{59}$$

Let $(y, y') \in \mathsf{Y}^2$. By (M1), h is surjective and we may pick $z \in h^{-1}(\{y\})$ and $z' \in h^{-1}(\{y'\})$. Using again (M1), it follows from (59) that, for all $(x, x') \in \mathsf{C}_\Delta(y) \times \mathsf{C}_\Delta(y')$,

$$|f(x) - x'| \leq |f(x) - f(z)| + |f(z) - z'| + |z' - x'| \leq a(b_0 + b\Delta) + D(y, y') + b_0 + b\Delta \ , \tag{60}$$

The proof follows from (40) and (60). □

Proof of Proposition 13.4. We will apply Theorem 13.1 by successively checking the assumptions (H1–3) and (27–30).

(H1) is satisfied. By (M2), for all $\eta > 0$, we may choose Δ_η large enough so that $\sup_{|s|>\Delta_\eta} \upsilon(s) \leq \eta \sup_{s\in\mathsf{X}} \upsilon(s)$. This, combined with (44), implies (H2) with $\mathsf{C}_\eta = \mathsf{C}_{\Delta_\eta}$ and $\Upsilon_\mathsf{X} = \sup_\mathsf{X} \upsilon$. (H3) is obvious.

To check (27–30), it will be needed to bound $\{D(Y_{k-1}, Y_k)\}_{k\geq 1}$ where D is defined in (41). For

$z,\ z' \in \mathsf{X}$ such that $h(z) = Y_{k-1},\ h(z') = Y_k$, it follows from (M1) that

$$\begin{aligned} |f(z) - z'| &\leq |f(z) - f(X_{k-1})| + |f(X_{k-1}) - X_k| + |X_k - z'| \ , \\ &\leq a(b_0 + b|\varepsilon_{k-1}|) + |\zeta_k| + b_0 + b|\varepsilon_k| \ . \end{aligned}$$

Therefore, for all integer $k \geq 1$,

$$D(Y_{k-1}, Y_k) \leq c + ab|\varepsilon_{k-1}| + |\zeta_k| + b|\varepsilon_k| \ . \tag{61}$$

We now consider conditions (27–29). Let η_0 be fixed and set $\Delta = \Delta_{\eta_0}$. Since by definition (42), γ^- is a nonincreasing function, it follows by plugging the bound (61) into (45) that

$$-n^{-1} \sum_{k=2}^{n} \log \varepsilon_\Delta^-(Y_{k-1}, Y_k) \leq n^{-1} \sum_{k=2}^{n} Z_k^\Delta \ , \tag{62}$$

where Z_k^Δ is defined in (46). Since the process $\{ab|\varepsilon_{k-1}| + |\zeta_k| + b|\varepsilon_k|\}_{k\geq 1}$ is stationary 2-dependent, the strong law of large numbers for m-dependent sequences and the integrability condition (47) yield

$$\lim_{n\to\infty} n^{-1} \sum_{k=2}^{n} Z_k^\Delta = \mathbb{E}(Z_1^\Delta) < \infty \ , \qquad \mathbb{P}_{\nu_0} - a.s. \tag{63}$$

By combining (62) and (63), the first condition (27) of Theorem 13.1 is satisfied. By assumption (M2), the density υ is bounded which implies that $\sup_{y\in\mathsf{Y}} \Upsilon_{\mathsf{X}}(y) \leq \sup \upsilon$. Hence, the second condition (28) of Theorem 13.1 is satisfied. We now consider the third condition (29). Since the measure appearing in the definition of the LD-set function does not depend on $y,\ y'$, the function $(y, y') \mapsto \Psi_{\mathsf{C}_\Delta(y')}(y, y')$, defined in (19), does not depend on y and is given by

$$\Psi_{\mathsf{C}_\Delta(y')}(y, y') = \int_{\mathsf{C}_\Delta(y')} \upsilon[y' - h(x)]\, \lambda^{\mathrm{Leb}}(\mathrm{d}x) \geq \lambda^{\mathrm{Leb}}[\mathsf{C}_\Delta(y')] \times \inf_{|s|\leq\Delta} \upsilon(s) \ .$$

Since the function h is uniformly continuous, for any fixed $\Delta > 0$, there exists $\delta > 0$ such that, for all $x,\ x' \in \mathsf{X}$ satisfying $|x - x'| \leq \delta$, we have $|h(x) - h(x')| \leq \Delta$. Since h is surjective, it follows that $\lambda^{\mathrm{Leb}}[\mathsf{C}_\Delta(y')]$ is bounded below by the volume of a ball of radius δ in $\mathbb{R}^{d_{\mathsf{X}}}$. Thus, we have, for all $y,\ y' \in \mathsf{Y}$,

$$\Psi_{\mathsf{C}_\Delta(y')}(y, y') \geq \varrho_\Delta \ , \tag{64}$$

for some $\varrho_\Delta > 0$, depending only on Δ. The third condition (29) of Theorem 13.1 follows.

We now prove (30). We have

$$\log \Lambda_\eta(Y_{0:n}, \alpha) \leq \max \left\{ \sum_{k=1}^{n} \delta_k U_k : \{\delta_k\}_{k=1}^n \in \{0,1\}^n, \sum_{k=1}^{n} \delta_k \geq \alpha n \right\} \ .$$

where

$$\begin{aligned} R_{\Delta_\eta}(x) &\stackrel{\mathrm{def}}{=} \log\left\{1 - (\gamma^-/\gamma^+)^2 \left[2c + d\Delta_\eta + x\right]\right\} \ , \\ U_k &\stackrel{\mathrm{def}}{=} R_{\Delta_\eta}(ab|\varepsilon_{k-1}| + |\zeta_k| + b|\varepsilon_k|) \ . \end{aligned} \tag{65}$$

Then, since for any $\upsilon > 0$, $U_k \leq -\upsilon \mathbb{1}\{U_k < -\upsilon\}$, we have

$$\begin{aligned}\log \Lambda_\eta(Y_{0:n}, \alpha) &\leq \max_{\{\{\delta_k\}_{k=1}^n \in \{0,1\}^n : \sum_{k=1}^n \delta_k \geq \alpha n\}} -\upsilon \sum_{k=1}^n \delta_k \mathbb{1}\{U_k < -\upsilon\} \\ &\leq -\upsilon \lfloor (\alpha n - \sum_{k=1}^n \mathbb{1}\{U_k \geq -\upsilon\})^+ \rfloor \ . \end{aligned} \tag{66}$$

Dividing by n and letting n goes to infinity, the strong LLN for 2-dependent sequences yields that $\mathbb{P}_{\nu_0} - a.s.$,

$$\limsup_{n\to\infty} n^{-1} \log \Lambda_\eta(Y_{0:n}, \alpha) \leq -\upsilon(\alpha - \mathbb{P}(U_1 > -\upsilon))) \ . \tag{67}$$

Note that U_1 is non positive and $\mathbb{P}(U_1 = 0) = \mathbb{P}(|\varepsilon_{k-1}| + |\zeta_k| + |\varepsilon_k| = \infty) = 0$ by the integrability condition (47). Hence, U_1 is almost surely negative and $\lim_{\upsilon\to 0} \mathbb{P}(U_1 > -\upsilon) = 0$; we may thus choose υ small enough so that $\alpha - \mathbb{P}(U_1 > -\upsilon) > 0$. The rhs is then negative by taking υ sufficiently small.

□

Proof of Proposition 13.5. (51) implies that $r_0(\nu, n) \vee r_0(\nu', n) \leq c_0 \mathrm{e}^{-\delta_0 n}$ for some c_0, $\delta_0 > 0$. Now, recall that ψ_Z denotes the log-moment generating function of the random variable Z defined by $\psi_Z(\lambda) \stackrel{\text{def}}{=} \log \mathbb{E}[\mathrm{e}^{\lambda Z}]$ and we define its Legendre's transformation by

$$\psi_Z^*(x) = \sup_{\lambda \geq 0} \{x\lambda - \psi_Z(\lambda)\} \ .$$

We start by giving an exponential inequality for m-dependent variables whose proof is elementary.

Lemma 13.4. *Let $\{Z_k\}_{k\geq 0}$ be a sequence of m-dependent stationary random variables. Then, for all $M \geq 0$,*

$$\mathbb{P}\left(\sum_{k=1}^n Z_k \geq Mn\right) \leq m \exp[-\lfloor n/m \rfloor \psi_{Z_1}^*(M)] \ .$$

It follows by equation (62) that

$$\mathbb{P}\left(-n^{-1} \sum_{k=2}^n \log \varepsilon_\Delta^-(Y_{k-1}, Y_k) \geq M_1 n\right) \leq \mathbb{P}\left(\sum_{k=2}^n Z_k^\Delta \geq M_1 n\right) \ .$$

Thanks to (50), by applying Lemma 13.4, there exist some constant c_1, $\delta_1 > 0$ such that $r_1(n) \leq c_1 \mathrm{e}^{-\delta_1 n}$. Since υ is bounded, we can choose M_2 large enough such that $r_2(n) = 0$. By (64), for all $(y, y') \in \mathsf{Y}^2$, $\Psi_{\mathsf{C}_\Delta(y')}(y, y') \geq \varrho_\Delta$, for some $\varrho_\Delta > 0$. Then, by choosing M_3 large enough, we have $r_3(n) = 0$. Using (66), $r_4(n)$ is bounded by

$$r_4(n) \leq \mathbb{P}\left(-\upsilon \left\lfloor \left(\sum_{k=1}^n (\alpha - \mathbb{1}\{U_k \geq -\upsilon\})\right)^+ \right\rfloor \geq -\delta n\right)$$

Choosing υ such that $\alpha - \mathbb{P}(U_k > -\upsilon) > 0$ and then δ such that $\upsilon(\alpha - \mathbb{P}(U_k > -\upsilon)) > \delta$ and applying Lemma 13.4 with $Z_k = \alpha - \mathbb{1}\{U_k \geq -\upsilon\}$ which is bounded provides the existence of constants c_4, $\delta_4 > 0$ such that $r_4(n) \leq c_4 \mathrm{e}^{-\delta_4 n}$. Thus, Theorem 13.2 applies and provides a geometric rate. □

13.5 Proofs of Propositions 13.6 and 13.7

Proof of Lemma 13.3. For all integer $k \geq 1$, for $z,\ z' \in \mathsf{X}$ such that $h(z) = Y_{k-1},\ h(z') = Y_k$ and for $u,\ u' \in \mathsf{X}$ such that $h^*(u) = Y_{k-1},\ h^*(u') = Y_k$, we have

$$\begin{aligned} |f(z) - z'| &\leq |f(z) - f^*(z)| + |f^*(z) - f^*(u)| + |f^*(u) - u'| + |u' - z'| \ , \\ &\leq \|f - f^*\|_\infty + a^*|z - u| + |f^*(u) - u'| + |u' - z'| \ . \end{aligned} \tag{68}$$

Let us notice that

$$|z - u| \leq b_0 + b|h(z) - h(u)| \leq b_0 + b\underbrace{|h(z) - h^*(u)|}_{=0} + b|h^*(u) - h(u)| \ .$$

Then, by denoting $K = b_0 + b\,\|h^* - h\|_\infty$, it follows that $|z - u| \leq K$ and similarly, $|z' - u'| \leq K$. Combining these two upper bounds with (68) leads to

$$\begin{aligned} |f(z) - z'| &\leq \kappa + |f^*(u) - f^*(X_{k-1})| + |f^*(X_{k-1}) - X_k| + |X_k - u'| \ , \\ &\leq \kappa + a^*[b_0^* + b^*|h^*(z) - h^*(X_{k-1})|] + \tau_+(\zeta_k) + b_0^* + b^*|h^*(X_k) - h^*(u')| \ , \end{aligned}$$

where $\kappa = \|f - f^*\|_\infty + K(1 + a^*)$. Thus, it is proven that, for all integer $k \geq 1$,

$$D(Y_{k-1}, Y_k) \leq \kappa + (a^* + 1)b_0^* + a^*b^*|\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+(\zeta_k) \ .$$

□

Proof of Proposition 13.6. Define, for all $\Delta > 0$ and for all integer $k \geq 1$,

$$V_k^{*\Delta} = -\log q^- \left[c + d\Delta + \kappa + (a^* + 1)b_0^* + a^*b^*|\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+(\zeta_k)\right] \ . \tag{69}$$

Using the definitions (54), (55) of q^- and ε_Δ^-, Lemma 13.3 shows that

$$-\,n^{-1}\sum_{k=2}^{n} \log \varepsilon_\Delta^-(Y_{k-1}, Y_k) \leq n^{-1}\sum_{k=2}^{n} V_k^{*\Delta} \ . \tag{70}$$

Thus, (27) follows from LLN for 2-dependent sequences.

