



Département d'Electronique

Mémoire de Magister

Option : Traitement du signal & Communication

Présenté par : DRIS Lyamine

Thème

LES ALGORITHMES ADAPTATIFS DE POURSUITE DE CIBLE

Membres du Jury

M. GUERTI M'hania, Pr, ENP
M. BERKANI Daoud, Pr, ENP
Mme. HAMAMI Latifa, MC, ENP
M. LARBES Chérif, MC, ENP
M. ZERGUI Rachid, CC, ENP

Présidente
Rapporteur
Examinatrice
Examinateur
Examinateur

- 2008-

ملخص:

هذه المذكرة تخص الخوارزميات المتعلقة بمتابعة الأهداف. اهتمامنا يتمحور حول الأهداف المتحركة وبالأخص تقدير حركات هذه الأهداف, اللازمة للتطبيق الكافي لحالة المتابعة. نوعية المتابعة تكون أحسن بكثير بفضل التقديرات المتينة للحركة. من أجل إنجاز التقدير, اخترنا تقنيات الترشيح الإحصائي الذي يتعلق بمرشح كالمان في الحالة الخطية القوسية أو بطرق تعاقبية لمنتي كارلو في الحالة الغير الخطية. تمثيل الحالة مكيف دائما بدلالة الملاحظات الدائمة وهذا من أجل تمثيل أحسن لدناميكيات النظام. في هذه المذكرة, عملنا يهدف إلى مقارنة نتائج المعطيات بواسطة مرشح كالمان المعمم و مرشح منتي كارلو في إطار متابعة الهدف. الطرق المقدمة في هذه المذكرة تم تأكيدها بالتمثيل على الحاسوب. كلمات مفتاحيه: مرشح إحصائي, مرشح كالمان, مرشح كالمان المكيف, طرق منتي كارلو, مرشح منتي كارل التعاقبي, متابعة الهدف المتحرك.

Résumé :

Le travail de ce mémoire concerne plus particulièrement les algorithmes de poursuite de cible.

Nous nous sommes intéressé au cas d'objets mobiles, et plus particulièrement à l'estimation du mouvement de ces objets, nécessaire pour une exécution satisfaisante des tâches de poursuites. La qualité de poursuite est grandement améliorée grâce à une estimation robuste du mouvement. Pour réaliser l'estimation, nous avons choisi les techniques de filtrage stochastique. Il s'agit de filtre de Kalman dans le cas linéaire gaussien, ou des méthodes séquentielles de Monte Carlo dans le cas non linéaire. La représentation d'état est adaptée en permanence en fonction des observations courantes pour représenter au mieux la dynamique du système. Dans ce mémoire, notre travail consiste à faire de la comparaison des résultats donnés par le filtre de Kalman étendu et le filtre particulaire dans le cadre de la poursuite de cible.

Les méthodes présentées dans ce mémoire ont été validées en simulation.

Mots clés : Filtrage stochastique, filtre de Kalman, Filtre de Kalman adaptatif, Méthode de Monte Carlo, Filtre particulaire, Poursuite de cible manoeuvrante.

Summary :

The work of this memory more particularly relates to the algorithms of tracking target.

We were interested in the mobile cases of objects, and more particularly in the estimate of the movement of these objects, necessary for a satisfactory execution of the spots of tracking. The quality of tracking is largely improved thanks to a robust estimate of the movement. To carry out the estimate, we chose the techniques of stochastic filtering. It is about Kalman filter in the Gaussian linear case or of the sequential methods of Monte Carlo in the nonlinear case. The representation of state is adapted permanently according to the current observations to represent the dynamics of the system as well as possible. In this memory, our work consists in making comparison of the results given by the extended Kalman filter and the particle filter within the framework of the tracking target.

The methods presented in this memory were validated in simulation.

Key words: Stochastic filtering, Kalman filter, adaptive Kalman filter, Monte Carlo method, Particle Filter, Manoeuvring tracking target.

Remerciements

Je tiens vivement à remercier Monsieur **BERKANI Daoud**, professeur à l'ENP, pour avoir proposé un sujet aussi intéressant, pour la confiance et l'intérêt qu'il m'a témoigné tout au long de la réalisation de ce travail.

Je tiens à remercier Madame **GUERTI M'hania, Pr, ENP**, pour m'avoir fait l'honneur de présider le jury.

Je tiens également à remercier Madame **HAMAMI Latifa, MC, ENP**, pour l'honneur qu'elle me fait en participant au jury.

Je tiens également à remercier Monsieur **LARBES Chérif, MC, ENP**, pour l'honneur qu'il me fait en participant au jury.

Mes remerciements vont également à Monsieur **ZERGUI Rachid, CC, ENP**, pour l'honneur qu'il me fait en participant au jury.

Je tiens également à remercier tous ceux qui de près ou de loin ont participé à l'élaboration de ce travail et plus particulièrement à mes amis Mabrouk, Nassim, Redha, Larbi et ne compter que ces quatre.

Je voudrais remercier tout particulièrement ma chère mère ainsi que mes frères, mes deux sœurs et Remdhan, le mari de ma sœur, pour leurs soutiens et leurs encouragements tout au long de mon cursus. Je le dédie également à mon défunt père

Résumé

Remerciement

Table des Matières

Liste des Figures

Liste des Algorithmes

Liste des Notations générales

Liste des Abréviations

Introduction générale

Chapitre I – Le filtrage : état d’art des méthodes analytiques et numériques

Introduction.....	1
I. Description du problème de filtrage.....	1
Modèle général d’état et d’observation considéré.....	3
I.1. Evolution du filtrage optimal.....	5
II. Etat de l’art des méthodes classiques d’approximations.....	7
II.1. Méthodes analytiques.....	7
II.1.1. Filtre de Kalman-Bucy (KF).....	8
II.1.2. Filtre de Kalman Etendu (EKF).....	12
II.1.3. Filtre de Kalman sans parfum (UKF).....	15
II.2. Méthodes Numériques.....	16
III. Conclusion.....	17

Chapitre II – Méthodes de Monte Carlo pour le filtrage non linéaire

Introduction	18
I. Principe des méthodes d’approximation de Monte Carlo.....	18
I.1. Le principe de l’intégration de Monte Carlo.....	18
I.2. Méthode de simulation de variables aléatoires.....	19
I.2.1. Méthode d’acceptation / rejet.....	20
I.2.2. Echantillonnage pondéré.....	20
II. Application à la problématique du filtrage.....	23

II.1. Echantillonnage pondéré.....	23
II.2. Echantillonnage pondéré bayésien.....	24
II.3. Echantillonnage d'importance bayésien séquentiel (SIS).....	27
II.4. Choix d'une fonction d'importance adéquate.....	30
II.5. Techniques de rééchantillonnage.....	32
II.5.1. Rééchantillonnage multinomial.....	33
II.5.2. Rééchantillonnage résiduel.....	34
II.5.3. Rééchantillonnage à variance minimale.....	35
II.6. Echantillonnage pondéré séquentiel avec rééchantillonnage (SIR).....	35
II.6.1. Inconvénient du rééchantillonnage.....	38
II.7. Echantillonnage pondéré séquentiel avec rééchantillonnage adaptatif.....	39
II.8. Le Bootstrap.....	41
II.9. Limitation des méthodes SMC.....	43
II.10 Variantes du SIR.....	44
II.10.1. Echantillonnage pondéré séquentiel à variable auxiliaire (APF).....	44
II.10.2. Echantillonnage pondéré séquentiel régularisé (RPF).....	47
II.10.3. La Rao-Blackwellisation.....	51
III.3. Conclusion.....	52
Chapitre III - Modélisation des trajectoires d'une cible	53
Introduction	53
I. Modèle d'une droite (Trajectoire rectiligne).....	53
I.1. Mouvement rectiligne à vitesse constante.....	54
I.2. Mouvement rectiligne à accélération constante.....	56
II. Modèle d'un cercle (Trajectoire circulaire).....	57
II.1. Mouvement circulaire à vitesse constante.....	58
II.2. Mouvement circulaire à accélération constante.....	60
III. Conclusion.....	62
Chapitre IV- Application au problème de la poursuite radar	63
Introduction	63
I. Modèle d'état.....	63
I.1. poursuite radar.....	64

II. Simulation.....	65
III. Conclusion.....	72

Conclusion générale

Bibliographie

Liste des figures :

N° figure	Titre	Pages
Figure I-1	Graphe de dépendance de la chaîne de Markov cachée.	5
Figure I-2	Le schéma classique de la décomposition d'une itération d'un filtre non linéaire optimal.	6
Figure I-3	un simple exemple de la fonction de vraisemblance et de la densité prédite pour une observation précise ou incohérente par rapport à la prédiction.	7
Figure II-1	Mesure discrète pondérée de la probabilité a posteriori.	27
Figure II-2	Echantillonnage multinomial.	33
Figure II-3	Dégénérescence des particules : faible bruit d'état.	37
Figure II-4	Dégénérescence des particules : faible bruit d'observation.	37
Figure II-5	Evolution de la densité conditionnelle.	40
Figure II-6	Le principe de Bootstrap.	42
Figure II-7	Configuration critique de la densité a priori et la vraisemblance.	43
Figure II-8	Mélange de gaussiennes.	48
Figure III-1	Modélisation d'une droite.	53
Figure III-2	Modélisation d'un cercle.	57
Figure IV-1	La trajectoire vraie, estimée de la cible par l'EKF.	66
Figure IV-2	La position vraie, estimée suivant l'axe X et l'axe Y par l'EKF.	66
Figure IV-3	L'erreur entre X et Xest, Y et Yest.	66
Figure IV-4	L'angle vrai, estimé par SIR adaptatif avec $N = 150$ et $N_{Seuil} = 0.5 N$.	67
Figure IV-5	La distance vraie, estimée par SIR adaptatif avec $N = 150$ et $N_{Seuil} = 0.5 N$.	67
Figure IV-6	L'angle vrai, estimé par SIR systématique avec $N = 150$.	68
Figure IV-7	La distance vraie, estimés par SIR systématique avec $N = 150$.	68
Figure IV-8	L'angle vrai, estimé par SIR adaptatif avec $N = 500$ et $N_{Seuil} = 0.5 N$	69
Figure IV-9	La distance vraie, estimée par SIR adaptatif avec $N = 500$ et $N_{Seuil} = 0.5 N$.	69
Figure IV-10	L'angle vrai, estimé par SIR systématique avec $N = 500$.	69
Figure IV-11	La distance vraie, estimée par SIR systématique avec $N = 500$.	70
Figure IV-12	L'angle vrai, estimé par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.5 N$.	70
Figure IV-13	La distance vraie, estimée par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.5 N$.	71
Figure IV-14	L'angle vrai, estimé par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.8 N$.	71
Figure IV-15	La distance vraie, estimé par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.8 N$.	71

Liste des Algorithmes :

N° Algorithme	Titre	Pages
Alg. I-1	Filtre de Kalman pour le système linéaire/gaussien.	12
Alg. I-2	Filtre de Kalman Etendu pour le système non linéaire.	14
Alg. II-1	Générateur d'une loi $p(X)$ par méthode d'acceptation / rejet.	20
Alg. II-2	Algorithme de filtrage par l'échantillonnage pondéré séquentiel SIS.	29
Alg. II-3	Algorithme d'échantillonnage pondéré séquentiel avec rééchantillonnage systématique.	36
Alg. II-4	Algorithme d'échantillonnage pondéré séquentiel avec rééchantillonnage adaptatif.	40
Alg. II-5	L'algorithme Bootstrap.	42
Alg. II-6	L'algorithme ASIR.	47
Alg. II-7	Algorithme d'échantillonnage pondéré séquentiel régularisé.	50

Liste des Notations générales :

$X_{0:k} = (X_0, X_1, \dots, X_k)^T$	Trajectoire du processus markovien à estimer
$Y_{1:k} = (Y_1, Y_2, \dots, Y_k)^T$	Vecteur de mesure jusqu'à l'instant k
k	Le temps discrétisé
W_k	Bruit du système à l'instant k
V_k	Bruit de mesure à l'instant k
\hat{x}_k	Estimateur de l'état X_k
(Ω, f, P)	Espace probabilisé sur lequel sont définies les variables aléatoires utilisées
$Q_k(x_{k-1}, dx)$	Noyau de transition de X_k
$g_k(x, Y_k)$	Fonction de la vraisemblance associée
$E_p[\cdot]$	L'espérance par rapport à la densité p
$p_k(dx) = p(X_k Y_{1:k})$	Densité conditionnelle
$p_{k k-1}(dx) = p(X_k Y_{1:k-1})$	Densité prédite
F_k	Fonction d'évolution du vecteur d'état, non linéaire en général
H_k	Fonction de mesure reliant une mesure Y_k à l'état dont elle dépend
Q_k	Matrice de covariance du bruit dynamique
R_k	Matrice de covariance du bruit de mesure
$\nu_k = Y_k - \hat{Y}_{k k-1}$	Processus innovation
N	Le nombre de particules
$\delta_{X_{0:k}^i}$	Masse de Dirac située en $X_{0:k}^i$
w_k^i	Poids des particules X_k^i

\tilde{w}_k^i	Poids des particules X_k^i normalisés
$U[a, b]$	Loi de probabilité uniforme sur l'intervalle $[a, b]$
$N(0,1)$	Loi normale centrée réduite
$q(X)$	Densité d'importance
$p_N(X)$	L'approximation particulière de la loi $p(X)$
$\xrightarrow{N \rightarrow \infty}$	Convergence en probabilité
N_{eff}	La taille effective de l'échantillon
α	L'angle que fait la droite avec l'axe des abscisses
s	La distance parcourue le long de la droite
v	La vitesse
a	L'accélération
X_0	L'abscisse du centre du cercle
Y_0	L'ordonnée du centre du cercle
R_0	Rayon du cercle

Listes des Abréviations :

KF	Filtre de Kalman, Kalman Filter
EKF	Filtre de Kalman étendu, Extended Kalman Filter
UKF	Filtre de Kalman “Unscented“, Unscented Kalman Filter
FP	Filtre particulaire, Particle Filter
MSMC	Méthodes Séquentielles de Monte Carlo
SIS	Echantillonnage pondéré séquentiel, Sequential Importance Sampling
SIR	Echantillonnage pondéré séquentiel avec rééchantillonnage, Sequential Importance Resampling
SMC	Méthode de Monte Carlo séquentiel, Sequential Monte carlo
MCMC	Markov Chains Monte Carlo
APF	Echantillonnage pondéré séquentiel à variable auxiliaire, Auxiliary Particle Filter
RPF	Echantillonnage pondéré séquentiel régularisé, Regularized Particle Filter
MISE	Mean Integrated Square Error
i.i.d.	Independent Identically distributed
HMM	Modèles de Markov caché, Hidden Markov Models

Introduction Générale

Dans ce mémoire de magister, nous nous intéressons aux problèmes de filtrage. Considérons un système dynamique c'est-à-dire un système d'équations différentielles qui régit l'évolution (déterministe ou stochastique) d'un état inconnu au cours du temps. L'état peut représenter par exemple les paramètres cinématiques d'un mobile quelconque (position, vitesse, accélération,...). L'observateur dispose de mesures partielles entachées d'erreurs (mesures indirectes d'une partie du vecteur d'état) comme des mesures de distance ou d'angle. Le filtrage consiste à estimer, à restituer l'état d'un système dynamique au vu d'observations bruitées des états passés du système. Notons qu'il peut s'appliquer dans de nombreux domaines telles que l'évaluation de risque en finance, la localisation d'objet en robotique, le suivi de contours d'images et comme dans notre cas le pistage de cible (estimation d'une trajectoire d'un mobile en mouvement). C'est ce contexte applicatif riche qui motive pour l'essentiel la recherche dans ce domaine. Nous avons choisi de considérer le cas particulier du filtrage en temps discret et ce choix est naturellement justifié par le caractère nécessairement échantillonné des observations délivrées par les appareils de mesure. Précisément, nous cherchons à estimer, à chaque instant d'observation, la valeur d'un processus aléatoire markovien, de transitions connues, appelé signal ou processus d'état, indirectement lié à un processus observation dont on connaît une réalisation. Le but est alors de calculer (de manière approchée), à chaque instant d'observation, la distribution de probabilité conditionnelle de l'état à l'instant courant sachant les observations passées. La suite de mesures de probabilité ainsi formée est appelée filtre optimal. Elle évolue dans le temps et dans le cadre bayésien par une simple alternance d'étapes de prédictions et de correction. En filtrage, on cherche un algorithme permettant d'effectuer le calcul du filtre optimal qui satisfasse en outre des contraintes de temps réel. Concrètement, le temps de calcul doit être très court (de l'ordre du centième de seconde en pistage). Pour satisfaire la contrainte de temps et économiser la mémoire de l'ordinateur, on cherche de préférence un algorithme récursif. Ainsi, le calcul de l'estimée à l'instant courant ne doit dépendre que de l'observation courante et de l'estimée précédente. Ces contraintes sont primordiales et caractérisent, pour une précision d'estimation donnée, l'efficacité de l'algorithme proposé.

En 1964 Kalman et Bucy [1, 2] révolutionnent la théorie de l'estimation en fournissant le premier algorithme de filtrage récursif. Le filtre de Kalman-Bucy permet le calcul exact et rapide du filtre optimal lorsque les modèles d'état et d'observation ne font intervenir que des fonctions linéaires et des bruits additifs gaussiens. On montre, en fait, que dans ce cas la loi conditionnelle recherchée est gaussienne. Il suffit alors de propager sa moyenne et sa variance ; le problème est de dimension finie [3, 4]. En pratique, on continue à utiliser, dans ce cas, des versions dérivées du filtre de Kalman-bucy, comme le filtre de Kalman étendu qui linéarise le modèle autour de l'estimée courante. Bien que cet algorithme soit depuis une trentaine d'années le filtre non linéaire le plus utilisé, il n'a été justifié rigoureusement que très récemment grâce en particulier aux travaux de Picard [5] qui montre son efficacité dans certains cas d'observations très précise. Hors de ces cas, l'erreur commise par le filtre de Kalman étendu n'est en général pas connue. On observe d'ailleurs des résultats mauvais lorsque le système est trop non linéaire ou lorsque le filtre est mal initialisé. Une autres approche plus rigoureuse propose de calculer la solution des équations du filtrage par des techniques de maillage de l'espace d'état [6, 7]. Cependant la complexité de ces méthodes numériques augmentant fortement avec la dimension de l'espace les rend inutilisables en temps réel, à partir de la dimension trois. Plus récemment, les méthodes de Monte Carlo ont été proposées comme une alternative attrayante [8, 9, 10]. Connues pour leur quasi-

insensibilité à la dimension de l'espace et aux non linéarités du systèmes, leur utilisation semble particulièrement adaptée au filtrage non linéaire. Les méthodes particulières sont une version séquentielle des méthodes de Monte Carlo pour résoudre le filtrage. Elles ont été introduites par Del Moral, Rigal, Salut [9], Gordon, Salmond et Smith [8]. Elles proposent de représenter la loi conditionnelle de l'état par un nombre fini de masses de Dirac pondérées. Un ensemble de points appelés particules est généré, chacune de ces particules représente un état probable du système. Les coefficients de pondération (poids) sur chaque particule sont une mesure de degré de confiance que l'on peut avoir en ces dernières pour représenter effectivement l'état. Les particules évoluent suivant l'équation d'état du système (étape de prédiction) et les poids sont ajustés en fonction des observations (étape de correction).

Cependant, le filtre particulaire ainsi décrit a un défaut majeur ; les poids des particules ont tendance à dégénérer de sorte que, après un certain nombre de mesures, la plupart des particules ont un poids négligeable. Ce phénomène est connu sous le nom de dégénérescence des poids. Le système de particules est appauvri et donc ne peut plus représenter correctement la densité conditionnelle. Le filtre diverge. Pour éviter ce phénomène, une nouvelle étape dite de rééchantillonnage a été introduite. Celle-ci conduit à dupliquer les particules de poids fort et à éliminer les particules de poids faible. Plusieurs versions particulières ont été proposées dans la littérature comme ; le Sampling Importance Resampling Filter (SIR) proposé par Gordon [8], Auxiliary Particle Filter (APF) proposé par Pitt et Sheppard [11] et Regularized Particle Filter (RPF) développé par Musso et Oudjane [12]. Malgré ces améliorations on observe parfois la divergence de ces filtres particuliers. Ces problèmes sont dus aux approximations Monte Carlo en cascade aussi bien dans l'évaluation des intégrales que dans l'étape de rééchantillonnage.

Nous avons choisie l'application poursuite de cible pour mettre en œuvre les méthodes de filtrage décrites ci-dessus. Une comparaison entre les méthodes particulières et le EKF est réalisée.

Ce mémoire est organisé en cinq chapitres selon le plan suivant :

Dans le premier chapitre, on présente le problème de filtrage non linéaire, on décrit un état d'art des différentes méthodes analytiques (filtre de Kalman –Bucy, filtre de Kalman étendu, ...) et numériques classiquement utilisées dans le filtrage. Nous montrons les limites de ces méthodes de résolution.

Dans le second chapitre, on décrit le principe fondamental des méthodes de Monte Carlo et on compare certain des ces méthodes appliquées dans le filtrage non linéaire. La liste de ces méthodes n'est pas exhaustive. Nous montrons certaines faiblesses des méthodes particulières classiques. Nous présentons des variantes de ce filtre qui répondent aux contraintes.

Dans le troisième chapitre, nous décrivons les différents modèles de trajectoires des systèmes dynamiques non linéaires. La majorité des trajectoires ne sont que des successions de trajectoires élémentaires canoniques tels que des segments de droites, des arcs de cercles,

Dans le dernier chapitre, nous abordons la problématique du la poursuite de cible. Une mise en œuvre de filtres particuliers pour la poursuite d'une cible et une comparaison aux résultats d'un filtre de Kalman étendu clôt ce chapitre.

Une conclusion générale et le dégagement de plusieurs perspectives à ces travaux clôt ce document.

Introduction

Dans ce chapitre, nous commençons par exposer le problème du filtrage dans le cas général. Les différentes approches classiques utilisées pour résoudre ce problème sont décrites. Deux grandes classes de méthodes sont présentées : tout d'abord les méthodes analytiques, à savoir le célèbre filtre de Kalman et ses dérivées, le filtre de Kalman étendu et le filtre de Kalman unscented, puis la méthode numérique comme la méthode de maillage de l'espace d'état. Nous montrons les limites et faiblesses de ces méthodes.

En 1949, N. Wiener pose explicitement le problème du filtrage d'un signal stochastique. Il propose une méthode originale, qui se base sur des filtres linéaires minimisant une intégrale du carré de l'erreur entre le signal utile et le signal estimée [13]. A. N. Kolmogorov obtient indépendamment les mêmes résultats.

I. Description du problème de filtrage :

La recherche en filtrage est motivée par de nombreuses applications dans des domaines varies. Ce problème est en effet assez général puisqu'il consiste à déterminer des estimateurs des variables d'un système dynamique (variables d'états) sujet à des perturbations et observées partiellement. La modélisation de l'évolution dans le temps des états $(X_t)_{t \geq 0}$ du système considéré s'écrit alors sous la forme d'une partie d'évolution déterministe et d'une partie stochastique,

$$X_t = X_0 + \int_0^t f(s, X_s, W_s) ds, \quad (1.1)$$

où X_0 est de loi donnée, f est une fonction déterministe caractérisant la part déterministe de la dynamique et $(W_t)_{t \geq 0}$ est un processus de Wiener standard modélisant les perturbations aléatoires de la dynamique et notre méconnaissance du modèle. Cette équation de dynamique traduit la connaissance à priori que l'on a du système. Le processus X_t , (1.1) est un processus Markovien, on désigne par Q_k son noyau de transition. Dans le but de déterminer l'état précis du système, on est amené à construire une équation d'observation qui relie, à des instants k , l'observation Y_k à l'état courant X_k . Ces mesures sont, en général, entachées d'erreurs dues à l'imperfection du capteur de mesure. La suite des observations $(Y_k)_{k \geq 1}$ est modélisée par l'équation suivante,

$$Y_k = h_k(X_k) + V_k, \quad (1.2)$$

où h_k est la fonction d'observation connue et $(V_k)_{k \geq 0}$ une suite de variables aléatoires (v.a.) de statistique connue modélisant l'imperfection des observations. On cherche à déterminer l'état du système X_k à partir d'observations bruitées (Y_1, \dots, Y_n) . Suivant la valeur de n , on distingue :

- ✓ Si $n > k$, il s'agit d'un problème de lissage,
- ✓ Si $n = k$, il s'agit d'un problème de filtrage,
- ✓ Si $n < k$, il s'agit d'un problème de prédiction,

Afin d'assurer le caractère markovien du processus d'état, le bruit $(W_t)_{t \geq 0}$ est supposé être un bruit blanc et l'indépendance conditionnelle des observations est assurée par le fait que $(V_k)_{k \geq 0}$ soit également un bruit blanc, indépendant $(W_t)_{t \geq 0}$.

Les hypothèses classiques du filtrage non linéaire sont :

- ✓ Les observations $\{Y_k, k \geq 1\}$ sont mutuellement indépendantes, conditionnellement à $\{X_k, k \geq 0\}$.
- ✓ Le processus d'état X_k est Markovien de bruit de mesure V_k est une séquence de variables aléatoires indépendantes identiquement distribuées (i.i.d), indépendant du processus X_k .
- ✓ X_0 , l'état initial de densité p_0 est supposé de densité connue et indépendant de V_k et de W_k .

Le choix du critère qualifiant la précision de l'estimée n'est pas unique.

On cherche à minimiser la variance de l'erreur de filtrage. Ainsi, si \hat{x}_k désigne un estimateur de l'état X_k , celui-ci sera optimal s'il minimise l'erreur moyenne quadratique :

$$E \left[\|X_k - \hat{x}_k\|^2 \right]$$

Partant de l'hypothèse que seules informations dont on dispose sur le système sont les mesures $Y_{1,k} = Y_1, \dots, Y_k$, l'espérance conditionnelle donnée par,

$$\hat{x}_k = E[X_k | Y_1, \dots, Y_k] \tag{1.3}$$

donne l'estimateur non biaisé de variance minimale. Dans le cas où la loi conditionnelle de X_k sachant les observations (Y_1, \dots, Y_k) est fortement multimodale, il peut être souhaitable de prendre comme estimateur le maximum a posteriori,

$$\hat{x}_k = \text{ArgMax}_{x_k} [p(X_k | Y_1, \dots, Y_k)] \quad (1.4)$$

Dans tous les cas, l'information la plus riche disponible au vu des observations (Y_1, \dots, Y_n) est contenue dans la loi conditionnelle de X_k sachant les observations passées. Avec la donnée de cette loi, on est libre dans le choix de l'estimateur. Le problème de filtrage consistera donc à calculer (de manière exacte ou approchée) cette mesure de probabilité conditionnelle. En toute généralité, nous avons à faire à un problème de dimension infinie.

Dans notre travail, nous nous sommes particulièrement intéressés à l'application pistage du filtrage non linéaire. En pistage le système considéré est un mobile (par exemple un avion ou un navire, etc.) dont on cherche à estimer, à chaque instant k , l'état X_k , constitué par sa position et sa vitesse. Si l'avion n'est pas en phase de manœuvre, il est raisonnable de choisir un modèle de dynamique a priori rectiligne uniforme. Un radar fournit des mesures bruitées $(Y_k)_{k \geq 1}$ de sa distance à l'avion et de l'angle formé par un axe de référence et son axe de visée. La difficulté particulière de cette application est la contrainte de temps réel, c'est-à-dire que le temps de calcul de l'estimée fournie par le filtre doit être de l'ordre de la période de mesure $(k+1)T - kT$ (au plus fin de l'ordre du centième de seconde).

