



Méthodes statistiques fondées sur les groupes de Lie pour le suivi d'un amas de débris spatiaux.

Samy Labsir

► To cite this version:

Samy Labsir. Méthodes statistiques fondées sur les groupes de Lie pour le suivi d'un amas de débris spatiaux.. Traitement du signal et de l'image [eess.SP]. Université de Bordeaux, 2020. Français. NNT : 2020BORD0294 . tel-03103892

HAL Id: tel-03103892

<https://theses.hal.science/tel-03103892>

Submitted on 8 Jan 2021

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

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

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITE DE BORDEAUX

Ecole Doctorale des Sciences Physiques et de l'Ingénieur

Spécialité Automatique, Productique, Signal, Image et
Ingénierie Cognitive

Par **SAMY LABSIR**

Méthodes statistiques fondées sur les groupes de Lie pour le suivi d'un amas de débris spatiaux

Sous la direction de : **AUDREY GIREMUS**

Soutenue le 15/12/2020 devant le jury composé de :

Marc ARNAUDON Professeur, Université de Bordeaux	Examineur
Silvère BONNABEL Professeur, Mines ParisTech	Rapporteur
Hichem SNOUSSI Professeur, Université de technologie de Troyes	Rapporteur
Arnaud BRELOY Maître de conférences, Université de Paris-Nanterre	Examineur
Audrey GIREMUS Maître de conférences, Université de Bordeaux	Directrice de thèse
Brice YVER Ingénieur, CEA-CESTA	Co-encadrant de thèse
Thomas BENOUDIBA-CAMPANINI Ingénieur, CEA-CESTA	Co-encadrant de thèse

Préparée à l'université de Bordeaux
Laboratoire d'accueil : Laboratoire IMS
351, avenue de la libération - 33405 Talence cedex

Remerciements

Ce travail de thèse est le fruit de trois années de thèse effectué au sein du groupe SIGNAL du laboratoire IMS dans le cadre d'une collaboration avec le CEA-CESTA.

Pour commencer, je tiens à remercier Silvère BONNABEL et Hichem SNOUSSI d'avoir accepté de rapporter mon manuscrit de thèse ainsi que Marc ARNAUDON et Arnaud BRELOY d'avoir accepté de l'examiner.

Je remercie Audrey GIREMUS de m'avoir proposé ce sujet de thèse et de m'avoir accompagné et fait progresser scientifiquement tout au long de ces trois années. Je remercie également mes co-encadrants de thèse du CEA-CESTA, Brice YVER et Thomas BENOUDIBA-CAMPANINI, de m'avoir fourni des conseils avisés et de m'avoir fait penser plus "physique".

Je tiens à remercier l'ensemble des doctorants avec qui j'ai eu le plaisir de partager trois années au sein du même laboratoire.

Pour finir, je souhaite remercier l'ensemble de mes proches pour m'avoir toujours soutenu.

Table des matières

Remerciements	3
Table des figures	8
Liste des tableaux	11
Liste des publications	12
Glossaire	13
Introduction	14
1 Débris spatiaux : enjeux et méthodes de surveillance	17
1.1 Débris spatiaux : généralités	17
1.1.1 Contexte	17
1.1.2 Menace de la prolifération	18
1.1.3 Catalogage et systèmes de surveillance spatiaux	19
1.2 Systèmes d'observation des débris	20
1.2.1 Principe général d'un système radar	21
1.2.2 Traitement à la réception	22
1.2.2.1 Expression du signal radar	22
1.2.2.2 Filtrage adapté et échantillonnage du signal reçu	23
1.2.2.3 Détection d'une cible	24
1.2.3 Génération des mesures	25
1.2.3.1 Mesure de la distance	26
1.2.3.2 Mesure des angles d'arrivée	26
1.2.3.3 Modèle de mesure	29
1.3 Modèle dynamique des débris	29
1.3.1 Système de représentation spatiale	29
1.3.2 Définition du modèle de trajectoire	30
1.3.3 Simulation de la trajectoire d'un amas de débris	32
1.4 Pistage de débris spatiaux	34
1.4.1 Pistage d'une cible ponctuelle	34
1.4.1.1 Présentation du filtre bayésien	34
1.4.1.2 Filtre bayésien analytique : le filtre de Kalman	36
1.4.1.3 Filtres bayésien approchés	37
1.4.1.4 Remarques sur le pistage d'une cible ponctuelle	42
1.4.2 Pistage d'une cible étendue	42

1.4.2.1	Modélisation par des matrices aléatoires	43
1.4.2.2	Modélisation comme des formes convexes étoilées	47
1.5	Conclusions du chapitre	51
2	Généralités sur les groupes de Lie	53
2.1	Concepts fondamentaux	54
2.1.1	Définitions générales	54
2.1.2	Algèbre de Lie	55
2.1.3	Exemples de groupe de Lie	56
2.1.4	Opérateurs de non-commutativité	58
2.1.4.1	Adjoint de groupe	58
2.1.4.2	Adjoint sur l'algèbre	59
2.1.4.3	Formules de Baker-Campell-Hausdorff	60
2.2	Dérivation et intégration	61
2.2.1	Dérivées d'ordre 1	61
2.2.1.1	Lien entre différentielle et dérivée	62
2.2.2	Dérivées d'ordre 2	62
2.2.2.1	Connexion affine	62
2.2.2.2	Hessienne sur groupe de Lie	63
2.2.3	Intégration sur groupe de Lie	64
2.2.3.1	Mesure de Haar et fonction modulaire	65
2.2.3.2	Intégration sur groupe de Lie unimodulaire	65
2.3	Statistiques et incertitudes	66
2.3.1	Définir une distribution sur variété	66
2.3.2	État de l'art des distributions sur variétés	67
2.3.3	Distribution gaussienne sur groupe de Lie	67
2.3.3.1	Définition générale	67
2.3.3.2	Régions d'incertitude sur $SE(3)$	69
2.4	Algorithmes d'estimation	71
2.4.1	Algorithmes d'optimisation	71
2.4.1.1	État de l'art	71
2.4.1.2	Algorithme de Newton sur groupe de Lie (LG-N)	72
2.4.1.3	Algorithme de Gauss-Newton sur groupe de Lie (LG-GN)	73
2.4.2	Algorithmes de filtrage	73
2.4.2.1	État de l'art	73
2.4.2.2	Filtre de Kalman invariant (IKF)	74
2.4.2.3	Filtre de Kalman sur groupe de Lie : approche bayésienne	76
2.5	Conclusions du chapitre	81
3	Modèles et algorithmes sur groupe de Lie pour le pistage d'un amas de débris spatiaux	82
3.1	Modélisation de l'amas comme une cible étendue sur groupe de Lie	83
3.1.1	Démarche proposée	83
3.1.2	Modèle de dispersion des réflecteurs	84
3.1.3	Modèle d'observation	85
3.1.4	Illustrations du modèle	86
3.1.5	Problème d'estimation considéré	87
3.2	Estimation statique des paramètres de l'amas	89
3.2.1	Estimateur <i>a posteriori</i>	89

3.2.1.1	Expression de la vraisemblance	90
3.2.1.2	Expression des <i>a priori</i>	91
3.2.2	Calcul de l'estimateur par un algorithme d'optimisation	92
3.2.2.1	Remarques sur l'approche	95
3.3	Algorithme de pistage de l'amas	95
3.3.1	Modèles d'évolution	96
3.3.1.1	Modèle d'évolution du centroïde	96
3.3.1.2	Modèle d'évolution de la matrice d'étendue	98
3.3.1.3	Modèle d'état augmenté	98
3.3.2	Description de l'algorithme de pistage	98
3.3.2.1	Étape de prédiction	98
3.3.2.2	Étape de correction	99
3.3.2.3	Approximation de Laplace	100
3.4	Validation expérimentale des algorithmes	103
3.4.1	Définitions des métriques d'estimation	103
3.4.1.1	Métrique intrinsèque sur groupe de Lie	104
3.4.1.2	Métrique de comparaison de forme	105
3.4.2	Mise en oeuvre de l'algorithme statique	107
3.4.2.1	Premier scénario	108
3.4.2.2	Simulation selon le modèle physique	111
3.4.3	Mise en oeuvre de l'algorithme de pistage	115
3.4.3.1	Premier scénario	115
3.4.3.2	Second scénario : simulation selon le modèle gravitationnel	120
3.5	Conclusions du chapitre	124
4	Mise en place d'une borne bayésienne d'estimation sur groupe de Lie	125
4.1	État de l'art des bornes de performance	126
4.1.1	État de l'art des bornes euclidiennes	126
4.1.1.1	Borne de Cramér-Rao fréquentiste	126
4.1.1.2	Borne de Cramér-Rao bayésienne	127
4.1.2	État de l'art des bornes de performances sur variétés	127
4.2	Proposition d'une borne de Cramér-Rao sur groupe de Lie	128
4.2.1	Motivations	128
4.2.2	Expression générale	129
4.2.2.1	Démonstration	130
4.2.3	Hypothèses sur $\hat{\mathbf{X}}$ et expression approchée de la borne	133
4.2.4	Expression analytique	134
4.2.4.1	Cas de $SO(3)$	134
4.2.4.2	Cas de $SE(3)$	135
4.3	Mise en oeuvre sur le modèle d'observation proposé	138
4.3.1	Expression de $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$:	139
4.3.2	Expression de $\mathbb{E}(\mathbf{H}_{\mathbf{M}}(\mathbf{M}))$	142
4.3.2.1	Remarques sur la borne	143
4.3.3	Simulations	143
4.4	Conclusion du chapitre	146
	Conclusions et perspectives	147

A	Résolution et discrétisation d’une équation différentielle	151
A.1	Cas classique linéaire	151
A.2	Cas non linéaire	152
B	Calculs usuels sur groupe de Lie	155
B.1	Calculs usuels de dérivée de Lie	155
B.2	Calcul de hessienne sur groupe de Lie	158
B.3	Approximation de Gauss-Laplace et Laplace	158
B.3.1	Rappels dans le cas euclidien	158
B.3.1.1	Approximation de Gauss-Laplace	158
B.3.1.2	Approximation de Laplace	159
B.3.2	Sur groupe de Lie	159
B.3.2.1	Approximation de Gauss-Laplace	159
B.3.2.2	Approximation de Laplace	160
C	Calculs des dérivées du critère d’optimisation	162
C.1	Calcul de la jacobienne de $\varepsilon(\mathbf{X}_k^{v,Z})$	162
C.2	Calcul de la dérivée de Lie et de la hessienne de $J_{nq}(\mathbf{X}_k^{v,Z})$	165
D	Compléments sur le calcul de la borne de Cramér-Rao	167
D.1	Rappel de la démonstration de la borne de Cramér-Rao <i>a posteriori</i> (4.7) . . .	167
D.2	Détails du calcul de $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$	168
D.3	Détails du calcul de $\mathbb{E}(\mathbf{H}_{\mathbf{M}}(\mathbf{M}))$	169
	Bibliographie	170

Table des figures

1.1	Évolution du nombre de débris depuis 50 ans, source : NASA Orbital Debris Program Office.	18
1.2	Radar FPS-85 (Eglin, Floride, États-Unis), source : cmano-db.com.	19
1.3	Radar FPS-108 désigné sous le nom de Cobra Dane (Îles Aléoutiennes, Alaska, États-Unis) , source : technologiemedi.net.	19
1.4	Système radar du projet <i>Space Fence</i> (Kwajalein Atoll, Îles Marshall) source : spacenews.com.	20
1.5	Radar GRAVES (Montseugny, France), source : www.theagilityeffect.com.	20
1.6	Principe général d'un système radar bistatique.	21
1.7	Représentation des coordonnées de la cible dans le repère radar sphérique.	22
1.8	Illustration des cases radars et de l'échantillonnage spatio-temporel.	24
1.9	Illustration des différentes configurations de cibles sur l'axe distance.	24
1.10	Illustration de la détection radar : la vraisemblance du signal reçu dépend de l'hypothèse considérée.	25
1.11	Principe de mesure du retard sur le signal filtré.	26
1.12	Principe du sequential lobing (sur un plan en deux dimensions) : l'antenne corrige sa direction vers le lobe de plus forte amplitude (ici le lobe en bleu).	27
1.13	Principe de la mesure des angles d'arrivée par comparaison d'impulsion.	28
1.14	Trajectoires simulées de deux débris spatiaux dans le repère ECI pour $T = 0.1s$ et de vitesses initiales $[6.10^3, -10^3, 5.10^3] \text{ m s}^{-1}$	32
1.15	Évolution des trajectoires d'un amas de débris pour $N = 500$ débris et une dispersion initiale de 50 km.	33
1.16	Évolution des trajectoires d'un amas de débris pour $N = 500$ débris et une dispersion initiale de 500 km.	33
1.17	Illustration du pistage d'une cible ponctuelle.	35
1.18	Principe du filtre bayésien.	35
1.19	Approximation gaussienne.	38
1.20	Principe de la transformée sans parfum.	39
1.21	Représentation simplifiée de la génération de mesures issues d'une cible étendue pour un système radar.	43
1.22	Illustration de l'approche <i>matrice aléatoire</i>	44
1.23	Modélisation et pistage d'une forme convexe étoilée : $\phi_{k,i}$ correspond à l'angle d'orientation du réflecteur $\mathbf{z}_{k,i}$ et $r_{k,i}$ à la distance entre le centroïde $\mathbf{p}_k$ et $\mathbf{z}_{k,i}$	47
1.24	Méthode de pistage d'un objet étendu fondée sur des processus gaussiens.	50
2.1	Relation entre $T_{\mathbf{I}}G$ et $T_{\mathbf{X}}G$ à travers l'opérateur $L_{\mathbf{I}\mathbf{X}}^R$	55
2.2	Relation entre l'espace euclidien $\mathbb{R}^m$, l'algèbre de Lie $\mathfrak{g}$ et le groupe de Lie G	56
2.3	Construction d'un échantillon $\mathbf{Y}$ selon une distribution gaussienne concentrée à gauche sur un groupe de Lie G à partir d'un échantillon gaussien sur $\mathbb{R}^m$	69