The proof of assumptions (28) and (29) can be checked as in Proposition 13.4. It remains to check (30). Let

$$\begin{aligned} U_n(\upsilon) &\stackrel{\text{def}}{=} n^{-1}\sum_{k=1}^{n} \mathbb{1}\{G(\varepsilon_{k-1}, \varepsilon_k, \zeta_k) \geq -\upsilon\} \ , \\ V_n(\upsilon) &\stackrel{\text{def}}{=} n^{-1}\sum_{k=1}^{n} \mathbb{P}\left\{G(\varepsilon_{k-1}, \varepsilon_k, \zeta_k) \geq -\upsilon | \mathcal{F}_{k-1}\right\} \ , \end{aligned}$$

where $G(\varepsilon_{k-1}, \varepsilon_k, \zeta_k) \stackrel{\text{def}}{=} R_{\Delta_\eta}[\kappa + (a^* + 1)b_0^* + a^*b^*|\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+(\zeta_k)]$ and R_{Δ_η} is defined in (65). Similarly to (66), we have for any $\upsilon > 0$,

$$\log \Lambda_\eta(Y_{0:n}, \alpha) \leq -\upsilon\lfloor(\alpha n - nU_n(\upsilon)\})^+\rfloor \ .$$

so that $\mathbb{P}_\star$-a.s.

$$\limsup_n \log \Lambda_\eta(Y_{0:n}, \alpha) \leq -\upsilon \lfloor (\alpha - \limsup_n U_n(\upsilon)\})^+ \rfloor \ . \tag{71}$$

Moreover, using the LLN for 2-dependent sequences, we have that $\mathbb{P}_\star$-a.s.

$$\lim_{n\to\infty} U_n(\upsilon) = \lim_{n\to\infty} V_n(\upsilon) \leq \mu_+^* \mathbb{E}\left[\int \mathbb{1}\{G(\varepsilon_0, \varepsilon_1, w) \geq -\upsilon\}\psi^*(w)\,\mathrm{d}w\right] \ .$$

Since G is $\mathbb{P}_\star$-a.s. negative, the right-hand side of the above equation thus converges $\mathbb{P}_\star$-a.s. to 0 as υ tends to 0. Thus, the right-hand side of (71) is negative by choosing υ sufficiently small. □

Proof of Proposition 13.7. (58) implies that $r_0(\nu, n) \vee r_0(\nu', n) \leq c_0 \mathrm{e}^{-\delta_0 n}$ for some c_0, $\delta_0 > 0$. It follows, by definition of r_1 and Lemma 13.3 that

$$r_1(n) = \mathbb{P}_\star\left(-n^{-1}\sum_{k=2}^n \log q^-[c + d\Delta + D(Y_{k-1}, Y_k)] \geq M_1 n\right) \leq \mathbb{P}_\star\left(n^{-1}\sum_{k=2}^n Z_k^{*\Delta} \geq M_1 n\right) \ .$$

with $c_0 = c + d\Delta + \kappa + (a^*+1)b_0^*$. Then, applying Lemma 13.4, there exist some constants c_1, $\delta_1 > 0$ such that $r_1(n) \leq c_1 \mathrm{e}^{-\delta_1 n}$. By the same arguments as in proof of Proposition 13.5, the real numbers M_2 and M_3 can be chosen large enough such that $r_2(n) = 0$ and $r_3(n) = 0$. Let us denote by $\{U_k^+\}_{k\geq 0}$ the sequence defined by $U_k^+ \stackrel{\text{def}}{=} R_\Delta[\kappa + (a^*+1)b_0^* + a^* b^* |\varepsilon_{k-1}| + b^*|\varepsilon_k| + \tau_+(\zeta_k)]$, for all integer $k \geq 1$. Similarly to the proof of Proposition 13.5, for any $\delta > 0$,

$$r_4(n) \leq \mathbb{P}\left(-\upsilon \left\lfloor \left(\sum_{k=1}^n (\alpha - \mathbb{1}\{U_k^+ \geq -\upsilon\})\right)^+ \right\rfloor \geq -\delta n\right) \ .$$

We first choose υ small enough so that $\alpha - \mathbb{P}(U_k^+ > -\upsilon) > 0$ holds; then δ is chosen such that $\upsilon(\alpha - \mathbb{P}(U_k^+ > -\upsilon)) > \delta$. By applying Lemma 13.4 with $Z_k = \alpha - \mathbb{1}\{U_k^+ \geq -\upsilon\}$ which is a bounded random variable, there exist constants c_4, $\delta_4 > 0$ such that $r_4(n) \leq c_4 \mathrm{e}^{-\delta_4 n}$. Thus, Theorem 13.2 applies and provides a geometric rate. □

Chapitre 14

Filtre Particulaire

Sommaire

14.1 Introduction 175
14.2 Stabilité asymptotique pour le BOT 176
14.3 Un pas vers le filtrage particulaire 180
14.4 Dynamique manœuvrante 181
14.5 Simulations 182
14.6 Perspectives 186

14.1 Introduction

La trajectographie passive est une question que l'on trouve présente dès l'origine des filtres particulaires, dans l'article fondateur de gordon et al.[56]. Il est vraisemblable que ce problème a motivé pour partie l'introduction de cette nouvelle génération d'algorithmes. On trouve, dans la littérature scientifique, de multiples déclinaisons de filtres particulaires pour le BOT avec des dynamiques plus ou moins sophistiquées, voir par exemple [7], [90], [54], [16], [71], [21]. Toutefois, bien que les travaux relatifs au BOT soient nombreux, on ne trouve que peu de résultats théoriques sur le comportement des méthodes de Monte-Carlo séquentielles dans ce contexte.

Avant de considérer plus avant le filtrage particulaire appliqué au BOT, rappelons les enjeux de ces techniques et pourquoi la trajectographie passive demeure un relief particulier dans le paysage des HMM issues de situations réelles. La fonction du filtre particulaire est, par nature, d'approcher le filtre optimal d'une chaîne de Markov cachée. On dispose d'une borne pour l'estimation de la chaîne cachée $X_{0:n}$ appelée *Borne de Cramer-Rao a posteriori* (PCRB, Posterior Cramer-Rao Bound) [114]. Dans le cas précis du problème de filtrage, à savoir l'estimation de l'état courant X_n à partir des observations, cette borne est un minorant de la variance du filtre optimal. Un pas important a été franchi pour mesurer la performance des méthodes particulaires avec l'algorithme de calcul itératif de la PCRB décrit par Tichavsky et al.[89].[1] Les praticiens utilisent donc cette borne pour évaluer l'efficacité d'un filtre particulaire, selon que la variance de celui-ci s'approche ou non de ce minorant.

Toutefois, cette borne n'est un critère pertinent que si le filtre particulaire converge vers le filtre optimal. Cette question n'est, à notre connaissance, que partiellement résolue. Le cas ergodique

[1] Signalons également les travaux de Bréhard et al.[20] qui introduisent un jeu de coordonnées dans lequel le calcul de la PCRB peut être fait de manière itérative exacte. L'intérêt semble toutefois limité car, bien que le calcul soit exact, il semble que la complexité algorithmique soit terriblement pénalisante alors que la méthode d'origine qui requiert une approximation par Monte-Carlo est d'un coût raisonnable.

est traité sous des hypothèses assez restrictives de noyau fortement mélangeant, voir [36], [37], [35], [79], [78], [94], [75], [113].

En revanche, pour un signal non-ergodique, les résultats se raréfient considérablement. Compte-tenu des liens étroits qui existent entre la propriété d'oubli et la convergence des filtres à particules, ceci est à mettre en parallèle avec l'avancée des travaux sur la stabilité asymptotique. En effet, hormis le très classique cas linéaire gaussien, aucun travaux, à notre connaissance, avant ceux de Kleptsyna et al.[73], ne permettaient de couvrir le cas d'un noyau de transition gaussien. Tous les autres résultats sont récents [40],[113], [32] ce qui montre combien le BOT s'inscrit dans des problématiques actuelles à la fois pour le milieu industriel et pour le monde académique.

Mentionnons enfin que la convergence des algorithmes particulaires n'implique pas qu'ils vont fournir une information pertinente au le praticien dont la préoccupation est de localiser le mobile pisté. En effet, l'étalement du filtre optimal est fonction à la fois des manœuvres du porteur ainsi que de la dynamique supposée de la cible, ce que l'on pourrait mesurer comme un défaut d'ergodicité de l'équation d'état. Ceci relance la question très complexe de la commande mais pour le BOT formulé en tant que chaîne de Markov cachée [44].

Dans ce chapitre, on expose d'abord comment étendre le résultat de stabilité asymptotique en trajectographie passive puis on décrit comment adapter les techniques présentées dans Oudjane et al.[93] à notre problème. En dernière section, on présente les résultats de quelques simulations de filtres particulaires pour différents scénarios de trajectographie passive. Les résultats de ces simulations appellent à une certaine prudence sinon réserve car les phénomènes observés ne semblent pas aussi concluants que ceux que l'on trouve dans la littérature. Plusieurs hypothèses sont invoquées pour justifier ces différences.

14.2 Stabilité asymptotique pour le BOT

On reprend la formulation du problème de trajectographie passive comme chaîne de Markov cachée décrite en page 24. On note l'état du système à l'instant t_k par $X_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, où (x_k, y_k) désigne la position de la cible dans le plan et $(\dot{x_k}, \dot{y_k})$ le vecteur vitesse. Les équations d'état et d'observation sont

$$\begin{cases} X_k = FX_{k-1} + \zeta_k \ , & (1) \\ Y_k = h(X_k, t_k) + \varepsilon_k \ , & (2) \end{cases}$$

avec X_k la matrice colonne $X_k = \begin{pmatrix} x_k & y_k & \dot{x_k} & \dot{y_k} \end{pmatrix}^{\mathrm{T}}$, $F = \begin{pmatrix} I_2 & \Delta_T I_2 \\ 0 & I_2 \end{pmatrix}$ où Δ_T désigne la période d'échantillonnage, $\{\zeta_k\}_{k\geq 0}$ définie par $\zeta_k = \beta G V_k$ où $\beta > 0$, $G = \begin{pmatrix} \frac{\Delta_T^2}{2} I_2 \\ \Delta_T I_2 \end{pmatrix}$ et $\{V_k\}_{k\geq 0}$ est une suite de variables indépendantes gaussiennes standards dans $\mathbb{R}^2$, $\{\varepsilon_k\}_{k\geq 0}$ une suite de variables i.i.d. de loi $\mathcal{N}(0, \sigma^2)$ indépendantes de $\{\zeta_k\}_{k\geq 0}$ et h l'application définie, pour $x = \begin{pmatrix} x_1 & x_2 & \dot{x_1} & \dot{x_2} \end{pmatrix}^{\mathrm{T}}$, par

$$h(x, t) \stackrel{\text{def}}{=} \Psi[(x_1, x_2), t] \ ,$$

où $\Psi[(x_1,x_2),t]$ est défini par $\Psi[(0,0),t]=0$ et $\Psi[(x_1,x_2),t]=\mathrm{angle}[x_1-O_1(t),x_2-O_2(t)]$ pour $(x_1,x_2)\neq O(t)$ avec la fonction d'angle donnée par

$$\mathrm{angle}(x,y)\stackrel{\mathrm{def}}{=}\begin{cases}\arctan(x/y)+\pi\times\mathrm{sgn}(x)\times\mathbb{1}_{y<0} & \text{si } x\neq 0 \text{ et } y\neq 0,\\ \frac{\pi}{2}\times\mathrm{sgn}(x) & \text{si } x\neq 0 \text{ et } y=0,\\ \frac{\pi}{2}\times[1-\mathrm{sgn}(y)] & \text{si } x=0 \text{ et } y\neq 0.\end{cases}$$