Modèle général d'état et d'observation considéré :

Nous introduisons quelques notations et définitions qui serviront tout au long du ce mémoire. Nous considérons deux suites de variables aléatoires $(X_k)_{k \geq 0}$ et $(Y_k)_{k \geq 1}$, définies sur l'espace probabilisé (Ω, f, P) et à valeurs (respectivement) dans R^d et R^q . $(X_k)_{k \geq 0}$ est appelé signal ou processus état, et $(Y_k)_{k \geq 1}$ processus observation.

- ✓ Le processus état $(X_k)_{k \geq 0}$ est une chaîne de Markov inhomogène, avec pour noyaux de transition Q_k , i.e. pour tout entier $k \geq 1$

$$P[X_k \in dx | X_{0:k-1} = x_{0:k-1}] = P[X_k \in dx | X_{k-1} = x_{k-1}] = Q_k(x_{k-1}, dx),$$

Et pour distribution de probabilité initiale p_0 . La distribution de probabilité conjointe de $X_{0:k}$ est notée $p_{0:k}$, nous avons donc

$$p_{0:k}(dx_{0:k}) = \mathbb{P}(X_{0:k} \in dx_{0:k}) = p_0(dx_0)Q_1(x_0, dx_1) \dots Q_k(x_{k-1}, dx_k),$$

Pour tout $x_{0:k} = (x_0, \dots, x_k) \in R^{(k+1)d}$. Par exemple, $(X_k)_{k \geq 0}$ peut être défini par une équation du type

$$X_k = f_k(X_{k-1}, W_k),$$

Où $(W_t)_{t \geq 0}$ est une suite de variables aléatoires indépendantes, pas nécessairement gaussiennes et indépendantes de l'état initial X_0 .

- ✓ Le processus observation $(Y_k)_{k \geq 1}$ est lié au processus état de la manière suivante, pour tout entier $k \geq 1$,

$$Y_k = h_k(X_k, V_k),$$

Où $(V_k)_{k \geq 1}$ est une suite de variables aléatoires (pas nécessairement gaussiennes) et indépendantes du processus état $(X_k)_{k \geq 0}$. Nous supposons que pour chaque entier $k \geq 1$, la distribution de probabilité $\mathbb{P}[Y_k \in dy | X_k = x]$ sur R^q , paramétrée par $x \in R^d$, est dominée, i.e.

$$\mathbb{P}[Y_k \in dy | X_k = x] = g_k(x, y) \lambda_k(dy)$$

Pour une mesure positive λ_k , typiquement la mesure de Lebesgue). $g_k(x, Y_k)$ est la fonction du vraisemblance associée.

Pour tout $x \in R^d$. Notons que $g_k(x, Y_k)$ dépend implicitement de l'observation Y_k . Comme de plus les observations $\{Y_k, k \geq 1\}$ sont mutuellement indépendantes, conditionnellement à $\{X_k, k \geq 0\}$, la distribution de probabilité $\mathbb{P}[Y_{1:k} \in dy_{1:k} | X_{1:k} = x_{1:k}]$ sur $R^{k \times q}$ est dominée et possède la densité $g_{1:k}(x_{1:k}, \cdot)$, telle que

$$\begin{aligned} \mathbb{P}[Y_{1:k} \in dy_{1:k} | X_{1:k} = x_{1:k}] &= g_{1:k}(x_{1:k}, y_{1:k}) \lambda_{1:k}(dy_{1:k}) \\ &= g_1(x_1, y_1) \lambda_1(dy_1) \dots g_k(x_k, y_k) \lambda_k(dy_k) \end{aligned}$$

Où $dy_{1:k} = dy_1 \times \dots \times dy_k$ désigne l'hypercube élémentaire de $R^{k \times q}$.

Les propriétés d'indépendance conditionnelle des mesures et la structure markovienne du processus d'état conduisent à une structure de chaîne de Markov cachée, illustrée en figure I-1.

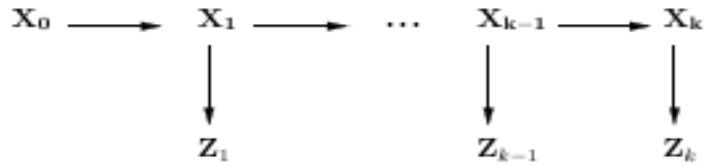


Figure I-1 Graphe de dépendance de la chaîne de Markov cachée constituée par les processus X_t et Y_t

I.1. Evolution du filtrage optimal :

Le problème de filtrage non linéaire est de calculer à chaque pas de temps la densité conditionnelle p_k de l'état X_k , sachant une réalisation des observations $Y_{1:k} = (Y_1, \dots, Y_k)$ jusqu'à l'instant courant k ,

$$p_k(dx) = P[X_k \in dx | Y_{1:k}]$$

La suite des lois de probabilités conditionnelles $(p_k)_k \geq 0$, est appelée filtre optimal. En pratique, l'algorithme proposé pour réaliser ce calcul doit satisfaire des contraintes de temps réel. Pour cette raison, on souhaite obtenir un algorithme récursif de calcul de p_k , i.e. le calcul de p_k ne doit être fonction que de la dernière observation Y_k et de la loi conditionnelle précédente p_{k-1} , sans avoir à réutiliser les observations entre les instant $t = 1$ et $t = k - 1$.

Le calcul effectif du filtre optimal est en général un problème très délicat, par contre son formalisme mathématique est aisé.

Le passage entre p_{k-1} à p_k fait intervenir la loi conditionnelle de $p_{k/k-1}$ de l'état X_k sachant les observations $Y_{1:k-1} = (Y_1, \dots, Y_{k-1})$ jusqu'à l'instant précédent $k - 1$, i.e.

$$p_{k|k-1}(dx) = P[X_k \in dx | Y_{1:k-1}]$$

La suite des lois de probabilités $(p_{k/k-1})_{k \geq 0}$, est appelée filtre de prédiction. L'évolution du filtre optimal se décompose en deux étapes essentielles.

Prédiction : L'étape de prédiction est la première étape d'évolution du filtre optimal, elle utilise la connaissance a priori du système, à travers le noyau de transition Q_k , en réalisant la transition de p_{k-1} à $p_{k|k-1}$ de la manière suivante,

$$p_{k|k-1}(dx) = \int_{R^d} p_{k-1}(dx') Q_k(x', dx) \equiv (Q_k p_{k-1})(dx) \tag{1.5}$$

où R^d désigne l'espace d'état.

L'étape de prédiction est linéaire.

Correction : La deuxième étape utilise l'observation Y_k à travers la densité conditionnelle de y_k sachant x_k . Cette densité notée $g_k(x) = p(y_k/x_k)$ est appelée vraisemblance, elle permet de corriger la densité prédite $p_{k|k-1}$, par application de la formule de bayes ;

$$p_k(dx) = \frac{g_k(x) p_{k|k-1}(dx)}{\int_{R^d} g_k(x') p_{k|k-1}(dx')} \tag{1.6}$$

L'étape de correction est donc non linéaire.

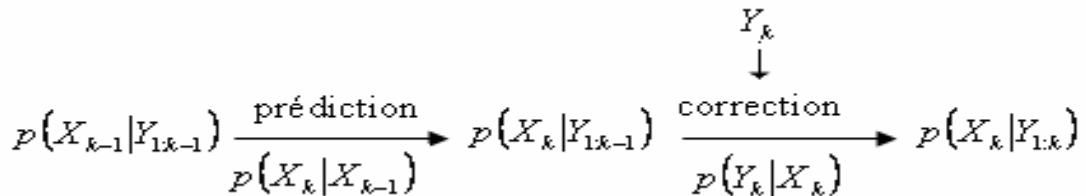


Figure I-2

La figure I-2 donne le schéma classique de la décomposition d'une itération d'un filtre non linéaire optimal en une étape de prédiction, faisant intervenir la loi de transition, et une étape de correction faisant intervenir la nouvelle observation et la fonction vraisemblance locale associée.

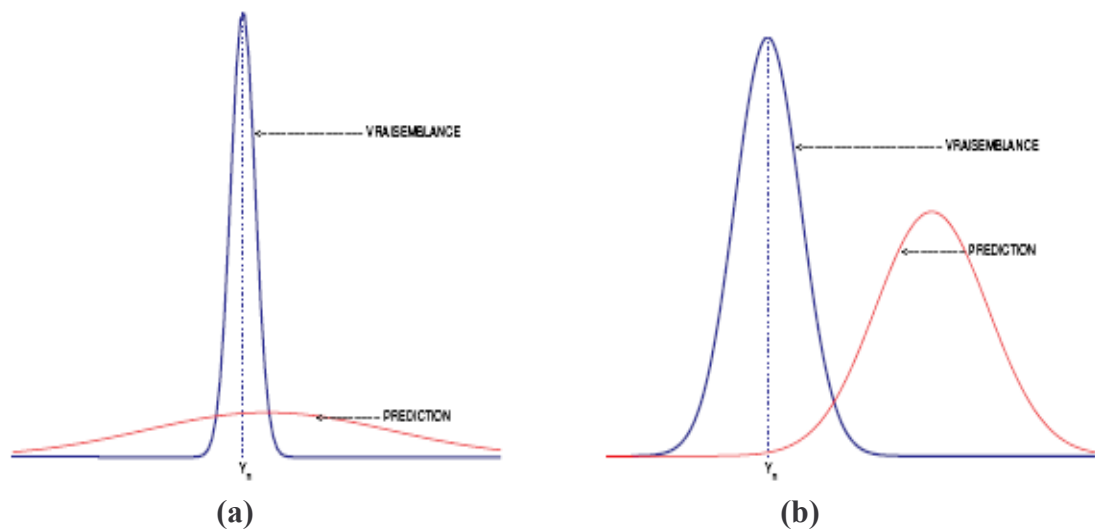


Figure I-3

La figure I-3 donne un simple exemple de la fonction de vraisemblance et de la densité prédite pour une observation précise (a) ou incohérente (b) par rapport à la prédiction, (dans le cas d'un état directement observé i.e. $(Y_k = X_k + V_k)$)

L'application de l'étape (1.5) donne la solution théorique du filtrage non linéaire optimal. Son calcul effectif exige le calcul d'intégrales multi-dimensionnelles. Malheureusement, dès que l'équation d'état et/ou l'équation d'observation est non linéaire, ces intégrales ne se calculent pas analytiquement. On a alors recours à des approximations qui sont des méthodes analytiques (par exemple le filtre de Kalman étendu qui considère le modèle linéarisé).

II. Etat de l'art des méthodes classiques d'approximations :

Dans cette section, nous présentons un historique des principales méthodes classiques proposées pour résoudre le problème de filtrage. Nous les avons regroupées en deux classes : les méthodes analytiques, à savoir le filtre de Kalman et ses dérivées, et les méthodes numériques qu'on décrit brièvement.

II.1. Méthodes analytiques :

On s'intéresse particulièrement aux filtres de Kalman-Bucy et de Kalman étendu. Ce sont les outils les plus communément utilisés pour résoudre les problèmes de filtrage linéaire et non linéaire.

II.1.1. Filtre de Kalman-Bucy (KF) :

Le filtre de Kalman-Bucy développé à l'origine par Kalman [1] en 1960 pour le temps discret puis repris en 1961 par Kalman et Bucy [2] pour le temps continu permet de résoudre les problèmes de filtrage linéaire lorsque les bruits sont additifs et gaussiens. Version moderne et probabiliste de la méthode des moindres carrés, le filtre de Kalman concerne essentiellement les systèmes linéaires. Une étude détaillée de filtre de Kalman-Bucy est développée dans [14]. On pourra trouver une introduction claire et concise dans [15]. On considère ici le modèle linéaire gaussien, où le processus état suit une évolution linéaire avec un bruit additif Gaussien et où l'observation est une fonction linéaire de l'état entachée d'un bruit additif Gaussien, i.e.,

$$\begin{cases} X_k = F_k X_{k-1} + B_k + W_k \\ Y_k = H_k X_k + D_k + V_k \end{cases}$$

Où

- ✓ F_k de dimension $(d \times d)$, $B_k (d \times 1)$, $H_k (q \times d)$ et $D_k (q \times 1)$ sont des matrices déterministes et connues.
- ✓ Les bruits d'état et d'observation W_k et V_k à valeurs respectivement dans R^d et R^q sont des bruits blancs Gaussiens de matrices de covariance respectives Q_k et R_k . Ces bruits sont mutuellement indépendants et indépendants de la condition initiale X_0 .
- ✓ La loi initiale X_0 est Gaussienne de moyenne $E[X_0] = \bar{X}_0$ et de matrice de covariance P_0 .

On montre dans ce cas que le couple $(X_k, Y_k)_{k \geq 1}$ est gaussien et que par conséquent les lois conditionnelles p_k sont elles aussi gaussiennes. Le problème est alors de dimension finie. Deux paramètres suffisent à caractériser la densité p_k : sa moyenne \hat{X}_k et sa matrice de covariance P_k . Ces deux paramètres sont calculés de manière récursive et exactement à partir des équations (1.5) et (1.6) du filtre de Kalman-Bucy.

On pose

$$p(X_k | Y_{1:k-1}) = N(\hat{X}_k, P_k) \text{ et } p(X_k | Y_{1:k-1}) = N(\hat{X}_{k|k-1}, P_{k|k-1})$$

avec :

$$\begin{aligned} \hat{X}_k &= E[X_k | Y_{1:k}], & P_k &= E[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T], \\ \hat{X}_{k|k-1} &= E[X_k | Y_{1:k-1}], & P_{k|k-1} &= E[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^T], \end{aligned}$$

Les conditions initiales du filtre sont données par

$$\begin{cases} \hat{X}_0 = \bar{X}_0 \\ P_0 \end{cases}$$

Puis \hat{X}_k et P_k sont calculés par les étapes de prédiction et de correction, la prédiction : on calcule $P_{k|k-1}$ à l'aide de l'équation d'état. La correction : on corrige la prédiction en tenant compte de la nouvelle observation Y_k et l'équation d'observation afin d'obtenir p_k .

Un point important est de déterminer ce que la nouvelle observation Y_k apporte de nouveau par rapport aux observations passées $Y_{1:k-1}$, la réponse est donnée par le processus innovation [16].

$$v_k = Y_k - \hat{Y}_{k|k-1} \quad \text{où} \quad \hat{Y}_{k|k-1} = E[Y_k | Y_{1:k-1}] \quad (1.7)$$

Donc elle s'exprime de la façon suivante,

$$v_k = Y_k - H_k \hat{X}_{k|k-1} \quad (1.8)$$

et elle forme une suite indépendante de variables aléatoires, avec

$$S_k = H_k P_{k|k-1} H_k^T + R_k$$

Ça matrice de covariance, et $v_k \perp Y_{1:k-1}$. Ainsi les informations contenues dans $(Y_{1:k-1}, Y_k)$ et dans $(Y_{1:k-1}, v_k)$ sont identiques.

➤ **Pour la phase de prédiction**, on a

$$\begin{aligned} \hat{X}_{k|k-1} &= E[X_k | Y_{1:k-1}] \\ &= F_{k-1} E[X_{k-1} | Y_{1:k-1}] + B_{k-1} + E[W_k | Y_{1:k-1}]. \end{aligned}$$

Comme $E[X_{k-1} | Y_{1:k-1}] = \hat{X}_{k-1}$ et $E[W_k | Y_{1:k-1}] = 0$

On en déduit

$$\hat{X}_{k|k-1} = F_{k-1}\hat{X}_{k-1} + B_{k-1} \quad (1.9)$$

d'une part et $X_k - \hat{X}_{k|k-1} = F_{k-1}(X_{k-1} - \hat{X}_{k-1}) + W_k$ D'autre part, d'où

$$\begin{aligned} P_{k|k-1} &= \mathbb{E}\left[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^T\right] \\ &= \mathbb{E}\left[\left(F_{k-1}(X_{k-1} - \hat{X}_{k-1}) + W_k\right)\left(F_{k-1}(X_{k-1} - \hat{X}_{k-1}) + W_k\right)^T\right] \\ &= F_{k-1}P_{k-1}F_{k-1}^T + Q_k \end{aligned} \quad (1.10)$$

En effet, $(X_k - \hat{X}_k) \perp W_k$ donc $\mathbb{E}\left[(X_{k-1} - \hat{X}_{k-1})W_k^T\right] = 0$.

➤ **Pour la phase de correction**, on a

$$\hat{X}_k = \mathbb{E}[X_k | Y_{1:k}] = \hat{X}_{k|k-1} + \mathbb{E}[X_k - \hat{X}_{k|k-1} | Y_{1:k}]$$

D'après la définition de l'innovation (1.8) on a :

$$\mathbb{E}[X_k - \hat{X}_{k|k-1} | Y_{1:k}] = \mathbb{E}[X_k - \hat{X}_{k|k-1} | Y_{1:k-1}, \nu_k]$$

et comme $Y_{1:k-1} \perp (X_k - \hat{X}_{k|k-1}, \nu_k)$, on obtient :

$$\mathbb{E}[X_k - \hat{X}_{k|k-1} | Y_{1:k-1}, \nu_k] = \mathbb{E}[X_k - \hat{X}_{k|k-1} | \nu_k].$$

Ainsi

$$\hat{X}_k = \hat{X}_{k|k-1} + \mathbb{E}[X_k - \hat{X}_{k|k-1} | \nu_k] \text{ et}$$

$$\begin{aligned} X_k - \hat{X}_k &= (X_k - \hat{X}_{k|k-1}) - (\hat{X}_k - \hat{X}_{k|k-1}) \\ &= (X_k - \hat{X}_{k|k-1}) - \mathbb{E}[X_k - \hat{X}_{k|k-1} | \nu_k]. \end{aligned}$$

On calcul maintenant la moyenne et la covariance de $(X_k - \hat{X}_{k|k-1})v_k$. Le vecteur $(X_k - \hat{X}_{k|k-1}, v_k)$ est gaussien, centré de covariance

$$\begin{pmatrix} P_{k|k-1} & P_{k|k-1}H_k^T \\ H_k P_{k|k-1} & H_k P_{k|k-1}H_k^T + R_k \end{pmatrix}.$$

On a donc la loi conditionnelle de $X_k - \hat{X}_{k|k-1}$ sachant v_k est gaussienne dont la moyenne et la covariance sont données par

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k [Y_k - (H_k \hat{X}_{k|k-1} + D_k)] \tag{1.11}$$

$$P_k = [I - K_k H_k] P_{k|k-1} \tag{1.12}$$

$$K_k = P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1} \tag{1.13}$$

Précisons que le filtre de Kalman est à la fois un théorème et un algorithme. L'algorithme de filtre de Kalman est défini comme suit

% Initialisation

$$\hat{X}_0 \leftarrow \bar{X}_0$$

$$P_0 \leftarrow Q_0$$

% Itérations

Pour k = 1, 2, 3, ... faire

%prédiction

$$\hat{X}_{k|k-1} \leftarrow F_{k-1} \hat{X}_{k-1} + B_{k-1}$$

$$P_{k|k-1} \leftarrow F_{k-1} P_{k-1} F_{k-1}^T + Q_k$$

% correction

$$\begin{aligned} K_k &\leftarrow P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1} \\ \hat{X}_k &\leftarrow \hat{X}_{k|k-1} + K_k [Y_k - (H_k \hat{X}_{k|k-1} + D_k)] \\ P_k &\leftarrow [I - K_k H_k] P_{k|k-1} \end{aligned}$$

Fin pour

Alg. I-1 Filtre de Kalman pour le système linéaire/gaussien

K_k est appelé le gain de Kalman. On remarque que les suites des matrices de covariances $(P_k)_{k \geq 0}$, $(P_{k|k-1})_{k \geq 0}$ (solution de l'équation de Riccati) et les gains $(K_k)_{k \geq 0}$ ne dépendent pas des observations $(Y_k)_{k \geq 1}$. On peut donc les précalculer afin de diminuer sensiblement la quantité de calculs à effectuer en temps réel. On peut voir le filtre de Kalman de trois manières différentes :

- ✓ Le filtre de Kalman est solution récursive du maximum de vraisemblance (moindres carrés récursifs).
- ✓ Le filtre de Kalman résout explicitement les équations bayésienne (1.5).
- ✓ Le filtre de Kalman est la projection du processus d'état sur l'espace des mesures de base orthonormale constituée par innovations $[y_k - (H_k \hat{x}_{k|k-1} + D_k)]$ démontré par Anderson et Moore [17] ainsi que Chui et Chen [18].

Il est intéressant de noter que, dans le cas de bruits non Gaussiens, et de modèles linéaires, le filtre de Kalman fournit toujours l'estimée linéaire à variance minimale de l'état du système. C'est notamment pour cette raison qu'il est le filtre le plus répandu. Dans le cas où les coefficients dépendent des observations passées, le modèle est dit "conditionnellement Gaussien". Le processus (x_k, y_k) n'est pas un processus Gaussien mais la loi conditionnelle p_k reste Gaussienne et le filtre de Kalman fonctionne toujours [19].

II.1.2. Filtre de Kalman Etendu (EKF) :

Considérons cette fois-ci un système dynamique non linéaire, avec toujours des bruits blancs Gaussiens additifs, indépendants entre eux et indépendants de la loi initiale.

$$\begin{cases} X_k = f_k(X_{k-1}) + W_k \\ Y_k = h_k(X_k) + V_k \end{cases}$$

Où les fonctions f_k et h_k sont non linéaires. Le filtre optimal associé à ce système n'est plus gaussien et il ne peut pas se résoudre explicitement comme dans le cas linéaire/gaussien. Une idée naturelle est alors de linéariser le système dynamique autour de l'état prédit $X_{k|k-1}$ et autour de l'état courant X_k , puis d'appliquer la technique de filtre de Kalman.

Pour l'initialisation on pose :

$$\hat{X}_0 = E[X_0], \text{ et } P_0 = \text{cov}(X_0)$$

Où X_0 n'est pas nécessairement gaussien.

➤ **Pour la phase de prédiction**, on linéarise l'équation d'état autour de \hat{X}_{k-1} :

$$\begin{aligned} X_k &= f_k(X_{k-1}) + W_k \\ &\cong F_k X_{k-1} + f_k(\hat{X}_{k-1}) - F_k \hat{X}_{k-1} + W_k \end{aligned}$$

Où

$$F_k \equiv \nabla f_k(\hat{X}_{k-1})$$

Ainsi

$$\begin{aligned} \hat{X}_{k|k-1} &= E[X_k | Y_{1:k-1}] \\ &\cong E[f_k(\hat{X}_{k-1}) | Y_{1:k-1}] + F_k E[X_{k-1} - \hat{X}_{k-1} | Y_{1:k-1}] + E[W_k | Y_{1:k-1}] \\ &= f_k(\hat{X}_{k-1}). \end{aligned}$$

La covariance prédite est donnée par :

$$\begin{aligned} P_{k|k-1} &= E\left[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^T \right] \\ &\cong F_{k-1}^T P_{k-1} F_{k-1} + Q_k. \end{aligned}$$

- **Pour la phase de correction**, on linéarise l'équation d'observation autour de l'estimée courante, c'est-à-dire $X_{k|k-1}$, on obtient :

$$Y_k = h_k(X_k) + V_k$$

$$\cong H_k X_k + h_k(\hat{X}_{k|k-1}) - H_k \hat{X}_{k|k-1} + V_k$$

Où $H_k \equiv \nabla h_k(\hat{X}_{k|k-1})$. En appliquant le filtre de Kalman on obtient

$$K_k = P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1}$$

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k [Y_k - h_k(\hat{X}_{k|k-1})]$$

$$P_k = [I - K_k H_k] P_{k|k-1}$$

En résumé, l'algorithme de filtre de Kalman étendu s'écrit de la façon suivante :

% Initialisation

$$\hat{X}_0 \leftarrow \bar{X}_0$$

$$P_0 \leftarrow Q_0$$

% Itérations

Pour $k = 1, 2, 3, \dots$ **faire**

%prédiction

$$F_k \leftarrow \nabla f_k(\hat{X}_{k-1})$$

$$\hat{X}_{k|k-1} \leftarrow f_k(\hat{X}_{k-1})$$

$$P_{k|k-1} \leftarrow F_k P_{k-1} F_k^T + Q_k$$

%correction

$$\begin{aligned}
 H_k &\leftarrow \nabla h_k(\hat{X}_{k|k-1}) \\
 K_k &\leftarrow P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1} \\
 \hat{X}_k &\leftarrow \hat{X}_{k|k-1} + K_k [Y_k - (H_k \hat{X}_{k|k-1})] \\
 P_k &\leftarrow [I - K_k H_k] P_{k|k-1}
 \end{aligned}$$

Fin pour

Alg. I-2 Filtre de Kalman Étendu pour le système non linéaire.

Le filtre de Kalman appliqué au système linéaire ainsi défini fournit les équations du filtre de Kalman étendu avec linéarisation au premier ordre. Plus la linéarisation sera fine, plus le filtre étendu donnera de bons résultats.

Dans la pratique, lorsque le système linéaire est une mauvaise approximation du fait de non linéarités trop importantes, le filtre de Kalman étendu a tendance à diverger, accordant trop de confiance au système linéaire. Pour y remédier, il a été proposé d'augmenter artificiellement, soit la covariance du bruit d'état pour compenser le fait que le modèle de prédiction n'est pas bien adapté, soit d'augmenter la covariance $P_{k|k-1}$, accordant par là-même moins d'importance au passé. Bien qu'il soit utilisé depuis une trentaine d'années, des études théoriques permettant de le justifier n'ont été développées que très récemment, essentiellement grâce aux travaux de Picard [5, 20, 21]. Ces travaux réalisés pour le modèle état en temps continu / observation en temps continu justifient l'utilisation d'algorithmes du type EKF lorsque l'observation est très bonne et lorsque le bruit d'observation et l'erreur initiale sont faibles. Ces résultats ont été étendus par et Getout-Petit [22] à certains cas où l'observation est partiellement très bonne, typiquement lorsqu'une coordonnée est totalement observée.

Cependant, l'EKF ne peut pas s'appliquer lorsque la densité conditionnelle est multimodale.

II.1.3. Filtre de Kalman sans parfum (Unscented Kalman Filter) (UKF) :

Nous présentons succinctement ici un dernier filtre du type filtre de Kalman, le filtre de Kalman sans parfum (Unscented Kalman filter, dénoté UKF en anglais), proposé par julier [23], et qui repose sur le principe suivant : on approxime à nouveau la loi conditionnelle a posteriori $p(X_{1:k} | Y_{0:k})$ par une loi gaussienne dont la moyenne et la covariance obtenues par le filtre sont notées respectivement $\bar{X}_{k|k}$ et P_k . Cependant, il n'y a pas de linéarisation des équations d'état et de mesures. La loi gaussienne est caractérisée par un ensemble de points qui captent entièrement sa moyenne et sa covariance et auquel on applique les vrais modèles non linéaires d'état et de mesure.

L'algorithme de l'UKF est le suivant :

- **Initialisation** : la première étape du filtre de Kalman Unscented consiste à générer un nuage de points répartis dans un volume paramétré par \bar{X}_0 et P_0 et qui constitue une approximation discrète de la loi $N(\bar{X}_0, P_0)$.
- **Prédiction** : à l'instant $t = k - 1$, chaque point du nuage est prédit selon la loi d'évolution non linéaire du système. Les estimées prédites $\hat{X}_{1:k|0:k-1}$ et $P_{1:k|0:k-1}$ sont alors calculées en fonction de l'échantillon de la loi $p(X_{1:k} | Y_{0:k-1})$ que constituent les points du nuage prédit. A chaque point du nuage est de plus associée une mesure prédite dont on calcule également la moyenne $\hat{Y}_{k|k-1}$.
- **Mise à jour (correction)** : les estimées prédites $\hat{X}_{1:k|0:k-1}$ et $P_{1:k|0:k-1}$ sont mise à jours proportionnellement à l'innovation $Y_k - \hat{Y}_{k|k-1}$ et à la matrice de covariance des mesures prédites.

