Coupled methods of nonlinear estimation and control applicable to terrain-aided navigation

Emilien Flayac

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

ENSTA IP PARIS

THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription : École nationale supérieure de techniques avancées

Établissement d'accueil : Office national d'études et de recherches aérospatiales

Laboratoire d'accueil : Unité de mathématiques appliquées, ENSTA-CNRS-INRIA

Spécialité de doctorat : Mathématiques appliquées

Emilien FLAYAC

Coupled methods of nonlinear control and estimation applicable to Terrain-Aided Navigation

Jury de soutenance :

DANIEL KUHN	(Ecole Polytechnique Fédérale de Lausanne) Rapporteur
SILVÈRE BONNABEL	(MINES Paristech) Rapporteur
FRANÇOIS LE GLAND	(INRIA Rennes) Président du Jury
TAREK HAMEL	(Université Nice Sophia Antipolis) Examinateur
PIERRE CARPENTIER	(ENSTA Paris) Examinateur
FRÉDÉRIC JEAN	(ENSTA Paris) Directeur de Thèse
KARIM DAHIA	(ONERA) Encadrant
BRUNO HÉRISSÉ	(ONERA) Encadrant

Fondation mathématique FMJH Jacques Hadamard

ONERA THE FRENCH AEROSPACE LAB

Acknowledgements

Je souhaiterais remercier toutes les personnes qui m'ont permis de réaliser cette thèse dans de très bonnes conditions, à commencer par Silvère Bonnabel et Daniel Kuhn qui ont accepté d'être mes rapporteur de thèse. Je remercie aussi Tarek Hamel, François Le Gland et Pierre Carpentier de bien avoir voulu être membres de mon jury de thèse.

Je voudrais exprimer ma sincère gratitude envers mes deux encadrants de thèse de l'ONERA, Bruno Hérissé et Karim Dahia qui m'ont soutenu et supporté pendant trois ans, ce qui ne fut pas de tout repos à mon avis. Cette gratitude s'étend bien sûr à mon directeur de thèse, Frédéric Jean, qui a toujours su m'apporter ses conseils et ses critiques aux bons moments depuis qu'il a commencé à suivre mon parcours.

Je me dois de mentionner mes amis de promotion de l'ENSTA que je ne cesserai d'épargner et qui en feront toujours de même, à juste titre. En particulier, bien qu'il soit docteur en Automatique, je remercie Benoît Bonnet pour son oreille scientifique et humaine attentive et toujours pertinente. J'ai une pensée impériale et impérieuse pour Florian Enaud, grand stratège du hasard de notre amitié. J'aime penser que Thomas Bourguenolle a rejoint ce monde pour compenser mon manque de "skill" et moi son manque de "fail". Je dois aussi parler d'Hadrien Godard alias "Banach" qui, faute d'avoir les compétences d'un mathématicien, en a le surnom. Je finirai cette liste par Rémi Lucas, passionné par ce qui n'est pas passionnant et par Rémi Feuillet adorateur des maillages, de René Coty et des maillages de René Coty.

Parmi les doctorants de l'ONERA que j'ai cotoyés pendant trois ans, qui m'ont aidé et que, plus rarement, j'ai aidé, je remercie Vincent, Iréna, Romain, Nicolas, i.e. les très vieux, puis Nathan, Léon, Sergio, Julien et Camille la grande, i.e. les vieux comme moi, et enfin les stagiaires ou assimilés, comme Ali, Esteban, Etienne, Rodolphe, Guillaume, Camille la jeune, Jean-Lynce, Baptiste, Antonella, Thomas (mon stagiaire) et bien d'autres dont j'ai oublié le nom (Enzo par exemple). Je voudrais aussi remercier tous les membres de l'unité NGPA, pour leur bonne humeur, leur sympathie et leurs afterworks mémorables. Je vais m'attarder sur Ioannis Sarras, car il m'a fourni beaucoup de conseils pratiques sur le déroulement d'une thèse et beaucoup d'aide scientifique sur des sujets que je ne maitrisais pas du tout, en arrivant à l'ONERA.

Finalement, je remercie les membres de ma famille ou ceux considérés comme tel, en commençant par Kester James Bradbury dont l'impertinence molle et la folie rêveuse me complètent depuis dix ans. J'ai une pensée particulière pour ma mamie, Monique, et son soutien indéfectible et pour mon oncle, Alain et ma cousine, Emma qui m'enjaillent au détour de soirées parisiennes par leur poésie du quotidien. Que dire de ma soeur, Justine et de mon beau-frère Nicolas à part qu'ils m'ont beaucoup invité chez eux et que je suis peu venu.

Mon dernier mot sera dedié à mes parents Serge et Isabelle à qui je dois d'être qui je suis et ce que je suis dans ce monde à savoir un don ou une malédiction.

Résumé

Dans cette thèse, nous étudions les problèmes de contrôle et d'estimation non linéaire couplés avec application à la navigation par corrélation de terrain (TAN en anglais). L'objectif est de guider et d'estimer la position 3D d'un drone survolant une zone connue. La principale difficulté de cette application est la nature de l'information disponible sur le système. En effet, on suppose que les seules données disponibles sont la vitesse du système, une mesure de hauteur/sol et une carte de la zone survolée. Comme la carte est généralement créée avec des données empiriques, cela pose un problème d'estimation très complexe en soi qui a été étudié de manière approfondie depuis les années 50.

Habituellement, les problèmes d'estimation et de contrôle sont résolus séparément en invoquant le classique principe de séparation qui stipule que de bonnes performances peuvent être obtenues en combinant un contrôleur avec information parfaite et un estimateur conçu indépendamment.

Cependant, le principe de séparation n'est applicable que pour des classes de systèmes très particulières, dont les systèmes linéaires. Ainsi, pour un système non linéaire général, le contrôle et l'estimation doivent être gérés conjointement. La TAN est un bon exemple d'application non linéaire où le principe de séparation ne peut pas être appliqué. En réalité, la qualité des observations dépend du contrôle et plus précisément de la zone survolée par le drone. En conséquence, Les sujets suivant ont été étudiés

- La conception d'observateurs non linéaires et contrôle en retour de sortie pour la TAN avec des cartes analytiques au sol dans un cadre temporel continu.
- Le problème couplé du filtrage optimal non linéaire et du contrôle optimal stochastique en temps discret avec des informations imparfaites.
- La conception de schémas de commande prédictive duale explicite couplés à un filtre particualrie et leur implémentation numérique vers une application non linéaire (y compris la TAN).

Concernant le premier sujet, la technique d'immersion et d'invariance a été utilisé pour concevoir des observateurs non linéaires capables de reconstruire la position 3D du drone dans plusieurs cas de cartes au sol (quadratiques, cubiques, gaussiennes). Il a été démontré que, sous condition de persistance d'excitation sur la vitesse horizontale, l'erreur d'estimation converge vers zéro. Le problème de contôle en retour de sortie a également été étudié en utilisant le concept de δ -persistance. Concernant le deuxième sujet, une reformulation des problèmes de filtrage optimal non linéaire et de contrôle optimal stochastique en temps discret avec information imparfaite en un seul problème d'optimisation a été proposée. Cela permet de justifier l'utilisation de deux étapes dans la résolution du problème sous des hypothèses naturelles sur la fonction de coût. La première étape consiste à résoudre un problème classique d'estimation optimale. La quasi-optimalité de la moyenne empirique d'un filtre particulaire modifié par rapport à l'erreur quadratique moyenne a été montrée dans cette thèse. Cela justifie l'utilisation du filtrage particulaire. La deuxième étape consiste à résoudre un problème de contrôle optimal modifié avec un nouveau terme issu d'une estimation optimale. Ceci établit une connexion avec le contrôle dual explicite dans lequel un nouveau terme représentant une mesure d'information est ajouté empiriquement au coût. En réalité, ce terme

empirique peut être vu comme une approximation du terme issu de l'estimation optimale. En ce qui concerne le troisième sujet, les filtres de Kalman sont répandus dans l'estimation d'états mais connus pour être inefficaces pour la TAN contrairement aux filtres particulaires. En conséquence, au cours de cette thèse, deux contrôleurs prédictifs stochastiques explicites duauxen retour de sortie pour des systèmes non linéaire généraux, couplés à un filtre particulaire, ont été conçus et appliqués à la TAN. Les deux schémas sont basés sur l'*integrated experiment design*, mais l'un contient une pénalisation de l'objectif de guidage dans le coût et l'autre utilise une contrainte de Lyapunov. Les problèmes d'optimisation résultants sont résolus grâce à une méthode d'approximation de Monte Carlo et les deux contrôleurs ont montré de bons résultats en simulation.

Contents

1 Introduction **1**

I Terrain-Aided Navigation with analytical ground maps **7**

2 Terrain models and observability conditions **9**
2.1 General nonlinear observability conditions . . . 9
2.1.1 General nonlinear controlled system . . . 9
2.1.2 General definitions of observability . . . 10
2.2 Observability properties in Terrain-Aided Navigation . . . 11
2.2.1 Dynamical models for Terrain-Aided Navigation (TAN) with closed-form ground maps . . . 12
2.2.2 Local weak observability of Terrain-Aided Navigation models . . . 17

3 Observer and controller design for TAN **29**
3.1 Nonlinear observer design by I&I for TAN . . . 29
3.1.1 Description of I&I for the design of observers of a general nonlinear system . . . 30
3.1.2 Application of I&I design to Terrain-Aided Navigation . . . 32
3.2 Persistent output-feedback control and stochastic excitation . . . 50
3.2.1 Output-feedback control . . . 50
3.2.2 An example of a piecewise constant almost surely persistent processes . . . 59

II Terrain-Aided Navigation with real ground maps **67**

4 Elements of stochastic control and stochastic filtering **69**
4.1 Discrete-time Stochastic Optimal Control . . . 69
4.1.1 Deterministic case . . . 70
4.1.2 Stochastic case with perfect information . . . 72
4.1.3 Stochastic case with imperfect information . . . 75
4.2 State Estimation, Nonlinear Filtering and Particle Filters . . . 79
4.2.1 Optimal estimation . . . 79
4.2.2 Linear Kalman filtering . . . 80
4.2.3 Nonlinear filtering . . . 82
4.3 Design of suboptimal policies . . . 88
4.3.1 Dual effect and classification of suboptimal policies . . . 88

4.3.2 Model Predictive Control 90

5 Modelling of joint problem of control and estimation 99
5.1 Definition of the coupling optimisation problem 99
5.1.1 Setup of the problem of optimal control and estimation 99
5.1.2 Dynamic programming principle 101
5.2 Near-optimal nonlinear filtering 103
5.2.1 Statement of the problem 103
5.2.2 Error bounds between the optimal MSE and the empirical MSE 104
5.3 Stochastic optimal control with an estimation-based cost 116

6 Dual Particle Fisher SMPC and its application to TAN 119
6.1 Design of the Dual Particle Fisher SMPC schemes 119
6.1.1 Implementation of the guiding objective in Explicit Dual SMPC 120
6.1.2 Penalising Fisher Dual Particle SMPC scheme 121
6.1.3 Lyapunov/Fisher Dual Particle SMPC scheme 122
6.2 Modelling of TAN with a real ground map 125
6.3 Numerical simulations 128

7 Conclusion 133

Chapter 1

Introduction

Problems of control and state estimation of dynamical systems with imperfect information are widespread, for instance, in chemical engineering [Kumar and Ahmad, 2012], in electrical and mechanical engineering [Bolton, 2003], in mathematical finance [Mamon and Elliott, 2007] and in aerospace engineering [Eren et al., 2017]. In this work, we study nonlinear coupled control and estimation problems in order to solve a challenging application that is Terrain-Aided Navigation (TAN). The objective is to guide and estimate the 3D position of a drone flying over a known area. The main difficulty of this application is the nature of the available information on the system. Indeed, it is assumed that the only available data are the speed of the system, a measurement of the height from the ground and a map of the area flied over. As the map is generally created with empirical data, this leads to a very complex estimation problem in itself which has been thoroughly studied since the 50's ([Dahia, 2005], [Melo and Matos, 2017], [Sabatino et al., 2003]). Thus, adding guiding to the problem makes it even harder and lacks of treatment in the literature.

Usually, estimation and control problems are solved separately invoking the classical separation principle which states that good performance can be achieved by combining a controller with perfect information on the state and an independently devised estimator. See [Hespanha, 2009] and [Bertsekas, 1995] for examples in different frameworks.

However, the separation principle is known to hold true only for very particular classes of systems including linear systems. Thus, for a general nonlinear system, control and estimation must be handled jointly. TAN is a good example of nonlinear application where the separation principle cannot be applied. Actually, the quality of the observations depends on the control and more precisely on the area that is flied over by the drone. Figure 1.1 depicts two typical trajectories of a drone in TAN. The ellipses represent the uncertainty on the final state. The red trajectory corresponds to flight over a flat area with constant altitude. In this case, one measurement of height matches a whole horizontal area and the estimation error on the horizontal position is of the same order of magnitude as the size of the area which can be very large. On the contrary, if the drone flies over a rough terrain, which corresponds to the green trajectory, then one measurement of height corresponds to a smaller area on the ground. The estimation error is then reduced, as shown in Figure 1.1 where the green ellipse is smaller than the red one.

The main objective of the thesis is to tackle this problem and to design coupled control and estimation methods for nonlinear dynamical systems applicable to terrain-aided navigation.

In Part I, we put ourselves in a simplified framework of TAN to understand how the control can influence observability and how to build observers and control laws accordingly. Firstly, in Chapter 2,

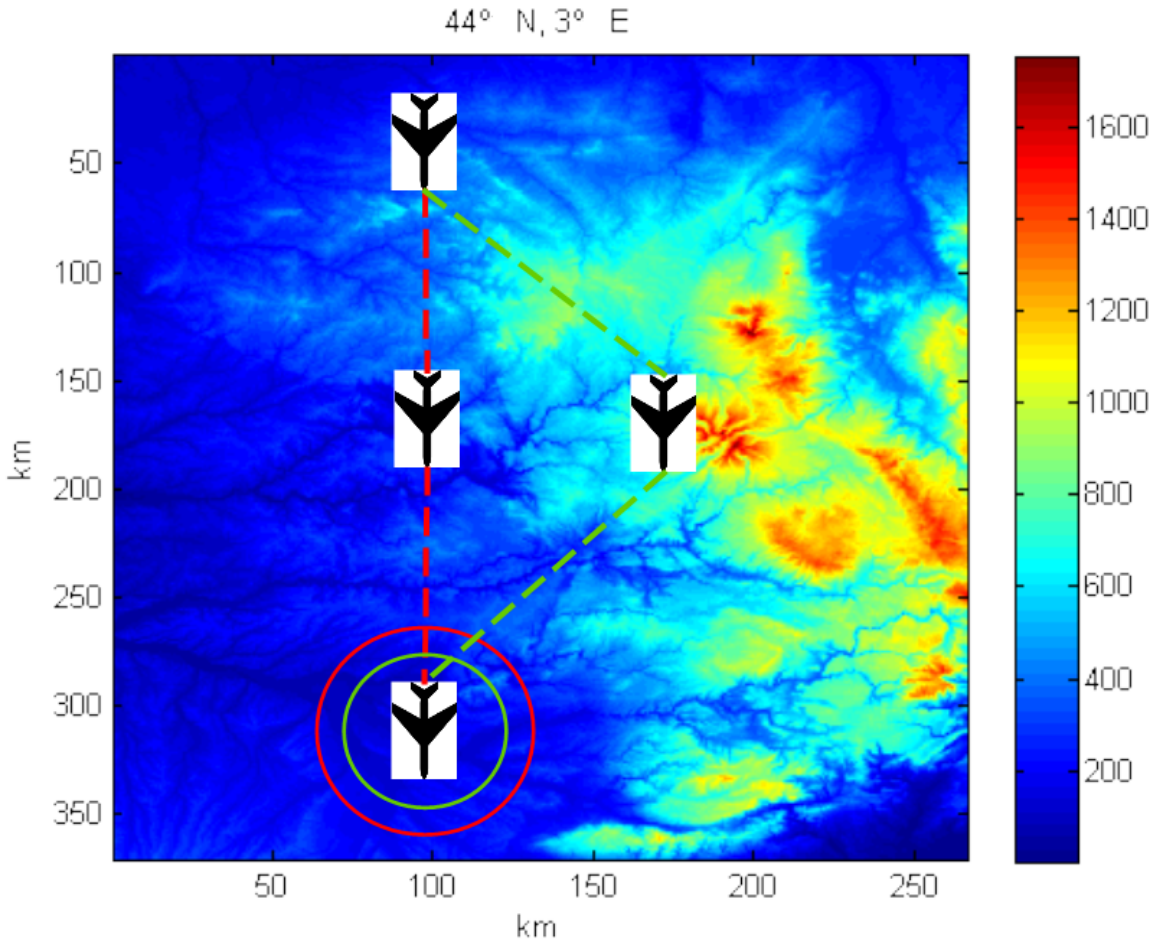


Figure 1.1: Example of trajectories of a UAV flying over a real terrain

we assume that the dynamics of the drone is a deterministic double integrator in continuous time. We represent the state of the system with the 3D position, the 3D speed and with the control being the 3D acceleration. We also assume that the altitude of the ground is represented by a map h_M from $\mathbb{R}^2$ to $\mathbb{R}$ which is a polynomial, a Gaussian, a sum of Gaussians or a Fourier serie. This modelling simplifications allow us to get insight on which trajectory the drone must follow to be able to reconstruct its position. It is done by computing the classical conditions for local weak observability [Besançon, 2007]. Actually, it emerges that observability is guaranteed if the horizontal speed and acceleration are not colinear.

This condition is related to persistence of excitation (PE) of the horizontal speed. In adaptive control, PE describes the fact that a vector signal must explore sufficiently many linearly independent directions in the state space. This allows one to excite the system and make the estimation of some unknown parameters possible (see [Narendra and Annaswamy, 2012]).

Following this concept in Chapter 3, we propose new observers for TAN based on two variants of the Immersion and Invariance (I&I) technique [Astolfi et al., 2008] and a controller designed accordingly. Observer design by I&I consists in defining a nonlinear estimation error in the augmented space (state/observer) so that the manifold where the error vanishes is invariant and attractive. Our first ap-

proach is to directly handle the nonlinearity of the map in the design of the nonlinear estimation error. This technique has been successfully applied to the case of a polynomial map of degree 2 and a Gaussian map with an additional measurement of the absolute altitude. In [Martin and Sarras, 2018], a similar technique is used to show global convergence of a state observer for a particular nonlinear system. However, some maps are too complex to be tackled directly. Our second approach is then to immerse the original system into a higher dimensional one for which observer design is easy. Eventually, one has to make sure that the higher dimensional observer can be used to estimate the original system. This method has been applied to the case of a cubic map and partially to the case of a trigonometric map.

In the direct approach, the resulting estimation error equation is very well-known in adaptive control. Moreover, global convergence of the estimation error to zero is proven under a condition of persistence of excitation on the horizontal speed. Intuitively, it states that the system must avoid to go in straight line. In the indirect approach, a PE condition is still required but it is harder to interpret. PE has been widely studied but it is still not obvious how to incorporate it in an output-feedback loop. Besides, continuous time almost sure stochastic PE conditions have not been extensively studied.

That is why, afterwards, we focus on finding interesting sufficient conditions for the convergence of the previous observers, in terms of output-feedback control and stochastic excitation.

It is rather simple to design a control law that ensures a PE condition on the speed. However, persistent signals generally do not converge, think of sinusoidal signals for example. It means that, if one tries to achieve output-feedback asymptotic stability, the speed cannot be persistent as classically defined. Moreover, a stabilising controller is generally not persistent enough to ensure the convergence of the estimation error to zero. Accordingly, an idea is to enforce the speed to verify a condition of δ-persistence inspired from [Loría et al., 1999] and [Loría et al., 2002] which allows the whole system to converge. More precisely, the speed is persistent with a time varying level of excitation. The latter is usually made to decrease as some norm of the state of the system. This allow the true state of the system to converge to an equilibrium point while ensuring that the estimation error goes to zero.

However, to apply existing results in [Loría et al., 2002], the level of excitation should be an increasing and zero at zero function of the norm the estimation error. Although, the latter is unknown at all time so it cannot be used in the design of the controller. Consequently, we choose to make the level of excitation of the speed diminish as the estimated trajectory converges to the equilibrium. The main drawback of the method is that, if the estimated trajectory converges too quickly, the level of excitation reaches zero before the estimation error become small. However, we show that if the estimated trajectory converges sufficiently slowly to the equilibrium then the whole system verifies the property of semi-global practical stability.

The observer-controller performances are validated by numerical simulations in above cited cases. In practice, the controller is composed of a stabilising term plus a term which is δ-persistent with respect to the speed and the state of the observer. The excitation can be introduced by any classical PE signal like sinusoidal ones. Although, in simulations, stochastic signals may also be sufficient. In [Loría et al., 2002], a Gaussian white noise is used. We give a new example of 2D rectangular processes that are almost surely PE with a random level of excitation.

To sum up Part I, the observability properties of TAN with several types of analytical ground maps are first studied in Chapter 2. Then, Chapter 3 is dedicated to the design of nonlinear observers and control laws with a focus on the condition of persistence of excitation of the horizontal speed. Restricting ourselves to analytical ground maps gives us an intuition and some avenues on how to deal with TAN for real maps. Still, the previous framework does not provide good results for more complicated maps like a sum of Gaussians even if the observability conditions are similar. Actually, the additional nonlinearity in these cases makes the problem of finding a globally convergent observer too difficult.

Therefore, in Part II, we decide to tackle the problem of estimation and control with terrain-aided measurement for a real ground map in a discrete time stochastic framework. Concretely, our first motivation is to be able to use stochastic filters like particle filters to estimate the state of the drone. Our second motivation is to use techniques from stochastic control with imperfect information, such as dual stochastic Model Predictive Control, to couple the design of the estimator and of the controller.

To do so, we consider a general non-linear Markov chain to represent the state of the system and a nonlinear observation equation to represent the noisy partial measurements. In this framework, as one considers general non vanishing stochastic perturbation in a nonlinear system, it is impossible to get exact output-feedback convergence of the state of the system or exact convergence of an estimator to the true state. That is why, we decide to choose an optimisation-based solution to deal with estimation and control.

In Chapter 4, we recall the fundamentals of stochastic control and filtering. On one hand, there is the state estimation problem. In this framework, it corresponds to finding an estimator of the state as a function of the available information that minimised some estimation error. Customarily, one tries to minimise the mean square error between the state and the estimator. In this case, the best estimator is the conditional expectation of the state with respect to all the available information. Additionally, if the system is linear Gaussian then the Kalman filter gives the optimal solution in closed form [Anderson and Moore, 1979]. Although, for a general nonlinear non-Gaussian system, there is a closed form neither for the conditional expectation nor for the conditional distribution (also called the filtering distribution). However, the latter verifies a recursive equation, the so-called filtering equation, which is itself intractable. Consequently, in practice, only approximations of the filtering equation are computed. Among the popular approximations are Extended Kalman filters, Unscented Kalman filters, Ensemble Kalman filters and particle filters. Recently, particle filters have demonstrated their performance in terrain-aided navigation [Dahia, 2005], [Murangira, 2014]. The principal reason is that particle filters (PF) can deal with nonlinear dynamics or observation equation and with multimodal uncertainties which both appears in TAN. That is why, we focus on PF in the following.

On the other hand, in the stochastic control framework, one does not look for control values but rather for control policies that are measurable functions that map the available information to a control value. An important property of the control when only partial information is available is that the controller must look for more information to ensure a good estimation and a better control in the end. In other words, for a general system, the controller may degrade the available information and prevent one from building good estimators. This property, referred as the dual effect property, was first brought to light by Feldbaum in his seminal work [Feldbaum, 1960]. It also means that the separation principle does not apply in general. In [Bar-Shalom and Tse, 1974], a classification of control policies according to their level of information use and probing is presented. It goes from open loop controls, where one looks for controls values depending only on the initial information, to closed loop controls, where the current information is used and the future one is predicted in some way. In the case of stochastic optimal control with imperfect information as described in the survey paper [Mesbah, 2017], one can see from the Dynamic Programming (DP) principle [Bertsekas, 1995] that optimal controls have the dual effect properties. In this case, it is called implicit dual effect because it is only due to the optimality of the controller without external excitation. However, as these problems are intractable in practice, suboptimal output-feedback control laws are computed instead, with the idea to keep the dual effect property. They are called dual controllers. There exist two kinds of dual controllers: implicit and explicit ones. The idea behind implicit dual controllers is to keep the natural implicit dual effect coming from optimal controls by approximating the DP equation. We do not consider implicit dual controllers in this work as they are very difficult to combine with PF and computationally costly. Instead, we consider explicit

dual controllers which are derived by solving a related open loop control problem where a measure of information is added to the cost. This new cost allows the controller to recover the dual effect lost by doing an open loop approximation. The main drawback of these controllers is that the measure of information is generally empirical and hard to make theoretically grounded.

Therefore, our first contribution, gathered in Chapter 5, consists in the definition a new infinite horizon multistage stochastic optimisation problem that gathers a stochastic optimal control problem with imperfect information and an optimal estimation problem. This problem can be recast as a classical infinite-time stochastic optimal control problem with imperfect information by considering an augmented control. The augmented controls are composed of an estimation policy and of the original control policy. By applying the DP principle to this problem and, under natural assumptions on the cost function, one can justify the use of two steps in the resolution of the problem:

- The first step is to solve a classical optimal estimation problem. From a more theoretical perspective, one would like to know if the Mean Square Error (MSE) associated with the empirical mean of a PF converges to the optimal MSE, the one associated with the true conditional expectation. There exist many results of convergence of PF to the true filtering distribution as the number of particles goes to infinity. In [Crisan and Doucet, 2002], and [Arnaud Doucet et al., 2001] several convergence results and error bounds are surveyed. The most classical error bounds involve integrals of bounded continuous functions of the PF and the filtering distribution. These results do not allow us to conclude as the mean is the integral of the identity which is unbounded. Surprisingly, very few results concerning integrals of unbounded functions exist. Nevertheless, using the results of [Hu et al., 2011], we provide an error bound between the MSE of the PF and the optimal one, for a specific PF algorithm and a rather large class of models that contains TAN. Thus, the modified PF solves approximately the optimal estimation problem for a square estimation error.

- Once optimal or near optimal estimation is achieved, we show that, to solve our new optimisation problem, the last step is to solve a stochastic optimal control problem with imperfect information where the cost function is composed of two terms that are contradictory in general. The first one is the estimation error associated to the optimal estimator seen as a function of the state and the available information. The second one is a classical cost function depending on the control and on the state which can be, for instance, an economic cost, a penalisation of some guiding goal or a combination of the two. Consequently, our claim is that classical explicit dual control problems are approximations of the former optimal control problem. The estimation-based term which is generally unknown is replaced by a related one usually independent of the observations such as the Fisher Information Matrix.

One of the most common method to solve approximately infinite time stochastic optimal control problems is Stochastic Model Predictive Control (SMPC). It consists first in solving a finite horizon problem at the current. Then, one keeps and applies only the first optimal control to the true system. Finally, these computations are repeated starting from the new current state in a receding horizon way. In the case of partial information, the current state is usually replaced by the current estimation. A lot work has been done using Kalman-like filters [Hovd and Bitmead, 2004], [Subramanian et al., 2015], [Heitsch and Römisch, 2009], [Heirung et al., 2015]. Although few MPC methods uses PF in the literature (see [Sehr and Bitmead, 2016]).

Following the previous algorithm scheme, our second main contribution, gathered in Chapter 6, consists in the design of two dual output-feedback SMPC methods which includes a particle filter for state estimation. The first one consists in coupling a particle filter with the resolution by a Monte Carlo method

of an explicit dual stochastic optimal control problem. The initial condition of each Monte Carlo trajectory is chosen to be a particle from the particle filter. This allows the control to be aware of the multimodal uncertainty on the state which is more difficult or even impossible with a Kalman-like filter. This has been used in [Sehr and Bitmead, 2016] but not in the framework of dual control. The main drawback of this method is that the cost function of the optimisation problem combines a stabilising term and an information probing term. This creates a trade-off between going toward the guiding goal and looking for more information. Indeed, if too much importance is given to the information term then probing will be efficient but the system may get stuck far from the target. Conversely, if too much importance is given to the stabilising term then the probing effect will not be sufficient and output-feedback performance may be poor. It is difficult to know which case will occur beforehand. As a consequence, we propose another dual MPC where the guiding objective is dealt with by a Lyapunov constraint coming from the theory of Markov chain stability. The main consequence is that there is no more trade-off between guiding and information probing inside the cost. We prioritise the guiding goal and we only look for stabilising controls that maximise the expected information.

Finally, these two algorithms are applied to TAN with a real ground map. The main challenge is to solve the stochastic optimal control problem numerically. To do so, we consider a Monte Carlo approximation and solve it with a nonlinear programming solver. Besides, the real ground map is usually obtained from discrete data. Thus, one has to interpolate the map during the resolution.

To sum up Part II, the basics of stochastic control and filtering are first described in Chapter 4. Secondly, the modelling of state estimation and stochastic optimal control for a general nonlinear system gathered in one global stochastic optimisation problem is presented in Chapter 5. Finally, in Chapter 6, two explicit dual output-feedback stochastic MPC schemes based on particle filtering are proposed and tested on TAN with a real ground map.

Part I

Terrain-Aided Navigation with analytical ground maps

Chapter 2

Terrain models and observability conditions

In whole Part I, we consider the problem of Terrain-Aided Navigation in a simplified deterministic continuous time framework. The ground profile is assumed to be a simple known functions or combinations of simple ones. The goal of this part is twofold. First, one would like to design observers and controllers for as many different models of TAN as possible. Secondly, one is looking for some insight on how to solve the problem of TAN with a complicated real maps. In Chapter 2, we define the several models of maps and study the intrinsic observability properties of the resulting controlled system. One is willing to identify the influence of the control on the observability of these systems.

2.1 General nonlinear observability conditions

2.1.1 General nonlinear controlled system

First, consider the following general nonlinear controlled dynamical system:

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), \\ y(t) &= h(x(t), u(t)), \end{aligned} \tag{2.1}$$

where $\forall t \geq 0$:

- $x(t) \in \mathbb{R}^n$ is the state of the system.
- $x(0) = x_0$ where $x_0 \in \mathbb{R}^n$ is an initial condition.
- $y(t) \in \mathbb{R}^p$ is the output of the system.
- $u(t) \in \mathbb{R}^m$ is the input of the system. We denote by $\mathcal{U}$, the set of possible input $u(\cdot)$.
- $f : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ represents the dynamics of the system and is assumed to be$\mathcal{C}^\infty$ in x and u.
- $h : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^p$ is the observation function that links the output $y(t)$ to the state $x(t)$. In the following, we also assume that h is $\mathcal{C}^\infty$.

This model will be used in the following to present some general concepts in nonlinear observability theory and nonlinear observer design.

2.1.2 General definitions of observability

We consider a general nonlinear model of the form (2.1). We fix an input $u(\cdot)$ and an initial condition x_0. we define $x(t; x_0, u(\cdot))$ as the solution of equation (2.1) starting from x_0 with input $u(\cdot)$. As the model is nonlinear, classical Kalman rank condition [Hespanha, 2009] does not apply. Moreover, the observability properties depend on the input.

2.1.2.1 Local weak observability

We would like to check if one is able to discriminate x_0 from all other initial conditions using the output $y(\cdot)$ and the control $u(\cdot)$. This is a problem of observability.

In the following, we recall several classical notions of nonlinear observability borrowed and adapted from those in [Besançon, 2007]. It is important to notice that, as the input is fixed, we study the observability of an uncontrolled system parametrised by $u(\cdot)$. First we consider the definition of distinguishable states.

Definition 2.1. *A pair* $(x_0, x_0') \in (\mathbb{R}^n)^2$ *is* distinguishable with input $u(\cdot)$ *if,* $\exists t \geq 0$ *such that:*

$$h(x(t; x_0, u(\cdot)), u(\cdot)) \neq h(x(t; x_0', u(\cdot), u(\cdot)).$$

A state x is said to be distinguishable from x_0 with input $u(\cdot)$ *if the pair (x, x_0) is distinguishable with input $u(\cdot)$.*

It is different from classical distinguishability defined in [Besançon, 2007]. Usually, one looks at every possible input and shows that there exists one that is able to separate two different initial conditions. Distinguishability from an particular initial condition with respect to some particular input is more concerned with checking that a proposal input allows to discriminate this initial condition from other ones. From this version of distinguishability, we define local weak observability at x_0 with input $u(\cdot)$.

Definition 2.2. *The system* (2.1) *is said to be* locally weakly observable at x_0 with input $u(\cdot)$ *if there exists a neighbourhood of x_0, $\mathbb{U}$, such that $\forall x_0' \in \mathbb{U}$, $\exists t \geq 0$ such that:*

$$\begin{aligned} h(x(t; x_0, u(\cdot)), u(\cdot)) &\neq h(x(t; x_0', u(\cdot), u(\cdot)), \\ x(t; x_0', u(\cdot)) &\in \mathbb{U}. \end{aligned}$$

This notion means that one can distinguish all states near x_0 while staying near x_0 using the input $u(\cdot)$. Therefore, this kind of property has a practical interest but is hard to check this form. That is why we define the notion of observability space. As we are concerned with local properties for the moment, we look for control values that locally allows observability. Thus we consider in the following definition that $u(\cdot) \equiv u_0$ with $u_0 \in \mathbb{R}^m$.

Definition 2.3. The observability space with constant input u_0, *denoted by $\mathcal{O}_{u_0}(h)$, is the smallest real vector space of $\mathcal{C}^\infty$ functions from $\mathbb{R}^n$ to $\mathbb{R}$ that contains the components of $h(\cdot, u_0)$ and is stable under the Lie derivative $\mathcal{L}_{u_0}$ which is defined by:*

$$\mathcal{L}_{u_0}\phi = d\phi f(x, u_0)$$

The notion of observation space allows to write a sufficient condition of local weak observability based on a rank condition.

Definition 2.4. *The system (2.1) is said to verify* the observability rank condition at x_0 with input u_0 *if:*

$$\dim(d\mathcal{O}_{u_0}(h)|_{x_0}) = n$$

where $d\mathcal{O}_{u_0}(h)|_{x_0}$ *is the set of* $d\phi(x_0)$ *with* $\phi \in \mathcal{O}_{u_0}(h)$

Proposition 2.1. *If the system (2.1) satisfy the observability rank condition at* x_0 *with input* u_0*, then it is locally weakly observable at* x_0 *with input* u_0*.*

Proof. The proof can be found for a more general case in [Hermann and Krener, 1977]. The idea is that the rank condition allows one to construct a local diffeomorphism around x_0 with functions from $\mathcal{O}_{u_0}(h)$. As these functions can be "observed", the result follows. □

2.1.2.2 Universal and persistent inputs

Conditions of local weak observability from the last section give insight on how a constant control must be chosen depending on the initial state x_0. We are now interested in finding conditions on the input $u(\cdot)$ to ensure that observability is preserved along the trajectory. This leads to the definition of universal inputs.

Definition 2.5. *An input* $u(\cdot)$ *is an* universal input on $[0, t]$ *if for any* $x_0 \neq x_0'$*, there exists* $\tau \in [0, t]$ *such that* $h(x(\tau; x_0, u(\cdot)), u(\cdot)) \neq h(x(\tau; x_0', u(\cdot), u(\cdot))$.

From Definition 2.5, one can derive an equivalent integral property:

Proposition 2.2. *An input* $u(\cdot)$ *is universal on on* $[0, t]$ *if and only if* $\forall x_0 \neq x_0'$*:*

$$\int_0^t \|h(x(\tau; x_0, u(\cdot)), u(\cdot)) - h(x(\tau; x_0', u(\cdot)), u(\cdot))\|^2 d\tau > 0.$$

Universal inputs are inputs that never destroy the observability of the system. A related property is the property of regular persistence of the input.

Definition 2.6. *A input* $u(\cdot)$ *is said to be* regularly persistent *if* $\exists t_0 \geq 0, T > 0$ *and there exists* $\kappa : \mathbb{R} \longrightarrow \mathbb{R}$ *increasing with* $\kappa(0) = 0$ *such that,* $\forall t \geq t_0$*,* $\forall x \neq x'$*:*

$$\int_t^{t+T} \|h(x(\tau; x, u(\cdot)), u(\cdot)) - h(x(\tau; x', u(\cdot)), u(\cdot))\|^2 d\tau \geq \kappa(\|x - x'\|).$$

Obviously, from Proposition 2.2, a regularly persistent input is an universal input. Informally, a regularly persistent input excites the system consistently to ensure observability on a receding time horizon. It happens to be a useful property to force the convergence of some observers.

2.2 Observability properties in Terrain-Aided Navigation

The objective of this section is to check the observability rank condition in Terrain-Aided Navigation. To do so, we first define several models of ground maps in closed form and then apply the theory from Section 2.1.1.

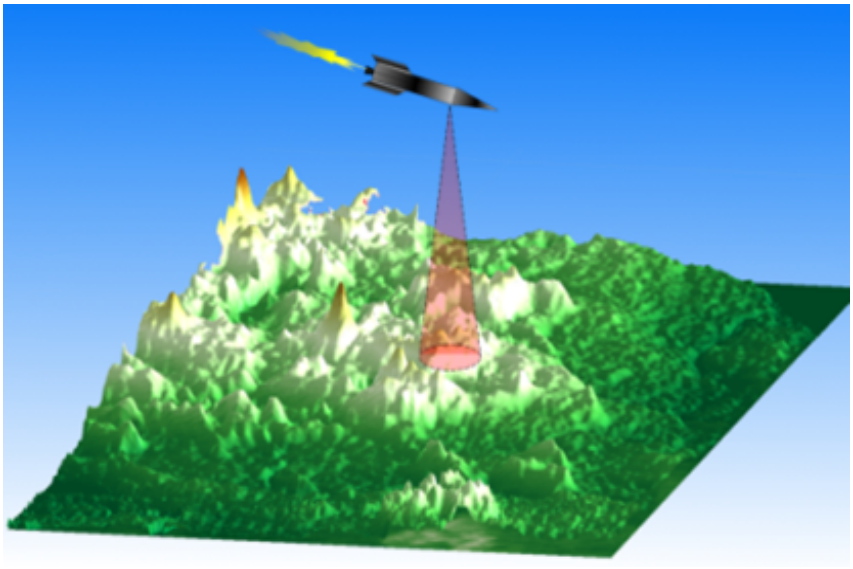


Figure 2.1: Example of a UAV localized by Terrain-Aided Navigation

2.2.1 Dynamical models for Terrain-Aided Navigation (TAN) with closed-form ground maps

We consider the scenario of a drone evolving in a 3D space without direct measurements of its 3D position. This scenario can occur after a GPS failure or simply because GPS devices are not used as they are easy to jam. It is assumed instead that the drone is equipped with a radar-altimeter. A radar-altimeter gives a measurement of the vertical distance between the drone and the ground. It is also assumed that the terrain flied over by the drone is known. The issue of the following is to determine if these two pieces of information are sufficient to reconstruct the horizontal and the vertical position of the drone. This problem is called Terrain-Aided Navigation and is represented in Figure 2.1. Figure 2.2 represents the test map that has been used throughout the thesis.

For the sake of simplicity, the drone is represented by a point in 3 dimensions and we suppose that either its 3D speed or its 3D acceleration is controlled. In any case, the inertial speed i.e. the speed in the inertial frame is supposed to be known. It is a substantial simplification as in many applications the inertial speed is not measured but the speed in the body frame is. Pose estimation is then generally needed to reconstruct the inertial speed. However, one can imagine that an external estimation loop is run such that the inertial speed is reconstructed precisely. Occasionally, we will assume that an additional measurement of altitude is available mainly to simplify the design of some observers. It is not a restrictive assumption as baro-altimeters are very common in aircrafts. We make the previous simplifying assumptions because we want to focus on the main difficulty of the problem of TAN which is the nature of the available information on the *position*. Actually, the available information depends on the nature of the area that is flied over by the drone and consequently on the input. For example, let us assume that the drone flies over an ambiguous area, like a flat or a periodic one, with constant altitude then one measurement of height matches a whole horizontal area. The resulting estimation error on the horizontal position is of the order of magnitude of the size of the area which can be very large. On the contrary, if the drone flies over a rough terrain, then one measurement of height corresponds to a smaller

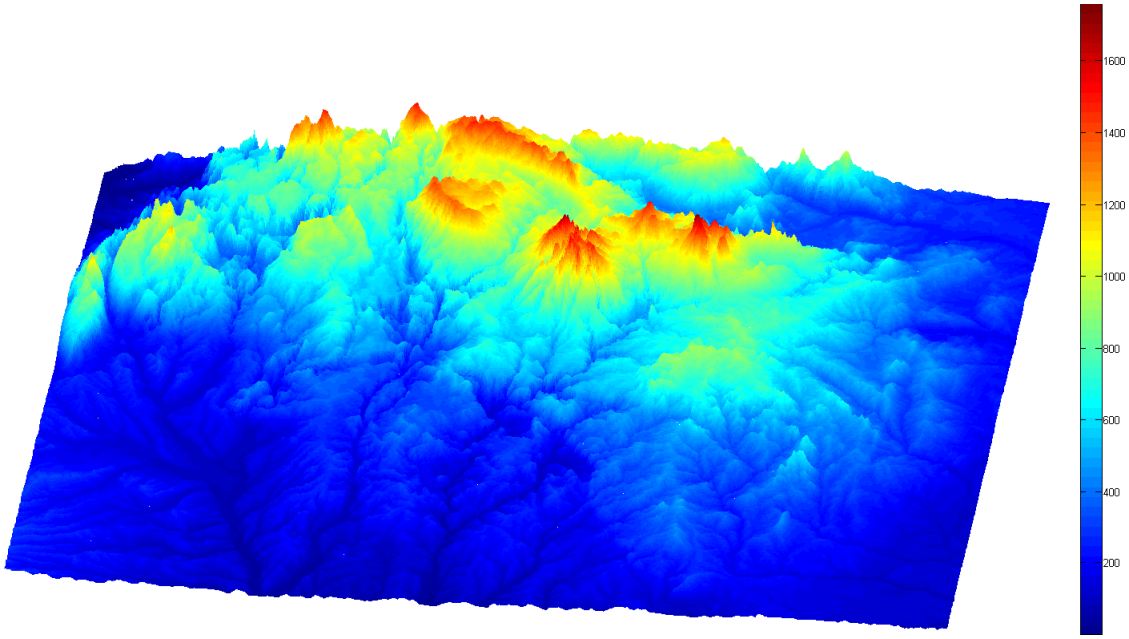


Figure 2.2: Example of a real terrain

area on the ground and the estimation error is reduced. Intuitively, the more perturbed the ground flied over is, the more information it contains. The idea of the following is to verify this intuition by studying the dynamical model of the drone with different types of maps.

Consequently, in the sequel, the dynamics of the drone is either a simple or a double integrator in 3D with control on the speed or on the acceleration.:

$$\dot{X} = V, \tag{2.2}$$

or

$$\begin{aligned} \dot{X} &= V, \\ \dot{V} &= U, \end{aligned} \tag{2.3}$$

where

- $X = (x_1, x_2, x_3)$ is the 3D position with (x_1, x_2) being the horizontal 2D position and x_3 the altitude of the drone,
- $V = (v_1, v_2, v_3)$ is the 3D speed with (v_1, v_2) being the 2D horizontal speed and v_3 the vertical speed. V is the input in the system (2.2),
- $U = (u_1, u_2, u_3)$ is the 3D acceleration with (u_1, u_2) being the 2D horizontal acceleration and u_3 the vertical acceleration. U is the input in the system (2.3).

In the case where the inertial speed of the drone and its height with respect to the ground are measured, the observation equation reads:

$$y = h(X, V), \tag{2.4}$$

with

$$h(X,V) = \begin{bmatrix} x_3 - h_M(x_1,x_2) \\ V \end{bmatrix}.$$

If the altitude is also measured, the observation equation is written as follow:

$$y = h_{alt}(X,V), \tag{2.5}$$

with

$$h_{alt}(X,V) = \begin{bmatrix} h_M(x_1,x_2) \\ x_3 \\ V \end{bmatrix},$$

where $h_M : \mathbb{R}^2 \longrightarrow \mathbb{R}$ represents the profile of the ground. In equation (2.5), it is supposed that h_M is measured because measuring x_3 and $x_3 - h_M$ is equivalent to measuring x_3 and h_M.

In part I, we consider maps h_M that can be written under closed form. In the sequel, we study in detail maps that have the following forms.

- A polynomial function of degree 2:

$$h_M(x_1,x_2) = a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2 + a_{10}x_1 + a_{01}x_2 + a_{00}, \tag{2.6}$$

  with $(a_{20}, a_{11}, a_{02}, a_{10}, a_{01}, a_{00}) \in \mathbb{R}^6$. Figure 2.3 represents a portion of the map presented in Figure 2.2 and its least-square polynomial approximation of order 2. It is a typical example of the possible uses of quadratic maps in TAN.

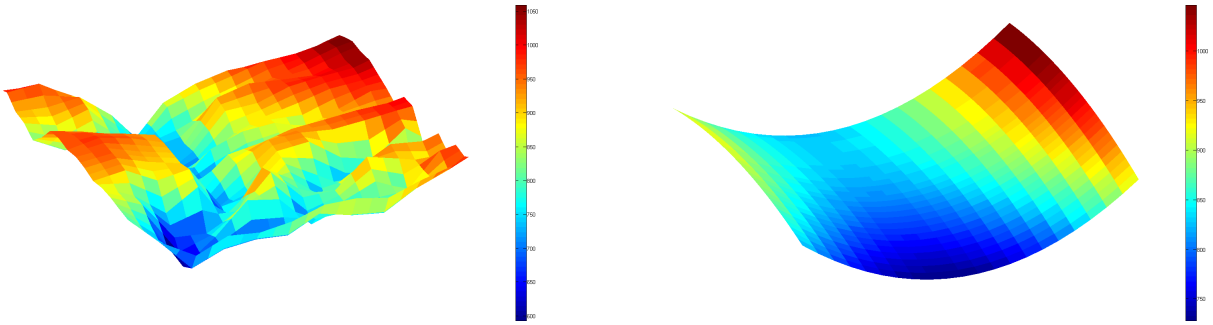


Figure 2.3: Representation of a part of the real map and its quadratic approximation

- A polynomial function of degree 3:

$$\begin{aligned} h_M(x_1,x_2) = {} & a_{30}x_1^3 + a_{21}x_1^2x_2 + a_{12}x_1x_2^2 + a_{03}x_2^3 \\ & + a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2 + a_{10}x_1 + a_{01}x_2 + a_{00}, \end{aligned} \tag{2.7}$$

  with $(a_{30}, a_{21}, a_{12}, a_{03}, a_{20}, a_{11}, a_{02}, a_{10}, a_{01}, a_{00}) \in \mathbb{R}^{10}$. As in the quadratic case, Figure 2.4 represent a cubic approximation of the real map. For the sake of clarity, the approximations in Figure 2.3 and 2.4 are done on a large area which make them rough. Indeed, in practice, one would need much more precise approximations.

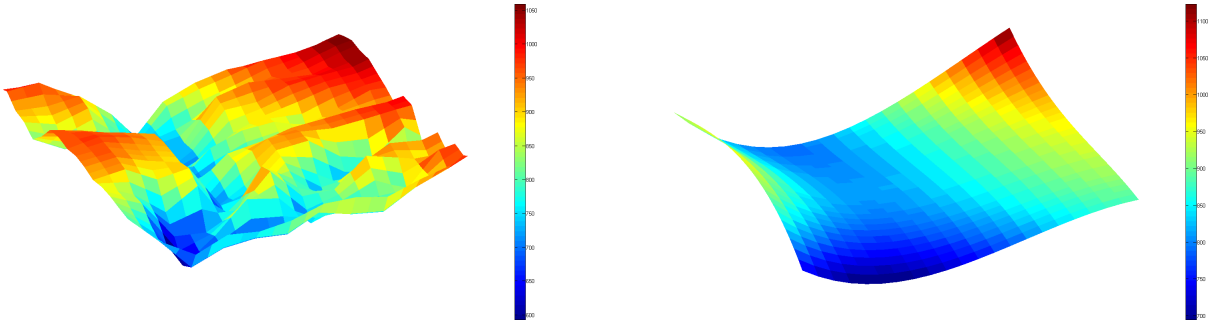


Figure 2.4: Representation of a part of the real map and its cubic approximation

- A Gaussian function:

$$h_M(x_1, x_2) = H^0 \exp\left(-\frac{1}{2}\begin{bmatrix} x_1 - x_1^0 & x_2 - x_2^0 \end{bmatrix} R \begin{bmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \end{bmatrix}\right), \tag{2.8}$$

  where $H^0 > 0$ is the height of the Gaussian, $X^0 = (x_1^0, x_2^0)^T \in \mathbb{R}^2$ is the center of the Gaussian and R is a symmetric positive definite 2×2 matrix that represents the width and the orientation of the Gaussian.

- A finite sum of Gaussian functions:

$$h_M^i(x_1, x_2) = H^i \exp\left(-\frac{1}{2}\begin{bmatrix} x_1 - x_1^i & x_2 - x_2^i \end{bmatrix} R^i \begin{bmatrix} x_1 - x_2^i \\ x_2 - x_2^i \end{bmatrix}\right),$$

$$h_M = \sum_{i=1}^{n_g} h_M^i, \tag{2.9}$$

  where $n_g \geq 2$ and for $i = 1..n_g$, H^i, $X^i = (x_1^i, x_2^i)^T$ and R^i are the parameters of the Gaussian h_M^i as described in equation (2.8). Figure 2.5 represents an approximation of another part of the real map with a sum of 3 Gaussians.

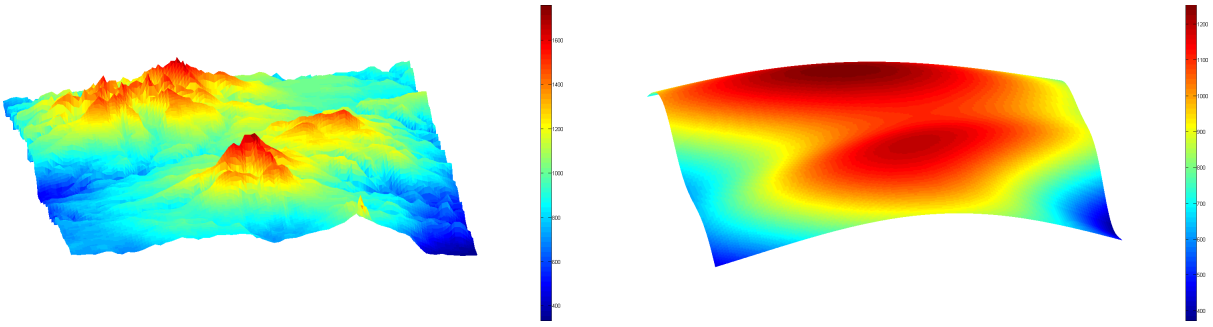


Figure 2.5: Representation of a part of the real map and one possible approximation by a sum of Gaussians

The main advantage of polynomial maps is their potential to represent complicated real maps through local or global interpolation and/or approximation while adding relatively simple nonlinearities in the

system as shown in Figure 2.3 and 2.4. The main advantage of using Gaussian maps is to represent multimodal maps, with several hills or peaks for instance as in Figure 2.5 with potentially one global approximation. However, Gaussian maps introduces non polynomial nonlinearity in the observation equation which makes observer design in this case very complex.

We also study with less detail the case of a map decomposed in spatial trigonometric functions.

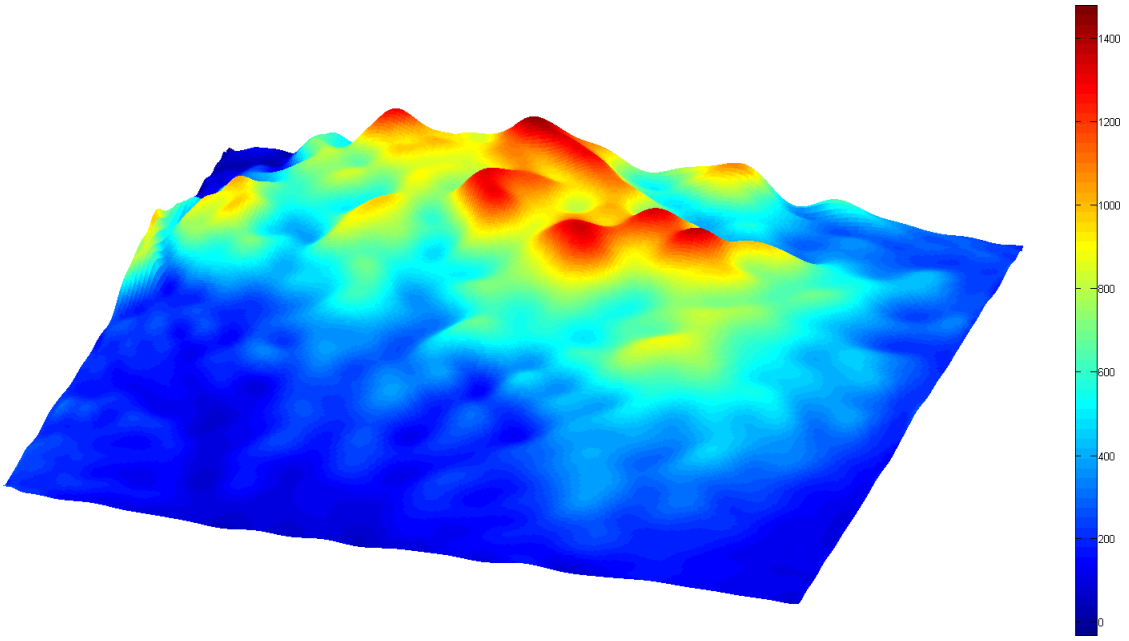


Figure 2.6: Fourier approximation of the real terrain with $n_{tri} = 1600$

- A spatial trigonometric function:

$$h_M(x_1, x_2) = \alpha \cos(\omega_1 x_1 + \phi_1)\cos(\omega_2 x_2 + \phi_2) \tag{2.10}$$

where $\alpha \in \mathbb{R}$ is a Fourier coefficient, $(\omega_1, \omega_2) \in (\mathbb{R}^+)^2$ are pulsations and $(\phi_1, \phi_2) \in \mathbb{R}^2$ are the phases at the origin.

- A sum of spatial trigonometric functions:

$$\begin{aligned} h^i_M(x_1, x_2) &= \alpha^i \cos(\omega^i_1 x_1 + \phi^i_1)\cos(\omega^i_2 x_2 + \phi^i_2) \\ h_M &= \sum_{i=1}^{n_{tri}} h^i_M, \end{aligned} \tag{2.11}$$

where $n_{tri} \geq 2$, for $i = 1, .., n_{tri}$, α^i, (ω^i_1, ω^i_2) and (ϕ^i_1, ϕ^i_2) are the parameters of trigonometric functions as in equation (2.10). Figure 2.6 represents a Fourier approximation of the map in Figure 2.2 with 1600 pulsations (ω^i_1, ω^i_2).

As for polynomial maps, it is well known that the potential of approximation of such maps is huge but dealing with the cosines in the dynamics is a big issue in observer design. That is why, the full treatment of this case is still open.

Actually, a problem of observability arises with the observation equation (2.4) and (2.5) as one wants to reconstruct a 3D position with a 2D measurement or a 2D horizontal position with a 1D measurement. Therefore, the first step of the study is to look for conditions on the input and on the map of the ground that makes the state observable.

2.2.2 Local weak observability of Terrain-Aided Navigation models

In this section, we fix a position X, a speed V and an acceleration U. Depending on the dynamics considered in the following, X is a point in the statespace and V a constant input or (X, V) is a point in the statespace and U a constant input. To check if a model is locally weakly observable, we evaluate the observability rank condition. To do so, The idea is to compute all the functions from the observation space $\mathcal{O}_U(h)$ (or $\mathcal{O}_V(h)$) by successively applying the Lie derivative on the component of the observation function.