Compte-tenu de la dépendance en temps, le filtre optimal est défini par

$$\begin{aligned}\phi_{\nu,n}(y_{0:n},t_{0:n})&\stackrel{\mathrm{def}}{=}\frac{\nu\left[g(\cdot,y_0,t_0)Qg(\cdot,y_1,t_1)Q\ldots Qg(\cdot,y_n,t_n)\mathbb{1}_A\right]}{\nu\left[g(\cdot,y_0,t_0)Qg(\cdot,y_1,t_1)Q\ldots Qg(\cdot,y_n,t_n)\right]}\\&=\frac{\int_{\mathsf{X}^{n+1}}\nu(\mathrm{d}x_0)g(x_0,y_0,t_0)\prod_{i=1}^n Q(x_{i-1},\mathrm{d}x_i)g(x_i,y_i,t_i)\mathbb{1}_A(x_n)}{\int_{\mathsf{X}^{n+1}}\nu(\mathrm{d}x_0)g(x_0,y_0,t_0)\prod_{i=1}^n Q(x_{i-1},\mathrm{d}x_i)g(x_i,y_i,t_i)}\ ,\end{aligned}$$

où $g(x,y,t)=\upsilon[y-h(x,t)]$ avec υ la densité de la loi $\mathcal{N}(0,\sigma^2)$.

Pour $u=(u_1,u_2)\in\mathbb{R}^2$, on note

$$\mathsf{K}[u]\stackrel{\mathrm{def}}{=}\left\{x\in\mathbb{R}^4:|(x_1,x_2)-(u_1,u_2)|\in[R_{\min},R_{\max}]\quad\text{et}\quad|(\dot{x_1},\dot{x_2})|\in[v_{\min},v_{\max}]\right\}\ .$$

Pour tout $t\in[0,1]$ et $(u,v)\in\mathsf{K}[O(t)]^2$, on a

$$|h(u,t)-h(v,t)|\leq\Delta\Longrightarrow|u-v|\leq\varphi(\Delta)+v_{\max}-v_{\min}\ , \tag{3}$$

où, par des considérations trigonométriques, la quantité $\varphi(\Delta)$ est définie

$$\varphi(\Delta)\stackrel{\mathrm{def}}{=}\left\{R_{\max}^2-2R_{\min}R_{\max}\cos\Delta+R_{\min}^2\right\}^{1/2}\ . \tag{4}$$

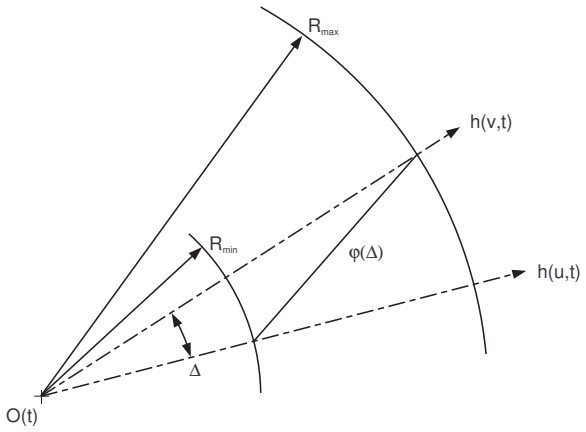


Fig. 14.1 – Contrôle pour un écart d'angle donné

Le contrôle décrit par l'équation (3) joue un rôle essentiel dans la localisation de l'état caché. En effet, pour deux mesures d'angles réalisées au même instant t distantes de Δ, la quantité $\varphi(\Delta)$ représente l'écart maximal en distance pour les états contraints dans $\mathsf{K}[O(t)]$ dont les azimuts sont ces mesures d'angles. Autrement dit, si deux états donnent des azimuts proches, leurs positions respectives ne peuvent être arbitrairement éloignées. Bien que ce fait puisse sembler naturel au praticien, on est contraint, sur le plan théorique, d'imposer des hypothèses très fortes au signal

$\{X_k\}_{k\geq 0}$ et qui sont vraisemblablement discutables.

Notons Q le noyau de transition de $\{X_k\}_{k\geq 0}$ avec $Q \ll \lambda^{\mathrm{Leb}}$ et $q = \mathrm{d}Q/\mathrm{d}\lambda^{\mathrm{Leb}}$. Alors, pour tout $A \in \mathcal{X}$ et tout $x \in \mathsf{X}$,

$$Q(x, A) = \int_A \gamma(x' - Fx)\lambda^{\mathrm{Leb}}(\mathrm{d}x') \ . \tag{5}$$

En vue d'appliquer les résultats de stabilité asymptotique obtenus au chapitre 13, on définit la fonction LD-set suivante :

$$(y, t) \longmapsto \mathsf{C}_\Delta(y, t) \stackrel{\mathrm{def}}{=} \{x \in \mathsf{K}[O(t)] : |h(x, t) - y| \leq \Delta\} \ .$$

L'hypothèse **H 1** formulée au chapitre 13 page 159 est clairement satisfaite. L'hypothèse **H 2** (page 162) doit être légèrement adaptée compte-tenu de la contrainte d'appartenance à $\mathsf{K}[O(t)]$ à l'instant t. Il s'agit de vérifier que, pour tout $\eta > 0$, il existe $\Delta > 0$ tel que

$$\sup_{x \in \mathsf{K}[O(t)] \setminus \mathsf{C}_\Delta(y,t)} \upsilon[y - h(x, t)] \leq \eta \sup \upsilon \ ,$$

ce qui découle clairement des caractéristiques du noyau gaussien υ.

Ainsi, pour tout $(x, x') \in \mathsf{C}_\Delta(y, t) \times \mathsf{C}_\Delta(y', t')$, en considérant des antécédents $z \in \mathsf{K}[O(t)]$ et $z' \in \mathsf{K}[O(t')]$ respectivement pour y par $h(\cdot, t)$ et y' par $h(\cdot, t')$, on a

$$|x' - Fx| \leq |x' - z'| + |z' - Fz| + \|F\| \, |z - x'| \ .$$

Il s'ensuit

$$|x' - Fx| \leq (1 + \|F\|)[\varphi(\Delta) + v_{\max} - v_{\min}] + D[(y, t), (y', t + \delta_t)] \ , \tag{6}$$

où D est défini par

$$D[(y, t), (y', t')] \stackrel{\mathrm{def}}{=} \sup \left\{|z' - Fz| : y = h(z, t), \ y' = h(z', t'), \ (z, z') \in \mathsf{K}[O(t)] \times \mathsf{K}[O(t')]\right\} \ .$$

En reprenant les notations du chapitre 13, on pose

$$\gamma^-(r) \stackrel{\mathrm{def}}{=} \inf_{|s| \leq r} \gamma(s) \ , \qquad \gamma^+(r) \stackrel{\mathrm{def}}{=} \sup_{|s| \leq r} \gamma(s) \ . \tag{7}$$

Par suite, d'après (5) et (6), pour tout $A \in \mathcal{X}$ et $x \in \mathsf{C}[(y, t), \Delta]$,

$$\begin{aligned} &\varepsilon^-_\Delta[(y, t), (y', t')]\lambda^{\mathrm{Leb}}\left[A \cap \mathsf{C}_\Delta(y', t')\right] \\ &\qquad \leq Q\left[x, A \cap \mathsf{C}_\Delta(y', t')\right] \leq \varepsilon^+_\Delta[(y, t), (y', t')]\lambda^{\mathrm{Leb}}\left[A \cap \mathsf{C}_\Delta(y', t')\right] \ , \end{aligned} \tag{8}$$

où

$$\begin{aligned} \varepsilon^-_\Delta[(y, t), (y', t')] &\stackrel{\mathrm{def}}{=} \gamma^-\left\{(1 + \|F\|)[\varphi(\Delta) + v_{\max} - v_{\min}] + D[(y, t), (y', t + \delta_t)]\right\} \ , \\ \varepsilon^+_\Delta[(y, t), (y', t')] &\stackrel{\mathrm{def}}{=} \gamma^+\left\{(1 + \|F\|)[\varphi(\Delta) + v_{\max} - v_{\min}] + D[(y, t), (y', t + \delta_t)]\right\} \ . \end{aligned}$$

Considérons la suite $\{Y_k\}_{k\geq 0}$ d'observations issues[2] de (2). Pour une observation Y_k, on note $\widehat{Y_k}$ la mesure d'angle correspondante[3]. Soit $u \in \mathsf{K}[O(t_k)]$ tel que $h(u, t_k) = \widehat{Y_k}$ et $u' \in \mathsf{K}[O(t_{k+1})]$

[2]On pourrait aussi considérer que les observations sont issues d'un autre système d'évolution de même type que (1-2), à l'instar de ce qui a été fait en section 13.3.2.

[3]Cette précaution permet de considérer tout type de bruit d'observation, notamment un bruit gaussien à support dans $\mathbb{R}$ tout entier

tel que $h(u', t_{k+1}) = \widehat{Y_{k+1}}$. Alors, il vient

$$|z' - Fz| \leq |z' - X_{k+1}| + |X_{k+1} - FX_k| + \|F\| \, |X_k - z| \; .$$

À présent, on souhaite établir, comme on l'a fait au chapitre 13, un contrôle de la quantité $D[(Y_k, t_k), (Y_{k+1}, t_{k+1})]$ dont le rôle est de localiser les états cachés afférents. Il s'agit donc de relier cette variable aléatoire avec le contrôle d'états par de mesures d'angles, contrôle fourni par l'implication (3). Ceci nous amène à formuler l'hypothèse suivante :

(H) La trajectoire porteur $t \mapsto O(t)$ est telle que

$$\{X_k\}_{k \geq 0} \in \prod_{k \geq 0} \mathsf{K}[O(t_k)] \quad \mathbb{P} - p.s.$$

Sous cette hypothèse, il vient

$$D[(Y_k, t_k), (Y_{k+1}, t_{k+1})] \leq \varphi(\varepsilon_{k+1}) + |\zeta_k| + \|F\| \, \varphi(\varepsilon_k) \; . \tag{9}$$

Utilisant la décroissance de γ^-, on a $\log \varepsilon_\Delta^-[(Y_{k-1}, t_{k-1}), (Y_k, t_k)] \geq -Z_k^\Delta$ où, pour tout entier $k \geq 1$ et $\Delta > 0$, Z_k^Δ est défini par

$$Z_k^\Delta \stackrel{\text{def}}{=} -\log \gamma^- \left\{ (1 + \|F\|)[\varphi(\Delta) + v_{\max} - v_{\min}] + \varphi(\varepsilon_{k+1}) + |\zeta_k| + \|F\| \, \varphi(\varepsilon_k) \right\} \; .$$

Proposition 14.1. *On considère la chaîne de Markov cachée $\{X_k, Y_k\}_{k \geq 1}$ associée au modèle de trajectographie passive décrit par les équations (1-2). On suppose que l'hypothèse* (H) *est satisfaite et que, pour tout $\Delta > 0$,*

$$\mathbb{E}|Z_1^\Delta| < \infty \; .$$

Alors, pour toutes lois de probabilités ν_0, ν et ν' sur $(\mathsf{X}, \mathcal{X})$, on a

$$\limsup_{n \to \infty} n^{-1} \log \left\| \phi_{\nu,n}[Y_{0:n}, t_{0:n}] - \phi_{\nu',n}[Y_{0:n}, t_{0:n}] \right\|_{\mathrm{TV}} < 0, \quad \mathbb{P}_{\nu_0} - p.s.$$

Démonstration. La preuve est identique à celle de la proposition 13.4 du chapitre 13 sous réserve d'avoir, pour une valeur $\Delta > 0$ fixée, $\lambda^{\mathrm{Leb}}[\mathsf{C}_\Delta(y, t)] \geq \delta$, pour tout $(y, t) \in \mathbb{R} \times [0, 1]$ avec $\delta > 0$ qui ne dépend que de Δ. Cette dernière condition est facilement vérifiée puisque l'on a

$$\lambda^{\mathrm{Leb}}[\mathsf{C}_\Delta(y, t)] = 2\Delta(R_{\max} - R_{\min}) \times (v_{\max} - v_{\min})^2 \; .$$

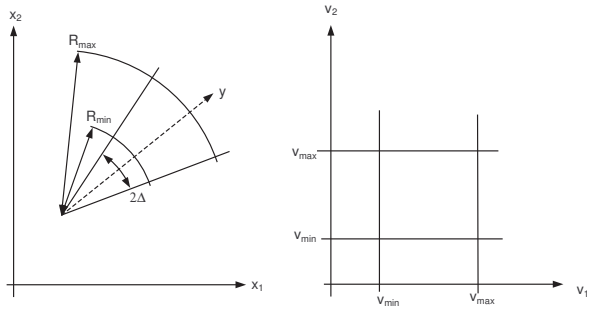


Fig. 14.2 – Représentation d'un LD-set $\mathsf{C}_\Delta(y, t)$

□

On peut adapter de manière identique la convergence en espérance décrite par la proposition 13.5.