2.4	Nuage de points observé pour $\mathbf{R}_i$ échantillonnée selon une distribution gaussienne concentrée sur $SO(3)$ de matrice de covariance $\mathbf{\Sigma}_R = \text{diag}([0.2^2, 0.1^2, 0.1^2])$	70
2.5	Nuage de points observé pour $\mathbf{R}_i$ échantillonnée selon une distribution gaussienne concentrée sur $SO(3)$ de matrice de covariance $\mathbf{\Sigma}_R = \text{diag}([0.5^2, 0.1^2, 0.1^2])$	70
2.6	Représentation géométrique de la mise à jour d'un paramètre par un algorithme d'optimisation sur groupe de Lie.	71
3.1	Trajectoire d'un amas de débris se courbant sous l'effet de la force de gravitation. L'amas est représenté toutes les 2000 secondes.	83
3.2	Modélisation géométrique de l'amas de débris.	84
3.3	Représentation géométrique du modèle proposé : la multiplication à gauche permet de transporter les réflecteurs sur une forme courbée approchant la forme vraie.	86
3.4	Exemples de formes générées selon le modèle (3.7).	87
3.5	Dispersion des réflecteurs générés selon les modèles (3.4) et (3.7).	87
3.6	Représentation hiérarchique du problème d'estimation	90
3.7	Architecture de l'algorithme proposé.	103
3.8	Représentation géométrique de l'écart intrinsèque.	105
3.9	Principe du calcul de la distance de Hausdorff dans le cas où deux ensembles S_1 et S_2 sont constitués de 3 points.	106
3.10	Comparaison des points prédits aux points issus de la forme vraie.	107
3.11	Forme et centroïde reconstruits à la première itération.	109
3.12	Forme et centroïde reconstruits à la dernière itération.	110
3.13	Évolution du critère d'optimisation au cours des 8 premières itérations.	111
3.14	Évolution de la RMSE intrinsèque sur $\mathbf{M}_k$ au cours des itérations.	111
3.15	Évolution de la RMSE intrinsèque sur $\mathbf{D}_k$ et $\mathbf{P}_k$ au cours des itérations.	111
3.16	Enveloppe de l'amas considéré pour le test de l'algorithme.	112
3.17	Forme estimée à l'itération initiale superposée à la forme de l'amas et aux réflecteurs.	113
3.18	Forme reconstruite selon un premier angle de vue.	113
3.19	Forme reconstruite selon un second angle de vue.	114
3.20	Forme reconstruite selon un troisième angle de vue.	114
3.21	Évolution du critère d'optimisation au cours des itérations.	115
3.22	Évolution de l'erreur d'estimation intrinsèque sur $\mathbf{M}_k$ au cours du temps.	117
3.23	Évolution de l'erreur d'estimation intrinsèque sur $\mathbf{P}_k$ et $\mathbf{D}_k$ au cours du temps.	117
3.24	Évolution des formes et de la trajectoire du centroïde estimées superposées aux formes et trajectoires du centroïde vraies.	118
3.25	Formes générées selon notre modèle pour différentes valeurs de α et $s = 10^5$ m.	118
3.26	Formes générées selon notre modèle pour différentes valeurs de s et $\alpha = 0.3$ rad.	119
3.27	Distance de Hausdorff moyenne en fonction de α pour notre approche et l'approche fondée sur les processus gaussiens pour $s = 10^2$ m.	119
3.28	Distance de Hausdorff moyenne en fonction de α pour notre approche et l'approche fondée sur les processus gaussiens pour $s = 10^4$ m.	120
3.29	Formes estimées superposées aux formes vraies pour différents instants : notre approche	122
3.30	Formes estimées superposées aux formes vraies aux mêmes instants : approche processus gaussiens	122
3.31	Représentation des distances de Hausdorff moyennées en fonction du temps.	123

3.32	Distances de Hausdorff obtenues pour différentes valeurs de σ_f comprises entre 1 et 300 et l entre 0 et 1. Le point bleu correspond à la distance de Hausdorff obtenue en déterminant les paramètres au sens du maximum de vraisemblance.	123
4.1	Construction d'une borne bayésienne sur groupe de Lie dans le cas où les observations et le paramètre inconnu vivent sur le même groupe de Lie.	129
4.2	Représentation du support de la distribution <i>a posteriori</i> $p(\mathbf{X} \mathbf{Z})$ sur G : l'estimateur $\hat{\mathbf{X}}$ est représenté en vert. $\mathbf{X}_1$ et $\mathbf{X}_2$ correspondant à des réalisations selon $p(\mathbf{X} \mathbf{Z})$	132
4.3	Polynôme fonction de $\sqrt{E_{int}}$: la borne sur $SE(3)$ est fournie par x_2^2	138
4.4	Évolution de la borne proposée superposée à l'erreur moyenne quadratique pour $Nr = 500$ réalisations et $\Sigma_M = \text{diag}([0.1, 0.1, 0.1, 100, 100, 100])$	144
4.5	Évolution de la borne proposée superposée à l'erreur moyenne quadratique en fonction de la variance <i>a priori</i> pour $n = 50$ et avec une covariance en rotation <i>a priori</i> égale à $\text{diag}([0.1, 0.1, 0.1])$	144
4.6	Évolution de la borne proposée en fonction de la distorsion avec $n = 50$, $Nr = 500$ et $\Sigma_M = \text{diag}([0.1, 0.1, 0.1, 100, 100, 100])$ superposée à l'erreur moyenne quadratique.	145

Liste des tableaux

3.1	Paramètres de simulation de l'algorithme.	109
3.2	Incertitudes de mesures.	112
3.3	Écart-types des bruits des modèles d'état.	116
3.4	Paramètres de l'algorithme.	116
3.5	Paramètres physiques de l'amas initial.	121
4.1	Paramètres de simulation.	143
4.2	Réglages algorithmiques.	143

Liste des publications

Conférences internationales :

Samy Labsir, Audrey Giremus, Guillaume Bourmaud, Brice Yver et Thomas Benoudiba–Campanini, *Tracking A cluster of space debris by filtering on Lie groups*, Brighton, RU, *ICASSP 2019*.

Samy Labsir, Audrey Giremus, Brice Yver, Thomas Benoudiba–Campanini, *A Lie-group based modelling for centroid and shape estimation of a cluster of space debris*, Amsterdam, NH, *EUSIPCO 2020*.

Conférence nationale :

Samy Labsir, Audrey Giremus, Brice Yver et Thomas Benoudiba–Campanini , *Modèle et algorithme sur groupe de Lie pour l'estimation de la forme d'un amas de débris spatiaux*, Lille, France, *GRETSI 2019*.

Article de revue :

Samy Labsir, Audrey Giremus, Brice Yver et Thomas Benoudiba–Campanini, *Joint shape and centroid position tracking of a cluster of space debris by filtering on Lie groups*, Elsevier Signal Processing, 2020 (en cours de révision).

Glossaire

ECEF	référentiel terrestre.
ECI	référentiel géocentrique.
PFD	principe fondamental de la dynamique.
FKEI	filtre de Kalman étendu itéré.
FKSP	filtre de Kalman sans parfum.
FKI	filtre de Kalman invariant.
LG-FKE	filtre de Kalman étendu sur groupe de Lie
LG-FKEI	filtre de Kalman étendu itéré sur groupe de Lie.
LG-N	algorithme de Newton sur groupe de Lie.
LG-GN	algorithme de Gauss-Newton sur groupe de Lie.
BCH	formule de Baker-Campbell-Hausdorff.
RMSE	racine carrée de l'erreur quadratique moyenne.

Introduction

Les débris spatiaux sont des objets artificiels mais non opérationnels en orbite autour de la Terre comme des fragments d'anciens satellites, de navettes ou de fusées. Généralement, ils sont issus de désintégrations opérées de manière volontaire ou bien causées de manière accidentelle (collision). Depuis une cinquantaine d'années, leur nombre augmente de manière continue et exponentielle telle que décrit par le syndrome de Kessler [1]. Ces débris représentent ainsi une menace pour l'ensemble des satellites fonctionnels et par conséquent pour l'activité humaine. De par leur potentielle dangerosité, il apparaît donc comme un défi majeur d'être capable de les détecter, d'estimer leur dynamique, ou encore de les classer, et ceci à des fins de catalogage. Pour y parvenir, des systèmes de détection et de mesure radar dédiés à la surveillance spatiale ont notamment été conçus et déployés sur le sol terrestre. Par l'émission d'une onde électromagnétique qui est rétrodiffusée par les cibles d'intérêt, ils nous fournissent des informations indirectes associées à ces cibles. À titre d'exemple, certains radars calculent la distance sol-cible ainsi que les angles d'arrivée de l'onde rétrodiffusée. Ces données fournies séquentiellement peuvent ainsi être utilisées pour réaliser un pistage des débris observés, c'est à dire estimer au cours du temps leur position et leur vitesse.

Dans le cadre de cette thèse, nous nous intéressons à un scénario précis : deux objets orbitaux se rencontrent et entrent en collision. À l'issue de cet événement, les débris résultant de ces deux objets sont très proches entre eux et forment un amas compact. Par conséquent, il paraît opportun d'estimer la dynamique de ces débris non pas individuellement mais plutôt dans leur ensemble.

Les systèmes de surveillance actuels possèdent des capteurs de très haute résolution. Par exemple, un système radar moderne peut détecter un débris de très petite dimension (inférieure à 50 m). Néanmoins, l'amas que nous cherchons à caractériser présente une dimension supérieure à la résolution du système considéré et donne lieu à plusieurs détections donc plusieurs observations radar. Ce type de cible est communément appelé cible étendue : sa description nécessite de prendre en compte, en plus de ses paramètres dynamiques, son étendue spatiale.

Ainsi, les questions qui se posent naturellement sont : comment décrire l'étendue de l'amas et comment réaliser le pistage de ce type de cible ? La manière la plus pertinente de caractériser la dispersion spatiale de l'amas est de proposer une modélisation paramétrique de sa forme. Étudier l'évolution de la cible au cours du temps revient alors à estimer les paramètres du modèle considéré en plus de la dynamique de son centroïde. Dans des applications terrestres, les objets étendus que l'on cherche à modéliser et pister peuvent être représentés par des formes géométriques simples. Par exemple, la forme d'un véhicule peut être modélisée par un rectangle, celle d'un avion par une ellipse ou une fusion de plusieurs ellipses. Dans ces cas précis, les paramètres utilisés peuvent être les dimensions (longueur, largeur, demi-axe) de ces figures. Lorsque l'objet étendu admet une forme plus complexe alors des paramétrisations plus avancées existent. En particulier, des méthodes de pistage fondées sur des approches bayésiennes permettent d'estimer récursivement les paramètres dynamiques ainsi que les paramètres de forme à partir de filtres de Kalman étendu. En particulier, nous savons que

les distributions gaussiennes donnent lieu à des réalisations qui se répartissent de manière ellipsoïdale. Par conséquent, ce type de modélisation se prête bien à des objets de forme elliptique et donnant lieu à des mesures réparties en leur volume. Dans notre application, l'amas de débris se modélise durant les premiers instants par une forme elliptique mais rapidement les trajectoires des débris s'infléchissent sous l'effet de la force de gravitation, ce qui entraîne une distorsion de l'amas. Un enjeu est alors d'obtenir une paramétrisation de l'étendue de l'amas prenant en compte cette courbure. De manière analogue, nous pouvons donc nous demander si des modèles probabilistes permettant de générer des réalisations prenant une forme courbée existent.

Une solution est de considérer des lois de probabilités sur des espaces non euclidiens et plus précisément sur des variétés. Ces structures relativement abstraites, issues de la géométrie différentielle, sont des espaces dont chaque élément est contraint par ses propriétés intrinsèques. Un exemple assez parlant physiquement de variété est une hypersphère. Dans le cas tridimensionnel, elle contient l'ensemble des points de $\mathbb{R}^3$ contraint par la propriété de norme unité. Dans notre contexte, la forme courbée de l'amas assure que chacun de ses éléments évolue sur un espace contraint. Dans cette thèse, nous nous focalisons sur une variété différentielle particulière nommée *groupe de Lie*. En particulier, le groupe de Lie $SE(3)$ permet de décrire le mouvement des objets étendus à la fois en termes de rotation et de translation. De plus, pour une loi de probabilité définie sur cet ensemble, l'incertitude sur les paramètres de translation prend la forme d'une bananoïde.