We focus on model (2.4) and we split the observation equation into a linear part h^l and nonlinear part h^{nl} such that:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^l(X,V) \end{bmatrix},$$

with $h^{nl}(X) = x_3 - h_M(x_1, x_2)$ and $h^l(X,V) = V$.

In the case of a double integrator, as the linear part is directly of the speed, its Lie derivatives are constant so the only difficulty comes from the nonlinear part. We denote by h_k^{nl} the k^{th} Lie derivative of h^{nl} in the direction of the dynamics considered. Then, $h_0^{nl} = h^{nl}$. Finally one only needs to check the rank of the family $\left(\nabla h^l, \nabla h_k^{nl}\right)_{k\geq 0}$ where ∇ is the gradient operator.

2.2.2.1 Polynomial map of degree 2

We consider the dynamical model (2.3) of the double integrator with the observation equation 2.4 and the map (2.6) which is polynomial of degree 2. The successive Lie derivatives of h^{nl} up to order 2 read:

$$\begin{aligned}
\nabla h_0^{nl} &= \begin{bmatrix} -(2a_{20}x_1 + a_{11}x_2 + a_{10}) \\ -(a_{11}x_1 + 2a_{02}x_2 + a_{01}) \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \\
h_1^{nl} &= \nabla^T h_0^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \\
&= -v_1(2a_{20}x_1 + a_{11}x_2 + a_{10}) - v_2(a_{11}x_1 + 2a_{02}x_2 + a_{01}) + v_3,
\end{aligned}$$

$$\nabla h_1^{nl} = \begin{bmatrix} -(2a_{20}v_1 + a_{11}v_2) \\ -(a_{11}v_1 + 2a_{02}v_2) \\ 0 \\ -(2a_{20}x_1 + a_{11}x_2 + a_{10}) \\ -(a_{11}x_1 + 2a_{02}x_2 + a_{01}) \\ 1 \end{bmatrix},$$

$$\begin{aligned} h_2^{nl} =& \nabla^T h_1^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \\ =& - v_1(2a_{20}v_1 + a_{11}v_2) - v_2(a_{11}v_1 + 2a_{02}v_2) \\ & - u_1(2a_{20}x_1 + a_{11}x_2 + a_{10}) - u_2(a_{11}x_1 + 2a_{02}x_2 + a_{01}) + u_3, \end{aligned}$$

$$\nabla h_2^{nl} = \begin{bmatrix} -(2a_{20}u_1 + a_{11}u_2) \\ -(a_{11}u_1 + 2a_{02}u_2) \\ 0 \\ -(4a_{20}v_1 + 2a_{11}v_2) \\ -(2a_{11}v_1 + 4a_{02}v_2) \\ 0 \end{bmatrix}.$$

We then define the observability matrix H by:

$$H = \begin{bmatrix} 0 & 0 & 0 & -(2a_{20}x_1 + a_{11}x_2 + a_{10}) & -(2a_{20}v_1 + a_{11}v_2) & -(2a_{20}u_1 + a_{11}u_2) \\ 0 & 0 & 0 & -(a_{11}x_1 + 2a_{02}x_2 + a_{01}) & -(a_{11}v_1 + 2a_{02}v_2) & -(a_{11}u_1 + 2a_{02}u_2) \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -(2a_{20}x_1 + a_{11}x_2 + a_{10}) & -(4a_{20}v_1 + 2a_{11}v_2) \\ 0 & 1 & 0 & 0 & -(a_{11}x_1 + 2a_{02}x_2 + a_{01}) & -(2a_{11}v_1 + 4a_{02}v_2) \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

It is not necessary to compute the other Lie derivatives of h because one can see that they do not depend on X so their gradient do not bring vectors that are linearly independent of those of H as V is fully measured. Therefore, as H is a square matrix, the observability conditions can be summed up in the determinant of H. Finally,

$$\begin{aligned} \det(H) =& \begin{vmatrix} -(2a_{20}v_1 + a_{11}v_2) & -(2a_{20}u_1 + a_{11}u_2) \\ -(a_{11}v_1 + 2a_{02}v_2) & -(a_{11}u_1 + 2a_{02}u_2) \end{vmatrix}, \\ =& (2a_{20}v_1 + a_{11}v_2)(a_{11}u_1 + 2a_{02}u_2) - (2a_{20}u_1 + a_{11}u_2)(a_{11}v_1 + 2a_{02}v_2), \\ =& 2a_{20}a_{11}v_1u_1 + 4a_{20}a_{02}v_1u_2 + a_{11}^2v_2u_1 + 2a_{11}a_{02}v_2u_2, \\ & -(2a_{20}a_{11}u_1v_1 + 4a_{20}a_{02}u_1v_2 + a_{11}^2u_2v_1 + 2a_{11}a_{02}u_2v_2), \\ =& (a_{11}^2v_2 - 4a_{20}a_{02}v_2)u_1 + (4a_{20}a_{02}v_1 - a_{11}^2v_1)u_2, \\ =& (a_{11}^2 - 4a_{20}a_{02})(u_1v_2 - u_2v_1). \end{aligned}$$

Then,

$$\begin{aligned} \dim\left(d\mathcal{O}_U(h)|_{(X,V)}\right) < 6 \Leftrightarrow \det(H) = 0 \Leftrightarrow (a_{11}^2 - 4a_{20}a_{02}) \det\left(\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}\right) = 0, \\ \Leftrightarrow a_{11}^2 - 4a_{20}a_{02} = 0 \text{ or } \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} // \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \end{aligned}$$

Finally, in this case, the rank condition is satisfied if:

- $a_{11}^2 - 4a_{20}a_{02} \neq 0$ which means that the quadrics that corresponds to the map is non degenerate;
- and the horizontal acceleration is chosen to be non colinear to the horizontal speed.

The case of a cubic polynomial leads to similar sufficient conditions involving higher order coefficients so we omit the details of the computations.

2.2.2.2 Gaussian map

We analyse now the case of the Gaussian map (2.8). we consider successively dynamics (2.2) and (2.3) to get a progressive insight because the conditions are less clear than in the polynomial case. Let us set:

$$X_{ce} = \begin{bmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \end{bmatrix}, \qquad V_{12} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

X_{ce} represents the horizontal position centred at the centre of the Gaussian and V_{12} is the horizontal speed. We first notice that: $\nabla_{(x_1,x_2)} h_M = -h_M R X_{ce}$. We start with the simple integrator.

Simple integrator

In this case, as V is an input $h = h_0 = h^{nl} = x_3 - h_M$. We compute several Lie derivatives:

$$\nabla h_0^{nl} = \begin{bmatrix} h_M R X_{ce} \\ 1 \end{bmatrix}, \tag{2.12}$$

$$\begin{aligned} h_1^{nl} &= \nabla h_0^{nl} V, \\ &= v_3 + X_{ce}^T R V_{12} h_M, \\ \nabla h_1^{nl} &= h_M \begin{bmatrix} -(X_{ce}^T R V_{12}) R X_{ce} + R V_{12} \\ 0 \end{bmatrix}, \end{aligned} \tag{2.13}$$

$$\begin{aligned} h_2^{nl} &= \nabla^T h_1^{nl} V, \\ &= h_M(-(X_{ce}^T R V_{12})^2 + V_{12}^T R V_{12}), \\ \nabla h_2^{nl} &= h_M \begin{bmatrix} -(-(X_{ce}^T R V_{12})^2 + V_{12}^T R V_{12}) R X_{ce} - 2(X_{ce}^T R V_{12}) R V_{12} \\ 0 \end{bmatrix}, \\ h_3^{nl} &= \nabla^T h_2^{nl} V, \\ &= h_M((X_{ce}^T R V_{12})^3 - 3(V_{12}^T R V_{12})(X_{ce}^T R V_{12})), \\ \nabla h_3^{nl} &= h_M \begin{bmatrix} -((X_{ce}^T R V_{12})^3 - 3(V_{12}^T R V_{12})(X_{ce}^T R V_{12})) R X_{ce} + (3(X_{ce}^T R V_{12})^2 - V_{12}^T R V_{12}) R V_{12} \\ 0 \end{bmatrix}. \end{aligned} \tag{2.14}$$

One can see from the previous computations that some structure appears in the Lie derivatives which leads to Proposition 2.3.

Proposition 2.3. *$\forall k \geq 1$, there exist polynomial functions $g_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R} \longrightarrow \mathbb{R}$ such that:*

$$h_k^{nl}(X, h_M, V) = f_k(v_3) + h_M(g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12})),$$

and $\forall (r,s) \in \mathbb{R}^+ \times \mathbb{R}$, $g_k(r,s)$ *of degree* k *in* s. *Moreover, the following formula holds for the functions* g_k, $\forall (r,s) \in \mathbb{R}^+ \times \mathbb{R}$ *and for* $k \geq 1$*:*

$$g_1(r,s) = s,$$
$$g_{k+1}(r,s) = -g_k(r,s)s + \frac{\partial g_k}{\partial s}(r,s)r.$$

Proof. The proposition is true for $k = 1$ according to equation (2.13) with $g_1(r,s) = s$ and $f_1(v_3) = v_3$

Let us assume that, for some $n \geq 1$, there exists $g_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R} \longrightarrow \mathbb{R}$ such that:

$$h_k^{nl}(X,V) = f_k(v_3) + h_M(g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12}))$$

Then by definition,

$$h_{k+1}^{nl} = V^T \nabla h_k^{nl}.$$

Moreover,

$$\nabla h_k^{nl} = h_M \begin{bmatrix} -g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) R X_{ce} + \frac{\partial g_k}{\partial s}(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) R V_{12} \\ 0 \end{bmatrix}. \tag{2.15}$$

Finally,

$$h_{k+1}^{nl} = h_M(-(X_{ce}^T R V_{12}) g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) + \frac{\partial g_k}{\partial s}(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) V_{12}^T R V_{12}).$$

By setting, $f_{k+1} = 0$ and $g_{k+1}(r,s) = -g_k(r,s)s + \frac{\partial g_k}{\partial s}(r,s)r$, and by noticing that if g_k is of degree k is s then g_{k+1} is of degree $k+1$ in s, the result is proved $\forall k \geq 1$.

□

We can deduce from Proposition 2.3 and equation (2.12) that:

$$d\mathcal{O}_V(h)|_X = Span\left(\begin{bmatrix} -h_M R X_{ce} \\ 1 \end{bmatrix}, \begin{bmatrix} -h_M((X_{ce}^T R V_{12}) R X_{ce} + R V_{12}) \\ 0 \end{bmatrix}, \right. \tag{2.16}$$
$$\left. \left(\begin{bmatrix} h_M g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) R X_{ce} + h_M \frac{\partial g_k}{\partial s}(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) R V_{12} \\ 0 \end{bmatrix} \right)_{k \geq 2} \right).$$

Equation (2.16) leads to the following sufficient conditions of local weak observability.

Proposition 2.4. *If* X_{ce} *is not colinear to* V_{12} *and* $\exists k \geq 1$, $\exists l \geq 1$ *such that* $k \neq l$ *and:*

$$\begin{vmatrix} g_k(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) & g_l(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) \\ \frac{\partial g_k}{\partial s}(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) & \frac{\partial g_l}{\partial s}(V_{12}^T R V_{12}, X_{ce}^T R V_{12}) \end{vmatrix} \neq 0. \tag{2.17}$$

then the system (2.2), (2.8) is locally weakly observable at X *with input* V.

Proof. The idea is to prove that the rank condition is verified at X with input V using the general form of $d\mathcal{O}_V(h)|_X$ in equation 2.16 . To do so, we pick k and l form the statement of the proposition and we compute the following determinant:

$$D := \begin{vmatrix} -h_M R x_{ce} & h_M g_k R X_{ce} + h_M \frac{\partial g_k}{\partial s} R V_{12} & h_M g_l R X_{ce} + h_M \frac{\partial g_l}{\partial s} R V_{12} \\ 1 & 0 & 0 \end{vmatrix},$$
$$D = \begin{vmatrix} h_M g_k R X_{ce} + h_M \frac{\partial g_k}{\partial s} R V_{12} & h_M g_l R X_{ce} + h_M \frac{\partial g_l}{\partial s} R V_{12} \end{vmatrix},$$
$$D = h_m^2 \det(R) \det\left(\begin{bmatrix} X_{ce} & V_{12} \end{bmatrix}\right) \begin{vmatrix} g_k & g_l \\ \frac{\partial g_k}{\partial s} & \frac{\partial g_l}{\partial s} \end{vmatrix}.$$

By construction, $h_M > 0$ and $\det(R) > 0$. Besides, by assumption, X_{ce} and V_{12} are not colinear so $\det\left(\begin{bmatrix} X_{ce} & V_{12} \end{bmatrix}\right) \neq 0$ and $\begin{vmatrix} g_k & g_l \\ \frac{\partial g_k}{\partial s} & \frac{\partial g_l}{\partial s} \end{vmatrix} \neq 0$.

Finally, $D \neq 0$ and $\dim(d\mathcal{O}_V(h)|_X) = 3$ which ensures that the system (2.2), (2.8) is locally weakly observable at X with input V by Proposition 2.1. □

Equation (2.17) is verified for a generical choice of V_{12} and U_{12} as $(V_{12}^T R V_{12}, X_{ce}^T R V_{12})$ is the root of some polynomial. Intuitively, Proposition 2.4 means that one can distinguish all 3D positions near X locally by going in straight line along a direction that is not parallel to X_{ce}.We now deal with the case of a double integrator.

Double integrator

In this case, (X, V) is the state and U the input. We define the horizontal acceleration as follows:

$$U_{12} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

As in the polynomial case the contribution of h^l is easy to determine. Thus, we focus on h^{nl}. Like in the simple integrator case one can derive the successive Lie derivatives:

$$\begin{aligned}
h_0^{nl} &= x_3 - h_M, \\
\nabla h_0^{nl} &= \begin{bmatrix} -h_M R X_{ce} \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (2.18) \\
h_1^{nl} &= \nabla^T h_0^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \qquad (2.19) \\
&= v_3 + X_{ce}^T R V_{12} h_M, \\
\nabla h_1^{nl} &= \begin{bmatrix} -h_M (X_{ce}^T R V_{12}) R X_{ce} + h_M R V_{12} \\ 0 \\ X_{ce}^T R V_{12} \\ h_M R X_{ce} \\ 1 \end{bmatrix}, \\
h_2^{nl} &= \nabla^T h_1^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \\
&= h_M(-(X_{ce}^T R V_{12})^2 + V_{12}^T R V_{12} + X_{ce}^T R U_{12}), \qquad (2.20)
\end{aligned}$$

$$\nabla h_2^{nl} = \begin{bmatrix} h_M((-(X_{ce}^T RV_{12})^2 + V_{12}^T RV_{12} + X_{ce}^T RU_{12})RX_{ce} - 2(X_{ce}^T RV_{12})RV_{12} + RU_{12}) \\ 0 \\ h_M(2RV_{12} - 2(X_{ce}^T RV_{12})RX_{ce}) \\ 1 \end{bmatrix},$$

$$\begin{aligned} h_3^{nl} &= \nabla^T h_2^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \\ &= h_M((X_{ce}^T RV_{12})^3 - 3(V_{12}^T RV_{12})(X_{ce}^T RV_{12}) \\ &+ 2U_{12}^T RV_{12} + (1 - 2(X_{ce}^T RV_{12}))(X_{ce}^T RU_{12})) + u_3. \end{aligned} \tag{2.21}$$

The expression of ∇h_3^{nl} becomes too long to be written here and as in the simple integrator, the Lie derivatives can be expressed in a general form with a sequence of multivariate polynomials.

Proposition 2.5. *$\forall k \geq 1$, there exist functions $g_k : \mathbb{R}^5 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that:*

$$h_k^{nl}(X, V, U) = f_k(v_3, u_3) + h_M(g_k(V_{12}^T RV_{12}, U_{12}^T RU_{12}, X_{ce}^T RV_{12}, X_{ce}^T RU_{12}, V_{12}^T RU_{12})),$$

and $\forall (r_u, r_v, s_{xv}, s_{xu}, s_{vu}) \in \mathbb{R}^{+2} \times \mathbb{R}^3$, $g_k(r_u, r_v, s_{xv}, s_{xu}, s_{vu})$ is polynomial in $(r_u, r_v, s_{xv}, s_{xu}, s_{vu})$. Moreover, a possible choice of the function sequence $(g_k)_{k\geq 1}$ is the one for which the following recursion holds for , $\forall (r_u, r_v, s_{xv}, s_{xu}, s_{vu}) \in \mathbb{R}^{+2} \times \mathbb{R}^3$ and for $k \geq 1$:

$$g_{k+1} = -g_k s_{xv} + \frac{\partial g_k}{\partial s_{xv}} r_v + \frac{\partial g_k}{\partial s_{vu}} r_u + \left(\frac{\partial g_k}{\partial s_{xu}} + 2\frac{\partial g_k}{\partial r_v} \right) s_{vu} + \frac{\partial g_k}{\partial s_{xv}} s_{xu},$$

Proof. The proposition is true for $k = 1$ according to equation (2.19) with $g_1(r_u, r_v, s_{xv}, s_{xu}, s_{vu}) = s_{xv}$ and $f_1(v_3, u_3) = v_3$

Let us assume that, for some $k \geq 1$, there exists $g_k : \mathbb{R}^5 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that:

$$h_k^{nl}(X, V) = f_k(v_3, u_3) + h_M g_k(V_{12}^T RV_{12}, U_{12}^T RU_{12}, X_{ce}^T RV_{12}, X_{ce}^T RU_{12}, V_{12}^T RU_{12}).$$

Then by definition,

$$h_{k+1}^{nl} = \begin{bmatrix} V^T & U^T \end{bmatrix} \nabla h_k^{nl}.$$

Moreover, for $k \geq 1$,

$$\nabla h_k^{nl} = \begin{bmatrix} h_M \left(-g_k RX_{ce} + \frac{\partial g_k}{\partial s_{xv}} RV_{12} + \frac{\partial g_k}{\partial s_{xu}} RU_{12} \right) \\ 0 \\ h_M \left(2\frac{\partial g_k}{\partial r_v} RV_{12} + \frac{\partial g_k}{\partial s_{xv}} RX_{ce} + \frac{\partial g_k}{\partial s_{vu}} RU_{12} \right) \\ \frac{\partial f_k}{\partial v_3} \end{bmatrix}.$$

Then,

$$\begin{aligned} h_{k+1}^{nl} =& h_M \left(-g_k X_{ce}^T RV_{12} + \frac{\partial g_k}{\partial s_{xv}} V_{12}^T RV_{12} + \frac{\partial g_k}{\partial s_{xu}} U_{12}^T RV_{12} + \right. \\ & \left. 2\frac{\partial g_k}{\partial r_v} V_{12}^T RU_{12} + \frac{\partial g_k}{\partial s_{xv}} X_{ce} RU_{12} + \frac{\partial g_k}{\partial s_{vu}} U_{12}^T RU_{12}. + \frac{\partial f_k}{\partial v_3} u_3 \right). \end{aligned}$$

By setting,

$$\begin{aligned} g_{k+1} =& - g_k X_{ce}^T R V_{12} + \frac{\partial g_k}{\partial s_{xv}} V_{12}^T R V_{12} + \frac{\partial g_k}{\partial s_{xu}} U_{12}^T R V_{12} \\ & + 2\frac{\partial g_k}{\partial r_v} V_{12}^T R U_{12} + \frac{\partial g_k}{\partial s_{xv}} X_{ce} R U_{12} + \frac{\partial g_k}{\partial s_{vu}} U_{12}^T R U_{12}, \\ f_{k+1} =& \frac{\partial f_k}{\partial v_3} u_3, \end{aligned}$$

one gets the result. □

From Proposition 2.5, one gets directly that:

$$\begin{aligned} d\mathcal{O}_U(h)|_{(X,V)} = Span & \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -h_M R X_{ce} \\ 1 \\ 0 \\ 0 \end{bmatrix}, \right. \\ & \left. \left(\begin{bmatrix} h_M \left(-g_k R X_{ce} + \frac{\partial g_k}{\partial s_{xv}} R V_{12} + \frac{\partial g_k}{\partial s_{xu}} R U_{12} \right) \\ 0 \\ h_M \left(\frac{\partial g_k}{\partial s_{xv}} R X_{ce} + 2\frac{\partial g_k}{\partial r_v} R V_{12} + \frac{\partial g_k}{\partial s_{vu}} R U_{12} \right) \\ \frac{\partial f_k}{\partial v_3} \end{bmatrix} \right)_{k \geq 1} \right). \end{aligned} \tag{2.22}$$

Suppose that V_{12} is not colinear to U_{12}. Let us define (x_v, x_u) as the coordinates of X_{ce} in the basis formed by (V_{12}, U_{12}). We can now state the equivalent of Proposition 2.5 for the double integrator:

Proposition 2.6. *If V_{12} is not colinear to U_{12}, $\exists k \geq 1$, $\exists l \geq 1$ such that $k \neq l$ and:*

$$\begin{vmatrix} -x_v g_k + \frac{\partial g_k}{\partial s_{xv}} & -x_v g_l + \frac{\partial g_l}{\partial s_{xv}} \\ -x_u g_k + \frac{\partial g_k}{\partial s_{vu}} & -x_u g_l + \frac{\partial g_l}{\partial s_{vu}} \end{vmatrix} \neq 0, \tag{2.23}$$

the system (2.3), (2.8) is locally weakly observable at (X, V) with input U.

Proof. Following the same reasoning as in Proposition 2.3, one only needs to compute the following determinant:

$$D := \begin{vmatrix} -h_M R X_{ce} & h_M(-g_k R X_{ce} + \frac{\partial g_k}{\partial s_{xv}} R V_{12} + \frac{\partial g_k}{\partial s_{xu}} R U_{12}) & h_M(-g_l R X_{ce} + \frac{\partial g_l}{\partial s_{xv}} R V_{12} + \frac{\partial g_l}{\partial s_{xu}} R U_{12}) \\ 1 & 0 & 0 \end{vmatrix},$$

$$D = h_m^2 \det(R) \det\left(\begin{bmatrix} V_{12} & U_{12} \end{bmatrix} \right) \begin{vmatrix} -x_v g_k + \frac{\partial g_k}{\partial s_{xv}} & -x_v g_l + \frac{\partial g_l}{\partial s_{xv}} \\ -x_u g_k + \frac{\partial g_k}{\partial s_{vu}} & -x_u g_l + \frac{\partial g_l}{\partial s_{vu}} \end{vmatrix}.$$

As in proposition (2.3), $D \neq 0$ and the system (2.3), (2.8) is locally weakly observable at (X, V) with input U. □

The property (2.23) is expected to be verified generically as (2.17). The property that V_{12} and U_{12} must not be colinear is more useful than its counterpart in the simple integrator case as both V_{12} and U_{12} are known. It is not surprising as more information is available in the double integrator case than in the simple integrator one. It also matches the condition in the polynomial case.

We now extend these results to the more realistic case of a map as sum of Gaussian.

2.2.2.3 Map as a sum of Gaussian

Let us set:

$$X_{ce}^i = \begin{bmatrix} x_1 - x_1^i \\ x_2 - x_2^i \end{bmatrix}.$$

We recall that:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^l(X,V) \end{bmatrix},$$

with $h^{nl}(X) = x_3 - \sum_{i=1}^{n_g} h_M^i(x_1,x_2)$ and $h^l(X,V) = V$.

We first notice that $\nabla_{(x_1,x_2)} h_M^i = -h_M^i R^i X_{ce}^i$.

Simple integrator

$$\nabla h_0^{nl} = \begin{bmatrix} \sum_{i=1}^{n_g} h_M^i R^i X_{ce}^i \\ 1 \end{bmatrix}, \tag{2.24}$$

$$\begin{aligned} h_1^{nl} &= \nabla^T h_0^{nl} V, \\ &= v_3 + \sum_{i=1}^{n_g} h_M^i X_{ce}^{iT} R^i V_{12}, \end{aligned} \tag{2.25}$$

$$\tag{2.26}$$

Following the same path as in the case of one Gaussian, proposition (2.7) is stated as follows.

Proposition 2.7. *$\forall k \geq 1$, $\exists\, g_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R} \longrightarrow \mathbb{R}$ such that:*

$$h_k^{nl}(X,V) = f_k(v_3) + \sum_{i=1}^{n_g} h_M^i(g_k(V_{12}^T R^i V_{12}, X_{ce}^{iT} R^i V_{12})),$$

and $\forall (r,s) \in \mathbb{R}^+ \times \mathbb{R}$, $g_k(r,s)$ is polynomial in (r,s) in s. Besides, $g_k(r,s)$ is of degree n. Moreover, the following recursion holds for the functions g_k, $\forall (r,s) \in \mathbb{R}^+ \times \mathbb{R}$ and for $n \geq 1$:

$$g_{k+1}(r,s) = -g_k(r,s)s + \frac{\partial g_k}{\partial s}(r,s)r$$

Proof. The proposition is true for $k = 1$ according to equation (2.25) with $g_1(r,s) = s$ and $f_1(v_3) = v_3$

Let us assume that, for some $k \geq 1$, there exists $g_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R} \longrightarrow \mathbb{R}$ such that:

$$h_k^{nl}(X,V) = f_k(v_3) + \sum_{i=1}^{n_g} h_M^i(g_k(V_{12}^T R^i V_{12}, X_{ce}^{iT} R^i V_{12})).$$

Then by definition,

$$h_{k+1}^{nl} = V^T \nabla h_k^{nl}.$$

To simplify the notations we denote $g_k(V_{12}^T R^i V_{12}, X_{ce}^{iT} R^i V_{12})$ by g_k^i and $\frac{\partial g_k}{\partial s}(V_{12}^T R^i V_{12}, X_{ce}^{iT} R^i V_{12})$ by $\frac{\partial g_k^i}{\partial s}$:

$$\nabla h_k^{nl} = \begin{bmatrix} -\sum_{i=1}^{n_g} g_k^i R^i X_{ce}^i + \sum_{i=1}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s} R^i V_{12} \\ 0 \end{bmatrix}. \tag{2.27}$$

Finally,

$$h_{k+1}^{nl} = \sum_{i=1}^{n_g} h_M^i (-(X_{ce}^{iT} R^i V_{12}) g_k^i + V_{12}^T R^i V_{12} \frac{\partial g_k^i}{\partial s}).$$

By setting, $f_{k+1} = 0$ and $g_{k+1}(r,s) = -g_k(r,s)s + \frac{\partial g_k}{\partial s}(r,s)r$, and by noticing that if g_k is of degree n is s then g_{k+1} is of degree $k+1$ in s, the result is proved $\forall k \geq 1$.

□

By equations and (2.24) and (2.27), one can see that:

$$d\mathcal{O}_U(h)|_X = Span\left(\begin{bmatrix} \sum_{i=1}^{n_g} h_M^i R^i X_{ce}^i \\ 1 \end{bmatrix}, \left(\begin{bmatrix} -\sum_{i=1}^{n_g} g_k^i R^i X_{ce}^i + \sum_{i=1}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s} R^i V_{12} \\ 0 \end{bmatrix} \right)_{n \geq 1} \right). \tag{2.28}$$

Simple analytic conditions for local weak observability in this case are harder to provide as the directions spanned by ∇h_k^{nl} are not a combination of X_{ce} and V_{12} but a combination between the X_{ce}^i and V_{12}. However, this new complexity seems beneficial, at least intuitively, as it means that more directions are spanned in the case of a sum of Gaussian than in the case of one Gaussian. For example, if $n_g = 2$, $R^1 = R^2$, and X_{ce}^1, X_{ce}^2 and V_{12} are colinear then one can show that rank condition is not satisfied, using that the same technique as in the single Gaussian case. However, it is a very particular case, with many symmetries. It means notably that the drone is located on the line that connects the centres of the two Gaussians and goes in that direction too. Moreover, if the drone deviates from this line the rank condition is satisfied. This example illustrates the fact that the rank condition seems to fail only in very special cases.

Double integrator

Using the same notations as in the single Gaussian case:

$$\begin{aligned} h_0^{nl} &= x_3 - \sum_{i=0}^{n_g} h_M^i, \\ \nabla h_0^{nl} &= \begin{bmatrix} -\sum_{i=0}^{n_g} h_M^i R^i X_{ce}^i \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \end{aligned} \tag{2.29}$$

$$h_1^{nl} = \nabla^T h_0^{nl} \begin{bmatrix} V \\ U \end{bmatrix}, \tag{2.30}$$
$$= v_3 + \sum_{i=0}^{n_g} X_{ce}^{iT} R^i V_{12} h_M^i,$$
$$\nabla h_1^{nl} = \begin{bmatrix} -\sum_{i=0}^{n_g} h_M^1 (X_{ce}^{iT} R^i V_{12}) R^i X_{ce}^i + \sum_{i=0}^{n_g} h_M^i R^i V_{12} \\ 0 \\ \sum_{i=0}^{n_g} h_M^i R^i X_{ce}^i \\ 1 \end{bmatrix}. \tag{2.31}$$

Proposition 2.8. *$\forall k \geq 1$, $\exists\ g_k : \mathbb{R}^5 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that:*

$$h_k^{nl}(X, V, U) = f_k(v_3, u_3) + \sum_{i=0}^{n_g} h_M^i (g_k(V_{12}^T R^i V_{12}, U_{12}^T R^i U_{12}, X_{ce}^{iT} R^i V_{12}, X_{ce}^{iT} R^i U_{12}, V_{12}^T R^i U_{12})).$$

and $\forall (r_u, r_v, s_{xv}, s_{xu}, s_{vu}) \in (\mathbb{R}^+)^2 \times \mathbb{R}^3$, $g_k(r_u, r_v, s_{xv}, s_{xu}, s_{vu})$ is polynomial in $(r_u, r_v, s_{xv}, s_{xu}, s_{vu})$. Moreover, a possible choice of the function sequence $(g_k)_{k\geq 1}$ is the one for which the following recursion holds for , $\forall (r_u, r_v, s_{xv}, s_{xu}, s_{vu}) \in \mathbb{R}^{+2} \times \mathbb{R}^3$ and for $k \geq 1$:

$$g_{k+1} = -g_k s_{xv} + \frac{\partial g_k}{\partial s_{xv}} r_v + \frac{\partial g_k}{\partial s_{vu}} r_u + \left(\frac{\partial g_k}{\partial s_{xu}} + 2 \frac{\partial g_k}{\partial r_v} \right) s_{vu} + \frac{\partial g_k}{\partial s_{xv}} s_{xu}.$$

Proof. The proposition is true for $k = 1$ according to equation (2.19) with $g_1(r_u, r_v, s_{xv}, s_{xu}, s_{vu}) = s_{xv}$ and $f_1(v_3, u_3) = v_3$

Let us assume that, for some $k \geq 1$, there exists $g_k : \mathbb{R}^5 \longrightarrow \mathbb{R}$ and $f_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that:

$$h_k^{nl}(X, V, U) = f_k(v_3, u_3) + \sum_{i=0}^{n_g} h_M^i (g_k(V_{12}^T R^i V_{12}, U_{12}^T R^i U_{12}, X_{ce}^{iT} R^i V_{12}, X_{ce}^{iT} R^i U_{12}, V_{12}^T R^i U_{12})).$$

Then by definition,

$$h_{k+1}^{nl} = \begin{bmatrix} V^T & U^T \end{bmatrix} \nabla h_k^{nl}.$$

Moreover, for $k \geq 1$,

$$\nabla h_k^{nl} = \begin{bmatrix} -\sum_{i=0}^{n_g} h_M^i g_k^i R^i X_{ce}^i + \sum_{i=0}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s_{xv}} R^i V_{12} + \sum_{i=0}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s_{xu}} R^i U_{12} \\ 0 \\ \sum_{i=0}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s_{xv}} R^i X_{ce}^i + 2 \sum_{i=0}^{n_g} h_M^i \frac{\partial g_k^i}{\partial r_v} R^i V_{12} + \sum_{i=0}^{n_g} h_M^i \frac{\partial g_k^i}{\partial s_{vu}} R^i U_{12} \\ \frac{\partial f_k}{\partial v_3} \end{bmatrix}.$$

The result follows as in (2.5), by setting, $\forall (r_u, r_v, s_{xv}, s_{xu}, s_{vu}) \in \mathbb{R}^{+2} \times \mathbb{R}^3$ and for $k \geq 1$:

$$g_{k+1} = -g_k s_{xv} + \frac{\partial g_k}{\partial s_{xv}} r_v + \frac{\partial g_k}{\partial s_{vu}} r_u + \left(\frac{\partial g_k}{\partial s_{xu}} + 2 \frac{\partial g_k}{\partial r_v} \right) s_{vu} + \frac{\partial g_k}{\partial s_{xv}} s_{xu}.$$

□

From proposition (2.8), one gets directly that:

$$d\mathcal{O}_U(h)|_{(X,V)}=Span\left(\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix},\begin{bmatrix}-\sum_{i=0}^{n_g}h_M^iR^iX_{ce}^i\\1\\0\\0\\0\\0\end{bmatrix},\right.$$
$$\left.\left(\begin{bmatrix}-\sum_{i=0}^{n_g}h_M^ig_k^iR^iX_{ce}^i+\sum_{i=0}^{n_g}h_M^i\frac{\partial g_k^i}{\partial s_{xv}}R^iV_{12}+\sum_{i=0}^{n_g}h_M^i\frac{\partial g_k^i}{\partial s_{xu}}R^iU_{12}\\0\\\sum_{i=0}^{n_g}h_M^i\frac{\partial g_k^i}{\partial s_{xv}}R^iX_{ce}^i+2\sum_{i=0}^{n_g}h_M^i\frac{\partial g_k^i}{\partial r_v}R^iV_{12}+\sum_{i=0}^{n_g}h_M^i\frac{\partial g_k^i}{\partial s_{vu}}R^iU_{12}\\\frac{\partial f_k}{\partial v_3}\end{bmatrix}\right)_{k\geq 1}\right). \tag{2.32}$$

As in the previous cases, the rank condition will be verified for a generical choice of vector V_{12} and U_{12} if the sum of Gaussian does not represent a pathological map, with concentric Gaussian for example.

The main conclusion that can be drawn from the study of the observability condition is that the more complex the map is the more information it contains and the more likely it is to lead to observable systems. It confirms the previous informal claim However, the conditions presented in this section, especially on the control, are qualitative but not quantitative. They give a good insight on the nature of the problem but no practical resolution scheme.

The goal of the next chapter is to find more quantitative and applicable conditions through observer and controller design.

Chapter 3

Observer and controller design for TAN with analytical maps under persistence of excitation

This chapter is firstly dedicated to observer design for the systems studied in Chapter 2. Actually, it turns out that these observers converge under a condition of persistence of excitation on the speed. This implies that there are also conditions on the control to ensure the convergence of the observer . Therefore, the second goal of this chapter is to design controllers that simultaneously verify these conditions and allow the system to reach his guiding objective. Finally, we study quite independently a special case of stochastic persistent horizontal speed.

3.1 Nonlinear observer design by Immersion and Invariance (I&I) for Terrain-Aided Navigation

Generally, an observer is defined as a dynamical system designed to asymptotically approach the unknown state of the original system. The main difficulty is that it must be constructed using only the available information.There is no explicit form of observers that handles any system so it must be adapted to the particular class of systems treated. The most popular nonlinear observer are the nonlinear Luenberger and Kalman observers. They are especially well suited for a large class of nonlinear systems that still exhibit some linear structure (see [Besançon, 2007]). Another family of nonlinear observers are high-gain observers. They are based on the idea of compensating the nonlinearity buy using a high gain on the innovation term. This also requires in general that the original system contains a linear detectable part (see [Khalil and Praly, 2014]). In the sequel, we prefer to use a more general method, called Immersion and Invariance (I&I), that is known to be able to deal with very nonlinear systems. Besides, it gives one many degrees of freedom to choose the estimation error equation. In this section, we first recall the general concepts of I&I applied to observer design. Then, we present the design of observers for the systems with a quadratic map, a cubic map, a Gaussian map and a trigonometric map.

3.1.1 Description of I&I for the design of observers of a general nonlinear system

To recall the general method of nonlinear observer design based on Immersion and Invariance, taken from [Astolfi et al., 2008], we first rewrite the system (2.1) under the following form:

$$\begin{aligned} \dot{\eta} &= f_1(\eta, y, u), \\ \dot{y} &= f_2(\eta, y, u), \end{aligned} \tag{3.1}$$

where $\eta \in \mathbb{R}^n$ is the unmeasured part of the state, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}^p$ is the measured part of the state and f_1 and f_2 correspond to the dynamics of η and y. It can always be done from the system (2.1) if h is sufficiently smooth. The solutions of equation (3.1) are supposed to be forward complete for simplicity. In the sequel, we recall the general definition of an observer from [Astolfi et al., 2008] so that we can deal with all the different kinds of model of TAN in the same framework.

Definition 3.1. *The dynamical system*

$$\dot{\xi} = \alpha(\xi, y, u), \tag{3.2}$$

with $\xi \in \mathbb{R}^q$, $q \geq n$, is called a observer for the system (3.1) if there exist mappings $\beta : \mathbb{R}^q \times \mathbb{R}^p \times \mathbb{R} \longrightarrow \mathbb{R}^q$ and $\phi : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R} \longrightarrow \mathbb{R}^q$ that are left-invertible with respect to their first argument and such that the manifold:

$$\mathcal{M} = \{(\eta, y, \xi, t) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R} : \beta(\xi, y, t) = \phi(\eta, y, t)\}, \tag{3.3}$$

satisfies the following properties:

(i) all trajectories of the extended system (3.1), (3.2) that start in $\mathcal{M}$ stay in $\mathcal{M}$ at all times i.e. it is positively invariant;

(ii) all trajectories of the extended system (3.1), (3.2) starting in a neighbourhood of $\mathcal{M}$ asymptotically converge to $\mathcal{M}$ i.e $\mathcal{M}$ is locally attractive;

If the attractivity property (ii) in Definition 3.1 holds for any $(\eta(t_0), y(t_0), \xi(t_0), t_0) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}$ then the system (3.2) is called a global observer.

The term, $\beta(\xi, y, t) - \phi(\eta, y, t)$ is in fact a nonlinear estimation error written in a very general form and $\mathcal{M}$ is the subset of the extended space state/observer where this estimation error vanishes. From the design of an observer ξ in definition (3.1) one can define an estimator of η as follows:

$$\hat{\eta} = \phi^L(\beta(\xi, y, t), y, t), \tag{3.4}$$

where ϕ^L is a left-inverse of ϕ.

The following proposition gives a general technique of construction of observers as described in Definition 3.1.

Proposition 3.1. *Consider the system (3.1), (3.2) and assume that there exist $\mathcal{C}^1$ mappings β and ϕ as in definition (3.1) with the left-inverse of ϕ being ϕ^L. Assume in addition that:*

1. *For any y, t, $\beta(\cdot, y, t)$ is left-invertible and:*

$$det\left(\frac{\partial \beta}{\partial \xi}\right) \neq 0.$$

2. *The system*

$$\dot{z} = -\frac{\partial \beta}{\partial y}(f_2(\hat{\eta}, y, u) - f_2(\eta, y, u)) + \left.\frac{\partial \phi}{\partial y}\right|_{\eta=\hat{\eta}} f_2(\hat{\eta}, y, u) - \frac{\partial \phi}{\partial y} f_2(\eta, y, u) \tag{3.5}$$
$$+ \left.\frac{\partial \phi}{\partial \eta}\right|_{\eta=\hat{\eta}} f_1(\hat{\eta}, y, u) - \frac{\partial \phi}{\partial \eta} f_1(\eta, y, u) + \left.\frac{\partial \phi}{\partial t}\right|_{\eta=\hat{\eta}} - \frac{\partial \phi}{\partial t},$$

with $\hat{\eta} = \phi^L(\phi(\eta, y, t) + z)$ has a locally (resp. globally) asymptotically stable equilibrium at $z = 0$, uniformly in η, y, and t.

Then the system (3.2) with:

$$\alpha(\xi, y, t) = -\left(\frac{\partial \beta}{\partial \xi}\right)^{-1} \left(\frac{\partial \beta}{\partial y} f_2(\hat{\eta}, y, u) + \frac{\partial \beta}{\partial t} - \left.\frac{\partial \phi}{\partial y}\right|_{\eta=\hat{\eta}} f_2(\hat{\eta}, y, u) \right. \tag{3.6}$$
$$\left. - \left.\frac{\partial \phi}{\partial \eta}\right|_{\eta=\hat{\eta}} f_1(\hat{\eta}, y, u) - \left.\frac{\partial \phi}{\partial t}\right|_{\eta=\hat{\eta}} \right), \tag{3.7}$$

where $\hat{\eta} = \phi^L(\beta(\xi, y, t))$, is a an observer (resp. a global observer) for the system (3.1).

Proof. See [Astolfi et al., 2008] for the complete proof. However, here is an outline of the proof. Consider the nonlinear estimation error variables (also called off-the-manifold variables) $z = \beta(\xi, y, t) - \phi(\eta, y, t)$. Then, by computing $\dot{z}$ and substituting with equation (3.1), one can show that z is of the form (3.5). One can conclude, by assumption 2, that the associated manifold $\mathcal{M}$ is positively invariant and attractive which leads to the result. □

Remark 3.1.

- Unlike Luenberger observers or others classical observers, the quantity that is defined as an observer ξ does not directly approach η. We build instead an estimator $\hat{\eta}$ that is a function of ξ and y as described in equation (3.4). This usually allows one to incorporate additional nonlinear terms in the observer and deal with very nonlinear systems.

- Actually, left-invertibility of ϕ implies that $\hat{\eta} - \eta = 0$ on $\mathcal{M}$ and attractivity of $\mathcal{M}$ implies that $\beta(\xi, y, t) - \phi(\eta, y, t)$ tends to 0 as $t \to +\infty$. However, attractivity of $\mathcal{M}$ does not necessarily imply that $\hat{\eta} - \eta$ tends to 0 as $t \to +\infty$. For this to be true, one needs an additional regularity assumption on ϕ^L. Typically, if ϕ^L is globally Lipschitz with respect to η and uniformly in t and y, then $\exists C > 0$ such that $\forall t$, $\forall y$, $\forall \xi$:

$$\|\hat{\eta} - \eta\| \leq C \|\beta(\xi, y, t) - \phi(\eta, y, t)\|,$$

and clearly $\hat{\eta} - \eta \to 0$ if $\beta(\xi, y, t) - \phi(\eta, y, t) \to 0$ which makes $\hat{\eta}$ a converging estimator. For instance, if $\phi^L(\chi, y, t) = E(y, t)\chi + F(y, t)$, where E is a bounded matrix with respect to y and t, then ϕ^L is globally Lipschitz with respect to η and uniformly in t and y. The assumption of uniformity w.r.t. y and t is important. For example, consider the dynamics of a simple integrator (2.2) with known speed with $\eta = X$:

$$\dot{\eta} = V.$$

It is well known that the speed is not enough to reconstruct the position. Nevertheless, if one chooses $\beta(\xi, y, t) = \xi$, $\phi(\eta, y, t) = e^{-t}\eta$ and $\alpha(\xi, y, t) = -\xi + e^{-t}V$, one gets that the dynamics of $z = \beta(\xi, y, t) - \phi(\eta, y, t)$ is :

$$\dot{z} = -z.$$

Obviously, $z \to 0$ as $t \to +\infty$ and $\hat{\eta} = e^t\xi$. However, by construction:

$$\begin{aligned}\dot{\hat{\eta}} &= e^t\dot{\xi} + e^t\xi, \\ &= -e^t\xi + V + e^t\xi, \\ &= V.\end{aligned}$$

So, $\hat{\eta} - \eta$ is constant. Usually, $\hat{\eta} - \eta \neq 0$ initially so $\hat{\eta} - \eta$ does not tend to zero while z always does. In this case, $n = q$, ϕ is invertible and $\phi^L(\xi, y, t) = e^t\xi$. Notice that ϕ^L is clearly Lipschitz w.r.t. ξ but not uniformly in t as its Lipschitz constant is e^t.

As α is uniquely defined by ϕ, β and the dynamics of (3.1), the problem of finding a observer is reduced to finding ϕ and β such that Assumption 2 is satisfied. It is a very complicated problem in general but in the following we will use particular cases of observers:

In the sequel, we apply the general technique of Proposition 3.1 to the design of observers for Terrain-Aided Navigation.

3.1.2 Application of I&I design to Terrain-Aided Navigation

In this section, we present the construction of observers for the dynamics (2.3) with the models of maps (2.6) to (2.11). Two different versions of I&I are applied:

- The first one consists in considering the nonlinear systems in its original state space, i.e setting $q = n$, and attacking directly the nonlinearity such that:

  $$\begin{aligned}\beta(\xi, y, t) &= \xi + \psi(y), \\ \phi(\eta, y, t) &= \eta.\end{aligned}$$

  In this case, one keeps the original variables η and uses $\psi(y)$ as a corrective term to the observer ξ. The addition of ψ is meant to give enough supplementary freedom in the total design to deal with complex nonlinear systems.

- The second approach is to immerse the system into a higher dimensional space where the new system has more structure i.e. choosing $q > n$, such that:

  $$\begin{aligned}\chi &:= \phi(\eta, y, t) = \phi(\eta), \\ \hat{\chi} &:= \beta(\xi, y, t) = \xi.\end{aligned}$$

  Hopefully, with an appropriate choice of ϕ, the dynamics of χ is simple and $\hat{\chi}$ can be chosen as an existing observers.

The first approach is successfully applied to the case of polynomial map of degree 2 and to the Gaussian case with an additional measurement of altitude. This method fails as the maps becomes more complicated e.g. a sum of Gaussian, a polynomial of degree 3 or sinusoidal. The second approach is successfully

applied to the polynomial map of degree 3 the map represented by a truncated Fourier series. We recall that the dynamics considered in the following is:

$$\dot{X} = V,$$
$$\dot{V} = U.$$

3.1.2.1 Direct method

We restrict ourselves to functions ϕ and β of the following form:

$$\beta(\xi, y, t) = \xi + \psi(y),$$
$$\phi(\eta, y, t) = \eta$$

where $\psi : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ and $\dot{\xi} = \alpha(\xi, y, U)$ are to be chosen. It appears that such a restriction is usual and allows one to solve many problems (see [Astolfi et al., 2008]) These choices of ϕ and β clearly satisfiy the required invertibility properties.

Polynomial map of degree 2

In this section, we consider the following observation equation:

$$h(X, V) = \begin{bmatrix} h^{nl}(X, V) \\ h^{l}(X, V) \end{bmatrix},$$

with $h^l(X, V) = V$, $h^{nl}(X, V) = x_3 - h_M(x_1, x_2)$, and

$$h_M(x_1, x_2) = a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2 + a_{10}x_1 + a_{01}x_2 + a_{00}.$$

We set $\eta = X = (x_1, x_2, x_3)$.

The objective is to define an observer of η for the model of map (2.6) denoted by $\hat{\eta} = (\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_2)$. To do so, we define the estimation error, z, such that:

$$z := \xi + \psi(y) - \eta := \widehat{\eta} - \eta. \tag{3.8}$$

The dynamics of z can be deduced:

$$\dot{z} = \dot{\xi} + \frac{\partial \psi}{\partial y}\dot{y} - \dot{\eta},$$
$$\dot{z} = \dot{\xi} + \frac{\partial \psi}{\partial y}\begin{bmatrix} v_3 - (2a_{20}x_1 + a_{11}x_2 + a_{10})v_1 - (a_{11}x_1 + 2a_{02}x_2 + a_{01})v_2 \\ U \end{bmatrix} - V,$$
$$\dot{z} = \dot{\xi} + \frac{\partial \psi}{\partial y}\begin{bmatrix} -(2a_{20}v_1 + a_{11}v_2) & -(a_{11}v_1 + 2a_{02}v_2) & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\eta + \frac{\partial \psi}{\partial y}\begin{bmatrix} v_3 - a_{10}v_1 - a_{01}v_2 \\ U \end{bmatrix} - V.$$

We recall that $\frac{\partial \psi}{\partial y} = \begin{bmatrix} \frac{\partial \psi}{\partial h^{nl}} & \frac{\partial \psi}{\partial V} \end{bmatrix}$. This leads to:

$$\dot{z} = \dot{\xi} + \frac{\partial \psi}{\partial h^{nl}}\begin{bmatrix} -(2a_{20}v_1 + a_{11}v_2) & -(a_{11}v_1 + 2a_{02}v_2) & 0 \end{bmatrix}\eta + \frac{\partial \psi}{\partial y}\begin{bmatrix} v_3 - a_{10}v_1 - a_{01}v_2 \\ U \end{bmatrix} - V.$$

Thus, ξ is defined by:

$$\dot{\xi} = \alpha(\xi, y, U) := -\frac{\partial \psi}{\partial h^{nl}} \begin{bmatrix} -(2a_{20}v_1 + a_{11}v_2) & -(a_{11}v_1 + 2a_{02}v_2) & 0 \end{bmatrix} \hat{\eta} - \frac{\partial \psi}{\partial y} \begin{bmatrix} v_3 - a_{10}v_1 - a_{01}v_2 \\ U \end{bmatrix} + V + \zeta.$$

where $\zeta : \mathbb{R}^+ \longrightarrow \mathbb{R}^6$ is a term added to deal with altitude estimation. Moreover we set:

$$\Sigma := \begin{bmatrix} -(2a_{20}v_1 + a_{11}v_2) & -(a_{11}v_1 + 2a_{02}v_2) & 0 \end{bmatrix}^T, \tag{3.9}$$

$$\frac{\partial \psi}{\partial h^{nl}} := \kappa_{12}\Sigma, \tag{3.10}$$

with $\kappa_{12} > 0$. The idea behind this choice is to get the following error dynamics:

$$\dot{z} = -\kappa_{12}\Sigma\Sigma^T z + \zeta, \tag{3.11}$$

$$\dot{z} = -\kappa_{12} \begin{bmatrix} (2a_{20}v_1 + a_{11}v_2)^2 & (2a_{20}v_1 + a_{11}v_2)(a_{11}v_1 + 2a_{02}v_2) & 0 \\ (2a_{20}v_1 + a_{11}v_2)(a_{11}v_1 + 2a_{02}v_2) & (a_{11}v_1 + 2a_{02}v_2)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} z + \zeta. \tag{3.12}$$

Let us define $\hat{h}^{nl} := \hat{\eta}_3 - h_M(\hat{\eta}_1, \hat{\eta}_2)$. Then,

$$\hat{h}^{nl} - h^{nl} = (\hat{\eta}_3 - \eta_3) - a_{20}(\hat{\eta}_1^2 - \eta_1^2) - a_{11}(\hat{\eta}_1\hat{\eta}_2 - \eta_1\eta_2) - a_{02}(\hat{\eta}_2^2 - \eta_2^2) - a_{10}(\hat{\eta}_1 - \eta_1) - a_{01}(\hat{\eta}_2 - \eta_2),$$

$$\hat{h}^{nl} - h^{nl} = (\hat{\eta}_3 - \eta_3) - a_{20}(\hat{\eta}_1^2 - \eta_1^2) - a_{11}(\hat{\eta}_1\hat{\eta}_2 - \eta_1\eta_2) - a_{02}(\hat{\eta}_2^2 - \eta_2^2) + \begin{bmatrix} -a_{10} & -a_{01} & 1 \end{bmatrix} z.$$

It can be seen from equation (3.12) that the estimation in (x_1, x_2) can be treated independently of the estimation of x_3 with an appropriate choice of ζ. That is why, ζ is defined as follows:

$$\zeta := \begin{bmatrix} 0 \\ 0 \\ -\kappa_3(\hat{h}^{nl} - h^{nl}) \end{bmatrix},$$

with $\kappa_3 > 0$. Finally, the error dynamics reads:

$$\dot{z} = -\begin{bmatrix} \kappa_{12} \begin{bmatrix} (2a_{20}v_1 + a_{11}v_2)^2 & (2a_{20}v_1 + a_{11}v_2)(a_{11}v_1 + 2a_{02}v_2) \\ (2a_{20}v_1 + a_{11}v_2)(a_{11}v_1 + 2a_{02}v_2) & (a_{11}v_1 + 2a_{02}v_2)^2 \end{bmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} -\kappa_3 a_{10} & -\kappa_3 a_{01} \end{matrix} & \kappa_3 \end{bmatrix} z + \kappa_3 \begin{bmatrix} 0 \\ 0 \\ -a_{20}(\hat{\eta}_1^2 - \eta_1^2) - a_{11}(\hat{\eta}_1\hat{\eta}_2 - \eta_1\eta_2) - a_{02}(\hat{\eta}_2^2 - \eta_2^2) \end{bmatrix}. \tag{3.13}$$

Moreover, define the matrix

$$\Sigma_0 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \Sigma. \tag{3.14}$$

The goal of the sequel is to show the convergence of the system (3.13) to 0 by a Lyapunov method. Note that the subsystem of (3.13) that corresponds to the error in horizontal position is a linear time varying system that is well known in adaptive control. It is known to converge under an assumption of boundedness and of persistence of excitation of Σ_0.

Assumption 3.1. *Persistence of excitation*
There exist some constants $T, \mu > 0$, such that for all $t \geq 0$,

$$\frac{1}{T}\int_t^{t+T} \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau \succeq \mu I, \tag{3.15}$$

where $\succeq$ represents the classical partial order on symmetric semi-definite positive matrix.

Assumption 3.2. *The functions $\Sigma_0(t)$, $\frac{\partial \Sigma_0}{\partial t}(t)$ are uniformly bounded, that is $\|\Sigma_0(t)\| \leq c_\Sigma$ and $\frac{\partial \Sigma_0}{\partial t}(t) \leq c_{\dot{\Sigma}}$.*

Assumption 3.3. *The horizontal position coordinates $(x_1, x_2) = (\eta_1, \eta_2)$ are bounded, i.e. $|\eta_1| \leq c_1$ $|\eta_2| \leq c_2$.*

Proposition 3.2. *Under Assumptions 3.1, 3.2, 3.3 and the gain selection (3.19), the system (3.13) is uniformly globally exponentially stable (UGES) at $z = 0$.*

Proof. Define the candidate Lyapunov function for the subystem $z_{12} = (z_1, z_2) := (\hat{\eta}_1 - \eta_1, \hat{\eta}_2 - \eta_2)$ as

$$\mathcal{V}_{12} := \kappa_{12} z_{12}^T \Psi(t) z_{12} + \frac{\kappa}{2}|z_{12}|^2,$$

where $\kappa > 0$ and $\Psi(t)$ is the 2×2 symmetric matrix defined by:

$$\Psi(t) := (1 + c_\Sigma^2)TI - \frac{1}{T}\int_t^{t+T}\int_t^s \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau ds,$$

with $c_\Sigma \geq \|\Sigma_0\|$ for the matrix norm induced by the 2-norm on $\mathbb{R}^3$ and T coming from Assumption 3.1.