L'hypothèse (H) peut sembler fortement restrictive puisqu'elle contraint la cible à évoluer dans une zone à distance bien délimitée par rapport à la trajectoire porteur. Ceci conditionne directement le noyau de transition de la chaîne $\{X_k\}_{k\geq 0}$ et exclut le cas le plus usité en trajectographie, à savoir un noyau gaussien. Cette condition relance une autre question qu'est celle de la commande du porteur. En effet, l'hypothèse n'a de sens que si le porteur agit pour que la cible soit dans une zone où la trajectographie est matériellement réalisable. Donc, l'hypothèse, bien que très contraignante au regard d'une modélisation théorique, est tout à fait raisonnable sur le plan pratique.

Intuitivement, cette contrainte n'est pas aberrante puisque, d'une certaine manière, elle permet de retrouver l'ergodicité que l'on n'exige pas dans la chaîne d'état. Cependant, en cherchant à suivre la trame des majorations mises en œuvre au chapitre 13, on s'oblige à des calculs fins alors qu'une majoration assez grossière est possible et qui permet d'établir les mêmes résultats. Pour tout $t \in [0,1]$ et $(u,v) \in \mathsf{K}[O(t)]^2$, on a

$$|h(u,t) - h(v,t)| \leq \Delta \Longrightarrow |u - v| \leq R_{\max} + R_{\min} + v_{\max} - v_{\min} \ ,$$

par des conditions géométriques évidentes. Avec cette nouvelle majoration, on peut réécrire la proposition 14.1 en modifiant de manière *ad hoc* la suite de variables aléatoires $\{Z_k^{\Delta}\}_{k\geq 0}$ et l'hypothèse de moment.

Ces considérations laissent penser que la formulation de l'hypothèse (H) est sans doute trop restrictive. Toutefois, nous ne voyons pas à ce jour comment relâcher cette hypothèse quitte à contraindre davantage la trajectoire porteur afin de pouvoir, par exemple, traiter le cas d'un noyau de transition gaussien .

14.3 Un pas vers le filtrage particulaire

Un des enjeux dans la formulation du BOT comme chaîne de Markov cachée est de savoir si oui ou non il y a convergence des filtres particulaires sur un horizon temporel infini. À notre connaissance, très peu de résultats s'appliquent actuellement aux cas des chaînes non-ergodiques. Dans les travaux de Oudjane et al.[93] et Crisan et al.[33], les convergences obtenues ne concernent pas directement les filtres particulaires mais des filtres tronqués sur les LD-set fonctions[4]. Les hypothèses de [33] sur la fonction d'observation qui doit être bijective, lipschitzienne et de réciproque lipschitzienne sont trop fortes pour que les résultats qui en découlent soient adaptés au cas de la trajectographie passive.

Il semble que les résultats de [93] puissent se transposer dans le cadre du BOT. La condition clé pour que tout fonctionne est l'hypothèse (E1) formulée au chapitre 13. Afin de l'adapter en trajectographie passive, on suppose l'hypothèse (H) à savoir

(H) La trajectoire porteur $t \mapsto O(t)$ est telle que

$$\{X_k\}_{k\geq 0} \in \prod_{k\geq 0} \mathsf{K}[O(t_k)] \quad \mathbb{P} - p.s.$$

Cette condition va nous permettre d'utiliser le contrôle suivant : pour tout $t \in [0,1]$ et $(u,v) \in$

[4]Ce concept n'était pas encore introduit lors de la publication de ces travaux mais les auteurs travaillent exactement sur les mêmes ensembles que ceux induits par les LD-set fonctions.

$\mathsf{K}[O(t)]^2$, on a

$$|h(u,t) - h(v,t)| \leq \Delta \Longrightarrow |u - v| \leq \varphi(\Delta) + v_{\max} - v_{\min} \ ,$$

où φ est défini par (4). Cette inégalité n'est pas exactement du même type que la condition (E1) mais les arguments de [93] s'adaptent sans difficulté avec cette inégalité.

14.4 Dynamique manœuvrante

Le modèle de trajectographie passive décrit par les équations (1-2) est relativement simple car la dynamique de base « $X_k = FX_{k-1}$ » est celle d'un MRU. On peut s'intéresser à des dynamiques plus complexes, dynamiques dites manœuvrantes où le régime peut changer de mode. Les modèles usuellement considérés pour le BOT ont pour équation

$$\begin{cases} X_k = F^{(r_k)}(X_{k-1}) \cdot X_{k-1} + \zeta_k \ , \\ Y_k = h(X_k, t_k) + \varepsilon_k \ , \end{cases}$$

avec $\{r_k\}_{k\geq 0}$ une chaîne de Markov dans un espace d'état fini et $(F^{(i)})_{i\in I}$ un famille de matrices qui correspondent chacune à un type de dynamique usuelle pour le mobile cible. Le cas le plus courant est celui où la cible peut suivre trois régimes de trajectoires :

- le mouvement rectiligne uniforme ;
- le mouvement circulaire uniforme à accélération de giration donnée dans le sens des aiguilles d'une montre ;
- le mouvement circulaire uniforme à accélération de giration donnée dans le sens inverse des aiguilles d'une montre.

La chaîne de Markov $\{r_k\}_{k\geq 0}$ est à valeur dans $\mathsf{S} = \{1, 2, 3\}$. Sa matrice de transition Q_r est choisie en fonction du contexte opérationnel. Une cible manœuvrante qui ne se sait pas observée sera le plus souvent en MRU. Typiquement, on choisit

$$Q_r = \begin{pmatrix} 0,9 & 0,05 & 0,05 \\ 0,4 & 0,5 & 0,1 \\ 0,4 & 0,1 & 0,5 \end{pmatrix} \ ,$$

et les matrices des dynamiques correspondantes sont données par

$$F^{(1)} = \begin{pmatrix} I_2 & \Delta_T I_2 \\ 0 & I_2 \end{pmatrix} \quad \text{et} \quad F^{(j)} = \begin{pmatrix} 1 & 0 & \frac{\sin(\omega_k \Delta_T)}{\omega_k} & -\varepsilon_j \frac{1-\cos(\omega_k \Delta_T)}{\omega_k} \\ 0 & 1 & \varepsilon_j \frac{1-\cos(\omega_k \Delta_T)}{\omega_k} & \frac{\sin(\omega_k \Delta_T)}{\omega_k} \\ 0 & 0 & \cos(\omega_k \Delta_T) & -\varepsilon_j \sin(\omega_k \Delta_T) \\ 0 & 0 & \varepsilon_j \sin(\omega_k \Delta_T) & \cos(\omega_k \Delta_T) \end{pmatrix} \tag{10}$$

pour $j \in \{2,\ 3\}$ avec $\varepsilon_2 = 1$ et $\varepsilon_3 = -1$ et $\omega_k = a_m / \sqrt{\dot{x_k}^2 + \dot{y_k}^2}$ avec a_m une accélération typique de manœuvre.

La stabilité du filtre optimal pour de telles dynamiques semble beaucoup plus délicate. L'état caché est désormais le couple $u = (x, i) \in \mathsf{X} \times \mathsf{S}$. Si l'on peut construire une LD-set fonction pour le noyau de transition de l'état caché, alors, on pourra suivre la même trame que précédemment pour vérifier la stabilité asymptotique.

On a

$$\begin{aligned} Q(u, A \times B) &= \sum_{j \in B} \pi_{i,j} Q^{(j)}(x, A) \ , \\ &= \sum_{j \in B} \pi_{i,j} \int_A \gamma[x' - F^{(j)}(x) \cdot x] \lambda^{\mathrm{Leb}}(\mathrm{d}x') \end{aligned}$$

On va considérer les ensembles $\mathsf{C}_\Delta(y,t) \times \mathsf{S}$. Fixons $j \in \mathsf{S}$. pour tout $(x, x') \in \mathsf{C}_\Delta(y,t) \times \mathsf{C}_\Delta(y',t')$, en considérant des antécédents $z \in \mathsf{K}[O(t)]$ et $z' \in \mathsf{K}[O(t')]$ respectivement pour y par $h(\cdot, t)$ et y' par $h(\cdot, t')$, on a

$$|x' - F^{(j)}(x) \cdot x| \leq |x' - z'| + |z' - F^{(j)}(z) \cdot z| + |F^{(j)}(z) \cdot z - F^{(j)}(x) \cdot x| \ .$$

Or, l'application $x \mapsto F^{(j)}(x) \cdot x$ est de classe $\mathscr{C}^1$ sur le compact $\mathsf{K}[O(t)]$ donc lipschitzienne sur cet ensemble. Il s'ensuit

$$|x' - Fx| \leq (1 + C_t)[\varphi(\Delta) + v_{\max} - v_{\min}] + D^{(j)}[(y,t),(y',t')] \ ,$$

où $D^{(j)}$ est défini par

$$D^{(j)}[(y,t),(y',t')] \stackrel{\text{def}}{=} \sup \left\{ |z' - F^{(j)}(z) \cdot z| : y = h(z,t),\ y' = h(z',t'),\ (z,z') \in \mathsf{K}[O(t)] \times \mathsf{K}[O(t')] \right\} \ .$$

On voit désormais ce qui va poser problème. Il faut pouvoir contrôler $D^{(j)}[(Y_k, t_k), (Y_{k+1}, t_{k+1})]$ en ne faisant intervenir que les termes de perturbation, à savoir les bruits d'état et d'observation. La difficulté vient du fait que l'on ne sait pas de quelle dynamique est extraite une observation Y_k ce qui rend inopérantes les techniques déployées précédemment.

Il n'est pas exclu que l'on puisse adapter les arguments du chapitre 13 pour obtenir les bonnes inégalités qui permettent, sous des hypothèses raisonnables, d'annoncer un résultat de stabilité asymptotique en dynamique manœuvrante. Les arguments précédents illustrent la complexité de ce problème et suggéreront peut-être des pistes pour des travaux futurs.

14.5 Simulations

La convergence des filtres particulaires en trajectographie est un problème ouvert de première importance. Il s'agit d'un point d'étude très difficile pour deux raisons majeures :

– la chaîne des états cachés est non ergodique ;
– la fonction d'observation est très fortement non-linéaire et non-inversible.