Les contributions majeures de cette thèse sont de trois ordres :

- La première est de définir un nouveau modèle des mesure radar et de l'amas, considéré comme une cible étendue, assurant une dispersion des réflecteurs adaptée à la forme de celui-ci. Ce modèle est basé sur une représentation sur groupe de Lie dans laquelle les débris sont supposés distribués selon une *loi gaussienne concentrée*.
- La seconde est de proposer, à partir de ce modèle, un nouvel algorithme permettant de réaliser au cours du temps l'estimation des paramètres cinématiques et de forme de l'amas de débris. Les grandeurs d'intérêts sont exprimées sur des groupes de Lie bien spécifiques afin de permettre leur estimation conjointe. À cet effet, nous proposons une variante du *filtre de Kalman étendu itéré sur groupe de Lie*. Il s'agit d'un algorithme de filtrage adapté pour des variables d'état et d'observations évoluant sur ces espaces. Il peut être vu comme une généralisation du filtre de Kalman étendu itéré classique, dans lequel la distribution *a posteriori* est approchée par une *distribution gaussienne concentrée*. Le problème d'estimation est alors résolu à chaque instant en identifiant les paramètres de cette dernière par optimisation.
- Une dernière contribution de cette thèse est d'évaluer théoriquement les performances d'estimation de notre algorithme. En effet, lorsque nous estimons des paramètres via un certain modèle, il est pertinent de quantifier l'erreur minimale atteignable en utilisant ce dernier. Dans un cas euclidien, cette information est fournie par les *bornes de Cramér-Rao* fréquentiste et bayésienne. Pour des paramètres évoluant sur variétés, il existe des bornes d'erreur minimales dites intrinsèques et nous proposons d'en définir une dans un contexte bayésien appropriée aux modèles que nous avons défini sur groupe de Lie.

Ce manuscrit est organisé de la manière suivante :

Le chapitre 1 est dédié à la présentation des débris spatiaux ainsi qu'aux motivations de leur catalogage. Les principales méthodes de l'état de l'art pour le pistage d'une cible ponctuelle et

d'une cible étendue sont ensuite présentées. Les approches modélisant dans un premier temps l'objet étendu par une ellipse et dans un second temps par une forme convexe étoilée, sont ainsi détaillées.

Le chapitre 2 présente la théorie des groupes de Lie et en particulier la généralisation des méthodes d'optimisation et de filtrage à ces espaces contraints. À ce titre, il est d'abord introduit aussi bien des outils d'analyse que le formalisme probabiliste sur groupe de Lie. Puis, les principales méthodes d'optimisation hors-ligne et de filtrage récursif intrinsèques aux groupes de Lie sont introduites.

Le chapitre 3 est dédié aux contributions théoriques et algorithmiques. Dans un premier temps, deux nouveaux modèles sont proposés, définis sur groupe de Lie, pour décrire d'une part l'amas et d'autre part les mesures radar. Puis, un algorithme d'optimisation sur groupe de Lie pour l'estimation des paramètres de l'amas, fondé sur ces modèles, est mis en place. Enfin, cet algorithme est généralisé dans un contexte récursif par un filtre de Kalman étendu itéré modifié sur groupe de Lie. En particulier, sont présentés les résultats de simulation de ces algorithmes obtenus pour un amas de débris à un instant précis puis lorsqu'il se disperse au cours du temps.

Le dernier chapitre fournit une contribution théorique dans laquelle nous proposons une nouvelle borne d'erreur de type Cramér-Rao intrinsèque dans un contexte bayésien. Cette borne est d'abord développée dans un cadre générique puis spécifiée dans le cas du groupe de Lie $SE(3)$. Elle est en particulier mise en oeuvre dans le cadre du modèle d'observation proposé.

Chapitre 1

Débris spatiaux : enjeux et méthodes de surveillance

Sommaire

1.1 Débris spatiaux : généralités	17
1.1.1 Contexte	17
1.1.2 Menace de la prolifération	18
1.1.3 Catalogage et systèmes de surveillance spatiaux	19
1.2 Systèmes d'observation des débris	20
1.2.1 Principe général d'un système radar	21
1.2.2 Traitement à la réception	22
1.2.3 Génération des mesures	25
1.3 Modèle dynamique des débris	29
1.3.1 Système de représentation spatiale	29
1.3.2 Définition du modèle de trajectoire	30
1.3.3 Simulation de la trajectoire d'un amas de débris	32
1.4 Pistage de débris spatiaux	34
1.4.1 Pistage d'une cible ponctuelle	34
1.4.2 Pistage d'une cible étendue	42
1.5 Conclusions du chapitre	51

L'objectif de ce chapitre est de présenter la problématique des débris spatiaux. Il résume notamment les différents systèmes et techniques employés pour pouvoir assurer leur surveillance. Dans un premier temps, des généralités sur les débris sont présentées et le danger de leur prolifération est mis en évidence à travers des exemples. Dans un second temps, l'enjeu de leur catalogage est pointé. À ce titre, les modèles dynamiques d'évolution des débris ainsi que les systèmes radar permettant de les observer sont détaillés. Enfin, les algorithmes de pistage permettant leur suivi sont rappelés, aussi bien dans un contexte de cible ponctuelle que de cible étendue.

1.1 Débris spatiaux : généralités

1.1.1 Contexte

Depuis plusieurs décennies, la mise en orbite de satellites joue un rôle central dans l'activité humaine. En effet, les programmes spatiaux sont motivés par des objectifs très variés : les

satellites de télécommunications permettent de relayer de l'information d'un émetteur jusqu'à un récepteur, les satellites de navigation permettent de localiser un utilisateur sur la surface terrestre ou encore, les satellites d'observation permettent de fournir des images de la Terre pouvant être utilisées dans diverses applications (météorologie, biologie, géographie,...). Ils sont devenus aujourd'hui essentiels, que ce soit dans les domaines civils ou militaires.

Durant ces dernières années, de par la recrudescence du nombre de satellites en orbite, plusieurs collisions avec divers engins spatiaux (missiles, lanceurs,...) se sont produites et ont abouti à la création de petits fragments de ces objets. Ils sont nommés débris spatiaux et sont définis de manière générale comme l'ensemble des objets artificiels inactifs en orbite terrestre [2]. Ils peuvent être aussi bien des morceaux d'anciens satellites que des petites particules chimiques et sont donc très hétéroclites. À titre d'exemple, nous pouvons citer la collision récente, survenue en 2009, entre les satellites de télécommunication Iridium-33 et Kosmos-2251, en orbite basse au dessus de la Sibérie qui en a généré près de 600 morceaux [3].

1.1.2 Menace de la prolifération

En raison de l'imprévisibilité de leurs apparitions et de la difficulté à déterminer leurs trajectoires, les débris spatiaux représentent une menace de premier ordre [4][5]. En effet, leur collision avec un satellite peut non seulement aboutir à son dysfonctionnement mais également engendrer la création d'un nouvel amas de débris. Ce scénario répété à plusieurs reprises pourrait entraîner une croissance exponentielle de leur nombre. Si ce dernier devenait trop important, il pourrait devenir compliqué voire impossible pour l'ensemble des satellites en orbite d'éviter un impact avec un nouveau débris. Cette hypothèse envisagée est connue sous le nom de syndrome de Kessler [1]. Les observations récentes corroborent ce risque : le nombre de débris ne cesse de croître et nourrit donc des inquiétudes dans la communauté spatiale.

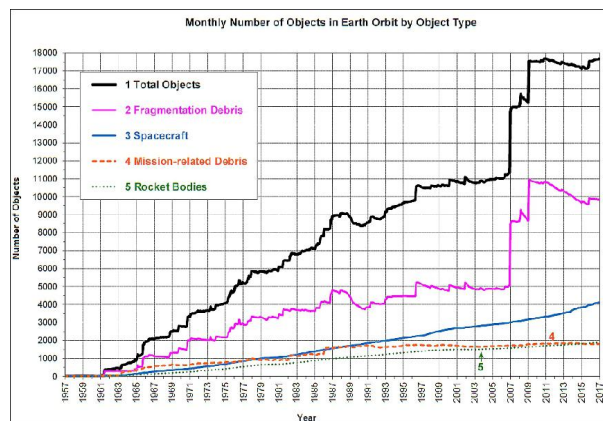


FIGURE 1.1 – Évolution du nombre de débris depuis 50 ans, source : NASA Orbital Debris Program Office.

Les débris de grande dimension sont le plus souvent des satellites inactifs, qui sont généralement de taille variant entre 10 cm et 10 m. Lorsqu'ils entrent en collision avec un autre objet, ils se fragmentent et peuvent générer de nouveaux débris, généralement de taille plus petite mais en nombre plus important. Il existe également des débris de taille infime, issus essentiellement de résidus de missiles ou de fusées. Nous pouvons citer en particulier les gouttelettes de liquide

de refroidissement ou les micro-particules provenant de moteurs de fusées (0.1 mm à 3 cm). Deux solutions ont été envisagées afin de limiter leur prolifération qui dépendent de leur dimension. Ainsi, renforcer la protection des satellites suffit à les protéger de l'impact des débris de petite taille, qui sont les moins dangereux. Néanmoins, ce n'est pas une solution pour les débris de plus grande taille, il n'existe pas actuellement d'autre moyen que de tenter de les éviter, par exemple en réalisant une manoeuvre du satellite menacé. Pour y parvenir, il est donc nécessaire de prédire leurs trajectoires pour être capable d'anticiper leurs déplacements.

1.1.3 Catalogage et systèmes de surveillance spatiaux

L'objectif du catalogage est de déterminer l'évolution dynamique des débris à partir d'un réseau de capteurs situé à la surface terrestre. Ainsi, prédire leurs positions futures permet de prendre une décision afin d'éviter d'éventuelles collisions. Les capteurs utilisés génèrent des informations fournies à des algorithmes de pistage, afin d'en déduire l'estimation des trajectoires des débris. Dans notre application, nous nous focalisons sur des débris en orbite basse, généralement situés à une altitude de l'ordre de 1000 km par rapport au sol terrestre.

De par les enjeux liés à la prolifération des débris, des systèmes de détection dédiés à la surveillance spatiale ont été largement développés par la plupart des grandes puissances mondiales. Il s'agit principalement de systèmes d'antennes radars. Des dispositifs optiques tels que des télescopes sont également utilisés mais pour l'observation de régions lointaines. Les États-Unis possèdent un réseau de surveillance permettant de sonder l'intégralité de l'orbite terrestre. Dans le cas d'orbites basses, les systèmes radars utilisés sont basés sur des réseaux d'antennes à commande de phase. Il en existe plusieurs répartis dans différentes zones géographiques et ils ont des capacités de détection de très haute performance. Le radar FPS-85 [6] (illustré en figure 1.2), mis en service en février 1959, est conçu et dédié à la surveillance de l'espace. Le radar FPS-108 [7] (illustré en figure 1.3), déployé en 1977 pour le pistage de cibles ballistiques, est également utilisé pour cette tâche. Plus récemment, la technologie radar à bande S, fondée sur l'émission d'onde électromagnétique allant de 2 à 4 GHz, a été développée dans le cadre du projet Space Fence, de l'US Air force (illustré en figure 1.4). Cette nouvelle génération de systèmes radars, déclarée opérationnelle le 28 mars 2020, est capable de pister un nombre d'objets bien plus important que les anciens dispositifs.



FIGURE 1.2 – Radar FPS-85 (Eglin, Floride, États-Unis), source : cmano-db.com.



FIGURE 1.3 – Radar FPS-108 désigné sous le nom de Cobra Dane (Îles Aléoutiennes, Alaska, États-Unis) , source : technologiemia.net.



FIGURE 1.4 – Système radar du projet *Space Fence* (Kwajalein Atoll, Îles Marshall)
source : spacenews.com.

Il existe également des systèmes de surveillance au niveau européen comme le radar Tracking and Imaging Radar (TIRA) mais dont les performances de détection en terme de précision ne peuvent concurrencer les systèmes américains. Le système radar français Grand Réseau Adapté à la VEille Spatiale GRAVES [8] (illustré en figure 1.5), développé par l'Office National d'Etudes et de Recherche Aérospatiale (ONERA) est l'unique système européen actuellement en place permettant de détecter et de cataloguer des débris en orbite basse et rivaliser avec les méthodes de catalogage américaines. Il est de type bistatique (fondé sur deux antennes différentes en émission et en réception) et émet une onde électromagnétique de manière continue. Par ailleurs, il fournit une information de type angulaire et Doppler mais pas d'information de distance.



FIGURE 1.5 – Radar GRAVES (Montseugny, France),
source : www.theagilityeffect.com.

La Russie possède également un programme pour la surveillance spatiale. Il est constitué d'un ensemble de plusieurs radars et est connu historiquement sous le nom de OS-1.

1.2 Systèmes d'observation des débris

Comme souligné dans la partie 1, la majorité des systèmes de surveillance spatiale sont des systèmes radars. Cette partie a donc pour objet de présenter succinctement leur principe. En particulier sont décrites la forme du signal radar reçu, et la manière dont le système en extrait des informations, à travers les différents traitements réalisés à la réception.

1.2.1 Principe général d'un système radar

Un système radar moderne est un réseau d'antennes émettant une onde électromagnétique afin de déterminer la position et la vitesse d'un objet en mouvement [9][10]. Pour y parvenir, une antenne émettrice envoie une onde électromagnétique dans une direction de l'espace. La fréquence d'émission peut se situer par exemple dans les bandes L [1 - 2] GHz et S [2 - 4] GHz. Lorsque l'onde rencontre un obstacle, une fraction de son énergie, liée à l'interaction onde/matière, est réfléchiée dans la direction du radar. Le signal retourné, nommé écho radar, va alors être reçu et traité afin d'en déduire des informations sur l'obstacle rencontré. Si l'antenne de réception n'est pas celle qui a servi à l'émission, alors le radar est dit bistatique. Nous nous plaçons dans ce cadre par la suite.