The form of $\mathcal{V}_{12}$ is inspired by the construction of strict Lyapunov functions for persistently excited time-varying systems in [Maghenem and Loría, 2017] and [Malisoff and Mazenc, 2009]. Notice $\Psi(t)$ satisfies

$$\begin{aligned} TI \preceq \Psi(t) &\preceq (1 + c_\Sigma^2)TI, \\ \dot{\Psi}(t) &= -\frac{1}{T}\int_t^{t+T} \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau + \Sigma_0(t)\Sigma_0^T(t). \end{aligned} \tag{3.16}$$

Under the working assumptions, $\mathcal{V}_{12}$ is clearly positive definite and radially unbounded. Its time derivative along trajectories of the error dynamics reads:

$$\begin{aligned} \dot{\mathcal{V}}_{12} = &- \kappa\kappa_{12} z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - \frac{\kappa_{12}}{T} z_{12}^T \left(\int_t^{t+T} \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau\right) z_{12} \\ &+ \kappa_{12} z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - 2\kappa_{12}^2 z_{12}^T \Psi(t)\Sigma_0(t)\Sigma_0^T(t) z_{12}. \end{aligned}$$

By definition of Ψ and using Assumption 3.1:

$$\begin{aligned} \dot{\mathcal{V}}_{12} \leq &- \kappa_{12}(\kappa - 1) z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - \kappa\mu|z_{12}|^2 - 2\kappa_{12}^2(1 + c_\Sigma^2)T z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} \\ &- \frac{2\kappa_{12}^2}{T} z_{12}^T \left(\int_t^{t+T}\int_t^s \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau ds\right)\Sigma_0(t)\Sigma_0^T(t) z_{12}, \end{aligned}$$

By Cauchy-Swartz inequality and using Assumption 3.2 and the bounds on Ψ:

$$\begin{aligned}\dot{\mathcal{V}}_{12} \leq & -\kappa_{12}\Big(\kappa + 2\kappa_{12}(1+c_\Sigma^2)T - 1\Big) z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - \kappa_{12}\mu|z_{12}|^2 \\ & + \frac{2\kappa_{12}^2}{T}|z_{12}|\left|\int_t^{t+T}\int_t^s \Sigma_0(\tau)\Sigma_0^T(\tau)d\tau ds\right| |\Sigma_0(t)||\Sigma_0^T(t)z_{12}|, \\ \leq & -\kappa_{12}\Big(\kappa + 2\kappa_{12}(1+c_\Sigma^2)T - 1\Big) z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - \kappa_{12}\mu|z_{12}|^2 \\ & + 2\kappa_{12}^2 c_\Sigma^2 T c_\Sigma |z_{12}||\Sigma_0^T(t)z_{12}|,\end{aligned}$$

By applying Young's inequality on $|z_{12}||\Sigma_0^T(t)z_{12}|$ with constant $\epsilon' > 0$:

$$\begin{aligned}\dot{\mathcal{V}}_{12} \leq & -\kappa_{12}\Big(\kappa + 2\kappa_{12}(1+c_\Sigma^2)T - 1\Big) z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - \kappa_{12}\mu|z_{12}|^2 \\ & + \epsilon'|z_{12}|^2 + \frac{\kappa_{12}^4 c_\Sigma^3 T^2}{\epsilon'} z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12}, \\ = & -\kappa_{12}\left(\kappa + 2\kappa_{12}(1+c_\Sigma^2)T - 1 - \frac{\kappa_{12}^4 c_\Sigma^3 T^2}{\epsilon'}\right) z_{12}^T \Sigma_0(t)\Sigma_0^T(t) z_{12} - (\kappa_{12}\mu - \epsilon')|z_{12}|^2.\end{aligned}$$

With κ selected such that:

$$\kappa \geq 1 + \frac{\kappa_{12}^4 c_\Sigma^3 T^2}{\epsilon'} - 2\kappa_{12}^2(1+c_\Sigma^2)T, \tag{3.17}$$

one finally obtains:

$$\dot{\mathcal{V}}_{12} \leq -(\kappa_{12}\mu - \epsilon')|z_{12}|^2. \tag{3.18}$$

This establishes the uniform global exponential stability of the equilibrium $(z_1, z_2) = (0, 0)$.

We proceed now with the stability of the origin for the complete error dynamics. Take the positive definite function

$$\mathcal{V}_{123} := \mathcal{V}_{12} + \frac{1}{2}|z_3|^2.$$

Its derivative along trajectories of the z-dynamics reads

$$\begin{aligned}\dot{\mathcal{V}}_{123} \leq & -\kappa_{12}\mu(z_1^2 + z_2^2) - \kappa_3 z_3^2 + \kappa_3|a_{10}||z_1||z_3| + \kappa_3|a_{01}||z_2||z_3| \\ & + \kappa_3|z_3|\left|a_{20}(\hat{\eta}_1^2 - \eta_1^2) + a_{11}(\hat{\eta}_1\hat{\eta}_2 - \eta_1\eta_2) + a_{02}(\hat{\eta}_2^2 - \eta_2^2)\right|, \\ \leq & -\kappa_{12}\mu(z_1^2 + z_2^2) - \kappa_3 z_3^2 + \frac{\epsilon}{2}z_3^2 + \frac{k_3^2 a_{10}^2}{2\epsilon}z_1^2 + \frac{\epsilon}{2}z_3^2 + \frac{\kappa_3^2 a_{01}^2}{2\epsilon}z_2^2 \\ & + \kappa_3|z_3|\left|a_{20}(\hat{\eta}_1^2 - \eta_1^2) + a_{11}(\hat{\eta}_1\hat{\eta}_2 - \eta_1\eta_2) + a_{02}(\hat{\eta}_2^2 - \eta_2^2)\right|,\end{aligned}$$

where for the second inequality we have simply applied Young's inequality with $\epsilon > 0$. We now bound properly the last term in the previous inequalities. By using successively Young's inequality with $\epsilon > 0$,

and the bounds on (x_1, x_2) from Assumption 3.3, we have for each separate term

$$\begin{aligned}
\kappa_3|z_3||a_{20}(\hat{\eta}_1^2-\eta_1^2)| &= \kappa_3|z_3||a_{20}||z_1(z_1+2\eta_1)| \leq \kappa_3|a_{20}|(2c_1|z_3||z_1|+|z_3||z_1|^2),\\
&\leq \frac{\epsilon}{2}z_3^2+\frac{2c_1^2\kappa_3^2a_{20}^2}{\epsilon}z_1^2+ \leq \frac{\epsilon}{2}z_3^2+\frac{\kappa_3^2a_{20}^2}{2\epsilon}z_1^4,\\
&= \epsilon z_3^2+\frac{2c_1^2k_z^2a_{20}^2}{\epsilon}z_1^2+\frac{\kappa_3^2a_{20}^2}{2\epsilon}z_2^4,\\
\kappa_3|z_3||a_{02}(\hat{\eta}_2^2-\eta_2^2)| &\leq \epsilon z_3^2+\frac{2c_2^2\kappa_3^2a_{02}^2}{\epsilon}z_2^2+\frac{\kappa_3^2a_{02}^2}{2\epsilon}z_2^4,\\
\kappa_3|z_3||a_{11}(\hat{\eta}_1\hat{\eta}_2-\eta_1\eta_2)| &= \kappa_3|z_3||a_{11}||(\eta_2z_1+z_1z_2+\eta_1z_2)| \leq \kappa_3|z_3||a_{11}|(c_2|z_1|+|z_1||z_2|+c_1|z_2|),\\
&\leq \frac{\epsilon}{2}z_3^2+\frac{\kappa_3^2a_{11}^2c_2^2}{2\epsilon}z_1^2+\epsilon z_3^2+\frac{1}{4\epsilon}z_1^2z_2^2+\frac{\epsilon}{2}z_3^2+\frac{\kappa_3^2a_{11}^2c_1^2}{2\epsilon}z_2^2,\\
&\leq 2\epsilon z_3^2+\frac{\kappa_3^2a_{11}^2c_2^2}{2\epsilon}z_1^2+\frac{1}{8\epsilon}z_1^4+\frac{1}{8\epsilon}z_2^4+\frac{\kappa_3^2a_{11}^2c_1^2}{2\epsilon}z_2^2.
\end{aligned}$$

Using these bounds and gathering terms, we have that $\dot{\mathcal{V}}_{123}$ now reads:

$$\begin{aligned}
\dot{\mathcal{V}}_{123} \leq &-\Big(\kappa_{12}\mu-\frac{\kappa_3^2}{2\epsilon}(4c_1^2a_{20}^2+c_2^2a_{11}^2+a_{10}^2)\Big)z_1^2-\Big(\kappa_{12}\mu-\frac{\kappa_3^2}{2\epsilon}(4c_2^2a_{02}^2+c_1^2a_{11}^2+a_{01}^2)\Big)z_2^2,\\
&-(\kappa_3-5\epsilon)z_3^2+\frac{\kappa_3^2}{8\epsilon}\Big(4a_{20}^2+a_{11}^2\Big)z_1^4+\frac{\kappa_3^2}{8\epsilon}\Big(4a_{02}^2+a_{11}^2\Big)z_2^4,\\
\leq &-\Big(\kappa_{12}\mu-\frac{\kappa_3^2}{2\epsilon}(4c_1^2a_{20}^2+c_2^2a_{11}^2+a_{10}^2)\Big)z_1^2-\Big(k\mu-\frac{\kappa_3^2}{2\epsilon}(4c_2^2a_{02}^2+c_1^2a_{11}^2+a_{01}^2)\Big)z_2^2,\\
&-(\kappa_3-5\epsilon)z_3^2+\frac{\kappa_3^2}{8\epsilon}\max\Big(4a_{20}^2+a_{11}^2,4a_{02}^2+a_{11}^2\Big)\|z_{12}\|^4.
\end{aligned}$$

To conclude uniform global exponential stability, we take the candidate Lyapunov function

$$\mathcal{V} := \mathcal{V}_{123}+\left(\kappa'+\frac{\kappa_3^2}{32\kappa_{12}\mu\epsilon}\max\Big(4a_{20}^2+a_{11}^2,4a_{02}^2+a_{11}^2\Big)\right)\mathcal{V}_{12}^2.$$

with $\kappa' > 0$, that after straightforward calculations gives

$$\begin{aligned}
\dot{\mathcal{V}} \leq &-\Big(\kappa_{12}\mu-\frac{\kappa_3^2}{2\epsilon}(4c_1^2a_{20}^2+c_2^2a_{11}^2+a_{10}^2)\Big)z_1^2-\Big(\kappa_{12}\mu-\frac{\kappa_3^2}{2\epsilon}(4c_2^2a_{02}^2+c_1^2a_{11}^2+a_{01}^2)\Big)z_2^2\\
&-(\kappa_3-5\epsilon)z_3^2-\kappa'\|z_{12}\|^4.
\end{aligned}$$

The claim is proven with the gain selection

$$\begin{aligned}
\kappa_3 &> 5\epsilon,\\
\kappa_{12}\mu &> \frac{\kappa_3^2}{2\epsilon}\max\Big(4c_1^2a_{20}^2+c_2^2a_{11}^2+a_{10}^2,4c_2^2a_{02}^2+c_1^2a_{11}^2+a_{01}^2\Big),\\
\kappa' &> 0.
\end{aligned} \tag{3.19}$$

□

Remark 3.2.

- $\Sigma_0 = MV_{12}$ with $M = \begin{bmatrix} 2a_{20} & a_{11} \\ a_{11} & 2a_{02} \end{bmatrix}$ so Assumption 3.1 is an assumption on the horizontal speed. It is then clear that $\det(M) \neq 0$ is a necessary condition of Assumption 3.1 which matches the sufficient condition of observability for the polynomial map of degree 2. Moreover, the condition of persistence of excitation on the speed means that the drone must change directions consistently during a moving time window. It means that, necessarily, the acceleration vector U_{12} must not be colinear to V_{12} sufficiently frequently which matches the sufficient conditions of observability. One can conclude that Assumption 3.1 is very natural from a mathematical point of view.

- Assume that the map h_M represents any smooth ground profile and that one has a good guess on the actual horizontal position of the drone, denoted by $\bar{X}_{12} = (\bar{x}_1, \bar{x}_2)$. Then, by the Taylor formula:

$$h_M(x_1, x_2) = h_{quad}(x_1, x_2) + o(\|X_{12} - \bar{X}_{12}\|^2),$$

with

$$h_{quad}(x_1, x_2) = h_M(\bar{x}_1, \bar{x}_2) + \frac{\partial h_M}{\partial X_{12}}(X_{12} - \bar{X}_{12}) + \frac{1}{2}(X_{12} - \bar{X}_{12})^T \frac{\partial^2 h_M}{\partial {X_{12}}^2}(X_{12} - \bar{X}_{12}),$$

where $X_{12} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\frac{\partial^2 h_M}{\partial {X_{12}}^2}$ is the hessian matrix of h_M. With this representation one can consider the quadratic form h_{quad} to be the true ground profile and apply the previous observer design with $h^{nl} = x_3 - h_{quad}$. Proposition 3.2 shows that the resulting observer is uniformly locally exponentially stable around $\bar{X}_{12}$. This observer is interesting, for example, after some long time control planning where $\bar{X}_{12}$ is the guiding objective and when one already knows that X_{12} is near $\bar{X}_{12}$. In this case one only needs a local controller/observer and could use our quadratic approximation of the map.

- By definition of Σ_0, the PE condition can be rewritten as follows

$$\frac{1}{T}\int_t^{t+T} V_{12}(\tau)V_{12}^T(\tau)d\tau \succeq \mu M^{-2}.$$

It implies that if M is close to be singular, V_{12} must be very persistent. Actually if M is close to be singular, the map is close to be invariant under a translation in at least one direction. It follows that V_{12} must be very persistent precisely in that direction because of the term M^{-2}. For example, suppose that $M = -\begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix}$ with $\epsilon > 0$ close to zero and that V_{12} is chosen as follows:

$$V_{12} = \begin{bmatrix} \alpha_1 \cos(\frac{2\pi}{T}t) \\ \alpha_2 \sin(\frac{2\pi}{T}t) \end{bmatrix},$$

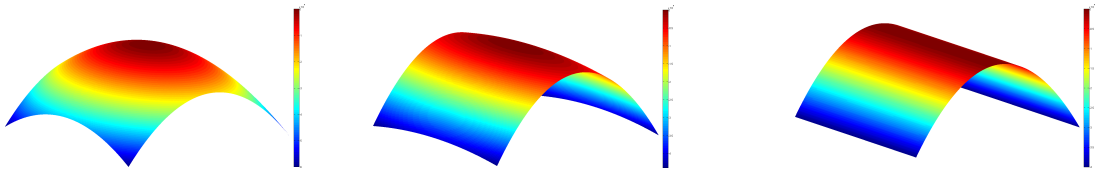
with $(\alpha_1, \alpha_2) \in \mathbb{R}^2$. Then one gets:

$$\frac{1}{T}\int_t^{t+T} V_{12}(\tau)V_{12}^T(\tau)d\tau = \begin{bmatrix} \alpha_1^2 & 0 \\ 0 & \alpha_2^2 \end{bmatrix}.$$

Thus, in this example, the PE condition is equivalent to:

$$|\alpha_1| \geq \mu^{\frac{1}{2}},$$
$$|\alpha_2| \geq \frac{\mu^{\frac{1}{2}}}{\epsilon}.$$

Figure 3.1 shows examples of h_{nl} with quadratic form M and decreasing values of ϵ from left to right. It is clear from Figure 3.1c that, as ϵ goes to 0, the map becomes invariant in the direction of x_2 and $|\alpha_2|$ must be larger and larger. It implies that V_{12} must be very persistent in that direction.



(a) $\epsilon = 0.5$ (b) $\epsilon = 0.1$ (c) $\epsilon = 0.001$

Figure 3.1: Three examples of quadrics with quadratic form M with different values of ϵ

Numerical simulations

We present here the numerical implementation of the observer described in Proposition 3.2. Figure 3.2 shows the trajectory of the 3D position and its estimator and Figure 3.3 their time evolution. The test ground map satisfies $a_{20} = -2$, $a_{11} = -1$, $a_{02} = 3$, $a_{10} = a_{01} = a_{00} = 0$ leading to a hyperbolic paraboloid map. The system (2.3) is run with an open loop sinusoidal control such that the 2D horizontal orbits are ellipses and the altitude stays constant. This implies in particular that the PE condition 3.1 is verified. κ_{12} and κ_3 are chosen to satisfy the conditions (3.19). Actually, the condition (3.19) is only needed to ensure the convergence of $\widehat{\eta}_3$. Figures 3.3c and 3.3d confirm the exponential convergence showed in Proposition 3.2.

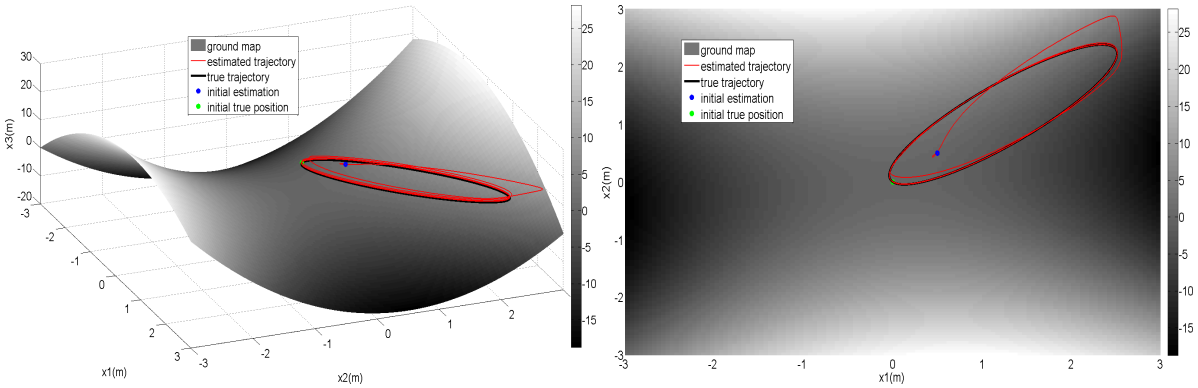
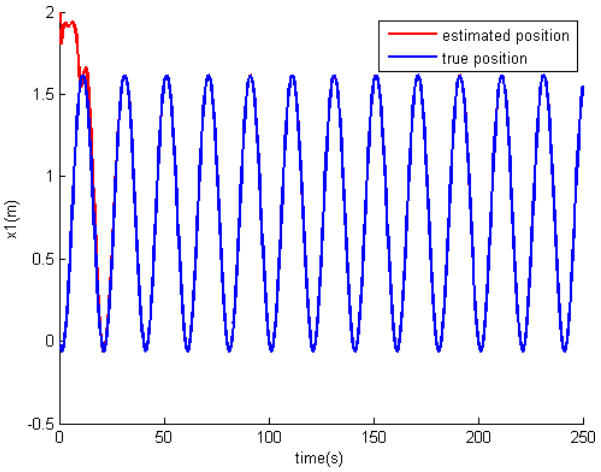
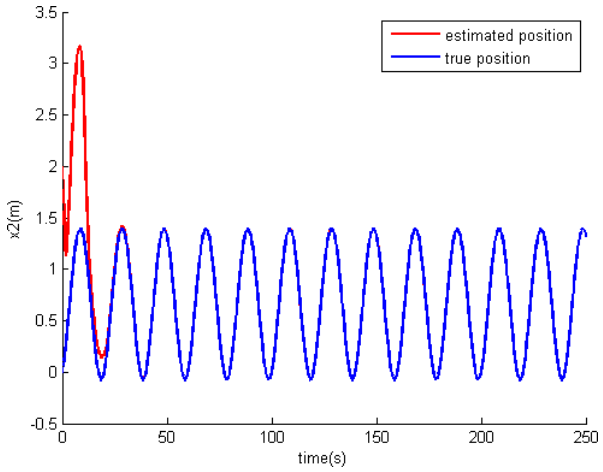


Figure 3.2: Plot of the test quadratic ground map, the trajectory of the system and of the observer

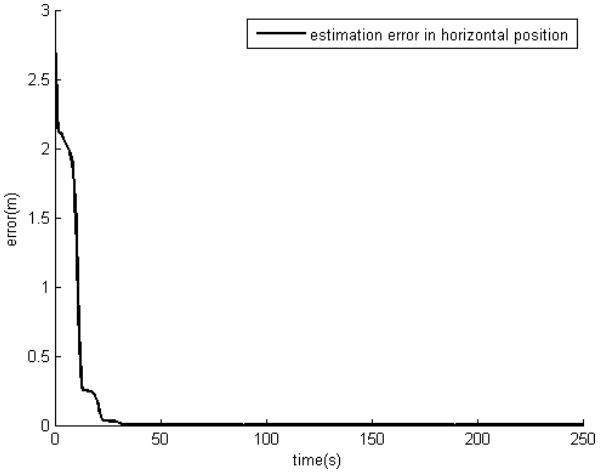
The success of the presented global observer design is entirely due to the fact that the dynamics of the horizontal estimation error can be easily decoupled from the vertical variables and easily made linear. The reason behind this is that the third order term of h_M vanishes. As it is not true for the other models



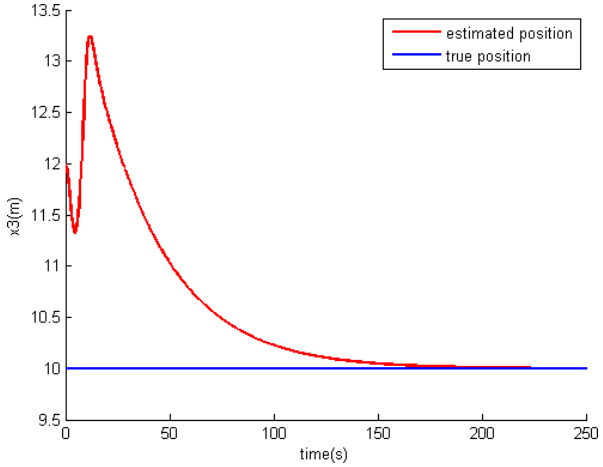
(a) Plot of $\eta_1(t)$ and $\widehat{\eta}_1(t)$



(b) Plot of $\eta_2(t)$ and $\widehat{\eta}_2(t)$



(c) Plot of $\|z_{12}(t)\|$



(d) Plot of $\eta_3(t)$ and $\widehat{\eta}_3(t)$

Figure 3.3: Time evolution of η, $\widehat{\eta}$ and z_{12} in the case of an ellipsoidal trajectory

of map we consider, this type of observer is hard to extend. For instance, for a polynomial map of degree 3, there would be additional quadratic terms in equation (3.13) that are hard to deal with directly. In the case of a Gaussian map, as $\dot{h}_M = X_{ce}^T R V_{12} h_M$, there still remains the nonlinear term h_M in the dynamics of the estimation error. In fact, the issue is that h_M is unknown with the single measurement $h^{nl} = x_3 - h_M$. However, with a supplementary measurement of altitude, i.e considering model (2.5), h_M is known and a similar observer can be designed. It is the topic of the next section.

Gaussian map with an altitude measurement

In this section, we consider the following observation equation:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^{l}(X,V) \end{bmatrix},$$

with $h^l(X,V) = \begin{bmatrix} x_3 \\ V \end{bmatrix}$, $h^{nl}(X,V) = h_M(x_1,x_2)$, and

$$h_M(x_1,x_2) = H^0 \text{exp}\left(-\frac{1}{2}\begin{bmatrix} x_1 - x_1^0 & x_2 - x_2^0 \end{bmatrix} R \begin{bmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \end{bmatrix}\right).$$

To lighten the computation, we consider normalised and centred horizontal state and control variables η, $\widetilde{V}_{12}$ and $\widetilde{U}_{12}$ defined by:

$$\begin{aligned} \eta &= R^{\frac{1}{2}} X_{ce}, \\ \widetilde{V}_{12} &= R^{\frac{1}{2}} V_{12}, \\ \widetilde{U}_{12} &= R^{\frac{1}{2}} U_{12}, \end{aligned} \tag{3.20}$$

where $R^{\frac{1}{2}}$ is the matrix square root of R. Thus, the new dynamics of the system can be expressed as follows:

$$\begin{aligned} \dot{\eta} &= \widetilde{V}_{12}, \\ \dot{x}_3 &= v_3, \\ \dot{h}_M &= -\eta^T \widetilde{V}_{12} h_M, \\ \dot{\widetilde{V}}_{12} &= \widetilde{U}_{12}, \\ \dot{v}_3 &= u_3. \end{aligned}$$

Therefore, one can define the estimation error like in the previous sections:

$$z := \xi + \phi(h_M, \widetilde{V}_{12}) - \eta := \widehat{\eta} - \eta, \tag{3.21}$$

where $\xi : \mathbb{R}^+ \longrightarrow \mathbb{R}^2$ and $\phi : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ are to be chosen. The dynamics of z can be deduced as follows:

$$\dot{z} = \dot{\xi} + \frac{\partial \psi}{\partial h_M}(-\eta^T \widetilde{V}_{12} h_M) + \frac{\partial \psi}{\partial \widetilde{V}_{12}} \widetilde{U}_{12} - \widetilde{V}_{12}.$$

Then, ξ is chosen such that:

$$\dot{\xi} = \alpha(\xi, y, U) := \frac{\partial \psi}{\partial h_M}(\hat{\eta}^T \widetilde{V}_{12} h_M) - \frac{\partial \psi}{\partial \widetilde{V}_{12}} \widetilde{U}_{12} + \widetilde{V}_{12}.$$

Finally the dynamics of z becomes:

$$\dot{z} = \frac{\partial \psi}{\partial h_M}(h_M \widetilde{V}_{12}^T) z.$$

As $h_m > 0$, one can choose $\frac{\partial \psi}{\partial h_M} = -\frac{\kappa_{12}}{h_M} \widetilde{V}_{12}$ with $\kappa > 0$ which leads to the following error equation:

$$\dot{z} = -\kappa_{12} \widetilde{V}_{12} \widetilde{V}_{12}^T z. \tag{3.22}$$

Moreover, one possible choice of ψ is:

$$\begin{aligned}\psi(h_M, \widetilde{V}_{12}) &= -\kappa_{12}\ln(h_M)\widetilde{V}_{12}, \\ &= \frac{\kappa_{12}}{2}\|\eta\|^2\widetilde{V}_{12}.\end{aligned} \tag{3.23}$$

As equation (3.13) and (3.22) are the same up to a linear change of coordinates, we make the following assumptions on $\widetilde{V}_{12}$:

Assumption 3.4. *There exist some constants $T, \mu > 0$, such that for all $t \geq 0$,*

$$\frac{1}{T}\int_t^{t+T} \widetilde{V}_{12}(\tau)\widetilde{V}_{12}^T(\tau)d\tau \succeq \mu I, \tag{3.24}$$

Assumption 3.5. *There exists $c_V > 0$, such that for all $t \geq 0$, $\|\widetilde{V}_{12}\| \leq c_V$.*

Consquently, one can prove exponential convergence of the estimation error:

Proposition 3.3. *Under Assumptions 3.4 and 3.5, the system (3.22) is UGES at $z = 0$.*

Proof. The result follows as in Proposition 3.2 by considering the candidate Lyapunov function:

$$\mathcal{V}_{12} := \kappa_{12} z^T \Psi(t) z + \frac{\kappa}{2}\|z\|^2,$$

with $\kappa > 0$. $\Psi(t)$ is the 2×2 symmetric matrix defined by

$$\Psi(t) := (1 + c_V^2)TI - \frac{1}{T}\int_t^{t+T}\int_t^s \widetilde{V}_{12}(\tau)\widetilde{V}_{12}^T(\tau)d\tau ds.$$

□

Remark 3.3.

- Note that for the two proposed observers, the estimation error does not depend explicitly on the control U. However, it depends on U through V. Indeed, in practice, U must be designed so that V satisfies the PE conditions. This the is subject of a later section.
- As in the quadratic case, the case of a Gaussian with an altitude measurement has a very particular structure. Actually, the use of the logarithm in (3.23) allows one to recover a quadratic structure and to obtain a linear error equation. However, it is no longer possible if the altitude is not measured $h^{nl} = x_3 - h_M$ or for a sum of two Gaussian $h^{nl} = h_M^1 + h_M^2$. For this reason, the representation of the map with Gaussians is hard to deal with the presented technique. For similar reasons, the representation of the map with trigonometric functions is hard to handle directly.

Numerical simulations

Figure 3.2 depicts the trajectory of a drone flying over a hill represented by a Gaussian and its estimation by the observer from Proposition 2.4. Figure 3.5 represents the time evolution of this trajectory and its estimation. As in the polynomial case, Figure 3.5c confirms the exponential convergence of the estimation error to zero. Unlike the polynomial case considered previously, as x_3 is measured here, there is no condition on the gain κ_{12} in this case. Proposition 2.4 ensures global convergence of the observer.

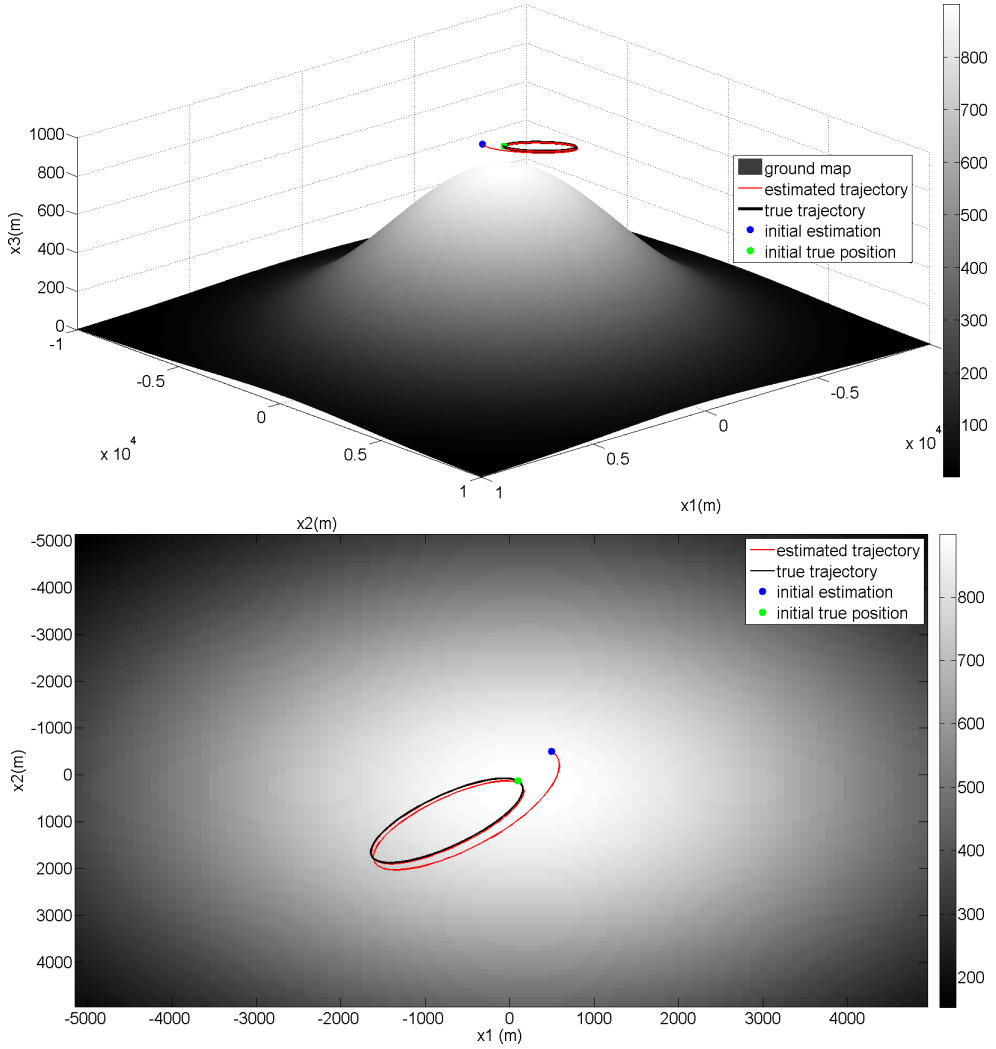


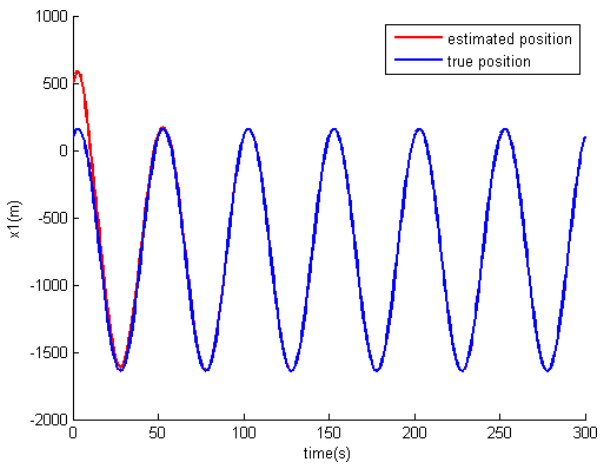
Figure 3.4: Plot of the test Gaussian map, the trajectory of the system and of the observer

However, if the system is situated far from the centre of the Gaussian then the term $\ln(h_M)$ in (3.23) is ill-defined numerically and the method fails. Intuitively, this system is poorly observable far from the centre of the Gaussian because of the increasing flatness of the terrain. In other words, the local observer described in Remark 3.2 could not work in this case as the hessian of h_M would be close to singular.

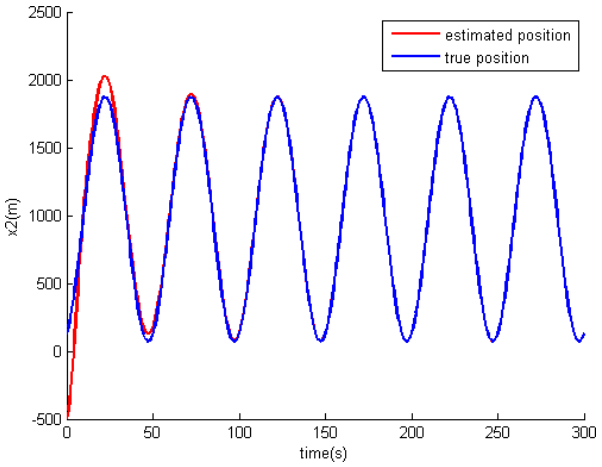
The method presented in this section seems not to be suited for the treatment of more complex maps. Therefore, in the following, we present and apply another class of methods based on an immersion of the original into a higher dimensional space where the immersed dynamics is state affine for which classical Kalman-like observers can be applied.

3.1.2.2 Method based on an immersion into an affine system

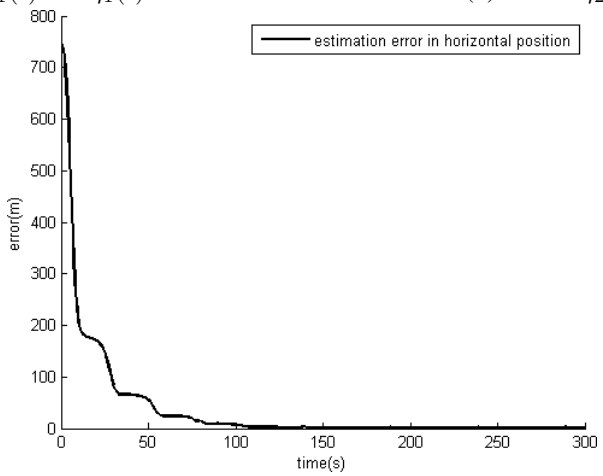
The goal of this section is to design observers using (3.1) with different choices of β and ϕ compared to the last section. The idea is to work with an augmented system and new variables $\chi := \phi(\eta)$ with $q > n$



(a) Plot of $\eta_1(t)$ and $\widehat{\eta}_1(t)$



(b) Plot of $\eta_2(t)$ and $\widehat{\eta}_2(t)$



(c) Plot of $\|z_{12}(t)\|$

Figure 3.5: Time evolution of η and $\widehat{\eta}$ the the case of an ellipsoidal trajectory

using the notations corresponding to (3.1) and (3.1). ϕ must be chosen such that the dynamics of χ is simpler than the one of η so that an observer $\widehat{\chi} = \xi$ can be easily found. An estimator of η is deduced by taking $\hat{\eta} = \phi^L(\widehat{\chi})$. Thus we restrict ourselves to the following choice of ϕ and β :

$$\begin{aligned} \chi &:= \phi(\eta, y, t) = \phi(\eta), \\ \widehat{\chi} &:= \beta(\xi, y, t) = \xi. \end{aligned}$$

with $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^q$ and $q > n$.

Kalman-like observer for state affine system

We recall the construction of a Kalman-like observer for state-affine systems.

Definition 3.2. *The system (2.1) is said to be* state-affine in the unmeasured state *if:*

$$\begin{aligned} f(x,u) &= A(u,Cx)x + B(u,Cx), \\ h(x) &= Cx, \end{aligned}$$

where $A : \mathbb{R}^m \times \mathbb{R}^p \longrightarrow \mathbb{R}^{n\times n}$, $B : \mathbb{R}^m \times \mathbb{R}^p \longrightarrow \mathbb{R}^n$ *and* $C \in \mathbb{R}^{p\times n}$ *leading to the following system:*

$$\begin{aligned} \dot{x} &= A(u,y)x + B(u,y), \\ y &= Cx. \end{aligned} \tag{3.25}$$

In [Besançon, 2007], it is showed that, under a condition of regular persistence of y and u, a Kalman-like observer can be designed.

Proposition 3.4. *Assume that* u *is chosen such that* $v(t) := (u(t), Cx(t, x_0, u(\cdot)))^T$ *is regularly persistent as in Definition 2.6 for the system:*

$$\begin{aligned} \dot{x} &= A(v(t))x, \\ y &= Cx, \end{aligned}$$

then, the system (3.26) admits an observer of the following form:

$$\dot{\hat{x}}(t) = A(u(t), y(t))x(t) + B(u(t), y(t)) - K(t)(C\hat{x}(t) - y(t)),$$

with $K(t)$ *given by:*

$$K(t) = P(t)CW^{-1},$$

and $P(t)$ *such that:*

$$\dot{P}(t) = P(t)A^T(u(t), y(t)) + A(u(t), y(t))P(t) - P(t)C^TW^{-1}CP(t) + V + \delta P(t),$$

where:

- $P(0) = P^T(0)$ *and* $W = W^T \succ 0$.
- $2\|A(u(t), y(t))\| < \delta$ *or* $V = V^T \succ 0$.

Polynomial map of degree 3

In this section, we consider dynamics (2.3) with the following observation equation:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^{l}(X,V) \end{bmatrix},$$

with

$$\begin{aligned} h^l(X,V) &= V, \\ h^{nl}(X,V) &= x_3 - h_M(x_1,x_2), \\ h_M(x_1,x_2) &= a_{30}x_1^3 + a_{21}x_1^2x_2 + a_{12}x_1x_2^2 + a_{03}x_2^3 \\ &\quad + a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2 + a_{10}x_1 + a_{01}x_2 + a_{00}. \end{aligned}$$

We set $\eta = X = (x_1, x_2, x_3)$. The idea is to consider the following overparametrization:

$$\chi = \phi(\eta) = \begin{bmatrix} \eta_3 - h_M(\eta_1, \eta_2) \\ \frac{\partial h_M}{\partial x_1} \\ \frac{\partial h_M}{\partial x_2} \\ \frac{\partial^2 h_M}{\partial x_1^2} \\ \frac{\partial^2 h_M}{\partial x_1 x_2} \\ \frac{\partial^2 h_M}{\partial x_2^2} \\ \eta_3 \end{bmatrix}. \tag{3.26}$$

The dynamics of χ and V can be derived by simple calculations and reads:

$$\begin{aligned} \dot{\chi} &= A_{cub}(V)\chi + B_{cub}V, \\ \dot{V} &= U. \end{aligned} \tag{3.27}$$

where:

$$A_{cub}(V) = \begin{bmatrix} 0 & -v_1 & -v_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & v_1 & v_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & v_1 & v_2 & 0 \\ & & & & & & \\ & & & \mathbf{0} & & & \\ & & & & & & \end{bmatrix},$$

and

$$B_{cub} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 6a_{30} & 2a_{21} & 0 \\ 2a_{21} & 2a_{12} & 0 \\ 2a_{12} & 6a_{03} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

By construction, one has $y = \begin{bmatrix} \chi_1 \\ V \end{bmatrix}$.

Clearly the system (3.27) is affine with respect to χ so $\hat{\chi}$ can be chosen as the Kalman observer from Proposition (3.4) and it will converge if U and y are regularly persistent as described in Proposition (3.4).

The next step is to make sure that ϕ left invertible. Actually, from the definition of h_M and ϕ:

$$\begin{bmatrix} \chi_4 \\ \chi_5 \\ \chi_6 \\ \chi_7 \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 h_M}{\partial x_1^2} \\ \frac{\partial^2 h_M}{\partial x_1 x_2} \\ \frac{\partial^2 h_M}{\partial x_2^2} \\ \eta_3 \end{bmatrix} = \begin{bmatrix} 6a_{30} & 2a_{21} & 0 \\ 2a_{21} & 2a_{12} & 0 \\ 2a_{12} & 6a_{03} & 0 \\ 0 & 0 & 1 \end{bmatrix} \eta + \begin{bmatrix} 2a_{20} \\ a_{11} \\ 2a_{02} \\ 0 \end{bmatrix}. \tag{3.28}$$

Thus, if rank $\left(\begin{bmatrix} 6a_{30} & 2a_{21} \\ 2a_{21} & 2a_{12} \\ 2a_{12} & 6a_{03} \end{bmatrix}\right) = 2$ then ϕ is left invertible and ϕ^L can be expressed as follows:

$$\phi^L(\chi) = E\chi - \eta_0,$$

where E is a matrix of appropriate dimensions depending on $\begin{bmatrix} 6a_{30} & 2a_{21} \\ 2a_{21} & 2a_{12} \\ 2a_{12} & 6a_{03} \end{bmatrix}$ and η_0 is vector of dimension 3 extracted from $\begin{bmatrix} 2a_{20} \\ a_{11} \\ 2a_{02} \\ 0 \end{bmatrix}$.

Finally, by setting $\hat{\eta} = E(\hat{\chi}) - \eta_0$, the estimation error can be written as follows:

$$\hat{\eta} - \eta = E(\hat{\chi} - \chi),$$

so if $\hat{\chi} - \chi$ converge to 0, then $\hat{\eta} - \eta$ converges to 0.

Remark 3.4.

- In this case, ϕ^L satisfies the Lipschitz condition mentioned in Remark 3.1.
- The previous design can easily be extended to the case of polynomial maps of degree higher than 3. The idea is to add more components to ϕ that represent higher order derivatives. The resulting dynamics would still be state-affine and a Kalman observer can be used. However, the dimension of χ grows rapidly with the degree of h_M which makes the practical implementation rapidly difficult. With this potential extension in mind, it seems unnecessary to use higher order polynomials as cubic splines are very efficient in practice. That is why, we only present the cubic case.
- Another possible extension is to use the previous observer for unstructured maps. Assume, as in Remark 3.2 that, h_M represents any realistic map. One can approximate h_M with a polynomial of degree 3, h_{cub}. The approximation error can be seen as an additive noise on the measurement i.e:

  $$h^{nl} = x_3 - h_M = x_3 - h_{cub} + v(t).$$

  As Kalman observer are known to be robust to measurement noise see, [Hespanha, 2009], the previous observer is expected to have to good performances if properly tuned. However, global polynomial approximations are rarely precise if the original map has non-stationary variations. Nevertheless, local approximations like cubic splines are much more precise. The final extension would then be to construct an observer of the system (3.27) where the coefficients $(a_{30}, a_{21}, a_{12}, a_{03})$ may switch from different values depending on the horizontal position. To apply our observer, one would have to identity the good coefficient by running several Kalman observers with all the possible values of the coefficients simultaneously and keep the most likely estimate as it is done in Interacting Multiple Model filtering for instance.

In the sequel we present a partial extension of this design to the case of a spatial trigonometric map.

Map as a sum of spatial trigonometric functions

We start by defining an overparametrisation for the case of a single cosine. In this section, we consider dynamics (2.3) with the following observation equation:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^{l}(X,V) \end{bmatrix},$$

with

$$\begin{aligned} h^{l}(X,V) &= V, \\ h^{nl}(X,V) &= x_3 - h_M(x_1,x_2), \\ h_M(x_1,x_2) &= \alpha \cos(\omega_1 x_1 + \phi_1)\cos(\omega_2 x_2 + \phi_2). \end{aligned}$$

We also set $\eta = X = (x_1, x_2, x_3)$ and, similarly to the cubic case, we set:

$$\chi = \phi(\eta) = \begin{bmatrix} h_M(\eta_1,\eta_2) \\ \frac{\partial h_M}{\partial x_1} \\ \frac{\partial h_M}{\partial x_2} \\ \frac{\partial^2 h_M}{\partial x_1 x_2} \\ \eta_3 \end{bmatrix}. \tag{3.29}$$

As $\frac{\partial^2 h_M}{\partial x_1^2} = -\omega_1^2 h_M$ and $\frac{\partial^2 h_M}{\partial x_2^2} = -\omega_2^2 h_M$, the dynamics of χ reads:

$$\begin{aligned} \dot{\chi} &= A_{trig}(V)\chi + B_{trig}V, \\ \dot{V} &= U, \end{aligned} \tag{3.30}$$

where

$$A_{trig}(V) = \begin{bmatrix} 0 & v_1 & v_2 & 0 & 0 \\ -\omega_1^2 v_1 & 0 & 0 & v_2 & 0 \\ -\omega_2^2 v_2 & 0 & 0 & v_1 & 0 \\ 0 & -\omega_2^2 v_2 & -\omega_1^2 v_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$B_{trig} = \begin{bmatrix} 0 \\ \\ 0 \quad 0 \quad 1 \end{bmatrix}.$$

The dynamics (3.30) is clearly affine w.r.t χ and $y = \begin{bmatrix} \chi_5 - \chi_1 \\ V \end{bmatrix}$ so a Kalman observer could be considered. However, the model of map (2.10) is periodic i.e. invariant under a particular translation on (η_1, η_2) and the dynamics (2.3) is invariant under any translation on (η_1, η_2). It means that it is

impossible to distinguish two different horizontal position (η_1, η_2) and $(\eta_1 + n_1\frac{2\pi}{\omega_1}, \eta_2 + n_2\frac{2\pi}{\omega_2})$ for any $(n_1, n_2) \in (\mathbb{N}^*)^2$ and whatever control is applied. Therefore, there is no hope of building a global observer. Nevertheless, with the model of sum of cosines (2.11), it seems more reasonable. As a matter of fact, a sum of cosines can be non-periodic depending on the different spatial pulsations. In the following we show that the same overparametrization can be done to each trigonometrical term of the map.

Consider now, the observation equation:

$$h(X,V) = \begin{bmatrix} h^{nl}(X,V) \\ h^{l}(X,V) \end{bmatrix},$$

with

$$\begin{aligned} h^{l}(X,V) &= V, \\ h^{nl}(X,V) &= x_3 - h_M(x_1,x_2), \\ h_M &= \sum_{i=1}^{n_{tri}} h_M^i, \\ h_M^i(x_1,x_2) &= \alpha^i \cos(\omega_1^i x_1 + \phi_1^i)\cos(\omega_2^i x_2 + \phi_2^i). \end{aligned}$$

We also set $\eta = X = (x_1, x_2, x_3)$ and, similarly to the single cosine case, we set:

$$\chi = \phi(\eta) = \begin{bmatrix} \chi^1 \\ \vdots \\ \chi^{n_{trig}} \end{bmatrix} = \begin{bmatrix} \vdots \\ h_M^i(\eta_1,\eta_2) \\ \frac{\partial h_M^i}{\partial x_1} \\ \frac{\partial h_M^i}{\partial x_2} \\ \frac{\partial^2 h_M^i}{\partial x_1 x_2} \\ \vdots \\ \eta_3 \end{bmatrix}. \tag{3.31}$$

Then the dynamics of χ reads:

$$\dot{\chi} = \begin{bmatrix} A_{trig}^1(V) & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & & & A_{trig}^{n_{trig}}(V) \end{bmatrix} \chi + \begin{bmatrix} B_{trig} \\ \vdots \\ B_{trig} \end{bmatrix} V,$$

where, for $i = 1..n_{trig}$:

$$A_{trig}^i(V) = \begin{bmatrix} 0 & v_1 & v_2 & 0 & 0 \\ -\omega_1^{i2} v_1 & 0 & 0 & v_2 & 0 \\ -\omega_2^{i2} v_2 & 0 & 0 & v_1 & 0 \\ 0 & -\omega_2^{i2} v_2 & -\omega_1^{i2} v_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this case, $y = \begin{bmatrix} \chi_{4n_{trig}+1} - \sum_{i=1}^{n_{trig}} \chi_1^i \\ V \end{bmatrix}$. As this kind of map is in general non-periodic with no symmetries, one can consider to use a Kalman observer with success as in the cubic case. However, there remains the issue of finding the left inverse of ϕ for this type of map. As trigonometric functions are involved in (3.29) and (3.31), a global inverse of ϕ does not exist. Finally, a solution could be to use two compatible inverse and change between the two dynamically.

In conclusion of the whole section, we have provided observer designs through I&I with a direct method in the case of a quadratic map and a Gaussian map with altitude measurement and with an over-parametrisation and Kalman observers in the case of a cubic and a trigonometric map. The former observers are proved to converge if the horizontal speed satisfies a well-known condition a persistence of excitation while the sufficient conditions of convergence for the latter are less explicit. That is why, in the following section, we focus on studying ways to ensure the classical persistence of excitation of the horizontal via outputfeedback control and stochastic excitation.

3.2 Persistent output-feedback control and stochastic excitation

Assumption 3.1 concerns a priori only the horizontal speed. However, in the output-feedback loop of a double integrator, one chooses the acceleration and not the speed. This means that the control law must ensure, at the same time, the convergence of the true system and the persistence of the speed. The design and study of such a controller is the first topic of the section. Besides in [Loría et al., 2002], a continuous-time stochastic process is used as a persistent signal and convergence is obtained numerically. It appears that almost sure PE in a continuous time framework has not been studied extensively. Thus, the second topic of the section is to show that a class of 2-dimensional rectangular signal are almost surely persistent.

3.2.1 Output-feedback control

In this section, we restrict ourselves to the 2D system

$$\begin{aligned} \dot{X}_{12} &= V_{12}, \\ \dot{V}_{12} &= U_{12}. \end{aligned}$$

We also consider that an estimator from Section 3.1.2.1 has already been designed leading to the following error equation in 2D:

$$\dot{z}_{12} = -\kappa_{12} V_{12} V_{12}^T z_{12}.$$

As we consider only 2D positions in this section, we simplify the notations by setting $\kappa = \kappa_{12}$, $z = z_{12}$ and $V = V_{12}$. To match the notations of the previous sections, we also set $\eta = X_{12}$. The equilibrium of the system is translated to 0.

This leads to the following reformulation:

$$\begin{aligned} \dot{\eta} &= V, \\ \dot{V} &= U, \\ \dot{z} &= -\kappa V V^T z. \end{aligned} \tag{3.32}$$

The idea of the section is to find a control law that ensure the convergence of the full system (3.32). From (3.32) and Proposition 3.2, one can guess that the control U should be designed to enforce the convergence of $\zeta = \begin{bmatrix} \eta \\ V \end{bmatrix}$ to 0 while ensuring some condition of persistence of excitation so that z also converges to 0. This property is known as δ-persistence and has been studied notably in [Loría et al., 1999] and [Loría et al., 2002]. Notice that the control will depend explicitly on the time, so we consider an arbitrary initial time $t_0 \geq 0$ and $t \geq t_0$. Thus, the control law, inspired from [Loría et al., 2002], is chosen as follows:

$$U = -K_x\hat{\eta} - K_vV + \alpha\left(\|\hat{\zeta}\|\right)\phi(t), \tag{3.33}$$

where:

- $\hat{\eta} = \eta + z$ is the estimator of η and $\hat{\zeta} = \zeta + \begin{bmatrix} z \\ 0 \end{bmatrix}$ represents the augmented state (observer/horizontal speed),
- $K_x >, K_v > 0$ are gains to tune,
- ϕ: $\mathbb{R}^+ \longrightarrow \mathbb{R}^2$ is of norm 1, is differentiable with bounded derivative, and satisfies the classical condition of persistence of excitation, i.e. $\exists T > 0$, $\exists \mu > 0$, $\exists C_{\dot{\phi}} > 0$ such that $\forall t \geq 0$:

$$\frac{1}{T}\int_t^{t+T} \phi(\tau)\phi^T(\tau)d\tau \succeq \mu, \tag{3.34}$$

$$\|\phi(t)\| = 1, \tag{3.35}$$

$$\|\dot{\phi}(t)\| \leq C_{\dot{\phi}}. \tag{3.36}$$

- α: $\mathbb{R}^+ \longrightarrow \mathbb{R}$ is C^1 and globally Lipschitz with constant $L > 0$, increasing and $\alpha(0) = 0$ such that, $\forall (r_1, r_2) \in \mathbb{R}^2$:

$$|\alpha(r_2) - \alpha(r_1)| \leq L|r_2 - r_1|, \tag{3.37}$$

$$\alpha'(r_1) \leq L. \tag{3.38}$$

From equation (3.32), the output feedback system can be written as follows:

$$\begin{aligned} \dot{\eta} &= V, \\ \dot{V} &= -K_x\eta - K_vV + \alpha\left(\|\zeta\|\right)\phi(t) + \left(\alpha\left(\left\|\zeta + \begin{bmatrix} z \\ 0 \end{bmatrix}\right\|\right) - \alpha\left(\|\zeta\|\right)\right)\phi(t) - K_xz, \\ \dot{z} &= -\kappa VV^Tz. \end{aligned} \tag{3.39}$$

Intuitively, one would like the speed to be very persistent if the estimation error is large and less persistent if it is small. This would allow the speed to be less and less persistent and to converge eventually. Actually, this would be the case if the term $\|\hat{\zeta}\|$ was replaced by $\|z\|$ in (3.33). In this case, one could have applied the result form [Loría et al., 2002]. However, z is unknown so it cannot be used directly in the controller. Therefore, the main issue of the controller (3.33)is that U depends on $\|\hat{\zeta}\|$ and not on $\|z\|$. It means

that if $\hat{\zeta}$ goes to zero too quickly then the speed is not persistent during enough time and the estimation error does not go to zero. The goal of the following is then to find sufficient conditions on the speed of convergence of $\|z\|$ that allows system (3.39) to converge asymptotically. This starts by proving the boundedness of the trajectories of (3.39)

3.2.1.1 Proof of boundedness

In the rest of section 3.2.1, we make the following assumption.

Assumption 3.6. *K_x and K_v are chosen so that there exist two 4×4 matrices $P \succ 0$ and $Q \succ 0$ such that the following Lyapunov equation is verified:*

$$PA + A^T P = -Q, \tag{3.40}$$

$$\lambda_{min}(Q) - 2L\|P\| > 0, \tag{3.41}$$

where $A = \begin{bmatrix} 0 & I_2 \\ -K_x I_2 & -K_v I_2 \end{bmatrix}$.

Assumption 3.6 means informally that the potentially destabilising effects of the δ-persistent input $\alpha\phi$ are compensated by the gains K_x and K_v. Besides, (3.41) is rather classical in input-to-state stability theory and has been studied in the Chapter 6 of [Khalil, 2002]. The idea of the following is then to prove the boundedness of ζ in system (3.39) by arguments of input-to-state stability. This result is summed up in Proposition 3.5.

Proposition 3.5. *Under Assumption 3.6, the solutions of equation (3.39) are bounded. In particular,* $\zeta = \begin{bmatrix} \eta \\ V \end{bmatrix}$ *is bounded.*

Proof. First, the boundedness of z is obtained by showing that $\mathcal{V}_z = \|z\|^2$ is a non strict Lyapunov function for the estimation error equation in (3.39):

$$\begin{aligned} \dot{\mathcal{V}}_e &= -\kappa z^T V V^T z, \\ &= -\kappa \|V^T z\|^2 \leq 0. \end{aligned}$$

Thus, $\forall t \geq t_0$,

$$\|z(t)\| \leq \|z(t_0)\| \equiv M_z. \tag{3.42}$$

Secondly, one consider $\mathcal{V}_\zeta = \zeta^T P \zeta$ as a candidate Lyapunov function for ζ:

$$\begin{aligned} \dot{\mathcal{V}}_\zeta = \zeta^T (PA + A^T P)\zeta + 2\zeta^T P \begin{bmatrix} 0 \\ \alpha\left(\|\zeta\|\right)\phi(t) \end{bmatrix} \\ + 2\zeta^T P \begin{bmatrix} 0 \\ \left(\alpha\left(\left\|\zeta + \begin{bmatrix} z \\ 0 \end{bmatrix}\right\|\right) - \alpha\left(\|\zeta\|\right)\right)\phi(t) \end{bmatrix} - 2K_x \zeta^T P \begin{bmatrix} 0 \\ z \end{bmatrix}. \end{aligned} \tag{3.43}$$

From the Lipschitz property of α and the reverse triangle inequality, one gets:

$$\left|\alpha\left(\left\|\zeta + \begin{bmatrix} z \\ 0 \end{bmatrix}\right\|\right) - \alpha\left(\|\zeta\|\right)\right| \leq L\|z\| \tag{3.44}$$

Therefore, combining (3.35), (3.42), (3.43), and (3.44):

$$\begin{aligned}\dot{\mathcal{V}}_\zeta &\leq -\zeta^T Q\zeta + 2L\|P\|\|\zeta\|^2 + 2L\|P\|\|z\|\|\zeta\| + 2K_x\|P\|\|z\|\|\zeta\|,\\ &\leq -(\lambda_{min}(Q) - 2L\|P\|)\|\zeta\|^2 + 2(L+K_x)\|P\|\|z\|\|\zeta\|,\\ &\leq -(\lambda_{min}(Q) - 2L\|P\|)\|\zeta\|^2 + 2M_z(L+K_x)\|P\|\|\zeta\|,\end{aligned}$$

where $\lambda_{min}(Q)$ is the lowest eigenvalue of Q and $\|P\| = \lambda_{max}(P)$ is the highest eigenvalue of P. Thus, for $\epsilon > 0$, if $2(L+K_x)\|P\|\|z\| \leq 2(L+K_x)\|P\|M_z \leq \epsilon\|\zeta\|$ then

$$\dot{\mathcal{V}}_\zeta \leq -(\lambda_{min}(Q) - 2L\|P\| - \epsilon)\|\zeta\|^2. \tag{3.45}$$

Under Assumption 3.6, as ϵ can be taken arbitrarily small, equation (3.43) shows that ζ is input-to-state stable with respect to z. In particular, ζ is bounded as z is bounded.