Ce filtre repose ainsi sur l'idée qu'il est plus aisé d'approximer une gaussienne par un nuage de points que de linéariser une fonction. Les performances des l'UKF seraient ainsi meilleures que celles de l'EKF, pour une complexité équivalente [24], du fait de la meilleure estimation des moments. Il montre cependant ses limites pour des phénomènes fortement non linéaires et / ou non gaussiens.

Les filtres analytiques présentés ci-dessus ne sont pas adaptés lorsque la densité conditionnelle est multimodale.

II.2. Méthodes Numériques :

Considérons un système dynamique à temps continu d'état $X_t \in R^d$ et observé par la variable $Y_t \in R^q$ dont les équations d'évolution sont définies par :

$$\begin{cases} dX_t = f(X_t)dt + \sigma(X_t)dW_t \\ dY_t = h(X_t)dt + dV_t \end{cases}$$

Où dW_t et dV_t sont deux processus de Wiener indépendants. Parmi les approches utilisées pour traiter le problème du filtrage non linéaire, l'approche numérique qui consiste à voir le filtre optimal p_t comme la solution d'une équation aux dérivées partielles stochastique, l'équation de Zakai (sous forme continue) [25, 26].

$$dp_t = L^T p_t dt + p_t h^T R^{-1} dY_t \quad (1.14)$$

Avec L est un opérateur associé au processus X_t , $t \geq 0$

$$L = \frac{1}{2} \sum_{i,j=1}^q f_{i,j}(\cdot) \frac{\partial^2}{\partial X_i \partial X_j} + \sum_{i,j=1}^q g_i(\cdot) \frac{\partial}{\partial X_i} \quad (1.15)$$

Avec $f = (f_{i,j}) = \sigma \sigma^T$.

On discrétise l'équation de Zakai par des méthodes de maillage de l'espace d'état (schéma d'Euler). Il existe deux types de méthodes de maillage : méthode de maillage à pas fixe qui a l'inconvénient de mailler des zones où la densité de probabilité prend des valeurs négligeables. Pour cette méthode, la taille de la maille sera prise identique pour des zones de forte probabilité. Les méthodes de maillage à pas variable développées par Cai, LeGland et Zhang [6], sont plus intéressantes car elles permettent de raffiner le pas de maillage dans des zones de forte probabilité. Pour de faibles dimensions de l'espace d'état ($d \leq 3$), ces méthodes donnent de bons résultats pour un temps de calcul acceptable. Mais au-delà de la dimension 3 ces méthodes sont inutilisables car, pour une précision donnée, le temps de calcul croît exponentiellement avec d .

III. Conclusion

Dans ce chapitre, nous avons décrit les deux principales méthodes utilisées dans le filtrage qui sont les méthodes analytiques et numériques. Dans les méthodes analytiques, le filtre de Kalman étendu donne de très bons résultats dans le cadre des systèmes faiblement non linéaires. Mais lorsque les systèmes sont fortement non linéaires, les performances du filtre de Kalman étendu sont mauvaises. Quant aux méthodes numériques, elles fournissent des résultats d'une bonne précision au problème du filtrage non linéaire. Cependant, elles restent limitées à des espaces de dimension au plus égale à 3. Au delà de cette dimension, le temps de calcul augmente d'une façon exponentielle avec la dimension de l'espace, ce qui les rend impraticables en temps réel. C'est pour cela qu'une nouvelle catégorie de méthodes a été introduite : les méthodes de Monte Carlo. Elles sont mieux adaptées au problème du filtrage non linéaire car elles sont peu sensibles à la dimension de l'espace d'état. Nous appliquerons ce type de méthodes, appelées filtrage particulière au problème de la poursuite de cibles.

Introduction

Dans ce chapitre on se propose d'introduire les bases d'une technique d'approximation dite particulière. Cette technique fait appel à la méthode de Monte Carlo que nous présentons dans un premier temps. Après avoir présenté le formalisme et le principe des méthodes de Monte Carlo, nous abordons les difficultés liées à ce type d'estimateur ainsi que les techniques permettant de pallier ces problèmes. Ensuite, nous décrivons l'algorithme générique de filtrage particulière qui peut en être déduit.

Les méthodes de Monte Carlo sont apparues dans les années 50 avec l'article de Metropolis et Ulam [27] suivi d'autres travaux comme [28] ou [29]. Puis dans les années 60-70, les faibles puissances de calcul ainsi que la dégénérescence des algorithmes due à leur implémentation basée sur un échantillonnage pondéré séquentiel brut ont conduit à une perte d'intérêt pour ces méthodes inutilisables en pratique. Seuls quelques développements ponctuels tels que [30] ou [31] ont continué à explorer ces idées. Plus récemment, dans les années 90, les méthodes de Monte Carlo ont été de nouveau explorées, avec l'introduction de rééchantillonnage par Gordon et al [8] qui permet de limiter la dégénérescence des algorithmes basés sur l'échantillonnage pondéré séquentiel. Cette contribution majeure associée à l'augmentation importante des capacités de calcul a eu un impact déterminant dans la communauté Traitement du signal en rendant les filtres Monte Carlo séquentiels utilisables en pratique pour la première fois.

Depuis, les activités de recherche dans ce domaine ont énormément augmenté [32] conduisant à de nombreuses contributions améliorant l'efficacité des filtres de Monte Carlo séquentiel ou les filtres particuliers. Ces méthodes ont été largement exploitées dans divers domaines comme par exemple le traitement du signal [33, 34], le traitement de la parole [35], la robotique mobile [36], ou le suivi visuel [37, 38], ...etc.

I. Principe des méthodes d'approximation de Monte Carlo :

Le principe de Monte-Carlo consiste à utiliser un échantillon discret obtenu par simulation pour approcher une distribution de probabilités qu'il est généralement impossible de calculer analytiquement. La concentration des échantillons dans une région est supposée représentative de la probabilité de cette région.

I.1. Le principe de l'intégration de Monte Carlo :

Le principe de l'intégration de Monte-Carlo repose sur l'idée simple que la probabilité d'une région est proportionnelle à la densité des échantillons de cette région.

Soit $p(X)$ une distribution de probabilité d'une variable aléatoire X que l'on cherche à approcher numériquement afin de calculer l'intégrale :

$$I(\varphi(X)) = \int \varphi(X)p(X)dX \quad (2.1)$$

Pour une fonction φ donnée.

L'idée est de donner une interprétation probabiliste à l'intégrale $I(\varphi(X))$. En l'occurrence cette intégrale peut s'écrire sous la forme $I(\varphi(X)) = E[\varphi(X)]$. La densité $p(X)$

peut être approchée par une loi discrète $p_N(X)$ à l'aide d'un échantillon de N variables aléatoires $(X^i)_{i=1,\dots,N}$ indépendantes et identiquement distribuées (i.i.d) de loi de densité $p(X)$,

$$p_N(X) = \frac{1}{N} \sum_1^N \delta_{X^{(i)}}(X) \quad (2.2)$$

Les intégrales $I(\varphi(X))$ peuvent alors être approchées par des sommes :

$$I(\varphi(X)) \approx I_N(\varphi(X)) = \frac{1}{N} \sum_1^N \varphi(X^{(i)}) \quad (2.3)$$

D'après la loi des grands nombres, cette approximation est convergente $I_N \xrightarrow{N \rightarrow \infty} I$ p.s. La vitesse de convergence est donnée par le théorème central limite :

$$\frac{\sqrt{N}}{\sigma} (I_N - I) \xrightarrow{N \rightarrow \infty} N(0,1) \quad (2.4)$$

Où $\sigma^2 = \text{var}(\varphi(X))$. Ce dernier résultat est très important en pratique. Avec $N(0,1)$ est la loi normale centrée réduite. Ceci montre que l'erreur moyenne d'une quadrature Monte Carlo est de l'ordre de σ/\sqrt{N} . Ces méthodes sont basées sur la simulation de variables aléatoires pour calculer approximativement des intégrales de grande dimension. Elles tirent leur justification de la loi des grands nombres qui permet d'approcher une mesure de probabilité par la mesure empirique calculée à l'aide d'un échantillon.

Un point positif est que la vitesse de convergence $1/\sqrt{N}$ de l'estimateur ne dépend donc que de N , et pas de la taille de l'espace. Cette vitesse étant donnée, on peut toutefois améliorer cette méthode en simulant de façon efficace le N échantillon, c'est le problème de simulation. Par ailleurs on peut se demander s'il est possible de diminuer σ^2 dans (2.4), c'est le problème de réduction de variance. Nous allons introduire dans le paragraphe suivant les méthodes d'échantillonnage.

I.2. Méthode de simulation de variables aléatoires :

Comment générer une suite de nombres X^1, \dots, X^N , dites pseudo aléatoires, qui ressemblent à un échantillon de taille N de la loi de densité $p(X)$? Ce problème est complexe et on se contente ici de présenter quelques méthodes. La résolution de ce problème peut être envisagée à partir de deux grandes classes d'approches : les méthodes non itératives et les méthodes itératives. Les méthodes itératives, basées sur les chaînes de Markov (MCMC) ne seront pas approfondies en raison de leur coût en termes de puissance de calcul jugé élevé et en raison d'une vitesse de convergence mal maîtrisée [54]. Les méthodes non itératives de génération de nombres aléatoires distribués selon une loi donnée dépendent des propriétés et des connaissances que l'on a de la loi objective.

I.2.1. Méthode d'acceptation / rejet :

Supposons maintenant que l'on s'intéresse à une loi de densité $p(X)$ selon laquelle on désire générer l'échantillon $(X^{1:N})$: d'une part on ne sait pas simuler directement selon elle, d'autre part on ne sait pas inverser sa fonction de répartition. En revanche on suppose que l'expression de la densité $p(X)$ est connue et que l'on sait simuler selon une autre densité $q(X)$ telle que $\frac{p(X)}{q(X)} \leq M < \infty$. Cette méthode consiste à générer un échantillon de variables aléatoires (v.a) qui suit exactement la loi $p(X)$ en acceptant ou rejetant les variables aléatoires (v.a) générées suivant $q(X)$.

L'algorithme de la méthode d'acceptation / rejet est définie de cette façon :

Répéter

$$u \sim U[0,1]$$

$$X' \sim q(X)$$

Jusqu'à $u \leq \frac{p(X')}{Mq(X')}$

Rendre X'

Alg. II-1 Générateur d'une loi $p(X)$ par méthode d'acceptation / rejet.

Dans la pratique, cet algorithme est limité par, d'une part, le fait que la détermination de la constante M soit souvent impossible car cette dernière s'exprime, notamment dans le contexte de l'estimation bayésienne, en fonction de densités inconnues a priori. La deuxième limitation de l'algorithme d'acceptation / rejet provient de la valeur de la constante M . Lorsque cette dernière est trop grande, ce qui intervient en particulier lorsque la loi $q(X)$ est trop différente de la loi objective $p(X)$, la probabilité d'acceptation des variables générées est trop faible et l'algorithme se révèle trop lent pour produire un échantillon de taille suffisante.

Si, maintenant, on abandonne l'hypothèse qui consiste à considérer la constante M comme connue, on aboutit au concept de l'échantillonnage pondéré (l'échantillonnage d'importance).

I.2.2. Echantillonnage pondéré :

D'abord, essayons de comprendre ce que l'on entend par "approximation particulière" d'une loi $p(X)$. L'approximation (2.2) de (2.1) est valide pour toute fonction $\varphi(X)$. Il s'agit donc d'une approximation de la loi $p(X)$, en effet :

$$p(X) \approx p_N(X) = \frac{1}{N} \sum_{i=1}^N \delta_{X^i}(X) \quad (2.5)$$

Où $X^{1:N} \sim p(X)$ (i.i.d).

Ainsi $p_N(X)$ est une approximation particulière de la loi $p(X)$. Plus généralement, une approximation particulière sera de la forme

$$p_N(X) = \sum_{i=1}^N w^i \delta_{X^i}(X) \quad (2.6)$$

Avec $w^i \geq 0$ et $\sum_{i=1}^N w^i = 1$. L'avantage de ce genre d'approximation réside dans le fait que les calculs d'intégrales sous cette loi sont très simples, en effet

$$E[\varphi(X)] = \int \varphi(X) p(X) dX \approx \int \varphi(X) p_N(X) = \sum_{i=1}^N w^i \varphi(X^i) \quad (2.7)$$

On cherche une approximation de la forme (2.6) d'une loi $p(x)$ donnée. Si l'on sait échantillonner selon $p(X)$ on peut utiliser l'approximation (2.5). Si on ne sait pas (ou ne veut pas) échantillonner selon $p(X)$, on peut faire appel à une méthode dite d'échantillonnage d'importance [39] qui s'appuie sur la remarque suivante : supposons que l'on dispose d'une autre loi $q(X)$ telle que

$$p(X) = 0 \Rightarrow q(X) = 0 \quad (2.8)$$

Plus précisément, supposons que les particules X^1, \dots, X^N soient échantillonnées de manière indépendante selon une densité (loi d'importance) $q(X)$ telle que $p(X) > 0$ implique $q(X) > 0$. L'intégrale $E[\varphi(X)]$ s'écrit

$$E[\varphi(X)] = \int \varphi(X) p(X) dX = \int \varphi(X) \frac{p(X)}{q(X)} q(X) dX .$$

de sorte qu'elle peut être approximée par la somme $\sum_{i=1}^N w^i \varphi(X^i)$.

où l'expression des poids est donnée par :

$$w^i = \frac{p(X^i)}{q(X^i)} .$$

Cependant, l'estimateur $\bar{E}[\varphi(X)]$ ne satisfait pas la condition de normalisation $\sum_1^N w^i = 1$.

C'est pourquoi il faut poser

$$\tilde{w}^i = \frac{w^i}{\sum_{i=1}^N w^i}. \quad (2.9)$$

En résumé, l'ensemble de particules pondérées $\{X^i, \tilde{w}^i\}$ constitue une description cohérente de $p(X)$. i.e. la loi discrète obtenue converge en loi vers la loi objective $p(X)$ [39].

$$p_N(X) = \sum_{i=1}^N \tilde{w}^i \delta_{X^i}(X) \quad (2.10)$$

La formulation de l'algorithme est comme suit :

- ✓ Générer un échantillon $\{X^i\}_{i=1:N}$ selon la loi $q(X)$.
- ✓ Pondérer chacune des réalisations selon le poids d'importance normalisé correspondant :

$$\tilde{w}^i = \frac{w^i}{\sum_{i=1}^N w^i}.$$

Ce n'est pas une méthode d'acceptation/rejet : ici toutes les simulations sont utilisées et on n'a pas à majorer $p(X)/q(X)$.

L'intérêt de l'échantillonnage pondéré est double [40, 41]. D'une part, il peut permettre une approximation de $p(X)$ plus efficace, par la réduction de la variance de son estimateur. Ceci implique que le rapport $p(X)/q(X)$ soit borné, sous peine que les poids \tilde{w}^i varient beaucoup et ne soient élevés que pour un nombre restreint de particules. D'autre part, comme cela a été indiqué plus haut, il permet de ne pas échantillonner selon $p(X)$. Ceci est particulièrement intéressant lorsque la loi $p(X)$ n'est connue qu'à une constante de normalisation près. Il convient toutefois de mentionner que ces propriétés nécessitent un effort calculatoire plus important au niveau de l'évaluation des poids.

II. Application à la problématique du filtrage :

II.1. Echantillonnage pondéré :

Dans ce contexte, il s'agit donc d'établir une représentation particulière de la loi conjointe a posteriori $p(X_{0:k}|Y_{1:k})$ ou de la loi marginale $p(X_k|Y_{1:k})$. Du fait que ces densités ne sont généralement connues qu'à une constante de normalisation près, l'échantillonnage d'importance est invoqué.

On choisit pour cela une distribution d'importance $q(X_{0:k}|Y_{1:k})$ pouvant être facilement échantillonnée et dont le support inclut le support de $p(X_{0:k}|Y_{1:k})$, i.e.

$$p(X_{0:k}|Y_{1:k}) > 0 \Rightarrow q(X_{0:k}|Y_{1:k}) > 0 \quad (2.11)$$

Conformément à (2.2), on peut écrire

$$p(X_{0:k}|Y_{1:k}) \approx p_N(X_{0:k}|Y_{1:k}) = \frac{1}{N} \sum_{i=1}^N \delta(X_{0:k} - X_{0:k}^i), \quad (2.12)$$

On pose le poids $w_{0:k}^i = \frac{1}{N}$, telle que $\sum_{i=1}^N w_{0:k}^i = 1$.

Dés lors que les échantillons $X_{0:k}^i$ sont i.i.d. et affecté de poids $w_{0:k}^i$ selon

$$X_{0:k}^i \sim q(X_{0:k}|Y_{1:k}), \quad w_{0:k}^i \propto \frac{p(X_{0:k}^i|Y_{1:k})}{q(X_{0:k}^i|Y_{1:k})}, \quad \sum_{i=1}^N w_{0:k}^i \quad (2.13)$$

L'intégrale

$$\begin{aligned} I(\varphi) &= \int \varphi(X_{0:k}) p(X_{0:k}|Y_{1:k}) dX_{0:k} \\ &= \int \varphi(X_{0:k}) \frac{p(X_{0:k}|Y_{1:k})}{q(X_{0:k}|Y_{1:k})} q(X_{0:k}|Y_{1:k}) dX_{0:k} \end{aligned} \quad (2.14)$$

relative à l'espérance a posteriori d'une fonction $\varphi(X_{0:k})$ de la trajectoire d'état depuis l'instant 0 jusqu'à l'instant k est par conséquent approximée au moyen de l'estimateur

$$I_N(\varphi) = \sum_{i=1}^N w_{0:k}^i \varphi(X_{0:k}^i) \quad (2.15)$$

dont les propriétés ont déjà été évoquées au I.1.

Notons qu'une marginalisation triviale de (2.12) par rapport à $X_{0:k-1}$ permet de déduire l'approximation particulière de la densité de filtrage, qui s'écrit

$$p(X_k | Y_{1:k}) \approx p_N(X_k | Y_{1:k}) = \sum_{i=1}^N w_k^i \delta(X_k - X_k^i), \quad (2.16)$$

Où $w_k^i = w_{0:k}^i$ et $\delta_{X_k^i}(X_k) = \delta(X_k - X_k^i)$.

L'échantillonnage pondéré permet de procéder, dans un cas général, à une estimation des états par intégration de Monte-Carlo. Cependant, la nécessité de la prise en compte du cadre bayésien, dictée par la formulation d'état adoptée pour notre problématique qui est la poursuite de cible manœuvrante, suppose d'adapter la procédure en conséquence.

Le paragraphe qui suit traite de l'adaptation de l'échantillonnage d'importance ou pondéré au cas bayésien.

II.2. Echantillonnage pondéré bayésien :

Le principe d'échantillonnage pondéré peut être appliqué au cadre du filtrage bayésien. Des phénomènes physiques peuvent être représentés par un modèle mathématique qui décrit l'évolution du système au cours du temps. Dans cette classe de modèle, un processus décrit le signal d'intérêt non observé, dit caché, et nous n'avons accès qu'à des observations liées statistiquement à ce processus. A partir de ces observations, nous désirons estimer le processus. Un cadre statistique naturel pour aborder l'estimation de ces modèles est le cadre bayésien [42].

$$X_k \sim p(X_k | X_{0:k-1}).$$

$$Y_k \sim p(Y_k | X_k).$$

X_k et Y_k représentent respectivement les états cachés et les observations et $p(X_k | X_{0:k-1})$ et $p(Y_k | X_k)$ leurs densités de probabilité. On considère également comme la densité à priori $p(X_0)$.

On fait, en général, deux hypothèses supplémentaires concernant le modèle

✓ Le modèle d'évolution est un processus markovien $p(X_{0:k}) = p(X_0) \prod_{j=1}^k p(X_j | X_{j-1})$.

✓ Les observations sont indépendantes conditionnellement aux états

$$p(Y_{1:k} | X_{0:k}) = \prod_{j=1}^k p(Y_j | X_j).$$

L'introduction de l'échantillonnage pondéré bayésien est issu de [42]. L'objectif est d'estimer la distribution à posteriori $p(X_{0:k}|Y_{1:k})$. Nous choisirons donc la distribution d'importance multivariée $q(X_{0:k}|Y_{1:k})$, à partir de laquelle la simulation d'échantillons est aisée.

L'espérance de la fonction $\varphi(X_{0:k})$ s'écrit, en appliquant la règle de bayes, de la manière suivante

$$\begin{aligned} E[\varphi(X_{0:k})] &= \int \varphi(X_{0:k}) \frac{p(X_{0:k}|Y_{1:k})}{q(X_{0:k}|Y_{1:k})} q(X_{0:k}|Y_{1:k}) dX_{0:k} \\ &= \int \varphi(X_{0:k}) \frac{p(Y_{1:k}|X_{0:k})p(X_{0:k})}{p(Y_{1:k})q(X_{0:k}|Y_{1:k})} q(X_{0:k}|Y_{1:k}) dX_{0:k} \\ &= \int \varphi(X_{0:k}) \frac{w_{0:k}}{p(Y_{0:k})} q(X_{0:k}|Y_{1:k}) dX_{0:k} \end{aligned} \quad (2.17)$$

Où les variables $w_{0:k}$ sont les poids d'importance non normalisés

$$w_{0:k} = \frac{p(Y_{1:k}|X_{0:k})p(X_{0:k})}{q(X_{0:k}|Y_{1:k})} \quad (2.18)$$

On peut se débarrasser de la constante de normalisation inconnue $p(Y_{1:k})$ dans l'équation (2.17), qui dans le cas générale n'admet pas de solution analytique.

Donc $E[\varphi(X_{0:k})]$ peut s'écrire de la façon suivante

$$E[\varphi(X_{0:k})] = \frac{1}{p(Y_{1:k})} \int \varphi(X_{0:k}) w_{0:k} q(X_{0:k}|Y_{1:k}), \quad (2.19)$$

De même, en appliquant la règle de bayes, $p(Y_{1:k})$ peut s'écrire

$$\begin{aligned} p(Y_{1:k}) &= \int p(Y_{1:k}|X_{0:k})p(X_{0:k})dX_{0:k} \\ &= \int \frac{p(Y_{1:k}|X_{0:k})p(X_{0:k})}{q(X_{0:k}|Y_{1:k})} q(X_{0:k}|Y_{1:k})dX_{0:k} \\ &= \int w_{0:k} q(X_{0:k}|Y_{1:k})dX_{0:k} \end{aligned} \quad (2.20)$$

Et finalement, en remplaçant (2.20) dans (2.19) on obtient

$$E[\varphi(X_{0:k})] = \frac{\int \varphi(X_{0:k}) w_{0:k} q(X_{0:k} | Y_{1:k}) dX_{0:k}}{\int w_{0:k} q(X_{0:k} | Y_{1:k}) dX_{0:k}}$$

$$E[\varphi(X_{0:k})] = \frac{E_{q(\cdot|Y_{1:k})}[\varphi(X_{0:k}) w_{0:k}]}{E_{q(\cdot|Y_{1:k})}[w_{0:k}]} \quad (2.21)$$

On obtient un estimateur $E_N(\varphi)$ de $E[\varphi(X_{0:k})]$ simplement à partir des estimateurs des numérateur et dénominateur du quotient donné par la relation (2.21).

$$E_N[\varphi(X_{0:k})] = \frac{1/N \sum_{i=1}^N \varphi(X_{0:k}^i) w_{0:k}^i}{1/N \sum_{i=1}^N w_{0:k}^i}$$

$$= \sum \tilde{w}_{0:k}^i \varphi(X_{0:k}^i) \quad (2.22)$$

Avec les poids d'importance $\tilde{w}_{0:k}^i$ normalisés donnés par

$$\tilde{w}_{0:k}^i = \frac{w_{0:k}^i}{\sum_{i=1}^N w_{0:k}^i} \quad (2.23)$$

La distribution a posteriori tend vers la loi suivante, où les échantillons $X_{0:k}^i$ sont générés selon la loi d'importance $q(X_{0:k} | Y_{1:k})$, elle est assimilée à des mesures discrètes pondérées.

$$p_N(X_{0:k} | Y_{1:k}) = \sum w_{0:k}^i \delta(X_{0:k} - X_{0:k}^i) \quad (2.24)$$

L'échantillonnage d'importance constitue donc une véritable méthode de simulation de lois.

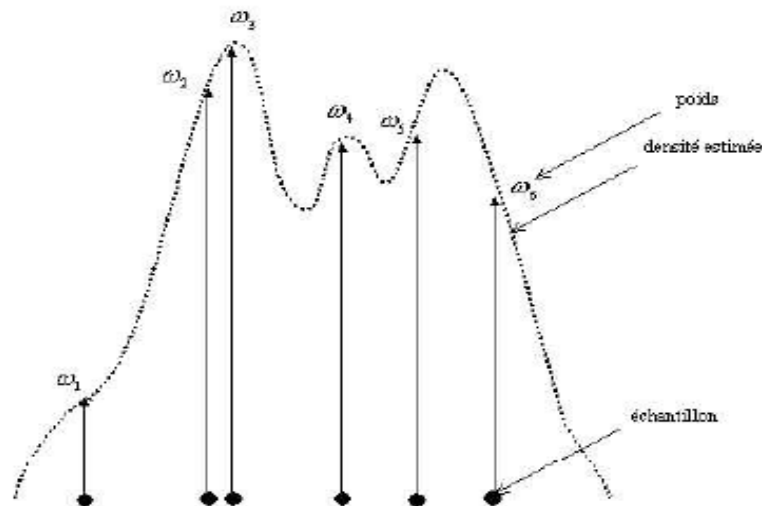


Figure II-1 Mesure discrète pondérée de la probabilité a posteriori

La limitation d'emploi de cet algorithme provient de la nécessité de calculer, à chaque récurrence du filtre, l'ensemble de la séquence d'état pour la mise à jour des poids d'importance. On aboutit rapidement à une explosion de la charge de calcul qui a pour effet de rendre inapplicable, en pratique, un tel algorithme. Il est cependant possible, moyennant une perte substantielle de généralité concernant la fonction d'importance, d'obtenir une formulation séquentielle de l'échantillonnage pondéré.

II.3. Echantillonnage d'importance bayésien séquentiel (SIS) :

Il est possible de modifier la méthode d'échantillonnage pondéré permettant l'estimation de la distribution a posteriori sans engendrer de modification de trajectoires passées $\{X_{0:k-1}^i; i = 1, \dots, N\}$. Il suffit pour cela que la distribution d'importance à l'indice k admette comme distribution marginale à l'index $k-1$ la fonction d'importance $q(X_{0:k-1}|Y_{1:k-1})$, d'où

$$q(X_{0:k}|Y_{1:k}) = q(X_{0:k-1}|Y_{1:k-1})q(X_k|X_{0:k-1}, Y_{1:k}) \quad (2.25)$$

Par itération, on obtient

$$q(X_{0:k}|Y_{1:k}) = q(X_0) \prod_{i=1}^k q(X_i|X_{0:k-i}, Y_{1:i}) \quad (2.26)$$

Ce qui, compte tenu de l'indépendance des observations conditionnellement aux états et du caractère markovien du processus d'état, permet d'exprimer les poids d'importance sous une forme récursive.