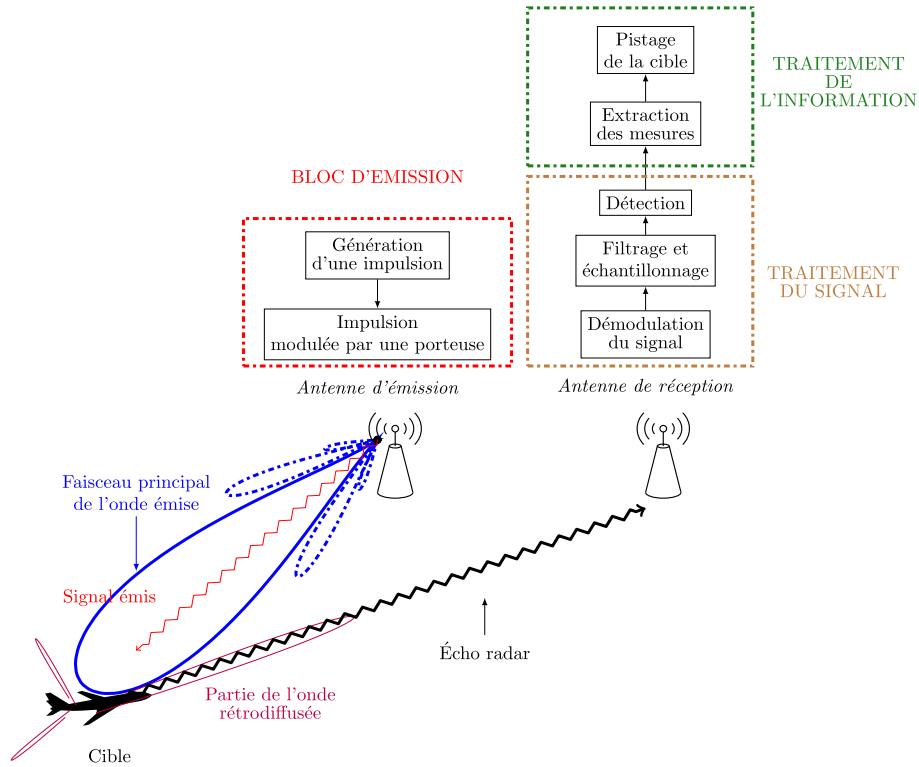


FIGURE 1.6 – Principe général d'un système radar bistatique.

La forme du signal émis dépend du type de radar. Il peut être continu, continu modulé en fréquence ou constitué d'une série d'impulsions.

Comme illustrée en figure 1.6, les caractéristiques de l'onde rétrodiffusée sont étudiées dans un repère associé au radar, en utilisant généralement un système de coordonnées sphériques. Ainsi, la cible en ligne de mire du radar peut être localisée à partir de sa distance par rapport au centre du radar ainsi que par deux angles, dits d'élévation et d'azimuth. Quatre paramètres sont classiquement mesurés :

- Le premier fournit la distance radar-cible R et son calcul est réalisé à partir de la mesure du temps de trajet aller-retour du signal réfléchi par la cible τ . Compte tenu de la vitesse de propagation des ondes dans l'espace c , la relation qui lie les deux paramètres est :

$$R = \frac{c \tau}{2}. \quad (1.1)$$

- Le second et le troisième sont les deux angles d'arrivée de l'écho (élévation θ et azimuth ϕ) qui sont obtenus grâce à la directivité de l'antenne radar.
- Le quatrième fournit le décalage de fréquence Doppler, qui permet d'en déduire une mesure de la vitesse radiale v_r de la cible à partir de la formule :

$$f_d = \frac{2v_r}{\lambda} \quad (1.2)$$

où λ est la longueur d'onde de l'onde émise.

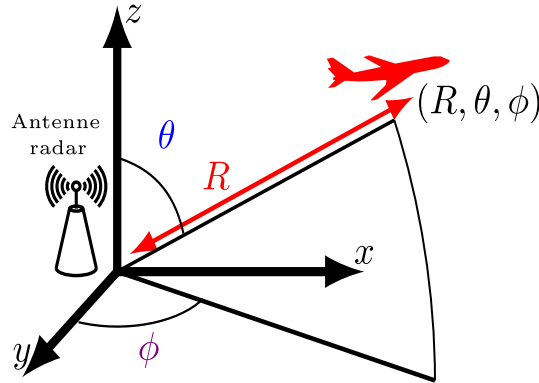


FIGURE 1.7 – Représentation des coordonnées de la cible dans le repère radar sphérique.

Le traitement de l'écho à la réception nécessite deux principales étapes afin d'aboutir à l'estimation des paramètres d'intérêt (position et vitesse), comme illustré sur la figure 1.6 :

- l'étape de traitement du signal radar dans laquelle plusieurs actions sont réalisées :
 - premièrement, le contraste entre le plancher de bruit inhérent au récepteur radar et le niveau de l'écho de cible est amélioré par un filtrage adapté. Cette étape permet d'atténuer le bruit galactique et le bruit thermique lié aux composants radio-fréquences du récepteur.
 - À partir de l'écho filtré, un seuil est déterminé et associé à une probabilité de fausse alarme donnée. Lorsque l'écho d'une cible est supérieur à ce seuil, elle est alors détectée par le radar.
- L'étape de traitement de l'information : en présence d'une détection, le radar génère un *plot* constitué des mesures (distance, angles, fréquence Doppler) dépendant du type de radar. Ce *plot* est associé à une piste (nouvelle ou déjà existante), représentant un objet détecté. Par la suite, il alimente un algorithme de pistage fournissant une estimation de la trajectoire de cette piste.

1.2.2 Traitement à la réception

1.2.2.1 Expression du signal radar

Lorsque l'onde émise est réfléchiée par une cible, l'écho radar peut s'exprimer comme une version atténuée et retardée du signal envoyé. Dans le cas où aucune cible n'est rencontrée par l'onde, il consiste uniquement en un bruit de réception. Ainsi, si nous notons $s(t)$ le signal

émis à l'instant t , l'expression du signal reçu à l'instant t est [11] :

$$\begin{cases} y(t) = \alpha(t) s(t - \tau) \exp(-j(2\pi(f_c + f_d)t + \psi)) + n(t) & \text{si une cible est en vue} \\ = n(t) & \text{sinon} \end{cases} \quad (1.3)$$

Dans cette formule, plusieurs termes apparaissent :

- le coefficient $\alpha(t)$ représente l'atténuation de la puissance émise en réception. Elle est due à de nombreuses causes (portée du signal, puissance des antennes) et s'écrit de la manière suivante [11] :

$$\alpha(t) = \sqrt{\frac{\sigma P_e(t)}{(4\pi^3)}} \frac{\lambda}{R(t)^2} G. \quad (1.4)$$

Cette équation comporte plusieurs quantités physiques :

- $P_e(t)$ (W) correspond à la puissance de l'onde émise.
 - G est le gain de l'antenne supposé le même en émission et en réception. Sa valeur dépend notamment de la position angulaire du faisceau radar.
 - $R(t)$ (m) est la distance entre le centre du radar et la cible.
 - σ (m²) correspond à la *surface équivalente radar* de la cible. Cette quantité caractérise sa capacité à rétrodiffuser l'onde émise.
- Le terme exponentiel fait apparaître la fréquence du signal porteur f_c , la phase à l'origine du signal ψ ainsi que la fréquence Doppler f_d . Cette dernière modifie la fréquence de l'écho.
- $n(t)$ correspond à un bruit thermique provenant de l'électronique des composants du système ainsi qu'au bruit galactique. Il peut se modéliser par un bruit gaussien dont la puissance s'écrit $P_n = k_b T B_r$ avec k_b la constante de Boltzmann (J.K⁻¹), B_r la bande passante du récepteur radar (Hz) et T sa température (K).

1.2.2.2 Filtrage adapté et échantillonnage du signal reçu

À la réception, l'écho reçu est tout d'abord démodulé. Ensuite, un filtrage adapté est réalisé afin d'atténuer le bruit pouvant être induit par les composants électroniques du système [9]. L'allure du signal filtré dépend du type d'impulsion émise. Pour pouvoir être traité, le signal est ensuite échantillonné à une certaine cadence : cette discrétisation s'effectue d'un point de vue temporel mais également d'un point de vue spatial puisque la cible est décrite par une distance ainsi que deux angles dans le repère radar. Comme il est possible de faire correspondre un temps à une distance via la formule (1.1), échantillonner temporellement est équivalent à discrétiser en distance le segment radar/cible. La capacité du senseur à échantillonner l'axe distance ainsi qu'angulairement permet alors de constituer un ensemble de cellules tridimensionnelles pouvant contenir ou non des cibles et appelées cases de résolution radar. Elles définissent la notion de résolution en distance et angulaire du radar qui traduit sa capacité à distinguer deux cibles proches entre elles aussi bien temporellement que spatialement [12][13]. La résolution en distance dépend explicitement de la bande spectrale B (Hz) occupée par le signal en sortie du filtre adapté :

$$\delta D = \frac{c}{2B} \quad (1.5)$$

Nous remarquons alors qu'elle est d'autant plus précise que la largeur spectrale du signal émis sera étroite.

Les résolutions angulaires dépendent des dimensions physiques du panneau radar, ainsi que de λ . Si L correspond à la longueur du panneau et H à sa hauteur, il vient :

$$\begin{cases} \delta_\theta \simeq \frac{65 \lambda}{L} \\ \delta_\phi \simeq \frac{65 \lambda}{H} \end{cases} \quad (1.6)$$

$$(1.7)$$

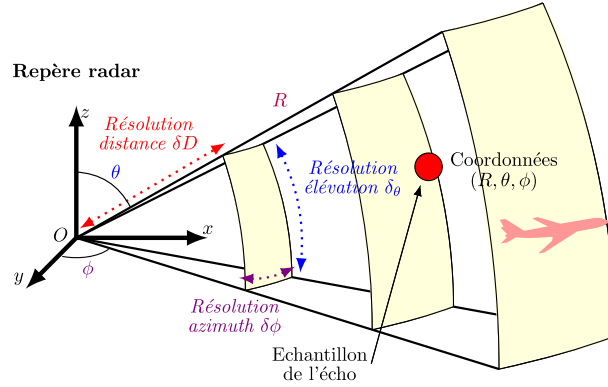


FIGURE 1.8 – Illustration des cases radars et de l'échantillonnage spatio-temporel.

Selon leur dimension, les objets réfléchissant le signal radar peuvent occuper une ou plusieurs cases radar, ils sont respectivement dits ponctuels ou étendus. Dans le cas où deux cibles sont très proches entre elles, elles sont dites non résolues si l'échantillonnage n'est pas suffisamment précis pour pouvoir les distinguer sur deux cases distances : l'écho reçu ne peut être associé à deux cibles différentes.

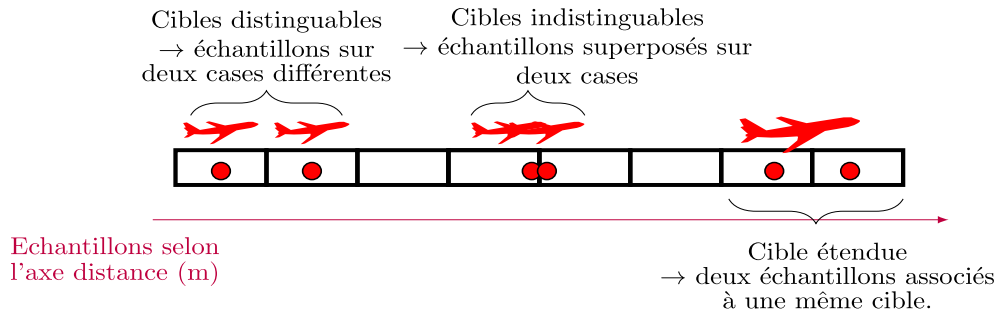


FIGURE 1.9 – Illustration des différentes configurations de cibles sur l'axe distance.

Dans le cas d'une cible ponctuelle, le signal échantillonné après filtrage adapté peut se représenter dans une direction angulaire donnée sous la forme de M échantillons dans un vecteur $\mathbf{s} = [s_1, \dots, s_M]^T$ où M correspond au nombre de cases distance.

1.2.2.3 Détection d'une cible

À partir de la puissance du bruit, il est possible de définir un critère afin de décider si le signal reçu provient d'une cible. Il est basé sur un test statistique dans lequel il est supposé que $\mathcal{H}_0$

correspond à l'hypothèse "cible non présente" et $\mathcal{H}_1$ à l'hypothèse "cible présente".