□

3.2.1.2 Convergence of $\hat{\eta}$ and V

In this section, we reformulate the system (3.39) into the following equivalent form:

$$\begin{aligned}\dot{\hat{\eta}} &= V - \kappa VV^T z,\\ \dot{V} &= -K_x\hat{\eta} - K_v V + \alpha\left(\|\hat{\zeta}\|\right)\phi(t),\\ \dot{z} &= -\kappa VV^T z.\end{aligned} \tag{3.46}$$

The formulation (3.46) happens to be easier to interpret and to deal with afterwards. Notably, the results of the last section will allow us to show that system (3.46) is uniformly globally stable (UGS) and that $\hat{\zeta} = \begin{bmatrix}\hat{\eta}\\ V\end{bmatrix}$ converges to 0. This is the topic of Proposition 3.6.

Proposition 3.6. *Under Assumption 3.6, the system (3.46) is UGS at 0 and* $\lim_{t\to+\infty}\|\hat{\zeta}(t)\| = 0$.

Proof. One considers the following non-strict candidate Lyapunov function:

$$\mathcal{V} = \hat{\zeta}^T P\hat{\zeta} + \kappa_1\|z\|^2,$$

where P is defined as in (3.40). One has:

$$\dot{\mathcal{V}} = -\hat{\zeta}Q\hat{\zeta} + 2\begin{bmatrix}-\kappa VV^T z\\ 0\end{bmatrix}P\hat{\zeta} + 2\begin{bmatrix}0\\ \alpha\left(\|\hat{\zeta}\|\right)\phi(t)\end{bmatrix}P\hat{\zeta} - \kappa_1\kappa|z^T V|^2, \tag{3.47}$$

Using (3.35) and (3.37):

$$\dot{\mathcal{V}} \leq -\lambda_{min}(Q)\|\hat{\zeta}\|^2 + 2L\|P\|\|\hat{\zeta}\|^2 + 2\kappa\|P\|\|V\||z^T V|\|\hat{\zeta}\| - \kappa_1\kappa|z^T V|^2.$$

Using Young's inequality for $\epsilon > 0$:

$$\begin{aligned}\dot{\mathcal{V}} &\leq -(\lambda_{min}(Q) + 2L\|P\| - \epsilon)\|\hat{\zeta}\|^2 + \frac{4\kappa^2\|P\|^2\|V\|^2}{\epsilon}|z^T V|^2 - \kappa_1\kappa|z^T V|^2,\\ &\leq -(\lambda_{min}(Q) + 2L\|P\| - \epsilon)\|\hat{\zeta}\|^2 - \left(\kappa_1\kappa - \frac{4\kappa^2\|P\|^2\|V\|_\infty^2}{\epsilon}\right)|z^T V|^2,\end{aligned}$$

where $\|V\|_\infty$ is the L_∞ norm of V in time. Note that, under Assumption 3.6, V is bounded according to Proposition (3.5). In particular, $\|V\|_\infty < +\infty$.Therefore, for κ_1 sufficiently large, one has $\kappa_1\kappa - \frac{4\kappa^2\|P\|^2\|V\|_\infty^2}{\epsilon} > 0$ and then:

$$\dot{\mathcal{V}} \leq -(\lambda_{min}(Q) + 2L\|P\| - \epsilon)\|\hat{\zeta}\|^2.$$

Similarly to the previous section, under Assumption 3.6 for sufficiently small ϵ:

$$\dot{\mathcal{V}} \leq -\lambda\|\hat{\zeta}\|^2 \leq 0.$$

with $\lambda = \lambda_{min}(Q) + 2L\|P\| - \epsilon > 0$. As $\mathcal{V}$ is positive definite and decrescent, the system (3.46) is UGS. In particular, we obtain again that $\hat{\zeta}$ and z are bounded. For the convergence of $\hat{\zeta}$, one would like to use Barbalat lemma (lemma 2.12 [Narendra and Annaswamy, 2012]). To do so, first notice that, $\forall t \geq 0$:

$$\int_0^t \|\hat{\zeta}\|^2 \leq \frac{\mathcal{V}(0)}{\lambda}.$$

Thus,

$$\int_0^{+\infty} \|\hat{\zeta}\|^2 \leq \frac{\mathcal{V}(0)}{\lambda} < +\infty.$$

Secondly, as $\hat{\zeta}$, z and ϕ are bounded, one can see from system (3.46) that $\frac{d\|\hat{\zeta}\|^2}{dt}$ is bounded. Consequently by Barbalat's lemma:

$$\lim_{t\to+\infty} \|\hat{\zeta}(t)\| = 0$$

□

The next objective is to find conditions under which the estimation error z also converges to 0. The first step is to characterise the δ-persistence of V w.r.t. $\hat{\zeta}$.

3.2.1.3 δ-persistence of V

Let $r > 0$. Set $\widehat{\zeta}_{tot} = (\hat{\zeta}, z)$. In the following, it is assumed that $\|\widehat{\zeta}_{tot}(t_0)\| \leq r$. First, we define the notion of δ-persistence of a signal ψ w.r.t. $\hat{\zeta}$.

Definition 3.3. *A function of time ψ is said to be* δ-persistence w.r.t. $\hat{\zeta}$ *if* $\forall\delta > 0$, $\exists T_\delta > 0$, $\exists\mu_\delta > 0$, $\forall t \geq t_0$.

$$\underset{\tau\in[t,t+T_\delta]}{min}\|\hat{\zeta}(\tau)\| \geq \delta \Rightarrow \frac{1}{T_\delta}\int_t^{t+T_\delta} \psi(\tau)\psi^T(\tau)d\tau \succeq \mu_\delta I_2,$$

With a slight abuse of notations, we define $\psi : \mathbb{R}\times\mathbb{R}^4 \longrightarrow \mathbb{R}^2$ by, $\forall(t,\zeta)$:

$$\psi(t,\zeta) = \alpha(\|\zeta\|)\phi(t) \tag{3.48}$$

Then, $t \to \psi(t, \hat{\zeta}(t))$ verifies the property of δ-persistence w.r.t $\hat{\zeta}$ described in Proposition 3.7.

Proposition 3.7. *The function $t \to \psi(t, \hat{\zeta}(t))$ defined in equation (3.48) is δ-persistent w.r.t. $\hat{\zeta}$.*

Proof. Fix $\delta > 0$. Define $T_\delta = T$ with T from (3.34). Then, for $t \geq t_0$, if $\forall \tau \in [t, t+T_\delta]$, $\|\hat{\zeta}(\tau)\| \geq \delta$ then:

$$\frac{1}{T_\delta}\int_t^{t+T_\delta} \psi(\tau,\hat{\zeta}(\tau))\psi^T(\tau,\hat{\zeta}(\tau))d\tau = \frac{1}{T}\int_t^{t+T} \alpha(\|\hat{\zeta}(\tau)\|)^2\phi(\tau)\phi^T(\tau)d\tau$$

As α is increasing:

$$\frac{1}{T_\delta}\int_t^{t+T_\delta} \psi(\tau,\hat{\zeta}(\tau))\psi^T(\tau,\hat{\zeta}(\tau))d\tau \succeq \frac{1}{T}\int_t^{t+T} \alpha(\delta)^2\phi(\tau)\phi^T(\tau)d\tau,$$

From (3.34):

$$\frac{1}{T_\delta}\int_t^{t+T_\delta} \psi(\tau,\hat{\zeta}(\tau))\psi^T(\tau,\hat{\zeta}(\tau))d\tau \succeq \mu_\delta I_2,$$

with $\mu_\delta = \alpha(\delta)^2\mu > 0$ □

Proposition summarises the main result of this section

Proposition 3.8. *Under Assumption 3.7, the function $t \to V(t)$ solution of system (3.46) is δ-persistent w.r.t. $\hat{\zeta}$.*

Proof. To show that V is δ-persistent w.r.t. $\hat{\zeta}$, we use the technique of proof from [Panteley et al., 2001]. Let us fix $\xi \in \mathbb{R}^2$ such that $\|\xi\| = 1$, our goal is then to compare the quantities defined, $\forall t \geq t_0$, by:

$$\begin{aligned} a(t) \equiv& \xi^T\psi(t,\hat{\zeta}(t)),\\ a_f(t) \equiv& \xi^T V(t). \end{aligned}$$

Consider $w = aa_f$ then:

$$\begin{aligned} \dot{w} &= \dot{a}a_f + \dot{a_f}a,\\ &= a_f\xi^T\dot{\psi} + a\xi^T(-K_x\hat{\eta} - K_vV + \psi),\\ &= a_f\xi^T\left(\alpha\left(\|\hat{\zeta}\|\right)\dot{\phi} + \frac{\alpha'(\|\hat{\zeta}\|)}{\|\hat{\zeta}\|}\hat{\zeta}^T\dot{\hat{\zeta}}\phi\right) + a\xi^T(-K_x\hat{\eta} - K_vV + \psi),\\ &= a_f\xi^T\left(\alpha\left(\|\hat{\zeta}\|\right)\dot{\phi} + \frac{\alpha'(\|\hat{\zeta}\|)}{\|\hat{\zeta}\|}(\hat{\eta}^TV - \kappa(z^TV)V^T\hat{\eta} + V^T(-K_x\hat{\eta} - K_vV + \psi))\phi\right) \\ &+ a\xi^T(-K_x\hat{\eta} - K_vV + \psi). \end{aligned} \tag{3.49}$$

By applying Cauchy-Swartz inequality and using (3.35), (3.36), (3.37) and (3.38)

$$\begin{aligned} \dot{w} \geq& -|a_f|\Big(C_{\dot{\phi}}L\|\hat{\zeta}\| + L(\|\hat{\zeta}\| + \kappa\|\hat{\zeta}_{tot}\|\|\hat{\zeta}\|^2 + (K_x + K_v + L)\|\hat{\zeta}\|)\Big) \\ &- K_vL\|\hat{\zeta}\||a_f| - |a|K_x\|\hat{\zeta}\| + |a|^2. \end{aligned} \tag{3.50}$$

As the system (3.46) is UGS, is is known that there exists a C^0, increasing function γ such that, $\gamma(0) = 0$ and $\forall t \geq t_0$:

$$\|\widehat{\zeta}_{tot}(t)\| \leq \gamma(\|\widehat{\zeta}_{tot}(0)\|),$$

then, as $\|\widehat{\zeta}_{tot}(0)\| \leq r$ and γ is increasing, $\forall t \geq t_0$:

$$\|\hat{\zeta}(t)\| \leq \|\widehat{\zeta}_{tot}(t)\| \leq \gamma(r). \tag{3.51}$$

Using (3.51), one gets:

$$\begin{aligned} \dot{w} &\geq -|a_f|L((C_{\dot{\psi}} + K_x + 2K_v + L + 1)\gamma(r) + \kappa(\gamma(r))^3) - LK_x(\gamma(r))^2 + |a|^2, \\ &\geq -f(r)|a_f| - g(r) + |a|^2, \end{aligned} \tag{3.52}$$

with $f(r) = L((C_{\dot{\psi}} + K_x + 2K_v + L + 1)\gamma(r) + \kappa(\gamma(r))^3)$ and $g(r) = LK_x(\gamma(r))^2$.

Consider $T_f = (k+1)T$ a persistence horizon for V where $k \in \mathbb{N}$ is to be defined. Then, from (3.52) and by using Cauchy-Swartz inequality, $\forall t \geq 0$:

$$\begin{aligned} w(t+T_f) - w(t) &\geq \int_t^{t+T_f} |a(\tau)|^2 d\tau - f(r)T_f^{\frac{1}{2}} \left(\int_t^{t+T_f} |a_f(\tau)|^2 d\tau \right)^{\frac{1}{2}} - T_f g(r), \\ &\geq \xi^T \left(\int_t^{t+T_f} \psi(\tau, \hat{\zeta}(\tau)) \psi^T(\tau, \hat{\zeta}(\tau)) d\tau \right) \xi \\ &\quad - f(r)T_f^{\frac{1}{2}} \left(\int_t^{t+T_f} |a_f(\tau)|^2 d\tau \right)^{\frac{1}{2}} - T_f g(r). \end{aligned}$$

Fix $\delta > 0$ and $t \geq 0$. Suppose that $\forall \tau \in [t, t+T_\delta]$, $\|\hat{\zeta}(\tau)\| \geq \delta$ then from Proposition 3.7:

$$w(t+T_f) - w(t) \geq (k+1)T\mu_\delta - f(r)T_f^{\frac{1}{2}} \left(\int_t^{t+T_f} |a_f(\tau)|^2 d\tau \right)^{\frac{1}{2}} - (k+1)Tg(r),$$

$$f(r)T_f^{\frac{1}{2}} \left(\int_t^{t+T_f} |a_f(\tau)|^2 d\tau \right)^{\frac{1}{2}} \geq (k+1)T(\mu_\delta - g(r)) + w(t) - w(t+T_f).$$

From (3.37) and (3.51):

$$|w| \geq -L(\gamma(r))^2.$$

Thus,

$$f(r)T_f^{\frac{1}{2}} \left(\int_t^{t+T_f} |a_f(\tau)|^2 d\tau \right)^{\frac{1}{2}} \geq (k+1)T(\mu_\delta - g(r)) - 2L(\gamma(r))^2.$$

To show that V is δ-persistent, one needs the following assumption:

Assumption 3.7. $\mu_\delta > g(r) = LK_x(\gamma(r))^2$.

Assumption 3.7 is quite difficult to check in practice because γ is not explicit but it can be reformulated as follows:

$$K_x < \frac{\mu\alpha^2(\delta)}{L(\gamma(r))^2},$$

where $0 < \mu < 1$ and $0 < \delta < r$. Thus, under Assumption 3.7, one can do as in [Panteley et al., 2001] and choose k such that:

$$(k+1)T(\mu_\delta - g(r)) - 2L(\gamma(r))^2 \geq T(\mu_\delta - g(r))$$

i.e.

$$k = \left\lfloor \frac{2L(\gamma(r))^2}{T(\mu_\delta - g(r))} \right\rfloor + 1.$$

Finally, $\forall \xi \in \mathbb{R}^2$ such that $\|\xi\| = 1$

$$\begin{aligned} \frac{1}{T_f}\int_t^{t+T_f} |a_f(\tau)|^2 d\tau &\geq \mu_f, \\ \frac{1}{T_f}\xi^T\left(\int_t^{t+T_f} V(\tau)V^T(\tau)d\tau\right)\xi &\geq \mu_f, \\ \frac{1}{T_f}\left(\int_t^{t+T_f} V(\tau)V^T(\tau)d\tau\right) &\succeq \mu_f I_2, \end{aligned} \tag{3.53}$$

with $\mu_f(r,\delta) = \frac{T^2(\mu_\delta - g(r))^2}{(f(r))^2 T_f^2}$.

Therefore, under Assumption 3.7, V is δ-persistent w.r.t. $\hat{\zeta}$. □

This means, as expected, that the level of persistence of V decreases as $\hat{\zeta}$ tends to 0. Moreover, Assumption 3.7 means informally that K_x must be sufficient small for fixed μ, δ and r. This is consistent with our remark at the beginning of the section because K_x small implies that $\hat{\eta}$ converges slowly.

3.2.1.4 Convergence of z

The goal of the section is to obtain a result of local practical stability for the system (3.46) described in Proposition 3.9.

Proposition 3.9. *Assume that for some $r > 0$ and $\epsilon_1 > 0$, there exist $\delta(\epsilon_1) > 0$ and a controller U from (3.33) satisfying Assumptions 3.6, 3.7 and 3.8 at $\delta(\epsilon_1)$ and r. Therefore, if $\|\widehat{\zeta}_{tot}(t_0)\| \leq r$, then there exists $t_1 \geq t_0$ such that $\forall t \geq t_1$:*

$$\|\widehat{\zeta}_{tot}(t)\| \leq \epsilon_1.$$

Proof. As the convergence of $\hat{\eta}$ has been shown, there remains only to show the convergence of z. In the following, we summarise the results established previous to precisely point out what kind of result must be shown on z. Fix $\epsilon_1 > 0$. According to Section 3.2.1.2, the system (3.46) is UGS so $\exists \delta' > 0$ such that, $\forall t_1 \geq 0$:

$$\|\widehat{\zeta}_{tot}(t_1)\| \leq \delta' \Rightarrow \forall t \geq t_1,\ \|\widehat{\zeta}_{tot}(t)\| \leq \epsilon_1. \tag{3.54}$$

Because of the equivalence of norms, $\exists \delta > 0$ such that:

$$\max(\|\hat{\zeta}\|, \|z\|) \leq \delta \Rightarrow \|\widehat{\zeta}_{tot}\| \leq \delta' \tag{3.55}$$

We recall that we consider $r > 0$ such that $\|\widehat{\zeta}_{tot}(t_0)\| \leq r$. (3.46). We assume in the sequel that Assumptions 3.6 holds and that Assumption 3.7 holds at δ and r. From the proof of Proposition 3.5, one knows that, $\forall t_2 \geq t_1$:

$$\|z(t_2)\| \leq \|z(t_1)\|, \tag{3.56}$$

From Proposition 3.6, one knows that $\|\hat{\zeta}\|$ converges to 0, so $\exists t_\zeta \geq t_0$ such that:

$$\|\hat{\zeta}(t_\zeta)\| \leq \delta \tag{3.57}$$

Combining the properties (3.54) to (3.57) one can see that it is now sufficient to show that, under suitable additional assumptions, $\exists t_z \geq t_0$ such that:

$$\|z(t_z)\| \leq \delta \tag{3.58}$$

Equation (3.58) means that one only has to show that z reaches a neighbourhood of 0 of size δ. To do so, one defines the reaching times $T_{\hat{\zeta}}$ and T_z by:

$$T_{\hat{\zeta}} = \inf\{t \geq t_0,\ \|\hat{\zeta}(t)\| \leq \delta\}, \tag{3.59}$$

$$T_z = \inf\{t \geq t_0,\ \|z(t)\| \leq \delta\}, \tag{3.60}$$

with the convention that $\inf \varnothing = +\infty$. From, (3.57), one gets that $T_{\hat{\zeta}} < +\infty$ and one wants to show that $T_z < +\infty$

From Section 3.2.1.3, one can see that if, for $t_0 \leq t_1 \leq t_2$, $\min\limits_{\tau \in [t_1, t_2+T_f]} \|\hat{\zeta}(\tau)\| \geq \delta$, then $\forall t \in [t_1, t_2]$:

$$\frac{1}{T_f}\left(\int_t^{t+T_f} V(\tau)V^T(\tau)d\tau\right) \succeq \mu_f I_2, \tag{3.61}$$

From (3.16) and (3.18), note that while (3.61) holds true there exists a strict time dependent quadratic Lyapunov function $\mathcal{V}_z$ for the subsystem of z such that:

$$(T_f + \kappa_1)I \preceq \mathcal{V}_z \preceq ((1 + (\gamma(r))^2)T_f + \kappa_1)I.$$
$$\dot{\mathcal{V}}_z \leq -(\kappa\mu_f - \epsilon_2)\|z\|^2,$$

with $\kappa_1 > 0$ and $0 < \epsilon_2 < \kappa\mu_f$.

Then, for $t_0 \leq t_1 \leq t_2$:

$$\min_{\tau \in [t_1, t_2+T_f]} \|\hat{\zeta}(\tau)\| \geq \delta \Rightarrow\ \forall t \in [t_1, t_2],\ \|z(t)\| \leq z_{max}(t_1, t) \tag{3.62}$$

with $z_{max}(t_1, t) = \left(\frac{(1+(\gamma(r))^2)T_f+\kappa_1}{T_f+\kappa_1}\right)^{\frac{1}{2}} \|z(t_1)\| \exp\left(-\frac{\kappa\mu_f-\epsilon_2}{2T_f}(t - t_1)\right)$. Equation (3.62) simply means that, if $\|\hat{\zeta}\|$ is sufficiently large on the interval $[t_1, t_2+T_f]$, then $\|z(t)\|$ decreasing exponentially according to the bound z_{max}. Let us define $T_{z_{max}}$ by:

$$T_{z_{max}} = \inf\{t \geq t_0,\ z_{max}(t_0, t) \leq \delta\}.$$

Clearly, $\lim\limits_{t \to +\infty} z_{max}(t_0, t) = 0$ so $T_{z_{max}<+\infty}$. Consequently, we make the following assumption:

Assumption 3.8. $T_{z_{max}} + T_f \leq T_{\hat{\zeta}}$

Intuitively, Assumption 3.8 means that $\hat{\zeta}$ has to reach the neighbourhood of size δ long enough after z_{max} to keep V persistent and allow z to converge. Thus, under Assumption 3.8, by replacing t_1 by t_0 and t_2 by $t_0 + T_{\hat{\zeta}} - T_f \geq t_0$ in (3.62), one gets:

$$\min_{\tau\in[t_0,T_{\hat{\zeta}}]} \|\hat{\zeta}(\tau)\| \geq \delta \Rightarrow \forall t \in [t_0, t_0 + T_{\hat{\zeta}} - T_f],\ \|z(t)\| \leq z_{max}(t_0, t),$$

$$\min_{\tau\in[t_0,T_{\hat{\zeta}}]} \|\hat{\zeta}(\tau)\| \geq \delta \Rightarrow T_z \leq T_{z_{max}} < +\infty.$$

By definition $\min_{\tau\in[t_0,T_{\hat{\zeta}}]} \|\hat{\zeta}(\tau)\| \geq \delta$ holds true. So $T_z < +\infty$ and one gets the result. □

This is a result of practical stability because the controller (3.33) depends on δ which itself depends on ϵ_1. Besides, it is only local because it not clear if Assumption 3.7 is true for any $r > 0$. One would need a more precise result on the δ-persistence of V. However, it is a substantial result as system (3.39) has not been studied formally in the literature.

3.2.1.5 Numerical simulations

Figure 3.6 depicts an example of trajectory of η and $\widehat{\eta}$ from the system (3.39) with the control law (3.33). Figure 3.7 represents the time evolution of η, $\widehat{\eta}$ and z. A target point in the state space has been selected and the parameters in (3.33) have been chosen to ensure Assumption 3.8 in practice. It is confirmed by Figure 3.7, as one can see that the estimation error converges to 0 much earlier than the estimator and the true state converge to the target point. ϕ is a 2-dimensional vector of out-of-phase sinusoidal and $\alpha(r) = \epsilon_{per}\tanh(r)$ with $\epsilon_{per} > 0$ and tanh is the hyperbolic tangent. The choice of α is inspired by [Loría et al., 2002]. The boundedness of tanh prevents the system from being subjected to unreasonably high oscillations when the system starts far from the target. We have chosen the case of a Gaussian map to demonstrate the performances of the controller but the same can be undertaken with a quadratic map as the error equations are the same up to a linear change of coordinate.

3.2.2 An example of a piecewise constant almost surely persistent processes

The previous analysis of the system (3.32) has been undertaken in a deterministic framework. However, it appears that, in practice, if the persistent term ϕ from (3.33) is stochastic, then the system (3.32) also converges and in particular the estimation error tends to zero. In the sequel, let $(\Omega, \mathcal{F}, P)$ be a probability space. As we study stochastic processes, Ω can be seen as the classical set of càdlàg functions from $\mathbb{R}$ to $\mathbb{R}^d$ with $d \geq 1$ and P the canonical associated probability distribution. In this section, we study the particular 2D system:

$$\dot{z} = -\kappa V V^T z \tag{3.63}$$

where $V = (V(t,\omega))_{(t,\omega)\in[t_0,+\infty(\times\Omega} = (v_1(t,\omega), v_2(t,\omega))_{(t,\omega)\in[t_0,+\infty(\times\Omega}$ is a $\mathbb{R}^2$-valued stochastic process. We suppose that V belongs to the class of piecewise constant processes with time intervals of constant size. We also suppose that v_1 and v_2 can only take two opposite values.

Formally, we assume that $\exists\Delta > 0$, $\exists a > 0$ and there exists a discrete time stochastic process $(W^i(\omega))_{i\geq 0,\omega\in\Omega} = (w_1^i(\omega), w_2^i(\omega))_{i\geq 0,\omega\in\Omega}$, such that for almost all ω, $\forall i \in \mathbb{N}$, $\forall t \in [t_0 + i\Delta, t_0 +$

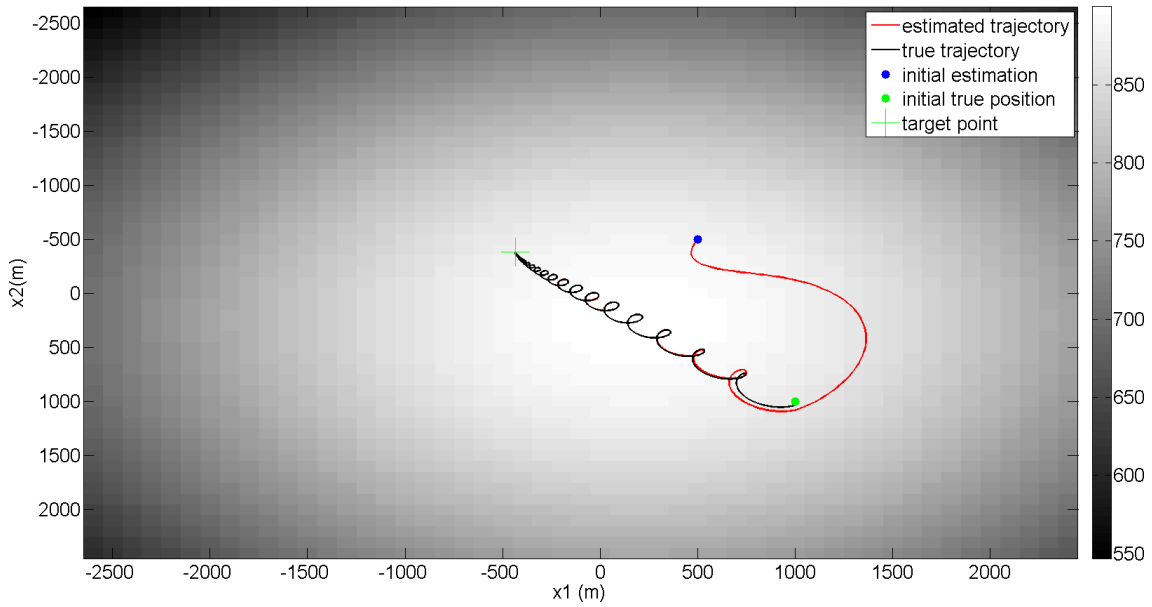


Figure 3.6: Plot of the test Gaussian map, the trajectory of the outputfeedback system and the observer

$(i+1)\Delta)$:

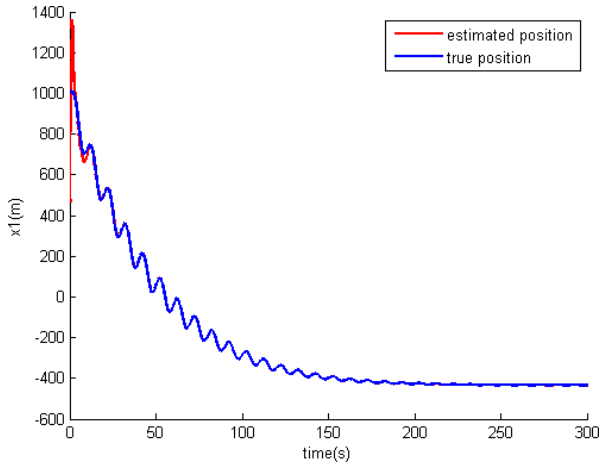
$$\begin{aligned} v_1(t,\omega) &= w_1^i(\omega) \in \{-a, a\}, \\ v_2(t,\omega) &= w_2^i(\omega) \in \{-a, a\}. \end{aligned} \tag{3.64}$$

These hypotheses make system (3.63) a switched system. PE conditions in switched systems has been treated recently in [Lee and Jiang, 2008] and [Lee et al., 2014] but only in a deterministic framework. Stochastic PE in a discrete-time framework has also been studied and review in [Bitmead, 1984]. However, sufficient conditions of almost sure PE in continuous-time as described in Assumption 3.1 seem not to have been studied since. One may argue that since we assume that V is piecewise constant then discrete-time results apply directly. However, the subtlety lies in the fact that Assumption 3.1 must hold for any $t \geq t_0$ which cannot be treated by discrete-time results directly. Moreover, we focus on the classical PE condition because it implies exponential stability. There are more recent and weaker conditions from [Barabanov and Ortega, 2017] and [Efimov et al., 2018]. They only require the receding horizon integral to be bounded from below at a countable number of time points but they only imply non-uniform asymptotic stability.

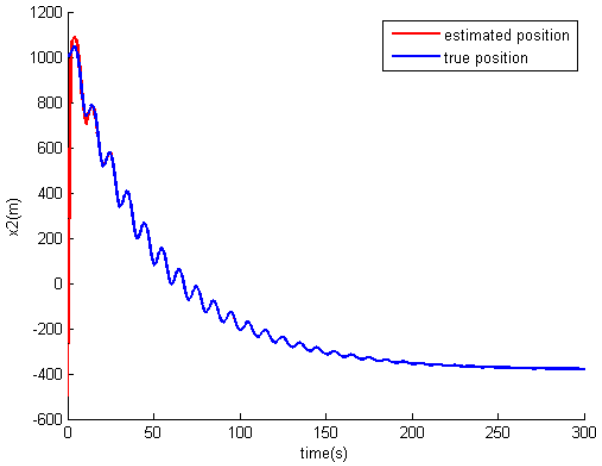
Therefore, the goal of the following is to find conditions on V that ensures for almost all ω, $\exists T(\omega) > 0$, $\exists \mu(\omega) > 0$ such that $\forall t \geq t_0$:

$$\frac{1}{T(\omega)} \left(\int_t^{t+T(\omega)} V(s,\omega) V^T(s,\omega) ds \right) \succeq \mu(\omega) I_2. \tag{3.65}$$

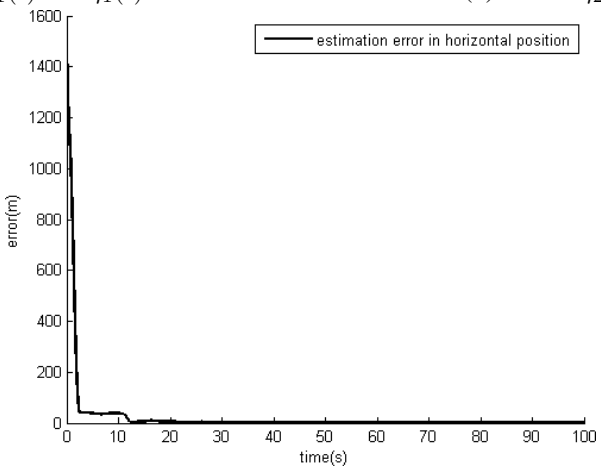
If (3.65) is satisfied, then, for almost all ω, Proposition 3.2 can be applied and the system (3.63) converges to 0 almost surely.



(a) Plot of $\eta_1(t)$ and $\widehat{\eta}_1(t)$



(b) Plot of $\eta_2(t)$ and $\widehat{\eta}_2(t)$



(c) Plot of $\|z_{12}(t)\|$

Figure 3.7: Time evolution of η and $\widehat{\eta}$ the the case of an ellipsoidal trajectory

3.2.2.1 Persistence of a 2D piecewice constant signal

The first step is to fix ω and to reformulate the persistence of excitation condition (3.65) on the function of time $V(\cdot, \omega)$ into a condition involving only the discrete time process W. To do so, we rewrite V in the polar coordinates. Actually, in two dimensions, V is persistent if it switches between two independent directions. As $\|V\|$ is constant from (3.64), it means that the corresponding angle must consistently switch between two values. We omit ω as an argument when there is no ambiguity. It is clear from (3.64) that, $\forall t \geq t_0$:

$$V(t) = r(t)u(t), \tag{3.66}$$

where $r(t) = \sqrt{2}a$ and u is defined such that:

$$u(t) = \begin{bmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{bmatrix}, \tag{3.67}$$

where θ is such that $\exists \theta_0 \in [0, \pi[$ such that $\forall i \in \mathbb{N}, \forall t \in [t_0 + i\Delta, t_0 + (i+1)\Delta)$:

$$\theta(t) = \theta_i,$$

where $(\theta_i)_{i\geq 0}$ is a sequence of angles. As $u(t)u^T(t) = \begin{bmatrix} \cos^2(\theta(t)) & \cos(\theta(t))\sin(\theta(t)) \\ \cos(\theta(t))\sin(\theta(t)) & \sin^2(\theta(t)) \end{bmatrix}$, we can limit ourselves to angles in $[0, \pi)$. This means that we can assume that θ_i takes only two values in such a way that $\theta_i \in \left\{\theta_0, \theta_0 + \frac{\pi}{2}\right\}$. It appears that the quantities that play a crucial role are the switching times of θ. We define the sequence of switching of θ, $(T_\theta^k)_{k\geq 0}$, recursively as follows, $\forall k \geq 0$:

$$\begin{aligned} T_\theta^0 &= \inf\{t \geq t_0 |\ \theta(t) = \theta_0 + \frac{\pi}{2}\}, \\ T_\theta^{2k+1} &= \inf\{t \geq T_\theta^{2k} |\ \theta(t) = \theta_0\}, \\ T_\theta^{2k+2} &= \inf\{t \geq T_\theta^{2k+1} |\ \theta(t) = \theta_0 + \frac{\pi}{2}\}, \end{aligned}$$

where we use the convention $\inf \varnothing = +\infty$. Note that if for some $k \geq 0$, $T_\theta^{2k} < +\infty$ (resp $T_\theta^{2k+1} < +\infty$) then $\forall t \in [T_\theta^{2k}, T_\theta^{2k+1})$, (resp $\forall t \in [T_\theta^{2k+1}, T_\theta^{2k+2})$), $\theta(t) = \theta_0 + \frac{\pi}{2}$ (resp $\theta(t) = \theta_0$). The following proposition gives a characterisation of persistent signal of the form (3.64).

Proposition 3.10. *If V is piecewise constant as described in (3.64) and decomposed as in (3.67), then the following are equivalent:*

(i) $\forall k \geq 0$, $T_\theta^k < +\infty$, *sup*$\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\} < +\infty$, *and sup*$\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\} < +\infty$,

(ii) $\exists T > 0$, $\exists \mu > 0$ *such that* $\forall t \geq t_0$*:*

$$\frac{1}{T}\left(\int_t^{t+T} V(s)V^T(s)ds\right) \succeq \mu I_2. \tag{3.68}$$

Proof. First from (3.67), one can deduce that for any $t \geq t_0$ and $T > 0$:

$$\int_t^{t+T} V(s)V^T(s)ds = 2a^2 \int_t^{t+T} u(s)u^T(s)ds$$

Thus u is persistent if and only if V is.

$(ii) \Rightarrow (i)$: Suppose $\exists T > 0$, $\exists \mu > 0$ such that $\forall t \geq t_0$:

$$\frac{1}{T}\left(\int_t^{t+T} u(s)u^T(s)ds\right) \succeq \mu I_2. \tag{3.69}$$

First, it is clear that $\forall k \geq 0$, $T_\theta^k < +\infty$ if and only if $T_\theta^0 < +\infty$ and $\forall k \geq 0$, $T_\theta^k < +\infty \Rightarrow T_\theta^{k+1} < +\infty$. Assume now by contradiction, that (i) does not hold true. Then

$$T_\theta^0 = +\infty,$$

or $\exists k \geq 0$ such that either

$$T_\theta^k < +\infty \text{ and } T_\theta^{k+1} = +\infty,$$

or

$$\sup\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\} = +\infty,$$

or

$$\sup\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\} = +\infty.$$

If $T_\theta^0 = +\infty$, then $\forall t \geq t_0$, $\theta(t) = \theta_0$ so $\forall t \geq t_0$:

$$\det\left(\int_t^{t+T} u(s)u^T(s)ds\right) = \det(u(t_0)u^T(t_0)) = 0,$$

which contradicts (ii). Similarly if $\exists k \geq 0$ such that $T_\theta^k < +\infty$ and $T_\theta^{k+1} = +\infty$ then θ is constant after T_θ^k which leads $\forall t \geq T_\theta^k$ to:

$$\det\left(\int_t^{t+T} u(s)u^T(s)ds\right) = 0$$

Which also contradicts (ii).

If $\sup\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\} = +\infty$, then there exists a subsequence $(T_\theta^{2n(k)+1} - T_\theta^{2n(k)})_{k\geq 0}$ such that $\lim_{k\to+\infty} T_\theta^{2n(k)+1} - T_\theta^{2n(k)} = +\infty$. So $\exists k_0 \in \mathbb{N}$, such that $T_\theta^{2n(k_0)+1} - T_\theta^{2n(k_0)} \geq T+1$. In particular, $u \equiv u(T_\theta^{2n(k_0)})$ on $[T_\theta^{2n(k_0)}, T_\theta^{2n(k_0)} + T)$ which leads to:

$$\det\left(\int_{T_\theta^{2n(k_0)}}^{T_\theta^{2n(k_0)}+T} u(s)u^T(s)ds\right) = \det\left(u\left(T_\theta^{2n(k_0)}\right)u^T\left(T_\theta^{2n(k_0)}\right)\right) = 0$$

which contradicts (ii). We get similarly a contradiction if $\sup\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\} = +\infty$. One gets to a contradiction in each case so (i) holds true.

$(i) \Rightarrow (ii)$: Assume $\forall k \geq 0, T_\theta^k < +\infty$, $\sup\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\} < +\infty$, and $\sup\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\} < +\infty$. Set $T_{\max,\frac{\pi}{2}} = \sup\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\}$ and $T_{\max,0} = \sup\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\}$.

From these, we define a candidate horizon of persistence of excitation $T = \max(T_{\max,\frac{\pi}{2}}, T_{\max,0}) + 2\Delta$. Intuitively, θ cannot stay constant during more than $\max(T_{\max,\frac{\pi}{2}}, T_{\max,0})$. Therefore, if $T > \max(T_{\max,\frac{\pi}{2}}, T_{\max,0})$ then the value of θ changes at least once which happens to be sufficient to get persistence of excitation. We add specifically 2Δ to ensure that $\int_t^{t+T} u(s)u^T(s)ds$ can be bounded from below by a matrix independent of t.

To do so precisely, we fix $t \geq t_0$ and and $i \in \mathbb{N}$ such that $t_i \leq t \leq t_{i+1}$ with $t_i = t_0 + i\Delta$. We also set $M_i = u(t_i)u^T(t_i)$. From (3.67), one gets, $\forall i \geq 0$:

$$M_i = \begin{bmatrix} \cos^2(\theta_i) & \cos(\theta_i)\sin(\theta_i) \\ \cos(\theta_i)\sin(\theta_i) & \sin^2(\theta_i) \end{bmatrix},$$
$$M_i \in \{A_1, A_2\},$$

with $A_1 = \begin{bmatrix} \cos^2(\theta_0) & \cos(\theta_0)\sin(\theta_0) \\ \cos(\theta_0)\sin(\theta_0) & \sin^2(\theta_0) \end{bmatrix}$ and $A_2 = \begin{bmatrix} \sin^2(\theta_0) & -\cos(\theta_0)\sin(\theta_0) \\ -\cos(\theta_0)\sin(\theta_0) & \cos^2(\theta_0) \end{bmatrix}$.

By (3.64), one can say that the switching times T_θ^k are exactly of the following form, $\forall k \geq 0$:

$$T_\theta^k = t_0 + p_k \Delta, \tag{3.70}$$

with $p_k \in \mathbb{N}$. This is due to the fact that V is right continuous and that all the switches happen at a time of the form (3.70). Thus, since the suprema in (i) are finite, they are equal to suprema on a bounded subset on $\mathbb{N}$ so they are attained. Therefore, from the definition of T and equation (3.70), $p = \frac{T}{\Delta} \in \mathbb{N}$. Moreover as, clearly, $T_{\max,\frac{\pi}{2}} \geq \Delta$ and $T_{\max,0} \geq \Delta$, $T \geq 3\Delta$ and $p \geq 3$ This allows us to compute the receding horizon mean of uu^T in the following way:

$$\begin{aligned}
\frac{1}{T}\left(\int_t^{t+T} u(s)u^T(s)ds\right) &= \frac{1}{T}\left(\int_t^{t_{i+1}} u(s)u^T(s)\right) + \frac{1}{T}\left(\int_{t_{i+1}}^{t_{i+p-1}} u(s)u^T(s)\right) + \frac{1}{T}\left(\int_{t_{i+p}}^{t+T} u(s)u^T(s)\right), \\
&= \frac{1}{T}\left((t_{i+1}-t)M_i + \Delta\sum_{j=1}^{p-1} M_{i+j} + (t+T-t_{i+p})M_{i+p}\right) \\
&\succeq \frac{\Delta}{T}\sum_{j=1}^{p-1} M_{i+j}
\end{aligned}$$

As mentioned earlier, u cannot be constant on $[t_{i+1}, t_{i+p-1})$ because $t_{i+p-1} - t_{i+1} = T - \Delta > \max(T_{\max,\frac{\pi}{2}}, T_{\max,0})$. This means that there exist $1 \leq j_1, j_2 \leq p-1$, with $j_1 \neq j_2$ such that $M_{i+j_1} = A_1$ and $M_{i+j_2} = A_2$. Finally,

$$\begin{aligned}
\frac{1}{T}\left(\int_t^{t+T} u(s)u^T(s)ds\right) &\succeq \frac{\Delta}{T}(M_{i+j_1} + M_{i+j_2}), \\
&\succeq \frac{\Delta}{T}(A_1 + A_2),
\end{aligned}$$

By noticing that $A_1 + A_2 = I_2$, one obtains:

$$\begin{aligned}
\frac{1}{T}\left(\int_t^{t+T} u(s)u^T(s)ds\right) &\succeq \frac{\Delta}{T} I_2, \\
\frac{1}{T}\left(\int_t^{t+T} V(s)V^T(s)ds\right) &\succeq 2a^2\frac{\Delta}{T} I_2,
\end{aligned}$$

and (ii) is proved with $\mu = 2a^2\frac{\Delta}{T}$. □

3.2.2.2 Almost sure persistence of excitation

The main consequence of Proposition 3.10 is that, in order to show almost sure PE, one only needs to check that the angle θ satisfies property (i) in Proposition 3.10 almost surely. It is a substantial simplification because condition (i) in Proposition 3.10 only involves discrete-time quantities. As a result, the idea of this section is to give an example of process of the form (3.64) that satisfies (i) from Proposition 3.10.

To do so, consider a stochastic process $(\theta(\omega,t))_{\omega\in\Omega,\, t\geq t_0}$ that is almost surely of the form described in (3.67). It means that it can be represented by a discrete-time process $(\theta_i(w))_{\omega\in\Omega,\, i\geq 0}$. Let us set $A = \{\omega |\ \theta$ verifies (i) from Proposition 3.10$\}$. One can deduce that:

$$A = \{\omega |\ \forall k \geq 0, T_\theta^k < +\infty,\ \sup\{k \geq 0 |\ T_\theta^{2k+1} - T_\theta^{2k}\} < +\infty,\ \sup\{k \geq 0 |\ T_\theta^{2k+2} - T_\theta^{2k+1}\} < +\infty\},$$
$$\{\ \omega |\ \exists M > 0\ s.t.\ \forall k \geq 0, T_\theta^k < +\infty,\ T_\theta^{2k+1} - T_\theta^{2k} \leq M,\ T_\theta^{2k+2} - T_\theta^{2k+1} \leq M\}.$$

By definition of T_θ^{2k+1} and T_θ^{2k+2},

$$A = \Big\{\ \omega |\ \forall k \geq 0, T_\theta^k < +\infty,\ \text{and } \exists M > 0\ s.t.\ \forall i \geq 0\ (\theta_i, \ldots, \theta_{i+M}) \neq (\theta_0, \ldots, \theta_0),\ \text{and}$$
$$(\theta_i, \ldots, \theta_{i+M}) \neq \Big(\theta_0 + \frac{\pi}{2}, \ldots, \theta_0 + \frac{\pi}{2}\Big)\Big\}, \tag{3.71}$$

The objective of the following is to build an example of process (θ_i) such that $P(A) = 1$. To do so, first we fix $M \in \mathbb{N}^*$, and set:

$$A_M = \Big\{\ \omega |\ \forall k \geq 0, T_\theta^k < +\infty,\ \text{and } \forall i \geq 0\ (\theta_i, \ldots, \theta_{i+M}) \neq (\theta_0, \ldots, \theta_0),\ \text{and}$$
$$(\theta_i, \ldots, \theta_{i+M}) \neq \Big(\theta_0 + \frac{\pi}{2}, \ldots, \theta_0 + \frac{\pi}{2}\Big)\Big\}. \tag{3.72}$$

Clearly, $A_M \subset A$. Then, we consider $(\theta_i)_{i\geq 0}$ as the sample paths of a M-dependent process on $\{\theta_0, \theta_0 + \frac{\pi}{2}\}$ with the following probability distribution, $\forall i \geq 0$:

$$\begin{aligned} P(\theta_{i+M} = \theta_0 |\ \theta_{i+M-1} = \theta_0, \ldots, \theta_i = \theta_0) &= 0, \\ P\Big(\theta_{i+M} = \theta_0 + \frac{\pi}{2} |\ \theta_{i+M-1} = \theta_0 + \frac{\pi}{2}, \ldots, \theta_i = \theta_0 + \frac{\pi}{2}\Big) &= 0, \\ P\left(\theta_{i+M} = \theta_0 |\ \exists\ (i_1, i_2) \in 0, .., M-1\ s.t.\ i_1 \neq i_2 \text{ and } \theta_{i_1} \neq \theta_{i_2}\right) &= p, \end{aligned} \tag{3.73}$$

with $0 < p < 1$.

Intuitively, Equation (3.73) means that sample paths of $(\theta_i(w))_{\omega\in\Omega,\ i\geq 0}$ cannot be constant during a time window strictly greater than M. By simple manipulations of unions of set and of probabilities, one gets:

$$P(\exists i \geq 0,\ \theta_{i+M} = \theta_0, \ldots, \theta_i = \theta_0 \text{ or } \theta_{i+M} = \theta_0 + \frac{\pi}{2}, \ldots, \theta_i = \theta_0 + \frac{\pi}{2}) = 0,$$
$$P(\exists k \geq 0, T_\theta^k = +\infty) = 0.$$

and $P(A_M^c) = 0$ which leads to $P(A_M) = 1$ and $P(A) = 1$.

Finally, a process defined by (3.73) verifies (i) from Proposition 3.10 with probability one. In particular, by Proposition 3.10, the continuous time process $(\theta(\omega, t))_{\omega\in\Omega,\ t\geq t_0}$ defined by $(\theta_i(w))_{\omega\in\Omega,\ i\geq 0}$ is persistent almost surely and V verifies Assumption 3.1 almost surely.

As an extension of this result, on could imagine reproducing these arguments for higher dimensional vectors V using an argument of recursion on the dimension. One could also study a process V that does not have a constant norm or with time intervals of varying size.

To conclude Part I, one can say that the modelling of TAN provides tools to build nonlinear observers and adapted controllers for interesting classes of ground maps. The Remarks 3.2 and 3.4 also give hints on how to deal with real maps using analytical ones. However, this framework seems not to be able to provide a systematic treatment of real empirical ground. That is why we move to a discrete-time stochastic framework to deal with this issue in Part II.

Part II

Terrain-Aided Navigation with real ground maps

Chapter 4

Elements of stochastic control and stochastic filtering

We have seen in Part I that the problem of TAN with a real ground map is hard to treat directly with analytical methods because they require too much structure on the terrain profile and are not directly compatible with terrain coming from numerical data. It turns out that the deterministic setting of Part I is too rigid. In particular, the notion of observer of Proposition 3.1 is too demanding. Actually, terrain-based estimation is classically solved by Bayesian methods which means that the stochastic framework is natural for this application. Moreover, it is very hard to find explicit control laws in TAN with a real map for the same reasons as in estimation. Therefore, optimal control techniques that allow one to define a control implicitly for very general systems is much more promising. As in Part I, TAN with a real map also need control and estimation methods that depend on each other. Motivated by the last remarks, the goal of Part II is to present methods of optimal control and estimation methods for a general nonlinear discrete-time stochastic system that are applicable to TAN with a real ground map.

We start by reviewing the classical methods of control and estimation in a stochastic discrete-time framework, in Chapter 4, in the following way. First, the basics of optimal stochastic control are recalled. Secondly, the principles of nonlinear stochastic filtering are presented and, in particular, the concepts of particle filtering. Finally, the notions of dual effect and Stochastic Model Predictive Control (SMPC) detailed.

4.1 Discrete-time Stochastic Optimal Control

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For $i \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^i)$ denotes the Borel σ-algebra of $\mathbb{R}^i$. $\mathcal{P}(\mathbb{R}^i)$ denotes the set of probability measures on $\mathbb{R}^i$. In the following, random variables refer to measurable functions from Ω to $\mathbb{R}^i$. For a random variable X and a probability distribution, $X \sim p$ means that p is the probability law of X. $P(\cdot|\cdot)$ and $E(\cdot|\cdot)$ denotes the conditional probability and expectation. For $\mu \in \mathcal{P}(\mathbb{R}^i)$ and $f\colon \mathbb{R}^i \longrightarrow \mathbb{R}^j$ integrable, we set $\langle \mu, f \rangle = \int_{\mathbb{R}^i} f(x)\mu(\mathrm{d}x)$.

The goal of this section is to recall the basics of discrete-time deterministic optimal control, stochastic optimal control including perfect/imperfect information problems with finite and infinite horizon.

4.1.1 Deterministic case

4.1.1.1 Definitions

First, we recall the concept of deterministic discrete-time optimal control. The state of the controlled deterministic dynamical system we consider is represented by a sequence $(x_k)_{k\in\mathbb{N}}$ valued in $\mathbb{R}^{n_x}$. It is described by the following equation, $\forall k \in \mathbb{N}$:

$$x_{k+1} = f^{det}(x_k, u_k), \tag{4.1}$$

where:

- x_0 is a fixed initial condition.
- $(u_k)_{k\in\mathbb{N}}$ is the control sequence valued in $\mathcal{U} \subset \mathbb{R}^{n_u}$. $\mathcal{U}$ is the set of admissible control values.
- f^{det}: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \longrightarrow \mathbb{R}^{n_x}$ is (Borel)-measurable.

Additionally, to define an optimal control problem in this framework, one needs:

- A time horizon, $T \in \mathbb{N}^* \cup \{+\infty\}$. $T < +\infty$ corresponds to a finite horizon problem while $T = +\infty$ corresponds to a infinite horizon one.
- If $T < +\infty$, $\forall i = 0, .., T-1$, an instantaneous cost g_i: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \longrightarrow \mathbb{R}$ and a final cost g_F: $\mathbb{R}^{n_x} \longrightarrow \mathbb{R}^+$.
- If $T <= \infty$, we consider for simplicity a time homogeneous instantaneous cost g: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \longrightarrow \mathbb{R}^+$ and a discount factor $\alpha \in (0, 1]$.

Then, a finite horizon optimal control problem in this framework is of the following form, for any initial condition $x_0 \in \mathbb{R}^{n_x}$:

$$\begin{array}{rrcl} V_0^{det}(x_0) = \min\limits_{u_0,\dots,u_{T-1}\in U} & \sum_{k=0}^{T-1} g_k(x'_k, u_k) + g_F(x'_T) & & \\ \text{s.t.} & x'_{k+1} & = & f^{det}(x'_k, u_k),\ \forall k = 0, .., T-1, \\ & x'_0 & = & x_0. \end{array} \tag{4.2}$$

Similarly, a infinite horizon optimal control problem is of the following form:

$$\begin{array}{rrcl} V^{det}(x_0) = \min\limits_{u_0,\dots,u_k,\dots\in U} & \sum_{k=0}^{+\infty} \alpha^k g(x'_k, u_k) & & \\ \text{s.t.} & x'_{k+1} & = & f^{det}(x'_k, u_k),\ \forall k \geq 0, \\ & x'_0 & = & x_0. \end{array} \tag{4.3}$$

$V_0^{det}(x_0)$ and $V^{det}(x_0)$ are respectively the optimal values of the problems (4.2) and (4.3).

Remark 4.1.

- The cost function g is assumed to be non-negative to ensure that the problem (4.3) is well defined. In the sequeal, we write "min" instead of "inf"' in (4.2) and (4.3) because we assume that the problems always admit a solution.

- Under additional assumption of regularity on the cost function and the set of admissible controls, the problem (4.2) is indeed a classical nonlinear program which is nonconvex as soon as the dynamics (4.1) is nonlinear. Problem (4.2) appears typically after the discretisation of a continuous-time deterministic problem or as an approximation of a discrete-time stochastic one.

- In deterministic control, open-loop control and feedback control represent the same class of control law. More precisely, on one hand, an admissible open loop sequence of controls of the problem (4.2), $(u_0, \ldots, u_{T-1})$, can be written as a function of x_0 such that, for $k = 0, .., T-1$:

$$u_k = \pi'_k(x_0),$$

where π'_k: $\mathbb{R}^{n_x} \longrightarrow \mathcal{U} \subset \mathbb{R}^{n_u}$ is measurable. On the other hands, feedback controls are functions of present and the past of the state of the system meaning that for $k = 0, .., T-1$:

$$u_k = \pi_k(x_0, \ldots, x_k),$$

where π_k: $(\mathbb{R}^{n_x})^{k+1} \longrightarrow \mathcal{U} \subset \mathbb{R}^{n_u}$. It clear that any open loop control law can be seen as a feedback control law where π_k depends only on x_0. Conversely, from equation (4.1), the state is exactly predictable from the initial condition, x_0, so $(x_0, \ldots, x_k)$ can seen as a *deterministic* function of x_0 such that:

$$(x_0, \ldots, x_k) = F^{det}(x_0),$$

where F^{det}: $\mathbb{R}^{n_x} \longrightarrow (\mathbb{R}^{n_x})^{k+1}$ is measurable. By setting $\pi'_k = \pi_k \circ F^{det}$, one gets an open loop control law. In other words, the knowledge of the initial condition x_0 or of the whole past and present of the state of the system $(x_0, \ldots, x_k)$ are equivalent from the point of view of control design. Besides, due to the dependency structure of equation (4.1), feedback of the form $u_k = \pi(x_k)$ are sufficient to get optimal controls.