On peut résumer les hypothèses

- ✓ Une distribution d'importance telle que $q(X_{0:k+1}|Y_{1:k+1}) = q(X_{0:k}|Y_{1:k})q(X_{k+1}|X_{0:k}, Y_{1:k+1})$.
- ✓ Un système dynamique markovien $p(X_{0:k+1}) = p(X_0) \prod_{j=1}^{k+1} p(X_j|X_{j-1})$.
- ✓ Des observations indépendantes conditionnellement aux états $p(Y_{0:k+1}|X_{0:k+1}) = \prod_{j=1}^{k+1} p(Y_j|X_j)$.

Les poids d'importance non normalisés de l'échantillonnage d'importance bayésien sont donnés par (2.18). Après développement, en tenant compte des hypothèses précédentes, on aura

$$w_{0:k} = \frac{p(Y_k|X_k)p(X_k|X_{k-1})p(Y_{1:k-1}|X_{0:k-1})p(X_{0:k-1})}{q(X_k|X_{0:k-1}, Y_{1:k})q(X_{0:k-1}|Y_{1:k-1})}$$

$$= \frac{p(Y_k|X_k)p(X_k|X_{k-1})}{q(X_k|X_{0:k-1}, Y_{1:k})} w_{0:k-1} \tag{2.27}$$

Si, de plus, la fonction d'importance satisfait $q(X_k|X_{0:k-1}, Y_{1:k}) = q(X_k|X_{k-1}, Y_k)$, les poids (2.27) dépend uniquement de l'état précédent et de l'observation courante, au sens où

$$w_{0:k} = \frac{p(Y_k|X_k)p(X_k|X_{k-1})}{q(X_k|X_{k-1}, Y_k)} w_{0:k-1} \tag{2.28}$$

On établit ainsi une relation qui permet d'obtenir les poids d'importance de manière séquentielle sans qu'il soit nécessaire de recalculer, à chaque récurrence, l'ensemble des trajectoires.

L'estimation de la loi $p(X_{0:k}|Y_{1:k})$ s'exprime donc de manière séquentielle. Chaque trajectoire $X_{0:k-1}^i$ est prolongée d'un nouvel état X_k^i tiré selon la fonction d'importance. Le poids de chaque trajectoire est ensuite corrigé à partir de l'information apportée par la nouvelle observation. L'approximation de Monte Carlo permet alors d'approcher la distribution jointe $p(X_{0:k}|Y_{1:k})$.

$$p(X_{0:k}|Y_{1:k}) \approx \sum_{i=1}^N w_{0:k}^i \delta(X_{0:k} - X_{0:k}^i), \tag{2.29}$$

Ou la distribution de filtrage $p(X_k | Y_{1:k})$

$$p(X_k | Y_{1:k}) \approx \sum_{i=1}^N w_k^i \delta(X_k - X_k^i), \quad (2.30)$$

Ces développements permettent de définir l'algorithme d'échantillonnage pondéré bayésien séquentiel, encore appelé SIS (Sequential Importance Sampling) résumé ci-dessous, qui construit récursivement un nuage de particules pondérés approchant la loi de filtrage $p(X_k | Y_{1:k})$ à l'instant k . Chaque particule X_{k-1}^i est propagée selon la fonction d'importance, puis les poids sont mis à jour selon (2.28) préalablement à leur normalisation.

Initialisation

- Pour $i = 1, \dots, N$ générer $X_0^i \sim q(X_0)$ avec $w_0^i \propto \frac{p(Y_1 | X_0^i) p(X_0^i)}{q(X_0^i)}$
- Normaliser les poids d'importance $\tilde{w}_0^i = \frac{w_0^i}{\sum_{i=1}^N w_0^i}$

Pour $k = 1, \dots, K$

- Propager les particules
 - Pour $i = 1, \dots, N$, générer $X_k^i \sim q(X_k | X_{0:k-1}^i, Y_{1:k})$ et $X_{0:k}^i = (X_{0:k-1}^i, X_k^i)$
- Mettre à jour des poids
 - Pour $i = 1, \dots, N$, évaluer les poids d'importance

$$w_k^i \propto w_{k-1}^i \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | X_{0:k-1}^i, Y_{1:k})}$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$

- Sortie de l'algorithme : le nuage de particules associées à leurs poids respectifs $(X_{0:k}^i, \tilde{w}_k^i)_{i=1, \dots, N}$ permet d'approcher la loi à posteriori

$$p(X_k | Y_{1:k}) \approx \sum_{i=1}^N \tilde{w}_k^i \delta(X_k - X_k^i)$$

Alg. II-2 Algorithme de filtrage par l'échantillonnage pondéré séquentiel SIS

Cet algorithme très simple d'estimation de la loi de filtrage $p(X_k | Y_{1:k})$ possède l'avantage d'être parallélisable [43]. Toutefois, sa nature récursive soulève certains problèmes. Il souffre des limitations de l'échantillonnage d'importance et notamment d'un problème de dégénérescence du poids des particules dû à l'augmentation inévitable de leur variance mise en évidence par Kong, Liu et Wong [44]. Ainsi au bout de quelques récurrences, on constate que les poids normalisés des particules tendent vers 0 et que seul le poids d'une particule tend vers 1. Evidemment, dans ces conditions, l'estimation de la densité a posteriori en est gravement affectée.

Pour pallier cet inconvénient qui constitue un dysfonctionnement notoire du filtre, plusieurs approches ont été proposées. On a vu comment construire un estimateur récursif de la distribution a posteriori $p(X_{0:k} | Y_{1:k})$ ainsi que de la loi de filtrage $p(X_k | Y_{1:k})$. La précision de l'estimateur peut être qualifiée au moyen de sa variance. Dans cette section, nous nous proposons de discuter certains problèmes relatifs à ce critère, et présentons des méthodes et techniques permettant leur minimisation.

La plus intuitive consiste à augmenter artificiellement le nombre de particules, cette option est rarement couronnée de succès et ne fera donc pas l'objet de développements complémentaires. Deux autres méthodes sont cependant souvent utilisées : le choix d'une distribution d'importance adéquate et l'introduction d'une étape de rééchantillonnage sensée améliorer la diversité des états.

II.4. Choix d'une fonction d'importance adéquate :

En raison de la nature récursive de l'algorithme SIS, la variance inconditionnelle des poids dans le temps augmente [44], induisant une dégradation de la précision de l'estimation. Ceci se traduit par des phénomènes de dégénérescence du nuage, au sens où après un certain nombre d'étapes de récursions, la plupart des particules sont affectées d'un poids normalisé négligeable. La dégénérescence de l'algorithme est d'autant plus sensible que la distribution d'importance $q(X_{0:k} | Y_{1:k})$ diffère de $p(X_{0:k} | Y_{1:k})$, qui correspondrait au cas équipondéré.

Afin de limiter l'augmentation de la variance des poids d'importance, Doucet [42] cherche l'expression d'une distribution d'importance qui minimise cette variance conditionnellement aux mesures et aux états passés. D'après l'expression séquentielle (2.27) des poids d'importance non normalisés de l'échantillonnage bayésien, on établit

$$\begin{aligned}
 E_q [w_k^i] &= \int w_k^i q(X_k^i | Y_{1:k}, X_{0:k-1}^i) dX_k^i \\
 &= \int w_{k-1}^i \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | Y_{1:k}, X_{0:k-1}^i)} q(X_k^i | Y_{1:k}, X_{0:k-1}^i) dX_k^i \\
 &= w_{k-1}^i p(Y_k | X_{k-1}^i)
 \end{aligned} \tag{2.31}$$

D'une manière tout à fait analogue, il est possible d'établir l'expression de la variance conditionnelle des poids d'importance.

$$\begin{aligned}
 \text{Var}_q(w_k^i) &= \mathbb{E}_q \left[(w_k^i)^2 \right] - \left(\mathbb{E}_q [w_k^i] \right)^2 \\
 &= \int (w_k^i)^2 \frac{p^2(Y_k | X_k^i) p^2(X_k^i | X_{k-1}^i)}{q^2(X_k^i | Y_{1:k}, X_{0:k-1}^i)} q(X_k^i | Y_{1:k}, X_{0:k-1}^i) dX_k^i - (w_k^i)^2 p^2(Y_k | X_{k-1}^i) \\
 &= (w_{k-1}^i)^2 \left[\int \frac{p^2(Y_k | X_k^i) p^2(X_k^i | X_{k-1}^i)}{q^2(X_k^i | Y_{1:k}, X_{0:k-1}^i)} dX_k^i - p^2(Y_k | X_{k-1}^i) \right] \tag{2.32}
 \end{aligned}$$

Par ailleurs, on applique bayés, il vient

$$p(X_k^i | Y_k, X_{k-1}^i) = \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{p(Y_k | X_{k-1}^i)} \tag{2.33}$$

La variance est nulle et donc minimale si

$$\int \frac{p^2(Y_k | X_k^i) p^2(X_k^i | X_{k-1}^i)}{q^2(X_k^i | Y_{1:k}, X_{0:k-1}^i)} dX_k^i = p^2(Y_k | X_{k-1}^i)$$

C'est-à-dire,

$$\int \frac{p^2(Y_k | X_{k-1}^i) p^2(X_k^i | Y_k, X_{k-1}^i)}{q^2(X_k^i | Y_{1:k}, X_{0:k-1}^i)} dX_k^i = p^2(Y_k | X_{k-1}^i) \tag{2.34}$$

Soit,

$$\int \frac{p^2(X_k^i | Y_k, X_{k-1}^i)}{q^2(X_k^i | Y_{1:k}, X_{0:k-1}^i)} dX_k^i = 1$$

D'où

$$q(X_k^i | Y_{1:k}, X_{0:k-1}^i) = p(X_k^i | Y_k, X_{k-1}^i) \tag{2.35}$$

La distribution d'importance $q(X_k^i | Y_{1:k}, X_{0:k-1}^i) = p(X_k^i | Y_k, X_{k-1}^i)$ est connue comme étant la distribution d'importance optimale au sens de la minimisation de la variance, conditionnellement aux mesures et aux états passés, des poids d'importance. Donc, Il reste à établir l'expression des poids d'importance lorsque la fonction de proposition est la fonction d'importance optimale.

En reprenant la relation (2.27) qui donne l'expression séquentielle des poids d'importance adaptée à la fonction d'importance optimale, il vient :

$$w_k^i = \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{p(X_k^i | Y_k, X_{k-1}^i)} w_{k-1}^i \quad (2.36)$$

En utilisant (2.33) et en remplaçant dans (2.36), on obtient immédiatement l'expression des poids d'importance lorsque la distribution de proposition n'est autre que la distribution d'importance optimale.

$$w_k^i = w_{k-1}^i p(Y_k | X_{k-1}^i) \quad (2.37)$$

Une propriété remarquable de (2.37) est que les poids w_k^i associé à chaque particule X_k^i ne dépend que de la particule prédécesseur X_{k-1}^i . Il s'en suit que la variance des poids conditionnellement à $Y_{1:k}$ et $X_{0:k-1}^i$ est nulle. De même, le poids d'une particule à un instant k est indépendant de la valeur proposée de la particule au même instant. Les calculs de propagation du nuage particulaire peuvent donc être menés indépendamment de la détermination des poids d'importance. Cette indépendance est de nature à faciliter une implémentation algorithmique de type parallèle.

La mise en oeuvre du filtre à partir de la distribution d'importance optimale requiert d'être en mesure de simuler des échantillons distribués selon la loi $p(X_k | Y_k, X_{k-1})$. Par ailleurs, le calcul des poids d'importance non normalisés nécessite de pouvoir évaluer, à une constante de normalisation près, $p(Y_k | X_{k-1}^i)$. Lorsque le modèle est à bruit additif gaussien et que l'équation de mesure est linéaire, la distribution d'importance est alors une fonction gaussienne dont les paramètres admettent une expression analytique. La simulation d'échantillons ne pose donc guère de problèmes. De plus, le calcul des poids d'importance peut être effectué de manière explicite.

Un second avantage sous-jacent à l'utilisation de la fonction d'importance optimale, est que la propagation du nuage est conditionnée par la mesure à l'instant courant, ce qui contribue à guider les particules vers les zones à forte vraisemblance.

II.5. Techniques de rééchantillonnage :

La seconde grande catégorie de méthodes utilisées pour limiter l'augmentation de la variance des poids d'importance et ainsi éviter la dégénérescence de l'échantillonnage d'importance séquentiel bayésien est de procéder à un rééchantillonnage des particules. L'idée consiste à éliminer les trajectoires de poids normalisés faibles, et, à multiplier les trajectoires dont les poids normalisés sont importants afin de concentrer le nuage particulaire sur les particules de poids les plus forts.

Gordon et al [8] ont, les premiers, introduit une étape de rééchantillonnage dans le cadre de l'échantillonnage d'importance séquentiel. Cette étape de rééchantillonnage, aussi appelée

étape de sélection, permet d'associer à chaque particule $X_{0,k}^i$ un nombre de particules filles N_k^i tel que $\sum_{i=1}^N N_k^i = N$.

Les méthodes potentielles sont, a priori, extrêmement nombreuses. Cependant dans la plupart des publications, le schéma de sélection retenu respecte le postulat suivant $E[N_k^i] = N\tilde{w}_k^i$. Les méthodes possèdent des performances différentes concernant la variance $Var(N_k^i)$. Classiquement, trois principales techniques de rééchantillonnage sont utilisées [45] ; le rééchantillonnage multinomial, le rééchantillonnage résiduel et le rééchantillonnage à variance minimale.

II.5.1. Rééchantillonnage Multinomial :

Cette technique est la plus utilisée, chaque particule est sélectionnée en fonction de l'importance de son poids \tilde{w}_k^i .

La première étape de la procédure consiste à construire la fonction de répartition des particules du nuage. En abscisse, figure l'index de la particule et, en ordonnée, son poids cumulé à celui de toutes les particules précédentes, i.e. les particules de numéro d'index inférieur, voir figure II-2.

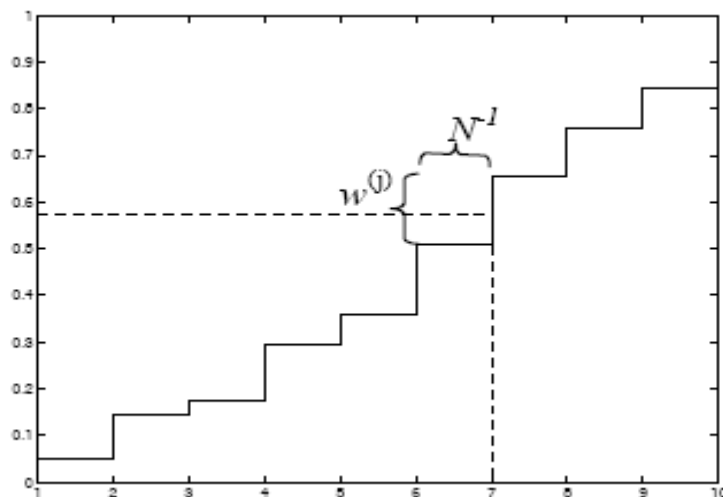


Figure II-2 Echantillonnage multinomial

L'étape suivante procède au tirage successif de l'index de N particules. Pour ce faire, on tire N nombres appartenant à l'intervalle $[0,1]$ et distribués selon une loi uniforme. Supposons que la situation représentée sur la figure II-2 corresponde au rééchantillonnage de la particule d'index i . En effectuant un tirage aléatoire, selon la loi uniforme sur $[0,1]$, le nombre extrait est environ 0.57, ce dernier est ensuite projeté sur la fonction de répartition, l'intersection donne l'index j (ici 7) d'une particule qui est conservée comme nouvel échantillon d'indice i . Son poids est réinitialisé à la valeur $1/N$. La probabilité de tirer une particule quelconque à chacun des tirages correspond à son poids normalisé \tilde{w}_k^j . Il apparaît

clairement que l'effectif de la descendance d'une particule donnée (après rééchantillonnage) est proportionnel au poids normalisé de la particule. Globalement les effectifs N_k^i des particules, à l'issue de rééchantillonnage, correspondent à une réalisation de la loi multinomiale des paramètres N et \tilde{w}_k^i avec $i = 1, \dots, N$. La variance de l'effectif de la descendance d'une particule i est donnée par la relation classique de la variance d'une loi binomiale de paramètres N et \tilde{w}_k^i .

$$Var(N_k^i) = N \tilde{w}_k^i (1 - \tilde{w}_k^i) \quad (2.38)$$

Cependant, il est possible de réduire la variance par d'autres schémas de sélection.

II.5.2. Rééchantillonnage résiduel :

Cette technique est une amélioration du tirage multinomial, dans le but de diminuer la variance $Var(N_k^i)$ des particules dupliquées. Elle comprend deux étapes successives : la première est déterministe [10] et la seconde est basée sur l'échantillonnage multinomial décrit précédemment.

- ✓ Dans un premier temps, le nombre de descendant d'une particule i est fixé de manière déterministe à la partie entière de l'espérance c'est-à-dire $\tilde{N}_k^i = \lfloor N \tilde{w}_k^i \rfloor$
- ✓ On réalise ensuite un rééchantillonnage multinomial des particules restantes au nombre de $N - \sum_{n=1}^N \tilde{N}_k^n$ afin de conserver un nombre de particules constant au cours des récurrences. Les poids respectif de chacune des particules est fixé à $\left(N - \sum_{n=1}^N \tilde{N}_k^n \right)^{-1} (N \tilde{w}_k^i - \tilde{N}_k^i)$, c'est-à-dire à une valeur proportionnelle à la partie non entière de l'espérance du nombre de descendants de la particule i .

Finalement, le nombre de descendants d'une particule i quelconque correspond à la partie entière de son espérance \tilde{N}_k^i ajouté au nombre de fois où la particule en question a été sélectionnée au cours de la seconde phase de rééchantillonnage.

La variance du nombre de descendants d'une particule est alors donnée uniquement par la variance introduite dans la seconde phase, compte tenu du caractère déterministe de la première phase.

$$Var(N_k^i) = \left(N - \sum_{n=1}^N \tilde{N}_k^n \right) \frac{(N \tilde{w}_k^i - \tilde{N}_k^i)}{\left(N - \sum_{n=1}^N \tilde{N}_k^n \right)} \left(1 - \frac{(N \tilde{w}_k^i - \tilde{N}_k^i)}{\left(N - \sum_{n=1}^N \tilde{N}_k^n \right)} \right) \quad (2.39)$$

La variance est obtenue d'une manière similaire à celle obtenue dans le cas du rééchantillonnage multinomial, c'est à dire à partir de l'expression générale de la variance d'une loi binomiale dont les paramètres sont $N - \sum_{n=1}^N \tilde{N}_k^n$ et $\left(N - \sum_{n=1}^N \tilde{N}_k^n\right)^{-1} (N\tilde{w}_k^i - \tilde{N}_k^i)$.

II.5.3. Rééchantillonnage à variance minimale :

Les méthodes considérées ici limitent la différence à 1 entre le nombre de descendants effectifs d'une particule et le nombre théorique. En guise d'exemple, on peut citer le rééchantillonnage stratifié [10] qui entre dans cette catégorie. La procédure est relativement simple : il suffit d'échantillonner l'intervalle $[0,1]$ à l'aide de N points espacés de N^{-1} , puis de considérer les poids d'importance normalisés comme des intervalles contigus de $[0,1]$. Le nombre de descendants N_k^i de chacune des particules est alors donné par le nombre de points contenus dans l'intervalle $\left[\sum_{j=1}^{i-1} \tilde{w}_k^j; \sum_{j=1}^i \tilde{w}_k^j\right]$.

En utilisant les notations définies pour le rééchantillonnage résiduel, pour chaque particule, le nombre de descendants peut prendre deux valeurs soit \tilde{N}_k^i , partie entière de l'espérance de N_k^i , soit $\tilde{N}_k^i + 1$. La loi suivie par le nombre de descendants est donc une loi de Bernoulli de paramètre $(N\tilde{w}_k^i - \tilde{N}_k^i)$ et qui représente la partie décimale de l'espérance de N_k^i . L'expression de la variance de descendants d'une particule s'en déduit

$$Var(N_k^i) = (N\tilde{w}_k^i - \tilde{N}_k^i) \left(1 - (N\tilde{w}_k^i - \tilde{N}_k^i)\right) \quad (2.40)$$

La variance obtenue par cette méthode est inférieure à celle obtenue par la méthode du rééchantillonnage résiduel.

Selon Merwe et al [46], la méthode de rééchantillonnage utilisée n'influe pas significativement sur le résultat final. Cependant, un autre critère semble plus problématique : il s'agit de la fréquence à laquelle est effectué ce rééchantillonnage.

Les filtres particuliers utilisant le rééchantillonnage s'appellent les filtres SIR (Sampling Importance Resampling) [8] ou les filtres IPF (Interacting Particle Filter) [9, 47]. Nous présentons ci-dessous l'algorithme du filtrage particulier SIR.

II.6. Echantillonnage pondéré séquentiel avec rééchantillonnage (SIR) :

En définissant une structure générique basée sur l'algorithme SIS, auquel on ajoute une étape de rééchantillonnage, l'algorithme obtenu est couramment désigné sous l'acronyme de SIR, et constitue le fondement des méthodes séquentielles de Monte-Carlo (MSMC) encore appelées filtres particuliers. Ces méthodes consistent à représenter la loi a posteriori $p(X_{0:k} | Y_{1:k})$ par une mesure discrète constituée d'une somme finie de masses de Dirac cumulées en des points dénommés particules et pondérées par un coefficient appelé poids d'importance. Les méthodes particulières permettent de faire évoluer le nuage de particules et

les poids associés en utilisant récursivement l'échantillonnage pondéré associé à une technique de rééchantillonnage. Un maillage stochastique et adaptatif de l'espace d'état est ainsi mis en oeuvre. La structure d'un filtre particulière générique dont l'étape de rééchantillonnage est réalisée à chaque récurrence est définie par Alg. II-3.

Initialisation

- Pour $i = 1, \dots, N$, générer $X_0^i \sim p(X_0)$ avec $w_0^i = 1/N$

Pour $k = 1, \dots, K$

- Propagation des particules
 - Pour $i = 1, \dots, N$, générer $X_k^i \sim q(X_k | X_{0:k-1}^i, Y_{1:k})$ et $\tilde{X}_{0:k}^i = (X_{0:k-1}^i, X_k^i)$
- Mise à jour des poids
 - Pour $i = 1, \dots, N$, évaluer les poids d'importance

$$w_k^i \propto w_{k-1}^i \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | X_{0:k-1}^i, Y_{1:k})}$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$
- Rééchantillonnage
 - Pour $i = 1, \dots, N$, tirer avec remise $X_{0:k}^i$ parmi $(\tilde{X}_{0:k}^i)_{i=0:N}$ avec une probabilité \tilde{w}_k^i et fixer le nouveau poids d'importance à $1/N$.

$$X_{0:k}^i \sim \sum_{i=1}^N \tilde{w}_k^i \delta(X_k - \tilde{X}_k^i) \text{ et } w_k^i = 1/N.$$

- Sorties de l'algorithme : Le nuage de particules associées à leurs poids respectifs $(X_{0:k}^i, N^{-1})_{i=1, \dots, N}$ permet d'approcher la loi a posteriori

$$p(X_{0:k} | Y_{1:k}) \approx \frac{1}{N} \sum_{i=1}^N \delta_{X_{0:k}^i} (dX_{0:k})$$

Alg. II-3 Algorithme d'échantillonnage pondéré séquentiel avec rééchantillonnage systématique.

Le filtre particulaire présente des comportements instables dues particulièrement aux cas de bruit de dynamique ou de mesure faible. En effet, après chaque rééchantillonnage, des particules peuvent occuper la même position, si le bruit de dynamique est faible. Les particules n'occupent alors qu'une petite zone restreinte de l'espace d'état. Ainsi, les particules perdent leur capacité d'exploration de l'espace d'état, ceci provoque la divergence du filtre.

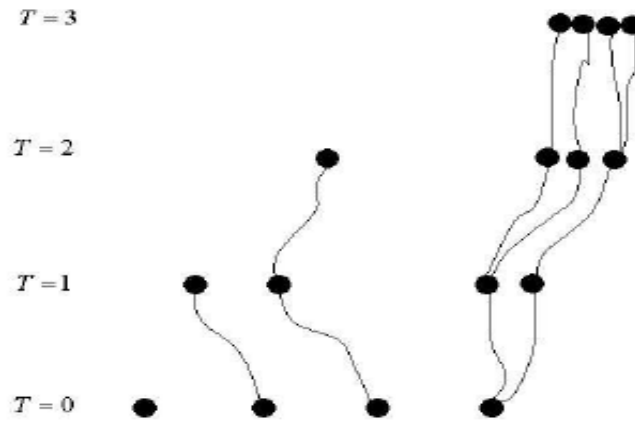


Figure II-3 Dégénérescence des particules : faible bruit d'état.

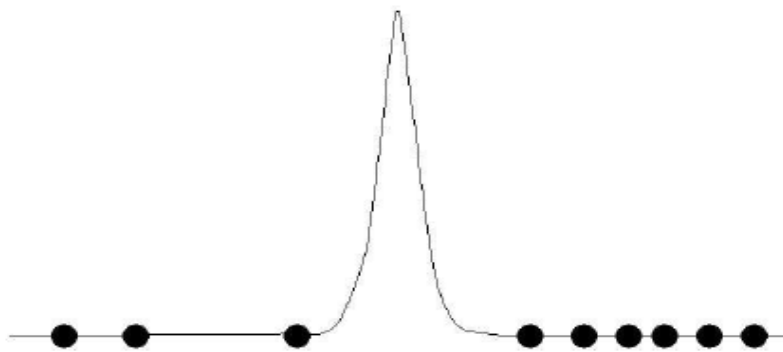


Figure II-4 Dégénérescence des particules : faible bruit d'observation.

Dans la figure II-3, chaque particule est pondérée suivant sa vraisemblance puis sélectionnée ou abandonnée. Les particules sélectionnées suivent une évolution suivant le noyau de transition. On constate, qu'à l'instant $T = 3$, toutes les particules sont regroupées, et viennent toutes de la même particule initiale. Le filtre particulaire a peu de chance de retrouver l'état, ce celui-ci n'est pas dans la petite zone couverte par les particules. Dans le cas de bruit de mesure faible figure II-4, la vraisemblance se concentre dans une petite zone dépourvue de particules. Il est dans ce cas impossible d'exploiter efficacement l'information apportée par l'observation.

Dans ces deux cas extrêmes, la nature discrète de l'approximation particulaire limite considérablement la capacité exploratoire du filtre particulaire.

Il convient de mentionner que le rééchantillonnage ne comporte pas uniquement des aspects positifs, car l'introduction d'une étape de rééchantillonnage dans l'algorithme SIR remet en cause sa parallélisation de manière aussi immédiate que pour le SIS, et complexifie considérablement les preuves relatives à sa convergence. En effet, les particules étant statistiquement dépendantes à l'issue du rééchantillonnage, les résultats classiques de convergence des méthodes de Monte Carlo, qui reposent sur l'hypothèse indépendant et identiquement distribué (i.i.d), ne permettent pas de conclure [48, 49].

II.6.1. Inconvénient du rééchantillonnage :

Bien que l'étape de rééchantillonnage permette de limiter la dégénérescence de l'algorithme en limitant l'augmentation de la variance des poids d'importance, son caractère stochastique introduit des erreurs. De plus, l'implémentation parallèle est rendue plus complexe en raison du rééchantillonnage qui rompt l'évolution structurée du nuage et impose par conséquent l'implémentation de procédures spécifiques (augmentation du temps de calcul) [42] afin d'assurer le transfert de particules entre modules de traitement. D'un point de vue plus théorique, la mise en oeuvre de l'étape de rééchantillonnage rompt l'indépendance des particules et ne permet plus d'utiliser les théorèmes classiques concernant la convergence (théorème central limite ou loi des grands nombres). Dans le cas particulier où le rééchantillonnage des particules est effectué de manière systématique, des résultats de convergence ont néanmoins été établis [50].