Le plus connu est celui de Neyman-Pearson [13]. Il consiste à comparer le rapport des vraisemblances du signal reçu $\mathbf{s}$ sous ces deux hypothèses à un certain seuil. Si la vraisemblance associée à l'hypothèse $\mathcal{H}_1$ est supérieure à celle de l'hypothèse $\mathcal{H}_0$ multiplié par le seuil, $\mathcal{H}_1$ est acceptée et une cible est bien détectée.

$$\begin{cases} \text{si } \frac{p(\mathbf{s}_i|\mathcal{H}_1)}{p(\mathbf{s}_i|\mathcal{H}_0)} > \epsilon_1 \Rightarrow \mathcal{H}_1 \text{ est acceptée, } \forall i \in \{1, \dots, M\} \\ \text{sinon } \mathcal{H}_0 \text{ est acceptée.} \end{cases} \quad (1.8)$$

$$(1.9)$$

Il est possible de montrer que cette comparaison est équivalente à comparer la puissance du signal en sortie du filtre à un certain seuil ϵ_2 [10].

$$\begin{cases} \text{si } |\mathbf{s}_i|^2 > \epsilon_2 \Rightarrow \mathcal{H}_1 \text{ est acceptée } \forall i \in \{1, \dots, M\} \\ \text{sinon } |\mathbf{s}_i|^2 < \epsilon_2 \Rightarrow \mathcal{H}_0 \text{ est acceptée } \forall i \in \{1, \dots, M\} \end{cases} \quad (1.10)$$

$$(1.11)$$

où ϵ_2 s'écrit en fonction de la probabilité de fausse alarme et de la variance du bruit thermique P_n :

$$\epsilon_2 = P_n \log(P_{fa}) \quad (1.12)$$

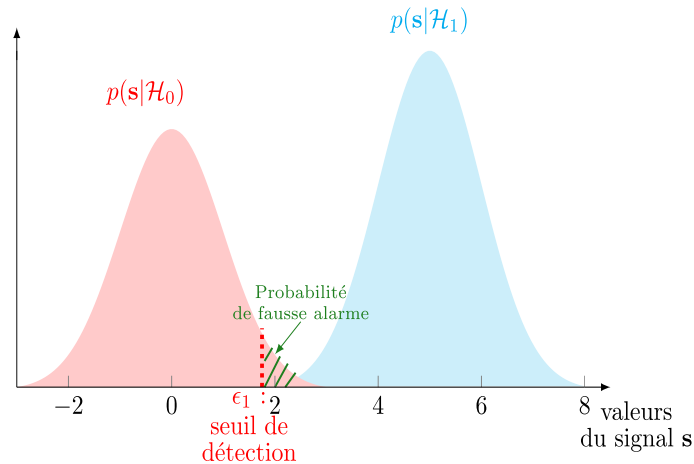


FIGURE 1.10 – Illustration de la détection radar : la vraisemblance du signal reçu dépend de l'hypothèse considérée.

Néanmoins, l'hypothèse d'un bruit de puissance connue est souvent restrictive et peu réaliste. De plus, celle-ci a tendance à varier spatialement et temporellement et ne peut être considérée comme étant la même selon la case radar. Dans ce cas, la détection se fait de manière adaptative. La puissance du bruit, pour chaque case, peut être estimée par exemple en faisant une moyenne de la puissance des cellules radars voisines. Cette méthode est connue sous le nom de *Détection à Taux de Fausse Alarme constant* [9].

1.2.3 Génération des mesures

Dans le cadre des radars à surveillance spatiale, nous nous intéressons exclusivement aux mesures de distance et d'angles d'arrivée dont nous détaillons le principe de génération.

1.2.3.1 Mesure de la distance

En présence d'une cible, le maximum du signal échantillonné après filtrage adapté, s'il a dépassé le seuil de détection, est atteint à un temps correspondant au retard de l'écho τ . Il est alors possible d'estimer celui-ci en récupérant la valeur de l'axe des temps correspondant à l'échantillon maximal. Grâce à la formule (1.1), l'estimation de la distance peut alors être obtenue.

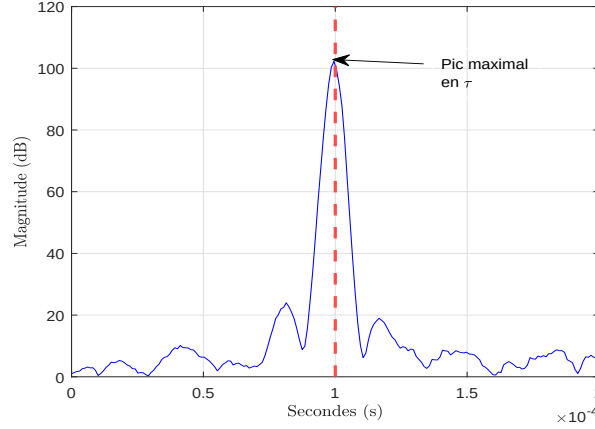


FIGURE 1.11 – Principe de mesure du retard sur le signal filtré.

Néanmoins, la valeur du retard peut potentiellement se situer entre deux échantillons, ce qui engendre une incertitude sur sa mesure. Pour améliorer la précision de son estimation, il existe des méthodes combinant plusieurs échantillons dans l'estimation. La méthode dite du centroïde moyenne un ensemble d'échantillons autour de l'échantillon maximal. Cette valeur est alors interpolée pour en déduire une estimation de la distance cible/radar. Il est démontré que l'écart-type minimal atteignable peut être approché par la formule de Woodward [11] :

$$\sigma_R = \frac{\delta D}{1.6 \sqrt{2r} B} \quad (1.13)$$

où r correspond au rapport signal à bruit du signal filtré. Cette valeur sera utilisée dans les simulations du chapitre 3.

1.2.3.2 Mesure des angles d'arrivée

Pour estimer les angles d'arrivée (élévation et azimuth), le système doit déterminer la direction dans laquelle le faisceau de l'onde reçue pointe. Il existe deux principales méthodes pour obtenir des mesures plus précises :

- la méthode dite du *sequential lobing*. Le principe est de déterminer une valeur précise des angles d'arrivée en utilisant l'information issue de la position prédite de la cible. Le radar pointe son antenne dans deux nouvelles directions situées symétriquement de part et d'autre de la direction prédite. La direction donnant lieu à un signal reçu de plus forte puissance est supposée plus proche de la direction réelle de la cible. Ce processus peut être réitéré pour ajuster séquentiellement la mesure de l'angle d'arrivée.

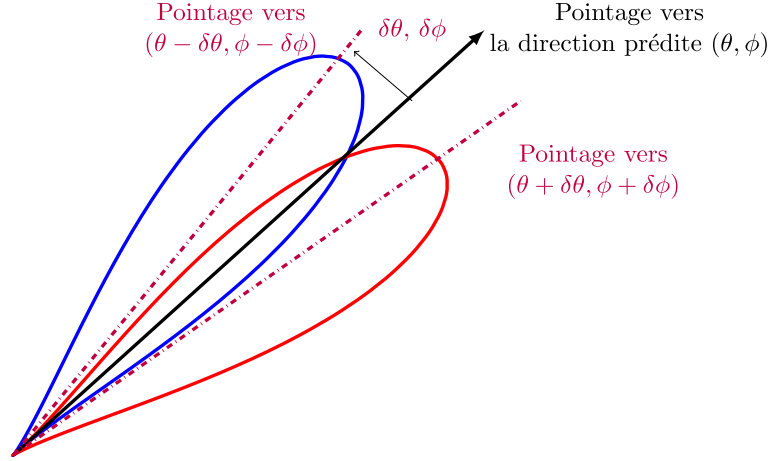


FIGURE 1.12 – Principe du sequential lobing (sur un plan en deux dimensions) : l’antenne corrige sa direction vers le lobe de plus forte amplitude (ici le lobe en bleu).

Cependant, cette approche séquentielle est sensible aux fluctuations d’amplitude impulsion à impulsion.

→ Dans le cadre de radars à mono-impulsion, la méthode dite de *la comparaison d’impulsion* est une alternative plus robuste basée sur le principe d’écartométrie [11]. Dans ce cas, le panneau radar est décomposé en quatre sous-panneaux où l’écho reçu se divise en quatre “sous-signaux” complexes $\{s_i(t)\}_{i=1}^4$ comme illustré sur la figure 1.13. Trois signaux sont alors définis :

→ un signal “somme” s’écrivant :

$$s_{\Sigma} = \sum_{i=1}^4 s_i, \quad (1.14)$$

→ un signal des différences verticales :

$$s_{\theta} = (s_1 + s_3) - (s_2 + s_4), \quad (1.15)$$

→ un signal des différences horizontales :

$$s_{\phi} = (s_1 + s_2) - (s_3 + s_4). \quad (1.16)$$

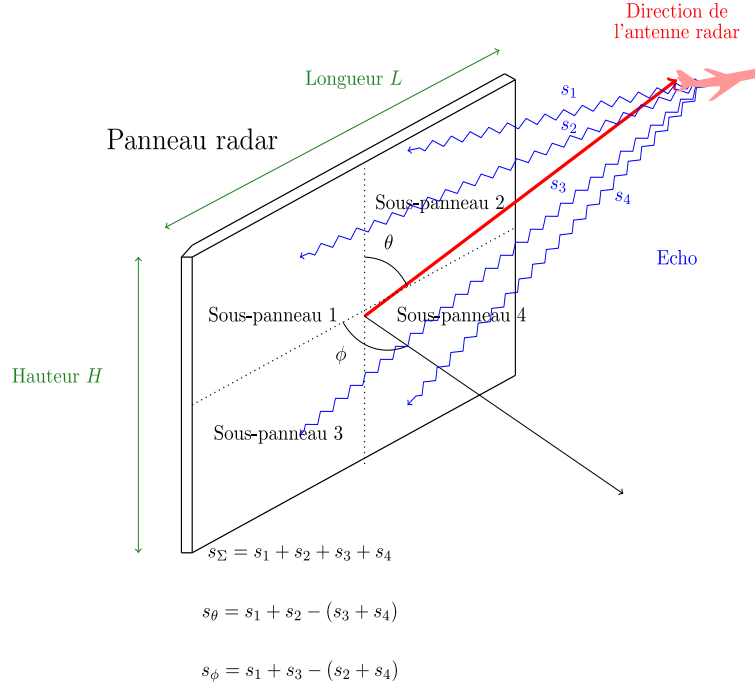


FIGURE 1.13 – Principe de la mesure des angles d'arrivée par comparaison d'impulsion.

Chacun de ces signaux étant complexe, ils peuvent être divisés en une partie réelle I et une partie imaginaire Q :

$$s_{\Sigma,I}(t) = \alpha(t) \cos(\psi) + w_{\Sigma,I} \quad s_{\Sigma,Q}(t) = \alpha(t) \sin(\psi) + w_{\Sigma,Q} \quad (1.17)$$

$$s_{\theta,I}(t) = \alpha(t) \eta_v \cos(\psi) + w_{\theta,I} \quad s_{\theta,Q}(t) = \alpha(t) \eta_v \sin(\psi) + w_{\theta,Q} \quad (1.18)$$

$$s_{\phi,I}(t) = \alpha(t) \eta_h \cos(\psi) + w_{\phi,I} \quad s_{\phi,Q}(t) = \alpha(t) \eta_h \sin(\psi) + w_{\phi,Q} \quad (1.19)$$

où ψ est la phase à l'origine de l'équation (1.3), η_h et η_v correspondent respectivement au rapport du gain de réception de $s_{\theta}(t)$ et de $s_{\phi}(t)$ avec le gain de réception de $s_{\Sigma}(t)$. $w_{\Sigma,I}, w_{\Sigma,Q}, w_{\theta,I}, w_{\theta,Q}, w_{\phi,I}, w_{\phi,Q}$ sont des bruits supposés blancs gaussiens.

Une approximation des angles d'arrivée peut alors être obtenue par les formules suivantes :

$$\begin{cases} \theta \simeq \frac{\eta_v \theta_3}{k} \\ \phi \simeq \frac{\eta_h \phi_3}{k} \end{cases} \quad (1.20)$$

$$\quad (1.21)$$

où θ_3 et ϕ_3 correspondent aux angles d'arrivée du faisceau d'onde dont l'amplitude est atténuée de 3 dB par rapport à son maximum. k est un facteur compris entre 1 et 2 [11]. Les incertitudes sur les deux angles mesurés peuvent alors être quantifiées et approchées par les formules de Woodward données par :

$$\begin{cases} \sigma_{\theta} \simeq \frac{65 \lambda}{H 1.6 \sqrt{2} r} \\ \sigma_{\phi} \simeq \frac{65 \lambda}{L 1.6 \sqrt{2} r} \end{cases} \quad (1.22)$$

$$\quad (1.23)$$

1.2.3.3 Modèle de mesure

À partir des formules de changement de système entre les coordonnées cartésiennes et les coordonnées sphériques, il est possible d'en déduire les relations entre les mesures (R, θ, ϕ) et (x, y, z) les coordonnées de la cible dans un repère local attaché au radar comme montré sur la figure 1.7. En intégrant les incertitudes sur chaque composante, nous obtenons les mesures entachées d'erreurs suivantes :

$$R = \sqrt{x^2 + y^2 + z^2} + b_R \quad (1.24)$$

$$\theta = \arccos \left(\frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) + b_\theta \quad (1.25)$$

$$\phi = \arctan \left(\frac{y}{x} \right) + b_\phi \quad (1.26)$$

où b_R , b_θ et b_ϕ sont des bruits blancs gaussiens centrés respectivement de variance σ_R^2 , σ_θ^2 et σ_ϕ^2 .