4.1.1.2 Dynamic programming principle

We set additionally, for $l = 0..T-1$ the *cost-to-go* of problem (4.2) and $x \in \mathbb{R}^{n_x}$ as follows:

$$\begin{aligned} V_\ell^{det}(x) = \min_{u_\ell, \ldots, u_{T-1} \in U} \quad & \textstyle\sum_{k=\ell}^{T-1} g_k(x'_k, u_k) + g_F(x'_T) \\ \text{s.t.} \quad & x'_{k+1} = f^{det}(x'_k, u_k),\ \forall k = \ell, .., T-1, \\ & x'_\ell = x, \\ V_T^{det}(x) = g_F(x). & \end{aligned} \tag{4.4}$$

The cost-to-go of the problem (4.2) satisfy the following backward recursive equation known as the Bellman equation or the Dynamic Programming (DP) principle [Bellman, 1966]. For $l = 0..T-1$, and $x \in \mathbb{R}^{n_x}$:

$$\begin{aligned} V_\ell^{det}(x) = \min_{u \in U} \quad & g_\ell(x'_\ell, u) + V_{\ell+1}^{det}(x'_{\ell+1}) \\ \text{s.t.} \quad & x'_{\ell+1} = f^{det}(x'_\ell, u), \\ & x'_\ell = x. \\ V_T^{det}(x) = g_F(x). & \end{aligned} \tag{4.5}$$

The major usefulness of this formulation is that, if, for $\ell = 0, .., T-1$, u_ℓ^* is a solution of (4.5) then the vector of these controls, $(u_0^*, \dots, u_{T-1}^*)$, is a solution of the problem (4.2). In the infinite horizon case, because of the time-homogeneity hypothesis, the DP principle can be written as one fixed point equation, for $x \in \mathbb{R}^{n_x}$:

$$\begin{aligned} V^{det}(x) = \min_{u \in U} \quad & g(x,u) + \alpha V^{det}(x^+) \\ \text{s.t.} \quad & x^+ = f^{det}(x,u). \end{aligned} \tag{4.6}$$

Similarly to the finite horizon case, if $\pi(x) \in \underset{u \in U}{\operatorname{argmin}}\ g(x,u) + \alpha V^{det}(x^+)$, the sequence $(\pi(x_0), \dots)$ is optimal for (4.3).

The main inconvenient of Dynamic Programming is that it requires to be able to compute the optimal values $V_\ell^{det}(x)$ and $V^{det}(x)$ for any x. In practice, it induces some space discretisation procedure and if n_x is too large (typically, $n_x \geq 5$), the computations become intractable. This issue is called the *curse of dimensionality* [Bertsekas, 2011].

4.1.2 Stochastic case with perfect information

4.1.2.1 Definition

We consider now a discrete-time process $X = (X_k)_{k \in \mathbb{N}}$ valued in $\mathbb{R}^{n_x}$ representing the state of a controlled stochastic dynamical system described by the following equation, $\forall k \in \mathbb{N}$:

$$\begin{aligned} X_{k+1} &= f(X_k, U_k, \xi_k), \\ X_0 &\sim p_0, \end{aligned} \tag{4.7}$$

where:

- p_0 is a probability law on $\mathbb{R}^{n_x}$.
- $(U_k)_{k \in \mathbb{N}}$ is the control process valued in $\mathcal{U} \subset \mathbb{R}^{n_u}$. $\mathcal{U}$ is the set of admissible control values.
- $(\xi_k)_{k \in \mathbb{N}}$ are i.i.d. random variables valued in $\mathbb{R}^{n_\xi}$ distributed according to p_ξ. For each $k \in \mathbb{N}$, ξ_k represents an external disturbance on the dynamics.
- f: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \longrightarrow \mathbb{R}^{n_x}$ is measurable.

In fact, equation (4.7) defines $(X_k)_{k \in \mathbb{N}}$ as a Markov Decision Process on $\mathbb{R}^{n_x}$. Its transition kernel is denoted by K meaning that $\forall A \in \mathcal{B}(\mathbb{R}^{n_x})$, for almost all (a.a.), $x \in \mathbb{R}^{n_x}$ and a.a. $u \in \mathbb{R}^{n_u}$:

$$P(X_{k+1} \in A | X_k = x,\ U_k = u) = K(A, x, u), \tag{4.8}$$

where K: $\mathcal{B}(\mathbb{R}^{n_x}) \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \longrightarrow \mathbb{R}^+$ is such that:

- $\forall A \in \mathcal{B}(\mathbb{R}^{n_x})$, $K(A, \cdot, \cdot)$ is measurable.
- $\forall x \in \mathbb{R}^{n_x}$, $\forall u \in \mathbb{R}^{n_u}$, $K(\cdot, x, u)$ is a probability measure on $\mathbb{R}^{n_x}$.

Note that "almost all" in this case means almost surely with respect to the joint distribution of X_k and U_k. We omit this in the sequel concerning conditional probability and expectation if no ambiguity appears. For any fixed control process $U = (U_k)_{k\in\mathbb{N}}$, we denote the probability distribution of X with control U when $X_0 \sim p_0$ and the resulting expectation operator respectively by $P^U_{p_0}$ and $E^U_{p_0}$. Likewise, for $x \in \mathbb{R}^{n_x}$, P^U_x and E^U_x denote the the probability and expectation of X with control U conditionally to $X_0 = x$. Sometimes in the sequel we will assume that K has a density w.r.t. the Lebesgue measure. In this case, with a slight abuse of notation, (4.8) is written $\forall A \in \mathcal{B}(\mathbb{R}^{n_x})$, for $x \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$:

$$P(X_{k+1} \in A | X_k = x,\ U_k = u) = \int_A K(y, x, u)\mathrm{d}y, \tag{4.9}$$

where dy is the Lebesgue measure.

In this section, we assume that X_k is exactly known at any time k and therefore available for the computation of U_k. In stochastic control, the nature of the decision, i.e. the choice of the control U, is structurally different from the one in deterministic control as optimal feedback controls are better than optimal open loop controls. Another way to describe it is to say that feedback control cannot be seen as open loop control anymore in the stochastic case. Indeed, as in the previous section, open loop controls can be written as follows, for $k \geq 0$:

$$U_k = \pi'_k(X_0), \tag{4.10}$$

and feedback controls in the following way:

$$U_k = \pi_k(X_0, \ldots, X_k). \tag{4.11}$$

It is clear that open-loop controls are still feedback controls but the converse no longer holds. Actually, $(X_0, \ldots, X_k)$ is not a *deterministic* function of X_0 anymore. The knowledge of $(X_0, \ldots, X_k)$ covers the knowledge of X_0 but is not equivalent to it because of the unforeseeable nature of the stochastic disturbances in (4.7). This makes the class of feedback controls larger than the open loop control and confirms our first claim that optimal feedback controls are better than optimal open-loop ones. This phenomenon can be summarised as the fact that more information leads to better controls. It is called the *value of information*. Illustrative examples of this property are given in [Bertsekas, 2011]. Due to the Markov property of the system (4.7), the feedback controls (4.11) can be reduced to controls depending only on the current state X_k, (see [Hernández-Lerma and Lasserre, 1996] for the details). The controls are then chosen as follows for $k \geq 0$:

$$U_k = \pi_k(X_k), \tag{4.12}$$

where π_k: $\mathbb{R}^{n_x} \longrightarrow \mathcal{U} \subset \mathbb{R}^{n_u}$ is measurable.

Similarly to the deterministic case, one needs the following additional elements to define a stochastic optimal control problem:

- A time horizon, $T \in \mathbb{N}^* \cup \{+\infty\}$. $T < +\infty$ corresponds to a finite horizon problem while $T = +\infty$ corresponds to a infinite horizon one.
- If $T < +\infty$, $\forall i = 0, .., T-1$, an instantaneous cost g_i: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \longrightarrow \mathbb{R}^+$ and a final cost g_F: $\mathbb{R}^{n_x} \longrightarrow \mathbb{R}^+$.
- If $T = \infty$, we consider also a time homogeneous instantaneous cost g: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \longrightarrow \mathbb{R}^+$ and a discount factor $\alpha \in]0, 1]$.

All the costs functions are assumed to be nonnegative for simplicity. A finite sequence of measurable functions as in (4.12), $\pi = (\pi_0, \dots, \pi_T)$ with $T \in \mathbb{N}^*$ is called a finite horizon control *policy*. A infinite sequence $\pi = (\pi_k)_{k\geq 0}$ is called a infinite horizon control policy. Policies with elements of the form (4.12) are called non-randomised Markov policies.

The goal of finite horizon discrete-time stochastic control is then to find a finite horizon non-randomised Markov control policy that minimises an expected cost under constraint on the control. We omit hard and soft constraints on the state of the system because they are not of paramount importance in our treatment. However, they could be added with an additional assumption of admissibility.

When the control U is chosen according to a policy π, we will denote the probability and expectations P^π and E^π instead of P^U and E^U. In the infinite horizon case, due to time homogeneity, the optimal policy can be looked for as the repetition of the same function π_0 so that $\pi = (\pi_0, \pi_0, \dots)$. This leads to the following optimisation problem, with a finite horizon, for $x_0 \in \mathbb{R}^{n_x}$:

$$\begin{array}{rrcl} V_0(x_0) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^\pi_{x_0}\left[\sum_{k=0}^{T-1} g_k(X_k, U_k, \xi_k) + g_F(X_T)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k, U_k, \xi_k), \\ & U_k & = & \pi_k(X_k),\ \forall k = 0, .., T-1, \\ & X_0 & = & x_0. \end{array} \tag{4.13}$$

and its equivalent with an infinite horizon:

$$\begin{array}{rrcl} V(x_0) = \min\limits_{\pi_0} & E^\pi_{x_0}\left[\sum_{k=0}^{+\infty} \alpha^k g(X_k, U_k, \xi_k)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k, U_k, \xi_k), \\ & U_k & = & \pi_0(X_k),\ \forall k \geq 0, \\ & X_0 & = & x_0, \end{array} \tag{4.14}$$

where $V_0(x_0)$ and $V(x_0)$ are then the optimal values. When $X_0 \sim p_0$, the two previous problems read:

$$\begin{array}{rrcl} V_0(p_0) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^\pi_{p_0}\left[\sum_{k=0}^{T-1} g_k(X_k, U_k, \xi_k) + g_F(X_T)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k, U_k, \xi_k), \\ & U_k & = & \pi_k(X_k),\ \forall k = 0, .., T-1, \\ & X_0 & \sim & p_0. \end{array} \tag{4.15}$$

and its equivalent with an infinite horizon:

$$\begin{array}{rrcl} V(p_0) = \min\limits_{\pi_0} & E^\pi_{p_0}\left[\sum_{k=0}^{+\infty} \alpha g(X_k, U_k, \xi_k)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k, U_k, \xi_k), \\ & U_k & = & \pi_0(X_k),\ \forall k \geq 0, \\ & X_0 & \sim & p_0. \end{array} \tag{4.16}$$

Remark 4.2.

- We recall that we assume that the problems (4.13) and (4.14) have solutions. Classical sufficient condition of existence of solutions for those problems are given in [Hernández-Lerma and Lasserre, 1996].
- The problems (4.13) and (4.14) are infinite dimensional so they are hard to solve directly. To do so, an alternative method is to use the DP principle adapted to this framework.

4.1.2.2 Dynamic Programming principle

We present here the version of the DP principle in the stochastic case with imperfect information. The finite horizon cost-to-go is defined as follows, for $\ell = 1, .., T-1$ and $x \in \mathbb{R}^{n_x}$:

$$\begin{aligned} V_\ell(x) = \min_{\pi_\ell,\dots,\pi_{T-1}} \quad & E^\pi\left[\sum_{k=\ell}^{T-1} g_k(X_k, U_k, \xi_k) + g_F(X_T) | X_\ell = x\right] \\ \text{s.t.} \quad & X_{k+1} = f(X_k, U_k, \xi_k), \\ & U_k = \pi_k(X_k),\ \forall k = \ell, .., T-1, \\ V_T(x) = g_F(x). & \end{aligned} \tag{4.17}$$

As in the deterministic case, the cost-to-go, for $\ell = 1, .., T-1$ and $x \in \mathbb{R}^{n_x}$, V_ℓ satisfies the following Bellman equation:

$$\begin{aligned} V_\ell(x) = \min_{u \in U} \quad & E\left[g_\ell(X_\ell, u, \xi_\ell) + V_{\ell+1}(X_{\ell+1}) | X_\ell = x\right] \\ \text{s.t.} \quad & X_{\ell+1} = f(X_\ell, u, \xi_\ell). \\ V_T(x) = g_F(x). & \end{aligned} \tag{4.18}$$

As in the deterministic case, one can compute optimal solutions of (4.13) by solving iteratively (4.18) More precisely, if $\pi_\ell^*(x) \in \underset{u \in U}{\operatorname{argmin}}\ E\left[g_\ell(X_\ell, u, \xi_\ell) + V_{\ell+1}(X_{\ell+1}) | X_\ell = x\right]$ then $(\pi_0^*, \dots, \pi_{T-1}^*)$ is a solution of (4.13).

In the infinite horizon case, the DP principle reads:

$$\begin{aligned} V(x) = \min_{u \in U} \quad & E\left[g(X_\ell, u, \xi_\ell) + \alpha V(X_{\ell+1}) | X_\ell = x\right] \\ \text{s.t.} \quad & X_{\ell+1} = f(X_\ell, u, \xi_\ell). \end{aligned} \tag{4.19}$$

As for the finite horizon case, if $\pi_0^*(x) \in \underset{u \in U}{\operatorname{argmin}}\ E\left[g(X_\ell, u, \xi_\ell) + \alpha V(X_{\ell+1}) | X_\ell = x\right]$ then $(\pi_0^*, \pi_0^*, \dots)$ is a solution of (4.14). Obviously, the curse of dimensionality is still a problem in the stochastic case. Besides, equation (4.18) and (4.19) require the resolution of a stochastic nonlinear program which is usually more difficult than solving a classical nonlinear program.

Assuming perfect information can be very limiting in many applications. In the following we recall the extension of the previous results to the case of imperfect information.

4.1.3 Stochastic case with imperfect information

4.1.3.1 Information vector formulation

In this section, we still assume that the state of the system is represented by $X = (X_k)_{k \in \mathbb{N}}$ and ruled by equation (4.7). However, we assume that the state of the system is only available through some observations represented by a stochastic process $Y = (Y_k)_{k \in \mathbb{N}}$ valued in $\mathbb{R}^{n_y}$ which verifies, $\forall k \in \mathbb{N}$:

$$Y_k = h(X_k, \eta_k), \tag{4.20}$$

where:

- $(\eta_k)_{k \in \mathbb{N}}$ are i.i.d. random variables valued in $\mathbb{R}^{n_\eta}$ distributed according to p_η. For each $k \in \mathbb{N}$, η_k represents an external disturbance on the observations.

- $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\eta} \longrightarrow \mathbb{R}^{n_y}$ is measurable.

Notice that, in general, discrete-time dynamics are seen as approximations of continuous-time ones which are more realistic. However, concerning observations, the converse is true. In practice, observations are available at discrete points in time. If the rate of measurement is high, then a model of continuous-time observations can used as an approximation of a discrete-time one.

For $k \in \mathbb{N}$, we define the vector of *available information* I_k as follows:

$$I_k = (Y_0, U_0, \ldots, Y_{k-1}, U_{k-1}, Y_k), \tag{4.21}$$

or equivalently, for $k \in \mathbb{N}$:

$$\begin{aligned} I_0 &= Y_0, \\ I_{k+1} &= (I_k, U_k, Y_{k+1}). \end{aligned} \tag{4.22}$$

Intuitively, I_k gathers all the values that are available to one to compute the control. In general, I_k does not contain the vector $(X_0, \ldots, X_k)$ so policies of the form (4.11) are not admissible. Therefore, one cannot do better than choosing U_k as a function of I_k, such that for $k \geq 0$:

$$U_k = \pi_k(I_k), \tag{4.23}$$

where π_k: $(\mathbb{R}^{n_y} \times \mathbb{R}^{n_u})^k \times \mathbb{R}^{n_y} \longrightarrow \mathcal{U} \subset \mathbb{R}^{n_u}$ is measurable. Using the same notations as in the perfect information case, stochastic optimal control problems with imperfect information can be formulated as follows, for $i_0 \in \mathbb{R}^{n_y}$:

$$\begin{aligned} V_0(i_0) = \min_{\pi_0,\ldots,\pi_{T-1}} \quad & E^{\pi}_{p_0}\left[\textstyle\sum_{k=0}^{T-1} g_k(X_k, U_k, \xi_k) + g_F(X_T) | I_0 = i_0\right] \\ \text{s.t.} \quad X_{k+1} &= f(X_k, U_k, \xi_k), \\ Y_k &= h(X_k, \eta_k), \\ I_{k+1} &= (I_k, U_k, Y_{k+1}), \\ U_k &= \pi_k(I_k),\ \forall k = 0, .., T-1, \end{aligned} \tag{4.24}$$

where the expectation is taken over the distribution of $(X_k, Y_k)_{k\in\mathbb{N}}$. The problem with infinite horizon reads for $i_0 \in \mathbb{R}^{n_y}$:

$$\begin{aligned} V_0(i_0) = \min_{\pi_0} \quad & E^{\pi}_{p_0}\left[\textstyle\sum_{k=0}^{+\infty} \alpha^k g(X_k, U_k, \xi_k) | I_0 = i_0\right] \\ \text{s.t.} \quad X_{k+1} &= f(X_k, U_k, \xi_k), \\ Y_k &= h(X_k, \eta_k), \\ I_{k+1} &= (I_k, U_k, Y_{k+1}), \\ U_k &= \pi_0(I_k),\ \forall k \geq 0. \end{aligned} \tag{4.25}$$

The DP principle cannot be applied directly to the problems (4.24) and (4.25) as the process $(X_k, Y_k)_{k\in\mathbb{N}}$ with control policies of the form (4.23) is no longer a Markov Chain. Actually, one can reformulate the problem (4.24) as a perfect information problem on the information space with I_k as the new state of the system and using equation (4.22) as the new dynamics. In fact, equation (4.22) describes a Markov chain on the information space (see [Bertsekas and Shreve, 2004] for a formal justification). This reformulation reads:

$$\begin{array}{rcl} V_0(i_0) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^{\pi}_{p_0}\left[\sum_{k=0}^{T-1}\hat{g}_k(I_k,U_k)+\hat{g}_F(I_T)|I_0=i_0\right] \\ \text{s.t.} & I_{k+1} = (I_k,U_k,Y_{k+1}), \\ & U_k = \pi_k(I_k),\ \forall k=0,..,T-1, \end{array} \tag{4.26}$$

with $\hat{g}_k(i_k,u) = E[g_k(X_k,u,\xi_k)|I_k=i_k]$ and $\hat{g}_F(i_T) = E[g_F(X_T)|I_T=i_T]$. Similarly, one gets the reformulation in the infinite horizon case:

$$\begin{array}{rcl} V(i_0) = \min\limits_{\pi_0} & E^{\pi}_{p_0}\left[\sum_{k=0}^{+\infty}\alpha^k\hat{g}(I_k,U_k)|I_0=i_0\right] \\ \text{s.t.} & I_{k+1} = (I_k,U_k,Y_{k+1}), \\ & U_k = \pi_0(I_k),\ \forall k\geq 0, \end{array} \tag{4.27}$$

where $\hat{g}(i_k,u) = E[g(X_k,u,\xi_k)|I_k=i_k]$.

A DP principle can be written on the information space but we omit it as it is not very useful. Actually, in practice, it is hard to exploit directly the information vector to gain knowledge on X_k. Besides, the dimension of I_k increases with time which makes the use of DP even less interesting. In the sequel, we use an alternative representation of I_k called the filtering distribution.

4.1.3.2 Filtering distribution formulation

From I_k, one can derive two important quantities in stochastic control with imperfect information that are the conditional distribution of X_k given I_k, called the filtering distribution, denoted by μ_k and the conditional distribution of X_{k+i} given $(I_k,U_k,\dots,U_{k+i-1})$ for any $i\in\mathbb{N}^*$, denoted by $\mu_{k+i|k}$. Moreover, in the following, we assume that the conditional distribution defined by equation (4.20) has a density with respect to the Lebesgue measure such that there exists a likelihood function denoted by ρ. Therefore, for $k\in\mathbb{N}$, $A\in\mathcal{B}(\mathbb{R}^{n_x})$ and $B\in\mathcal{B}(\mathbb{R}^{n_y})$:

$$P(Y_k\in B|X_k=x_k) = \int_B \rho(y_k,x_k)\mathrm{d}y_k.$$

μ_0 is supposed to be known. Thus, $\forall(k,i)\in\mathbb{N}^2$, and $\forall A\in\mathcal{B}(\mathbb{R}^{n_x})$, μ_k and $\mu_{k+i|k}$ verify the following nonlinear filtering equations:

$$\mu_{k+1|k}(A) = \int_{\mathbb{R}^{n_x}} K(A,x_k,U_k)\mu_k(\mathrm{d}x_k), \tag{4.28}$$

$$\mu_{k+1|k+1}(A) = \frac{\int_A \rho(Y_{k+1},x_{k+1})\mu_{k+1|k}(\mathrm{d}x_{k+1})}{\int_{\mathbb{R}^{n_x}} \rho(Y_{k+1},x_{k+1})\mu_{k+1|k}(\mathrm{d}x_{k+1})}, \tag{4.29}$$

$$\mu_{k+i+1|k}(A) = \int_{\mathbb{R}^{n_x}} K(A,x_{k+i},U_{k+i})\mu_{k+i|k}(\mathrm{d}x_{k+i}). \tag{4.30}$$

One can sum up equations (4.28) to (4.30), in the following way, $\forall(k,i)\in\mathbb{N}^2$:

$$\mu_{k+1} = F\left(\mu_k,Y_{k+1},U_k\right), \tag{4.31}$$

$$\mu_{k+i+1|k} = G\left(\mu_{k+i|k},U_{k+i}\right), \tag{4.32}$$

where $F:\mathcal{P}(\mathbb{R}^{n_x})\times\mathbb{R}^{n_y}\times\mathbb{R}^{n_u}\longrightarrow\mathcal{P}(\mathbb{R}^{n_x})$ and $G:\mathcal{P}(\mathbb{R}^{n_x})\times\mathbb{R}^{n_u}\longrightarrow\mathcal{P}(\mathbb{R}^{n_x})$.

μ_k is of central importance in Bayesian filtering as it contains and weighs the possible values of the current state X_k knowing only the value of I_k. $\mu_{k+i+1|k}$ is more useful in predictive control because it represents the future possible values of the state knowing only the current information.

It can be shown that μ_k carries as much information as I_k. More precisely, looking at μ_k as random variable on $\mathcal{P}(\mathbb{R}^{n_x})$ equipped with the Borel sigma-algebra for the weak topology, it is a sufficient statistics (see [Bertsekas and Shreve, 2004], [Bertsekas, 2011]). It means that the control policies can be looked for as functions of μ_k instead of I_k such that, for $k \geq 0$:

$$U_k = \pi_k(\mu_k), \tag{4.33}$$

where π_k: $\mathcal{P}(\mathbb{R}^{n_x}) \longrightarrow \mathcal{U} \subset \mathbb{R}^{n_u}$ is measurable. Moreover, another classical result in [Stettner, 1989] and [Bertsekas and Shreve, 2004] is that equations (4.28) and (4.29) describe a Markov Chain on $\mathcal{P}(\mathbb{R}^{n_x})$. This implies that the problem (4.24) can be seen as a perfect information problem in the space of probability measures such that, for $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} V_0(\mu) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^{\pi}\left[\sum_{k=0}^{T-1} \tilde{g}_k(\mu_k, U_k) + \tilde{g}_F(\mu_T)|\mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F(\mu_k, Y_{k+1}, U_k), \\ & U_k & = & \pi_k(\mu_k),\ \forall k = 0, .., T-1, \end{array} \tag{4.34}$$

with $\tilde{g}_k(\mu_k, u) = \langle \mu_k, c_k \rangle$ and $\tilde{g}_T(\mu_T) = \langle \mu_T, g_F \rangle$ where c_k depends on g_k and the conditional distribution of ξ_k knowing X_k. Its counterpart in the infinite horizon case reads:

$$\begin{array}{rrcl} V(\mu) = \min\limits_{\pi_0} & E^{\pi}\left[\sum_{k=0}^{+\infty} \alpha \tilde{g}(\mu_k, U_k)|\mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F(\mu_k, Y_{k+1}, U_k), \\ & U_k & = & \pi_0(\mu_k),\ \forall k \geq 0, \end{array} \tag{4.35}$$

with $\tilde{g}(\mu_k, u) = \langle \mu_k, c(\cdot, u) \rangle$ where c depends on g and the conditional distribution of ξ_k knowing X_k. With a slight abuse of notation, we also write the optimal value of the problems (4.24) and (4.25), respectively V_0 and V.

4.1.3.3 Dynamic Programming principle

The DP principle in this framework is also described in [Bertsekas and Shreve, 2004]. The cost-to-go can written as, for $\ell = 0, .., T-1$, and $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} V_\ell(\mu) = \min\limits_{\pi_\ell,\dots,\pi_{T-1}} & E^{\pi}\left[\sum_{k=\ell}^{T-1} \tilde{g}_k(\mu_k, U_k) + \tilde{g}_F(\mu_T)|\mu_\ell = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F(\mu_k, Y_{k+1}, U_k), \\ & U_k & = & \pi_k(\mu_k),\ \forall k = \ell, .., T-1, \end{array}, \tag{4.36}$$

$$V_T(\mu) = \tilde{g}_F(\mu). \tag{4.37}$$

Then the corresponding Bellman equation is for $\ell = 0, .., T-1$, and $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} V_\ell(\mu) = \min\limits_{u \in U} & E\left[\tilde{g}_\ell(\mu, u) + V_\ell(\mu_{\ell+1})|\mu_\ell = \mu\right] & & \\ \text{s.t.} & \mu_{\ell+1} & = & F(\mu_\ell, Y_{\ell+1}, u), \end{array}, \tag{4.38}$$

In the infinite horizon case, the Bellman equation is:

$$\begin{array}{rrcl} V(\mu) = \min\limits_{u \in U} & E\left[\tilde{g}(\mu, u) + \alpha V(\mu_{\ell+1})|\mu_\ell = \mu\right] & & \\ \text{s.t.} & \mu_{\ell+1} & = & F(\mu_\ell, Y_{\ell+1}, u), \end{array}, \tag{4.39}$$

The formulation (4.34) and (4.35) put to light two classical issues:

- First, the problems (4.34) and (4.35) are theoretically perfect information ones and μ_k is a function of I_k. However, in practice, for a general nonlinear or non Gaussian system (4.7), μ_k cannot be written in closed form and approximations are needed. Thus, to design coupled control and estimation method we also have to recall classical state estimation method. It can be stated as finding an function of I_k that approaches X_k. This problem is strongly related to the problem of approximating μ_k. These problems are the topic of Section 4.2.

- Secondly, the formulation (4.34) does not simplify the resolution of the problem (4.24) a priori as the new state space is $\mathcal{P}(\mathbb{R}^{n_x})$ which is only a metric space. However, if μ_k can be represented by a finite number of parameters then one can come back to classical finite dimensional DP with perfect information. Nevertheless, it is not true in general and optimal policies cannot be found exactly in most cases. Suboptimal policies have to be computed instead. This is where the latter formulations are useful because μ_k and the filtering equation are well known equations and many finite dimensional approximations of it exist. Consequently, this has notably allowed to build many suboptimal policies. There exist many ways to build suboptimal policies in a imperfect information setting. This is the topic of Section 4.3.

4.2 State Estimation, Nonlinear Filtering and Particle Filters

The goal of state estimation in this framework is to build an estimator, denoted by $\widehat{X}_k$, that approaches the true state of the system X_k in some sense, using only the available information I_k. Historically, this has been done in an optimisation context where one tries to minimise the variance of $\widehat{X}_k$. This leads to the well-known optimal estimation problem, see [Gelb, 1974] and [Anderson and Moore, 1979]. Actually, it is known that the expectation of X_k conditionally to I_k minimises the variance. This expectation can be computed in closed form only in the linear Gaussian case. As the linear and Gaussian assumptions are very limiting in practice, nonlinear filtering methods have appeared as extensions of linear Kalman filtering ones or not. In nonlinear filtering, the conditional expectation is not accessible explicitly so approximations have to be done. To do so, one usually tries to first approximate μ_k with, for example, Gaussian, sum of Gaussians or Monte-Carlo approximations. This approximation of μ_k allows then one to build a suboptimal estimator of X_k and assess a measure of the estimation error like a covariance matrix for instance. We focus our attention on particle filters as they are the only filters in the literature able to deal with the nonlinearity in TAN with a real map. Therefore, in the following, we recall the main features of linear and nonlinear stochastic filtering.

4.2.1 Optimal estimation

Classically, optimal estimation is concerned with finding an estimator of X_k as a function of I_k that minimises the conditional variance also named the conditional Mean Square Error (MSE). This optimisation problem reads:

$$\min_{\hat{x}\in R^{n_x}} E\left[\|\hat{x} - X_k\|^2 | I_k\right]. \tag{4.40}$$

By simple calculations, one gets that the almost surely optimal estimator w.r.t. the distribution of I_k is the expectation of X_k conditionally to I_k, denoted in the following by X_k^*:

$$\widehat{X}_k^* = E[X_k | I_k].$$

Notice that $\widehat{X}_k^*$ is also a solution of the problem of minimisation of the total MSE which reads:

$$\begin{array}{ll} \min\limits_{\pi^e} & E\left[\|\widehat{X}_k - X_k\|^2\right] \\ \text{s.t.} & \widehat{X}_k = \pi^e(I_k), \end{array} \tag{4.41}$$

where π^e: $(\mathbb{R}^{n_y} \times \mathbb{R}^{n_u})^k \times \mathbb{R}^{n_y} \longrightarrow \mathbb{R}^{n_x}$ is a measurable estimation law. To show this, consider any π^e and notice that $\pi^e(I_k)$ is admissible for the problem (4.40). Thus, by optimality of $\widehat{X}_k^*$:

$$E\left[\|\widehat{X}_k^* - X_k\|^2 | I_k\right] \leq E\left[\|\pi^e(I_k) - X_k\|^2 | I_k\right].$$

By integrating over I_k, one gets:

$$E\left[\|\widehat{X}_k^* - X_k\|^2\right] \leq E\left[\|\pi^e(I_k) - X_k\|^2\right].$$

By noticing that $\widehat{X}_k^*$ is measurable w.r.t. I_k, it is clear that $\widehat{X}_k^*$ is optimal in (4.41).

$\widehat{X}_k^*$ can also be written as the expectation of the distribution μ_k such that:

$$\widehat{X}_k^* = \langle \mu_k, Id \rangle,$$

where Id is the Identity map. Thus, $\widehat{X}_k^*$ is also optimal for the version of problem (4.41) where π^e is chosen as a function of μ_k. This problem reads:

$$\begin{array}{ll} \min\limits_{\pi^e} & E\left[\left\langle \mu_k, g^{mse}\left(\cdot, \widehat{X}_k\right)\right\rangle\right] \\ \text{s.t.} & \widehat{X}_k = \pi^e(\mu_k), \end{array} \tag{4.42}$$

where the expectation is taken over I_k and $g^{mse}(x, \hat{x}) = \|\hat{x} - x\|^2$.

Computing $\widehat{X}_k^*$ is a complicated issue a priori because it requires the knowledge of μ_k which is itself hard to compute. However, in the case of a linear dynamics and observation equation with Gaussian noise, the Kalman filter gives a recursive equation for $\widehat{X}_k^*$ and the conditional covariance matrix.

4.2.2 Linear Kalman filtering

In this section, we assume that the dynamics and the observation equation of the system have the following form:

$$X_{k+1} = FX_k + GU_k + \xi_k, \tag{4.43}$$
$$Y_k = HX_k + \eta_k, \tag{4.44}$$
$$X_0 \sim p_0,$$

where:

- p_0 is a Gaussian distribution on $\mathbb{R}^{n_x}$ with mean $\bar{X}_0$ and covariance matrix $P_0 \succ 0$.
- $F \in \mathbb{R}^{n_x \times n_x}$, $G \in \mathbb{R}^{n_x \times n_u}$ and $H \in \mathbb{R}^{n_y \times n_x}$.
- $(\xi_k)_{k \in \mathbb{N}}$ are Gaussian i.i.d. random variables valued in $\mathbb{R}^{n_\xi}$ with zero mean and covariance matrix $Q \succeq 0$.

- $(\eta_k)_{k\in\mathbb{N}}$ are Gaussian i.i.d. random variables and independent of $(\xi_k)_{k\in\mathbb{N}}$ valued in $\mathbb{R}^{n_\eta}$ with zero mean and covariance matrix $R \succ 0$.

It can be shown that under those assumptions, the conditional distributions μ_k and $\mu_{k|k-1}$ are Gaussian, see [Anderson and Moore, 1979] for the details. Then, they are characterised by their first and second moments that are denoted respectively by $\widehat{X}_k^*$ and P_k^* and by $\widehat{X}_{k|k-1}^*$ and $P_{k|k-1}^*$. More precisely, for $k \geq 1$:

$$\begin{aligned}
\widehat{X}_k^* &= E[X_k|I_k], \\
P_k^* &= E[(X_k - \widehat{X}_k^*)(X_k - \widehat{X}_k^*)^T|I_k], \\
\widehat{X}_{k|k-1}^* &= E[X_k|I_{k-1}, U_{k-1}], \\
P_{k|k-1}^* &= E[(X_k - \widehat{X}_{k|k-1}^*)(X_k - \widehat{X}_{k|k-1}^*)^T|I_{k-1}, U_{k-1}].
\end{aligned}$$

It happens that the previous quantities follow recursive equations known as the Kalman filter. These equation can be split into two steps coming from equations (4.28) and (4.29) called the prediction step and the correction step. During the prediction step, one propagates $\widehat{X}_{k-1}^*$ and P_{k-1} using the knowledge of U_{k-1} and the dynamics (4.43) to obtain $\widehat{X}_{k|k-1}^*$ and $P_{k|k-1}^*$. For simplicity, it is assumed that Y_0 is deterministic so that $\widehat{X}_0^* = \bar{X}_0$ and $P_0^* = P_0$. The prediction step reads:

$$\widehat{X}_{k|k-1}^* = F\widehat{X}_{k-1}^* + GU_{k-1}, \tag{4.45}$$

$$P_{k|k-1}^* = FP_{k-1}^*F^T + Q. \tag{4.46}$$

During the correction step, one incorporates the knowledge of Y_k and the observation equation (4.44) through a gain K_k called the Kalman gain such that:

$$S_k = HP_{k|k-1}^*H^T + R, \tag{4.47}$$

$$K_k = P_{k|k-1}^*H^TS_k^{-1}, \tag{4.48}$$

$$\widehat{X}_k^* = \widehat{X}_{k|k-1}^* + K_k(Y_k - H\widehat{X}_{k|k-1}^*), \tag{4.49}$$

$$P_k^* = (I - K_kH)P_{k|k-1}^*. \tag{4.50}$$

K_k can be seen from equations (4.46) and (4.48) as the optimal trade-off of confidence between the prediction from the dynamics (4.45) and the incoming observation Y_k. The optimality properties of the Kalman filter are studied in more detail in [Anderson and Moore, 1979].

It is also known and easily seen from equation (4.46) and (4.50), that P_k actually does not depend on I_k. It means that in this case, the conditional MSE, which is the trace of P_k^* coincides with the total MSE. It also means that the control has no impact on the MSE. That property is not true anymore for a general nonlinear system.

As said earlier, the Kalman filter describes exactly the conditional density. More precisely, it can be written as follows:

$$\mu_k = \mathcal{N}\left(\widehat{X}_k^*, P_k^*\right),$$

where $\mathcal{N}$ stands for the normal distribution. Unfortunately, for a non Gaussian nonlinear dynamics and observation equation, there is no exact representation of $\widehat{X}_k^*$ and P_k^* and approximations are required which is the topic of next section.

4.2.3 Nonlinear filtering

There exist countless suboptimal nonlinear filters in the literature hence we only mention here the most commonly used. The optimality properties of the linear Kalman filter are very appealing in practice. Therefore, a natural idea is to try to extend it to nonlinear systems. Such filters have had a lot of success until now but still cannot deal with certain very nonlinear systems. Following on that fact, alternative nonlinear filters based on Sequential Monte Carlo methods called particle filters have appeared. In the following, we present essentially these two types of filters with a focus on convergence results of particle filters.

4.2.3.1 Nonlinear Kalman filters

We assume that f and h have the following form:

$$f(x,u,\xi) = f^{det}(x,u) + \xi, \tag{4.51}$$
$$h(x,\eta) = h^{det}(x) + \eta, \tag{4.52}$$
$$p_\xi = \mathcal{N}(0,Q),$$
$$p_\eta = \mathcal{N}(0,R),$$

where f^{det} and h^{det} are differentiable.

The Extended Kalman filter

The most basic and most used nonlinear Kalman filter is the Extended Kalman filter (EKF). It is, in fact, a direct copy of the Linear Kalman filter where the matrix from the linear dynamics are replaced by the linearisation of the nonlinear one at the current estimate. More precisely, even if μ_k is not Gaussian in general, it is still assumed that it can be represented by its first and second moment, respectively called $\widehat{X}_k$ and P_k for simplicity. The EKF can then be defined as follows, for $k \geq 1$, first by its prediction step:

$$\widehat{X}_{k|k-1} = F_k \widehat{X}_{k-1} + G_k U_{k-1}, \tag{4.53}$$
$$P_{k|k-1} = F_k P_{k-1} F_k^T + Q. \tag{4.54}$$

and secondly by its correction step:

$$S_k = H_k P_{k|k-1} H_k^T + R, \tag{4.55}$$
$$K_k = P_{k|k-1} H_k^T S_k^{-1}, \tag{4.56}$$
$$\widehat{X}_k = \widehat{X}_{k|k-1} + K_k(Y_k - H_k \widehat{X}_{k|k-1}), \tag{4.57}$$
$$P_k = (I - K_k H_k) P_{k|k-1}, \tag{4.58}$$

where $F_k = \nabla_x f^{det}(\widehat{X}_k, U_k)$, $G_k = \nabla_u f^{det}(\widehat{X}_k, U_k)$, $H_k = \nabla h^{det}(\widehat{X}_{k|k-1})$, $Q \succeq 0$ and $R \succ 0$.

Notice that the matrices Q and R are data of the problem that are chosen by the user during the modelling phase. In the linear case, the Kalman filter is optimal whatever the value of the matrices Q and R is. It means, for instance, that if R is specified a priori as the mean sensor error, then it is optimal to use this precise value of R in the filter. However, in a nonlinear case, the EKF is not optimal so the matrices Q and R are not necessarily related to real quantities. In practice, Q and R can then be seen as

tunable parameters. The tuning of the distribution of the noise is actually a recurring issue in nonlinear filtering. In this case the approximation of μ_k, denoted by μ_k^{EKF} reads:

$$\mu_k^{EKF} = \mathcal{N}\left(\widehat{X}_k, P_k\right).$$

The conceptual simplicity of the EKF has one major drawback: it requires the linearisation of the original system. The main consequence of this is that the EKF diverges if the initial error is too large or if the system is too nonlinear. Besides, it also requires the computation of derivative which can be costly in practice.

To avoid this issue, more elaborated Kalman filters have been designed. Among them, we present the Unscented Kalman filter (UKF) [Julier and Uhlmann, 2004] and the Ensemble Kalman filter (EnKF) [Evensen, 2003] due to their clear resemblance to particle filters. The main idea is not to propagate the covariance matrix explicitly but to approach it with a sampling technique instead. For simplicity, we assume that the dynamics is of the form (4.51) and the observation equation as in (4.52) with h^{det} linear such that $h^{det}(x) = Hx$.

The Unscented Kalman filter

First, in the UKF, μ_k is also approximated by a Gaussian distribution such that $\mu_k^{UKF} = \mathcal{N}(\widehat{X}_k, P_k)$. The main difference with the EKF is that μ_k^{UKF} is sampled in a deterministic way using the *unscented transform* leading to $2n_x + 1$ particles denoted by $(x_k^i)_{i=0,..,2n_x}$. The unscented transform is usually defined as follows, for $k \geq 1$:

$$x_{k-1}^0 = \widehat{X}_{k-1}, \tag{4.59}$$

for $i = 1, .., n_x$:

$$x_{k-1}^i = \widehat{X}_{k-1} + \left(\sqrt{\frac{n_x}{1-\omega^0} P_{k-1}}\right)_i, \tag{4.60}$$

for $i = n_x + 1, .., 2n_x$:

$$x_{k-1}^i = \widehat{X}_{k-1} - \left(\sqrt{\frac{n_x}{1-\omega^0} P_{k-1}}\right)_{i-n_x}, \tag{4.61}$$

where $\left(\sqrt{n_x P_{k-1}}\right)_i$ is the ith column of the matrix square root of $n_x P_{k-1}$. The resulting particles are then propagated using the nominal nonlinear dynamics such that, for $i = 0, .., 2n_x$:

$$x_{k|k-1}^i = f^{det}(x_{k-1}^i, U_k), \tag{4.62}$$

The predicted mean and covariance are computed as follows:

$$\widehat{X}_{k|k-1} = \sum_{i=0}^{2n_x} \omega^i x_{k|k-1}^i, \tag{4.63}$$

$$P_{k|k-1} = \sum_{i=0}^{2n_x} \omega^i (x_{k|k-1}^i - \widehat{X}_{k|k-1})(x_{k|k-1}^i - \widehat{X}_{k|k-1})^T + Q, \tag{4.64}$$

where $-1 < \omega^0 < 1$ and $\omega^i = \frac{1-\omega^0}{2n_x}$ for $i = 1, ..2n_x$. The correction step is basically the same as in the original Kalman filter:

$$S_k = HP_{k|k-1}H^T + R, \tag{4.65}$$
$$K_k = P_{k|k-1}H^T S_k^{-1}, \tag{4.66}$$
$$\widehat{X}_k = \widehat{X}_{k|k-1} + K_k(Y_k - H\widehat{X}_{k|k-1}), \tag{4.67}$$
$$P_k = (I - K_k H_k)P_{k|k-1}. \tag{4.68}$$

The Ensemble Kalman filter

The Ensemble Kalman filter is based on a random sampling technique *à la Monte Carlo* rather than a deterministic one in the UKF. This requires to simulate several trajectories of the system with different realisations of the noise leading to a set of particles $X_k^N = (x_k^i)_{i=1,..,N}$ with $N \geq 1$. During the prediction step, the N particles are propagated according to the nonlinear dynamics, and the covariance is computed with an empirical unbiaised estimator. It reads, for $i = 1, .., N$ and $k \geq 1$:

$$x_{k|k-1}^i = f^{det}(x_{k-1}^i, U_{k-1}) + \xi_k^i, \tag{4.69}$$
$$m_{k|k-1} = \frac{1}{N}\sum_{i=1}^{N} x_{k|k-1}^i, \tag{4.70}$$
$$P_{k|k-1}^N = \frac{1}{N-1}\sum_{i=1}^{N}(x_{k|k-1}^i - m_{k|k-1})(x_{k|k-1}^i - m_{k|k-1})^T, \tag{4.71}$$

where $(\xi_k^i)_{i=1,..,N}$ are i.i.d. random variables of law p_ξ. The correction step is analogous to the one in the linear case and reads, for $i = 1, .., N$:

$$S_k^N = HP_{k|k-1}^N H^T + R, \tag{4.72}$$
$$K_k^N = P_{k|k-1}^N H^T (S_k^N)^{-1}, \tag{4.73}$$
$$x_k^i = x_{k|k-1}^i + K_k^N(Y_k + \eta_k^i - H\widehat{X}_{k|k-1}^*), \tag{4.74}$$

where $(\eta_k^i)_{i=1,..,N}$ are i.i.d. random variables of law p_η. The approximation of μ_k, denoted by μ_k^N is defined as follows:

$$\mu_k^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_k^i},$$

where δ_x is the Dirac measure at x. The main advantage of the UKF and the EnKF compared to the EKF is that the linearisation step is not required because the covariance matrix is computed in both cases using the particles. Consequently, both filters are more precise than the EKF for nonlinear systems that differ too much from their linear approximation. See [Evensen, 2003] for a review on EnKF and some extensions, and [Julier and Uhlmann, 2004] for a review of the UKF. Under assumptions of uniform observability, the mean square stability of the presented nonlinear filters have been proven. See [Reif et al., 1999] for the EKF, [Xiong et al., 2006] for the UKF and [Tong et al., 2016] the EnKF and [Karvonen, 2014] for a review.

However, this assumptions do not fit in the framework of a very nonlinear system especially with a complicated map h^{det}. This leads to the main drawback of UKF and EnKF which is that, at some point, μ_k or $\mu_{k|k-1}$ is assumed to be Gaussian. It is explicit in the UKF. In the EnKF, it is assumed during the correction step (4.73) to (4.74) that the predicted density $\mu_{k|k-1}$ is Gaussian. First, this may lead to poor performance if μ_k is multimodal. One needs to use more involved Kalman filters that uses Gaussian mixture, see [Dovera and Della Rossa, 2011] for example, to deal with multimodality. Secondly, even with those improvements, the approximations made cannot be controlled. For instance, in the case of the EnKF, it cannot be shown that μ_k^N tends to μ_k when $N \to +\infty$ for a non linear system. That is why, in the following, we focus on particle filters, which do not assume the Gaussianity of the conditional distribution.

4.2.3.2 Particle filtering

Definition

Algorithm 1 Particle filter (SIR) with adaptive resampling

1: Create a sample of N particles x_0^i according to the law μ_0 and initialize the weights ω_0^i with $\frac{1}{N}$.
2: **for** $k = 0, 1, 2 \dots$ **do**
3: **Prediction:**
4: Given a control u_k and a set of particles $\left(x_k^i\right)_{i=1,..,N}$ and a set of normalised weights $\left(\omega_k^i\right)_{i=1,..,N}$, compute the predicted particles by drawing samples from K:

$$x_{k+1|k}^i \sim K(\mathrm{d}x_{k+1|k}, x_k^i, u_k), \text{ for } i = 1, .., N,$$
$$\omega_{k+1|k}^i = \omega_k^i.$$

5: **Correction**:
6: Get the new observation Y_{k+1}.
7: Compute the unnormalised updated weights $\left(\tilde{\omega}_k^i\right)_{i=1,..,N}$ thanks to the likelihood function ρ:

$$\tilde{\omega}_{k+1}^i \propto \omega_{k+1|k}^i \rho(y_{k+1}, x_{k+1|k}^i).$$

8: **if** *Resampling criterion satisfied* **then**
9: Draw the *a posteriori* particles $\left(x_{k+1}^i\right)_{i=1,..,N}$ from the set $\left(x_{k+1|k}^i\right)_{i=1,..,N}$ and $\left(\tilde{\omega}_{k+1}^i\right)_{i=1,..,N}$ using a resampling technique and set $\omega_{k+1}^i = \frac{1}{N}$.
10: **else**
11: Set $x_{k+1}^i = x_{k+1|k}^i$ and $\omega_{k+1}^i = \frac{\tilde{\omega}_{k+1}^i}{\sum_{i=1}^N \tilde{\omega}_{k+1}^i}$.
12: **end if**
13: **end for**

A particle filter approximates the posterior distribution μ_k by a set of N particles, $\left(x_k^i\right)_{i=1,..,N}$ valued in $\mathbb{R}^{n_x}$, associated with nonnegative and normalized weights $\left(\omega_k^i\right)_{i=1,..,N}$. This approximation is denoted

by μ_k^N and reads:

$$\mu_k^N = \sum_{i=1}^{N} \omega_k^i \delta_{x_k^i}, \tag{4.75}$$

The predicted distribution, $\mu_{k|k-1}$, is also approximated, as an intermediate step, by the particles and weights: $\left(x_{k|k-1}^i\right)_{i=1,..,N}$ and $\left(\omega_{k|k-1}^i\right)_{i=1,..,N}$. This approximate distribution is denoted by $\mu_{k|k-1}^N$, and is defined as follows:

$$\mu_{k|k-1}^N = \sum_{i=1}^{N} \omega_{k|k-1}^i \delta_{x_{k|k-1}^i}.$$

As previously for Kalman filters, a particle filter is computed recursively following two steps: prediction and correction. During the prediction step, the particles are propagated using an importance distribution that is often chosen as the Markov kernel from the dynamics, K. During the correction step, the weights are updated thanks to Bayes' rule and the incoming observation Y_k. However, if one applies naively Bayes' rule at each step to update the weights, they tend to be concentrated on a few particles letting the majority of them with a very low weight which makes them useless. This fundamental flaw of particle filtering is called *sample degeneracy* [Li et al., 2014]. Usually, to remedy this issue, the predicted particles are resampled according to a resampling technique involving the updated weights. The principle of resampling is to remove the particles with low weight and keep mostly the particles with high weights. See [Li et al., 2015] for a review of existing resampling techniques. The resampling step tends to keep particles that are very similar and hopefully near the true state of the system. In other words, it concentrates them in some small areas in the state space. This phenomenon is known as *sample impoverishment*, see again [Li et al., 2014]. It may be problematic because too much confidence may be given to a potentially unreliable observation and the particles may get stuck on unlikely areas. In this case, the filter loses robustness because of the lack of diversity in the cloud of particles. There exist many techniques, surveyed in [Li et al., 2014], to deal with sample degeneracy and impoverishment. The most common one is not to do the resampling step at each time step. This allows one to make a trade off between degeneracy and impoverishment. Actually, criteria have been developed to decide whether or not resampling is needed. A classical particle filter algorithm with adaptive resampling is summed up in Algorithm 1.

In fact, sample degeneracy and impoverishment can be seen as causes of a more global problem in particle filtering which is the mismatch between the predicted approximate distribution $\mu_{k|k-1}^N$ and the likelihood at the new observation Y_k, $\rho(Y_k, \cdot)$. In practice, this mismatch is known to be the cause of divergence of the particle filter. Informally, it means that the information contained in the predicted $\mu_{k|k-1}^N$ is not coherent with the new information contained in Y_k. It can be characterised by the that the integral term $\langle \mu_{k|k-1}^N, \rho(Y_k, \cdot)\rangle$ is too small (see [Le Gland et al., 2004]). From a theoretical point of view, it has been shown in [Le Gland et al., 2004] that if $\langle \mu_{k|k-1}^N, \rho(Y_k, \cdot)\rangle$ is too small then some error bounds deteriorate with time and allow the divergence of the filter which confirms the practical considerations. The three main causes of mismatch can be summed up and analysed intuitively as follows:

- First, assume that the observations are too precise, for example that $\|R\|$ small in the Gaussian case. Then, the support of the likelihood is small. It implies that it is very common that only few particles have a high likelihood $\rho(Y_k, x_{k|k-1}^i)$. If additionally, those particles have a small predicted weight $\omega_{k|k-1}^i$ then $\langle \mu_{k|k-1}^N, \rho(Y_k, \cdot)\rangle$ is small indicating a mismatch, (see [Le Gland et al., 2004]).

- Secondly, if the sample is degenerate, then, even if the likelihood is widespread and the particles are well distributed in the state space, their mass is concentrated on few of them. Furthermore, it is likely that these high-weighed particles have a low likelihood at some time because of their intrinsic randomness. Therefore, sample degeneracy may cause mismatch.

- Finally, if the sample is impoverished, the situation is reversed compared to a degenerate one. The total mass of the sample may be equally spread on the particles but the support of the predicted approximation is small and the particles themselves are gathered in a small area in the state space. Therefore, sample impoverishment may also cause mismatch.

The mismatch between the true predicted distribution and the likelihood is also an issue but it is more complicated to analyse because one does not know μ_k. Nevertheless, this problem is also mentioned in [Hu et al., 2011] and [Oudjane, 2000].

Note that Algorithm 1 does not require K and ρ to have a specific structure which makes it appealing to deal with arbitrarily nonlinear system. The other main advantage of particle filtering is the zoo of theoretical convergence results and error bounds that can be proven between μ_k and μ_k^N.

Convergence results and error bounds in particle filtering

In this section, we review the main classes of convergence results in particle filtering. The first crucial step is the choice of the metric. We focus here on results of convergence and error bounds between $\langle \mu_k^N, \phi \rangle$ and $\langle \mu_k, \phi \rangle$ when $N \to +\infty$ for any k, for any scalar test function ϕ in a particular class. The main reason is that most of the popular metrics on probability measures can be written in this form. See [Bilingsley, 2013] for the definitions of the most classical metrics on probability measures. We review here three types of results: almost sure convergence, Central Limit Theorems, and L^p bounds.

In [Crisan and Doucet, 2002] and [Crisan, 2001], a classical particle filter is represented recursively similarly to equation (4.31) such that, for $k \geq 0$ and $N \geq 1$:

$$\mu_{k+1}^N = F^N \left(\mu_k^N, U_k, Y_{k+1} \right), \tag{4.76}$$

where $F^N : \mathcal{P}(\mathbb{R}^{n_x}) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \longrightarrow \mathcal{P}(\mathbb{R}^{n_x})$ gathers the prediction, correction and resampling steps. In [Crisan and Doucet, 2002], it is shown that under an assumption of continuity of F^N w.r.t. the weak convergence, $\langle \mu_k^N, \phi \rangle$ converges to $\langle \mu_k, \phi \rangle$ as $N \to +\infty$ almost surely w.r.t. the randomness of the particles, for any continuous bounded ϕ. Furthermore, using the Portmanteau theorem [Bilingsley, 2013], these results are useful to justify the computation of empirical probabilities. However, no rate of convergence is given in this case.

In [Del Moral and Guionnet, 1999] and [Del Moral and Jacod, 2002], the authors show a Central Limit Theorem (CLT) for a particle filter. The concept of the results is to show that, for any fixed i_k, the random variable $W_k^N = \sqrt{N} \langle \mu_k^N - \mu_k, \phi \rangle$ converges in distribution to a centred univariate Gaussian with variance σ_k^2 depending mainly on ρ and ϕ. [Del Moral and Guionnet, 1999] considers only bounded functions ϕ while [Del Moral and Jacod, 2002] considers classes of unbounded functions. More recent results can be found in [Whiteley, 2013]. The limitations of CLT is that, due the definition of the convergence in distribution, one can only deal with terms of the form $E[\psi(W_k^N) | I_k = i_k] = E\left[\psi\left(\sqrt{N} \langle \mu_k^N - \mu_k, \phi \rangle \right) | I_k = i_k \right]$ with ψ bounded continuous.

In fact, in Chapter 5, we will be interested in controlling the MSE generated by the empirical mean,

denoted by $\widehat{X}_k^N$ and defined by:

$$\widehat{X}_k^N = \sum_{i=1}^{N} \omega_k^i x_k^i = \langle \mu_k^N, Id \rangle,$$

as it is very commonly used as an approximation of $\widehat{X}_k^*$. Thus, we would like to choose ψ as a norm which is unbounded and ϕ as a coordinate function. By doing so, we would typically be able to control a term like $E\left[|\langle \mu_k, \phi \rangle - \langle \mu_k^N, \phi \rangle|^2 | I_k = i_k\right]$.

This leads us naturally to L^2 and more generally L^p-bounds. It is recalled in the survey paper [Crisan and Doucet, 2002] that, if $\rho(Y_k, \cdot)$ is bounded, then for any bounded measurable ϕ, for any $k \geq 0$, for any i_k, there exists $C_k > 0$ such that for $N \geq 1$:

$$E\left[|\langle \mu_k, \phi \rangle - \langle \mu_k^N, \phi \rangle|^2 | I_k = i_k\right] \leq \frac{C_k \|\phi\|_\infty^2}{N},$$

where $\|\phi\|_\infty$ denotes the L^∞-norm of ϕ. As said earlier one would like to take ϕ as a coordinate function which is unbounded so the previous classical bound cannot be applied. However, in [Hu et al., 2008] and [Hu et al., 2011], L^p bounds with ϕ potentially unbounded are proven under additional assumptions. These results will be used at some point in the sequel to study the optimality properties of $\widehat{X}_k^N$ w.r.t. the MSE.

To summarise, in this section, we have seen how one can compute an approximation of μ_k and an estimator $\widehat{X}_k$ of X_k as a function of I_k. As described at the end of Section 4.1.3, the next logical step is to determine the control variable U_k as function of I_k.

4.3 Design of suboptimal policies

In view of the fact that optimal policies are unreachable in practice, we recall the classical classification of suboptimal policies of the problems (4.34) and (4.35). Furthermore, we review the main optimisation-based design technique: Model Predictive Control.