Toutefois, la puissance de calcul des ordinateurs actuels permet de réaliser à volonté des simulations de ces algorithmes pour tous les scénarios souhaités. Pour les praticiens, la modélisation par chaîne de Markov cachée sert à traduire la confiance qu'ils ont respectivement dans la dynamique supposée de la cible et dans leurs observations.

Les simulations présentées dans cette section ont été réalisées avec le filtre *Bootstrap*. D'autres filtres plus élaborées, comme le filtre *SIS* (Sequential Importance Sampling), ont été testés sans qu'on observe de différences majeures dans les résultats[5], d'où le choix d'exposer les simulations par filtre Bootstrap qui sont facilement reproductibles. Ces algorithmes sont décrits dans [25], [45]. Pour toutes ces simulations, le nombre de particules est 5000, la durée d'observation est

[5] Le grand nombre d'observations et le faible niveau de bruit de mesure expliquent sans doute pourquoi le gain d'un filtre comme le SIS est mineur.

de $20\,\mathrm{s}$ et la période d'échantillonnage de $0.01\,\mathrm{s}$. Le Le bruit d'observation du filtre est celui des observations réellement acquises à savoir $\sigma = 5 \times 10^{-3}\,\mathrm{rad}$ et le bruit d'état est fixé à $1\,\mathrm{m/s^2}$.

Les résultats de ces simulations sont nuancés et nous invitent à une certaine réserve car les comportements diffèrent de ceux que l'on trouve décrits dans la littérature.Pour chacun des scénarios envisagés, on présente les particules à l'instant initial et les particules à l'instant final.

Sur les figures 14.3 et 14.4, le porteur suit une trajectoire constituée d'une section MRU suivi d'une section MCU puis à nouveau en MRU. Sa vitesse est constante égale à $250\,\mathrm{m/s}$ et l'accélération de manœuvre de $100\,\mathrm{m/s^2}$ (valeur très élevée mais encore plausible). La trajectoire fait donc un coude très marqué ce qui fournit un cadre de trajectographie très favorable (bon niveau d'observabilité, robustesse de l'estimation). On initialise d'abord avec une loi gaussienne centrée en $S(0)$.

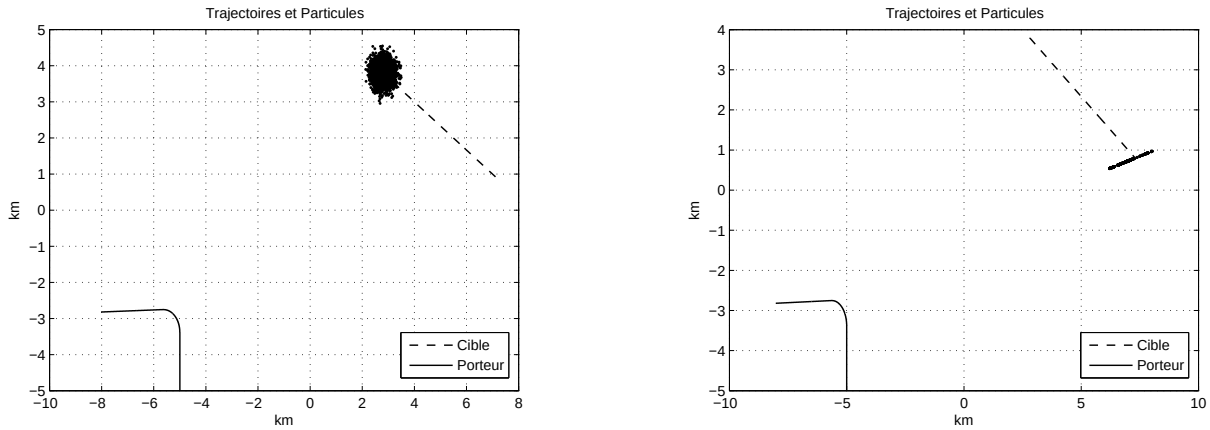


Fig. 14.3 – Cible en MRU, initialisation gaussienne centrée en $S(0)$

On procède ensuite avec une loi uniforme sur une portion de l'axe $(S(0), O(0))$.

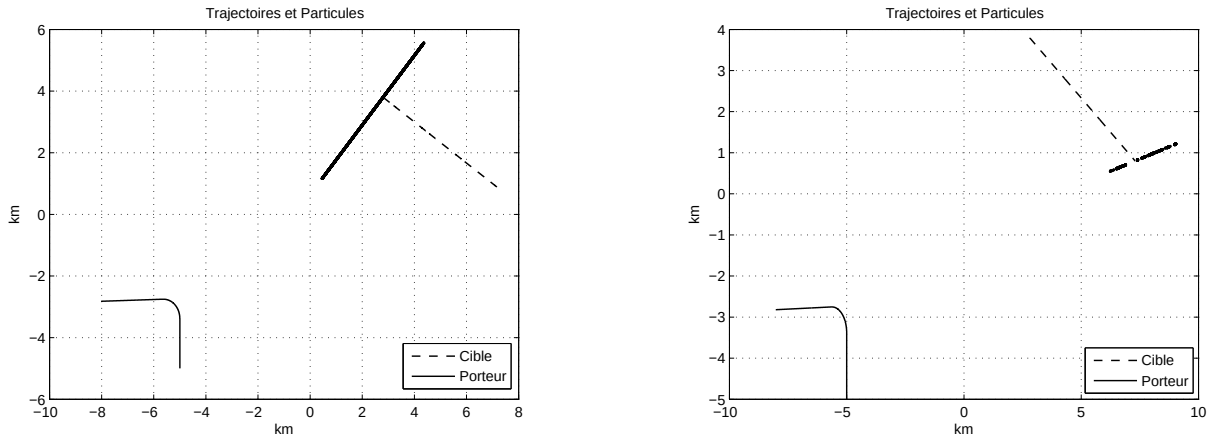


Fig. 14.4 – Cible en MRU, initialisation uniforme sur l'axe $(S(0), O(0))$

Les nuages de particules à l'instant final sont proches pour les deux stratégies d'initialisation ce qui porte à croire qu'on observe la stabilité asymptotique. Cependant, dans un contexte aussi favorable, on pouvait s'attendre à avoir des particules très concentrées autour de la position réelle de la cible ce qui n'est pas véritablement le cas.

Sur les figures 14.5 et 14.6, le porteur suit une trajectoire MRU-MCU par morceaux à vitesse constante égale à $250\,\mathrm{m/s}$ et avec des accélérations de manœuvre de l'ordre de $50\,\mathrm{m/s^2}$. On initialise d'abord avec une loi gaussienne centrée en $S(0)$.

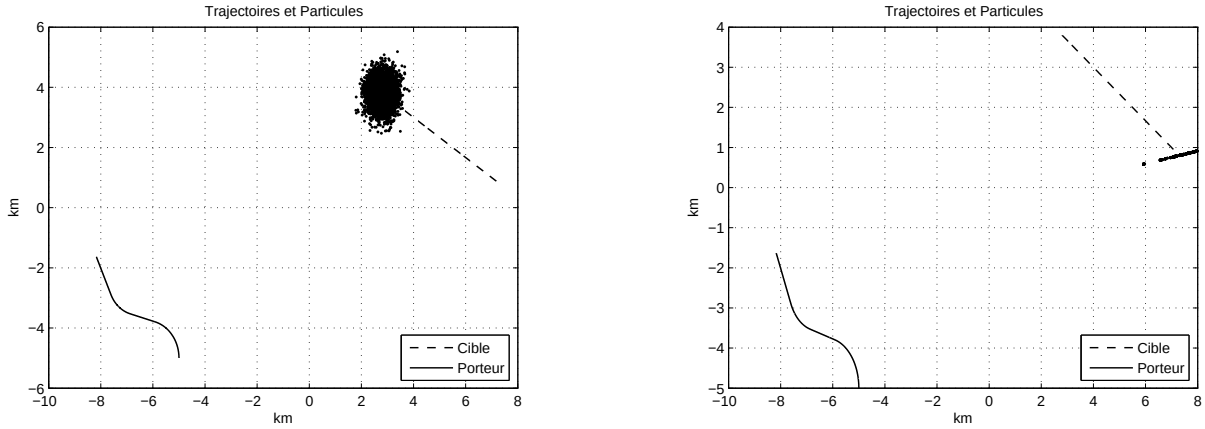


Fig. 14.5 – Cible en MRU, initialisation gaussienne centrée en $S(0)$

On procède ensuite avec une loi uniforme sur une portion de l'axe $(S(0), O(0))$.

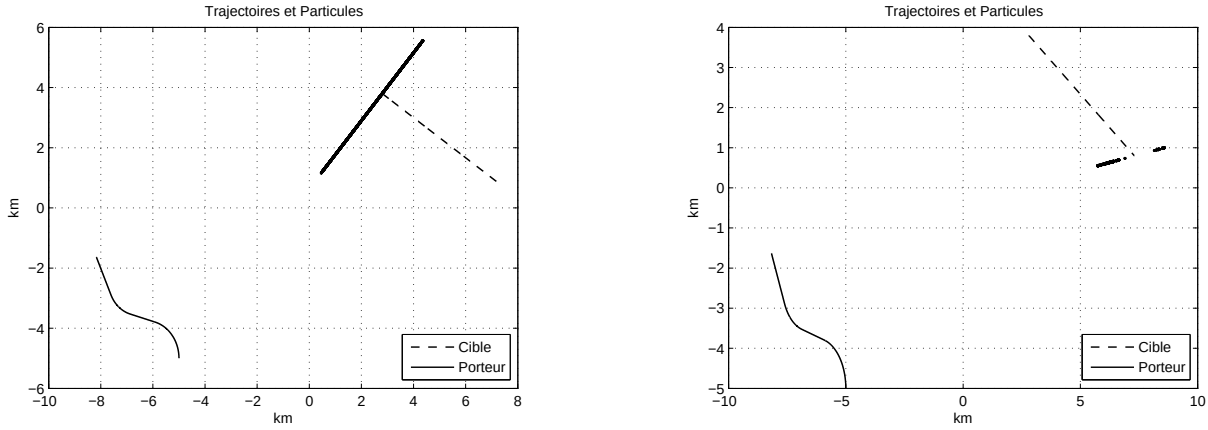


Fig. 14.6 – Cible en MRU, initialisation uniforme sur l'axe $(S(0), O(0))$

La stabilité asymptotique est moins flagrante pour cette simulation bien que les conditions de la simulation soient encore très favorables.

Pour les simulations suivantes, le porteur suit la même trajectoire que celle des simulations figure 14.5 et 14.6.

Sur les figures 14.7 et 14.8, la cible sui un MCU de courbure faible à vitesse constante égale à $270\,\mathrm{m/s}$. On initialise d'abord avec une loi gaussienne centrée en $S(0)$.

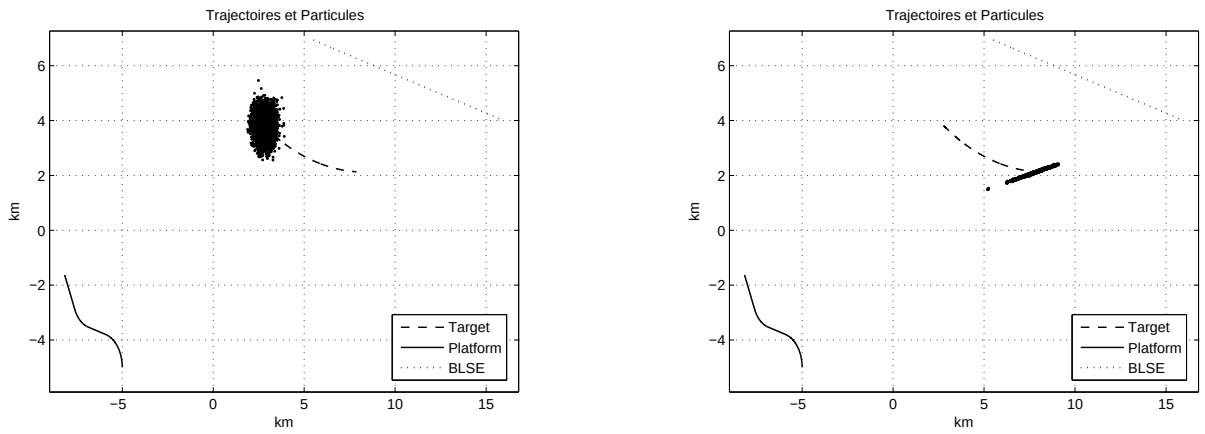


FIG. 14.7 – Cible en MRU, initialisation gaussienne centrée en $S(0)$

On initialise ensuite avec une loi uniforme entre la position réelle $S(0)$ et celle donnée par l'estimation $S_{\bar{\theta}}(0)$.