De plus, Hue [51] a montré que l'échantillonnage systématique effectué à chaque récurrence est entaché, par rapport à un échantillonnage adaptatif, de défauts importants

- ✓ Appauvrissement plus prononcé du nuage.
- ✓ Estimées fournies par le filtre moins fiables.
- ✓ Charge de calcul notoirement plus important.

L'idéal est de réaliser un compromis concernant la fréquence du rééchantillonnage en ne faisant intervenir ce dernier que lorsque cela s'avère nécessaire. Cette manière de procéder permet aux particules d'explorer l'espace au cours de plusieurs récurrences avant d'être éventuellement supprimées. Le critère le plus couramment utilisé, qui est plus simple et moins coûteuse en temps de calcul, et qui conditionne l'étape de rééchantillonnage est la taille effective de nuage N_{eff} définie par

$$N_{eff} = \frac{N}{1 + Var_q(w_k^i)} \quad (2.41)$$

L'expression de N_{eff} ne peut pas être calculée directement, mais une estimation est donnée par :

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{w}_{0:k}^i)^2} \quad (2.42)$$

Il donne la taille équivalente du nuage engendré par la distribution d'importance au nuage généré à partir de la distribution objectif. Les valeurs maximale $N_{eff} = N$ et minimale $N_{eff} = 1$ de N_{eff} correspondent respectivement à l'équipondération des particules $w_k^i = \frac{1}{N}$ et au cas de dégénérescence extrême.

L'échantillonnage adaptatif consiste alors à ne rééchantillonner les particules que lorsque l'estimateur de la taille effective est inférieur à un seuil donné : $\hat{N}_{eff} < N_{seuil}$.

Notons que, $1 \leq \hat{N}_{eff} \leq N$, où la limite supérieure est atteinte quand toutes les particules ont le même poids, tandis que la limite inférieure est due quand toute la masse de probabilité est en une seul particule. Dans [46], $N_{seuil} = 2N/3$.

II.7. Echantillonnage pondéré séquentiel avec rééchantillonnage adaptatif :

Le SIR à rééchantillonnage adaptatif a le même algorithme que celui de SIR standard avec une étape de rééchantillonnage adaptatif.

Initialisation

- Pour $i = 1, \dots, N$, générer $X_0^i \sim p(X_0)$ avec $w_0^i = 1/N$

Pour $k = 1, \dots, K$

- Propagation des particules
 - Pour $i = 1, \dots, N$, générer $X_k^i \sim q(X_k^i | X_{0:k-1}^i, Y_{1:k})$ et $\tilde{X}_{0:k}^i = (X_{0:k-1}^i, X_k^i)$
- Mise à jour des poids
 - Pour $i = 1, \dots, N$, évaluer les poids d'importance

$$w_k^i \propto w_{k-1}^i \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | X_{0:k-1}^i, Y_{1:k})}$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$.

- Rééchantillonnage

➤ Evaluer la taille effective de l'échantillon $\hat{N}_{eff} = \frac{1}{\sum_{n=1}^N (\tilde{w}_k^n)^2}$

➤ Si $\hat{N}_{eff} < N_{seuil}$ alors pour $i = 1, \dots, N$,

$$X_{0:k}^i \sim \sum_{i=1}^N \tilde{w}_k^i \delta(X_k - \tilde{X}_k^i) \text{ et } w_k^i = 1/N.$$

Sinon $X_{0:k}^i = \tilde{X}_{0:k}^i$ et les poids d'importance $w_k^i = \tilde{w}_k^i$ reste inchangés.

- Sorties de l'algorithme : Le nuage de particules associées à leurs poids respectifs $(X_{0:k}^i, w_k^i)_{i=1, \dots, N}$ permet d'approcher la loi a posteriori

$$p(X_{0:k} | Y_{1:k}) \approx \sum_{i=1}^N w_k^i \delta_{X_{0:k}^i} (dX_{0:k}).$$

Alg. II-4 Algorithme d'échantillonnage pondéré séquentiel avec rééchantillonnage adaptatif.

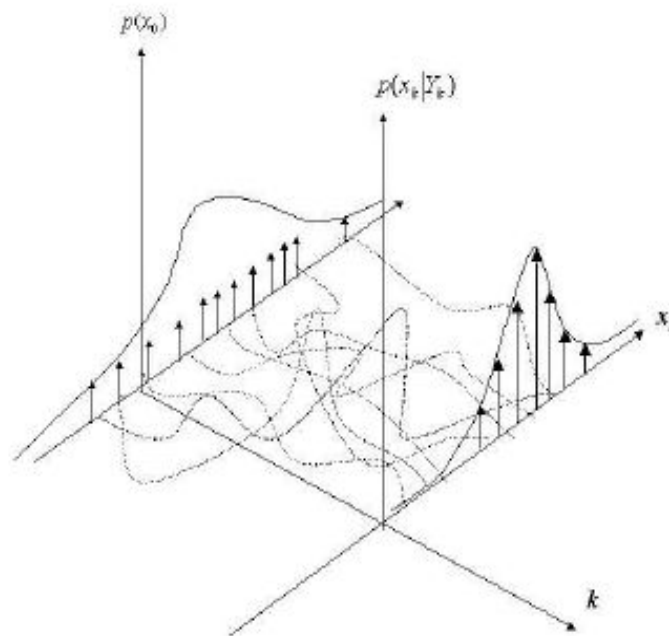


Figure II-5 Evolution de la densité conditionnelle

Tous les filtres particuliers peuvent être considérés comme une instance de SIR et ne diffèrent que par le choix de la fonction d'importance et par le choix de la technique de rééchantillonnage mise en œuvre. Il convient toutefois de signaler, pour une meilleure estimation de l'intégrale $I(\varphi)$ définie en (2.14) au moyen de (2.15), qu'il est préférable de ne pas procéder au rééchantillonnage à chaque instant et que l'évaluation de (2.15) doit être effectuée de préférence avant le rééchantillonnage.

Avant de se pencher sur les différentes évolutions et variantes proposées pour le SIR, nous détaillons les étapes du filtre particulier connu sous la terminologie de bootstrap.

II.8. Le Bootstrap :

De l'avis de certains auteurs [53, 8], et dans le cadre de l'échantillonnage pondéré séquentiel, on choisit comme distribution d'importance, la loi d'évolution a priori du système

$$q(X_k | X_{0:k-1}, Y_{1:k}) = p(X_k | X_{k-1}) \quad (2.43)$$

L'expression des poids d'importance des mesures devient alors

$$w_k^i = p(Y_k | X_k^i) w_{k-1}^i \quad (2.44)$$

En adaptant la procédure SIR à échantillonnage systématique aux hypothèses présentes ci-dessus, on déduit l'algorithme Bootstrap, qui est très utilisé dans de multiples domaines.

Initialisation

- Pour $i = 1, \dots, N$, générer $X_0^i \sim p(X_0)$ avec $w_0^i = 1/N$

Pour $k = 1, \dots, K$

- Propagation des particules
 - Pour $i = 1, \dots, N$, générer $X_k^i \sim q(X_k | X_{0:k-1}^i)$ et $\tilde{X}_{0:k}^i = (X_{0:k-1}^i, X_k^i)$
- Mise à jour des poids
 - Pour $i = 1, \dots, N$, évaluer les poids d'importance

$$w_k^i \propto w_{k-1}^i p(Y_k | \tilde{X}_k^i)$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$

- Rééchantillonnage

➤ Pour $i = 1, \dots, N$, tirer avec remise $X_{0:k}^i$ parmi $(\tilde{X}_{0:k}^i)_{i=0:N}$ avec une probabilité \tilde{w}_k^i et fixer le nouveau poids d'importance à $1/N$.

$$X_{0:k}^i \sim \sum_{i=1}^N \tilde{w}_k^i \delta(X_k - \tilde{X}_k^i) \text{ et } w_k^i = 1/N.$$

- Sorties de l'algorithme : Le nuage de particules associées à leurs poids respectifs $(X_{0:k}^i, N^{-1})_{i=1, \dots, N}$ permet d'approcher la loi a posteriori

$$p(X_{0:k} | Y_{1:k}) \approx \frac{1}{N} \sum_{i=1}^N \delta_{X_{0:k}^i} (dX_{0:k})$$

Alg. II-5 L'algorithme Bootstrap

Le schéma ci-dessous représente le principe de Bootstrap.

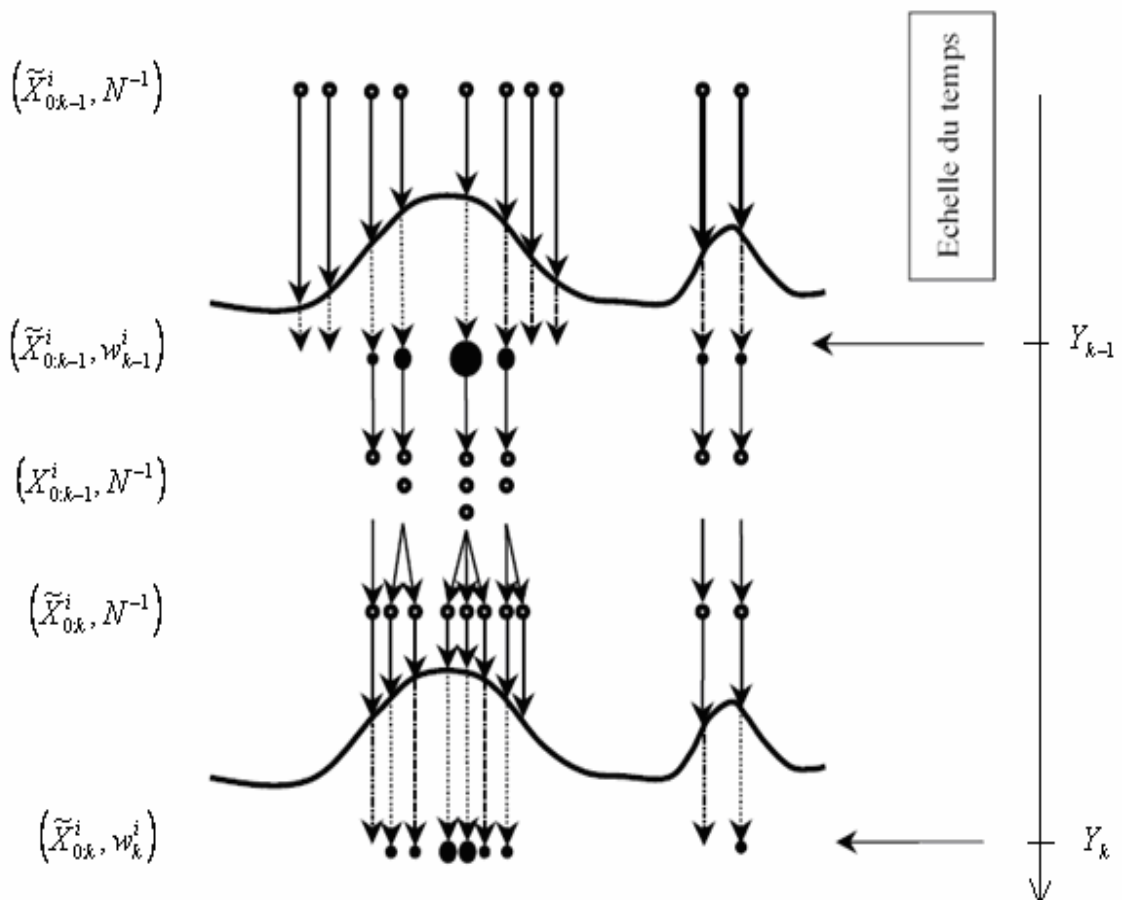


Figure II-6 Le principe de Bootstrap

Le filtre Bootstrap possède quelques propriétés intéressantes qui justifient l'intérêt qui lui est porté

- ✓ Il est rapidement et facilement implantable.
- ✓ Il est aisément adaptable à un nouveau problème, il suffit de modifier, au niveau du programme, la distribution d'importance ainsi que les poids d'importance.

L'inconvénient qu'on peut avoir avec de telle algorithmes, est dans le cas où la distribution prédite ne coïncide pas avec la vraisemblance, le risque d'assister à une dégénérescence est irréversible à cause de la concentration des particules dans des zones de faibles vraisemblance. Cependant, il existe des techniques spécifiques qui permettent de limiter l'impacte de tels comportements. Elles consistent à diriger les particules vers des zones privilégiées en utilisant des distributions intermédiaires entre la distribution prédite et la vraisemblance.

II.9. Limitations des méthodes SMC :

Le bon fonctionnement de toute méthode particulière est lié à la capacité qu'a la loi d'évolution à prédire les particules dans les régions de forte vraisemblance.

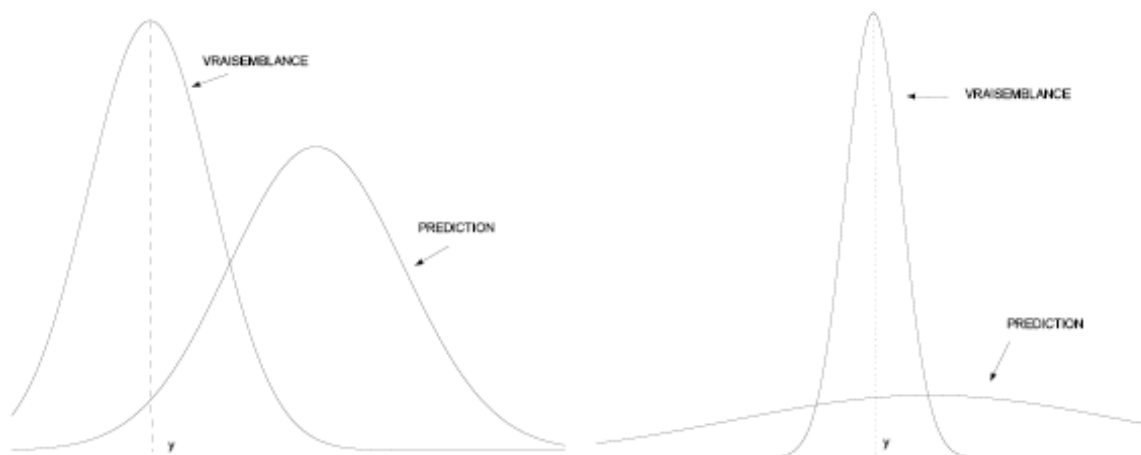


Figure II-7 Configuration critique de la densité a priori et la vraisemblance

La figure II-7 présente alors deux configurations critiques de la loi prédite par rapport à la vraisemblance de la mesure.

La figure II-7 à gauche, suppose un modèle de mesure pour lequel le bruit de mesure est faible et, au contraire, un modèle d'état approximatif ce qui se traduit par une grande dispersion des particules prédites par rapport à la mesure. Dans ces conditions, seules les particules prédites à proximité immédiate de la mesure auront un poids d'importance significatif, si bien qu'au cours de la phase de rééchantillonnage de nombreuses particules seront éliminées et quelques particules seront dupliquées un grand nombre de fois. Ce phénomène conduit irrémédiablement à un appauvrissement des états. Ce déficit de diversité des particules est hautement préjudiciable aux performances du filtre en matière d'estimation.

La figure II-7 à droite, illustre le second dysfonctionnement majeur des méthodes particulières. Il intervient lorsque les supports respectifs de la distribution de propagation a priori d'une part et la vraisemblance ne coïncident pas. Avant l'étape de rééchantillonnage, les poids d'importance sont tous à peu près égaux mais possèdent une valeur faible. Le rééchantillonnage dans ces conditions, en présence de particules sensiblement équipondérées, bien que conservant la diversité des états, ne permet pas de sélectionner des particules correspondant à des hypothèses réalistes. Ce phénomène de non coïncidence de l'a priori et de la vraisemblance, est particulièrement sensible en présence d'une mesure aberrante, ou encore lorsque les variances des bruits dynamique et de mesure sont faibles. On parle alors de dégénérescence des particules. [54, 55]

Afin de tenter, d'apporter des solutions ou au moins de réduire les inconvénients évoqués précédemment et inhérents à la mise en oeuvre des algorithmes séquentiels de Monte-Carlo, des variantes du schéma originel du SIR ont été présentées dans la littérature. Ces améliorations sont destinées avant tout à limiter l'appauvrissement des états.

II.10. Variantes du SIR :

Afin d'apporter des améliorations au S.I.R plusieurs variantes du filtrage particulières ont été proposées, il s'agit du filtre particulière à variable auxiliaire, du filtre particulière régularisé et une méthode de réduction de variance connue sous le nom de Rao-Blackwellisation qu'on va étudier en détails une par une.

II.10.1. Echantillonnage pondéré séquentiel à variable auxiliaire (APF):

Les méthodes de Monte-Carlo séquentielles utilisant le concept de variable auxiliaire ASIR, provient des travaux de [12].

A un instant k , l'Algorithme SIS fournit une approximation de la densité prédite suivant

$$p_N(X_{k+1}|Y_{1:k}) = \sum_{i=1}^N w_k^i p(X_{k+1}|X_k^i) \quad (2.45)$$

La densité de filtrage définie comme la marginale de la loi a posteriori s'écrit, compte tenu de la propriété de Bayes

$$p(X_{k+1}|Y_{1:k+1}) = \frac{p(Y_{k+1}|X_{k+1}, Y_{1:k})p(X_k|Y_{1:k})}{p(Y_k|Y_{1:k-1})} \propto p(Y_{k+1}|X_{k+1}, Y_{1:k})p(X_{k+1}|Y_{1:k}) \quad (2.46)$$

Compte tenu de l'indépendance des observations conditionnellement aux états, l'expression d'un estimateur de la densité de filtrage devient

$$p_N(X_{k+1}|Y_{1:k+1}) \propto p(Y_{k+1}|X_{k+1}) \sum_{i=1}^N w_k^i p(X_{k+1}|X_k^i) \quad (2.47)$$

En reprenant l'idée classique qui consiste, lorsque l'on désire générer un échantillon distribué selon un mélange de lois, à indexer les composantes du mélange par une variable dite

auxiliaire, on peut considérer que le mélange en question est celui donné par la relation (2.45). Compte tenu des remarques précédentes, la nouvelle loi objective est donc la loi conjointe $p(X_k, i | Y_{1:k})$. Les autres hypothèses demeurant inchangées, i représente l'index des particules. On a

$$p(X_{k+1}, i | Y_{1:k+1}) \propto w_k^i p(Y_{k+1} | X_{k+1}) p(X_{k+1} | X_k^i) \quad (2.48)$$

On appelle i la variable auxiliaire, son rôle consiste uniquement à faciliter la simulation d'un échantillon constitué des paires $\{X_{k+1}^j, i^j\}_{j=1}^N$, i^j est l'indice de la particule à l'instant k . En éliminant l'indice i , on obtient un échantillon $\{X_{k+1}^j\}_{j=1}^N$ distribué selon la marginale $p(X_k | Y_{1:k})$ de la loi à posteriori.

L'utilisation d'une variable auxiliaire est un principe général pouvant être mis en oeuvre à partir d'autres méthodes de simulation de lois. Dans le cas du SIR, on génère un échantillon tel que

$$X_{k+1}^j, i^j \sim q(X_{k+1}, i | Y_{1:k+1}) \quad (2.49)$$

Les poids d'importance sont alors donnés par

$$w_{k+1}^j \propto w_k^{i^j} \frac{p(Y_{k+1} | X_{k+1}^j) p(X_{k+1}^j | X_k^{i^j})}{q(X_{k+1}^j, i^j | Y_{1:k+1})}, \text{ et } \tilde{w}_{k+1}^j = \frac{w_{k+1}^j}{\sum_{n=1}^N w_{k+1}^n} \quad (2.50)$$

Le choix de la fonction d'importance $q(X_{k+1}, i | Y_{1:k+1})$ peut être extrêmement large. Néanmoins pitt et shephard [12] adoptent la fonction d'importance suivante et potentiellement bien adaptée à de nombreux cas.

$$q(X_{k+1}, i) \propto p(Y_{k+1} | \mu_{k+1}^i) p(X_{k+1} | X_k^i) w_k^i \quad (2.51)$$

μ_{k+1}^i est une valeur caractéristique de la densité de $X_{k+1} | X_k^i$ et peut être, par exemple la moyenne $\mu_{k+1}^i = E[X_{k+1} | X_k^i]$ ou encore un échantillon $\mu_{k+1}^i \sim p(X_{k+1} | X_k^i)$.

D'après Bayes, on a $q(X_{k+1}, i) = q(i | Y_{1:k+1}) q(X_{k+1} | i, Y_{1:k+1})$.

De plus, en posant $q(X_{k+1} | i, Y_{1:k+1}) \hat{=} p(X_{k+1} | X_k^i)$, on obtient $q(i | Y_{1:k+1}) \propto p(Y_{k+1} | \mu_{k+1}^i) w_k^i$.

La procédure d'échantillonnage suivante de la loi conjointe $q(X_{k+1}, i | Y_{1:k+1})$ comprend alors deux étapes, la première consiste à échantillonner un index i selon la loi $p(Y_{k+1} | \mu_{k+1}^i) w_k^i$ puis simuler (procédure SIS) selon la densité de transition $p(X_{k+1} | X_k^i)$ conditionnellement à la

particule origine i . En considérant l'équation (2.45) et en remplaçant la fonction d'importance conjointe par son expression explicitée, on obtient les poids d'importance

$$w_{k+1}^j \propto \frac{p(Y_{k+1} | X_{k+1}^j)}{p(Y_{k+1} | \mu_{k+1}^{i^j})} \quad (2.52)$$

La simulation s'effectue à partir de particules dont la vraisemblance prédite est grande, ce qui, en principe, doit guider les particules vers les zones d'intérêt et ainsi limiter la dégénérescence des poids d'importance.

L'algorithme ASIR qui découle de paragraphe précédent est le suivant.

Initialisation

- Pour $i = 1, \dots, N$, générer $X_0^i \sim p(X_0)$ avec $w_0^i = 1/N$

Pour $k = 1, \dots, K$

- Mise à jour des poids d'importance
 - Pour $i = 1, \dots, N$, calculer μ_k^i et les poids d'importance

$$w_k^i = q(i | Y_{1:k}) \propto w_{k-1}^i p(Y_k | \mu_k^i)$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$

- Rééchantillonnage des indices

- Pour $j = 1, \dots, N$, tirer avec remise i^j proportionnellement aux poids \tilde{w}_k^i

- Pour $j = 1, \dots, N$

- Générer

$$X_{0:k}^i \sim q(X_k | i^j, Y_{1:k}) = p(X_k | X_{k-1}^{i^j})$$

- Calculer les poids d'importance

$$w_k^j \propto \frac{p(Y_k | X_k^j)}{p(Y_k | \mu_k^{i^j})}$$

- Pour $j = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^j = \frac{w_k^j}{\sum_{j=1}^N w_k^j}$
- Sorties de l'algorithme : Le nuage de particules associées à leurs poids respectifs $(X_{0:k}^j, \tilde{w}_k^j)_{j=1, \dots, N}$ permet d'approcher la loi a posteriori

$$p(X_{0:k} | Y_{1:k}) \approx \sum_{i=1}^N \tilde{w}_k^i \delta_{X_{0:k}^i} (dX_{0:k})$$

Alg. II-6 L'algorithme APF

Comme nous l'avion déjà précisé, l'échantillonnage selon la loi conjointe $q(X_{k+1}, i | Y_{1:k+1})$ s'effectue en deux étapes. Au cours de ces deux étapes, la taille des échantillons produits peut éventuellement être différente. Dans ce cas, il faut rajouter une étape de rééchantillonnage à la fin du l'algorithme avant de poursuivre par une nouvelle récurrence.

Dans les cas difficiles où la vraisemblance et la densité a priori ne coïncident pas, l'ASIR donne de meilleurs résultats que le SIR. Par contre lorsque les lois coïncident, il est alors préférable de privilégier le SIR qui bénéficie d'une variance d'échantillons moins élevée en raison de l'absence de variable auxiliaire.

II.10.2. Echantillonnage pondéré séquentiel régularisé (RPF) :

Le fait que l'étape de rééchantillonnage, dans le cas du SIR, soit entreprise à partir d'une distribution de probabilité discrète a pour effet de concentrer les particules vers des points. On parlera alors de points d'accumulation. Si, de plus, le bruit d'état ou de mesure est faible, le risque d'une raréfaction des points d'accumulation devient sévère et entraîne fatalement un appauvrissement des états. Le filtre particulaire régularisé RPF préconisé par Musso, Oudjane et Legland [11] apporte une réponse à ce problème.

La différence essentielle avec le SIR réside dans la procédure de rééchantillonnage qui s'effectue pour le SIR à partir d'une marginale de la loi discrète de probabilité a posteriori

$$p(X_k | Y_{1:k}) \approx \frac{1}{N} \sum_{i=1}^N \delta_{X_k^i} (dX_k)$$

tandis que, pour le RPF, le rééchantillonnage est réalisé à partir

d'une approximation continue de la densité de filtrage $p(X_k | Y_{1:k})$. La théorie correspondante fait appel à l'estimation de la densité par noyaux. Le but est de procéder à la régularisation d'une mesure discrète par une mesure continue obtenue par la convolution de la mesure discrète par un noyau.

Un noyau de régularisation K est une densité de probabilité symétrique définie sur R^n , avec n la dimension du vecteur X , et jouissant des propriétés suivantes

$$K \geq 0, \quad \int K(X) dX = 1, \quad \int X K(X) dX = 0, \quad \int \|X\|^2 K(X) dX < \infty \tag{2.53}$$

De plus, pour toute largeur $h > 0$, on définit le noyau redimensionné à partir de

$$K_h(X) = \frac{1}{h^n} K\left(\frac{X}{h}\right) \tag{2.54}$$

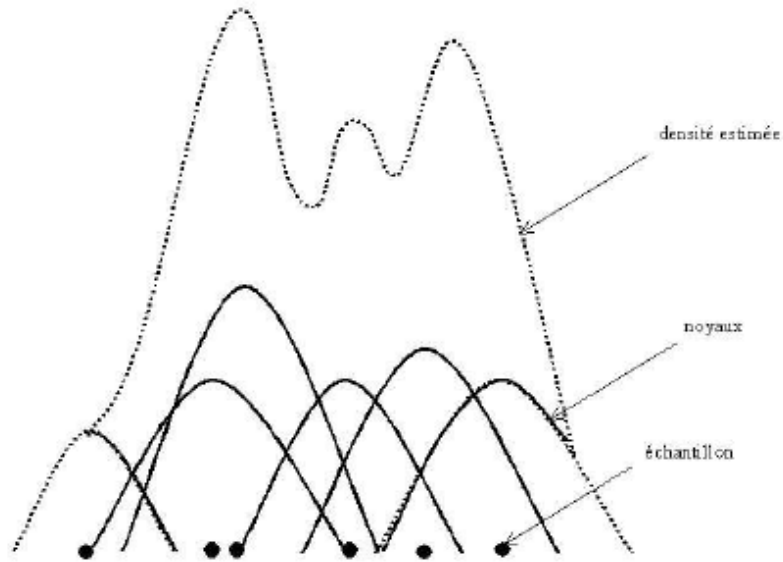


Figure II-8 Mélange de gaussiennes

Si on considère la loi discrète de probabilité $\sum_{i=1}^N \tilde{w}_k^i \delta_{X_k^i}(dX_k)$, on obtient la régularisation en effectuant la convolution de la loi discrète par le noyau de régularisation. On a

$$p_N(X_k | Y_{1:k}) = \int K_h(X_k - u) \sum_{i=1}^N \tilde{w}_k^i \delta_{X_k^i}(dX_k) du = \sum_{i=1}^N \tilde{w}_k^i K_h(X_k - X_k^i) \tag{2.55}$$

Le noyau K et la largeur sont déterminés à partir de la minimisation de l'erreur quadratique intégrée moyenne, définie par

$$MISE = E \left[\int P_N(X_k | Y_{1:k}) - p(X_k | Y_{1:k}) \right]^2 dX_k \tag{2.56}$$

Dans le cas où la loi discrète est équipondérée, il s'avère que le noyau optimal, au sens de la minimisation du critère MISE, correspond au noyau optimal d'Epanechnikov généralement défini par :

$$K_{opt} = \begin{cases} \frac{n+2}{2c_n} (1 - \|X\|^2), & \text{Si } \|X\| < 1 \\ 0, & \text{Sinon} \end{cases} \quad (2.57)$$

c_n est le volume de la sphère unité dans l'hyperespace R^n . Par ailleurs, dans le cas où la densité de filtrage sous-jacente est gaussienne et de variance unitaire, la valeur optimale de la largeur est donnée par

$$h_{opt} = A(K)N^{1/(n+4)} \quad \text{avec} \quad A(K) = \left[8c_n^{-1} (n+4) (2\sqrt{\pi})^n \right]^{1/(n+4)} \quad (2.58)$$

Les résultats donnés par (2.57) et (2.58) sont uniquement optimaux dans les cas d'une part d'une loi discrète équipondérée et d'autre part si la loi de filtrage $p(X_k | Y_{1:k})$ est gaussienne. Dans le cas général, il suffit de ne pas perdre de vue que les résultats sont alors forcément sous optimaux en raison des hypothèses effectuées sur la détermination du noyau optimal et la largeur de bande.