4.3.1 Dual effect and classification of suboptimal policies

In the perfect information setting, we have already seen that there exist only two interesting classes of control policies: open-loop policies and feedback policies. Open-loop policies use only the knowledge of X_0 while feedback policies use the knowledge of $(X_0, \ldots, X_k)$. In particular, a feedback policy already uses the maximal amount of information on the system at each time so the control has no impact on the quality of the future available information. It is no longer true in an imperfect information setting where the control can actively look for better information. The control is then said to have the *dual effect property*. This essential property has first been put to light by Feldbaum in his seminal work [Feldbaum, 1960]. He states that the control must have two roles:

- It guides the system in a standard way, for instance, it can drive it to some target in the state space.
- It actively probes information to improve the quality of the future observations.

Control laws that exhibit dual effect are called dual controllers. In [Bar-Shalom and Tse, 1974], a classification of control policies with imperfect information is defined according to the quantity of information used and the level of anticipation of the future. These classes of policies are defined as follows:

- *Open Loop* (OL) policies. In this case, the current control, U_k depends only on the initial information I_0, the knowledge of dynamics (4.7) and of (p_ξ). The sequence is determined once for all at time $k = 0$ and never adapts itself to the available information. As said previously in Section 4.1.2, OL policies perform the worst but are the easiest to compute. OL policies in a stochastic framework are used in robust control and lead to robust optimisation problems [Ben-Tal et al., 2009]. An application in robust path planning using a scenario approach is described in [Blackmore et al., 2010].

- *Feedback* (F) policies. In this class, U_k depends on I_k, the dynamics (4.7), p_ξ, the observation equations (4.20) up to time k and of p_η. A F policy incorporates the current available information but never anticipates the fact that observations will be available at instants strictly greater than k. F policies are typically built using the *certainty equivalence* principle [Bar-Shalom and Tse, 1974], [Bertsekas, 2011]. The idea of certainty equivalence is to combine an estimation law π_k^e and an independently devised perfect information control law π_k^c that usually stabilises the system with full state information. The resulting control and estimator $(U_k, \widehat{X}_k)$ read:

$$\begin{aligned} \widehat{X}_k &= \pi_k^e(I_k), \\ U_k &= \pi^c(\widehat{X}_k). \end{aligned} \tag{4.77}$$

  It can be shown that, in the linear quadratic Gaussian case, optimal policies of the problem (4.24) are of the form (4.77) where π^c is linear and $\widehat{X}_k = \widehat{X}_k^* = E[X_k|I_k]$. Moreover, under classical Kalman observability conditions, this control law stabilises the output-feedback system. The fact that the controller and the estimator can be designed completely independently with no alteration of the global performance is called the *separation principle*. See [Hespanha, 2009] and [Åström, 2012] for a complete review of the separation principle in the linear deterministic and stochastic case. In the nonlinear deterministic case, it is known that stabilizability and detectability are not sufficient for output feedback stability in general, see [Andrieu and Praly, 2009]. However, in [Khalil and Praly, 2014], a separation principle for a class of deterministic nonlinear systems is given based on high gain observers and saturated control laws. A separation principle is proven for a class bilinear stochastic systems in [Mohler and Kolodziej, 1981] and [Tenno and Nõmm, 2011]. In [Barty et al., 2006], a characterisation of the class of control that do not influence the information is given. However, the separation principle does not hold in general and the control must interact with the future information.

  Policies of the form (4.77) occur very commonly in *Model Predictive Control* (MPC). MPC is the subject of a Section 4.3.2.

- *m-Measurement Feedback* (m-MF). In this class, U_k depends on I_k, the dynamics (4.7), p_ξ and, for $1 \leq m \leq T$, on the probability distribution of the future observations from time k to time $k+m$, denoted by $p_{Y_{k:k+m}}$. Similarly to F policies, m-MF policies can adapt themselves to the current situation and also anticipate new observations up to m instants after k. A classical way to design m-MF policies is to build an estimation law π_k^e and a control policy π_k^d in the following way:

$$\begin{aligned} \widehat{X}_k &= \pi_k^e(I_k), \\ U_k &= \pi_k^d(\widehat{X}_k; p_{Y_{k:k+m}}). \end{aligned} \tag{4.78}$$

From (4.78), it is clear that m-MF policies are dual controllers. Actually, such policies are often the results of a dual stochastic MPC scheme [Mesbah, 2017], where the dependency on $p_{Y_{k:k+m}}$ is hidden inside an optimisation problem.

- *Closed Loop* (CL) policies. This class is the extension of the m-MF class up to time T for T finite or infinite such that:

$$\begin{aligned} \widehat{X}_k &= \pi_k^e(I_k), \\ U_k &= \pi_k^d(\widehat{X}_k; p_{Y_{k:T}}). \end{aligned} \tag{4.79}$$

  Optimal policies obtained from Dynamic Programming belong to this class. Indeed, each policy obtained from the backward Bellman equation (4.38) or (4.39) minimises an instantaneous cost plus a cost-to-go including the statistics of all the future possible observations. In other words, optimal policies have the dual effect property. More precisely, it is called *implicit* dual effect as it entirely comes from the optimality of the policy and the information feedback structure and not from an explicit excitation.

The previous classification provides a theoretical guideline of how to build suboptimal control policies. In the following, we present the main optimisation-based technique to actually build suboptimal control policies in this framework: Model Predictive Control.

4.3.2 Model Predictive Control

Dynamic planning problems are generally modelled as finite horizon optimal control problems. Intuitively, it is clear that the main difficulty appears when the prediction horizon is long because it requires a lot of computational resources. Concurrently, stabilisation problems can be modelled by infinite time optimal control problems. In both cases, the optimal control problem is generally intractable in itself. A very popular idea is then to solve a short horizon optimal control problem, starting from the current state. Afterwards, one keeps only the first optimal control and applies it to the real system. Finally, these steps are repeated on the new state and on the next ones recursively. This method is called Model Predictive Control (MPC) essentially because it involves specifically a prediction of the future states, unlike classical closed-form feedback control laws. Due to the use of a moving time window, MPC is also called Receding Horizon Control.

Independently of optimisation considerations, MPC appears to be an efficient technique to develop nonlinear control laws that deals with input and state constraints, see [Mayne, 2014] for a general review on MPC. In the following, we review MPC and, in particular, Stochastic MPC (SMPC) in the optimisation framework described in Section 4.1.

4.3.2.1 Deterministic MPC

We recall the deterministic infinite horizon problem from Section 4.1, for $\alpha = 1$ and $x_0 \in \mathbb{R}^{n_x}$:

$$\begin{aligned} V(x_0) = \min_{u_0,\dots,u_k,\dots \in U} \quad & \textstyle\sum_{k=0}^{+\infty} g(x'_k, u_k) \\ \text{s.t.} \quad x'_{k+1} &= f^{det}(x'_k, u_k),\ \forall k \geq 0, \\ x'_0 &= x_0. \end{aligned} \tag{4.80}$$

As said previously, in MPC, one would like to solve a finite horizon version of this problem of the form (4.3). Actually, following the DP principle, the problem (4.82) can be reformulated as a finite horizon problem in the following way, for any $1 \leq T < +\infty$:

$$\begin{array}{rrcl} V(x_0) = \min\limits_{u_0,\dots,u_{T-1}\in U} & \sum_{k=0}^{T-1} g_k(x'_k,u_k) + g_F^*(x'_T) & & \\ \text{s.t.} & x'_{k+1} & = & f^{det}(x'_k,u_k),\ \forall k=0,..,T-1, \\ & x'_0 & = & x_0, \end{array} \tag{4.81}$$

where $g_F^*(x) = V(x)$. This formulation does not remove the difficulty of the problem as g_F^* is not known. One would prefer a formulation like problem (4.3):

$$\begin{array}{rrcl} V_0(x_0,T) = \min\limits_{u_0,\dots,u_{T-1}\in U} & \sum_{k=0}^{T-1} g_k(x'_k,u_k) + g_F(x'_T) & & \\ \text{s.t.} & x'_{k+1} & = & f^{det}(x'_k,u_k),\ \forall k=0,..,T-1, \\ & x'_0 & = & x_0, \end{array} \tag{4.82}$$

where g_F is properly chosen. Note that g_F can be seen as a tractable approximation of g_F^*. We omit state constraints in this formulation for simplicity. The idea of MPC is first to solve the problem (4.82) starting from the current state x_k. Note that problem (4.82) is a nonlinear program that can be solved by classical numerical optimisation techniques [Nocedal and Wright, 2006]. The optimal control sequence is denoted by $(u_0^*,\dots,u_T)$. Secondly, one keeps the first control $u_0^*(x_0)$ and discards the others. Finally, the control $u_0^*(x_0)$ is applied to real systems and the process is repeated on the next state. Thus, the resulting deterministic feedback law, denoted by π_{MPC} is defined as follows, for any $x \in \mathbb{R}^{n_x}$:

$$\pi_{MPC}(x) = u_0^*(x).$$

Notice that π_{MPC} is independent of time because the problem (4.82) is time homogeneous. The sequence of controls applied to the system is then $u_{MPC} = (\pi_{MPC}(x_0),\dots,\pi_{MPC}(x_k),\dots)$. The huge success of deterministic MPC in practice, see for example [Qin and Badgwell, 2003], has given rise to the theoretical issues of optimality and stability of the MPC policy u_{MPC}. It has been studied notably in [Grune and Rantzer, 2008] and [Grune, 2013]. It is proven that, under assumptions on the cost functions, g_k and g_F, and the dynamics (4.1), the optimal value of the MPC finite horizon problem $V_0(x_0,T)$ tends to the optimal value of the infinite horizon problem $V(x_0)$, when $T \to +\infty$. It is also shown that the closed-loop system converges to an equilibrium point. This is typically done by showing that $V_0(x_0,T)$ is a discrete-time Lyapunov function.

4.3.2.2 Stochastic MPC with perfect information

Stochastic MPC with perfect information is the version of MPC where the dynamics of the system is random as in (4.7) with full knowledge of the current state X_k at each time k. As in the deterministic case, we study the infinite horizon problem, for $\alpha = 1$ and $x \in \mathbb{R}^{n_x}$:

$$\begin{array}{rrcl} V(x) = \min\limits_{\pi_0} & E_x^{\pi}\left[\sum_{k=0}^{+\infty} g(X_k,U_k,\xi_k)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k,U_k,\xi_k), \\ & U_k & = & \pi_0(X_k),\ \forall k \geq 0, \\ & X_0 & = & x. \end{array} \tag{4.83}$$

As in the deterministic case, using similar arguments to the ones in the proof of the DP principle, one can show that the problem (4.83) can be reformulate as a finite horizon one:

$$\begin{array}{rrcl} V(x_0) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^{\pi}_{x_0}\left[\sum_{k=0}^{T-1} g_k(X_k,U_k,\xi_k) + g_F^*(X_T)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k,U_k,\xi_k), \\ & U_k & = & \pi_k(X_k),\ \forall k = 0,..,T-1, \\ & X_0 & = & x_0, \end{array} \tag{4.84}$$

where $g_F^*(x) = V(x)$. Again, g_F^* is not known in practice so another cost g_F is chosen which leads to the following problem:

$$\begin{array}{rrcl} V_0(x_0,T) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^{\pi}_{x_0}\left[\sum_{k=0}^{T-1} g_k(X_k,U_k,\xi_k) + g_F(X_T)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k,U_k,\xi_k), \\ & U_k & = & \pi_k(X_k),\ \forall k = 0,..,T-1, \\ & X_0 & = & x_0. \end{array} \tag{4.85}$$

The idea of Stochastic MPC is basically the same as in classical deterministic MPC. One wants to solve the problem (4.85), get the optimal control policy $(\pi_0^*,\dots,\pi_T^*)$ and keep only π_0^*. In principle, since π_0^* is already an control law, one only needs to solve the problem (4.85) once. One can define the ideal Stochastic MPC law, denoted by π_{MPC}^*, such that, for $x \in \mathbb{R}^{n_x}$:

$$\pi_{MPC}^*(x) = \pi_0^*(x).$$

The stability and performance of the ideal SMPC policy are studied in [Chatterjee and Lygeros, 2015].

Unfortunately, solving (4.85) is already a hard problem. Closed form solution exists only for a few models while numerical solutions require the discretisation of the state space and suffer from the curse of dimensionality. See [Bertsekas, 2005b] for a review of Approximate Dynamic Programming and its link to MPC, and [Dupacovà et al., 2000] for a review on methods based on scenario trees. We present, in the following, the two simplest approximations of problem (4.85) in the literature: the deterministic nominal approximation and the open-loop approximation. The deterministic nominal approximation consists in a deterministic version of problem (4.85) where the random excitation ξ_k is replaced by a deterministic statistics $\bar{\xi}$ such as the mean. It reads, for $x_0 \in \mathbb{R}^{n_x}$:

$$\begin{array}{rrcl} \bar{V}_0(x_0,T) = \min\limits_{u_0,\dots,u_{T-1}\in U} & \sum_{k=0}^{T-1} g_k(x'_k,u_k,\bar{\xi}) + g_F(x'_T) & & \\ \text{s.t.} & x'_{k+1} & = & f(x'_k,u_k,\bar{\xi}),\ \forall k = 0,..,T-1, \\ & x'_0 & = & x_0. \end{array} \tag{4.86}$$

The open-loop approximation is simply a copy of problem (4.85) where the control policies are restrained to deterministic sequences. It reads, for $x_0 \in \mathbb{R}^{n_x}$:

$$\begin{array}{rrcl} V_0^{OL}(x_0,T) = \min\limits_{u_0,\dots,u_{T-1}\in U} & E^{\pi}_{x_0}\left[\sum_{k=0}^{T-1} g_k(X_k,u_k,\xi_k) + g_F(X_T)\right] & & \\ \text{s.t.} & X_{k+1} & = & f(X_k,u_k,\xi_k), \\ & X_0 & = & x_0. \end{array} \tag{4.87}$$

The problem (4.86) is purely deterministic so it is relatively simple to solve. However, it only takes the randomness of the system into account very slightly so it is not robust in general. On the contrary, the problem (4.87) is a stochastic program and typically a robust optimal control problem seen from the probabilistic perspective. Usually, soft state constraints in probability are added but we omit them for simplicity. This kind of problem can be handled by numerical stochastic optimisation techniques. See

[Fu, 2015] for a review on such techniques and [Calafiore and Fagiano, 2013] for an application of the scenario approach to MPC.

Both the problems (4.86) and (4.87) lead to OL policies if solved once. However, if they are solved in a receding horizon way, they lead to the following control law, for $x \in \mathbb{R}^{n_x}$:

$$\pi_{MPC}(x) = u_0^*(x),$$

where $u_0^*(x)$ is the first optimal control in (4.86) or (4.87). See [Mesbah, 2016] for a review of the main applications of SMPC.

As discussed in Section 4.3.1, in the perfect information case, there exist only two classes of policies, Feedback policies and Open Loop policies. In the imperfect information case, there exist more of them which enables more types of MPC schemes.

4.3.2.3 Stochastic MPC with imperfect information

Similarly to the perfect information, SMPC with imperfect information is concerned with solving approximately the infinite horizon problem:

$$\begin{array}{rrcl} V(\mu) = \min\limits_{\pi_0} & E^{\pi}\left[\sum_{k=0}^{+\infty} \tilde{g}(\mu_k, U_k)|\mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, U_k\right), \\ & U_k & = & \pi_0(\mu_k),\ \forall k \geq 0. \end{array} \tag{4.88}$$

It can also be reformulated as a finite horizon problem, for any $1 \leq T < +\infty$, and any $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} V(\mu) = \min\limits_{\pi_0,\ldots,\pi_{T-1}} & E^{\pi}\left[\sum_{k=0}^{T-1} \tilde{g}_k(\mu_k, U_k) + \tilde{g}_F^*(\mu_T)|\mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, U_k\right), \\ & U_k & = & \pi_k(\mu_k),\ \forall k = 0, .., T-1, \end{array} \tag{4.89}$$

where $\tilde{g}_F^* = V$.

As previously, $\tilde{g}_F^*$ is unknown and replaced by a tractable final cost $\tilde{g}_F$. The finite horizon problem to solve is then, for any $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} V_0(\mu, T) = \min\limits_{\pi_0,\ldots,\pi_{T-1}} & E^{\pi}\left[\sum_{k=0}^{T-1} \tilde{g}_k(\mu_k, U_k) + \tilde{g}_F(\mu_T)|\mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, U_k\right), \\ & U_k & = & \pi_k(\mu_k),\ \forall k = 0, .., T-1. \end{array} \tag{4.90}$$

As in the perfect information case, the ideal imperfect information MPC scheme requires to solve the problem (4.90) exactly and keep the first element of the optimal policy, π_0^*. The MPC control law reads, for any $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\pi_{MPC}^*(\mu) = \pi_0^*(\mu).$$

The output-feedback stability and performance of this scheme have been studied in [Sehr and Bitmead, 2017a] and [Sehr and Bitmead, 2017b] where it is applied to a finite state space case. Actually, building alternative tractable MPC policies is harder than in the perfect information case, especially because of the need of dual effect. In the following, we present the general form of the main type of tractable imperfect information SMPC schemes.

Open-Loop Feedback SMPC

Classically in the literature and in many practical applications, the link between the control and the future available information is ignored in the control design, invoking the separation principle. An example is presented in [Bertsekas, 2005a]. According to the classification from Section 4.3.1, this corresponds to F policies, also called Open-Loop Feedback (OLF) policies in the context of MPC. As described by equation (4.77), F policies are usually defined through a certainty equivalence principle using a perfect information controller computed at a state estimator. In the MPC framework, the control and estimator read:

$$\begin{aligned} \widehat{X}_k &= \pi_k^e(I_k), \\ U_k &= \pi_{MPC}(\widehat{X}_k). \end{aligned} \tag{4.91}$$

where π_{MPC} is a SMPC controller as described in Section 4.3.2.2. Examples of such output feedback controllers are gathered in [Mesbah, 2016]. In [Hokayem et al., 2012], [Mishra et al., 2017] and [Homer and Mhaskar, 2017], three variants of the previous scheme are proposed for a class of linear Gaussian systems with bounded controls and a class of continuous-time stochastic nonlinear system with a nonlinear observer. Additionally, stochastic output feedback stability is proven but a separation principle is assumed in both cases meaning that the controller is supposed to have no impact on the observer convergence.

Those controllers are very common in industrial applications because they are easy to implement, as long as one is able to compute the MPC law. However, in this form, the controller does not take into account the current uncertainty on the estimator which reduces its robustness. An alternative idea is to use the framework of problem (4.88) but, instead of propagating the full filter μ_k, one would only propagate the predicted distribution $\mu_{k+\ell|k}$ from equation (4.30). Roughly speaking, the idea is to use fully the knowledge of μ_k and use only the dynamics for the prediction of future states. The resulting optimal control problem reads for any $k \geq 1$:

$$\begin{aligned} V_{OLF}(\mu_k, T) = \min_{u_0,\dots,u_{T-1}\in U} \quad & E^{\pi}\left[\sum_{\ell=0}^{T-1} \tilde{g}_\ell(\tilde{\mu}_\ell, u_\ell) + \tilde{g}_F(\tilde{\mu}_T)|\tilde{\mu}_0 = \mu_k\right] \\ \text{s.t.} \quad & \tilde{\mu}_{\ell+1|0} = G\left(\tilde{\mu}_{\ell|0}, u_\ell\right) \ \forall \ell = 0, .., T-1. \end{aligned} \tag{4.92}$$

One aims a practical resolution in problem (4.92) thus the controls are open loop. In problem (4.92), we emphasise the fact that the controller starts specifically at μ_k to show that the whole distribution μ_k is used and not only an estimator $\widehat{X}_k$ like in (4.91). The resulting MPC scheme reads:

$$\pi_{OLF}(\mu_k) = u_0^*(\mu_k),$$

where $u_0^*(\mu_k)$ is the first optimal control in (4.92). An OLF MPC scheme combining a particle filter to approximate μ_k and a Monte Carlo sampling technique to approach the problem (4.92) is presented in [Sehr and Bitmead, 2016]

Regrettably, OLF MPC controllers are only passively learning. This means that it is completely fortuitous that the uncertainty on the state is reduced after using a OLF MPC controller. It may happen in practice but it is not designed for. Conversely, it may also increase the uncertainty on the state and cause poor output-feedback performance in the end. The last remark drives us naturally to consider dual MPC controller which are actively learning.

Dual Stochastic MPC

Dual Stochastic MPC controllers indicate MPC schemes that belong to the m-MF or the CL class. There are two main class of dual controllers: implicit dual controllers and explicit dual controllers. To understand the concepts behind these two types of controller, one needs to go back to the problem (4.88). As it was already briefly mentioned in Section 4.3.1, optimal policies from the problem (4.88) are CL policies and show a property of dual effect that is called *implicit dual effect*. In fact, to find an optimal policy, one needs to propagate the future information through equation (4.31) By the Dynamic Programming principle, any optimal policy drives the future filtering distribution toward a minimum of the sum of the instantaneous cost and the next cost-to-go. In this sense, an optimal policy is dual. It is called implicitly dual because the dual effect only comes from considerations that are internal to the original optimisation problem. More precisely, it comes from the feedback structure on I_k or μ_k of the problem (4.88). This dual effect is, by definition of optimality, improving control performance compared to other policies. However, the fact that the trajectory of filtering distributions $(\mu_k)_{k=0,..,T}$ is optimal does not necessarily mean that state estimation is successful. This implication is not clear for a general system with a general cost function, but it is true in many applications and notably in the case study of [Sehr and Bitmead, 2017a]. That is why, in the literature, it is often tacitly assumed that going toward the optimality of μ_k is beneficial for state estimation in some sense.

Implicit dual controllers can then be defined as suboptimal dual controllers designed to maintain the implicit dual effect coming from optimality in (4.88). To do so, the first idea is to solve approximately the Bellman equation (4.38). See [Bayard and Schumitzky, 2008] for an example of an implicit dual controller based on policy iteration and particle filtering. In [Weissel et al., 2008], approximate DP is achieved thanks to a branch-and-bound algorithm and a new nonlinear filter. The second main idea is to propagate the filtering distribution using scenario trees. It is a complicated problem because in the imperfect information case, the structure of the tree must depend on the control to reproduce accurately the feedback constraint. In [Subramanian et al., 2014] and [Subramanian et al., 2015], μ_k is approximated respectively by a EKF and a UKF. These filters are used as nodes of a scenario tree to anticipate the future information inside an optimisation problem. The issue of the dynamic tree structure is handled by assuming that the innovation process has a fixed distribution though it is not mentioned explicity. In [Hanssen and Foss, 2015], the future filtering distributions are simulated by a tree of EnKFs inside the optimisation problem. However, the issue of the dynamic tree structure is not discussed. The main flaw of the previous techniques is their numerical cost. In [Bayard and Schumitzky, 2008] and [Weissel et al., 2008], only finite control spaces can be treated. In [Subramanian et al., 2016] and [Hanssen and Foss, 2015], the size of the resulting nonlinear program grows very rapidly with the time horizon.

The computational burden of implicit dual controllers comes mostly from the approximate propagation of μ_k in the optimisation problem. To avoid this expensive step, an idea is to propagate only the predicted distribution like in problem (4.92). The feedback structure is then removed and so is the implicit dual effect. The new concept for making the controller actively learn is to add an explicit excitation in the original problem. It can take the form of a constraint of persistence of excitation as in [Marafioti et al., 2014]. However, the resulting controller is only adapted to slightly nonlinear systems. In [Hovd and Bitmead, 2004], the covariance matrix of an EKF is minimised online inside the optimisation problem. It applies only on slightly nonlinear systems too. Besides, one reason to do explicit dual control is precisely to avoid simulating the uncertainty on the state inside the optimisation problem.

That is why, we focus on the literature where the change in the original problem takes the form of a

new constraint or a new cost, depending on the Fisher Information Matrix (FIM). Its inverse is a lower bound of the estimation error covariance matrix of any unbiased estimator and is denoted by J_k in the sequel. In particular, as the conditional expectation $\widehat{X}_k^*$ is an unbiased estimator of X_k, the following inequality holds, for $k \geq 0$:

$$J_k^{-1} \preceq E\left[(\widehat{X}_k^* - X_k)(\widehat{X}_k^* - X_k)^T\right]. \tag{4.93}$$

It is an a priori quantitative measure of the available information at each time that is commonly used in Optimal Design [Fedorov and Hackl, 2012], because it does not depend on the estimator used. In [Tichavsky et al., 1998], a recursive way to compute the FIM, denoted by J_k, for a system of the form (4.51)-(4.52) is presented. It reads:

$$\begin{aligned} J_0 &= P_0^{-1}, \\ J_{k-1}^+ &= J_{k-1} + E\left[F_{k-1}^T Q^{-1} F_{k-1}\right], \\ J_k &= E\left[H_k R^{-1} H_k^T\right] \\ &\quad - Q^{-1} E\left[F_{k-1}\right]\left(J_{k-1}^+\right)^{-1} E\left[F_{k-1}\right]^T Q^{-1}, \end{aligned} \tag{4.94}$$

where P_0 is the covariance matrix associated with the initial distribution p_0, $F_{k-1} = \nabla_x f^{det}(X_{k-1}, U_{k-1})$ and $H_k = \nabla h^{det}(X_k)$. In [Telen et al., 2017], a controller with a constraint on J_k is proposed. Usually, such controllers lead to infeasibilty problems that are also hard to anticipate in nonlinear cases. In [Telen et al., 2017], this infeasibility issue is addressed but only in the deterministic framework. Adding a scalar measure of information in the cost to create an explicit excitation is called integrated experiment design. In integrated experiment design, one minimises a new cost denoted by g_k^{ex} that realises a trade off between the original costs g_k and g_F and a measure of information denoted by g_k^{info} and g_F^{info}, depending on J_k in practice, such that, for $k = 0, .., T-1$:

$$\begin{aligned} g_k^{ex} &= g_k + g_k^{info}, \\ g_F^{ex} &= g_F + g_F^{info}. \end{aligned} \tag{4.95}$$

After considering that J_k can be added to the state of the system, on can see g_k^{info} as function from $\mathbb{R}^{n_x}$ to $\mathbb{R}^+$. The resulting optimisation problem can be written as follows:

$$\begin{aligned} V_{EX}(\mu_k, T) = \min_{u_0,\ldots,u_{T-1}\in U} \quad & E^\pi\left[\textstyle\sum_{\ell=0}^{T-1} \tilde{g}_\ell^{ex}(\tilde{\mu}_{\ell|0}, u_\ell) + \tilde{g}_F^{ex}(\tilde{\mu}_{T|0})|\tilde{\mu}_0 = \mu_k\right] \\ \text{s.t.} \quad & \tilde{\mu}_{\ell+1|0} = G\left(\tilde{\mu}_{\ell|0}, u_\ell\right) \ \forall \ell = 0, .., T-1, \end{aligned} \tag{4.96}$$

where $\tilde{g}_\ell^{ex} = \tilde{g}_\ell + \tilde{g}_\ell^{info}$ and $\tilde{g}_\ell^{info} = \langle \mu_\ell, g_\ell^{info} \rangle$. The MPC law in this case reads:

$$\pi_{EX}(\mu_k) = u_0^*(\mu_k),$$

where $u_0^*(\mu_k)$ is the first optimal control in (4.96). In [Telen et al., 2017] and [La et al., 2017], two integrated experimental controllers based on the FIM or a variant are proposed. However, both of them are incompatible with particle filtering. See [Mesbah, 2017], for a complete review of all the different flavors of explicit dual MPC.

Contrary to their implicit counterpart, explicit MPC controllers do not try to reach closed-loop optimality directly. They rather try to increase the quantity of available information and improve state

estimation. There is then a tacit assumption that improving state estimation also improves closed-loop control performance. As in the implicit case, it is not clear for a general system but confirmed in many applications included TAN.

To understand the fundamental differences between implicit and explicit dual MPC, we present another way of analysing the classification from Section 4.3.1 through the prism of caution, also called conservativeness. In fact, in the imperfect information setting, pure OL policies are typically either not robust, and fail in guiding the system, or overly robust, and lead to extremely cautious controls also named small controls. Cautious controls are usually very costly in the end, because, for instance, the system is too slow or its trajectory is too long. The extreme caution of OL policies is generally due to their overestimation of the risk of poor future performance which is itself caused by the lack of information. In others words, robust OL policies are too safe. A very general example of method to build a finite horizon OL-policies is to solve problem (4.92) starting from μ_0 and apply the whole optimal sequence. In this case, the overestimation of the risk comes from the fact that only $\mu_{\ell|0}$ is used to compute state uncertainty. Intuitively, it is known that the prediction step represented by (4.28) tends to spread $\mu_{\ell|0}$ such that the uncertainty on the state grows. One can see this phenomenon from (4.94). Clearly, the higher the covariance matrix Q is, the less information is accumulated.

On the contrary, F-policies are, by nature, less cautious than OL policies because of the feedback structure on the current information. OLF policies presented earlier use the knowledge of the current level of uncertainty on the state, represented by μ_k. They also use a prediction of the state through the dynamics, represented by $\mu_{k+\ell|k}$. It means that the propagation of the uncertainty is the same in OLF policies that in OL policies. The major difference comes from the fact that in OLF the predicted distribution is propagated from μ_k. Thus, at each time k, OL policies use the distribution $\mu_{k+\ell|0}$ whereas OLF policies uses $\mu_{k+\ell|k}$. Consequently, OLF policies are better than OL-policies in assessing the risk of poor performance and end up being less cautious.

However, OLF policies still can be too conservative or simply fail and cause the divergence of the system. We have already mentioned the risk of divergence caused by OLF polices in the dedicated section. Output-feedback divergence typically appears when the controller makes the quantity of available information diminish. See [Anderson, 1985] for an example in adaptive control. Explicit dual MPC seems well suited to fight this issue by construction. Nevertheless, it is clear from (4.96) that the propagation of uncertainty in explicit dual MPC is as in OLF. In particular, an explicit dual controller does not take into account the fact that more observations are available in the future unlike in implicit ones. It focuses on keeping enough information. It is a very important point because, in some applications, improving state estimation is irrelevant or even impossible but anticipate the future information is crucial.

For instance, in [Weissel et al., 2008], a 2D robot path-planning sheme with obstacles in an uncertain environment and with imperfect information is proposed. The target area and the obstacles are handled by a penalisation term in the cost. Figure 4.1 is taken from [Weissel et al., 2008] and represents trajectories coming from 3 types of MPC schemes. The green curve corresponds to a deterministic MPC, the red curve to an OLF MPC scheme and the blue curve to a implicit dual one Roughly speaking, two paths are considered: the first one is safe and long (in red), the second one is shorter but more risky (in blue) because it requires to pass in a tiny space between two obstacles. It is assumed that the sensor are sufficiently precise for the robot to able to follow the risky path if it uses the measurements. It is also assumed that if it does not use any measurements it is too hazardous for it to follow the risky path. As expected, it is shown in simulation that, if one uses the OLF controller, the robot always follows the safe path. Still, with the implicit dual controller, the robot follows the risky path. The reason for this is that the OLF ignores the fact that the future uncertainty will be small enough for the robot to follow the second path

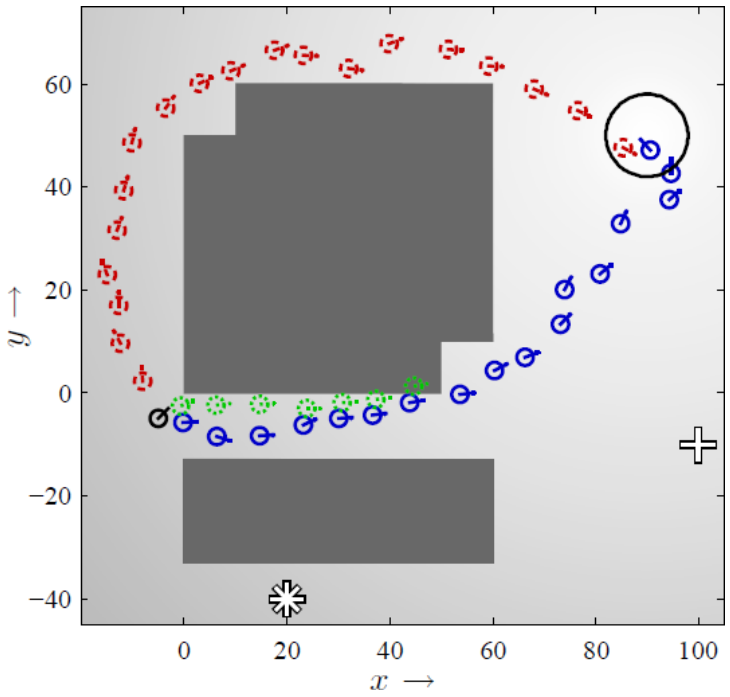


Figure 4.1: Example of the use of 3 types of MPC schemes on a wheeled-robot in a 2D environment with obstacles.

with an acceptable risk. On the contrary, their implicit dual policy do anticipate the future observations and is aware of their precision. It is not a surprise that the implicit dual controller performs better than the OLF policy. Our main claim is that an explicit dual controller defined by problem (4.96) is of no use in this example. The most important task in this example is not to improve the future information using some metric but to assess dynamically the future state uncertainty. About the last point, OLF policies and explicit dual ones behave in the same way because they both propagate only $\mu_{k+\ell|k}$ and not $\mu_{k+\ell}$. Consequently, an explicit dual MPC scheme is expected to produce the same result as an OLF policy in the example of the article.

This illustrates the main difference between the explicit and the implicit dual effect in terms of uncertainty propagation. To sum up, implicit dual effect is oriented toward improving directly the closed-loop control performance through information propagation. The potential contraction of the state uncertainty is only a bonus. On the contrary, explicit dual effect is focused on increasing the quantity of future information and improving state estimation, as a way of improving the closed-loop control performance.

It is clear that implicit dual policies have solid theoretical foundations because there come from an approximation of the Bellman equation. However, explicit dual policies are harder to justify formally because they require a change in the original problem that is empirical. It is true, in particular, in integrated experimental design where the cost is generally modified to match practical considerations only. In the following, we give a justification of the additional term in the cost using optimal estimation arguments.

Besides, in TAN with a real ground map, it is crucial to go over rough terrain to gain information. As a result, explicit dual controllers seem well suited to TAN. This also justifies our focus on explicit dual controllers in the following. Finally, it also appears in the literature that there are no explicit dual controllers able to use directly a particle filter as a starting point in (4.96). In the sequel, we provide two explicit dual MPC scheme applied to TAN that are able two use the full potential of a particle filter.

Chapter 5

Modelling of joint problem of optimal control and estimation

In this section, we would like to provide a new vision on the problem of estimation and control of a nonlinear stochastic system based on the definition of a multistage stochastic optimisation problem that gathers optimal control and optimal estimation. Through this, we give a justification of the combined use of particle filtering and Explicit Dual SMPC in coupled control and estimation methods. The first motivation for this is that particle filters are not very used in stochastic control as we have seen in the Section 4.3.2 dedicated to SMPC. It probably comes from the fact that particle filters are very recent compared to Kalman filters and not very well known in the control community. Besides, particle filter are complex and not necessarily easy to incorporate in a feedback. However, it surely has the potential to solve very nonlinear practical problem. In fact, our first justification of the use of particle filtering is to be able to deal with TAN with real ground maps. In this section, we give a more theoretical justification oriented toward optimal estimation. Secondly, as we have seen in the Section 4.3.2 that Explicit Dual SMPC is, by essence, hard to justify rigorously as it considers an optimal control problem that is modified empirically. However, we give a more theoretically grounded justification in this section also coming from optimal estimation.

5.1 Definition of the coupling optimisation problem

5.1.1 Setup of the problem of optimal control and estimation

In this section we consider the system (4.7) with the observation equation (4.20) and focus on the infinite horizon case. The idea is to add an estimator as a variable in (4.35) that is also looked for as a function of the available information. To do so we consider an augmented control $W_k = (U_k, \widehat{X}_k)$ and the corresponding augmented information vector $\tilde{I}_k$ defined recursively as follows:

$$\begin{aligned} \tilde{I}_0 &= Y_0, \\ \tilde{I}_{k+1} &= (\tilde{I}_k, V_k, Y_{k+1}). \end{aligned} \tag{5.1}$$

W_k is chosen as a function of $\tilde{I}_k$ such that for $k \geq 0$:

$$\begin{aligned} W_k &= (U_k, \widehat{X}_k), \\ U_k &= \pi_0^c(\tilde{I}_k), \\ \widehat{X}_k &= \pi_0^e(\tilde{I}_k), \\ \pi^{aug} &= (\pi_0^c, \pi_0^e). \end{aligned}$$

We also define the new augmented cost function g^{aug} and dynamics f^{aug} in the following way, for $x \in \mathbb{R}^{n_x}$, $v = (u, \hat{x}) \in \mathcal{U} \times \mathbb{R}^{n_x}$, $\xi \in \mathbb{R}^{n_\xi}$:

$$g^{aug}(x, v, \xi) = g^c(x, u, \xi) + g^e(x, \hat{x}), \tag{5.2}$$
$$f^{aug}(x, v, \xi) = f(x, u, \xi), \tag{5.3}$$

where

- g^c: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \longrightarrow \mathbb{R}^+$ is the cost function of a classical stochastic optimal control,
- g^e: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \longrightarrow \mathbb{R}^+$ is a measure of the estimation error. Generally, g^e is taken as the square error, i.e. $g^e(x, \hat{x}) = \|\hat{x} - x\|^2$ where $\|\cdot\|$ stands for the Euclidean norm on $\mathbb{R}^{n_x}$. However, the L_1-norm is very popular in machine learning and statistics [Vidaurre et al., 2013] and could be considered as an alternative.

Note that the augmented dynamics f^{aug} does not depend on $\hat{x}$ The main assumption in (5.2) is that the total cost g^{aug} can be separated as a sum of a control-oriented term and an estimation-oriented term. It is a mild assumption as one can consider that the true underlying problem would be a bi-objective one with g^c and g^e being the two cost functions. Intuitively, g^c and g^e are often anti-correlated because one often needs to trade some control performance for a better estimation.

With this in mind, (5.2) can be seen as a trade-off coming from the conversion of a biobjective problem into a mono objective one. The last remarks lead to the following optimisation problem for $\tilde{i}_0 \in \mathbb{R}^{n_y}$:

$$\begin{aligned} \widetilde{V}(\tilde{i}_0) = \min_{\pi_0^c, \pi_0^e} \quad & E_{p_0}^{\pi}\left[\textstyle\sum_{k=0}^{+\infty} \alpha^k g^{aug}(X_k, W_k, \xi_k) | \tilde{I}_0 = \tilde{i}_0\right] \\ \text{s.t.} \quad X_{k+1} &= f^{aug}(X_k, V_k, \xi_k), \\ Y_k &= h(X_k, \eta_k), \\ \tilde{I}_{k+1} &= (\tilde{I}_k, V_k, Y_{k+1}), \\ W_k &= (\pi_0^c(\tilde{I}_k), \pi_0^e(\tilde{I}_k)), \ \forall k \geq 0. \end{aligned} \tag{5.4}$$

If we write the problem (5.4) in terms of U_k and $\widehat{X}_k$ then we get:

$$\begin{aligned} \widetilde{V}(i_0) = \min_{\pi_0^c, \pi_0^e} \quad & E_{p_0}^{\pi}\left[\textstyle\sum_{k=0}^{+\infty} \alpha^k (g^c(X_k, U_k, \xi_k) + g^e(X_k, \widehat{X}_k)) | \tilde{I}_0 = \tilde{i}_0\right] \\ \text{s.t.} \quad X_{k+1} &= f(X_k, U_k, \xi_k), \\ Y_k &= h(X_k, \eta_k), \\ \tilde{I}_{k+1} &= (\tilde{I}_k, U_k, \widehat{X}_k, Y_{k+1}), \\ U_k &= \pi_0^c(\tilde{I}_k), \\ \widehat{X}_k &= \pi_0^e(\tilde{I}_k), \ \forall k \geq 0. \end{aligned} \tag{5.5}$$

It is clear from the formulation (5.5) that the problem (5.4) combines a general problem of stochastic optimal control with imperfect information with a general problem of optimal estimation. The study of this

formulation has been started in [Flayac et al., 2017]. It is actually inspired from [Copp and Hespanha, 2014] in which a similar gathering is done in a min-max optimisation framework. We would like to write the DP principle for the problem (5.4) to make explicit connections with classical optimal control and optimal estimation. Note that the fact that the dynamics represented by f^{aug} actually does not depend on $\hat{X}_k$ will be crucial when we write the Bellman equation for this problem.

As we did for the problem (4.25), we can rewrite the problem (5.4) in terms of the conditional distribution of X_k knowing $\tilde{I}_k$, denoted $\tilde{\mu}_k$. As for μ_k, one can derive the dynamics of $\tilde{\mu}_k$ from equation (4.28) and (4.29) using equation (5.3) and (4.20):

$$\tilde{\mu}_{k+1} = F^{aug}\left(\tilde{\mu}_k, Y_{k+1}, W_k\right), \tag{5.6}$$

with $F^{aug} : \mathcal{P}(\mathbb{R}^{n_x}) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \longrightarrow \mathcal{P}(\mathbb{R}^{n_x})$. This leads to the following reformulation:

$$\begin{array}{rrcl} \widetilde{V}(\mu) = \min\limits_{\pi_0^{aug}} & E^{\pi}\left[\sum_{k=0}^{+\infty} \alpha^k \tilde{g}^{aug}(\tilde{\mu}_k, W_k) | \tilde{\mu}_0 = \mu\right] \\ \text{s.t.} & \tilde{\mu}_{k+1} & = & F^{aug}\left(\tilde{\mu}_k, Y_{k+1}, W_k\right), \\ & W_k & = & \pi_0^{aug}(\tilde{\mu}_k),\ \forall k \geq 0, \end{array} \tag{5.7}$$

where, for $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$ and $v = (u, \hat{x}) \in \mathcal{U} \times \mathbb{R}^{n_x}$:

$$\begin{aligned} \tilde{g}^{aug}(\mu, v) &= \tilde{g}_c(\mu, u) + \tilde{g}_e(\mu, \hat{x}), \\ \tilde{g}_c(\mu, u) &= \langle \mu, c(\cdot, u) \rangle, \\ \tilde{g}_e(\mu, \hat{x}) &= \langle \mu, g^e(\cdot, \hat{x}) \rangle, \end{aligned}$$

with c: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \longrightarrow \mathbb{R}^+$ depending on g^c and the conditional distribution of ξ_k knowing X_k as in (4.35).

One would like to use the structure of the cost to separate the problem of control and estimation. In this sense, the formulation (5.7) is not practical because it involves $\tilde{\mu}_k$ which depends on the estimator whereas f^{aug} actually does not. To split the two problems, we start by noticing that, by writing equation (4.28) and (4.29) for the augmented system, one gets:

$$F^{aug}\left(\tilde{\mu}_k, Y_{k+1}, W_k\right) = F\left(\mu_k, Y_{k+1}, U_k\right).$$

Note that $\tilde{\mu}_0 = \mu_0$. This leads, by recursion on k, to $\tilde{\mu}_k = \mu_k$ almost surely. Finally, we can write our coupled control and estimation problem as a perfect information problem with μ_k as the state:

$$\begin{array}{rrcl} \widetilde{V}(\mu) = \min\limits_{\pi_0^c, \pi_0^e} & E^{\pi}\left[\sum_{k=0}^{+\infty} \alpha^k (\tilde{g}^c(\mu_k, U_k) + \tilde{g}^e(\mu_k, \hat{X}_k)) | \mu_0 = \mu\right] \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, U_k\right), \\ & U_k & = & \pi_0^c(\mu_k), \\ & \hat{X}_k & = & \pi_0^e(\mu_k),\ \forall k \geq 0, \end{array} \tag{5.8}$$

5.1.2 Dynamic programming principle

From (5.8), the Bellman equation of the coupled problem reads:

$$\begin{array}{rrcl} \widetilde{V}(\mu) = \min\limits_{(u,\hat{x}) \in U \times R^{n_x}} & E\left[\tilde{g}^c(\mu, u) + \tilde{g}^e(\mu, \hat{x}) + \alpha \widetilde{V}(\mu_{\ell+1}) | \mu_\ell = \mu\right] \\ \text{s.t.} & \mu_{\ell+1} = F\left(\mu_\ell, Y_{\ell+1}, u\right), \end{array}$$

$$\begin{array}{rrcl} \widetilde{V}(\mu) = \min\limits_{(u,\hat{x}) \in U \times R^{n_x}} & \tilde{g}^c(\mu, u) + \tilde{g}^e(\mu, \hat{x}) + \alpha E\left[\widetilde{V}(\mu_{\ell+1}) | \mu_\ell = \mu\right] \\ \text{s.t.} & \mu_{\ell+1} = F\left(\mu_\ell, Y_{\ell+1}, u\right), \end{array}$$

As F does not depend on $\hat{x}$:

$$\begin{aligned} \widetilde{V}(\mu) = \min_{u\in U} \quad & \left(\min_{\hat{x}\in R^{n_x}} \tilde{g}^e(\mu,\hat{x})\right) + \tilde{g}^c(\mu,u) + \alpha E\left[\widetilde{V}(\mu_{\ell+1})|\mu_\ell = \mu\right] \\ \text{s.t.} \quad & \mu_{\ell+1} = F\left(\mu_\ell, Y_{\ell+1}, u\right). \end{aligned} \tag{5.9}$$

Equation (5.9) illustrates the fact that the problem (5.4), which gathers optimal control and optimal estimation, can actually be split back into a hierarchy of two problems. Indeed, it justifies the use of a resolution scheme in two steps:

1. First, one solves the inner minimisation problem in (5.9):

$$\min_{\hat{x}\in R^{n_x}} \tilde{g}^e(\mu,\hat{x}). \tag{5.10}$$

   Any solution of (5.10), which is an optimal estimation problem, gives a time homogeneous optimal estimation policy denoted by $\pi^e_*(\mu)$. In particular:

$$\tilde{g}^e(\mu,\pi^e_*(\mu)) = \min_{\hat{x}\in R^{n_x}} \tilde{g}^e(\mu,\hat{x}) \tag{5.11}$$

   We set $\tilde{g}^e_*(\mu) = \tilde{g}^e(\mu,\pi^e_*(\mu))$.

2. Secondly, by substituting (5.11) in (5.9) one gets:

$$\begin{aligned} \widetilde{V}(\mu) = \min_{u\in U} \quad & \tilde{g}^e_*(\mu) + \tilde{g}^c(\mu,u) + \alpha E\left[\widetilde{V}(\mu_{\ell+1})|\mu_\ell = \mu\right] \\ \text{s.t.} \quad & \mu_{\ell+1} = F\left(\mu_\ell, Y_{\ell+1}, u\right). \end{aligned} \tag{5.12}$$

   Equation (5.12) can be interpreted as the Bellman equation of a stochastic optimal control problem on $\mathcal{P}(\mathbb{R}^{n_x})$ which has the same optimal value as problem (5.7). The second step is then to solve this problem:

$$\begin{aligned} \widetilde{V}(\mu) = \min_{\pi^c_0} \quad & E^\pi\left[\textstyle\sum_{k=0}^{+\infty} \alpha^k(\tilde{g}^c(\mu_k,U_k) + \tilde{g}^e_*(\mu_k))|\mu_0 = \mu\right] \\ \text{s.t.} \quad \mu_{k+1} &= F\left(\mu_k, Y_{k+1}, U_k\right), \\ U_k &= \pi^c_0(\mu_k),\ \forall k \geq 0. \end{aligned} \tag{5.13}$$

   Note that $\tilde{g}^e_*(\mu) = \langle \mu, g^e(\cdot,\pi^e_*(\mu))\rangle$ so $\tilde{g}^e_*$ is generally nonlinear in μ and cannot be represented as an integral as in (4.35). Therefore, the problem (5.13) cannot be written back as a stochastic optimal control problem with imperfect information. Nevertheless, it can be rewritten in terms of I_k as follows, for $i_0 \in \mathbb{R}^{n_y}$:

$$\begin{aligned} \widetilde{V}(i_0) = \min_{\pi^c_0} \quad & E^\pi\left[\textstyle\sum_{k=0}^{+\infty} \alpha^k(E[g^c(X_k,U_k)|I_k] + E[g^e(X_k,\pi^e_*(I_k)))|I_k]|I_0 = i_0\right] \\ \text{s.t.} \quad I_{k+1} &= (I_k, U_k, Y_{k+1}), \\ U_k &= \pi^c_0(I_k), \end{aligned} \tag{5.14}$$

   where π^e_* denotes here the optimal estimation policy written as a function of I_k.

This two-step scheme is not new in some sense because we have seen in section 4.3.1 that most of the policies are built from an estimation step and control step. However, there are several major differences with a classical scheme. First, it appears naturally from the coupled problem (5.4) meaning that the

splitting is structural in this case and does not come from an assumption of separation. Secondly, the actual value of $\widehat{X}_k$ is not directly involved in the control problem (5.14) only μ_k. In practice however, the same approximation of μ_k is used both in the estimation and the control step. Finally the control problem (5.14) has new interesting properties as we will see in a Section 5.3.

The goal of the next two sections is to justify formally the use of particle filtering and Explicit Dual SMPC in the practical resolution of the scheme described in this section.

5.2 Near-optimal nonlinear filtering

5.2.1 Statement of the problem

In this section, we suppose that the control process is a fixed deterministic sequence $(u_k)_{k\geq 0}$. For $k \geq 0$, we also consider a fixed deterministic sequence of observations $(y_0, \ldots, y_k)$. This is equivalent to considering a fixed vector of information i_k. μ_k is then the distribution of X_k conditionally to $I_k = i_k$ and $\mu_{k|k-1}$ is then the distribution of X_k conditionally to $I_{k-1} = i_{k-1}$.

In the previous section, we have seen that if we model the problem of control and estimation as an optimisation problem then a step of optimal estimation is required at each time k. The optimal estimation problem (5.10) is written at any μ but actually, it is only required to solve it at μ_k. The problem at μ_k reads:

$$\min_{\hat{x}\in R^{n_x}} \tilde{g}^e(\mu_k, \hat{x}). \tag{5.15}$$

In terms of I_k, it reads:

$$\min_{\hat{x}\in R^{n_x}} E[g^e(X_k, \hat{x})|I_k = i_k]. \tag{5.16}$$

The problem (5.15) can be hard to solve and to analyse if g_e is a complex nonlinear nonconvex cost function. Thus, we focus on the Mean Square Error (MSE) minimisation problem i.e. we assume that $g^e(x, \hat{x}) = \|\hat{x} - x\|^2$. We recall the problem seen in Section 4.2:

$$\min_{\hat{x}\in R^{n_x}} E[\|\hat{x} - X_k\|^2|I_k = i_k]. \tag{5.17}$$

We are concerned with conditional MSE in problem (5.17) precisely because it appears in step 1 in Section 5.1.2. However, we are also interested in the minimisation of the total MSE in which one integrates also over i_k. It is a more relevant a priori measure of the estimation error than the conditional MSE because it does not depend on i_k which is unknown initially. It is defined as follows:

$$\begin{array}{ll} \min\limits_{\pi^e} & E\left[\|\widehat{X}_k - X_k\|^2\right] \\ \text{s.t.} & \widehat{X}_k = \pi^e(I_k), \end{array} \tag{5.18}$$

where the expectation is also taken over I_k. It is known that the optimal solution of the problems (5.17) and (5.18) is given by the conditional expectation of X_k conditionally to I_k. We recall the definition of the conditional expectation, denoted by $\widehat{X}_k^*$, in terms of μ_k:

$$\widehat{X}_k^* = E[X_k|I_k = i_k] = \langle \mu_k, Id \rangle, \tag{5.19}$$

where Id is the identity map on $\mathbb{R}^{n_x}$. From this, we define the conditional and total optimal MSE at time k denoted respectively by $e^{cond}_{k,*}$ and $e^{tot}_{k,*}$, as follows, for any i_k:

$$e^{cond}_{k,*}(i_k) = E\left[\|X_k - \widehat{X}^*_k\|^2 | I_k = i_k\right], \tag{5.20}$$

$$e^{tot}_{k,*} = E\left[\|X_k - \widehat{X}^*_k\|^2\right]. \tag{5.21}$$

The simpler problems (5.17) and (5.18) are still impossible to solve analytically because the conditional distribution μ_k is not exactly known in general. Therefore, the issue of solving the problem (5.15) approximately appears naturally. A related issue is to determine if it is possible to use an approximation of μ_k coming from a particle filter to solve the problem (5.15) approximately. To define a candidate approximate solution, we consider a Monte Carlo approximation of μ_k and of $\mu_{k+1|k}$ denoted respectively by μ^N_k and $\mu^N_{k+1|k}$. We recall their general formulation, for $N \geq 1$ and $k \geq 0$:

$$\mu^N_k = \sum_{i=1}^N \omega^i_k \delta_{x^i_k}, \tag{5.22}$$

$$\mu^N_{k+1|k} = \sum_{i=1}^N \omega^i_{k+1|k} \delta_{x^i_{k+1|k}}, \tag{5.23}$$

where δ_x denotes the Dirac measure at x, $\left(\omega^i_k\right)_{i=1,..,N}$ are normalised weights in $\mathbb{R}^+$ and $\left(x^i_k\right)_{i=1,..,N}$ are random particles in $\mathbb{R}^n$. We also assume that this particle filter follows the algorithm from [Hu et al., 2011] and recalled in Algorithm 2 with slight modifications to fit it in our controlled framework. The main feature of this particle filter algorithm is that it has an additional step that forces the mean likelihood of the particles to be sufficiently high. A classical approximation of $\widehat{X}^*_k$ is the empirical mean of the particle filter that we denote by $\widehat{X}^N_k$. We recall its definition in the sequel:

$$\widehat{X}^N_k = \sum_{i=1}^N \omega^i_k x^i_k = \langle \mu^N_k, Id \rangle. \tag{5.24}$$

Similarly to the optimal MSE, we define the empirical conditional and total MSE associated with $\widehat{X}^N_k$ denoted respectively by $e^{cond}_{k,N}$ and $e^{tot}_{k,N}$, as follows, for any i_k:

$$e^{cond}_{k,N}(i_k) = E\left[\|X_k - \widehat{X}^N_k\|^2 | I_k = i_k\right], \tag{5.25}$$

$$e^{tot}_{k,N} = E\left[\|X_k - \widehat{X}^N_k\|^2\right], \tag{5.26}$$

where the expectation is also taken over the randomness of the particles.

The main contribution of Section 5.2 is to show that under suitable assumptions on the dynamics (4.7), the observation equation (4.20) and the particle filter from Algorithm 2, the empirical MSE converges to the optimal MSE as the number of particle goes to infinity. More precisely, we prove, in teh sequel, new error bounds between $e^{cond}_{k,N}(i_k)$ and $e^{cond}_{k,*}(i_k)$ and between $e^{tot}_{k,N}$ and $e^{tot}_{k,*}$.

5.2.2 Error bounds between the optimal MSE and the empirical MSE

The main difficulty in the following comes from the fact that, even if $\widehat{X}^N_k$ is very commonly used as an approximation of $\widehat{X}^*_k$, estimating rigorously the convergence of $\widehat{X}^N_k$ to $\widehat{X}^*_k$ cannot be achieved by

Algorithm 2 Particle filter with a likelihood test

1: Create a sample of N particles x_0^i according to the law μ_0 and initialize the weights ω_0^i with $\frac{1}{N}$.
2: **for** $k = 0, 1, 2 \dots$ **do**
3: **Prediction:**
4: Given a control u_k and a set of particles $\left(x_k^i\right)_{i=1,..,N}$ and a set of normalised weights $\left(\omega_k^i\right)_{i=1,..,N}$, compute the predicted particles by drawing samples from K:

$$\bar{x}_{k+1|k}^i \sim K(\mathrm{d}x_{k+1|k}, x_k^i, u_k), \text{ for } i = 1, .., N,$$
$$\bar{\omega}_{k+1|k}^i = \omega_k^i.$$

5: **Likelihood test**:
6: Get the new observation y_{k+1}.
7: if $\frac{1}{N}\sum_{i=1}^N \rho(y_{k+1}, \bar{x}_{k+1|k}^i) \geq \gamma_{k+1}$ where $\gamma_{k+1} > 0$ is a threshold then proceed to Step 8 otherwise return to Step 4.
8: Rename $x_{k+1|k}^i = \bar{x}_{k+1|k}^i$ and $\omega_{k+1|k}^i = \bar{\omega}_{k+1|k}^i$.
9: **Correction**:
10: Compute the unnormalised updated weights $\left(\tilde{\omega}_k^i\right)_{i=1,..,N}$ thanks to the likelihood function ρ:

$$\tilde{\omega}_{k+1}^i \propto \bar{\omega}_{k+1|k}^i \rho(y_{k+1}, x_{k+1|k}^i).$$

11: **if** *Resampling* **then**
12: Draw the *a posteriori* particles $\left(x_{k+1}^i\right)_{i=1,..,N}$ from the set $\left(x_{k+1|k}^i\right)_{i=1,..,N}$ and $\left(\tilde{\omega}_k^i\right)_{i=1,..,N}$ using a resampling technique and set $\omega_{k+1}^i = \frac{1}{N}$.
13: **else**
14: Set $x_{k+1}^i = x_{k+1|k}^i$ and $\omega_{k+1}^i = \frac{\tilde{\omega}_{k+1}^i}{\sum_{i=1}^N \tilde{\omega}_{k+1}^i}$.
15: **end if**
16: **end for**

classical error bounds on particle filters seen in Section 4.2.3.2. In fact, one can see from equations (5.19) and (5.24) that $\widehat{X}_k^*$ and $\widehat{X}_k^N$ can be seen as the integral of an unbounded function, Id, w.r.t. μ_k and μ_k^N. Therefore, it does not fit in the classical framework of weak error bounds reviewed in Section 4.2.3.2 and new ones are required.