1.3 Modèle dynamique des débris

1.3.1 Système de représentation spatiale

L'ensemble des débris spatiaux évolue en orbite autour de la Terre, et au même titre que les satellites, leur mouvement est déterminé par les lois de la mécanique classique [14]. Plus précisément, les lois de Kepler stipulent que tout corps en orbite terrestre suit une trajectoire elliptique dont la Terre est l'un des foyers. Le mouvement des objets en orbite peut être décrit dans deux principaux référentiels cartésiens :

- le référentiel géocentrique : son origine est le centre de masse de la Terre et ses trois axes pointent vers des étoiles lointaines apparaissant fixes. L'axe z est en particulier dirigé suivant l'axe de rotation de la Terre. Ainsi, tout objet fixe sur le sol est en rotation selon ce repère, à une vitesse de rotation correspondant à celle de la Terre. Dans le système international, il est communément nommé *Earth-Centered Inertial (ECI)*.
- Le référentiel terrestre correspond au repère dont l'origine est également au centre de la Terre mais dont les trois axes sont liés au globe terrestre. Son axe z coïncide avec celui du repère *ECI*. Les axes x et y sont situés sur le plan de l'équateur : l'axe x est localisé à l'intersection avec le méridien de Greenwich. Chacun de ces axes se déplace selon la rotation de la Terre, et par conséquent, tout objet sur le sol terrestre paraît immobile dans ce repère. Il est donc en rotation par rapport au référentiel géocentrique. Son appellation dans le système international est *Earth-Centered Earth-Fixed (ECEF)*. Dans notre application, nous utiliserons ce référentiel puisque le système radar est fixé à la surface de la terre.

Il est également possible d'exprimer les coordonnées d'un objet à partir d'un système géodésique mondial nommé *WGS84*. Dans ce système, un point en orbite terrestre est localisé par ses coordonnées géographiques que sont sa latitude, sa longitude et son altitude. Les systèmes géodésiques sont particulièrement utilisés pour la navigation inertielle pour identifier la position d'un mobile.

1.3.2 Définition du modèle de trajectoire

Dans cette partie est défini le modèle de mouvement d'un débris à partir des équations de la mécanique classique. En particulier, le principe fondamental de la dynamique (PFD) s'exprime sous la forme d'une équation différentielle dont la résolution permet de déterminer l'évolution de la position et de la vitesse du débris au cours du temps.

Nous savons qu'un objet orbital au sens général, peut être soumis à deux types de forces : l'attraction gravitationnelle et les forces de frottements. Dans le cas où son altitude se situe entre cent et mille kilomètres, les frottements peuvent être considérés comme négligeables.

Pour déterminer le modèle de trajectoire, nous supposons que l'orbite est exprimée dans le référentiel ECEF. Comme ce dernier est en rotation constante, il ne peut être considéré comme galiléen. Cette rotation induit alors la présence de deux forces supplémentaires dans le PFD que sont la force de Coriolis ainsi que la force d'entraînement.

Si nous notons $\mathcal{R}(\mathbf{O}, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ le référentiel ECEF et $\mathbf{P}$ le centroïde du débris, considéré ici comme ponctuel et de masse m , sa position est représentée par le vecteur $\mathbf{OP}(t) = \mathbf{p}(t) = [x(t), y(t), z(t)]^\top$ et son vecteur vitesse est $\mathbf{v}(t) = [v_x(t), v_y(t), v_z(t)]^\top$ dans ce repère. En supposant le débris ponctuel, l'application de la seconde loi de Newton nous permet d'écrire que :

$$m \mathbf{a}(\mathbf{P}/\mathcal{R})(t) = \mathbf{F}(t) + \mathbf{F}_e(t) + \mathbf{F}_c(t) \quad (1.27)$$

Dans cette équation fondamentale, quatre termes interviennent :

- $\mathbf{a}(\mathbf{P}/\mathcal{R})$ définit l'accélération du point P dans le référentiel $\mathcal{R}$.
- $\mathbf{F}(t)$ correspond à la force gravitationnelle. Elle modélise l'attraction terrestre que subit l'objet en orbite et s'écrit explicitement :

$$\mathbf{F}(t) = -\frac{G M m}{(x(t)^2 + y(t)^2 + z(t)^2)^{3/2}} \mathbf{p}(t) \quad (1.28)$$

où G est la constante gravitationnelle de la Terre et M sa masse.

- $\mathbf{F}_e(t)$ est la force d'entraînement s'écrivant :

$$\mathbf{F}_e(t) = -m \left(\mathbf{a}(\mathbf{O}/\mathcal{R}) + \frac{d\boldsymbol{\Omega}}{dt} \wedge \mathbf{p}(t) + \boldsymbol{\Omega} \wedge \boldsymbol{\Omega} \wedge \mathbf{p}(t) \right) \quad (1.29)$$

où $\boldsymbol{\Omega} = \Omega \mathbf{e}_z$ représente le vecteur vitesse de rotation de la Terre et $\mathbf{a}(\mathbf{O}/\mathcal{R})$ représente l'accélération relative du point $\mathbf{O}$. Cette expression est très générique et se simplifie dans notre contexte. En effet, puisque la vitesse de rotation est constante, nous avons $\frac{d\boldsymbol{\Omega}}{dt} = \mathbf{0}$. Par ailleurs, l'accélération relative de $\mathbf{O}$ est également nulle puisque $\mathbf{O}$, l'origine du repère, ne se déplace pas. Il s'ensuit :

$$\mathbf{F}_e(t) = -m (\boldsymbol{\Omega} \wedge \boldsymbol{\Omega} \wedge \mathbf{p}(t)) \quad (1.30)$$

- $\mathbf{F}_c(t)$ correspond à la force de Coriolis. Elle admet une expression plus simple et fait uniquement intervenir la vitesse de rotation de la Terre, sa masse et la vitesse de translation du débris :

$$\mathbf{F}_c(t) = -2m \boldsymbol{\Omega} \wedge \mathbf{v}(t), \quad (1.31)$$

La combinaison des équations (1.28), (1.30) et (1.31) nous permet de réécrire la seconde loi de Newton comme une égalité entre l'accélération et une fonction non linéaire dépendant de

$\mathbf{p}(t)$, G , m et $\boldsymbol{\Omega} = \Omega \mathbf{e}_z$. Plus précisément, nous obtenons :

$$\mathbf{a}(t) = \mathbf{f}(\mathbf{p}(t), \mathbf{v}(t), G, \boldsymbol{\Omega}) \quad (1.32)$$

avec :

$$\mathbf{f}(\mathbf{p}(t), \mathbf{v}(t), G, \boldsymbol{\Omega}) = \begin{bmatrix} \Omega^2 x(t) + 2\Omega v_y(t) - \frac{GM x(t)}{(x(t)^2 + y(t)^2 + z(t)^2)^{3/2}} \\ \Omega^2 y(t) - 2\Omega v_x(t) - \frac{GM y(t)}{(x(t)^2 + y(t)^2 + z(t)^2)^{3/2}} \\ - \frac{GM z(t)}{(x(t)^2 + y(t)^2 + z(t)^2)^{3/2}} \end{bmatrix} \quad (1.33)$$

Comme les variables d'intérêt sont la position ainsi que la vitesse, il est judicieux de réécrire l'expression (1.33) sous forme vectorielle.

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}_a(\mathbf{x}(t)) \quad (1.34)$$

où : $\mathbf{x}(t) = [\mathbf{p}(t)^\top, \mathbf{v}(t)^\top]^\top$ et $\mathbf{f}_a(\mathbf{x}(t)) = [\mathbf{v}(t)^\top, \mathbf{f}(\mathbf{p}(t), \mathbf{v}(t), G, \boldsymbol{\Omega})^\top]^\top$.

Cette formule décrit donc une équation différentielle non linéaire dont la variable à déterminer $\mathbf{x}(t)$ contient le vecteur position ainsi que le vecteur vitesse. Néanmoins, le modèle établi ci-dessus ne prend pas en compte les éventuelles perturbations non modélisables analytiquement. À ce titre, nous pouvons citer l'attraction lunaire ou les vents solaires. Elles peuvent être vues comme des incertitudes sur le modèle physique et modélisable par un bruit aléatoire agissant sur l'accélération et la vitesse. Le modèle incertain s'écrit alors :

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}_a(\mathbf{x}(t)) + \mathbf{b}(t) \quad (1.35)$$

où $\mathbf{b}(t)$ est un bruit gaussien centré de matrice de covariance $\mathbf{Q}(t)$. Cette matrice de covariance représente donc l'incertitude sur le modèle.

À des fins de mise en oeuvre numérique, il est nécessaire de discrétiser cette équation différentielle. Pour y parvenir, il est d'abord possible de linéariser $\mathbf{f}_a$ à l'ordre 0, de sorte à se ramener à la résolution d'une équation différentielle linéaire. Ainsi, le modèle d'évolution discret de $\mathbf{x}(t)$ peut être déduit par échantillonnage temporel de la solution de celle-ci, selon le procédé décrit dans l'annexe C :

$$\mathbf{x}(t_k) = \mathbf{x}(t_{k-1}) + \mathbf{f}_a(\mathbf{x}(t_{k-1})) T + \mathbf{b}(t_k) T \quad (1.36)$$

où T correspond à la période d'échantillonnage et t_k au $k^{\text{ième}}$ échantillon temporel. Il est intéressant de remarquer que cette approximation à l'ordre 1 est équivalent à effectuer une résolution de l'équation différentielle par un schéma d'Euler explicite.

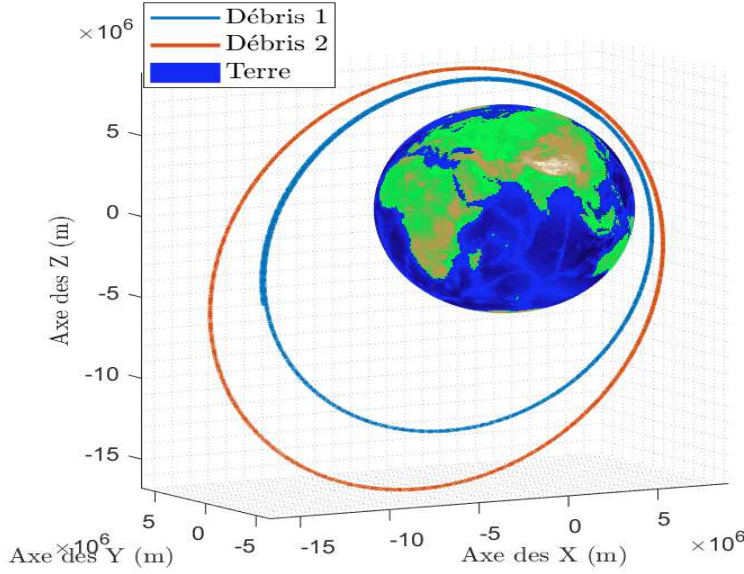


FIGURE 1.14 – Trajectoires simulées de deux débris spatiaux dans le repère ECI pour $T = 0.1s$ et de vitesses initiales $[6 \cdot 10^3, -10^3, 5 \cdot 10^3] \text{ m s}^{-1}$.

En observant la figure 1.14, nous constatons que les débris décrivent une trajectoire elliptique dans le repère ECI conformément au mouvement décrit par les lois de Kepler.

1.3.3 Simulation de la trajectoire d'un amas de débris

À présent, nous nous intéressons à simuler l'évolution dynamique d'un ensemble de débris dont le nombre est suffisamment important et qui sont suffisamment proches les uns des autres pour être considéré comme un amas.

Nous supposons un scénario dans lequel chacun des débris est issu d'une collision entre deux engins spatiaux conduisant à leur éclatement. Les débris générés se dispersent spatialement : ils sont expulsés dans des directions aléatoires de l'espace à partir du lieu de la collision.

Les débris ont des vitesses différentes qui déterminent la forme que prendra l'amas. Ce comportement peut se représenter statistiquement par un ensemble de positions et de vitesses tirées selon une distribution gaussienne de premier paramètre le lieu de la collision et une vitesse de 8 km s^{-1} respectivement. Le second paramètre, l'écart-type, représente les dispersions en vitesse et en position des différents débris. Ainsi, le nuage initial de débris aura une forme caractéristique de type ellipsoïdale.

Nous supposons que la collision a lieu en orbite basse, donc à une altitude de l'ordre de 1000 km. Comme précédemment, la période d'échantillonnage est prise à $T = 0.1s$ et la fenêtre de simulation est suffisamment longue pour observer l'évolution de l'amas au cours de l'orbite.

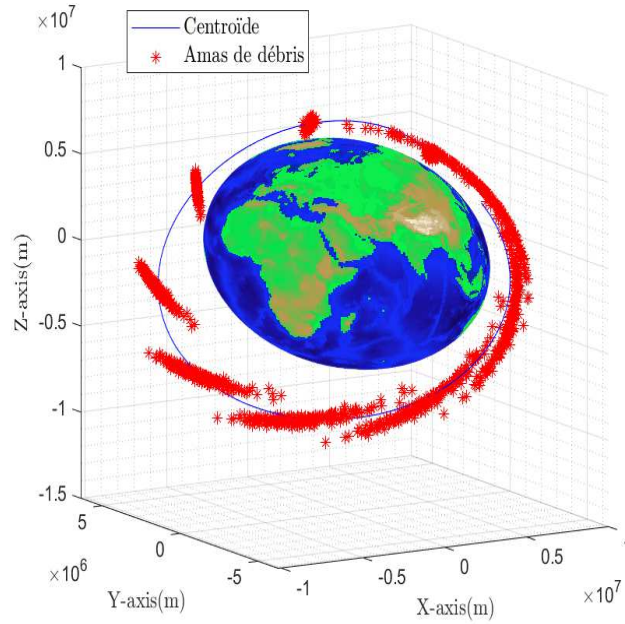


FIGURE 1.15 – Évolution des trajectoires d'un amas de débris pour $N = 500$ débris et une dispersion initiale de 50 km.