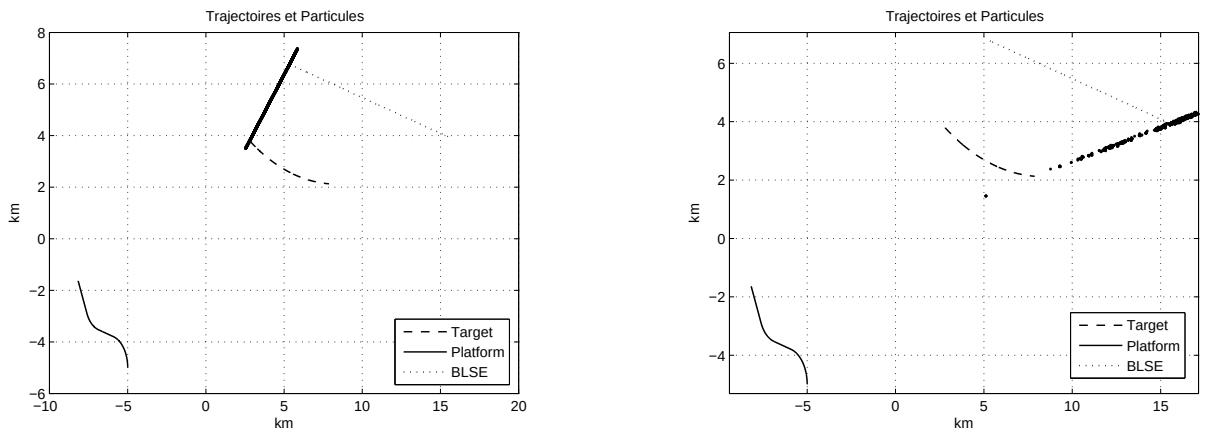


FIG. 14.8 – Cible en MRU, initialisation uniforme sur l'axe $(S(0), O(0))$

Il est difficile de se prononcer à la vue des résultats. On n'observe pas la stabilité du filtre, peut être à cause du caractère asymptotique de cette propriété.

Sur les figures 14.9 et 14.10, la cible suit une trajectoire constituée d'une section MCU puis d'une section MRU. On initialise d'abord avec une loi gaussienne centrée en $S(0)$.

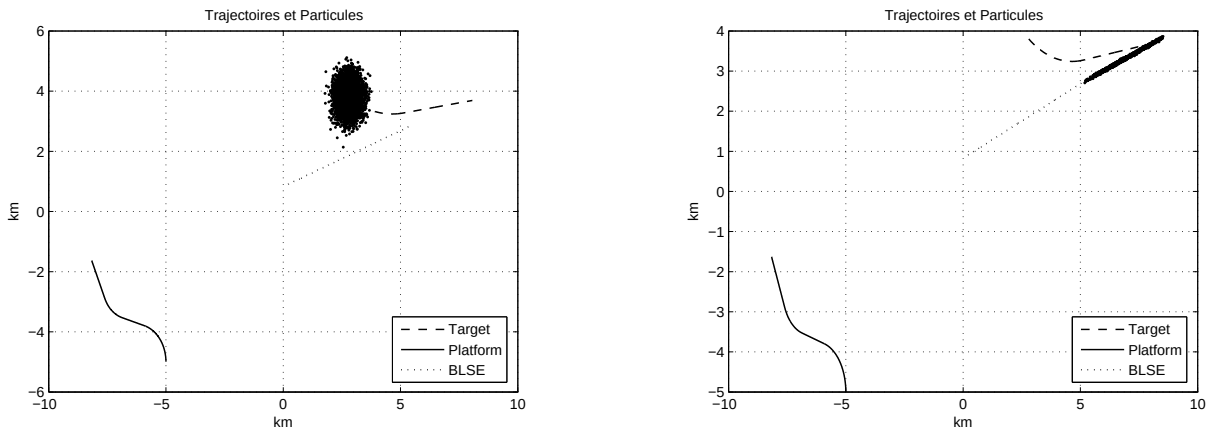


FIG. 14.9 – Cible en MRU, initialisation gaussienne centrée en $S(0)$

On procède ensuite avec une loi uniforme sur une portion de l'axe $(S(0), O(0))$.

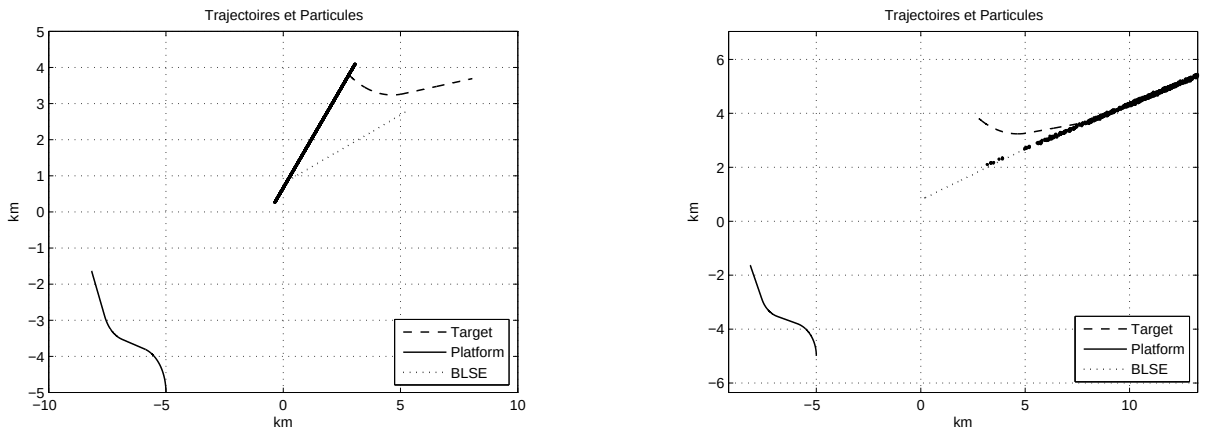


FIG. 14.10 – Cible en MRU, initialisation uniforme sur l'axe $(S(0), O(0))$

Là encore, il est difficile de conclure quoi que ce soit.

Dans ces cas où il est délicat d'affirmer qu'il y a ou pas convergence des filtres particulaires, on peut utiliser les résultats comme des indicateurs de non-robustesse pour les autres méthodes d'estimation.

14.6 Perspectives

La stabilité du filtre optimal est établie pour un modèle non-manœuvrant HMM de trajectographie passive sous une hypothèse réaliste sur un plan opérationnel. Ce résultat ouvre la voie pour envisager d'autres approches applicables à des modèles de dynamiques manœuvrantes mais aussi pour des noyaux de transition plus généraux. L'autre chantier à venir est sans nul doute la preuve de la convergence, en horizon temporel infini, des filtres particulaires en trajectographie passive.

Cinquième partie

Conclusion

À son origine, l'ambition de ce travail était simultanément la création d'algorithmes de pistage innovants et performants pour des cibles manœuvrantes et la production de résultats théoriques afférents à ces méthodes. La spécificité du problème de trajectographie passive motivé par la société Thales Optronique, à savoir un contexte optronique avec des mesures de grande précision acquises à cadence élevée, laissait croire que des méthodes habituellement délaissées en pratique comme le sont les méthodes *batch*, pourraient, dans ce contexte, apporter des résultats satisfaisants. Les techniques de sélection de modèles, initialement perçues comme très prometteuses, se sont avérées quasi impraticables pour les raisons détaillées dans le chapitre 9.

Cependant, l'étude de robustesse dans les situations de faible observabilité constitue une avancée importante dans la compréhension du problème. Elle a permis de mettre en avant le caractère particulièrement mal posé de la trajectographie passive. Par ailleurs, cette étude de robustesse a un impact majeur pour l'industriel qui s'engage sur une prestation de trajectographie passive. Spontanément, on pourrait croire que l'estimation d'un MRU sur une trajectoire proche d'un MRU à partir de ses mesures angulaires va renvoyer une trajectoire voisine de l'approximation des moindres carrés. On sait désormais que la prudence est de mise pour toute technique d'estimation ou de filtrage hors des modèles paramétriques. Ceci est d'une importance cruciale, à la fois pour cibler les fonctionnalités qu'un industriel peut proposer à ses clients, mais aussi pour décider de l'orientation des études amont. L'étude de robustesse laisse de nombreuses questions ouvertes et notamment l'une des plus immédiates : peut-on décider que l'estimation est robuste ou pas ? Des pistes sont évoquées dans les chapitres 10 et 14 qui présentent l'utilisation des filtres particulaires comme indicateur de robustesse. Pour les situations d'estimation non robustes, on constate l'étalement des filtres entre la trajectoire réelle et celle estimée. Ce phénomène pourrait faire l'objet d'une étude approfondie et servir, en temps réel ou pas, de critère de confiance. On peut également envisager son utilisation pour décider d'une commande sur le porteur pour améliorer la robustesse des filtres et de l'estimation. Dans cette thèse, on ne s'est pas penché, lors de l'étude de robustesse, sur la dépendance en la trajectoire porteur. Mais peut-être peut-on envisager d'identifier clairement des classes de trajectoires qui amènent à des situations non-robustes, ou au contraire à des situations robustes. Le lien intrinsèque entre observabilité et robustesse peut sans doute être précisé davantage. Si l'équivalence entre la qualité de l'observabilité et la robustesse était avérée, on aurait une convergence des travaux de commande optimale pour l'observabilité et la robustesse et peut être verrait-on l'émergence de nouvelles stratégies de commande.

La technique de reconstruction des trajectoires à vitesse constante par un schéma d'intégration présentée au chapitre 6 pourrait être envisagée comme une stratégie d'estimation de la trajectoire cible. Le problème d'une telle approche est bien évidemment son coût algorithmique puisqu'il faut balayer l'ensemble des conditions initiales et pour chaque jeu de conditions, construire toute la trajectoire. Si l'on parvient à déterminer les caractéristiques pertinentes d'une trajectoire conforme aux mesures angulaires, à savoir les non-simultanéités de changement de régime avec le porteur, sans avoir à la déterminer complètement, alors on disposerait d'un critère pratique utilisable pour discriminer les trajectoires et ne retenir que les candidates les plus probantes. Une telle stratégie pourrait être couplé à d'autres techniques : on pourrait mélanger des critères type *batch* avec des algorithmes de filtrage.