To begin with, without assumptions, we can compare the several MSE using the optimality of $e_{k,*}^{cond}$ and $e_{k,*}^{tot}$. This is the topic of Proposition 5.1.

Proposition 5.1. *The following inequalities hold, for any i_k:*

$$e_{k,*}^{cond}(i_k) \leq e_{k,N}^{cond}(i_k),$$
$$e_{k,*}^{tot} \leq e_{k,N}^{tot}.$$

Proof. It is clear from Algorithm 2 that $\left(\omega_k^i\right)_{i=1,..,N}$ and $\left(x_k^i\right)_{i=1,..,N}$ are independent of X_k conditionally

to i_k so:

$$\widehat{X}_k^* = E\left[X_k | I_k = i_k\right] = E\left[X_k | I_k = i_k, \left(\omega_k^i\right)_{i=1,..,N}, \left(x_k^i\right)_{i=1,..,N}\right].$$

It means that $\widehat{X}_k^*$ is the solution of the following problem:

$$\min_{\hat{x} \in R^{n_x}} E[\|\hat{x} - X_k\|^2 | I_k = i_k, \left(\omega_k^i\right)_{i=1,..,N}, \left(x_k^i\right)_{i=1,..,N}].$$

Moreover, $\widehat{X}_k^N$ is a function of i_k, $\left(\omega_k^i\right)_{i=1,..,N}$ and $\left(x_k^i\right)_{i=1,..,N}$ so it is admissible for the previous problem. Therefore, by optimality of $\widehat{X}_k^*$:

$$E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k, \left(\omega_k^i\right)_{i=1,..,N}, \left(x_k^i\right)_{i=1,..,N}\right] \leq E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k, \left(\omega_k^i\right)_{i=1,..,N}, \left(x_k^i\right)_{i=1,..,N}\right].$$

Then, by integrating conditionally to i_k:

$$\begin{aligned} E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] &\leq E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right], \\ e_{k,*}^{cond}(i_k) &\leq e_{k,N}^{cond}(i_k). \end{aligned} \tag{5.27}$$

By integrating over i_k:

$$e_{k,*}^{tot} \leq e_{k,N}^{tot}, \tag{5.28}$$

and the result is proven. □

We also assume that X_k is square integrable, i.e $\forall k \geq 0$:

$$\mathbb{E}[\|X_k\|^2] < +\infty.$$

This ensures that $e_{k,*}^{cond}(i_k) < +\infty$ and $e_{k,*}^{tot} < +\infty$. The goal of this assumption is to avoid infinite errors which make all the subsequent inequalities trivial.

Proposition 5.1 gives a natural lower bound on both the conditional and total MSE. To find upper bounds, we will first treat the conditional case and then extend it to the total case by integrating over i_k

5.2.2.1 Bound on the conditional MSE

This section is dedicated to the proof of an upper bound on $e_{k,N}^{cond}(i_k)$ and of the convergence of $e_{k,N}^{cond}(i_k)$ to $e_{k,*}^{cond}(i_k)$. The main result of this section is contained in Theorem 5.3.

First,we notice that $e_{k,N}^{cond}(i_k)$ can be rewritten as follows:

$$e_{k,N}^{cond}(i_k) = E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] = E\left[\|X_k - \widehat{X}_k^* + \widehat{X}_k^* - \widehat{X}_k^N\|^2 | I_k = i_k\right], \tag{5.29}$$

By Young's inequality, we get for any $\epsilon > 0$:

$$\begin{aligned} E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] &\leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] + \left(1 + \frac{1}{\epsilon}\right) E\left[\|\widehat{X}_k^* - \widehat{X}_k^N\|^2 | I_k = i_k\right], \\ &\leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] \\ &+ \left(1 + \frac{1}{\epsilon}\right) E\left[\|\langle \mu_k, Id\rangle - \langle \mu_k^N, Id\rangle\|^2 | I_k = i_k\right]. \end{aligned} \tag{5.30}$$

The last inequality can be rewritten as follows:

$$e_{k,N}^{cond}(i_k) \leq (1+\epsilon)e_{k,*}^{cond}(i_k) + \left(1+\frac{1}{\epsilon}\right) e_{k,N}^{filter}(i_k), \tag{5.31}$$

where $e_{k,N}^{filter}(i_k) = E\left[\|\langle \mu_k, Id\rangle - \langle \mu_k^N, Id\rangle\|^2 | I_k = i_k\right]$. One can deduce from Equation (5.31) that, up to ϵ, it is sufficient to control the term $e_{k,N}^{filter}(i_k)$. For $x = (x_1, \ldots, x_{n_x})$ in the canonical basis of $\mathbb{R}^{n_x}$, and $j = 1, .., n_x$ one defines the j^{th} coordinate function, ϕ_j, such that, $\forall x \in \mathbb{R}^{n_x}$:

$$\phi_j(x) = x_j.$$

Then, $e_{k,N}^{filter}(i_k)$ can be decomposed as follows:

$$\begin{aligned} e_{k,N}^{filter}(i_k) &= E\left[\sum_{j=1}^{n_x} |\langle \mu_k, \phi_j\rangle - \langle \mu_k^N, \phi_j\rangle|^2 | I_k = i_k\right], \\ &= \sum_{j=1}^{n_x} E\left[|\langle \mu_k, \phi_j\rangle - \langle \mu_k^N, \phi_j\rangle|^2 | I_k = i_k\right]. \end{aligned} \tag{5.32}$$

This term can be seen as a quadratic error term of the particle filter μ_k^N where the scalar test functions are the coordinate maps. As said earlier, $e_{k,N}^{filter}(i_k)$ cannot be bounded by classical error bounds because Id is not bounded. However, in [Hu et al., 2011], a bound on the L^p-norm for a class of potentially unbounded test functions is given. The error bound is written conditionally to i_k in the following form:

$$E\left[|\langle \mu_k, \phi\rangle - \langle \mu_k^N, \phi\rangle|^p | I_k = i_k\right] \leq C_{k|k} \frac{\|\phi\|_{k,p}^p}{N^{p-p/r}}, \tag{5.33}$$

where $p \geq 2$, $1 \leq r \leq 2$, ϕ is a test function, $\|\phi_j\|_{k,2} = \max(1, \langle \mu_0, |\phi_j|^2\rangle^{\frac{1}{2}}, \ldots, \langle \mu_k, |\phi_j|^2\rangle^{\frac{1}{2}})$, and $C_{k|k}$ is a coefficient depending on i_k. It was originally written conditionally to a sequence of observation $y_{0:k}$ but the extension conditionally to i_k is straightforward.

In the sequel, we would like to apply the bound (5.33) to $e_{k,N}^{filter}(i_k)$. To do so, we present the adapted assumptions of [Hu et al., 2011] for the particular case $\phi = \phi_j$, $p = 2$ and $r = 2$.

Assumption 5.1. *For any $k \geq 1$, for $0 < \epsilon_k < 1$, for almost all i_k, there exists $N_k(i_k) > 0$ such that, for $N \geq N_k(i_k)$:*

$$\gamma_k = \frac{1}{2} \inf_{i_k} \langle \mu_{k|k-1}, \rho\rangle > 0, \tag{5.34}$$

$$P(\langle \mu_{k|k-1}^N, \rho\rangle > \gamma_k | I_k = i_k) \geq 1 - \epsilon_k. \tag{5.35}$$

In particular, under Assumption 5.1, for almost all i_k:

$$\langle \mu_{k|k-1}, \rho\rangle > \gamma_k.$$

Assumption 5.2. *For $k \geq 1$ and for almost all y_k, x_k, x_{k-1} and u_{k-1}:*

$$\begin{aligned} \rho(y_k, x_k) &< +\infty, \\ K(x_k, x_{k-1}, u_{k-1}) &< +\infty. \end{aligned}$$

For $j = 1, .., n_x$, we denote respectively the L^∞-norm of K, ρ and $\rho\phi_j$ w.r.t. x by $\|K\|$, $\|\rho\|$ and $\|\rho\phi_j\|$ i.e, for almost all i_k:

$$\begin{aligned}\|K\| &= \sup_{x_0,x_1,} K(x_1, x_0, u_{k-1}),\\ \|\rho\| &= \sup_x \rho(y_k, x),\\ \|\rho\phi_j^2\| &= \sup_x |\phi_j^2(x)\rho(y_k, x)|.\end{aligned}$$

Assumption 5.3. *For $k \geq 1$ and for almost all i_k,*

$$\begin{aligned}\|K\| &< +\infty,\\ \|\rho\| &< +\infty,\\ \|\rho\phi_j^2\| &< +\infty.\end{aligned}$$

As $\|\rho\phi_j\| \leq \|\rho\phi_j^2\|^{\frac{1}{2}}\|\rho\|^{\frac{1}{2}}$, Assumption 5.3 implies that $\|\rho\phi_j\| < +\infty$. Proposition 5.2 presents then an upper bound of $e_{k,N}^{cond}(i_k)$.

Proposition 5.2. *Under Assumption 5.1, 5.2 and 5.3, for $\epsilon > 0$, for $j = 1, .., n_x$ and $k \geq 0$, for almost all i_k, there exist $C_{k|k,j} > 0$ and $M_{k|k,j} > 0$, such that for $N \geq N_k(i_k)$:*

$$e_{k,N}^{cond}(i_k) \leq (1+\epsilon)e_{k,*}^{cond}(i_k) + \left(1 + \frac{1}{\epsilon}\right)\frac{\sum_{j=1}^n C_{k|k,j}\|\phi_j\|_{k,2}^2}{N}, \tag{5.36}$$

where $\|\phi_j\|_{k,2} = \max(1, \langle\mu_0, |\phi_j|^2\rangle^{\frac{1}{2}}, \ldots, \langle\mu_k, |\phi_j|^2\rangle^{\frac{1}{2}})$. Besides, $C_{k|k,j}$ and $M_{k|k,j}$ follow the following coupled recursion, for almost all i_k:

$$M_{0|0,j} = 3, \tag{5.37}$$

$$C_{0|0,j} = 8\widetilde{C}, \tag{5.38}$$

$$M_{k|k,j} = 2 + \alpha_{k,j}\left(1 + \left(\frac{4-\epsilon_k}{1-\epsilon_k} + 1\right)M_{k-1|k-1,j}\right) \tag{5.39}$$

$$\begin{aligned}C_{k|k,j}^{\frac{1}{2}} &= 2^{\frac{3}{2}}(\widetilde{C})^{\frac{1}{2}}M_{k|k,j}^{\frac{1}{2}} + \frac{2^{\frac{3}{2}}(\widetilde{C})^{\frac{1}{2}}\beta_{k,j}}{(1-\epsilon_k)^{\frac{1}{2}}}M_{k-1|k-1,j}^{\frac{1}{2}} \\ &+ \frac{\|K\|^{\frac{3}{2}}\|\rho\|_{k,2}\beta_{k,j}}{(1-\epsilon_k)|\frac{\gamma_k}{2} - \langle\mu_{k|k-1}, \rho\rangle|}M_{k-1|k-1,j}^{\frac{1}{2}}C_{k-1|k-1,j}^{\frac{1}{2}} + \|K\|\beta_{k,j}C_{k-1|k-1,j}^{\frac{1}{2}},\end{aligned} \tag{5.40}$$

$$\beta_{k,j} = \frac{\|\rho\|(\|\phi_j\rho\| + \frac{\gamma_k}{2})}{\frac{\gamma_k}{2}\langle\mu_{k|k-1}, \rho\rangle}, \tag{5.41}$$

$$\alpha_{k,j} = \|K\|^2\frac{\|\rho\|(\|\phi_j^2\rho\| + \frac{\gamma_k}{2})}{\frac{\gamma_k}{2}\langle\mu_{k|k-1}, \rho\rangle}, \tag{5.42}$$

where:

- $\widetilde{C} > 0$;

- $\|\rho\|_{k,2} = \langle \mu_k, \rho \rangle \leq \|\rho\|$,
- $N_k(i_k) \geq \frac{\|\rho\|_{k,2}^2 \|K\|^2 \underset{j}{max}\, C_{k-1|k-1,j}}{|\gamma_k - \langle \mu_{k|k-1}, \rho \rangle|^2 \epsilon_k}$, *with* $0 < \epsilon_k < 1$ *is fixed independently of* i_k.

Proof. We recall that under Assumption 5.1:

$$\langle \mu_{k|k-1}, \rho \rangle > \gamma_k > \frac{\gamma_k}{2}.$$

Besides, $\forall k \geq 1$, for $0 < \epsilon_k < 1$, for almost all i_k, there exists $N_k(i_k) > 0$ such that, $\forall N \geq N_k(i_k)$:

$$P(\langle \mu^N_{k|k-1}, \rho \rangle > \frac{\gamma_k}{2} | I_k = i_k) \geq 1 - \epsilon_k.$$

Therefore, under the additional assumptions 5.2 and 5.3, we consider theorem 3.1 of [Hu et al., 2011] with $\phi = \phi_j$, with $p = r = 2$ and with $\frac{\gamma_k}{2}$ instead of γ_k. It implies that, for $j = 1, .., n_x$ and $k \geq 1$, for almost all i_k, $\exists C_{k|k,j} > 0$, such that $\forall N \geq N_k(i_k)$:

$$\begin{aligned} E\left[|\langle \mu_k, \phi_j \rangle - \langle \mu^N_k, \phi_j \rangle|^2 | I_k = i_k \right] &\leq \frac{C_{k|k,j} \|\phi_j\|^2_{k,2}}{N}, \\ E\left[\|\langle \mu_k, Id \rangle - \langle \mu^N_k, Id \rangle\|^2 | I_k = i_k \right] &\leq \frac{\sum_{j=1}^n C_{k|k,j} \|\phi_j\|^2_{k,2}}{N}. \end{aligned} \tag{5.43}$$

By putting end to end the computations in [Hu et al., 2011], one can show that, for $k \geq 1$, and $j = 1, .., n_x$, $C_{k|k,j}$ follow the coupled equations (5.37) to (5.42) that involves another coefficient denoted by $M_{k|k,j}$. From the computation of $C_{k|k,j}$ and Assumption 5.1, one can also show that $N_k(i_k)$ can be chosen as such that $N_k(i_k) \geq \frac{\|\rho\|^2_{k,2} \|K\|^2 \underset{j}{max}\, C_{k-1|k-1,j}}{|\gamma_k - \langle \mu_{k|k-1}, \rho \rangle|^2 \epsilon_k}$, with $0 < \epsilon_k < 1$.

Then, by combining equations (5.31) and (5.43), one obtains $\forall \epsilon > 0$, for almost all i_k, $\forall N \geq N_k(i_k)$:

$$e^{cond}_{k,N}(i_k) \leq (1 + \epsilon) e^{cond}_{k,*}(i_k) + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C_{k|k,j} \|\phi_j\|^2_{k,2}}{N},$$

and one gets the result. □

By combining Proposition 5.1 and 5.2, one finally gets Theorem 5.3.

Theorem 5.3. *Under Assumption 5.1, 5.2 and 5.3, for* $\epsilon > 0$, *for* $k \geq 0$, *for almost all* i_k, $\exists C_{k|k,j} > 0$, $\forall N \geq N_k(i_k)$*:*

$$e^{cond}_{k,*}(i_k) \leq e^{cond}_{k,N}(i_k) \leq (1 + \epsilon) e^{cond}_{k,*}(i_k) + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C_{k|k,j} \|\phi_j\|^2_{k,2}}{N}.$$

In particular, for almost all i_k*:*

$$E\left[\|X_k - \widehat{X}^N_k\|^2 | I_k = i_k \right] \underset{N \to +\infty}{\longrightarrow} E\left[\|X_k - \widehat{X}^*_k\|^2 | I_k = i_k \right]. \tag{5.44}$$

Proof. As Proposition 5.1 requires no assumptions, Proposition 5.1 and 5.2 hold under Assumption 5.1, 5.2 and 5.3 and one gets directly that $\forall \epsilon > 0$, for almost all i_k, for any $N \geq N_k(i_k)$:

$$e_{k,*}^{cond}(i_k) \leq e_{k,N}^{cond}(i_k) \leq (1+\epsilon) e_{k,*}^{cond}(i_k) + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^{n} C_{k|k,j} \|\phi_j\|_{k,2}^2}{N}.$$

By choosing $\epsilon = \frac{1}{N^q}$ with $0 < q < 1$, one can obtain from (5.45) that, for almost all i_k, $\forall N \geq N_k(i_k)$:

$$e_{k,*}^{cond}(i_k) \leq e_{k,N}^{cond}(i_k) \leq \left(1 + \frac{1}{N^q}\right) e_{k,*}^{cond}(i_k) + (1+N^q) \frac{\sum_{j=1}^{n} C_{k|k,j} \|\phi_j\|_{k,2}^2}{N}. \tag{5.45}$$

Moreover, the right-hand side converges such that:

$$\left(1 + \frac{1}{N^q}\right) e_{k,*}^{cond}(i_k) + (1+N^q) \frac{\sum_{j=1}^{n} C_{k|k,j} \|\phi_j\|_{k,2}^2}{N} \underset{N \to +\infty}{\longrightarrow} e_{k,*}^{cond}(i_k). \tag{5.46}$$

Thus, it is now clear from (5.45) and (5.46) that for almost all i_k:

$$e_{k,N}^{cond}(i_k) \underset{N \to +\infty}{\longrightarrow} e_{k,*}^{cond}(i_k),$$

which means that:

$$E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] \underset{N \to +\infty}{\longrightarrow} E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right].$$

□

Theorem 5.3 basically means that $\widehat{X}_k^N$ is near-optimal with respect to the conditional MSE. Besides, equation (5.45) gives an estimation of the speed of convergence of the empirical MSE. For example, for $q = \frac{1}{2}$, one can see that the speed of convergence of this MSE is of order $\frac{1}{\sqrt{N}}$. It is slower than usual in Monte Carlo methods. One would rather expect a convergence rate of order $\frac{1}{N}$ as it is the case in equation (5.43). The conservativeness of the bound (5.45) comes from our use of Young's inequality instead of Minkowsky's inequality. Actually, a very similar reasoning could be undertaken using Minkowsky's inequality and one would get a better convergence rate but it would involve the conditional Root Mean Square Errors (RMSE), $(e_{k,*}^{cond})^{\frac{1}{2}}$ and $(e_{k,N}^{cond})^{\frac{1}{2}}$, and not the MSE. The RMSE is easier to interpret than the MSE in practice in an estimation context for the same reasons that the standard deviation is easier to relate to concrete data than the variance. However, minimising a MSE, which is the mean of a quadratic form, is more adapted to the context of stochastic optimisation defined in Section 5.1.1. That is why, we focus on MSEs and not RMSEs in this section even if we lose some precision in the error bounds.

In the sequel, we would like to extend the result of Theorem 5.3 to the total MSE. A intuitive way would be to integrate equation (5.45) over i_k. However, it is not possible in its current form because $C_{k|k,j}$ and $M_{k|k,j}$ depend on i_k which makes the integrability on the right-hand of equation (5.45) hard to evaluate. Moreover, the threshold $N_k(i_k)$ also depends on i_k meaning that one would have to integrate equation (5.45) conditionally to $N \geq N_k(i_k)$ which is not satisfying. One cannot apply the Dominated Convergence theorem either precisely because equation (5.45) does not hold true for sufficiently large N with a threshold independent of i_k.

5.2.2.2 Bound on the total MSE

The main contribution of this section is the extension of the result of Theorem 5.3 to the total MSE. This result is presented in Theorem 5.5.

As in the conditional case, Proposition 5.1 provides a lower bound of $e^{tot}_{k,N}$, we are then looking for an upper bound of $e^{tot}_{k,N}$. As suggested earlier, one would like to integrate the right-hand side of equation (5.36) w.r.t. i_k, which is defined, for $\epsilon > 0$, for almost all i_k and for any $N \geq N_k(i_k)$ by:

$$(1+\epsilon)e^{cond}_{k,*}(i_k) + \left(1+\frac{1}{\epsilon}\right)\frac{\sum_{j=1}^{n} C_{k|k,j}\|\phi_j\|^2_{k,2}}{N}.$$

Note first that $E[e^{cond}_{k,*}(i_k)] = e^{tot}_{k,*} < +\infty$ as X_k is square integrable. Secondly, one can notice that $\forall k \geq 0, \forall j = 1,..,n$, $\langle \mu_k, |\phi_j|^2 \rangle$ is integrable because X_k is square integrable, by assumption. Thus, for $j = 1,..,n_x$, $\|\phi_j\|^2_{k,2}$ is integrable.

Actually, the first issue lays in the fact that $C_{k|k,j}$ depends on i_k and it is not clear at all from equations (5.37) to (5.42) that each term $C_{k|k,j}\|\phi_j\|^2_{k,2}$ is integrable w.r.t. i_k. To bypass this issue, we show in the following that $C_{k|k,j}$ and $M_{k|k,j}$ can be bounded by related coefficients that are constant with respect to i_k under an additional assumption. The dependence of $C_{k|k,j}$ and $M_{k|k,j}$ on i_k come from $\|K\|$, $\|\rho\|$, $\|\rho\phi^2_j\|$ and $\|\rho\phi_j\|$. Therefore, to be able to remove this dependence, we assume that these quantities are bounded uniformly in i_k.

Assumption 5.4. $\forall k \geq 0$, $\forall j = 1,..,n_x$*:*

$$\begin{aligned}
\|K\|_\infty &= \sup_{x_1,x_0,u_0} K(x_0,x_1,u_0) < +\infty,\\
\|\rho\|_\infty &= \sup_{x,y} \rho(y,x) < +\infty,\\
\|\rho\phi^2_j\|_\infty &= \sup_{x,y} |\phi^2_j(x)\rho(y,x)| < +\infty.
\end{aligned}$$

It is clear that Assumption (5.4) implies (5.3). As $\|\rho\phi_j\|_\infty \leq \|\rho\phi^2_j\|^{\frac{1}{2}}_\infty \|\rho\|^{\frac{1}{2}}_\infty$, Assumption 5.4 implies that $\|\rho\phi_j\|_\infty < +\infty$

We can now state the following proposition:

Proposition 5.4. *Under Assumptions 5.1, 5.2 and 5.4 , $\forall k \geq 0$, $\forall j = 1,..,n_x$, there exist $C'_{k|k,j} > 0$ and $M'_{k|k,j} > 0$ such that, for almost all i_k:*

$$\begin{aligned}
C_{k|k,j} &\leq C'_{k|k,j} < +\infty,\\
M_{k|k,j} &\leq M'_{k|k,j} < +\infty.
\end{aligned}$$

Proof. One defines $C'_{k|k,j}$, and $M'_{k|k,j}$ recursively as follows, $\forall k \geq 0$, $\forall j = 1,..,n$:

$$M'_{0|0,j} = 3, \tag{5.47}$$

$$C'_{0|0,j} = 8\widetilde{C}, \tag{5.48}$$

$$M'_{k|k,j} = 2 + \alpha'_{k,j}\left(1 + \left(\frac{4-\epsilon_k}{1-\epsilon_k} + 1\right) M'_{k-1|k-1,j}\right), \tag{5.49}$$

$$\begin{aligned}(C'_{k|k,j})^{\frac{1}{2}} &= 2^{\frac{3}{2}}(\widetilde{C})^{\frac{1}{2}}(M'_{k|k,j})^{\frac{1}{2}} + \frac{2^{\frac{3}{2}}(\widetilde{C})^{\frac{1}{2}}\beta'_{k,j}}{(1-\epsilon_k)^{\frac{1}{2}}}(M'_{k-1|k-1,j})^{\frac{1}{2}} \\ &+ \frac{\|K\|_\infty^{\frac{3}{2}}\|\rho\|_\infty\beta'_{k,j}}{(1-\epsilon_k)^{\frac{\gamma_k}{2}}}(M'_{k-1|k-1,j})^{\frac{1}{2}}(C'_{k-1|k-1,j})^{\frac{1}{2}} + \|K\|_\infty\beta'_{k,j}(C'_{k-1|k-1,j})^{\frac{1}{2}},\end{aligned} \tag{5.50}$$

$$\beta'_{k,j} = \frac{\|\rho\|_\infty(\|\phi_j\rho\|_\infty + \frac{\gamma_k}{2})}{\frac{\gamma_k^2}{2}}, \tag{5.51}$$

$$\alpha'_{k,j} = \|K\|_\infty^2\frac{\|\rho\|_\infty(\|\phi_j^2\rho\|_\infty + \frac{\gamma_k}{2})}{\frac{\gamma_k^2}{2}}. \tag{5.52}$$

Because of Assumption 5.4, the following inequalities hold, $\forall j = 1, .., n$:

$$\begin{aligned}\|\rho\| &\leq \|\rho\|_\infty < +\infty, \\ \|\rho\phi_j^2\| &\leq \|\rho\phi_j^2\|_\infty < +\infty, \\ \|\rho\phi_j\| &\leq \|\rho\phi_j\|_\infty < +\infty.\end{aligned} \tag{5.53}$$

Thus, by recursion on k, $\forall k \geq 0$, $\forall j = 1, .., n$

$$\begin{aligned}C'_{k|k,j} &< +\infty, \\ M'_{k|k,j} &< +\infty.\end{aligned}$$

Moreover, from Assumption 5.1, for almost all i_k:

$$\langle\mu_{k|k-1}, \rho\rangle \geq \gamma_k \geq \frac{\gamma_k}{2}, \tag{5.54}$$

$$\langle\mu_{k|k-1}, \rho\rangle - \frac{\gamma_k}{2} \geq \frac{\gamma_k}{2} > 0,$$

$$\frac{1}{|\langle\mu_{k|k-1}, \rho\rangle - \frac{\gamma_k}{2}|} = \frac{1}{\langle\mu_{k|k-1}, \rho\rangle - \frac{\gamma_k}{2}} \leq \frac{1}{\frac{\gamma_k}{2}}. \tag{5.55}$$

One needs to be able to bound the term $\frac{1}{|\langle\mu_{k|k-1},\rho\rangle - \frac{\gamma_k}{2}|}$ from above uniformly in i_k. Consequently, from (5.53) and (5.55), $\forall k \geq 0$, $\forall j = 1, .., n_x$, for almost all i_k:

$$\begin{aligned}\alpha_{k,j} &\leq \alpha'_{k,j}, \\ \beta_{k,j} &\leq \beta'_{k,j}.\end{aligned} \tag{5.56}$$

Finally, from (5.54), (5.56), equations (5.37) to (5.41) and equations (5.47) to (5.52), one can show by recursion on k, that $\forall k \geq 0$, $\forall j = 1, .., n$, for almost all i_k:

$$\begin{aligned}C_{k|k,j} &\leq C'_{k|k,j}, \\ M_{k|k,j} &\leq M'_{k|k,j}.\end{aligned}$$

Besides, since γ_k, $\|\rho\|_\infty$, $\|\rho\phi_j^2\|_\infty$ and $\|\rho\phi_j\|_\infty$ do not depend on i_k, $C'_{k|k,j}$ and $M'_{k|k,j}$ do not depend on i_k either and one gets the result. □

Finally, each term $C'_{k|k,j}\|\phi_j\|_{k,2}^2$ is integrable which solves our first problem. Our second problem is that the threshold $N_k(i_k)$ also depends on i_k. Actually, under the same assumption, one can find a larger threshold that does not depend on i_k.

This is the topic of the next result which is the main one of the section:

Theorem 5.5. *Under Assumptions 5.1, 5.2 and 5.4, for $0 < q < 1$, for $k \geq 0$, there exists $\bar{N}_k > 0$, such that for any $N \geq \bar{N}_k$*

$$e_{k,*}^{tot} \leq e_{k,N}^{tot} \leq \left(1 + \frac{1}{N^q}\right) e_{k,*}^{tot} + (1 + N^q)\,\frac{\sum_{j=1}^n C'_{k|k,j} E\left[\|\phi_j\|_{k,2}^2\right]}{N} < +\infty. \tag{5.57}$$

In particular:

$$E\left[\|X_k - \widehat{X}_k^N\|^2\right] \underset{N\to+\infty}{\longrightarrow} E\left[\|X_k - \widehat{X}_k^*\|^2\right]. \tag{5.58}$$

Proof. Under Assumptions 5.1, 5.2 and 5.4, Proposition 5.2 holds and implies that, for almost all i_k, $\forall N \geq N_k(i_k)$, $\forall \epsilon > 0$:

$$E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] \leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C_{k|k,j}\|\phi_j\|_{k,2}^2}{N}, \tag{5.59}$$

where $N_k(i_k) = \dfrac{\|\rho\|_{k,2}^2 \|K\|^2 \max\limits_j C_{k-1|k-1,j}}{|\frac{\gamma_k}{2} - \langle \mu_{k|k-1}, \rho\rangle|^2 \epsilon_k}$.

First, one can notice from (5.53), (5.55) and Proposition 5.4 that, $\forall k \geq 1$, for almost all i_k:

$$N_k(i_k) \leq \frac{\|\rho\|_\infty^2 \|K\|_\infty^2}{(\frac{\gamma_k}{2})^2 \epsilon_k} \max_j C'_{k-1|k-1,j} \equiv \bar{N}_k.$$

$\bar{N}_k$ does not depend on i_k then (5.59) is true for a number of particles independent of i_k i.e for almost all i_k, $\forall N \geq \bar{N}_k$, $\forall \epsilon > 0$:

$$E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] \leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C_{k|k,j}\|\phi_j\|_{k,2}^2}{N}.$$

By using Proposition 5.4 again:

$$E\left[\|X_k - \widehat{X}_k^N\|^2 | I_k = i_k\right] \leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2 | I_k = i_k\right] + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C'_{k|k,j}\|\phi_j\|_{k,2}^2}{N}.$$

Now one is able to integrate over i_k, which leads to:

$$E\left[\|X_k - \widehat{X}_k^N\|^2\right] \leq (1+\epsilon)\, E\left[\|X_k - \widehat{X}_k^*\|^2\right] + \left(1 + \frac{1}{\epsilon}\right) \frac{\sum_{j=1}^n C'_{k|k,j} E\left[\|\phi_j\|_{k,2}^2\right]}{N}. \tag{5.60}$$

We recall that $E\left[\left\|X_k-\widehat{X}_k^*\right\|^2|I_k=i_k\right]$ and $\|\phi_j\|_{k,2}^2$ are integrable w.r.t. i_k because X_k is square-integrable. Moreover, from Proposition 5.1, and by taking $\epsilon=\frac{1}{N^q}$ with $0<q<1$, one can get additionally that $\forall k\geq 0, \forall N\geq \bar{N}_k$:

$$e_{k,*}^{tot}\leq e_{k,N}^{tot}\leq\left(1+\frac{1}{N^q}\right)e_{k,*}^{tot}+(1+N^q)\frac{\sum_{j=1}^n C'_{k|k,j}E\left[\|\phi_j\|_{k,2}^2\right]}{N}<+\infty.$$

The convergence result is straightforward from the previous equation. □

Similarly to Theorem 5.3, Theorem 5.5 means that the total MSE associated with $\widehat{X}_k^N$ is close to be optimal if the number of particle is sufficiently high. Theorem 5.5 also provides an estimation of the rate of convergence. This leads to several remarks and interpretations.

Remark 5.1.

- Assumption 5.2 is very mild because most systems have finite likelihood and transition kernel. Besides, in Section 6.2 dedicated to TAN with a real ground map, we give an example of model in TAN where Assumptions 5.3 and 5.4 are verified. It requires that the likelihood function ρ vanishes sufficiently when $\|x\|\to+\infty$ to counter the increasing effect of ϕ_j as explained in [Hu et al., 2008].

  Finally, Assumption 5.1 is natural in particle filtering. It requires that the predicted *distribution* $\mu_{k|k-1}$ and the predicted *particles* match the likelihood ρ for each information vector i_k through the following inequalities, for $k\geq 0$, for i_k:

$$\begin{aligned}\langle\mu_{k|k-1},\rho\rangle&>\gamma_k,\\ P(\langle\mu_{k|k-1}^N,\rho\rangle&>\gamma_k|I_k=i_k)\geq 1-\epsilon_k.\end{aligned}$$

  As said in Section 4.2, the failure of this property a well known issue in particle filtering that is studied in more depth in [Hu et al., 2008], [Le Gland et al., 2004] and [Crisan and Doucet, 2002]. It is notably showed that it has an impact on the precision of some error bounds. In [Hu et al., 2011], it is proven that the first inequality and Step 7 in Algorithm 2 imply the second inequality Intuitively, Step 7 in Algorithm 2 forces the particles to be positioned where the true state is the most likely to be with respect to the new observation y_k.

- The coefficients $M'_{k|k,j}$ and $C'_{k|k,j}$ tend to $+\infty$ with time. Actually, It can be easily seen from equation (5.49) that, for $j=1..n_x$, and $k\geq 1$,

$$M'_{k|k,j}\geq\alpha_{k,j}\theta_k M'_{k-1|k-1,j},$$

  with $\theta_k=1+\frac{4-\epsilon_k}{1-\epsilon_k}\geq 2$. If $\alpha_{k,j}\geq 1$ then $M'_{k|k,j}\geq 2M'_{k-1|k-1,j}$ and $M'_{k|k,j}$ goes to $+\infty$. From equation (5.50), $(C'_{k|k,j})^{\frac{1}{2}}\geq 2^{\frac{3}{2}}(\widetilde{C})^{\frac{1}{2}}(M'_{k|k,j})^{\frac{1}{2}}$ and $C'_{k|k,j}$ would go to $+\infty$ too. By looking closer to $\alpha_{k,j}$ in equation (5.52), one can get:

$$\begin{aligned}\alpha'_{k,j}&=\|K\|_\infty^2\frac{\|\rho\|_\infty(\|\phi_j^2\rho\|_\infty+\frac{\gamma_k}{2})}{\frac{\gamma_k^2}{2}},\\ \alpha'_{k,j}&=\|K\|_\infty^2\frac{\|\rho\|_\infty}{\gamma_k}\frac{\|\phi_j^2\rho\|_\infty+\frac{\gamma_k}{2}}{\frac{\gamma_k}{2}},\end{aligned}$$

with $\frac{\|\rho\|_\infty}{\gamma_k} \geq 1$ from Assumption 5.1 and $\frac{\|\phi_j^2\rho\|_\infty^{+\frac{\gamma_k}{2}}}{\frac{\gamma_k}{2}} \geq 1$. Assume that $\alpha_{k,j} < 1$. One gets in particular that $\|K\|_\infty^2 < 1$. It implies that for $x_0 \in \mathbb{R}^{n_x}$, $x_1 \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$:

$$K(x_1, x_0, u) < 1.$$

This means that the transition densities are very spread for each x_0 and u. This excludes systems with small noises whose distribution is a peak, like a Gaussian density with a small variance. For this reason, we consider that $\|K\|_\infty^2 \geq 1$ and consequently that $C'_{k|k,j}$ and $M'_{k|k,j}$ go to $+\infty$.

Thus, $\bar{N}_k$ tends to $+\infty$ too which means that the error bound from Theorem 5.5 is not uniform in k. This is classical for this type of error bound as described in [Crisan and Doucet, 2002]. To get uniformity in k, one typically need a mixing assumption on K, see [Le Gland et al., 2004] for an example of uniform bound in time.

- Theorem 5.3 justifies the use of particle in the framework of Section 5.1.2 because it shows that $\widehat{X}_k^N$ solves approximately the problem 5.17 which was the first objective of this section. Theorem 5.5 rather show the way to a proof of error bounds on particle filters oriented toward output-feedback stabilisation. Actually, let us assume that a stricly suboptimal estimator of X_k w.r.t. to the total MSE, denoted by $\widehat{X}_k^{sub}$, is available. For example, $\widehat{X}_k^{sub}$ may come from a Kalman-like filter in a nonlinear case. By optimality of $\widehat{X}_k^*$, one gets, for $k \geq 0$:

$$E\left[\|X_k - \widehat{X}_k^*\|^2\right] < E\left[\|X_k - \widehat{X}_k^{sub}\|^2\right].$$

By Theorem 5.5, one gets for a sufficiently large N that:

$$E\left[\|X_k - \widehat{X}_k^N\|^2\right] < E\left[\|X_k - \widehat{X}_k^{sub}\|^2\right],$$

meaning the particle filter has better performance than any other suboptimal filter if the number of particles is sufficiently high. This result is not surprising and observed in practice. However, a rigorous proof of such a result has never been made to the best of our knowledge.

A possible application could be, for example, to show the stability of the estimator $\widehat{X}^N$ in terms of the boundedness of $E\left[\|X_k - \widehat{X}_k^N\|^2\right]$ by showing it for $E\left[\|X_k - \widehat{X}_k^{sub}\|^2\right]$ for some $\widehat{X}_k^{sub}$. In fact in [Reif et al., 1999], under a small error assumption, the stability of the Extended Kalman filter is proven in terms of bounded MSE. Other results of stability of nonlinear filters can be found in [Karvonen, 2014]. The idea would be to show that under the suitable assumptions, $\exists C > 0$, such that $\forall k \geq 0$ and for sufficiently high N:

$$E\left[\|X_k - \widehat{X}_k^N\|^2\right] \leq E\left[\|X_k - \widehat{X}_k^{sub}\|^2\right] \leq C.$$

This statement is not rigorous for the moment because one still needs a number of particle increasing with k according to the previous remark. However, it seems to be a good alternative in order to show MSE boundedness for a particle filter. Note that this kind of result is hard to obtain if one considers directly the particle filter because one typically needs some nonlinear stochastic observability condition. See again [Karvonen, 2014]. Finally, this kind of result is very useful in an output feedback control perspective because it could be a first step toward showing a moment stability result of the true state of system, X_k, in a nonlinear framework. See [Hokayem et al., 2012] for an example of outputfeedback moment stability with bounded MSE in a linear context.

5.3 Stochastic optimal control with an estimation-based cost

Following the concepts of MPC, one would like to approach the problem (5.13) by a finite horizon one, for $1 \leq T < +\infty$ and any $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} \widetilde{V}_0(\mu, T) = \min\limits_{\pi_0,\dots,\pi_{T-1}} & E^{\pi}\left[\sum_{k=0}^{T-1} \tilde{g}^c(\mu_k, U_k) + \tilde{g}^e_*(\mu_k, U_k) + \tilde{g}^c_F(\mu_T) + \tilde{g}^e_F(\mu_T) | \mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, U_k\right), \\ & U_k & = & \pi_k(\mu_k),\ \forall k = 0, .., T-1, \end{array} \tag{5.61}$$

where $\tilde{g}^c_F(\mu_T)$ and $\tilde{g}^e_F(\mu_T)$ are two final costs. The corresponding MPC law can be written as follows:

$$\pi^*_{MPC}(\mu_k) = \pi^*_0(\mu_k),$$

where π^*_0 is the first optimal control law in (5.61).

The main interest of the problem (5.61) is that π^*_{MPC} exhibits both implicit and explicit dual effect. The implicit dual effect comes from the Bellman equation as in the classical case discussed in Section 4.3.1. It also manifests some explicit dual effect coming from the term $g^e(X_k, \pi^e_*(I_k)))$ which depends explicitly on I_k. The term $E[g^e(X_k, \pi^e_*(I_k)))|I_k]$ also written as $\tilde{g}^e_*(\mu_k)$ is the optimal mean estimation error and can indeed be viewed as a measure of the quality of the available information as in explicit dual SMPC from Section 4.3.2.3. As $\tilde{g}^e_*$ is nonlinear in μ_k, the open loop approximation of problem (5.61) still requires to propagate the filtering distribution. The open loop problem reads, for any $\mu \in \mathcal{P}(\mathbb{R}^{n_x})$:

$$\begin{array}{rrcl} \widetilde{V}_0(\mu, T) = \min\limits_{u_0,\dots,u_{T-1} \in U} & E\left[\sum_{k=0}^{T-1} \tilde{g}^c(\mu_k, u_k) + \tilde{g}^e_*(\mu_k, u_k) + \tilde{g}^c_F(\mu_T) + \tilde{g}^e_F(\mu_T) | \mu_0 = \mu\right] & & \\ \text{s.t.} & \mu_{k+1} & = & F\left(\mu_k, Y_{k+1}, u_k\right). \end{array} \tag{5.62}$$

More precisely, if one removes the feedback structure in problem (4.90) to get the open loop problem (4.92), then the dependency of the cost $\tilde{g}_k$ on μ_k can also be removed because $E[E[g_k(X_k, u_k)|I_k]] = E[g_k(X_k, u_k)]$. As a result, the cost depends only on $\mu_{k+\ell|k}$. In problem (4.87), the dependency cannot be omitted and $\mu_{k+\ell}$ must be computed inside the optimisation problem.

For this reason, $\tilde{g}^e_*(\mu_k)$ is not a very practical measure of information. One would prefer a measure of information that does not depend on the realisations of I_k. Therefore, it seems natural to approach $\tilde{g}^e_*(\mu_k)$ by classical measures of information used in explicit dual SMPC like a scalar function of the FIM. As in Section 4.3.2.3, we denote such a measure by g^{info}. We recall problem (4.96) in the framework of this section:

$$\begin{array}{rrcl} V_{EX}(\mu_k, T) = \min\limits_{u_0,\dots,u_{T-1} \in U} & E\left[\sum_{\ell=0}^{T-1} \tilde{g}^{ex}(\tilde{\mu}_{\ell|0}, u_\ell) + \tilde{g}^{ex}_F(\tilde{\mu}_{T|0}) | \tilde{\mu}_0 = \mu_k\right] & & \\ \text{s.t.} & \tilde{\mu}_{\ell+1|0} & = & G\left(\tilde{\mu}_{\ell|0}, u_\ell\right)\ \forall \ell = 0, .., T-1, \end{array} \tag{5.63}$$

where $\tilde{g}^{ex} = \tilde{g}^c + \tilde{g}^{info}$ and $\tilde{g}^{info}(\mu) = \langle \mu, g^{info} \rangle$. The MPC law in this case reads:

$$\pi_{EX}(\mu_k) = u^*_0(\mu_k),$$

where $u^*_0(\mu_k)$ is the first optimal control in problem (5.63) The main statement of this section is that explicit dual control problems can be considered as approximations of the problem (5.13) where $\tilde{g}^e_*$ is replaced by a reward on the a priori information $\tilde{g}^{info}_*$.

To conclude about joint modelling of control and estimation, we have seen that a coupled optimal control and estimation can be split into two problems: one optimal estimation problem and one optimal

control problem whose cost is based on optimal estimation. It is shown that in the case of MSE minimisation, the empirical mean from a modified particle is near optimal. Besides, explicit dual control problems appear to be a good tractable approximation of the estimation-based optimal control problem. This gives a more formal justification of the use particle filtering and explicit dual control in general.

In practice, it is still not clear in the literature how to use particle filtering in dual control. In the sequel, we present two explicit dual SMPC schemes based on particle filtering and the Fisher Information Matrix.

Chapter 6

Dual Particle Fisher Stochastic Model predictive control and its application to TAN

The goal of this chapter is to present two explicit dual SMPC nonlinear controllers which combine a particle filter for state estimation and the Fisher Information Matrix for active information probing. The main difference between the two controllers is the way in which the stabilising effect is implemented inside the optimisation problem. These controllers are applied to TAN with a real terrain and numerical simulations and are displayed at the end of the section.

6.1 Design of the Dual Particle Fisher SMPC schemes

The main issue left from Chapter 5 is the practical design of a SMPC scheme based on problem (5.63). To do so, we rewrite problem (5.63) in the original state space, for $k \geq 0$:

$$\begin{array}{rrcl} V_{EX}(\mu_k, T) = \min\limits_{u_0,\dots,u_{T-1}} & E_{\mu_k}\left[\sum_{\ell=0}^{T-1} g_\ell^{ex}(X_\ell, u_\ell, \xi_\ell) + g_F^{ex}(X_T)\right] & & \\ \text{s.t.} & X_{\ell+1} & = & f(X_\ell, u_\ell, \xi_\ell), \\ & u_\ell & \in & \mathbb{U}_\ell,\ \forall \ell = 0, .., T-1, \\ & X_0 & \sim & \mu_k, \end{array} \tag{6.1}$$

where $\forall \ell = 0, .., T-1$, $\mathbb{U}_\ell \in \mathcal{U}$ is set of constraints on the control. We introduce these new sets to include the case where one would like to restrain even more the values of the control. We recall that the costs g_ℓ^{ex} and g_F^{ex} are defined as follows, for $\ell = 0, .., T-1$:

$$\begin{aligned} g_\ell^{ex} &= g_\ell^c + g_\ell^{info}, \\ g_F^{ex} &= g_F^c + g_F^{info}. \end{aligned}$$

The issue of the measures of information g_ℓ^{info} and g_F^{info} has already been addressed in Section 4.3.2.3. There still remains to decide on the costs g_ℓ^c and g_F^c and the sets $\mathbb{U}_\ell$. There also remains to build a tractable approximation of problem (6.1) that allows the use of a particle filter.

6.1.1 Implementation of the guiding objective in Explicit Dual SMPC

The choice of the objective functions g_ℓ^c and g_F^c and of the control constraints $\mathbb{U}_\ell$ is a matter of modelling. Indeed, in receding horizon control, one often has an economic cost to minimise that comes from practical considerations, such as a price or a fuel consumption. Moreover, the control must be designed to attain some target in the state space: it is the guiding problem. There are three classical ways to represent and address this issue in MPC:

- In classical MPC, notably in tracking MPC, the target is fixed a priori. It is for example, a steady state, $\bar{x} \in \mathbb{R}^{n_x}$, in the deterministic setting. After choosing $\bar{x}$, one adds a new term to the cost to enforce stability in some sense. In this case, the general cost is decomposed in the following way, for $\ell = 0, .., T-1$:

$$\begin{aligned} g_\ell^c &= g_\ell^{stab} + g_\ell^{eco}, \\ g_F^c &= g_F^{stab} + g_F^{eco}. \end{aligned} \tag{6.2}$$

  Thus, g_ℓ and g_F realise a compromise between convergence and economic costs. For instance, in the LQG case, $g_i(x, u, \xi) = x^T M_x x + u^T M_u u$ where M_x and M_u are positive definite matrices. The first term drives the state of the system to zero and the second term penalizes high controls, that can be expensive. The compromise is dealt with by tuning the matrices M_x and M_u. There also exist methods based on stabilising state constraints but we omit them as they are not well adapted to the stochastic case. See [Mayne et al., 2000] for a review of stabilisation techniques in classical MPC.

- Conversely, in economic MPC, the economic costs g_ℓ^{eco} is fixed and the final cost is generally not included. The guiding target is then derived from these costs. For example, in the deterministic setting, one usually looks for an steady state $\bar{x}^*$ that is, in some way, optimal w.r.t. g_ℓ^{eco}. The main point here is that $\bar{x}^*$ is not known beforehand.

- Finally, in some cases the guiding goal is fixed and the cost cannot be modified to make the system converge. An alternative is then to add a drift constraint on the first control, u_0, that enforces the decreasing of some Lyapunov-like function, during the first time step only, such that, for $\ell = 0, .., T$:

$$\begin{aligned} &g_\ell^c = g_\ell^{eco}, \\ &\text{a negative drift condition on } u_0. \end{aligned} \tag{6.3}$$

  Actually, since u_0^* is the only control from the optimal sequence that is applied on the system, the Lyapunov function decreases along the whole trajectory and then stability is obtained. It is also known as Lyapunov Economic MPC, see [Ellis et al., 2014] for a review in the deterministic setting. In the stochastic setting, it has been applied with output feedback for continuous-time non-linear systems in [Homer and Mhaskar, 2017] and for a discrete-time linear system with bounded controls in [Hokayem et al., 2012] and [Mishra et al., 2017]. To the best of our knowledge, these concepts have not been used in the context of dual SMPC.

We consider, in a very broad sense, that our guiding objective is to reach some point in the state space with high probability. Thus, we exclude the case of pure economic MPC. In the sequel, we consider the two remaining approaches to build two new SMPC controllers.

6.1.2 Penalising Fisher Dual Particle SMPC scheme

The first SMPC scheme presented here is based on the idea of the stabilising cost. Therefore the total costs read a priori :

$$\begin{aligned} g_\ell^{ex} &= g_\ell^{stab} + g_\ell^{eco} + g_\ell^{info}, \\ g_F^{ex} &= g_F^{stab} + g_F^{eco} + g_F^{info}. \end{aligned} \tag{6.4}$$

Notice that (6.4) represents the sum of three terms that usually contain weights that must be tuned. By construction, the weights affect the convergence of the system to the guiding goal which makes their tuning difficult. Roughly speaking, if too much importance is given to g_i^{info} then probing will be efficient but the system may get stuck far from the target. On the contrary, if too much importance is given to g_i^{stab} then the probing effect will not be sufficient and output feedback performance may be poor. It is difficult to know which case will occur a priori depending of the value of the weights because the optimal costs are not known at first sight. More precisely, this means that the weight on $(g_\ell^{info})_{\ell=1,..T-1}$ and (g_F^{info}), denoted by α_k, must decrease with k. Here k denotes the actual time and ℓ the time inside the optimisation problem. In fact, if α_k is constant, then the system will usually converge toward an point that is optimal for the compromise realised by (6.4), like in Economic MPC. This compromise has to be oriented more and more toward g_{stab} to allow asymptotic stability. Therefore, by assuming that the weight on g_ℓ^{eco} is fixed at a low value, it is clear that α_k must decrease with k. As a result, we rather consider the following total costs:

$$\begin{aligned} g_{\ell,k}^{ex} &= g_\ell^{stab} + g_\ell^{eco} + \alpha_k g_\ell^{info}, \\ g_{F,k}^{ex} &= g_F^{stab} + g_F^{eco} + \alpha_k g_F^{info}, \end{aligned} \tag{6.5}$$

where $(\alpha_k)_{k\geq 0}$ is decreasing. No constraint on the control is added so that $\forall \ell = 0,..,T-1$, $\mathbb{U}_\ell = \mathcal{U}$. The theoretical problem to solve is the particular case of (6.1) defined as follows, for $k \geq 0$:

$$\begin{aligned} V_{EX}^{pen}(\mu_k, T) = \min_{u_0,\dots,u_{T-1}} \quad & E_{\mu_k}\left[\textstyle\sum_{\ell=0}^{T-1} g_{\ell,k}^{ex}(X_\ell, u_\ell, \xi_\ell) + g_{F,k}^{ex}(X_T)\right] \\ \text{s.t.} \quad X_{\ell+1} &= f(X_\ell, u_\ell, \xi_\ell), \\ u_\ell &\in \mathcal{U},\ \forall \ell = 0,..,T-1, \\ X_0 &\sim \mu_k. \end{aligned} \tag{6.6}$$

The last step is the numerical resolution of problem (6.6). Note that one cannot draw samples from μ_k directly in general because μ_k is unknown. An idea is to use a particle filter μ_k^N as the initial distribution in (6.6). Actually, it appears that, in output feedback MPC, there is a synergy between the scenario approach [Calafiore and Campi, 2006] and particle filtering. In fact, the initial condition for each independent scenario can be chosen as a particle from the current set of particles. It improves global performance compared to a similar technique involving a Kalman filter in which, the initials conditions are always drawn according to a unimodal law. This method was already used in [Sehr and Bitmead, 2016] but not in a dual MPC scheme. The approximation can be defined for $k \geq 0$ by:

$$V_{EX,N}^{pen}(\mu_k^N, T) = \min_{u_0\cdots u_{T-1}} \textstyle\sum_{i=1}^{N_s} \omega_k^i \left(\sum_{\ell=0}^{T-1} g_{\ell,k}^{ex}\left(X_\ell^l, u_\ell, \xi_\ell^i\right) + g_{F,k}^{ex}\left(X_T^i\right)\right)$$

$$\begin{aligned} \text{s.t.} \quad X_{\ell+1}^i &= f(X_\ell^i, u_\ell, \xi_\ell^i), \\ u_\ell &\in \mathcal{U}, \\ X_0^i &= x_k^i,\ \forall i = 1,..,N_s,\ \forall \ell = 0,..,T-1, \end{aligned} \tag{6.7}$$

where:

- $\left(x_k^l\right)_{l=1,..,N}$ and $\left(\omega_k^l\right)_{l=1,..,N}$ are computed according to Algorithm (1).
- $N_s < N$ is the number of scenarios considered. It is supposed to be less than the number of particles for computational reasons. Thus, N_s particles must be extracted from the original set.
- $(\xi_i^l)_{i=0,..,T-1,l=1,..,N_s}$ are i.i.d. random variables sampled from p_ξ.

The entire Outputfeedback SMPC controller described in this section is summed up in Algorithm 3 and has been presented in [Flayac et al., 2017].

Algorithm 3 Penalising Fisher Output-Feedback Control

1: Create a sample of N particles x_0^i according to the law μ_0 and initialize the weights ω_0^i with $\frac{1}{N}$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Solve problem (6.7) starting from the particles x_k^i and the weights ω_k^i.
4: Get an optimal sequence $(u_0^*, \ldots, u_{T-1}^*)$.
5: Set $U_k = u_0^*$.
6: Compute the *a posteriori* particles x_{k+1}^i and weights ω_{k+1}^i given U_k according to Algorithm 1
7: **end for**

As said earlier, the tuning of α_k is hard to realise and yet primordial because it directly impacts the stability of the system. A way to avoid this difficult tuning is to put the stabilising effect in a constraint which leads to our second dual SMPC scheme.

6.1.3 Lyapunov/Fisher Dual Particle SMPC scheme

Intuitively, the idea behind the controller of this section is to look for controls that minimise the loss of information among the stabilising ones. One gives priority to the guiding goal in the form of a hard constraint on the control as described in (6.3). There remains to decide on a drift constraint. In fact, in the following we assume that our system would be stabilisable if we had perfect information.

6.1.3.1 Foster-Lyapunov drift in case of perfect information

As seen in the Section 4.1.2, in the perfect information case, the control is computed from state feedback control policies, i.e. measurable maps, denoted by κ_k, that maps a state X_k to a control U_k. Since equation (4.7) is time homogeneous, it is sufficient to consider a fixed control policy κ. Thus, for any κ, one can define the corresponding closed loop system as follows, $\forall k \in \mathbb{N}$:

$$\begin{aligned} X_{k+1} &= f(X_k, \kappa(X_k), \xi_k), \\ X_0 &\sim p_0. \end{aligned} \tag{6.8}$$

A state feedback control policy is said to be *admissible* if $\forall x \in \mathbb{R}^n, \kappa(x) \in \mathcal{U}$. Therefore, equation (6.8) defines also a time-homogeneous Markov chain whose stability can be studied via the classical theory of negative drifts conditions discussed in [Meyn and Tweedie, 2009]. In Proposition 6.1, we focus on *geometric drifts conditions* that are closely related to Lyapunov conditions for exponential stability of continuous-time processes.