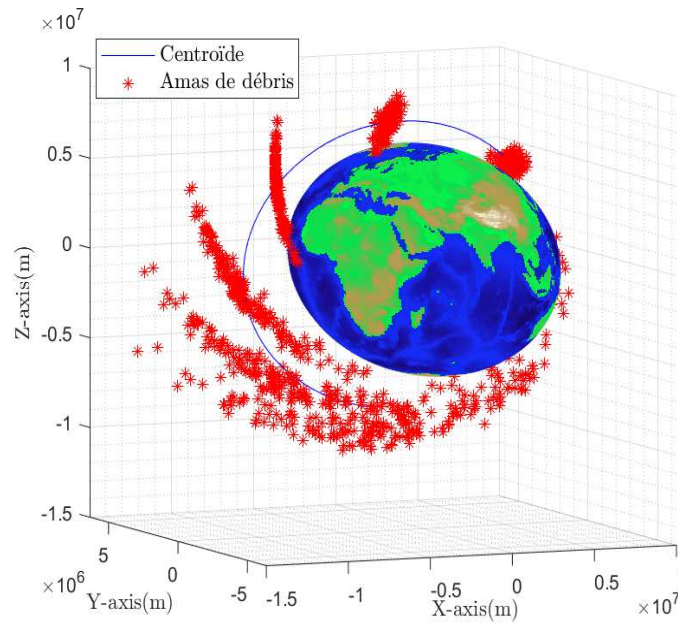


FIGURE 1.16 – Évolution des trajectoires d'un amas de débris pour $N = 500$ débris et une dispersion initiale de 500 km.

Sur les figures 1.15 et 1.16 est représentée l'évolution d'un amas de débris pour deux scénarios : le premier pour une dispersion initiale en position de l'ordre de 50 km dans chaque direction, le second en supposant celle-ci plus élevée, de l'ordre de 500 km. Néanmoins, pour chacun des deux cas, l'altitude initiale du centre de l'amas est fixée à 1000 km. Par l'effet de la force gravitationnelle, l'amas va avoir tendance à évoluer en prenant une forme spécifique. En effet, lors des premières secondes de simulation, le nuage de débris a une forme elliptique compacte et plutôt stable. Puis au bout de quelques minutes, chacun des débris vient à suivre

une trajectoire qui lui est propre et dépend de son vecteur vitesse initial. Il s’ensuit alors que le nuage elliptique se tord et s’étire pour prendre une forme courbée.

Durant les premiers instants, la trajectoire képlérienne étant presque linéaire, le nuage ne subit pas encore l’effet de courbure dû à la gravitation. À mesure que la non linéarité du modèle s’accroît, l’amas se courbe sous l’effet de la gravitation. Par ailleurs, son étendue s’accroît. Il est également intéressant de remarquer que cette forme caractéristique apparaît plus rapidement lorsque la dispersion en position de l’amas initial est plus élevée, donc lorsque l’incrément de vitesse de chaque débris est plus grande.

Ce comportement observé, déjà pointé dans différentes études [15][16], est fondamental et fondateur pour la suite de notre travail car il va nous permettre de motiver les modélisations et algorithmes mis en place. En effet, ce type de forme courbée est représentatif d’un ensemble de points contraints à évoluer sur un espace appelé *groupe de Lie*, dont nous discuterons dans le chapitre suivant.

D’un point de vue radar, lorsque l’amas de débris commence à se disperser, il forme un ensemble compact s’étendant sur plusieurs cellules de résolution. Par conséquent, il ne paraît pas judicieux de le traiter comme un ensemble d’objets mais plutôt comme une cible étendue pouvant être caractérisée par deux types de paramètres :

- les paramètres de son centre de gravité : sa position, son orientation et sa vitesse,
- un paramètre d’étendue caractérisant sa géométrie ainsi que sa dispersion spatiale.

1.4 Pistage de débris spatiaux

L’objectif de cette partie est de rappeler les fondements des principales approches de la littérature permettant de réaliser le pistage d’un ou d’un ensemble de débris spatiaux. Dans un premier temps, un contexte de cible ponctuelle est considéré, lorsque le système radar observe un unique débris. Dans un second temps, le formalisme des méthodes de pistage d’objets étendus est présenté, lorsque le radar génère plusieurs mesures associées à un seul objet, ce qui sera le cas dans notre application.

1.4.1 Pistage d’une cible ponctuelle

1.4.1.1 Présentation du filtre bayésien

Une cible ponctuelle est un objet qui génère à chaque instant une unique mesure capteur. Ainsi, elle est détectée dans une unique cellule de résolution capteur. L’objectif du pistage est alors d’estimer dynamiquement ses paramètres cinématiques (position, vitesse, accélération,...) à l’aide d’un algorithme dédié à ce type de tâche.

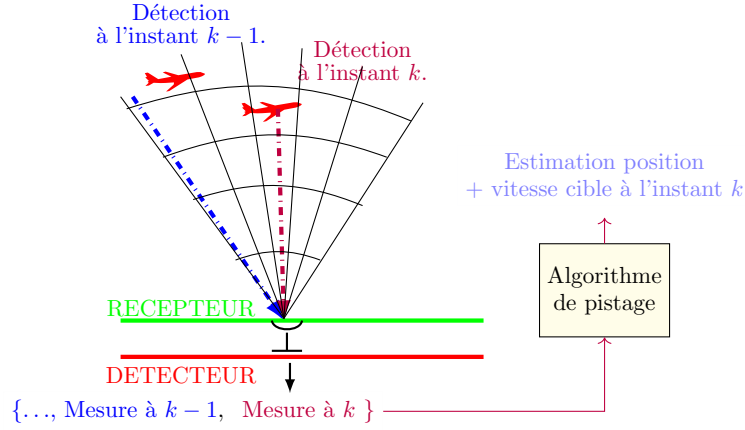


FIGURE 1.17 – Illustration du pistage d'une cible ponctuelle.

Les paramètres de la cible peuvent être décrits à l'instant k par un vecteur dit d'état $\mathbf{x}_k \in \mathbb{R}^p$ contenant sa position, sa vitesse et potentiellement son accélération et les dérivées d'ordre supérieur. Connaissant un modèle d'évolution dynamique incertain, appelé modèle *a priori*, décrit par une densité de probabilité de transition $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, ainsi que le modèle imparfait d'observation du capteur prenant la forme d'une vraisemblance $p(\mathbf{z}_k|\mathbf{x}_k)$, avec $\mathbf{z}_k \in \mathbb{R}^m$ un ensemble de mesures, il est possible d'estimer récursivement l'état de la cible par le processus de *filtrage bayésien*.

Il consiste à déterminer la distribution *a posteriori* des variables inconnues sachant les observations de l'instant initial à l'instant courant notée $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ où $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ [17]. Une estimation *a posteriori* des paramètres d'intérêt peut alors en être déduite, par exemple en prenant la moyenne de cette loi. Pour prévenir une augmentation de la charge calculatoire au cours du temps, le calcul est réalisé séquentiellement de sorte à obtenir une expression récursive de la loi courante à partir de la loi *a posteriori* précédente.

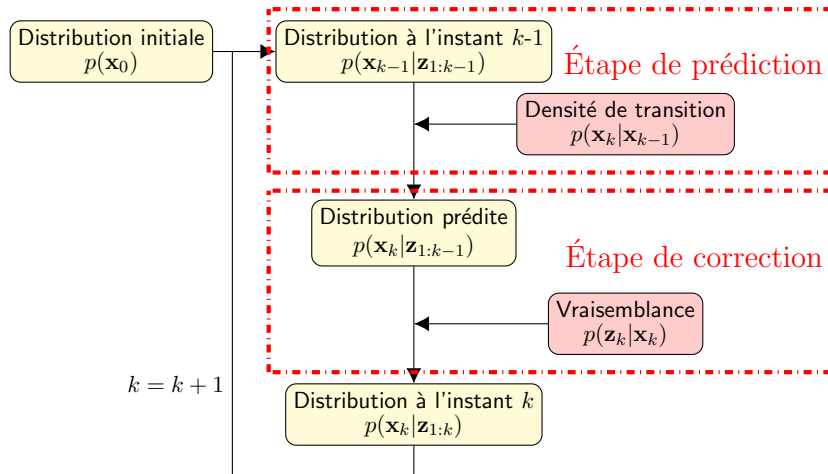


FIGURE 1.18 – Principe du filtre bayésien.

La mise à jour de la densité *a posteriori* est effectuée en deux étapes dont l'enchaînement est représenté en figure 1.18.

- La première étape, dite de prédiction, détermine la distribution prédite $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ grâce à la formule de Chapman-Kolgomorov :

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int_{\mathbf{x}_{k-1} \in \mathbb{R}^p} p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}, \quad (1.37)$$

- la seconde, appelée étape de correction, met à jour la distribution prédite en utilisant l'information apportée par la mesure au travers de la règle de Bayes :

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \quad (1.38)$$

Par souci de concision, nous utilisons par la suite les notations $p_{k-1|k-1}(\mathbf{x}_{k-1}) = p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$, $p_{k|k-1}(\mathbf{x}_k) = p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ et $p_{k|k}(\mathbf{x}_k) = p(\mathbf{x}_k|\mathbf{z}_{1:k})$.

1.4.1.2 Filtre bayésien analytique : le filtre de Kalman

Généralement, les expressions du filtre bayésien ne peuvent pas être calculées analytiquement, sauf si les modèles d'état et d'observations présentent certaines propriétés. Si le modèle d'évolution est linéaire gaussien, il vient que :

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{v}_k \quad (1.39)$$

où $\mathbf{F}_k$ se nomme la matrice de transition représentant l'évolution entre $\mathbf{x}_{k-1}$ et $\mathbf{x}_k$, et $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ est un bruit blanc supposé gaussien de matrice de covariance $\mathbf{Q}_k$. Si la même hypothèse de linéarité et de gaussianité est supposée pour le modèle d'observation, il peut s'écrire sous la forme suivante :

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k \quad (1.40)$$

où $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ est un bruit blanc gaussien de matrice de covariance $\mathbf{R}_k$. Selon ces deux modèles, la vraisemblance ainsi que la densité de transition sont gaussiennes : $p(\mathbf{z}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k)$ et $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_k \mathbf{x}_{k-1}, \mathbf{Q}_k)$. De plus, les bruits d'état et de mesure sont supposés indépendants.

Dans ce cas, les équations du filtrage bayésien (1.37) et (1.38) permettent d'en déduire que la distribution *a posteriori* est également gaussienne de moyenne $\hat{\mathbf{x}}_{k|k}$ et de matrice de covariance $\mathbf{P}_{k|k}$. Ainsi, déterminer cette distribution est équivalent à déterminer ses moments du premier et du second ordre. Les étapes de prédiction et de correction fournissent directement $\{\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\}$ issues de $p_{k|k-1}(\mathbf{x}_k)$ ainsi que $\{\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}\}$ issues de $p_{k|k}(\mathbf{x}_k)$. Ces équations définissent une itération du filtre de Kalman [18].

À l'étape de prédiction, nous avons :

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} \\ \mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k \end{cases} \quad (1.41)$$

$$\quad (1.42)$$

où $\{\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}\}$ sont les paramètres de $p_{k-1|k-1}(\mathbf{x}_{k-1})$.

L'étape de correction fournit :

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}) \\ \mathbf{P}_{k|k} = (\mathbf{I}_k - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \end{cases} \quad (1.43)$$

$$\quad (1.44)$$

où $\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{S}_k^{-1}$ définit le gain du filtre et $\mathbf{S}_k = \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top$ est la matrice de covariance de l'innovation définie par $\boldsymbol{\eta}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$. Cette dernière contient l'information nouvelle apportée par la mesure courante $\mathbf{z}_k$ par rapport à la mesure prédite $\mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$. Le gain représente la confiance qu'on peut lui accorder.

1.4.1.3 Filtres bayésien approchés

Les hypothèses du filtre de Kalman supposent que les modèles sont linéaires. En pratique, ce n'est pas toujours le cas, puisque les trajectoires d'objet peuvent suivre des dynamiques non linéaires. De plus, les modèles physiques des capteurs sont généralement aussi plus complexes comme établi dans le cas d'un système radar. Par conséquent, leurs expressions mathématiques sont généralement plus génériques et peuvent prendre la forme suivante :

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \quad (1.45)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k \quad \mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k) \quad (1.46)$$

où $\mathbf{f}_k : \mathbb{R}^p \rightarrow \mathbb{R}^p$ et $\mathbf{h}_k : \mathbb{R}^p \rightarrow \mathbb{R}^m$ sont deux fonctions non linéaires.

Les méthodes d'approximation analytiques du filtre bayésien ont pour objectif d'approcher la distribution *a posteriori* par une distribution paramétrique, généralement une distribution gaussienne. Il en existe plusieurs dans la littérature :

- Les méthodes d'approximation du premier ordre sont fondées sur une linéarisation des modèles par un développement de Taylor-Young, et aboutissent à une variante du filtre de Kalman : le *filtre de Kalman étendu (FKE)*.
- Des méthodes d'approximations du second ordre existent. En particulier, le *central difference Kalman Filter (CDKF)* est fondé sur une interpolation polynomiale de Sterling au second ordre des modèles [19].
- Des méthodes numériques peuvent également être appliquées pour en déduire par simulation une approximation de la distribution *a posteriori*. Nous pouvons citer en particulier les méthodes *grid-based* qui sont de type déterministes, les méthodes de Monte-Carlo séquentielles connues sous le nom de *filtre particulière* et basées sur un échantillonnage stochastique [17][20].
- Des méthodes fondées sur la transformée sans parfum existent et définissent le *filtre de Kalman sans parfum (FKSP)* [21]. Il utilise un petit nombre d'échantillons choisis de façon déterministe afin d'approcher les lois d'intérêts par des distributions gaussiennes.
- Dans le cas où la non-linéarité des modèles est conservée, il est possible d'obtenir une approximation gaussienne de la distribution *a posteriori* grâce à une autre variante du filtre de Kalman : le *filtre de Kalman étendu itéré (FKEI)* [22].