Alg.II-7 donne l'algorithme de filtrage particulaire régularisé.

Initialisation

- Pour $i = 1, \dots, N$, générer $X_0^i \sim p(X_0)$ avec $w_0^i = 1/N$

Pour $k = 1, \dots, K$

- Propagation des particules
 - Pour $i = 1, \dots, N$, générer $X_k^i \sim q(X_k^i | X_{0:k-1}^i, Y_{1:k})$ et $\tilde{X}_{0:k}^i = (X_{0:k-1}^i, X_k^i)$
- Mise à jour des poids
 - Pour $i = 1, \dots, N$, évaluer les poids d'importance

$$w_k^i \propto w_{k-1}^i \frac{p(Y_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | X_{0:k-1}^i, Y_{1:k})}$$

- Pour $i = 1, \dots, N$, normaliser les poids d'importance $\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^N w_k^i}$
- Rééchantillonnage

➤ Evaluer la taille effective de l'échantillon $\hat{N}_{eff} = \frac{1}{\sum_{n=1}^N (\tilde{w}_k^n)^2}$

➤ Si $\hat{N}_{eff} < N_{seuil}$

✓ Calculer la matrice de covariance empirique S_k du nuage $(\tilde{X}_{0:k}^i, \tilde{w}_k^i)_{i=1}^N$

✓ Calculer D_k tel que $S_k = D_k D_k^T$.

✓ Rééchantillonner, Pour $i = 1, \dots, N$

$$X_{0:k}^i \sim \sum_{i=1}^N \tilde{w}_k^i \delta(X_k - \tilde{X}_k^i) \quad \text{et} \quad w_k^i = 1/N.$$

✓ Pour $i = 1, \dots, N$,

▪ Simuler $\varepsilon^i \sim K$, noyau d'Epanechnikov.

▪ $X_k^i = X_k^i + h_{opt} D_k \varepsilon^i$ et $X_{0:k}^i = (X_{0:k-1}^i, X_k^i)$.

➤ Si $\hat{N}_{eff} > N_{seuil}$, $X_{0:k}^i = \tilde{X}_{0:k}^i$ et les poids d'importance $w_k^i = \tilde{w}_k^i$ restent inchangés.

- Sorties de l'algorithme : Le nuage de particules associées à leurs poids respectifs $S_k^N = (X_{0:k}^i, w_k^i)_{i=1, \dots, N}$ permet d'approcher la loi a posteriori

$$p(X_{0:k} | Y_{1:k}) \approx \sum_{i=1}^N w_k^i \delta_{X_{0:k}^i} (dX_{0:k})$$

Alg. II-7 Algorithme d'échantillonnage pondéré séquentiel régularisé (RPF).

Musso [11] propose de multiples améliorations à la version de base du RPF présenté, notamment lors de la phase de rééchantillonnage.

Le filtre particulaire régularisé possède de meilleures performances que le SIR lorsque la vraisemblance est très étroite (bruit de mesure et/ou d'état faibles), obtenues par l'amélioration de la diversité des particules lors de l'étape de rééchantillonnage au moyen d'une approximation continue de la densité de filtrage.

II.10.3. La Rao-Blackwellisation

Après s'être successivement intéressés au filtre à variable auxiliaire qui permet d'améliorer les performances du SIR par un meilleur guidage des particules vers les zones d'intérêt puis à la méthode d'amélioration de la diversité des états qu'est le RPF, nous examinons une méthode de réduction de variance des poids d'importance : la Rao-Blackwellisation, fondée sur la structure particulière de certains modèles.

Lorsqu'une partie du vecteur d'état, conditionnellement à d'autres composantes, peut être traitée par une méthode optimale telle que le filtre de Kalman, il est opportun de ne pas résoudre le problème de filtrage en adoptant une solution purement particulière. L'application de la méthode de Rao-Blackwellisation permet de tirer parti de cette propriété, et de réduire ainsi la variance du filtre [56, 45, 33].

Dans cette méthode, on suppose donc que le vecteur d'état X_k peut se diviser en deux parties $X1_k$ et $X2_k$ telle que

$$p(X1_k | X2_k, X1_{k-1}, X2_{k-1}) = p(X1_k | X1_{k-1}) \quad (2.59)$$

De sorte que l'équation de la dynamique du système s'écrit

$$p(X_k | X_{k-1}) = p(X1_k, X2_k | X1_{k-1}, X2_{k-1}) = p(X1_k | X1_{k-1}) p(X2_k | X1_{k-1}, X2_{k-1}) \quad (2.60)$$

Si on suppose également que la distribution a posteriori conditionnelle $p(X2_k | X1_k, Y_{1:k})$ peut être exprimée analytiquement, alors, en se basant sur la distribution objective $p(X_k | Y_{1:k})$ qui s'écrit

$$p(X_k | Y_{1:k}) = p(X1_k, X2_k | Y_{1:k}) = p(X1_k | Y_{1:k}) p(X2_k | X1_k, Y_{1:k}) \quad (2.61)$$

On peut simplement marginaliser sur $X2_k$ pour se focaliser sur l'estimation de $p(X1_k | Y_{1:k})$ dont l'espace est réduit.

Dans un premier temps, on construit une approximation particulière de $p(X1_k | Y_{1:k})$,

$$p(X1_k | Y_{1:k}) \approx \sum_{i=1}^N w_k^i \delta(X1_k - X1_k^i) \quad (2.62)$$

Ensuite, la densité marginale de $X2_k | Y_{1:k}$, qui s'écrit

$$p(X2_k | Y_{1:k}) = \int p(X2_k | X1_k, Y_{1:k}) p(X1_k | Y_{1:k}) dX1_k \quad (2.63)$$

est approchée, grâce à (2.62) par un mélange de lois calculables

$$\begin{aligned}
p(X_{2,k} | Y_{1:k}) &= \int p(X_{2,k} | X_{1,k}, Y_{1:k}) \sum_{i=1}^N w_k^i \delta(X_{1,k} - X_{1,k}^i) dX_{1,k} \\
&\approx \sum_{i=1}^N w_k^i p(X_{2,k} | X_{1,k}^i, Y_{1:k})
\end{aligned} \tag{2.64}$$

Des études théoriques dans [32] ont montré que la méthode de Rao-Blackwellisation apporte bien un gain sur l'efficacité du filtrage, en diminuant la variance par rapport à la variance obtenue par un estimateur standard non Rao-Blackwellisation.

III. Conclusion

Nous avons dressé dans ce chapitre un panorama des méthodes de Monte Carlo séquentielles les plus utilisées. Toutes découlent du principe d'échantillonnage pondéré. Dans le cas général, les méthodes particulières donnent une bonne approximation du filtre optimal. Elles permettent de construire récursivement un nuage de particules pondérées approchant la loi de filtrage. Cependant, il existe une limitation à l'utilisation du filtre particulaire, car sa nature récursive conduit à une dégénérescence du nuage induit par l'augmentation dans le temps de la variance inconditionnelle de ses poids et aussi dans le cas de bruits de dynamique faible et / ou le bruit de mesure faible, le filtre particulaire classique peut diverger. Nous détaillons ensuite la fonction d'importance dite optimale qui permet de positionner les particules en tenant compte de la dynamique du système et de l'observation à l'instant courant. On montre que pour une telle fonction, le phénomène de dégénérescence est compensé mais en pratique, cette fonction n'est généralement pas utilisable. D'autres améliorations sont alors envisagées. L'introduction d'un rééchantillonnage dans l'algorithme, permet de limiter la dégénérescence des poids d'importance en éliminant les particules de poids négligeables et en multipliant les particules les plus significatives. Le filtre résultant constitue l'algorithme générique de filtrage SIR à partir duquel toutes les stratégies de filtrage particulaire dérivent. D'autres méthodes telle que l'échantillonnage pondéré à variable auxiliaire utilise des poids auxiliaires pour sélectionner les particules les vraisemblables avant leur propagation tandis que le filtre particulaire régularisé possède de meilleures performances que le SIR lorsque la vraisemblance est très étroite (bruit de mesure et/ou d'état faibles), obtenue par l'amélioration de la diversité des particules lors de l'étape de rééchantillonnage au moyen d'une approximation continue de la densité de filtrage.

Nous n'avons pas non plus considéré les algorithmes de poursuite multi cible. Dans ce contexte, [57, 58] représentent la densité a posteriori par un mélange de gaussiennes afin de maintenir la multimodalité.

Introduction

Pour modéliser la trajectoire d'une cible, nous pouvons affirmer que quelque soit la complexité de cette dernière, elle pourra toujours être représentée par une succession de trajectoires élémentaires canoniques tels que des segments de droites, des arcs de cercles, des arcs elliptiques, A partir de la, nous pouvons dire que pour déterminer la trajectoire suivie par une cible donnée durant un temps t , il suffit de déterminer les paramètres des différentes trajectoires canoniques suivies par la cible durant ce temps.

Dans notre cas, nous allons restreindre les trajectoires canoniques choisies au segment de droites et à l'arc de cercle.

I. Modèle d'une droite (Trajectoire rectiligne) : [59]

Nous modéliserons une droite de la manière suivante :

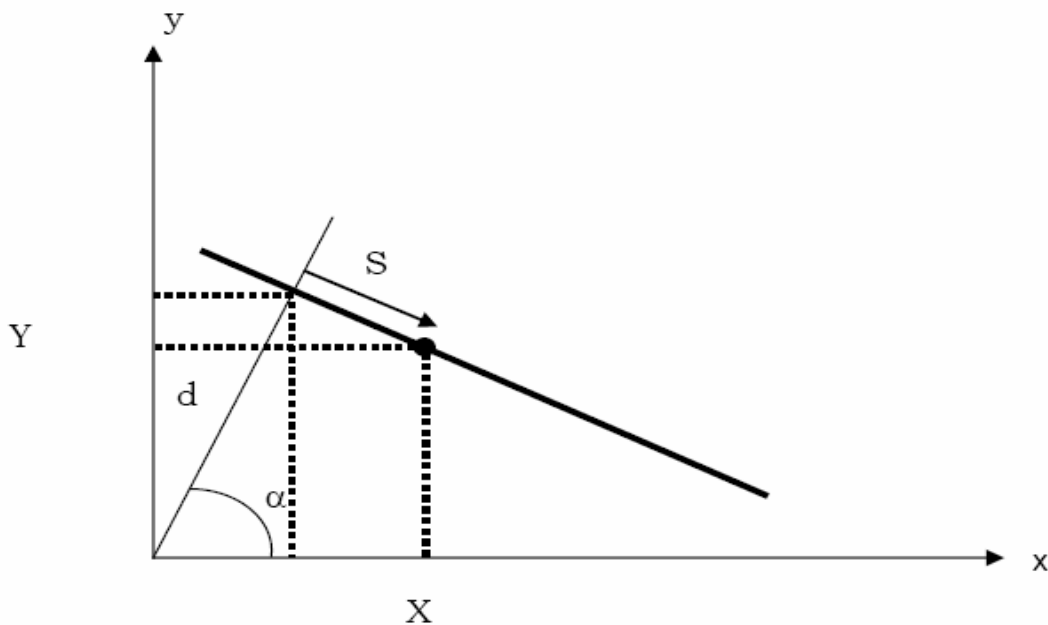


Figure III-1 Modélisation d'une droite

La représentation ci-dessus permet de définir toute droite (mouvement rectiligne) par ses coordonnées plucklériennes d et α , où

- d est la distance perpendiculaire à la droite passant par l'origine des coordonnées.
- α est l'angle que fait l'axe des abscisses avec la droite reliant l'origine à sa projection orthogonale sur la droite considérée.

D'autre part, on représente le mouvement sur une droite quelconque par les grandeurs suivantes :

- La distance parcourue s .
- La vitesse de déplacement $v = \dot{s}$.
- L'accélération $a = \ddot{s}$.

Cette représentation permet un découplage entre le mouvement et la trajectoire suivie.

I.1. Mouvement rectiligne à vitesse constante :

L'état de base d'un mouvement rectiligne à vitesse constante est :

$$\begin{cases} s = (t_k - t_{k-1})v + s_0 \\ v = \frac{ds}{dt} = \dot{s} \end{cases} \quad (3.1)$$

Où :

$(t_k - t_{k-1})$ est la différence de temps.

En temps discret, ces équations deviennent :

$$\begin{cases} s(k) = (t_k - t_{k-1})v(k-1) + s(k-1) \\ v(k) = v(k-1) \end{cases} \quad (3.2)$$

Posons $\begin{cases} x_1(k) = s(k) \\ x_2(k) = v(k) \end{cases}$,

on obtient

$$x(k) = \begin{cases} x_1(k) = x_1(k-1) + Tx_2(k-1) \\ x_2(k) = x_2(k-1) \end{cases} \quad (3.3)$$

Avec $T = t_k - t_{k-1}$

L'équation du vecteur des paramètres est $\theta(k) = \begin{cases} \alpha \\ d \end{cases} \quad (3.4)$

Posons $\begin{cases} x_3(k) = \alpha \\ x_4(k) = d \end{cases}$,

on obtient $\theta(k) = \begin{cases} x_3(k) = x_3(k-1) \\ x_4(k) = x_4(k-1) \end{cases} \quad (3.5)$

Parce qu'ils sont invariants dans le temps.

Donc l'état augmenté sera :

$$x(k) = \begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \quad (3.6)$$

$$x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} = \begin{bmatrix} x_1(k-1) + Tx_2(k-1) \\ x_2(k-1) \\ x_3(k-1) \\ x_4(k-1) \end{bmatrix} + v(k-1) \quad (3.7)$$

Avec $v(k)$: bruit d'état

$$x(k) = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + v(k-1) \quad (3.8)$$

$x(k)$ est Sous la forme $x(k+1) = Fx(k) + v(k)$,

Où F est la matrice de transition d'état :

$$F \stackrel{\Delta}{=} \exp(AT), \text{ avec } \dot{x}(t) = Ax(t) + \tilde{v}(t)$$

La fonction de mesure est une fonction non linéaire reliant l'état de base avec les paramètres inconnus du mouvement rectiligne (droite) selon les deux équations suivantes :

$$\begin{cases} X = d \cos(\alpha) + s \sin(\alpha) \\ Y = d \sin(\alpha) - s \cos(\alpha) \end{cases} \quad (3.9)$$

Qui sont la position X et Y de l'objet poursuivi.

Ces équations peuvent être écrites en fonction de variables d'état comme suit :

$$y(k) = \begin{cases} X(k) = x_4(k) \cos(x_3(k)) + x_1(k) \sin(x_3(k)) \\ Y(k) = x_4(k) \sin(x_3(k)) - x_1(k) \cos(x_3(k)) \end{cases} \quad (3.10)$$

$y(k)$ a la forme suivante $y(k) = H(x(k)) + w(k)$

Avec $w(k)$: bruit de mesure.

La linéarisation de $h(x)$ est réalisé en considérant une petite variation dans l'état :

$$H_x = \begin{bmatrix} \frac{\partial X}{\partial S} & \frac{\partial X}{\partial \dot{S}} & \frac{\partial X}{\partial \alpha} & \frac{\partial X}{\partial d} \\ \frac{\partial Y}{\partial S} & \frac{\partial Y}{\partial \dot{S}} & \frac{\partial Y}{\partial \alpha} & \frac{\partial Y}{\partial d} \\ \frac{\partial S}{\partial S} & \frac{\partial S}{\partial \dot{S}} & \frac{\partial S}{\partial \alpha} & \frac{\partial S}{\partial d} \end{bmatrix}$$

Dans ce cas on aura

$$H_x = \begin{bmatrix} \sin(\alpha) & 0 & -d \cos(\alpha) + S \cos(\alpha) & \cos(\alpha) \\ -\cos(\alpha) & 0 & d \cos(\alpha) - S \sin(\alpha) & \sin(\alpha) \end{bmatrix}$$

I.2. Mouvement rectiligne à accélération constante :

L'état de base d'un mouvement rectiligne à accélération constante est :

$$\begin{cases} s(k) = s(k-1) + Tv(k-1) + \frac{1}{2}T^2 a(k-1) \\ v(k) = v(k-1) + Ta(k-1) \\ a(k) = a(k-1) \end{cases} \quad (3.11)$$

L'équation du vecteur des paramètres reste la même que le mouvement rectiligne à vitesse constante.

Posons

$$\begin{cases} x_1(k) = s(k) \\ x_2(k) = v(k) \\ x_3(k) = a(k) \\ x_4(k) = \alpha \\ x_5(k) = d \end{cases} \quad (3.12)$$

L'état augmenté est donc :

$$x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \end{bmatrix} = \begin{bmatrix} x_1(k-1) + Tx_2(k-1) + 0.5T^2 x_3(k-1) \\ x_2(k-1) + Tx_3(k-1) \\ x_3(k-1) \\ x_4(k-1) \\ x_5(k-1) \end{bmatrix} + v(k-1) \quad (3.13)$$

$$x(k) = \begin{bmatrix} 1 & T & 0.5T^2 & 0 & 0 \\ 0 & 1 & T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + v(k-1) \quad (3.14)$$

$x(k)$ est sous la forme $x(k) = Fx(k-1) + v(k-1)$.

Par conséquent :

$$F = \begin{bmatrix} 1 & T & 0.5T^2 & 0 & 0 \\ 0 & 1 & T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

La fonction de mesure reste la même que le mouvement précédent et la seule différence se trouve dans l'ordre des variables d'état.

$$y(k) = \begin{cases} X(k) = x_5(k)\cos(x_4(k)) + x_1(k)\sin(x_4(k)) \\ Y(k) = x_5(k)\sin(x_4(k)) - x_1(k)\cos(x_4(k)) \end{cases} \quad (3.15)$$

Donc $y(k)$ est de la forme $y(k) = H(x(k)) + w(k)$

Avec $w(k)$: bruit de mesure.

La matrice de mesure linéarisée H_x se déduit par dérivation de $h(x)$

$$H_x = \begin{bmatrix} \sin(\alpha) & 0 & 0 & -d \sin(\alpha) + S \cos(\alpha) & \cos(\alpha) \\ -\cos(\alpha) & 0 & 0 & d \cos(\alpha) - S \sin(\alpha) & \sin(\alpha) \end{bmatrix}$$

II. Modèle d'un cercle (Trajectoire circulaire) : [59]

Comme pour le cas d'un mouvement rectiligne, les mouvements circulaires peuvent être représentés en fonction des coordonnées pulcklériennes R_0, X_0 et Y_0 .

Où : R_0 est le rayon du cercle.

(X_0, Y_0) est le centre de ce cercle.

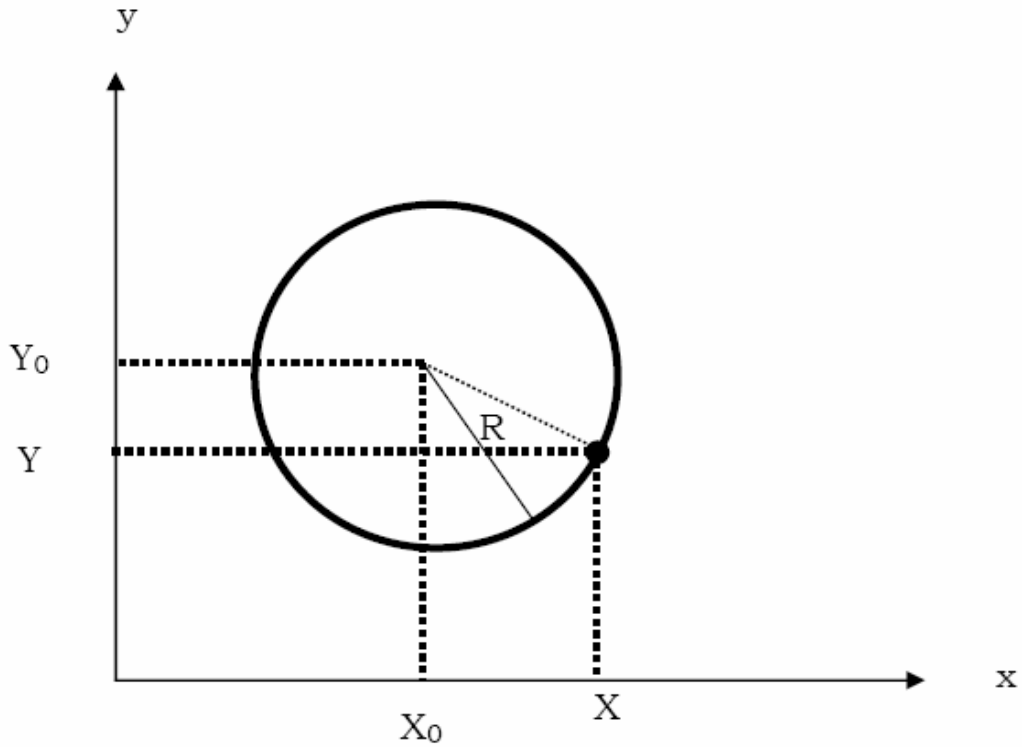


Figure III-2 Modélisation d'un cercle

II.1. Mouvement circulaire à vitesse constante :

L'état de base d'un mouvement circulaire à vitesse constante est identique a celui du mouvement rectiligne à vitesse constante, c'est-à-dire :

$$x(k) = \begin{cases} x_1(k) = x_1(k-1) + Tx_2(k-1) \\ x_2(k) = x_2(k-1) \end{cases} \tag{3.16}$$

L'équation de vecteur des paramètres est

$$\theta(k) = \begin{bmatrix} R_0 \\ X_0 \\ Y_0 \end{bmatrix} \tag{3.17}$$

Posons

$$\begin{cases} x_3(k) = R_0 \\ x_4(k) = X_0 \\ x_5(k) = Y_0 \end{cases} \tag{3.18}$$

L'état augmenté sera :

$$x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \end{bmatrix} = \begin{bmatrix} x_1(k-1) + Tx_2(k-1) \\ x_2(k-1) \\ x_3(k-1) \\ x_4(k-1) \\ x_5(k-1) \end{bmatrix} + v(k-1) \quad (3.19)$$

$$x(k) = \begin{bmatrix} 1 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + v(k-1) \quad (3.20)$$

$x(k)$ a la forme suivante $x(k) = Fx(k-1) + v(k-1)$

Avec $v(k-1)$: bruit d'état.

Et la matrice de transition F sera :

$$F = \begin{bmatrix} 1 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Les équations de mesure qui décrivent la position de l'objet poursuivi sont :

$$\begin{cases} X = X_0 + R_0 \cos\left(\frac{s}{R_0}\right) \\ Y = Y_0 + R_0 \sin\left(\frac{s}{R_0}\right) \end{cases} \quad (3.21)$$

Ces équations peuvent être écrites en fonction des variable d'état comme suit :

$$y(k) = \begin{cases} X(k) = x_4(k) + x_3(k) \cos\left(\frac{x_1(k)}{x_3(k)}\right) \\ Y(k) = x_5(k) + x_3(k) \sin\left(\frac{x_1(k)}{x_3(k)}\right) \end{cases} \quad (3.22)$$

$y(k)$ a la forme suivante $y(k) = H(x(k)) + w(k)$

Avec $w(k)$: bruit de mesure.

La matrice de mesure linéarisée se déduit par dérivation de $h(x)$

$$H_x = \begin{bmatrix} \frac{\partial X}{\partial S} & \frac{\partial X}{\partial \dot{S}} & \frac{\partial X}{\partial R} & \frac{\partial X}{\partial X_0} & \frac{\partial X}{\partial Y_0} \\ \frac{\partial Y}{\partial S} & \frac{\partial Y}{\partial \dot{S}} & \frac{\partial Y}{\partial R} & \frac{\partial Y}{\partial X_0} & \frac{\partial Y}{\partial Y_0} \end{bmatrix}$$

Nous aurons alors,

$$H_x = \begin{bmatrix} -\sin\left(\frac{S}{R}\right) & 0 & \cos\left(\frac{S}{R}\right) + \frac{S}{R} \sin\left(\frac{S}{R}\right) & 1 & 0 \\ \cos\left(\frac{S}{R}\right) & 0 & \sin\left(\frac{S}{R}\right) - \frac{S}{R} \cos\left(\frac{S}{R}\right) & 0 & 1 \end{bmatrix}$$

II.2. Mouvement circulaire à accélération constante :

L'état de base d'un mouvement circulaire à accélération constante est identique de celui du mouvement rectiligne à accélération constante.

$$x(k) = \begin{cases} s(k) = s(k-1) + T v(k-1) + \frac{1}{2} T^2 a(k-1) \\ v(k) = v(k-1) + T a(k-1) \\ a(k) = a(k-1) \end{cases} \quad (3.23)$$

Le vecteur des paramètres est le même que celui du mouvement circulaire à vitesse constante.

$$\theta(k) = \begin{bmatrix} R_0 \\ X_0 \\ Y_0 \end{bmatrix},$$

et posons

$$\begin{cases} x_1(k) = s(k) \\ x_2(k) = v(k) \\ x_3(k) = a(k) \\ x_4(k) = R_0 \\ x_5(k) = X_0 \\ x_6(k) = Y_0 \end{cases} \quad (3.24)$$

L'état augmenté sera

$$x(k) = \begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \end{bmatrix}$$

$$x(k) = \begin{bmatrix} 1 & T & 0.5T^2 & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + v(k-1) \quad (3.25)$$

$x(k)$ a la forme suivante $x(k) = Fx(k-1) + v(k-1)$

Les équations de mesures sont les même que celles du mouvement circulaire à vitesse constante et la seule différence se trouve dans l'ordre des variables d'état.

$$y(k) = \begin{cases} X(k) = x_5(k) + x_4(k) \cos\left(\frac{x_1(k)}{x_4(k)}\right) \\ Y(k) = x_6(k) + x_4(k) \cos\left(\frac{x_1(k)}{x_4(k)}\right) \end{cases} \quad (3.26)$$

$y(k)$ est de la forme suivant $y(k) = H(x(k)) + w(k)$

Avec $w(k)$: bruit de mesure.