Proposition 6.1. *Suppose that there exist $b > 0$ and $\lambda_{min} \in [0, 1)$, a measurable function $V : \mathbb{R}^{n_x} \longrightarrow [0, +\infty)$, a compact set $C \subset \mathbb{R}^{n_x}$ and an admissible state feedback control policy κ such that, for any $x \notin C$:*

$$E[V(f(x, \kappa(x), \xi_0)] \leq \lambda_{min} V(x),$$
$$\sup_{x' \in C} E_{x'}[V(X_1)] = b,$$

then, for any $\lambda \in [\lambda_{min}, 1[$,

$$E_x[V(X_k)] \leq \lambda^k V(x) + b(1-\lambda)^{-1}, \forall k \in \mathbb{N},\ \forall x \in \mathbb{R}^n,$$

where X_k is computed with equation (6.8).

Proposition 6.1 is a slightly different reformulation of Proposition 1 in [Chatterjee and Lygeros, 2015] that can be proven similarly. In particular, we consider that the parameter λ can be chosen arbitrarily in $[\lambda_{min}, 1[$. In practice, λ is a parameter to tune that determines the convergence speed of the system. Moreover, as explained in [Chatterjee and Lygeros, 2015], if Proposition 6.1 is verified for a norm-like function V then for $r > 0$, $P_x(\|X_k\| > r)$ decreases as the inverse of V so the distribution of state concentrates itself around 0. We suppose that the assumptions in Proposition 6.1 are fulfilled and notably that the system (4.7) can be stabilized with perfect information with some admissible state feedback control policy κ.

6.1.3.2 Design of the optimisation problem

The purpose of our MPC scheme is to compute controls that explicitly look for information by minimising some functions of the FIM, g_i^{info} and g_T^{info}, over the controls that stabilise a state estimator. This is guaranteed by the drift condition taken from Proposition 6.1 and applied to $\widehat{X}_k^*$ only when $\widehat{X}_k^* \notin C$. This means that the constraint sets on the control are defined such that for some $\lambda \in [\lambda_{min}, 1), \forall \ell = 1, .., T-1$:

$$\mathbb{U}_\ell = \mathcal{U},$$
$$\mathbb{U}_0 = \{u_0 \in \mathcal{U} \text{ s.t. } E_{\widehat{X}_k^*}\left[V\left(f(\widehat{X}_k^*, u_0, \xi_0)\right)\right] \leq \lambda V(\widehat{X}_k^*), \text{ when } \widehat{X}_k^* \notin C\}.$$

Thus, the stochastic optimal control problem to solve can be written as follows, $\forall \lambda \in [\lambda_{min}, 1)$:

$$\begin{aligned} V_{EX}^{lya} = \min_{u_{0:T-1}} \quad & E_{\mu_k}\left[E[\textstyle\sum_{\ell=0}^{T-1} g_\ell^{ex}(X_\ell, u_\ell, \xi_\ell) + g_T^{ex}(X_T)|I_0]\right] \\ \text{s.t.s.t.} \quad & X_{\ell+1} = f(X_\ell, u_\ell, \xi_\ell), \\ & u_i \in \mathcal{U},\ \forall \ell = 0, .., T-1, \\ & X_0 \sim \mu_k, \end{aligned} \tag{6.9}$$

$$E_{\widehat{X}_k^*}\left[V\left(f(\widehat{X}_k^*, u_0, \xi_0\right)\right] \leq \lambda V(\widehat{X}_k^*), \text{ when } \widehat{X}_k^* \notin C, \tag{6.10}$$

where:

$$g_\ell^{ex} = g_\ell^{eco} + g_\ell^{info},$$
$$g_T^{ex} = g_T^{eco} + g_T^{info}.$$

It is important to notice that the admissibility of the drift constraint is guaranteed by the existence of the stabilising admissible state feedback policy κ. As for the previous SMPC scheme, we would like to solve problem (6.6) using a Monte Carlo approximation. Keeping the notation from Section6.1.2, it reads:

$$V^{lya}_{EX,N}(\mu^N_k, T) = \min_{u_0 \cdots u_{T-1}} \sum_{i=1}^{N_s} \omega^i_k \left(\sum_{\ell=0}^{T-1} g^{ex}_\ell \left(X^l_\ell, u_\ell, \xi^i_\ell \right) + g^{ex}_F \left(X^i_T \right) \right)$$

$$\begin{aligned} \text{s.t.} \quad X^i_{\ell+1} &= f(X^i_\ell, u_\ell, \xi^i_\ell), \\ u_\ell &\in \mathcal{U}, \\ X^i_0 &= x^i_k, \ \forall i = 1, .., N_s, \ \forall \ell = 0, .., T-1, \end{aligned} \tag{6.11}$$

$$\frac{1}{N_{dr}} \sum_{j=1}^{N_{dr}} V(f(\widehat{X}^N_k, u_0, \tilde{\xi}^j_0) \leq \lambda V \left(\widehat{X}^N_k \right), \text{ when } \widehat{X}^N_k \notin C,$$

where:

- $N_{dr} \in \mathbb{N}^*$ is the size of the sample used to approximate the expectation in the drift constraint.
- $(\xi^i_\ell)_{\ell=0,..,T-1,i=1,..,N_s}$ and $(\tilde{\xi}^j_0)_{j=1,..,N_{dr}}$ are i.i.d. random variables sampled from p_ξ.

The entire Outputfeedback SMPC controller described in this section is summed up in Algorithm 4. Notice that, if $\widehat{X}^N_k \in C$, then the constraint (6.10) is not assured to be feasible so the full-state policy κ is used.

Algorithm 4 Lyapunov/Fisher Output Feedback Control

Create a sample of N particles x^i_0 according to the law μ_0 and initialize the weights ω^i_0 with $\frac{1}{N}$.
for $k = 0, 1, 2, \ldots$ **do**
  if $\widehat{X}^N_k \notin C$ **then**
    Solve problem (6.11) starting from the particles x^i_k and the weights ω^i_k.
    Get an optimal sequence $(u^*_0, \ldots, u^*_{T-1})$.
    Set $U_k = u^*_0$.
  else
    Set $U_k = \kappa \left(\widehat{X}^N_k \right)$
  end if
  Compute the *a posteriori* particles x^i_{k+1} and weights ω^i_{k+1} given U_k according to Algorithm 1
end for

This method has two main advantages compared to the one we presented in Algorithm 3:

- First, as mentioned earlier, in Algorithm 3, it is compulsory to decrease the weights on g^{info}_ℓ and g^{info}_T with time otherwise the system converged to a point that was far from the target. In fact, tuning the initial value and the decreasing of α_k is complicated a priori. In problem (6.9), the stability properties are much less influenced by the cost because of the drift condition which is a constraint. The most important parameter to tune is λ, and, in principle, it influences only the convergence speed of the system and not its qualitative properties of stability.

- Secondly, as stated in [Homer and Mhaskar, 2017], the stability properties are contained in one constraint and are easier to achieve in practice. In fact, the convergence of the system depends little on the quality of the solution of the optimisation problem and much more on its admissibility which is easy to obtain with classical solvers.

The main drawback of this method is that it requires to know a drift function and an admissible stabilising full state feedback policy. Both are not obvious to obtain for a general nonlinear system. Another drawback is that the drift constraint is usually not compatible with state constraints because of controllability issues.

In the following, the two new MPC scheme are validated on Terrain-Aided Navigation with a real ground map.

6.2 Modelling of TAN with a real ground map

In this section, Algorithms 3 and 4 are applied to the guidance and localisation of a drone by terrain-aided navigation. Our guiding goal is to drive a drone in a 3D space from an uncertain initial condition X_0 to a compact set centred around 0. If the original target is not 0 then a translation can be made to centre the problem around 0. In the Cartesian coordinates, we assume that the dynamics of the drone are described as follows:

- As in Chapter 2, the state is composed of a 3 dimensional position and a 3 dimensional speed: $X_k = (x_{1,k}, x_{2,k}, x_{3,k}, v_{1,k}, v_{2,k}, v_{3,k})$ and the control of a 3 dimensional acceleration $U_k = (u_{1,k}, u_{2,k}, u_{3,k})$. Note that $(x_{1,k}, x_{2,k})$ represents the horizontal position and $x_{3,k}$ the altitude.
- its dynamics (6.12) is linear with bounded controls, for $k \in \mathbb{N}$,

$$X_{k+1} = AX_k + BU_k + \xi_k, \tag{6.12}$$
$$\|U_k\| \leq U_{max}, \tag{6.13}$$

  where $U_{max} > 0$, $A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times n_u}$ correspond to a discrete-time second order system with damping on the speed, and $\xi_k \sim \mathcal{N}(0, Q)$, as in (4.51). Q is usually chosen in the following form:

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{v_1}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{v_2}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & q_{v_3}^2 \end{bmatrix}$$

  where $q_{v_1} > 0$, $q_{v_2} > 0$ and $q_{v_3} > 0$. In this setting, the position X_k can attained any position with a non-zero probability because of the Gaussian noise in our dynamical model (6.12). Thus, classical linear controllers do not lead to bounded controls almost surely and a nonlinear feedback controller must be designed.

We assume that the dynamics has a relatively simple form because the main difficulty of this application is the nature of the observations. Indeed, we recall that, as in Chapter 2, the speed is measured.

Besides, the only information on the position is a measurement of the difference between the altitude of the drone, $x_{3,k}$, and the altitude of the corresponding vertical point on the ground. We also suppose that the ground is represented by a map, h_M. In practice, h_M is determined by a smooth interpolation of data points on a bounded domain such that:

$$D = [x_{1,min}, x_{1,max}] \times [x_{2,min}, x_{2,max}],$$

where $(x_{1,min}, x_{1,max}, x_{2,min}, x_{2,max}) \in \mathbb{R}^4$. Formally, it means that $h_M : D \longrightarrow \mathbb{R}^+$ maps a horizontal position (x_1, x_2) in D to the corresponding height of the terrain in $\mathbb{R}^+$. However, as X_k can attain leave D with a non-zero probability, the map h_M should be defined on $\mathbb{R}^2$ and not only on D. Nevertheless, it is known in practice that the system will not leave D. To represent this additional information while keeping a Gaussian noise, we define an extended map $\tilde{h}_M : \mathbb{R}^2 \longrightarrow \mathbb{R}^+$ such that, for $(x_1, x_2) \in \mathbb{R}^2$:

$$\begin{aligned} &\tilde{h}_M(x_1, x_2) = h_M(x_1, x_2) \text{ if } (x_1, x_2) \in D, \\ &\tilde{h}_M(x_1, x_2) \geq \|(x_1, x_2)\| \text{ if } (x_1, x_2) \notin D. \end{aligned} \tag{6.14}$$

Therefore, the observation equation reads:

$$Y_k = \begin{bmatrix} y_{1,k} \\ y_{2,k} \\ y_{3,k} \\ y_{4,k} \end{bmatrix} = \begin{bmatrix} v_{1,k} \\ v_{2,k} \\ v_{3,k} \\ x_{3,k} - \tilde{h}_M(x_{1,k}, x_{2,k}) \end{bmatrix} + \eta_k, \tag{6.15}$$

where $\eta_k \sim \mathcal{N}(0, R)$ and R is defined as follows:

$$R = \begin{bmatrix} r_{v_1}^2 & 0 & 0 & 0 \\ 0 & r_{v_2}^2 & 0 & 0 \\ 0 & 0 & r_{v_3}^2 & 0 \\ 0 & 0 & 0 & r_h^2 \end{bmatrix},$$

with $r_{v_1} > 0$, $r_{v_2} > 0$, $r_{v_3} > 0$ and $r_h > 0$.

From the definition of $\tilde{h}_M$ and the Gaussian noise in (6.15), the likelihood function can be described as follows:

$$\rho(Y_k, X_k) \propto \exp\left(-\frac{1}{2r_h^2}(y_{4,k} - x_{3,k} + \tilde{h}_M(x_{1,k}, x_{2,k}))^2\right).$$

Notice that:

$$(y_{4,k} - x_{3,k} + \tilde{h}_M(x_{1,k}, x_{2,k}))^2 \geq \frac{1}{2}\left(y_{4,k}^2 + x_{3,k}^2 + (\tilde{h}_M(x_{1,k}, x_{2,k}))^2\right).$$

Besides, one gets from (6.14) that, if $\|(x_1, x_2)\|$ is sufficiently large:

$$(y_{4,k} - x_{3,k} + \tilde{h}_M(x_{1,k}, x_{2,k}))^2 \geq \frac{1}{2}\|(y_{4,k}, x_1, x_2, x_3)\|^2.$$

As a result, one gets that:

$$\rho(Y_k, X_k) = O\left(\exp\left(-\frac{1}{4r_h^2}\|(y_{4,k}, x_1, x_2, x_3)\|^2\right)\right) \text{ when } \|(y_{4,k}, x_1, x_2, x_3)\| \to +\infty \tag{6.16}$$

Intuitively, equation (6.16) means that the positions that are far outside D are very unlikely. As required, this represents the fact that one knows a priori that the state of the system is inside D. Besides, with a more formal perspective, the exponential decay in equation (6.16) ensures that the boundedness assumption 5.4 holds true. Furthermore, it is not very restrictive to suppose that the whole speed vector is measured because the speed on its own is not enough to reconstruct the whole state and the altitude measurement is still required. The use of particle filters is totally justified for most of the maps, included the one depicted in Figure 2.2. Actually, the conditional distribution μ_k is multimodal in this case and its modes are closely related to the level sets of h_M. In particular , Kalman filters cannot accurately deal with this problem.

Moreover, it appears very naturally that dual control is required in this application. Indeed, the quality of the observations depends on the control and more precisely on the area that is flied over by the drone. Let us assume that the drone flies over a flat area with constant altitude. Then, one measurement of height matches a whole horizontal area and the estimation error on $(x_{1,k}, x_{2,k})$ is of the order of magnitude of the size of the area, which can be very large. On the contrary, if the drone flies over a rough terrain, then one measurement of height corresponds to a smaller area on the ground and the estimation error is reduced. Therefore, in TAN, information probing consists in flying over informative areas on the ground.

Concerning the cost functions, we choose the economic cost to be the norm of the control for both algorithms such that $\forall k = 0, .., T-1$:

$$\begin{aligned} g_k^{eco}(X_k, U_k, \xi_k) &= \alpha_{eco}\|U_k\|_2^2, \\ g_F^{eco} &= 0. \end{aligned}$$

with $\alpha_{eco} > 0$ small. There exist many ways to define the loss of information that depends on the FIM, see [Fedorov and Hackl, 2012]. From (4.93), the trace of J_k^{-1} seems adapted if one wants initially to minimise the MSE. However, it requires to compute the inverse of a matrix which happens to be costly for an optimisation solver. That is why, we choose another loss of information defined as follows, $\forall k = 0, .., T-1$:

$$\begin{aligned} g_k^{info}(J_k) &= \frac{1}{\text{tr}(J_k^2)}, \\ g_F^{info}(J_T) &= \frac{1}{\text{tr}(J_T^2)}, \end{aligned} \tag{6.17}$$

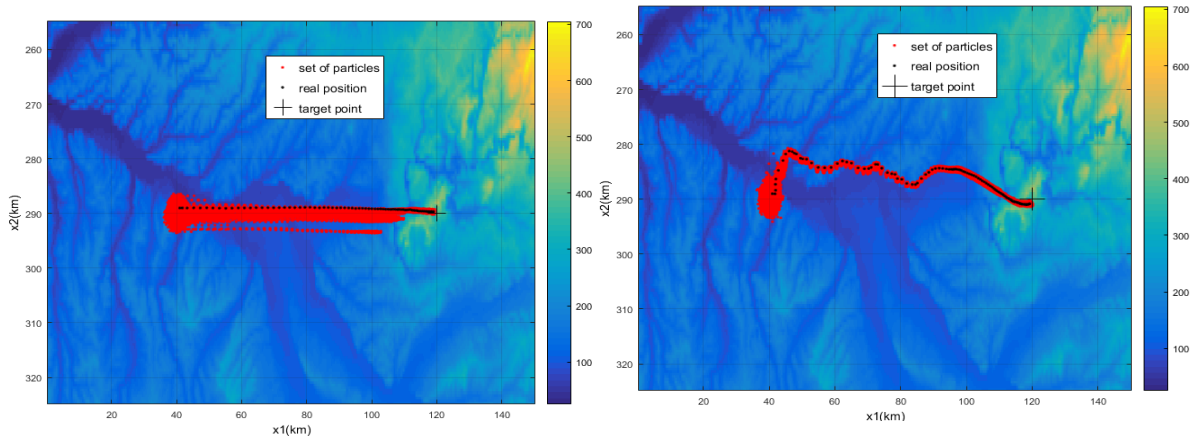
where tr denotes the trace operator. Note that g_k^{info} has the same monotony w.r.t. J_k as $\text{tr}(J_k^{-1})$ and does not require complex computations. Furthermore, the square in (6.17) forces g_k^{info} to be nonnegative even if J_k is not positive definite. This may happen during the resolution depending of the type of optimisation solver used.

In Algorithm 3, we have chosen the stabilising as the simple distance of the final state to the target. It reads $\forall k = 0, .., T-1$:

$$\begin{aligned} g_k^{stab} &= 0, \\ g_F(X_T) &= \|X_T\|_2^2. \end{aligned}$$

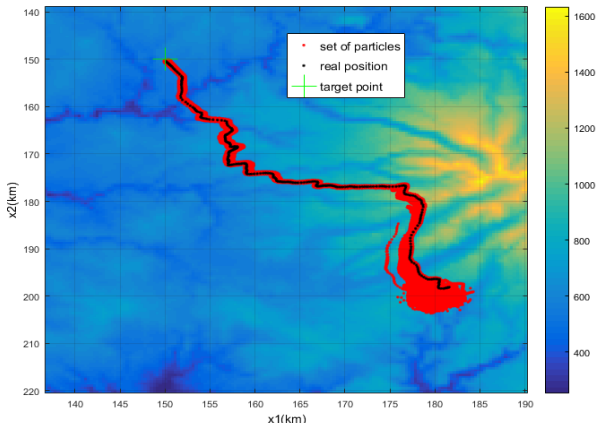
We recall that in Algorithm 4, the guiding problem is addressed with a drift constraint that depends on the dynamics. We have used the drift proposed in Proposition 8 of [Chatterjee and Lygeros, 2015] so that $V(X) = e^{\|X\|}$ and C is a ball centred around 0 whose radius depends only on the disturbances of the dynamics. To guarantee the existence of an admissible stabilising state feedback policy κ, the maximum control U_{max} must be sufficiently high. See [Chatterjee and Lygeros, 2015] for the precise definition of κ, U_{max} and C adapted to system (6.12). This drift constraint is only valid for orthonormal dynamics so we suppose that A is *neutrally stable*. It means that after a change of coordinates, it can be decomposed in a Schur-stable block and an orthonormal block. The drift is then considered only on the last block as the first one already ensures stability.

6.3 Numerical simulations



(a) Trajectory from the OLF controller over a flat area

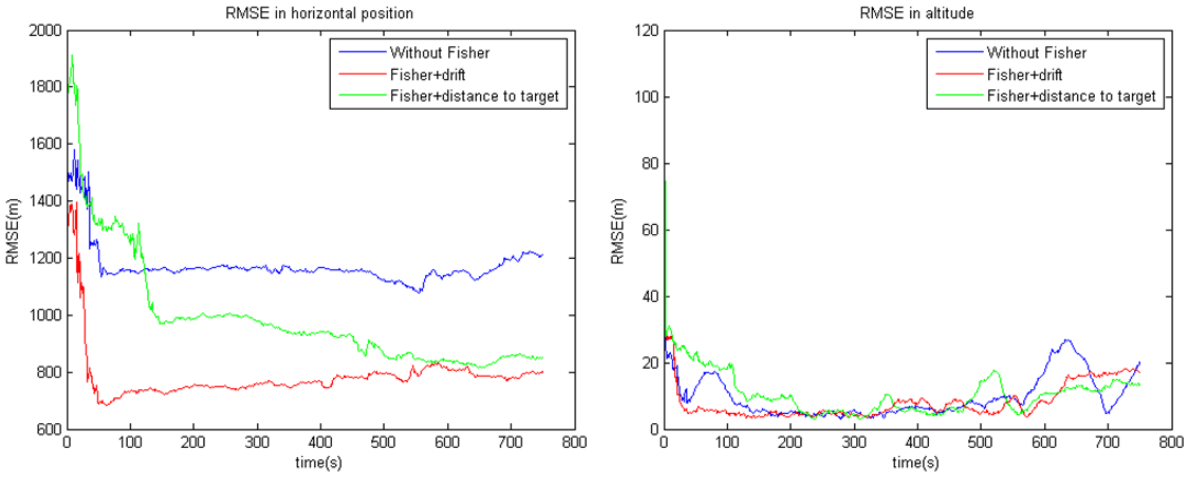
(b) Trajectory from the Penalising Fischer controller over a flat area



(c) Trajectory from the Lyapunov Fischer controller over a rough area

Figure 6.1: Realisation of a trajectory of the true state and the particles resulting from an OLF controller and the two Fisher controllers on a real map

Figure 6.1 represents the horizontal projection of a trajectory and the associated particles obtained by several controllers. The terrain map is at the background of the figure. Figure 6.1a represents a trajectory resulting from Algorithm 4 where the weight on the information is put to zero. The resulting policy is an OLF one. It is clear on Figure 6.1a that this policy is only passively learning. As a matter of fact, the set of particles is very spread during most of the travel because the drone flies over a flat area. It shrinks at the end because the target point happens to be located over a rougher area. We have tested the Penalising Fisher scheme from Algorithm 3 at the same location on the map and a sample trajectory is displayed in Figure 6.1b. As expected, the state estimation performed by the particle filter is much more efficient with the dual controller. We have also tested but not displayed the Lyapunov Fisher scheme and it gives similar results. Figure 6.1c depicts a sample trajectory resulting from the Lyapunov Fisher from Algorithm 4 scheme near a mountain. One can see that the drone makes a detour to gain information as in the previous case.



(a) RMSE in horizontal position

(b) RMSE in altitude

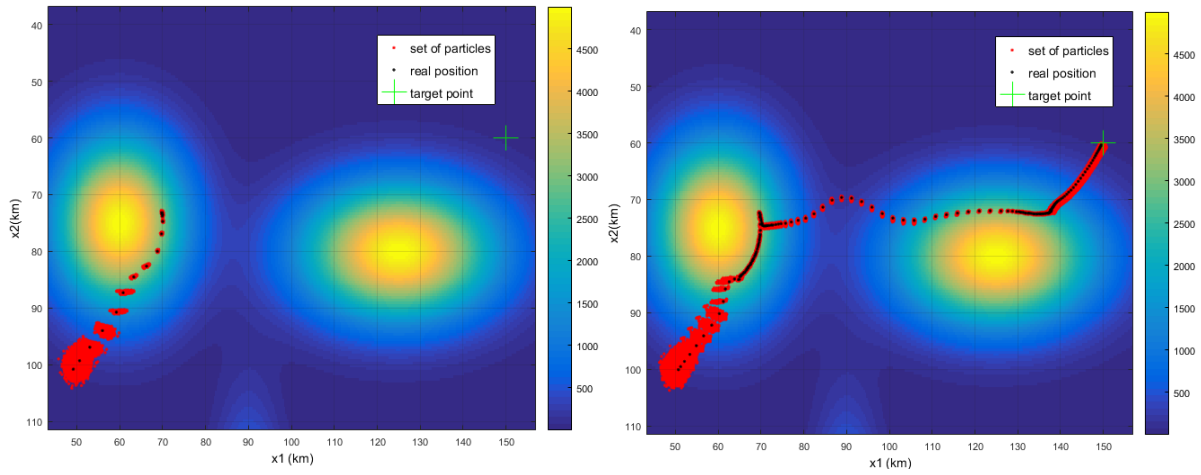
Figure 6.2: Plot of the RMSE in position for 3 policies: without the FIM (blue), with the FIM and the drift from Algorithm 4 (red), with the FIM and the distance to the target from Algorithm 3 (green)

Figure 6.2a and 6.2b shows RMSE respectively in horizontal position and altitude after 30 Monte Carlo simulations of the three different policies described previously. It appears that, on average, the horizontal estimation error is reduced with information probing. It is actually due to the fact that the filter diverges in 30% of the cases without the FIM and 15% with the FIM. Moreover, the mean distance between the final true state of the system and the target is of 267m with the Lyapunov Fisher scheme, of 385m with the Penalising Fisher scheme and of 457m without using the FIM in the controller.

To stress the difference between the two dual controllers, we have tested them on a simple artificial map. Figure 6.3 depicts the results of the Penalising Fisher policy with 3 different tuning of α_k. Figure 6.3 represents the results of the Lyapunov Fisher policy in the same scenario also with different tuning. One see in 6.3a, if the weight α_k is constant with k, then the system converges far from the target. This does not happen with the Lyapunov Fisher scheme in any case. To sum up, Figure 6.3 and 6.4 show that in the Penalising Fisher scheme, the weight on the FIM affect both the information probing and the stability of the system in a complicated interconnected way. However, in the Penalising Fisher scheme, the weight on the FIM affect only the information probing and cannot prevent the system from eventually

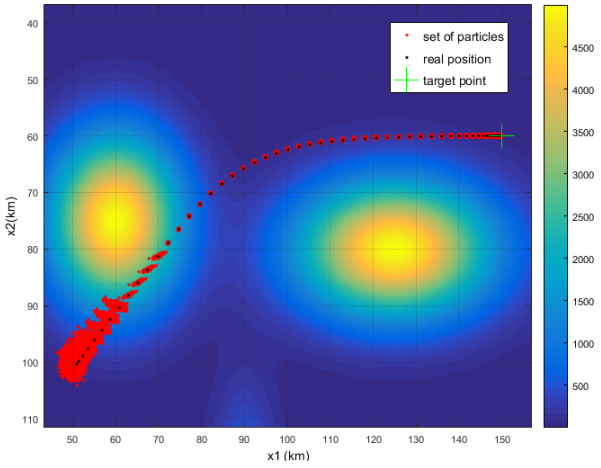
going to the target. Besides, the choice of λ determines the convergence speed of the system.

The simulations were run in MATLAB and the optimisation problems were solved using the modelling language AMPL and the solver Ipopt. Besides, thanks to user-defined functions in AMPL and a library in C for 3D-spline interpolation, we have managed to deal with the interpolation of a real map in AMPL. This is necessary for the computation of the FIM inside the optimisation problem.



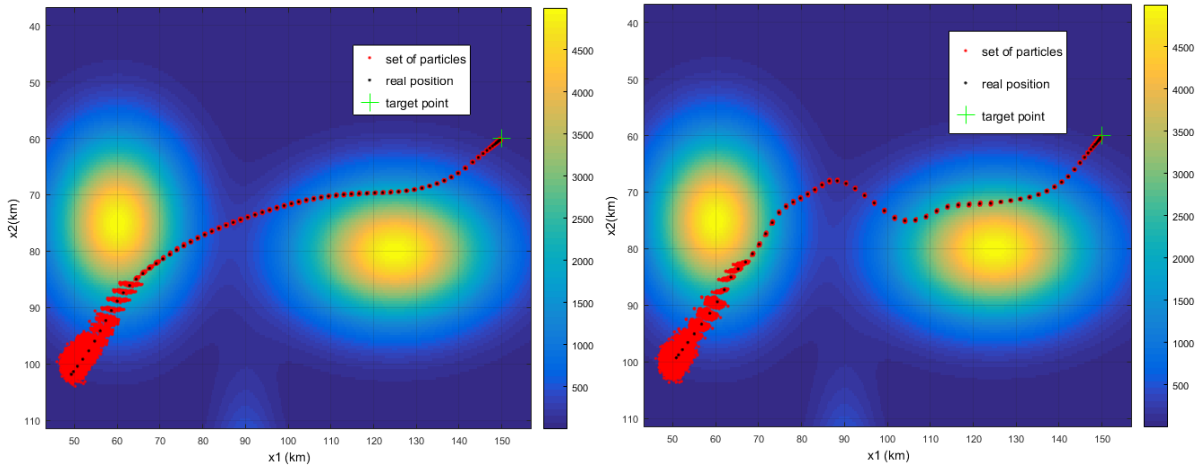
(a) Trajectory with a constant weight

(b) Trajectory with a slowly decreasing weight



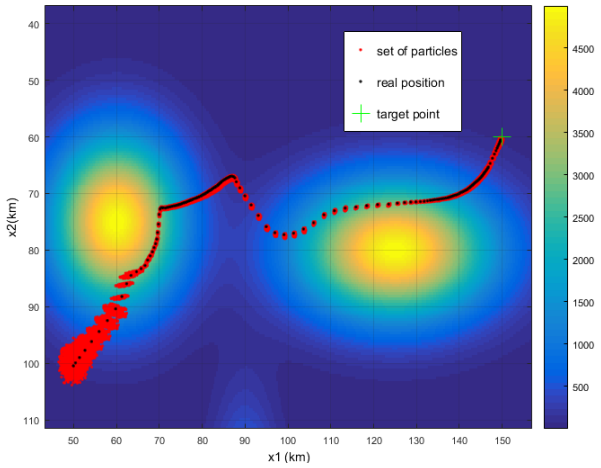
(c) Trajectory with a quickly decreasing weight

Figure 6.3: Realisation of a trajectory of the true state and the particles resulting from the Penalising Fisher controller on an artificial map with different tuning of α_k



(a) Trajectory with a small weight on the FIM and a small λ

(b) Trajectory with a large weight on the FIM and a small λ



(c) Trajectory with a large weight on the FIM and λ close to 1

Figure 6.4: Realisation of a trajectory of the true state and the particles resulting from the Lyapunov Fisher controller on an artificial map with and different tuning of the constant-in-time weight on the FIM and of λ

Chapter 7

Conclusion

The objective of this thesis is to design coupled control and estimation methods for nonlinear dynamical systems. The main target application is terrain-aided navigation, where the problem is to guide and estimate the 3D position of a drone flying over a known area. TAN is a typical example of nonlinear application where the separation principle cannot be applied. Therefore, there is a need for coupled estimation and control methods. It is to be noted that the estimation problem created by TAN is in itself difficult to analyse and solve. That is why, the contribution of this thesis is threefold:

- Nonlinear observability analysis and observer design for TAN with analytical ground maps has been proposed in a continuous time deterministic. It has been shown that under a condition of persistence of excitation on the horizontal speed, one can reconstruct the 3D position of the drone in several cases of ground maps (quadratic, cubic, Gaussian). The problem of output-feedback has also been studied using δ-persistence. Indeed, one needs the system to converge and to be persistent at the same time which leads to a complex nonlinear stabilisation problem.

- A general formalisation of the joint problem of nonlinear optimal filtering and discrete-time stochastic optimal control has been proposed. Under natural assumptions on the cost function one can justify the use of two steps in the resolution of the problem. The first step is to solve a classical optimal estimation problem. Near optimality of the empirical mean of a modified particle filter w.r.t. the mean square error has been shown in this thesis which justifies the use of particle filtering. The second step is to solve a modified optimal control problem with a new term coming from optimal estimation. This establishes a connection with explicit dual control where a new term representing a measure of information is empirically added to the cost. Actually, this empirical term can be seen as an approximation of the the term coming from optimal estimation.

- Following on the formalisation, two explicit dual stochastic MPC output-feedback controllers coupled with a particle filter have been designed and applied to TAN. Both are based on integrated experiment design but one contains with a penalisation of the guiding goal in the cost and the other one uses a drift constraint toward the target. The resulting optimisation problems are solved thanks to a Monte Carlo approximation method and both controllers have shown good results in simulation.

The presented work has left several perspectives that could be investigated in the future:

Perspective #1: On could foresee mixing the different models of analytical maps from Section 2.2.1 to approximate a real map and use the corresponding observers. For example, one may want to approach

the large scale variations of a real terrain with a Fourier approximation and its small scale variations with cubic splines. As for the simpler case, the idea would be to immerse the whole resulting system into a higher dimensional one and apply a Kalman observer.

Perspective #2: On could try to prove the output-feedback stability of the dual controllers in Section 6.1.3. It is a hard problem precisely because of the dual effect. Intuitively, one must ensure, in some way, that some stochastic observability properties are maintained by the controller which is already hard to make formal. However, the design of our Lyapunov controller has notably been inspired by the Lyapunov MPC scheme from [Homer and Mhaskar, 2017]. In [Homer and Mhaskar, 2017], the authors managed to prove the convergence of the output-feedback system by assuming the separation of the observer and the controller and the existence of Control Lyapunov function. One could apply a similar idea to show the convergence of the controller from Section 6.1.3 without assuming separation as it is not reasonable in this case. Actually, one could use the knowledge of a drift function for the system with perfect information and an assumption on the control forcing it to preserve the boundedness of the estimation error.

Perspective #3: In Section 5.2, one focuses on the near-optimality of the empirical mean w.r.t. the MSE. It could be interesting to study the more general problem (4.40) where the expectation in the cost is approximated by a particle filter. In [Chen et al., 1993], epi-convergence is used to show the convergence of solution of Monte Carlo approximations of an Open-Loop stochastic control problem to the optimal solution of the original problem. In [Vila and Gauchi, 2017], the previous result is extended to an Open-Loop Feedback SMPC involving the predicted distribution. Therefore, it seems promising to try to prove the convergence of the solution of the approximated estimation problem to the optimal estimator using epi-convergence techniques by adapting those results.

Perspective #4: The current numerical scheme in Algorithm 3 and 4 is rather simple and uses a convex programming solver in a heuristic way on a nonconvex problem. One could then work on improving the efficiency of these numerical methods. For example, to deal with the nonconvexity of the resulting nonlinear program more efficiently, one could try an Auxiliary-Function method. Besides, one could also incorporate a paralellisation scheme like decomposition/coordination methods, to speed up the Monte Carlo methods. Moreover, the stochastic optimisation problem (6.9) has a constraint in mean. Thus, an Arrow-Hurwitz technique potentially combined with a convex relaxation could be promising for the particular application of TAN.

Perspective #5: These methods have only been tested on an aerospace application but they seem well suited for any nonlinear adaptive control. In fact, adaptive control applications typically require a dual controllers to be able to estimate parameters. Moreover, the nonlinear identification part could be done by an adequately designed particle filter.

Bibliography

[Anderson, 1985] Anderson, B. D. (1985). Adaptive systems, lack of persistency of excitation and bursting phenomena. *Automatica*, 21(3):247–258.

[Anderson and Moore, 1979] Anderson, B. D. and Moore, J. B. (1979). Optimal filtering. *Englewood Cliffs*, 21:22–95.

[Andrieu and Praly, 2009] Andrieu, V. and Praly, L. (2009). A unifying point of view on output feedback designs for global asymptotic stabilization. *Automatica*, 45(8):1789–1798.

[Arnaud Doucet et al., 2001] Arnaud Doucet, de Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*. Springer New York, New York, NY.

[Astolfi et al., 2008] Astolfi, A., Karagiannis, D., and Ortega, R. (2008). *Nonlinear and Adaptive Control with Applications*. Communications and Control Engineering. Springer London, London.

[Bar-Shalom and Tse, 1974] Bar-Shalom, Y. and Tse, E. (1974). Dual Effect, Certainty Equivalence, and Separation in Stochastic Control. *IEEE Transactions on Automatic Control*, 19(5):494–500.

[Barabanov and Ortega, 2017] Barabanov, N. and Ortega, R. (2017). On global asymptotic stability of $dx/dt = \phi(t)\phi^T(t)x$ with ϕ not persistently exciting. *Systems & Control Letters*, 109:24–29.

[Barty et al., 2006] Barty, K., Chancelier, J.-P., Cohen, G., Lara, M. D., Guilbaud, T., and Carpentier, P. (2006). Dual effect free stochastic controls. *Annals of Operations Research*, 142(1):41–62.

[Bayard and Schumitzky, 2008] Bayard, D. S. and Schumitzky, A. (2008). Implicit dual control based on particle filtering and forward dynamic programming. *International Journal of Adaptive Control and Signal Processing*.

[Bellman, 1966] Bellman, R. (1966). Dynamic programming. *Science*, 153(3731):34–37.

[Ben-Tal et al., 2009] Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust optimization*. Princeton university press.

[Bertsekas, 1995] Bertsekas, D. P. (1995). *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA.

[Bertsekas, 2005a] Bertsekas, D. P. (2005a). *Dynamic programming and optimal control. Vol. 1: [...]*. Number 3 in Athena scientific optimization and computation series. Athena Scientific, Belmont, Mass, 3. ed edition. OCLC: 835987011.

[Bertsekas, 2005b] Bertsekas, D. P. (2005b). Dynamic programming and suboptimal control: A survey from ADP to MPC. *European Journal of Control*, 11(4-5):310–334.

[Bertsekas, 2011] Bertsekas, D. P. (2011). Dynamic programming and optimal control 3rd edition, volume II. *Belmont, MA: Athena Scientific*.

[Bertsekas and Shreve, 2004] Bertsekas, D. P. and Shreve, S. (2004). *Stochastic optimal control: the discrete-time case*.

[Besançon, 2007] Besançon, G., editor (2007). *Nonlinear observers and applications*. Number 363 in Lecture notes in control and information sciences. Springer, Berlin ; New York. OCLC: ocn166214280.

[Bilingsley, 2013] Bilingsley, P. (2013). *Convergence of probability measures*. John Wiley \& Sons.

[Bitmead, 1984] Bitmead, R. (1984). Persistence of excitation conditions and the convergence of adaptive schemes. *IEEE Transactions on Information Theory*, 30(2):183–191.

[Blackmore et al., 2010] Blackmore, L., Ono, M., Bektassov, A., and Williams, B. C. (2010). A Probabilistic Particle-Control Approximation of Chance-Constrained Stochastic Predictive Control. *IEEE Transactions on Robotics*, 26(3):502–517.

[Bolton, 2003] Bolton, W. (2003). *Mechatronics: electronic control systems in mechanical and electrical engineering*. Pearson Education.

[Calafiore and Campi, 2006] Calafiore, G. C. and Campi, M. C. (2006). The Scenario Approach to Robust Control Design. *IEEE Transactions on Automatic Control*, 51(5):742–753.

[Calafiore and Fagiano, 2013] Calafiore, G. C. and Fagiano, L. (2013). Robust Model Predictive Control via Scenario Optimization. *IEEE Transactions on Automatic Control*, 58(1):219–224.

[Chatterjee and Lygeros, 2015] Chatterjee, D. and Lygeros, J. (2015). On stability and performance of stochastic predictive control techniques. *IEEE Transactions on Automatic Control*, 60(2):509–514.

[Chen et al., 1993] Chen, L.-S., Geisser, S., and Geyer, C. J. (1993). Monte carlo minimization for sequential control. Technical report, University of Minnesota.

[Copp and Hespanha, 2014] Copp, D. A. and Hespanha, J. P. (2014). Nonlinear output-feedback model predictive control with moving horizon estimation. In *53rd IEEE Conference on Decision and Control*, pages 3511–3517. IEEE.

[Crisan, 2001] Crisan, D. (2001). Particle filter a theoretical perspective. In *Sequential Monte Carlo methods in practice*, pages 17–41. Springer.

[Crisan and Doucet, 2002] Crisan, D. and Doucet, A. (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on signal processing*, 50(3):736–746.

[Dahia, 2005] Dahia, K. (2005). *Nouvelles méthodes en filtrage particulaire-application au recalage de navigation inertielle par mesures altimétriques*. PhD thesis, Université Joseph Fourier.

[Del Moral and Guionnet, 1999] Del Moral, P. and Guionnet, A. (1999). Central limit theorem for nonlinear filtering and interacting particle systems. *Annals of Applied Probability*, pages 275–297.

[Del Moral and Jacod, 2002] Del Moral, P. and Jacod, J. (2002). The Monte-Carlo method for filtering with discrete-time observations: Central limit theorems. *Fields Institute Communications*, 34:29–53.

[Dovera and Della Rossa, 2011] Dovera, L. and Della Rossa, E. (2011). Multimodal ensemble Kalman filtering using Gaussian mixture models. *Computational Geosciences*, 15(2):307–323.

[Dupacovà et al., 2000] Dupacovà, J., Consigli, G., and Wallace, S. W. (2000). Scenarios for multistage stochastic programs. *Annals of operations research*, 100(1-4):25–53.

[Efimov et al., 2018] Efimov, D., Barabanov, N., and Ortega, R. (2018). Robust stability under relaxed persistent excitation conditions. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 7243–7248. IEEE.

[Ellis et al., 2014] Ellis, M., Durand, H., and Christofides, P. D. (2014). A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24(8):1156–1178.

[Eren et al., 2017] Eren, U., Prach, A., Koçer, B. B., Raković, S. V., Kayacan, E., and Açıkmeşe, B. (2017). Model Predictive Control in Aerospace Systems: Current State and Opportunities. *Journal of Guidance, Control, and Dynamics*, 40(7):1541–1566.

[Evensen, 2003] Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367.

[Fedorov and Hackl, 2012] Fedorov, V. V. and Hackl, P. (2012). *Model-oriented design of experiments*, volume 125. Springer Science & Business Media.

[Feldbaum, 1960] Feldbaum, A. (1960). The theory of dual Control, I-IV. *Automation remote control*, 2(11):1033.

[Flayac et al., 2017] Flayac, E., Dahia, K., Hérissé, B., and Jean, F. (2017). Nonlinear Fisher Particle Output Feedback Control and its application to Terrain Aided Navigation. In *56th IEEE Conference on Decision and Control*, Melbourne.

[Fu, 2015] Fu, M. C. a. o. (2015). *Handbook of simulation optimization*, volume 216. Springer.

[Gelb, 1974] Gelb, A. (1974). *Applied optimal estimation*. M.I.T. Press, Cambridge, Mass.

[Grune, 2013] Grune, L. (2013). Economic receding horizon control without terminal constraints. *Automatica*, 49(3):725–734.

[Grune and Rantzer, 2008] Grune, L. and Rantzer, A. (2008). On the Infinite Horizon Performance of Receding Horizon Controllers. *IEEE Transactions on Automatic Control*, 53(9):2100–2111.

[Hanssen and Foss, 2015] Hanssen, K. G. and Foss, B. (2015). Scenario Based Implicit Dual Model. In *5th IFAC Conference on Nonlinear Model Predictive Control NMPC 2015*, volume 48, pages 416–241, Seville, Spain. Elsevier.

[Heirung et al., 2015] Heirung, T. A. N., Foss, B., and Ydstie, B. E. (2015). MPC-based dual control with online experiment design. *Journal of Process Control*, 32:64–76.

[Heitsch and Römisch, 2009] Heitsch, H. and Römisch, W. (2009). Scenario tree reduction for multi-stage stochastic programs. *Computational Management Science*, 6(2):117–133.

[Hermann and Krener, 1977] Hermann, R. and Krener, A. (1977). Nonlinear controllability and observability. *IEEE Transactions on automatic control*, 22(5):728–740.

[Hernández-Lerma and Lasserre, 1996] Hernández-Lerma, O. and Lasserre, J. B. (1996). *Discrete-Time Markov Control Processes Basic Optimality Criteria*. Springer New York, New York, NY. OCLC: 869858938.

[Hespanha, 2009] Hespanha, J. P. (2009). *Linear systems theory*. Princeton University Press, Princeton. OCLC: ocn305421771.

[Hokayem et al., 2012] Hokayem, P., Cinquemani, E., Chatterjee, D., Ramponi, F., and Lygeros, J. (2012). Stochastic receding horizon control with output feedback and bounded controls. *Automatica*, 48(1):77–88.

[Homer and Mhaskar, 2017] Homer, T. and Mhaskar, P. (2017). Output-Feedback Lyapunov-based Predictive Control of Stochastic Nonlinear Systems. *IEEE Transactions on Automatic Control*, pages 1–1.

[Hovd and Bitmead, 2004] Hovd, M. and Bitmead, R. R. (2004). Interaction between control and state estimation in nonlinear MPC. *IFAC Proceedings Volumes*, 37(9):119–124.

[Hu et al., 2008] Hu, X.-L., Schon, T. B., and Ljung, L. (2008). A Basic Convergence Result for Particle Filtering. *IEEE Transactions on Signal Processing*, 56(4):1337–1348.

[Hu et al., 2011] Hu, X.-L., Schon, T. B., and Ljung, L. (2011). A General Convergence Result for Particle Filtering. *IEEE Transactions on Signal Processing*, 59(7):3424–3429.

[Julier and Uhlmann, 2004] Julier, S. J. and Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422.

[Karvonen, 2014] Karvonen, T. (2014). Stability of linear and non-linear Kalman filters.

[Khalil, 2002] Khalil, H. K. (2002). Nonlinear systems. *Upper Saddle River*.

[Khalil and Praly, 2014] Khalil, H. K. and Praly, L. (2014). High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24(6):993–1015.

[Kumar and Ahmad, 2012] Kumar, A. S. and Ahmad, Z. (2012). Model predictive control (mpc) and its current issues in chemical engineering. *Chemical Engineering Communications*, 199(4):472–511.

[La et al., 2017] La, H. C., Potschka, A., Schöder, J., and Bock, H. G. (2017). Dual control and online optimal experimental design. *SIAM Journal on Scientific Computing*, 39(4):B640–B657.

[Le Gland et al., 2004] Le Gland, F., Oudjane, N., and others (2004). Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *The Annals of Applied Probability*, 14(1):144–187.

[Lee and Jiang, 2008] Lee, T. and Jiang, Z.-P. (2008). Uniform Asymptotic Stability of Nonlinear Switched Systems With an Application to Mobile Robots. *IEEE Transactions on Automatic Control*, 53(5):1235–1252.

[Lee et al., 2014] Lee, T.-C., Tan, Y., and Nešić, D. (2014). Stability and Persistent Excitation in Signal Sets. *IEEE Transactions on Automatic Control*, 60(5):1188–1203.

[Li et al., 2015] Li, T., Bolic, M., and Djuric, P. M. (2015). Resampling Methods for Particle Filtering: Classification, implementation, and strategies. *IEEE Signal Processing Magazine*, 32(3):70–86.

[Li et al., 2014] Li, T., Sun, S., Sattar, T. P., and Corchado, J. M. (2014). Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches. *Expert Systems with applications*, 41(8):3944–3954.

[Loría et al., 2002] Loría, A., Panteley, E., and Melhem, K. (2002). UGAS of skew-symmetric time-varying systems: application to stabilization of chained form systems. *European Journal of Control*, 8(1):33–43.

[Loría et al., 1999] Loría, A., Panteley, E., and Teel, A. (1999). A new notion of persistency-of-excitation for UGAS of NLTV systems: Application to stabilisation of nonholonomic systems. In *Control Conference (ECC), 1999 European*, pages 1363–1368. IEEE.

[Maghenem and Loría, 2017] Maghenem, M. A. and Loría, A. (2017). Strict Lyapunov functions for time-varying systems with persistency of excitation. *Automatica*, 78:274–279.

[Malisoff and Mazenc, 2009] Malisoff, M. and Mazenc, F. (2009). *Constructions of strict Lyapunov functions*. Communications and control engineering. Springer, London ; New York. OCLC: ocn319210830.

[Mamon and Elliott, 2007] Mamon, R. S. and Elliott, R. J. (2007). *Hidden Markov models in finance*, volume 4. Springer.

[Marafioti et al., 2014] Marafioti, G., Bitmead, R. R., and Hovd, M. (2014). Persistently exciting model predictive control. *International Journal of Adaptive Control and Signal Processing*, 28(6):536–552.

[Martin and Sarras, 2018] Martin, P. and Sarras, I. (2018). Nonlinear attitude estimation from biased vector and gyro measurements. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 1385–1390. IEEE.

[Mayne, 2014] Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, 50(12):2967–2986.

[Mayne et al., 2000] Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Scokaert, P. O. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814.

[Melo and Matos, 2017] Melo, J. and Matos, A. (2017). Survey on advances on terrain based navigation for autonomous underwater vehicles. *Ocean Engineering*, 139:250–264.

[Mesbah, 2016] Mesbah, A. (2016). Stochastic Model Predictive Control: An Overview and Perspectives for Future Research. *IEEE Control Systems*, 36(6):30–44.

[Mesbah, 2017] Mesbah, A. (2017). Stochastic model predictive control with active uncertainty learning: A Survey on dual control. *Annual Reviews in Control*.

[Meyn and Tweedie, 2009] Meyn, S. P. and Tweedie, R. L. (2009). *Markov Chains and Stochastic stability*. Cambridge University Press, Cambridge, 2nd edition.

[Mishra et al., 2017] Mishra, P. K., Chatterjee, D., and Quevedo, D. E. (2017). Output Feedback Stable Stochastic Predictive Control With Hard Control Constraints. *IEEE Control Systems Letters*, 1(2):382–387.

[Mohler and Kolodziej, 1981] Mohler, R. and Kolodziej, W. (1981). Optimal control of a class of nonlinear stochastic systems. *IEEE Transactions on Automatic Control*, 26(5):1048–1054.

[Murangira, 2014] Murangira, A. (2014). *Nouvelles approches en filtrage particulaire. Application au recalage de la navigation inertielle*. PhD thesis, Université de Technologie de Troyes-UTT.

[Narendra and Annaswamy, 2012] Narendra, K. S. and Annaswamy, A. M. (2012). *Stable adaptive systems*. Courier Corporation.

[Nocedal and Wright, 2006] Nocedal, J. and Wright, S. J. (2006). *Numerical optimization*. Springer series in operations research. Springer, New York, 2nd ed edition. OCLC: ocm68629100.

[Oudjane, 2000] Oudjane, N. (2000). *Stabilite et approximations particulaires en filtrage non lineaire application au pistage*. PhD thesis, Université de Rennes 1.

[Panteley et al., 2001] Panteley, E., Loria, A., and Teel, A. (2001). Relaxed Persistency of Excitation for Uniform Asymptotic Stability. *Transaction on Automatic Control*, 46(12).

[Qin and Badgwell, 2003] Qin, S. J. and Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control engineering practice*, 11(7):733–764.

[Åström, 2012] Åström, K. J. (2012). *Introduction to stochastic control theory*. Courier Corporation.

[Reif et al., 1999] Reif, K., Gunther, S., Yaz, E., and Unbehauen, R. (1999). Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic control*, 44(4):714–728.

[Sabatino et al., 2003] Sabatino, A. E., Romero, L. G., Pyne, M. D., Martinez, S. D., and Cox, P. A. (2003). Method and system for terrain aided navigation.

[Sehr and Bitmead, 2016] Sehr, M. A. and Bitmead, R. R. (2016). Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC. *arXiv preprint arXiv:1612.00505*.

[Sehr and Bitmead, 2017a] Sehr, M. A. and Bitmead, R. R. (2017a). Stochastic model predictive control: Output-feedback, duality and guaranteed performance. *arXiv preprint arXiv:1706.00733*.

[Sehr and Bitmead, 2017b] Sehr, M. A. and Bitmead, R. R. (2017b). Tractable dual optimal stochastic model predictive control: An example in healthcare. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 1223–1228. IEEE.

[Stettner, 1989] Stettner, L. (1989). On invariant measures of filtering processes. In *Stochastic Differential Systems*, pages 279–292. Springer.

[Subramanian et al., 2014] Subramanian, S., Lucia, S., and Engell, S. (2014). Economic Multi-stage Output Nonlinear Model Predictive Control. In *2014 IEEE Conference on Control Applications (CCA)*, pages 1837–1842. IEEE.

[Subramanian et al., 2015] Subramanian, S., Lucia, S., and Engell, S. (2015). Economic Multi-stage Output Feedback NMPC using the Unscented Kalman Filter. *IFAC-PapersOnLine*, 48(8):38–43.

[Subramanian et al., 2016] Subramanian, S., Lucia, S., and Engell, S. (2016). A non-conservative robust output feedback MPC for constrained linear systems. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 2333–2338. IEEE.

[Telen et al., 2017] Telen, D., Houska, B., Vallerio, M., Logist, F., and Van Impe, J. (2017). A study of integrated experiment design for NMPC applied to the Droop model. *Chemical Engineering Science*, 160:370–383.

[Tenno and Nõmm, 2011] Tenno, R. and Nõmm, S. (2011). Partially observed stochastic bilinear system control. In *2011 9th IEEE International Conference on Control and Automation (ICCA)*, pages 951–956. IEEE.

[Tichavsky et al., 1998] Tichavsky, P., Muravchik, C. H., and Nehorai, A. (1998). Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on signal processing*, 46(5):1386–1396.

[Tong et al., 2016] Tong, X. T., Majda, A. J., and Kelly, D. (2016). Nonlinear stability and ergodicity of ensemble based Kalman filters. *Nonlinearity*, 29(2):657.

[Vidaurre et al., 2013] Vidaurre, D., Bielza, C., and Larrañaga, P. (2013). A survey of L1 regression. *Wiley Online Library*, 81:361–387.

[Vila and Gauchi, 2017] Vila, J.-P. and Gauchi, J.-P. (2017). Predictive Control of Discrete Time Stochastic Nonlinear State Space Dynamical Systems: A Particle Nonparametric Approach. *Applied Mathematics & Optimization*.

[Weissel et al., 2008] Weissel, F., Schreiter, T., Huber, M. F., and Hanebeck, U. D. (2008). Stochastic model predictive control of time-variant nonlinear systems with imperfect state information. In *Multi-sensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pages 40–46. IEEE.

[Whiteley, 2013] Whiteley, N. (2013). Stability properties of some particle filters. *The Annals of Applied Probability*, 23(6):2500–2537.

[Xiong et al., 2006] Xiong, K., Zhang, H., and Chan, C. (2006). Performance evaluation of UKF-based nonlinear filtering. *Automatica*, 42(2):261–270.

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Méthodes d'estimation et de contrôle nonlinéaires couplées avec application à la navigation par corrélation de terrain

Mots Clefs : estimation optimale, commande optimale stochastique, filtrage particulaire, commande prédictive, observateurs nonlinéaires, navigation par corrélation de terrain

Résumé : Dans cette thèse, nous étudions les problèmes de contrôle et d'estimation non linéaire couplés avec application à la navigation par corrélation de terrain (TAN en anglais). L'objectif est de guider et d'estimer la position 3D d'un drone survolant une zone connue. La principale difficulté est la nature de l'information disponible sur le système. En effet, les seules données disponibles sont la vitesse du système, une mesure de hauteur/sol et une carte de la zone survolée. Habituellement, les problèmes d'estimation et de contrôle sont résolus séparément en invoquant le principe de séparation. Cependant, il n'est applicable que pour des classes de systèmes très particulières, dont les systèmes linéaires. Ainsi, pour un système non linéaire général, le contrôle et l'estimation doivent être gérés conjointement. La TAN est une application non linéaire où le principe de séparation ne peut pas être appliqué. En réalité, la qualité des observations dépend du contrôle et plus précisément de la zone survolée par le drone. En conséquence, les principales contribution de cette thèse sont la conception d'observateurs non linéaires et contrôle en retour de sortie pour la TAN avec des cartes analytiques au sol dans un cadre temporel continu, la reformulation du problème couplé du filtrage optimal non linéaire et de contrôle optimal stochastique au sein d'un problème d'optimisation global et la conception de schémas de commande prédictive duale explicite couplés à un filtre particulaire et leur implémentation numérique pour la TAN.

Title : Coupled methods of nonlinear control and estimation applicable to Terrain-Aided Navigation

Keys words : optimal estimation, stochastic optimal control, particle filtering, model predictive control, nonlinear observers, terrain-aided navigation

Abstract : In this work, we study nonlinear coupled control and estimation problems in order to solve Terrain-Aided Navigation (TAN). The objective is to guide and estimate the 3D position of a drone flying over a known area. The main difficulty of this application is the nature of the available information on the system. Indeed, it is assumed that the only available data are the speed of the system, a measurement of the height from the ground and a map of the area flied over. Usually, estimation and control problems are solved separately invoking the classical separation principle. However, it is known to hold true only for very particular classes of systems. Thus, for a general nonlinear system, control and estimation must be handled jointly. TAN is a good example of nonlinear application where the separation principle cannot be applied. Actually, the quality of the observations depends on the control and more precisely on the area that is flied over by the drone. The main contribution of the thesis are the design of nonlinear observers and output-feedback control laws for TAN with analytical ground maps, the reformulation of the joint problem of nonlinear optimal filtering and discrete-time stochastic optimal control into one optimisation problem, the design of output-feedback Explicit dual stochastic MPC schemes coupled with a nonlinear filter and their numerical implementation to TAN.