Avec l'approche semi-paramétrique développée au chapitre 11, on introduit un problème d'estimation d'un genre nouveau. Ce travail confirme un des phénomène décrits lors de l'étude de robustesse : pour des perturbations isotropes hautes fréquences, l'estimateur des moindres carrés est satisfaisant. L'intérêt d'une telle étude est évident : pour une trajectoire cible qui vérifie les

caractéristiques mentionnées ci-avant, on dispose d'un estimateur aux bonnes propriétés asymptotiques pour une complexité algorithmique très raisonnable. Par ailleurs, même si la loi du bruit est inconnue, on dispose d'une région de confiance conservative, information particulièrement pertinente dans un contexte opérationnel. Des pistes vers des modèles de bruit d'état corrélé sont proposées dans le chapitre 11 et pourront faire l'objet d'études plus poussées et d'expérimentations numériques. On pourrait par exemple s'intéresser à considérer un bruit d'état modélisé par une série temporelle et étudier le comportement des estimateurs des moindres carrés et du maximum de vraisemblance en fonction des caractéristiques de la fonction d'auto-corrélation. Les questions ouvertes sur ce modèle semi-paramétrique sont multiples. Citons entre autres la détermination de l'information efficace, la construction d'un estimateur asymptotiquement efficace, le problème d'estimation non-paramétrique du bruit d'état, l'assouplissement des hypothèses de bruit d'état pour l'estimateur des moindres carrés. Par ailleurs, bien que conçu pour le BOT, il est vraisemblable que ce type de modèle trouvera d'autres champs d'applications et sans doute dans des contextes plus favorables que la trajectographie mono-porteur avec, par exemple, une fonction d'observation inversible, linéaire, situations pour lesquelles une étude semi-paramétrique plus poussée pourrait être menée et éventuellement avec des approches bayésiennes.

Enfin, le travail effectué sur les chaînes de Markov cachées non ergodiques illustre remarquablement l'intérêt commun du monde industriel et du milieu académique dans l'étude des problématiques appliquées. La question de la stabilité asymptotique, initialement motivée pour son application à la trajectographie passive, se trouve être une question complètement d'actualité dans la communauté du filtrage. Les résultats obtenus au chapitre 13 apportent une contribution d'importance à l'étude du filtre optimal et ouvrent la porte pour d'autres questions délicates comme la convergence en horizon temporel infini des filtres particulaires ou les propriétés asymptotiques d'estimateur dans des modèles de HMM non ergodiques.

Sixième partie

Annexe

Régularité de la fonction angle

La fonction angle est définie, pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, par

$$\mathrm{angle}(x,y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \mathrm{sgn}(x) \times \mathbb{1}_{y<0} & \text{si } x \neq 0 \text{ et } y \neq 0, \\ \frac{\pi}{2} \times \mathrm{sgn}(x) & \text{si } x \neq 0 \text{ et } y = 0, \\ \frac{\pi}{2} \times [1 - \mathrm{sgn}(y)] & \text{si } x = 0 \text{ et } y \neq 0, \end{cases}$$

où sgn représente la fonction signe. Pour un couple (x, y) de réels non simultanément nuls, la quantité $\mathrm{angle}(x, y)$ représente dans le plan l'angle à valeur dans $]-\pi, \pi]$ entre l'axe (Oy) et l'axe (OM) où M est le point de coordonnées (x, y).

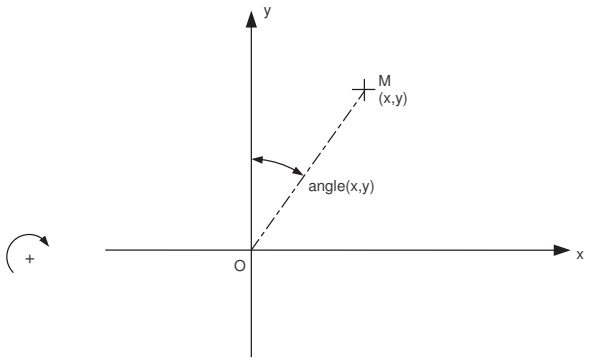


Fig. 14.11 – Fonction angle

On peut vérifier que la fonction angle est de classe $\mathscr{C}^\infty$ sur l'ouvert $\mathbb{R}^2 \setminus \{(0, y), y < 0\}$. En particulier, son vecteur gradient est défini par

$$\nabla_{(x,y)}\mathrm{angle} = \left(\frac{y}{x^2+y^2}, \quad \frac{-x}{x^2+y^2}\right) ,$$

et sa matrice hessienne

$$\nabla^2_{(x,y)}\mathrm{angle} = \frac{1}{(x^2+y^2)^2} \begin{pmatrix} -2xy & x^2 - y^2 \\ x^2 - y^2 & 2xy \end{pmatrix} .$$

Bibliographie

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards Applied Mathematics Series. 1964.

[2] V.J. Aidala. Kalman filter behavior in bearings-only tracking applications. *IEEE Trans. Aerospace Elect. Systems*, 15(1) :29–39, 1979.

[3] V.J. Aidala and S.C. Nardone. Observability criteria for bearings-only target motion analysis. *IEEE Trans. Aerospace Elect. Systems*, 17(2) :161–166, 1981.

[4] V.J. Aidala and S.C. Nardone. Biased Estimation Properties of the Pseudolinear Tracking Filter. *IEEE Transactions on Aerospace Electronic Systems*, 18 :432–441, July 1982.

[5] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, 1979.

[6] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.*, 50 :241–254, 2002.

[7] S. Arulampalam, B. Ristic, N. Gordon, and T. Mansell. Bearings-only tracking of manoeuvring targets using particle filters. *EURASIP J. Appl. Signal Process.*, 2004(1) :2351–2365, 2004.

[8] R. Atar and O. Zeitouni. Exponential stability for nonlinear filtering. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(6) :697–725, 1997.

[9] A. Averbuch, Y. Bar-Shalom, E. Mazor, and J. Dayan. Interacting multiple model methods in target tracking : A survey. *IEEE Trans. Aerospace Elect. Systems*, 34(1) :103–123, January 1998.

[10] Y. Bar-Shalom and T.E. Fortmann. *Tracking and Data Association*. Academic Press, 1998.

[11] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. Wiley-Interscience, 2001.

[12] J. S. Baras, A. Bensoussan, and M. R. James. Dynamic observers as asymptotic limits of recursive filters : special cases. *SIAM J. Appl. Math.*, 48(5) :1147–1158, 1988.

[13] L. E. Baum and T. P. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statist.*, 37 :1554–1563, 1966.

[14] L. E. Baum, T. P. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.*, 41(1) :164–171, 1970.

[15] V.E. Beneš and I. Karatzas. Estimation and control for linear, partially observable systems with non-Gaussian initial distribution. *Stochastic Process. Appl.*, 14 :233–248, 1983.

[16] N. Bergman, U. Forssell, F. Gunnarsson, F. Gustafsson, R. Karlsson, J. Jansson, and P.-J. Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Trans. Signal Process.*, 50(2) :425–437, 2002.

[17] S.S. Blackman and R.F. Popoli. *Modern Tracking Systems*. Artech House, 1999.

[18] H.A.P. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Automat. Control*, 33 :780–783, 1988.

[19] M. Boutayeb, H. Rafaralahy, and M. Darouach. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Trans. Automat. Control*, 42(4) :581–586, 1997.

[20] Th. Bréhard and J.-P. Le Cadre. Closed-form posterior Cramèr-Rao bound for bearings-only tracking. *IEEE Trans. Aerospace Elect. Systems*, 42(4) :1198–1223, 2006.

[21] Th. Bréhard and J.-P. Le Cadre. Hierarchical particle filter for bearings-only tracking. *IEEE Trans. Aerospace Elect. Systems*, 43(4) :1567–1585, 2007.

[22] R. Bucy and R.E. Kalman. New results in linear filtering and prediction theory. *J. Basic Eng., Trans. ASME, Series D*, 83(3) :95–108, 1961.

[23] A. Budhiraja and D. Ocone. Exponential stability of discrete-time filters for bounded observation noise. *Systems Control Lett.*, 30 :185–193, 1997.

[24] A. Budhiraja and D. Ocone. Exponential stability in discrete-time filtering for non-ergodic signals. *Stochastic Process. Appl.*, 82(2) :245–257, 1999.

[25] O. Cappé, S.J. Godsill, and E.Moulines. An overview of existing methods and recent advances in sequential Monte Carlo. *Proc. IEEE*, 95(5), May 2007.

[26] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer, 2005.

[27] I. Castillo. A semi-parametric Bernstein-von Mises theorem. 2008. Submitted.

[28] P. Chigansky, R. Liptser, and R. van Handel. Intrinsic methods in filter stability. In D. Crisan and B. Rozovsky, editors, *Handbook of Nonlinear Filtering*. Oxford University Press, 2009.

[29] C. K. Chui and G. Chen. *Kalman filtering with real-time applications*. Springer-Verlag New York, Inc., New York, NY, USA, 1987.

[30] G. Claeskens and N.L. Hjort. *Model Selection and Model Averaging*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge Press, 2009.

[31] J.B. Conway. *A course in functional analysis*. Springer, 1990. Graduate Texts in Mathematics 96.

[32] D. Crisan and K. Heine. Stability of the discrete time filter in terms of the tails of noise distributions. *J. London Math. Soc.*, 78(2) :441–458, 2008.

[33] D. Crisan and K. Heine. Uniform approximations of discrete-time filters. *Adv. in Appl. Probab.*, 40(4) :979–1001, 2008.

[34] P. Davidson and Oshman Y. Optimization of observer trajectories for bearings-only target localization. *IEEE Trans. Aerospace Elect. Systems*, 35(3) :892–902, July 1999.

[35] P. Del Moral. *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Springer, 2004.

[36] P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'Institut Henri Poincaré*, 37 :155–194, 2001.

[37] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV*, volume 1729 of *Lecture Notes in Math.*, pages 1–145. Springer, Berlin, 2000.

[38] B. Delyon and O. Zeitouni. Lyapunov exponents for filtering problem. In M. H. A. Davis and R. J. Elliot, editors, *Applied Stochastic Analysis*, pages 511–521. Gordon and Breach Science Publishers, London, 1991.

[39] K. Dogancay. On the bias of linear least squares algorithms for passive target localization. *Signal Processing*, 84(3) :475–486, March 2004.

[40] R. Douc, G. Fort, E. Moulines, and P. Priouret. Forgetting the initial distribution for hidden markov models. *Stochastic Processes and their Applications*, In Press, Corrected Proof :–, 2008.

[41] R. Douc, E. Moulines, and T. Rydén. Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime. *Ann. Statist.*, 32(5) :2254–2304, 2004.

[42] A. Doucet and Johansen A.M. A tutorial on particle filtering and smoothing : Fifteen years later. In D. Crisan and B. Rozovsky, editors, *Handbook of Nonlinear Filtering*. Oxford University Press, 2009.

[43] A. Doucet, N. De Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer, New York, 2001.

[44] A. Doucet, R.J. Evans, N. Kantas, S.S. Singh, and B.N. Vo. Simulation-based optimal sensor scheduling with application to observer trajectory planning. *Automatica*, 43(5) :817–830, 2007.

[45] A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte-Carlo sampling methods for Bayesian filtering. *Stat. Comput.*, 10 :197–208, 2000.

[46] K. Doğançay. On the efficiency of a bearings-only instrumental variable estimator for target motion analysis. *Signal Process.*, 85(3), 2005.

[47] P. Fearnhead. *Sequential Monte Carlo methods in filter theory*. PhD thesis, University of Oxford, 1998.

[48] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon. Sur les systèmes non linéaires différentiellement plats. *C.R. Acad. Sci. Paris*, I(315) :619–624, 1992.

[49] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon. Linéarisation par bouclage dynamique et transformations de Lie-Bäcklund. *C.R. Acad. Sci. Paris*, I(317) :981–986, 1993.

[50] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon. Flatness and defect of nonlinear system : introductory theory and examples. *Int. J. Control*, 61(6) :1327–1361, 1995.

[51] E. Fogel and M. Gavish. nth-order dynamics target observability from angle measurements. *IEEE Trans. Aerospace Elect. Systems*, 24(4) :305–308, 1988.

[52] A.R. Gallant. *Nonlinear Statistical Models*. John Wiley & Sons, 1987.

[53] A.R. Gallant and H. White. *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell, 1988.

[54] S.J. Godsill, J. Vermaak, K-F. Ng, and J-F. Li. Models and algorithms for tracking of manoeuvring objects using variable rate particle filters. *Proc. IEEE*, 95(5), May 2007.