Par conséquent, la matrice H_x s'écrit de la façon suivante

$$H_k = \begin{bmatrix} -\sin\left(\frac{S}{R}\right) & 0 & 0 & \cos\left(\frac{S}{R}\right) + \frac{S}{R}\sin\left(\frac{S}{R}\right) & 1 & 0 \\ \cos\left(\frac{S}{R}\right) & 0 & 0 & \sin\left(\frac{S}{R}\right) - \frac{S}{R}\cos\left(\frac{S}{R}\right) & 0 & 1 \end{bmatrix}$$

$\nu(k)$ et $w(k)$ Représentent les bruits du système et d'observation respectivement. Les variances des composantes du vecteur d'état qui représentent le mouvement sont prises de manière à satisfaire les conditions sur l'accélération maximale pour le mouvement à vitesse constante et l'incrément sur l'accélération maximale pour le mouvement à accélération constante.

III. Conclusion

La génération des trajectoires de la cible est basée sur les quatre modèles d'état cité ci-dessus, qui sont : mouvement rectiligne à vitesse constante, mouvement rectiligne à accélération constante, mouvement circulaire à vitesse constante et mouvement circulaire à accélération constante.

Introduction

Pour illustrer les algorithmes qui viennent d'être présentés, nous allons étudier le problème classique d'estimation de la position d'une cible par mesure de distance et d'angle. Le problème de la poursuite de cibles consiste à estimer les paramètres cinématiques d'un mobile (position, vitesse, accélération,...) à partir de mesures bruitées fonction de ces paramètres fournies par un capteur (radar, sonar, capteur optique,...). On possède en général un modèle dynamique décrivant l'évolution de l'état de la cible dans le temps découlant de la connaissance a priori que l'on a du système. Les modèles de la dynamique choisie sont souvent linéaires,

$$X_k = F_k X_{k-1} + W_k \quad (4.1)$$

En revanche, l'équation d'observation liant la mesure à l'état est très souvent non linéaire et dépend du capteur utilisé (mesure de distance, d'angle ou de doppler). La mesure de l'état X_k fournie par le capteur est modélisée par une équation du type,

$$Y_k = h_k(X_k) + V_k \quad (4.2)$$

W_k et V_k sont des bruits blancs gaussiens. On retrouve exactement l'énoncé d'un problème de filtrage classique.

Dans ce chapitre, nous nous intéressons aux comportements simulés des méthodes particulières sur des problèmes concrets de poursuite. Nous avons fait le choix d'applications réalistes. Ce chapitre présente néanmoins un intérêt pratique certain en montrant l'apport des méthodes particulières face aux méthodes classiques telles que l'EKF qui sont parfois mises en défaut et surtout en proposant des solutions simples à des problèmes difficiles de poursuite. Nous comparons les performances de PF et l'EKF sur le problème de la poursuite par mesure de distance et d'angle et nous verrons que la solution particulière de ce problème est simple et efficace comparée à l'algorithme classiquement utilisée pour résoudre ce problème.

Le cas tridimensionnel ne présentant aucune difficulté supplémentaire, nous nous sommes bornés par la suite à l'étude de modèle bidimensionnel.

I. Modèle d'état : [60]

On considère le problème de la poursuite d'un mobile se déplaçant dans un plan muni d'un repère cartésien de référence. L'évolution du mobile est ainsi décrite par l'évolution de ses coordonnées $(X(t), Y(t))$. Si on note \vec{F}_t la somme des forces agissant à l'instant t sur le mobile de masse m , alors on obtient, par les lois de la mécanique, une relation liant l'accélération du mobile à la force \vec{F}_t ,

$$m \begin{bmatrix} \ddot{X}(t) \\ \ddot{Y}(t) \end{bmatrix} = \vec{F}_t \quad (4.3)$$

La force \vec{F}_t se réduit aux frottements de l'air ou à d'autres perturbations extérieures lorsque le mobile n'est pas en phase de manœuvre i.e. lorsque qu'il n'accélère pas. Dans ce cas, on dit que le mobile est en mouvement rectiligne uniforme bruité. Lorsque le mobile opère une accélération volontaire (produite par la poussée des moteurs), la force \vec{F}_t résulte en plus de cette accélération. Dans ce cas on dit que le mobile est en phase de manœuvre. On fait l'hypothèse qu'entre deux instants de discrétisation t_k et t_{k+1} , l'accélération \vec{F}_t/m est supposée constante et égale à $[w_k^1, w_k^2]$ où elles sont indépendantes entre deux intervalles de discrétisation. Avec cette hypothèse, en intégrant par deux fois l'équation (4.3), entre t_k et t_{k+1} , on obtient

$$\begin{bmatrix} X(t_{k+1}) \\ \dot{X}(t_{k+1}) \\ Y(t_{k+1}) \\ \dot{Y}(t_{k+1}) \end{bmatrix} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X(t_k) \\ \dot{X}(t_k) \\ Y(t_k) \\ \dot{Y}(t_k) \end{bmatrix} + \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix} \quad (4.4)$$

Où $T = t_{k+1} - t_k$ est la période de discrétisation du modèle. Pour tout entier $k \geq 0$, on note

$$X_k = \begin{bmatrix} X1_k \\ X2_k \\ X3_k \\ X4_k \end{bmatrix} = \begin{bmatrix} X(t_k) \\ \dot{X}(t_k) \\ Y(t_k) \\ \dot{Y}(t_k) \end{bmatrix}$$

X_k constitue l'état du mobile à l'instant t_k , i.e. sa position et sa vitesse suivant chacune des coordonnées. L'équation (4.4) décrivant l'évolution de X_k dans le temps est appelée équation d'état. Le but de la poursuite est d'estimer la valeur de X_k , Il s'agit donc ici d'un problème de dimension 4.

I.1. Poursuite radar :

Nous rappelons brièvement le principe de mesure d'un radar de poursuite décrit en détails dans [61]. Le radar envoie un signal très bref ou impulsion dans une direction donnée. Ce signal se propage à la vitesse de la lumière, dans toutes les directions : il forme une onde sphérique. Au bout d'un intervalle de temps ΔT , le signal atteint la cible, ce temps est proportionnel à la distance radar-cible. Une partie du signal est réfléchi par la cible et au bout du même intervalle de temps ΔT , l'onde réfléchi atteint après atténuation et effet doppler à nouveau le radar. Celui-ci, après avoir traité le signal reçu (amplification, filtrage, détection, corrélation,...), fournit une estimation de la distance entre le radar et la cible et de l'angle formé par l'axe de visée du radar et un axe de référence. L'utilisation d'un tel radar de poursuite conduit au problème de pistage par mesure de distance et d'angle.

Dans le cas général, la matrice de covariance du bruit de mesure est fonction du rapport signal / bruit qui lui-même fonction de la distance radar-cible. Afin de faciliter les calculs, dans les applications considérées, on peut négliger cette contrainte. La matrice de covariance du bruit de mesure est considérée fixe.

II. Simulation :

Dans les simulations effectuées, le mobile suit un mouvement rectiligne uniforme. Le radar fournit des mesures bruitées de la distance entre l'observation et le mobile, ainsi que l'angle formé par l'axe vertical et la ligne de visée, i.e.,

$$h(X_k) = \begin{pmatrix} \sqrt{X1_k^2 + X3_k^2} \\ \arctan(X3_k / X1_k) \end{pmatrix} \quad (4.5)$$

Où $X = (X1, X2, X3, X4)$ désigne dans l'ordre, la position et la vitesse en abscisse et la position et la vitesse en ordonnée du mobile dont l'état est X .

Nous considérons le problème suivant de pistage par mesure de distance et d'angle.

- ✓ L'état initial est $X_0 = (5km, -20m/s, 5km, 20m/s)$ avec comme matrice de covariance $P_0[diag(2km, 20m/s, 2km, 20m/s)]^2$. La dynamique de la cible est bruitée selon l'équation (1.27) avec un bruit sur l'accélération $w^1 = w^2 = c$ variable.
- ✓ Le temps de poursuite varie avec une période de mesure égale à un.
- ✓ L'écart type des bruits de mesure sur la distance et l'angle sont respectivement de 50m et 1 degré.

Le nombre de particules utilisées par le PF varie entre $N = 100$ et $N = 1500$ particules.

Nous soulignons que les conditions garantissant le bon comportement de l'EKF ne sont pas vérifiées même si les mesures sont très précises.

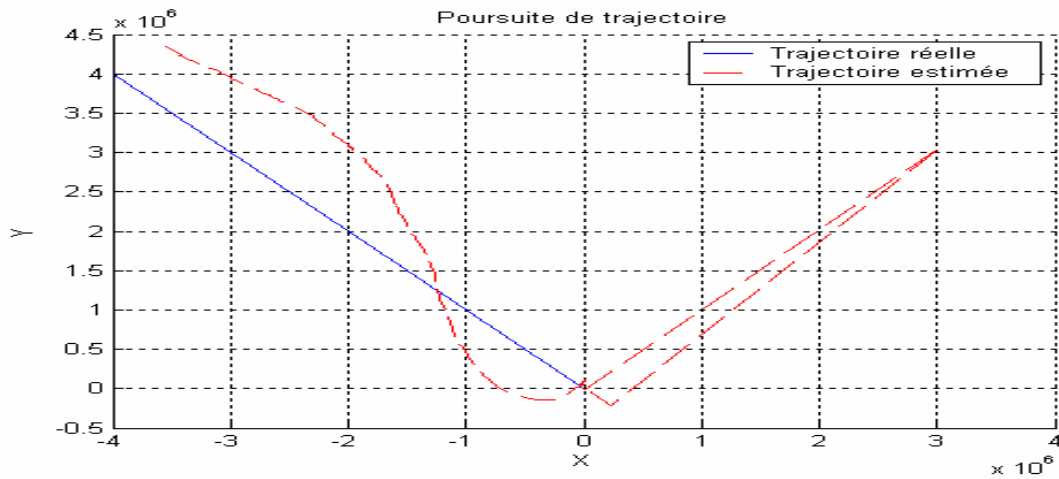


Figure IV-1 La trajectoire vraie, estimée de la cible par l'EKF

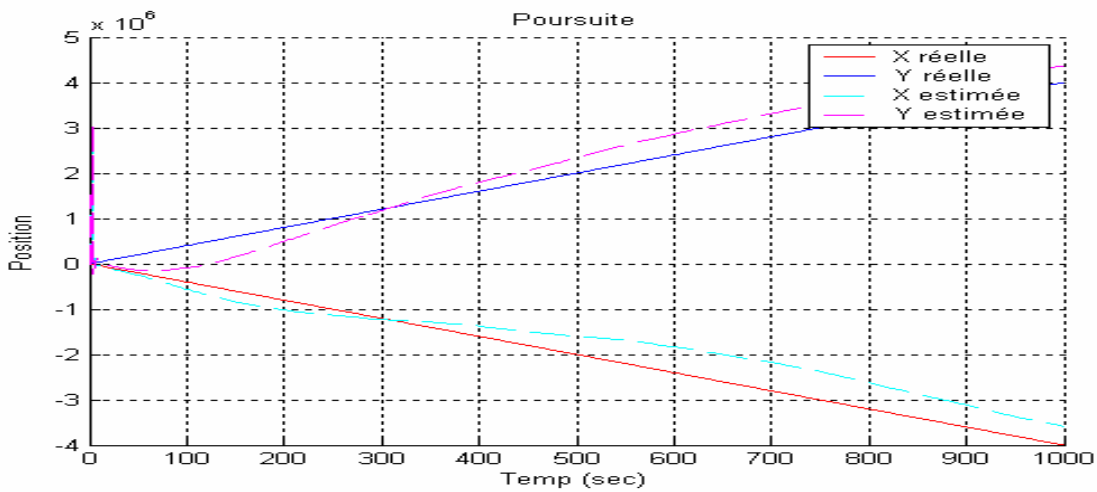


Figure IV-2 La position vraie, estimée suivant l'axe X et l'axe Y par l'EKF

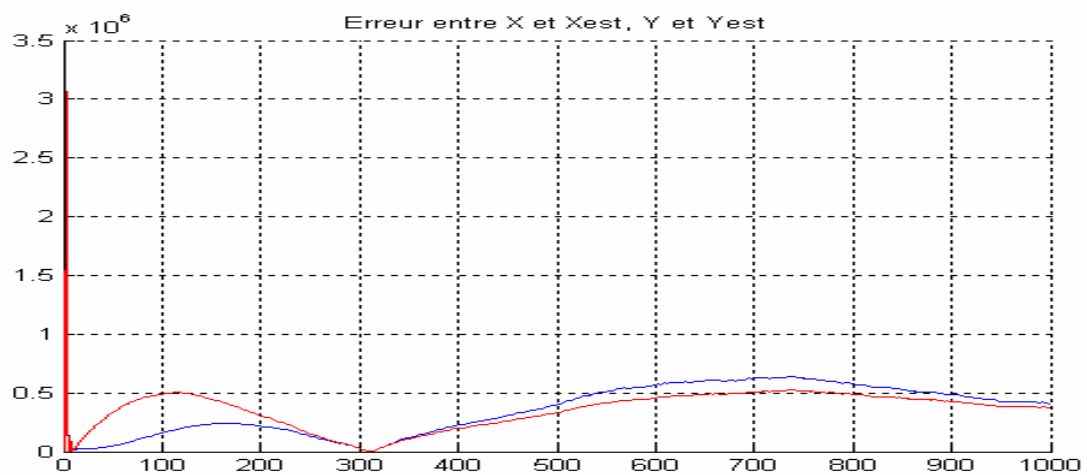


Figure IV-3 L'erreur entre X et Xest, Y et Yest.

Comme on peut le constater sur la figure IV-1, le filtre de Kalman étendu met un certain temps au début pour qu'il puisse suivre la trajectoire, tout ça à cause des conditions de départ mal initialisées. A partir de $t = 400s$ le filtre commence à diverger à cause de la forte non linéarité. Le bruit de dynamique joue aussi un rôle dans la divergence de filtre. On remarque que le filtre de Kalman présente des divergences alors que le modèle non linéaire pris est simple.

On remarque d'après les deux figures suivantes (IV-4, IV-5) que le filtre particulière adaptatif suit presque parfaitement la trajectoire de la cible donc donne de bons résultats, même avec un nombre d'échantillons petit et un seuil $N_{Seuil} = 0.5$, comparé aux résultats de filtre de Kalman ci-dessus.

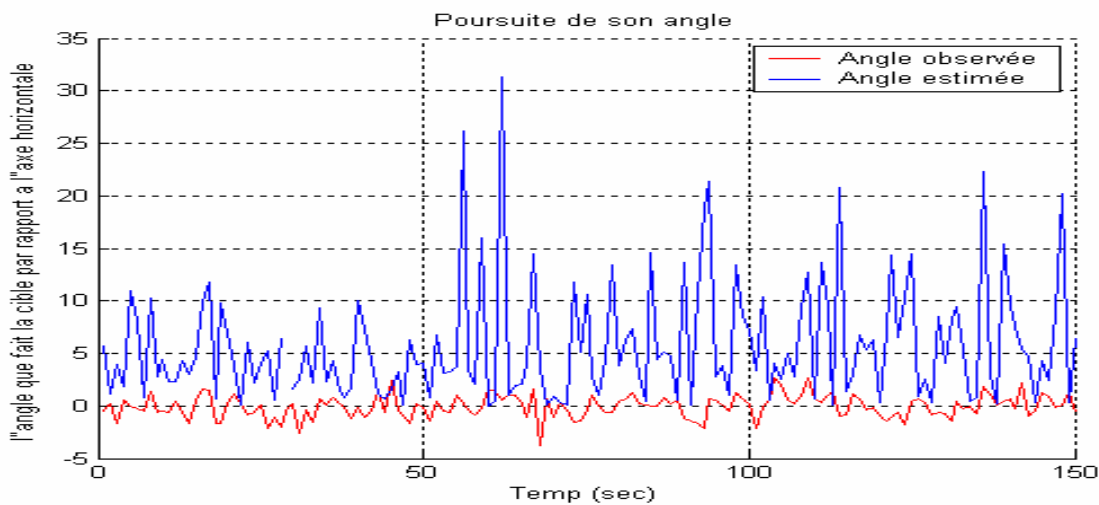


Figure IV-4 L'angle vrai, estimée par SIR adaptatif avec $N = 150$ et $N_{Seuil} = 0.5 N$

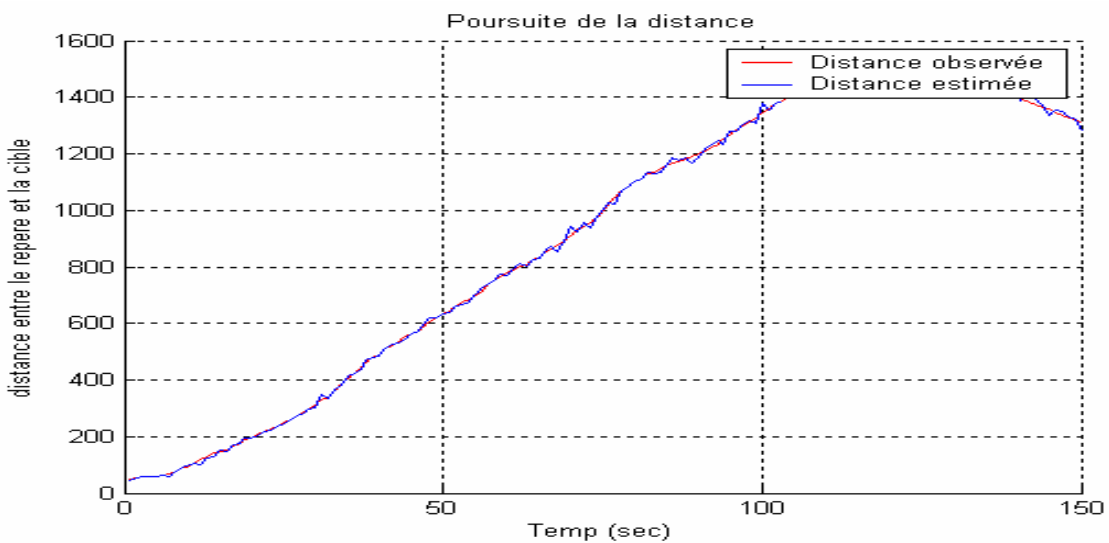


Figure IV-5 La distance vraie, estimée par SIR adaptatif avec $N = 150$ et $N_{Seuil} = 0.5 N$

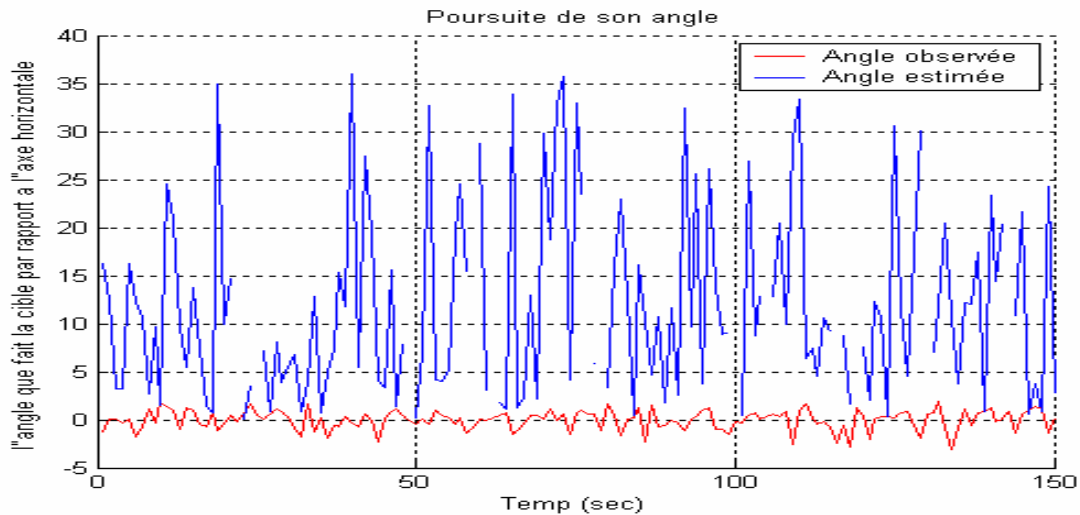


Figure IV-6 L'angle vrai, estimé par SIR systématique avec $N = 150$

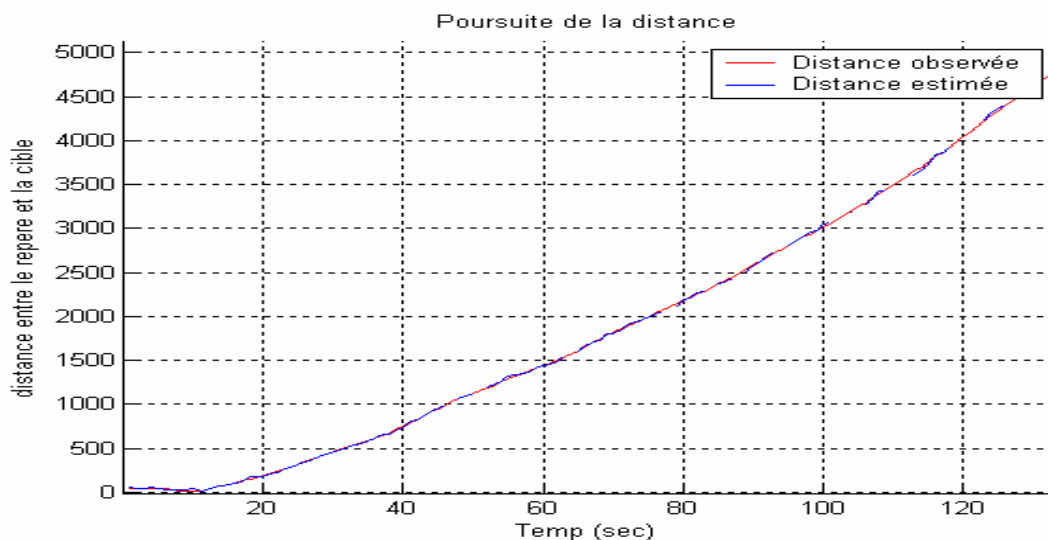


Figure IV-7 La distance vraie, estimés par SIR systématique avec $N = 150$

Comme on peut le constater sur les figures (IV-6, IV-7) de filtre SIR systématique, le nuage de particules est appauvri comparé aux figures (IV-4, IV-5) de filtre SIR adaptatif. Donc l'estimée donnée par le SIR systématique est moins fiable. Le temps de calcul pour le SIR systématique est plus important que celui de SIR adaptatif. On peut conclure que le SIR adaptatif donne de bons résultats comparés au SIR systématique.

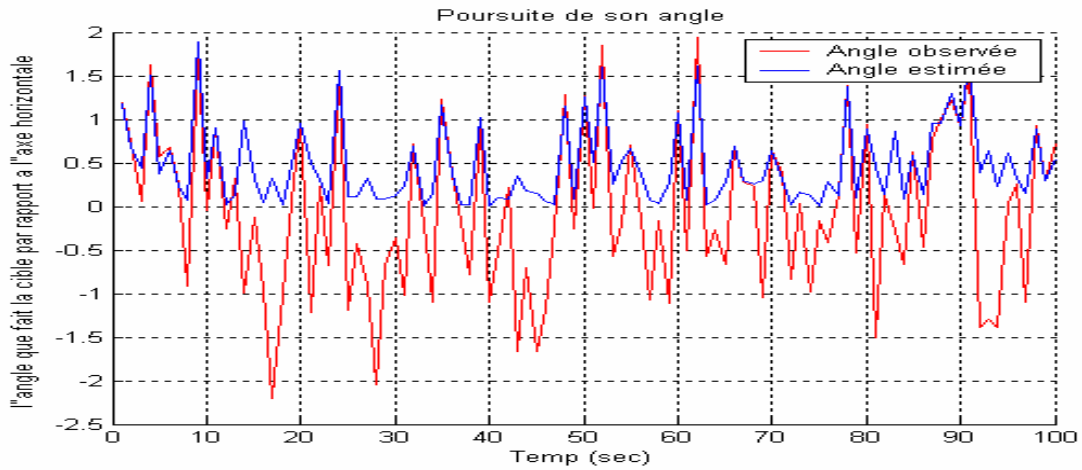


Figure IV-8 L'angle vrai, estimé par SIR adaptatif avec $N = 500$ et $N_{Seuil} = 0.5 N$

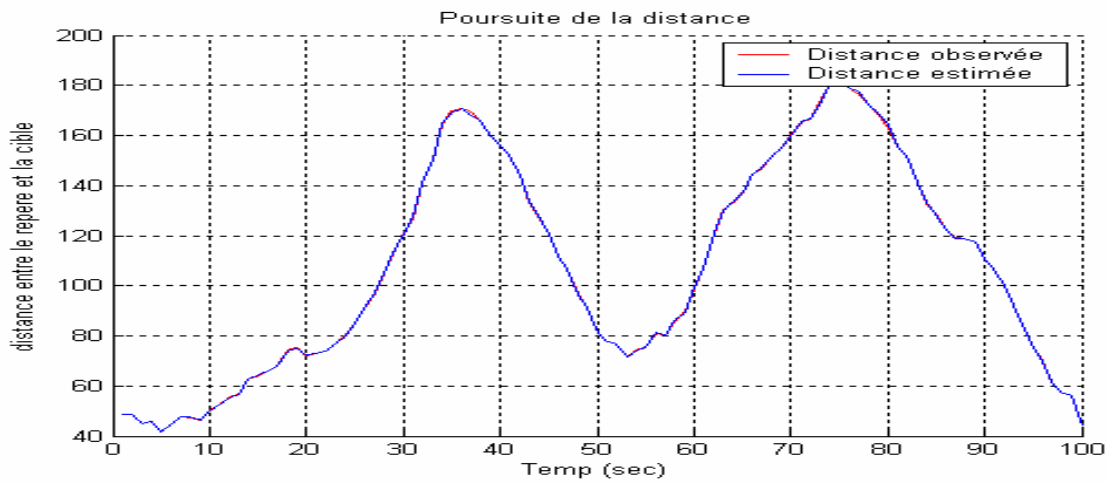


Figure IV-9 La distance vraie, estimée par SIR adaptatif avec $N = 500$ et $N_{Seuil} = 0.5$

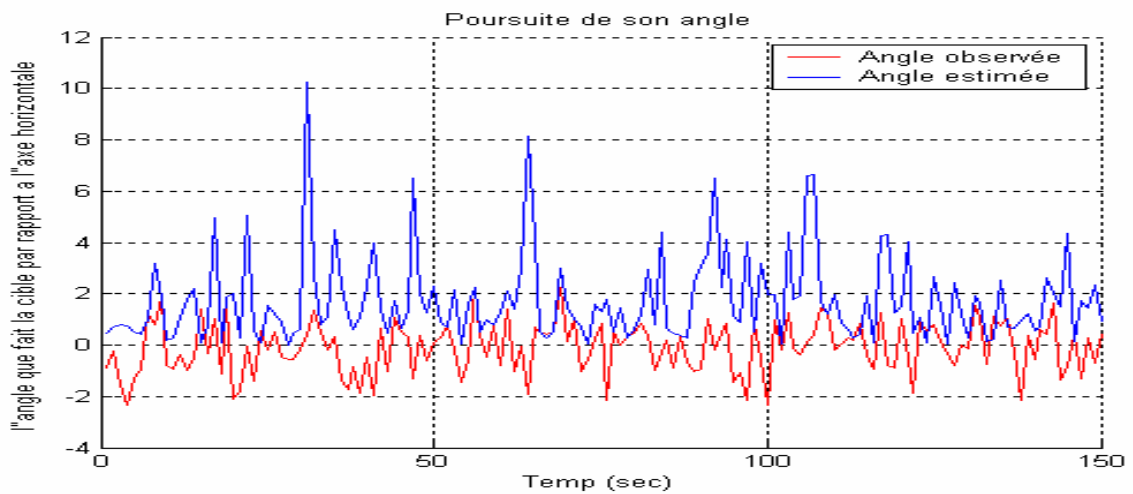


Figure IV-10 L'angle vrai, estimé par SIR systématique avec $N = 500$

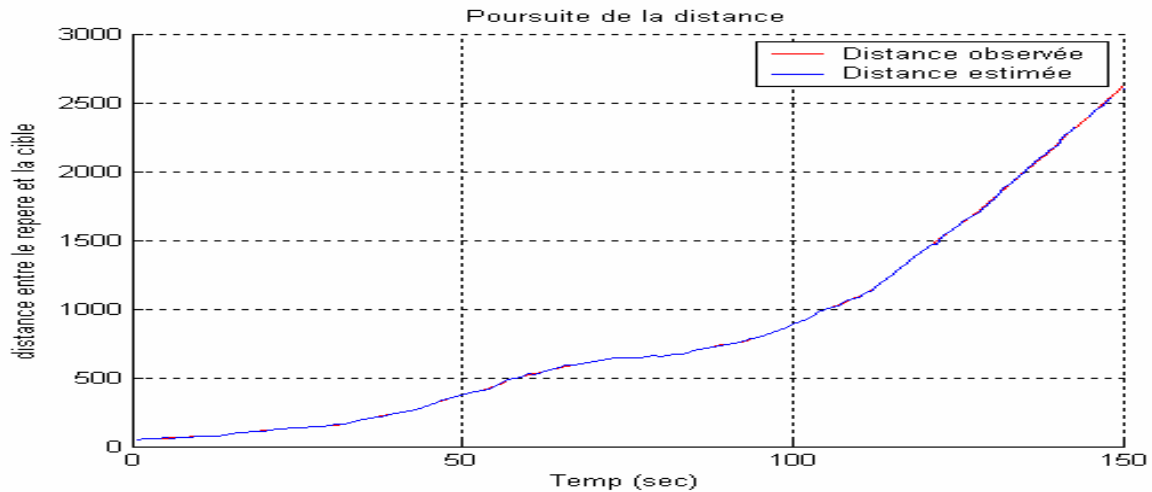


Figure IV-11 La distance vraie, estimée par SIR systématique avec $N = 500$

Les remarques faites pour les figures (IV-4, IV-5, IV-6, IV-7) s'appliquent sur les figures (IV-8, IV-9, IV-10, IV-11). La différence est que le nombre de particules augmente mais les résultats restent inchangés. Le résultat pour le SIR systématique est meilleur avec un nombre de particules élevé.