[55] K. F. Gong and A. G. Lindgren. Position and velocity estimation via bearing observations. *IEEE Trans. Aerospace Elect. Systems*, 14 :564–577, July 1978.

[56] N. Gordon, D. Salmond, and A. F. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F, Radar Signal Process.*, 140 :107–113, 1993.

[57] E. Guerre and P. Lavergne. Data-driven rate-optimal specification testing in regression models. *Ann. Statist.*, 33(2) :840–870, 2005.

[58] J. Hadamard. *Lecture on the Cauchy Problem in Linear Partial Differential Equations*. Yale University Press, 1923.

[59] Joel L. Horowitz and Vladimir G. Spokoiny. An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, 69(3) :599–631, 2001.

[60] Z.-C. Hu and W. Sun. A note on exponential stability of the nonlinear filter for denumerable Markov chains. *Systems Control Lett.*, 55(11) :955–960, 2006.

[61] I. A. Ibragimov and R. Z. Hasminskii. *Statistical Estimation. Asymptotic Theory*. Springer, New York, 1981.

[62] I.A. Ibragimov. Some limit theorems for stationary processes. *Theory of probability and its applications*, 7 :349–382, 1962.

[63] C. Jauffret. Observabilité et inversibilité de la matrice d'information de Fisher dans les problèmes de régression non-linéaire. In *Dix-septième colloque GRETSI*, pages 685–688, Vannes, France, September 1999.

[64] C. Jauffret and J.-P. Le Cadre. Discrete-time observability and estimability for bearings-only target motion analysis. *IEEE Trans. Aerospace Elect. Systems*, 33(1) :178–201, January 1997.

[65] C. Jauffret and D. Pillon. Observability in passive target motion analysis. *IEEE Trans. Aerospace Elect. Systems*, 32(4) :1290–1300, October 1996.

[66] C. Jauffret and D. Pillon. Trajectographie passive par mesure d'angle. (TE6705), 2005.

[67] R.I. Jennrich. Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.*, 40(2) :633–643, 1969.

[68] M. Joannides and F. Le Gland. Small noise asymptotics of the Bayesian estimator in nonidentifiable models. *Stat. Inference Stoch. Process.*, 5(1) :95–130, 2002.

[69] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *AeroSense : The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997.

[70] R.E. Kalman. A new approach to linear filtering and prediction theory. *J. Basic Eng., Trans. ASME*, 82 :34–45, 1960.

[71] R. Karlsson and F. Gustafsson. Recursive Bayesian estimation : bearings-only applications. *IEE Proceedings, Radar, Sonar and Navigation*, 152(5) :305–313, 2005.

[72] A. Kirsch. *An Introduction to the Mathematical Theory on Inverse Problems*, volume 120. Springer-Verlag, 1996.

[73] M. L. Kleptsyna and A. Yu. Veretennikov. On discrete time ergodic filters with wrong initial data. *Probab. Theory Related Fields*, 141(3-4) :411–444, 2008.

[74] Arthur J. Krener. The convergence of the extended Kalman filter. In *Directions in mathematical systems theory and optimization*, volume 286 of *Lecture Notes in Control and Inform. Sci.*, pages 173–182. Springer, Berlin, 2003.

[75] H. R. Künsch. Recursive Monte-Carlo filters : algorithms and theoretical analysis. *Ann. Statist.*, 33(5) :1983–2021, 2005.

[76] J.-P. Le Cadre and S. Laurent-Michel. Optimizing the receiver maneuvers for bearings-only tracking. *Automatica*, 35(4) :591–606, April 1999.

[77] L. Le Cam. *Asymptotic methods in statistical decision theory.* Springer Series in Statistics. New York etc. : Springer-Verlag. XXVI, 742 p. DM 138.00, 1986.

[78] F. Le Gland and N. Oudjane. A robustification approach to stability and to uniform particle approximation of nonlinear filters : the example of pseudo-mixing signals. *Stochastic Process. Appl.*, 106(2) :279–316, 2003.

[79] F. Le Gland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Ann. Appl. Probab.*, 14 :144–187, 2004.

[80] X. R. Li and V. P. Jilkov. Survey of maneuvering target tracking-part i : dynamic models. *IEEE Trans. Aerospace Elect. Systems*, 39(4) :1333–1364.

[81] R. S. Lipster and A. N. Shiryaev. *Statistics of Random Processes : I. General theory*. Springer, 2nd edition, 2001.

[82] R. S. Lipster and A. N. Shiryaev. *Statistics of Random Processes : II.* Springer, 2nd edition, 2001.

[83] P.T. Liu. An optimum approach in target tracking with bearing measurements. *Journal of Optimization Theory and Applications*, 56(2) :205–214, 1988.

[84] A.M. Makowski. Filtering formulae for partially observed linear systems with non-gaussian initial conditions. *Stochastics*, 16 :1–24, 1986.

[85] A.M. Makowski and Sowers R.B. Discrete time filtering for linear systems with non-gaussian initial conditions : Asymptotic behaviors of the difference between the MMLSE and the LMSE estimates. *IEEE Trans. Automat. Control*, 37 :114–121, 1992.

[86] P. Massart. *Concentration Inequalities and Model Selection*. Springer, 2003.

[87] Brad McNeney and Jon A. Wellner. Application of convolution theorems in semiparametric models with non-i.i.d. data. *J. Statist. Plann. Inference*, 91(2) :441–480, 2000.

[88] Hans-Georg Müller. Change-points in nonparametric regression analysis. *Ann. Statist.*, 20(2) :737–761, 1992.

[89] C.H. Muravchik, A. Nehorai, and P. Tichavsky. Posterior Cramer-Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.*, 45(5) :1386–1396, 1998.

[90] C. Musso, N. Oudjane, and F.Le Gland. Improving regularized particle filters. In N. de Freitas A. Doucet and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, pages 247–271. Springer–Verlag, 2001.

[91] Michael H. Neumann. Optimal change-point estimation in inverse problems. *Scand. J. Statist.*, 24(4) :503–521, 1997.

[92] D. Ocone and E. Pardoux. Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM J. Control*, 34 :226–243, 1996.

[93] N. Oudjane and S. Rubenthaler. Stability and uniform particle approximation of nonlinear filters in case of non ergodic signals. *Stoch. Anal. Appl.*, 23(3) :421–448, 2005.

[94] A. Papavasiliou. A uniformly convergent adaptive particle filter. *J. Appl. Probab.*, 42(4) :1053–1068, 2005.

[95] J.M. Passerieux and D. Van Cappel. Optimal observer maneuver for bearings-only tracking. *IEEE Trans. Aerospace Elect. Systems*, 34(3) :777–888, July 1998.

[96] R.C. Penney. *Linear algebra : ideas & applications*. Lavoisier, 2007.

[97] Jean Picard. Efficiency of the extended Kalman filter for nonlinear systems with small noise. *SIAM J. Appl. Math.*, 51(3) :843–885, 1991.

[98] Marc Raimondo. Minimax estimation of sharp change points. *Ann. Statist.*, 26(4) :1379–1397, 1998.

[99] E. Rio. About the Lindeberg method for strongly mixing sequences. *ESAIM : Probability and Statistics*, 1 :35–61, 1995.

[100] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*. Mathématiques et Applications. Springer, 2000.

[101] B. Ristic, S. Arulampalam, and A. Gordon. *Beyond Kalman Filters : Particle Filters for Target Tracking*. Artech House, 2004.

[102] T.L. Song. Observability of target with bearings-only measurements. *IEEE Trans. Aerospace Elect. Systems*, 32(4) :1468–1472, 1996.

[103] E.D. Sontag. *Mathematical Control Theory, Deterministic Finite Dimensional Systems*. 2nd edition, 1998.

[104] A.N. Tikhonov. Regularization of incorrectly posed problems. *Soviet. Math. Dokl.*, (4) :1624–1627, 1963.

[105] A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet. Math. Dokl.*, (4) :1035–1038, 1963.

[106] O. Trémois. *Étude de méthodes de trajectographie pour des sources manœuvrantes*. PhD thesis, Université de Rennes I, 1995.

[107] O. Trémois and J.-P. Le Cadre. Optimal observer trajectory in bearings-only tracking for maneuvering sources. *IEE Proceedings Radar, Sonar and Navigation*, 146(1) :31–39, February 1999.

[108] A.B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, 2009.

[109] S.A. Van de Geer. *Empirical Processes in M-Estimation*. Cambridge University Press, 1999.

[110] A. W. Van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

[111] R. van Handel. Discrete time nonlinear filters with informative observations are stable. *Electr. Comm. Probab.*, 13 :562, 2008.

[112] R. van Handel. Uniform observability of hidden markov models and filter stability for unstable signals, 2008.

[113] R. van Handel. Uniform time average consistency of monte carlo particle filters. 2009. submitted.

[114] H.L. VanTrees. *Detection, Estimation and Modulation Theory*, volume I. 1968.

[115] A. Yu. Veretennikov. Coupling method for Markov chains under integral Doeblin type condition. In *Proceedings of the Conference Dedicated to the 90th Anniversary of Boris Vladimirovich Gnedenko (Kyiv, 2002)*, volume 8, pages 383–390, 2002.

[116] Aidala V.J. and Hammel S.E. Utilization of modified polar coordinates for bearings-only tracking. *IEEE Trans. Aerospace Elect. Systems*, 28(3) :283–294, 1983.

[117] A. Wald. Asymptotic properties of the maximum likekihood estimate of an unknown parameter of a discrete stochastic process. *Ann. Math. Statist.*, 19 :40–46, 1948.

Résumé

Cette thèse présente une étude statistique du problème de la trajectographie passive. On s'intéresse dans une première partie à la question de l'observabilité pour des trajectoires paramétriques puis paramétriques par morceaux et ensuite des trajectoires à vitesse constante. La deuxième partie est consacrée à l'estimation : on présente les propriétés de l'estimateur du maximum de vraisemblance pour des trajectoires paramétriques et paramétriques par morceaux. On expose également le caractère non robuste de cette estimation en dépit de propriétés asymptotiques satisfaisantes. On s'intéresse alors à la sensibilité de l'estimation quand le modèle d'état n'est pas totalement spécifié. Son comportement est décrit pour des perturbations d'état déterministes puis aléatoires et un cadre semiparamétrique est considéré quand la loi du bruit d'état est inconnue. Dans la dernière partie, on aborde le problème de la trajectographie passive comme chaîne de Markov cachée. On s'intéresse à l'étude du filtre optimal et à sa résolution par des méthodes algorithmiques. Le filtre de Kalman étendu est expérimenté sous différentes conditions de bruit d'état. On présente ensuite des résultats de stabilité asymptotique du filtre optimal pour des chaînes de Markov cachées non ergodiques puis leur application en trajectographie passive.

Mots-clés : trajectographie passive, chaîne de Markov cachée, filtrage non linéaire, filtres particulaires, régression semiparamétrique, robustesse

Abstract

This thesis presents a statistical study of the bearings-only tracking problem. In a first part, we deal with the question of observability for parametric and piecewise parametric trajectories, and then for constant speed trajectories. The second part is dedicated to estimation : we give the properties of the maximum likelihood estimator for parametric and piecewise parametric trajectories. We also show the non-robust behaviour of this estimation, in spite of its good asymptotic properties . The sensitivity of estimation is then investigated when the state model is not completely specified. The cases of deterministic and stochastic state noise are considered and a semiparametric framework is used when the law of the state noise is unknown. In the last part, we consider the bearings-only tracking problem as a hidden Markov chain model. We focus on the study of the optimal filter and on its approximation by algorithmic methods. The extended Kalman filter is tested under different state noise conditions. We finally give asymptotic stability results for non-ergodic hidden Markov chain models and their application to bearings-only tracking.

Key-words : bearings-only tracking, hidden Markov chain model, nonlinear filtering, particle filters, semiparametric regression, robustness

AMS Classification : 60G35, 62C10, 62J02, 62F12, 62F25, 62P30, 93B07, 93E10, 93E11