Les figures suivantes montrent l'influence de la valeur du seuil sur les résultats des SIR adaptatifs.

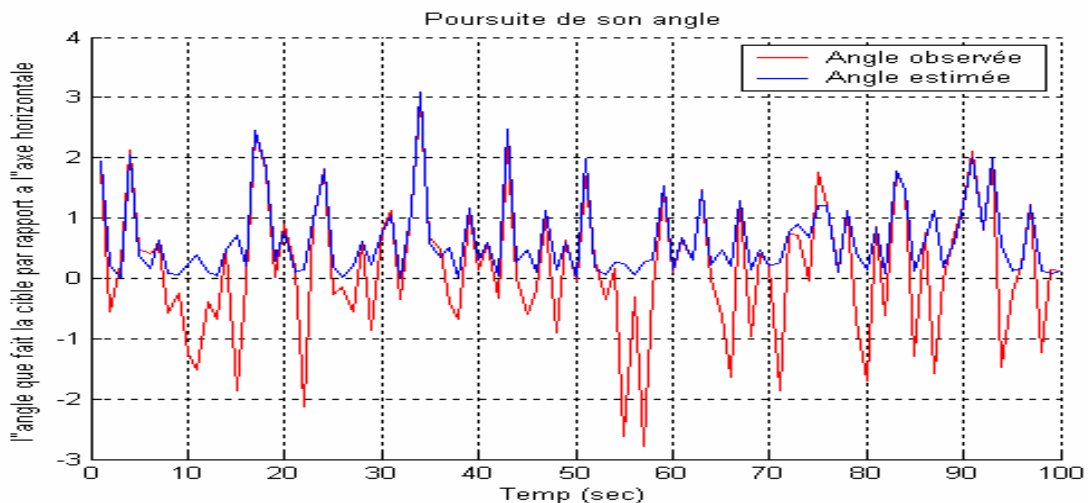


Figure IV-12 L'angle vrai, estimé par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.5$

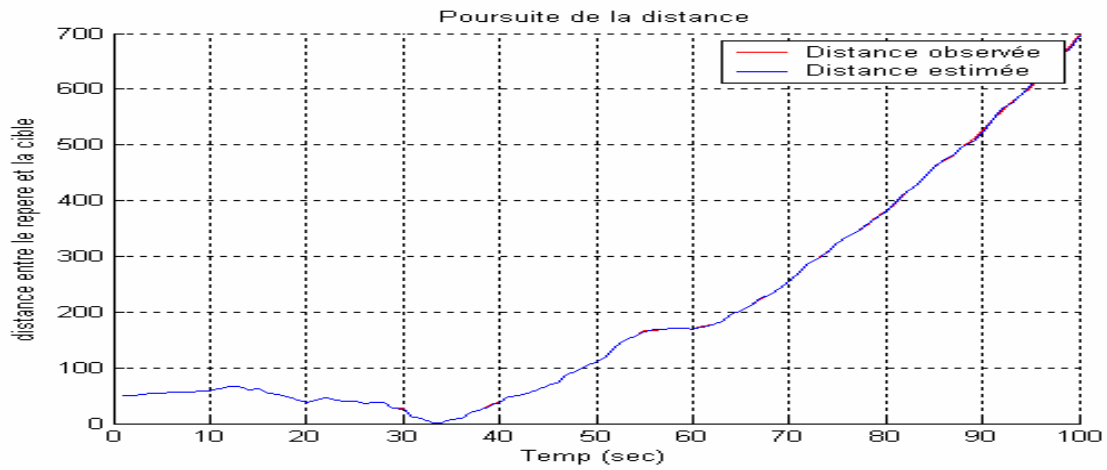


Figure IV-13 La distance vraie, estimée par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.5$

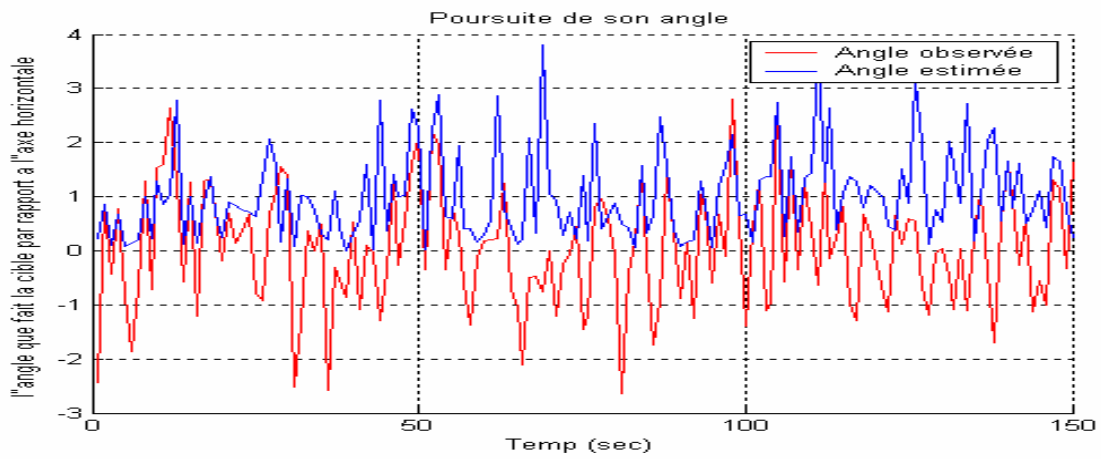


Figure IV-14 L'angle vrai, estimée par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.8$

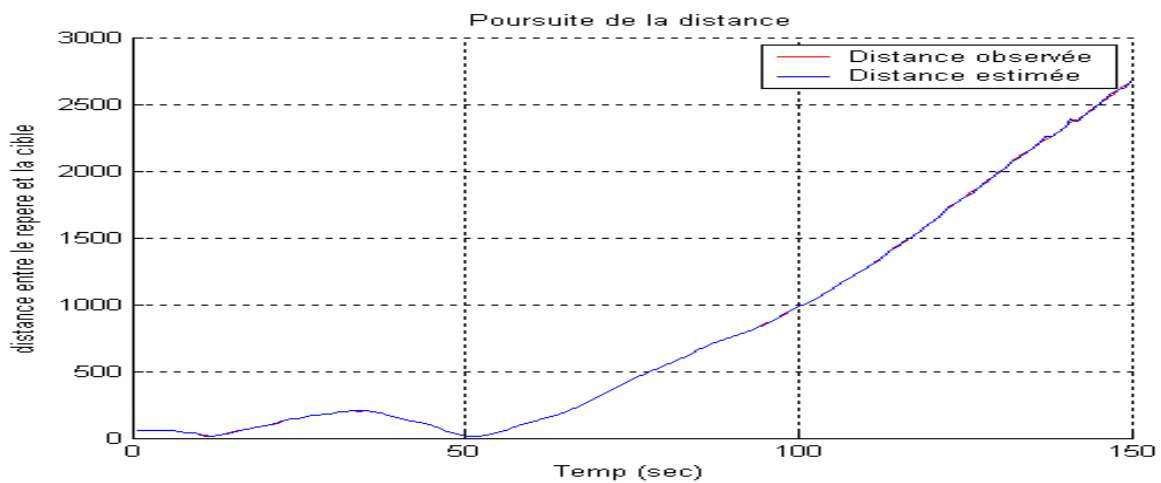


Figure IV-15 La distance vraie, estimée par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.8$

On remarque que les résultats donnés par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.5$ est meilleur que ceux donnés par SIR adaptatif avec $N = 1000$ et $N_{Seuil} = 0.8$.

On constate d'après les figures présentées ci-dessus que le filtre particulaire avec rééchantillonnage adaptatif donne de bons résultats comparés au filtre particulaire avec rééchantillonnage systématique. On obtient aussi de bons résultats, en augmentant le nombre d'échantillons. Pour $N = 150$, le filtre particulaire arrive quand même à suivre la trajectoire mais de temps en temps il la perd pendant des laps de temps. On remarque aussi des résultats différents pour des seuils différents.

Le bruit de dynamique et le bruit d'observation déterminent la fiabilité à la poursuite de filtre particulaire, car pour des valeurs faibles de bruits le filtre particulaire trouve des difficultés à suivre la trajectoire. L'initialisation du filtre joue aussi un rôle primordial dans la poursuite. Le temps de calcul du filtre particulaire adaptatif est beaucoup plus petit que celui du filtre de Kalman étendu.

III Conclusion

Le but de ce chapitre était d'étudier les performances en simulation des méthodes particulières adaptatifs SIR, sur des problèmes de poursuite d'une cible unique observée à travers des mesures de distance et d'angle (estimation de la position d'un mobile dans le plan) et de faire une comparaison avec le filtre de Kalman étendu. Nous avons montré l'avantage pratique du rééchantillonnage adaptatif retardant la dégénérescence des poids par rapport au rééchantillonnage systématique. Nous avons également montré la robustesse d'un tel filtre à une mauvaise initialisation et prouvé sa supériorité devant un filtre de Kalman étendu.

Enfin, nous soulignons que la mise en œuvre des algorithmes particuliers adaptatifs est simple et rapide.

Conclusion générale

Nous nous sommes intéressé dans ce mémoire de magister à l'estimation bayésienne du processus d'état markovien observé à travers le processus de mesures. Plus particulièrement, l'objectif était de fournir une estimation de la loi conditionnelle $p(X_k | Y_{0:k})$. Si le filtre optimal fournit une expression récursive de cette loi, il n'est en pratique calculable que dans des cas très restreints tel que le cas linéaire avec bruits additifs gaussiens, i.e., celui du filtre de Kalman dont nous avons rappelé les équations. Pour traiter le cas général, nous avons exposé dans un premier temps les méthodes déterministes développées depuis les années 70. les filtres de Kalman étendus et le filtre de Kalman unscented reposent sur l'approximation de la loi conditionnelle par une gaussienne. Ils proposent d'en approximer ses deux premiers moments respectivement par linéarisation locale du modèle d'état et de mesure et par approximation de la loi gaussienne par un nuage déterministe de points. Les filtres de Kalman étendus restent la méthode la plus utilisée mais aucun résultat de convergence n'est prouvé dans le cas général. Ils sont de toute façon un mauvais choix lorsque la loi conditionnelle est multimodale. Les méthodes numériques par maillage de l'espace d'état abandonnent l'hypothèse que la loi conditionnelle soit gaussienne et proposent d'approximer les intégrales intervenant dans le filtre bayésien optimal par une somme finie en des points disposés selon un maillage de l'espace d'état. Si cette approche est optimale lorsque l'espace d'état est discret, sa complexité devient beaucoup trop importante et sa convergence beaucoup trop lente lorsque la dimension de l'espace d'état continu dépasse trois ou quatre.

Abandonnant l'approche déterministe, les méthodes dites de Monte Carlo ont permis de développer notamment dans le cadre du filtrage une classe de méthodes très performantes, dénommées méthodes séquentielles de Monte Carlo. A chaque instant, la loi conditionnelle a posteriori est approximée par une somme pondérée de loi de Dirac centrée en des particules. Ces particules et leur poids évoluent au cours du temps selon un échantillonnage pondéré séquentiel et sont rééchantillonnées afin de n'en conserver que les meilleures au sens de la vraisemblance des observations. Devant la profusion d'études les concernant durant la dernière décennie, nous avons proposé un état de l'art présentant tout d'abord le filtre SIS et en suite le filtre SIR (méthode de Monte Carlo séquentiel avec rééchantillonnage) puis le filtre Bootstrap qui est le filtre le plus utilisé, enfin une classe générique englobant la plupart de ses variantes. Plusieurs choix de fonctions de proposition servant à guider au mieux les particules des régions de forte vraisemblance ont été décrits ainsi que plusieurs choix de techniques de rééchantillonnage limitant l'augmentation inutile de la variance des poids des particules. Ces méthodes sont théoriquement validées par l'existence de théorèmes probabilistes assurant notamment la convergence uniforme en temps de la solution particulière sous des conditions de mélange. Une propriété remarquable est leur vitesse de convergence qui ne dépend pas des propriétés de système générateur des processus d'état et de la mesure mais de celles du système utilisé pour la propagation des particules.

Nous avons enfin présenté une application du filtre Bootstrap que nous avons menée pour la poursuite d'une cible unique observée à travers des mesures de distance et d'angle. Nous avons en particulier montré l'avantage pratique du rééchantillonnage adaptatif retardant la dégénérescence des poids par rapport au rééchantillonnage systématique. Nous avons également montré la robustesse d'un tel filtre à une mauvaise initialisation et prouvé sa supériorité devant un filtre de Kalman étendu.

De nombreuses perspectives intéressantes peuvent être envisagées à ce présent travail, à savoir :

- Confronter nos algorithmes de poursuite de cible à des scénari réels,
- Etendre cette approche à la poursuite multicible,
- Appliquer cette approche à des scénari avec des cibles manoeuvrantes,

Bibliographie

- [1] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems." *Trans. ASME, Journal of Basic Engineering*, Vol. 82, pp. 34-45, 1960.
- [2] R. E. Kalman and R. Bucy. "A New Approach to Linear Filtering and Prediction Theory." *Trans. ASME, Journal of Basic Engineering*, Vol. 83, pp. 95-108, 1961.
- [3] Mireille Chaleyat-Maurel and Dominique Michel. "Des Résultats de Non Existence de Filtre de Dimension Finie." *Stochastics*, Vol. 13, pp. 83-102, 1984.
- [4] Daniel L. Ocone and Etienne Pardoux. "A Lie-algebraic Criterion for Non Existence of Finite Dimensionally Computable Filters." In Da Prato and Tubaro, Editors, *Stochastic Partial Differential Equations and Applications*, Volume 1390 of *Lecture Notes in Mathematics*. Springer, 1989.
- [5] Jean Picard, "Efficiency of the Extended Kalman filter." *SIAM Journal on applied mathematics*, Vol. 51, No. 3, pp. 843-885, 1991.
- [6] Zhiqiang Cai, François Le Gland, and Huilong Zhang. "An adaptive Local Grid Refinement Method for Nonlinear Filtering." *Tech. Rep. Technical Report 2679*, INRIA, 1995.
- [7] Harold J. Kushner and P. Dupuis. "Numerical Methods for Approximations in Stochastic Control Problems in Continuous Time." Volume 24 of *Applications of Mathematics*. Springer-Verlag, New-York, 1992.
- [8] Neil J. Gordon, David J. Salmond and Adrian F. M. Smith. "Novel Approach to Nonlinear / Non-Gaussian Bayesian State Estimation." *Radar and Signal processing, IEE Proceedings, Part F*, Vol. 140, No. 2, pp. 107-113, April 1993.
- [9] Pierre Del Moral, Gilles Rigal, and Gérard Salut. "Estimation et Commande Optimale Non-Linéaire." *Technical Report 2*, Convention D.R.E.T.-DIGILOG-LAAS/CNRS, SM.MCY/685.92/A, 89.34.553.00.470.75.01, Mars 1992.
- [10] G. Kitagawa. "Monte Carlo Filter and Smoother for Non Gaussian Nonlinear State Space Models." *Journal of Computational and Graphical Statistics*, Vol. 5, no. 1, pp. 1-25, 1996.
- [11] M. K. Pitt, N. Sheppard. "Filtering via simulation : Auxiliary particle filters." *Journal of the American Statistical Association*, Vol. 94, N° 446, pp. 590-599, 2000.
- [12] C. Musso, N. Oudjane, F. Legland. "Improving regularized particle filters." In *Sequential Monte-Carlo Methods and Practice* (A. Doucet, J.F.G. de Freitas, N.J. Gordon, Eds.), pp. 247-271, New-York, Springer-Verlag, Jan. 2001.
- [13] N. Wiener. "Extrapolation, Interpolation and Smoothing of Stationnary Time Series." *The M.I.T Press*, Cambridge, Massachusetts, 1966.

- [14] Andrew. H. Jazwinski. "Stochastic Process and Filtering Theory." Colloquium Publications. Academic Press, New-York and London, 1970.
- [15] François Le Gland. "Introduction au Filtrage en Temps Discret. Filtre de Kalman-Modèles de Markov Cachés." Dea stir, Université de Rennes I, 1995.
- [16] Peter S. Maybeck. "Stochastic Models, Estimation and Control." Academic Press, Vol. 3 1979.
- [17] B. D. O. Anderson and J. B. Moore. "Optimal Filtering." Providence, RI : Pentice-Hall, 1979.
- [18] C. K. Chui and G. Chen. "Kalman Filtering with Real-Time Applications." Springer-Verlag, New-York, 3rd edition, 1999.
- [19] R. S. Liptser and A. N. Shirayev. "Statistics of Random Processes." Springer, 1977.
- [20] Jean Picard. "Filtrage de Diffusions Vectorielles Faiblement Bruitées." In Analysis and Optimization of Systems, Volume 83 of Lecture Notes in Control and Information Sciences. Springer, 1986.
- [21] Jean Picard. "Asymptotic Study of Estimation Problems with Small Observation Noise." In Stochastic Modelling and Filtering, Volume 91 of Lecture Notes in Control and Information Sciences. Springer, 1987.
- [22] A. Getout-Petit. "Approximation Filter for the Conditional Law of Partially Observed Process in Nonlinear Filtering." SIAM Journal on Control and Optimization, Vol. 36, N°. 4, pp. 1423-1447, 1998.
- [23] S. J. Julier and J. K. Uhlmann. "A New Extension of the Kalman Filter to Nonlinear Systems." Proc. of Aerosense: the 11th International Symposium onj Aerospace/Defense Sensing, Simulation and Controls, Vol. Multi sensor fusion, Tracking and resource management II, Orlando, FL. 1997.
- [24] E. A. Wan and R. Van der Merve. "The Unscented Kalman Filter." At Kalman filtering and neural networks, S. Haykin. 2002.
- [25] A. H. Jazwinski. "Stochastic Processes and Filing Theory." New-York : Academic Press, 1970.
- [26] E. Pardoux. "Filtrage Non Linéaire et Equations aux Dérivées Partielles Stochastiques Associées." Ecole d'été de Probabilités de Saint-Flour XIX, volume 1464 of in lecture Notes in Mathematics. Springer-Verlag, 1989.
- [27] N. Metropolis and S. Ulam. "The Monte Carlo Method." Journal of the American Statistical Association, Vol. 44, pp. 335-341. 1949.
- [28] N. Metropolis, A. Rosenbluth, M. Rosebluth, A. Tell and E. Teller. "Equations of State Calculations by Fast Computing Machines. Journal of Chemical Physics, Vol. 21, pp. 1087-1092. 1953.

- [29] J. Hammersley and K. Morton. "Poor man's Monte Carlo." *Journal of the Royal Statistical Society B*, Vol. 16, pp. 23-38. 1954.
- [30] J. Handschin and D. Mayne. "Monte Carlo Technics to Estimate the Conditional Expectation in Multi-stage Non-linear Filtering." *International Journal of Control*, Vol. 9, N° 5, pp. 547-559. 1970.
- [31] H. Akashi and H. Kumamoto. "Random Sampling Approach to State Estimation in Switching Environments." *Automatica*, Vol. 13, pp. 429-434. 1977.
- [32] A. Doucet, N. De Freitas, and N. J. Gordon. "Sequential Monte Carlo Methods in Practice. Series Statistics for Engineering and Information Science. Springer-Verlag, New York. 2001.
- [33] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Janson, R. Karlsson, and P.-J. Nordlund. "Particule Filters for Positioning, Navigation and Tracking." *IEEE Transactions on Signal Processing*, Vol. 50, N° 2, pp. 425-437. 2002.
- [34] C. Hue, J. Le Cadre, and Perez. "Tracking Multiple Objects with Particule Filtering." *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, N° 3, pp. 791-812. 2002.
- [35] J. Vermaak, C. Andrieu, A. Doucet, and S. Godsill. "Particle Methods for Bayesian Modeling and Enhancement of Speech Signals." *IEEE Transaction on Speech and Audio Processing*, Vol. 10, N° 3, pp. 173-185. 2002.
- [36] C. Kwok, D. Fox, and M. Meil. "Real-Time Particle Filter." *Proceeding of IEEE (Issue on State Estimation)*, Vol. 92, N° 2. 2004.
- [37] M. Isard and A. Blake. "Condensation-Conditional Density Propagation for Visual Tracking." *Int. J. Comput. Vision*, Vol. 29, N° 1, pp. 5-28. 1998.
- [38] P. Pérez, J. Vermaak, et A. Blake. "Data Fusion for Visual Tracking with Particles." *Proc. IEEE*, Vol. 92, N° 3, pp. 495-513. 2004.
- [39] A.W. Marshall. "The Use of Multi-stage Sampling Schemes in Monte-Carlo Computations." Edité par M.A. Meyer, *Symposium on Monte-Carlo Methods*, John Wiley, New-York, pp. 123-140. 1956.
- [40] A. Millet. "Méthodes de Monte Carlo." Université Paris 6. 2005.
- [41] E. Moulines. "Hidden Markov Models and Particle Filters." 2002.
- [42] A. Doucet. "Algorithmes Monte-Carlo pour l'Estimation Bayésienne de Modèles Markoviens Cachés. Application au Traitement de Signaux de Rayonnements." Thèse de l'Université de Paris-Sud, Orsay, Décembre 1997.
- [43] Ludovic Brèthes. "Suivi Visuel Par Filtrage Particulaire. Application à l'Interaction Homme-Robot." Thèse de Doctorat de l'Université Paul Sabatier de Toulouse, Décembre 2005.

- [44] A. Kong, J.S. Liu and W.H. Wong. "Sequential Imputations and Bayesian Missing Data Problems." *Journal of American Statistical Association*, vol 89, N°425, pp. 278-288. 1994.
- [45] A. Doucet. "On Sequential Simulation Based Methods for Bayesian Filtering." Tech. Rep. Technical report CUFD/ F-INFENG/ TR. 310, Signal Processing group, Department of Engineering, University of Cambridge, 1998.
- [46] R. Van Der Merwe, A. Doucet, J.F.G. De Freitas and E. Wan. "The Unscented Particle Filter." Report CUED/F-INFENG/TR 380, Department of Engineering, Cambridge University, August 2000.
- [47] Pierre Del Moral. "Nonlinear Filtering : Interacting Particle Solution." *Markov Processus and Related Fields*, vol. 2, No. 4, pp. 555-580. 1996.
- [48] D. Crisan and A. Doucet. "A Survey of Convergence Results on Particle Filtering for Practitioners." *IEEE Trans. Signal Processing*, Vol. 50, N° 3, pp. 736-746. 2002.
- [49] J. MacCormick. "Probabilistic Modelling and Stochastic Algorithms for Visual Localisation and Tracking. Thèse de Doctorat, Departement of Engineering Science, University of Oxford. 2000.
- [50] C. Berzuini, N. Best, W. Gilks and C. Larizza. "Dynamic Conditionnal Independence Models and Markov Chain Monte-Carlo Methods." *J. Amer Stat Assoc*, 1997.
- [51] C. Hue. "Méthodes Séquentielles de Monte-Carlo pour le Filtrage non Linéaire Multi-objets dans un Environnement Bruité. Applications au Pistage Multi-cibles et à la Trajectographie d'Entités dans des Séquences d'Images 2D." Thèse de doctorat de l'Université de Rennes I, IRISA, Janvier 2003.
- [52] Richard Karlsson. "Simulation Based Methods for Target Tracking." Thesis, division of Automatic Control, Departement of Electrical Engineering, Linkopiings university, Sweden, 2002.
- [53] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp. "A Tutorial on Particle Filters for on Line Nonlinear / Non-Gaussian Bayesian Tracking." Qinetiq Ltd, DSTO, IEEE, 2001.
- [54] Jean-jacques Szkolnik. "Application des Méthodes de Monte Carlo Séquentielles à l'Extraction de Trames Radar." Thèse de doctorat, Université de Bretagne Occidentale, Novembre 2004.
- [55] Karim Dahia. "Nouvelles Méthodes en Filtrage Particulaires. Application au Recalage de Navigation Inertielle par Mesures Altimétriques." Thèse de Doctorat, Université Joseph Fourier, Janvier 2005.
- [56] G. Casella et C. Robert (1996). Rao-Blackwillisation of Sampling Schemes. *Biometrika*, Pages 81-94.
- [57] J. Kotecha and P.Djuric. "Gaussian Sum Particle Filtering." *Signal Processing, IEEE Transactions on / See also Acoustics, Speech, and Signal Processing, IEEE Transactions on*, Vol. 51, N° 10, pp. 2602-2612. 2003.

[58] Van Der R. Merwe and E. Wan. “Gaussian Mixture Sigma-point Particle Filters for Sequential Probabilistic Inference in Dynamic State-space Models.” In Proceeding of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hong Kong. IEEE. 2003.

[59] Mohand Said Djouadi. “Poursuite de Cibles Manoeuvrantes par Vision Active.” Thèse de Doctorat, Ecole Nationale Polytechnique d’Alger, Décembre 2005.

[60] Nadia Oudjane. “Stabilité et Approximations Particulières en Filtrage Non Linéaire. Application au Pistage.” Thèse de Doctorat, Université de Rennes 1, Décembre 2000.

[61] J. Darricau. “Physique et Théorie du Radar.” Concept de Traitement du Signal. Sodipe, 1994.