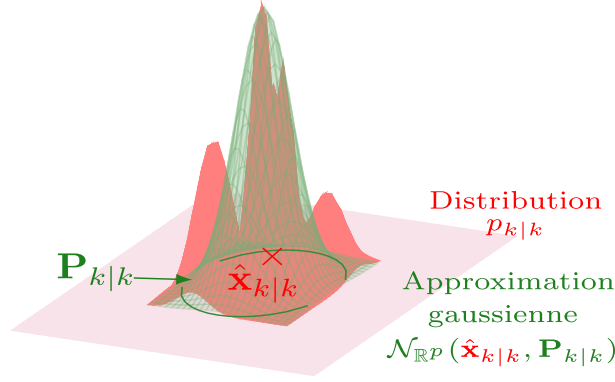


FIGURE 1.19 – Approximation gaussienne.

* Filtre de Kalman étendu :

Le *FKE* approche la loi de transition ainsi que la vraisemblance par deux distributions gaussiennes centrées respectivement autour de points de linéarisations des fonctions $\mathbf{f}_k$ et $\mathbf{h}_k$. De plus, la distribution *a posteriori* à l'instant $k-1$ est approchée par une distribution gaussienne. Les points de linéarisations généralement considérés sont le paramètre $\hat{\mathbf{x}}_{k-1|k-1}$ de la distribution $p_{k-1|k-1}$, pour $\mathbf{f}_k$ et le paramètre $\hat{\mathbf{x}}_{k|k-1}(\mathbf{x}_{k-1})$ de la distribution $p_{k|k-1}(\mathbf{x}_k)$ pour $\mathbf{h}_k$. Les matrices de covariance sont alors obtenues à partir des matrices jacobiniennes issues de ces deux approximations. Plus formellement, les modèles (1.45) et (1.46) se linéarisent sous les formes suivantes :

$$\mathbf{x}_k \simeq \mathbf{f}_k(\hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{J}_{\mathbf{f}_k}(\mathbf{x}_k - \hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k \quad (1.47)$$

$$\mathbf{z}_k \simeq \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{J}_{\mathbf{h}_k}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{n}_k \quad (1.48)$$

où $\mathbf{J}_{\mathbf{f}_k}$ et $\mathbf{J}_{\mathbf{h}_k}$ correspondent aux matrices jacobiniennes respectivement de $\mathbf{f}_k$ calculée en $\hat{\mathbf{x}}_{k-1|k-1}$ et de $\mathbf{h}_k$ en $\hat{\mathbf{x}}_{k|k-1}$. $p(\mathbf{z}_k|\mathbf{x}_k)$ et $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ s'approchent ainsi respectivement par $\mathcal{N}(\mathbf{z}_k; \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{J}_{\mathbf{h}_k}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}), \mathbf{R}_k)$ et $\mathcal{N}(\mathbf{x}_k; \mathbf{f}_k(\hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{J}_{\mathbf{f}_k}(\mathbf{x}_k - \hat{\mathbf{x}}_{k-1|k-1}), \mathbf{Q}_k)$. Ces deux approximations fournissent en sortie de l'étape de prédiction et de correction deux distributions gaussiennes. Ainsi, à l'étape de prédiction, nous avons :

$$p_{k|k-1}(\mathbf{x}_k) \simeq \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \quad (1.49)$$

avec :

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} = \mathbf{f}_k(\hat{\mathbf{x}}_{k-1|k-1}) \\ \mathbf{P}_{k|k-1} = \mathbf{J}_{\mathbf{f}_k} \mathbf{P}_{k-1|k-1} \mathbf{J}_{\mathbf{f}_k}^\top + \mathbf{Q}_k \end{cases} \quad (1.50)$$

$$\quad (1.51)$$

L'étape de correction aboutit à :

$$p_{k|k}(\mathbf{x}_k) \simeq \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) \quad (1.52)$$

avec :

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1})) \\ \mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}_k}) \mathbf{P}_{k|k-1} \end{cases} \quad (1.53)$$

$$\quad (1.54)$$

où :

$$\begin{cases} \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{J}_{\mathbf{h}_k}^\top \mathbf{S}_k^{-1} \\ \mathbf{S}_k = \mathbf{J}_{\mathbf{h}_k} \mathbf{P}_{k|k-1} \mathbf{J}_{\mathbf{h}_k}^\top + \mathbf{R}_k \end{cases} \quad (1.55)$$

$$(1.56)$$

Nous remarquons que les équations du *FKE* sont similaires à celle du *FK* puisque les mêmes hypothèses gaussiennes sont utilisées pour le calcul de $p_{k|k}(\mathbf{x}_k)$ et $p_{k|k-1}(\mathbf{x}_k)$.

* Filtre de Kalman sans parfum :

Tout comme le *FKE*, le *FKSP* est un filtre approchant la distribution *a posteriori* par une distribution gaussienne. Néanmoins, dans ce cas, les statistiques de cette distribution sont obtenues en utilisant une transformation non linéaire, nommée *Unscented Transform* (UT) [23]. Il est alors prouvé que cette méthode revient à approcher à l'ordre deux les modèles en utilisant des différences finies afin d'approcher les jacobiniennes et les dérivées d'ordre supérieur [24]. L'idée est d'utiliser des points d'intérêts, nommés *sigma points*, représentatifs de la distribution *a posteriori* précédente, et de les propager à travers les fonctions non linéaires de l'état et de la mesure pour en déduire une approximation de la moyenne et de la covariance des distributions de prédiction et de correction. Les moments d'ordre 1 et 2 de celles-ci sont obtenus en moyennant ces sigma points pondérés par des poids calculés de sorte à minimiser l'erreur d'estimation [21].

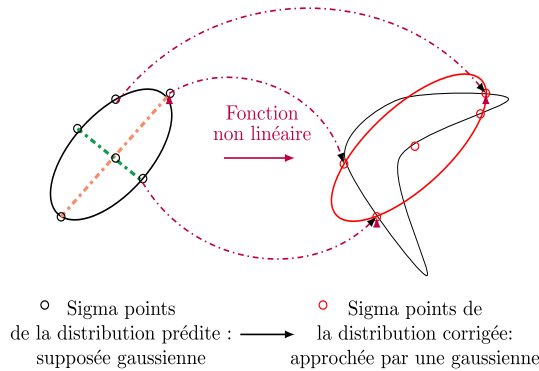


FIGURE 1.20 – Principe de la transformée sans parfum.

* Filtre particulaire :

Communément appelé *filtre de Monte-Carlo*, ou *méthode de Monte-Carlo séquentielle*, ce filtre réalise une approximation de la distribution *a posteriori* à l'instant courant par un ensemble pondéré d'échantillons, appelés particules. Ils sont générés selon une loi de proposition et les poids qui leur sont attribués permettent de corriger l'écart entre cette dernière et la loi cible. Ce procédé est nommé dans la littérature échantillonnage d'importance [25]. En appliquant séquentiellement cette approche, les équations du filtre bayésien peuvent s'écrire de manière approchée. Toutes les intégrales définies peuvent alors être remplacées par des sommes effectuées sur les vecteurs d'état des particules.

- L'étape de prédiction peut être vue comme celle permettant d'échantillonner les particules.
- L'étape de correction permet de mettre à jour les poids par introduction de l'information de vraisemblance.

L'estimation de l'état est finalement obtenu comme une moyenne pondérée des particules avec chacun des poids. La matrice de covariance de l'erreur d'estimation est obtenue par le même procédé. Au bout de quelques itérations de l'algorithme, il est montré que la variance des poids devient très grande. Par conséquent, une unique particule sera associée à un poids égal à 1 et toutes les autres seront assorties d'un poids nul. Pour pallier ce problème dit de dégénérescence, une étape de ré-échantillonnage des particules est effectuée de sorte à conserver uniquement les particules de poids non négligeables, et à affecter à chacune un poids équiprobable.

Nous n'entrerons pas en détails dans les explications théoriques puisque les méthodes de filtrage par simulation de Monte-Carlo n'interviendront pas explicitement dans nos travaux. Pour des explications plus précises et plus détaillées, le lecteur peut se référer à [20][26].

* Filtre de Kalman étendu itéré :

Le *FKEI* est une amélioration du *FKE* permettant d'obtenir une approximation de $p_{k|k-1}(\mathbf{x}_k)$ et de $p_{k|k}(\mathbf{x}_k)$ plus précise [22][27]. Il détermine une approximation gaussienne directe de $p_{k|k-1}(\mathbf{x}_k)$ et de $p_{k|k}(\mathbf{x}_k)$ dans laquelle la non-linéarité des modèles est préservée. Les modes de ces deux distributions sont alors utilisés comme moyennes respectives des deux approximations. Puis, deux linéarisations au premier ordre autour de ces modes sont alors réalisées pour en déduire une estimation des deux matrices de covariance par une approximation de Gauss-Laplace [28]. De par les non-linéarités des modèles, trouver le mode de ces distributions n'est pas réalisable analytiquement. Il doit être déterminé numériquement par minimisation d'un critère des moindres carrés non-linéaire. Par conséquent, le problème d'estimation devient un problème d'optimisation quadratique non-linéaire et peut être résolu par l'algorithme de Gauss-Newton [29]. Les deux étapes du filtre se réécrivent alors en deux étapes d'optimisation que nous proposons de détailler.

À l'étape de prédiction, le critère à optimiser est fondé sur la distribution jointe du couple $\{\mathbf{x}_k, \mathbf{x}_{k-1}\}$. D'après la règle de Bayes :

$$p(\mathbf{x}_k, \mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}) \underbrace{p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})}_{\simeq \mathcal{N}(\mathbf{x}_{k-1}; \tilde{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})} \quad (1.57)$$

L'approximation gaussienne de $p_{k|k-1}$ est alors obtenue en supposant que la loi jointe est gaussienne, puis en la marginalisant selon $\mathbf{x}_{k-1}$.

Sous cette hypothèse, déterminer le mode de la loi jointe est équivalent à déterminer sa moyenne. La moyenne de la loi prédite $\mathbf{x}_{k|k-1}^*$ de $p_{k|k-1}$ peut alors être obtenue à partir de la moyenne $\{\mathbf{x}_{k|k-1}^*, \mathbf{x}_{k-1|k-1}^*\}$ de la distribution jointe. Quant à elle, la matrice de covariance peut être obtenue à partir de la matrice de covariance $\mathbf{P}^a$ du couple $\{\mathbf{x}_k, \mathbf{x}_{k-1}\}$. En effet, la matrice de covariance de $\mathbf{x}_k$ est la sous-matrice de $\mathbf{P}^a$ correspondant à son bloc diagonal haut. Dans un premier temps, l'objectif est donc de résoudre :

$$\{\mathbf{x}_{k|k-1}^*, \mathbf{x}_{k-1|k-1}^*\} = \underset{\tilde{\mathbf{x}}_k \in \mathbb{R}^p, \tilde{\mathbf{x}}_{k-1} \in \mathbb{R}^p}{\operatorname{argmax}} p(\tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_{k-1} | \mathbf{z}_{1:k-1}) \quad (1.58)$$

D'après l'équation (1.57), cette distribution peut se réécrire :

$$p(\mathbf{x}_k, \mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) \propto \exp\left(-\frac{1}{2} \|\mathbf{x}_k - \mathbf{f}_k(\mathbf{x}_{k-1})\|_{\mathbf{Q}_k}^2\right) \exp\left(-\frac{1}{2} \|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}\|_{\mathbf{P}_{k-1|k-1}}^2\right) \quad (1.59)$$

Par passage au logarithme, la fonction exponentielle s'élimine et nous obtenons :

$$\log p(\mathbf{x}_k, \mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) = C_1 + \frac{1}{2} \|\mathbf{x}_k - \mathbf{f}_k(\mathbf{x}_{k-1})\|_{\mathbf{Q}_k}^2 + \frac{1}{2} \|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}\|_{\mathbf{P}_{k-1|k-1}}^2 \quad (1.60)$$

où $C_1 \in \mathbb{R}$ est une constante. Par conséquent, le problème de maximisation défini par l'équation (1.58) peut se réécrire comme un problème de minimisation :

$$\{\mathbf{x}_{k|k-1}^*, \mathbf{x}_{k-1|k-1}^*\} = \underset{\tilde{\mathbf{x}}_k \in \mathbb{R}^p, \tilde{\mathbf{x}}_{k-1} \in \mathbb{R}^p}{\operatorname{argmin}} -\log p(\tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_{k-1} | \mathbf{z}_{1:k-1}) \quad (1.61)$$

$$= \underset{\tilde{\mathbf{x}}_k \in \mathbb{R}^p, \tilde{\mathbf{x}}_{k-1} \in \mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2} \left(\|\tilde{\mathbf{x}}_k - \mathbf{f}_k(\tilde{\mathbf{x}}_{k-1})\|_{\mathbf{Q}_k}^2 + \|\tilde{\mathbf{x}}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}\|_{\mathbf{P}_{k-1|k-1}}^2 \right) \quad (1.62)$$

Les solutions à ce problème d'optimisation sont triviales et s'écrivent :

$$\begin{cases} \mathbf{x}_{k-1|k-1}^* = \hat{\mathbf{x}}_{k-1|k-1} \\ \mathbf{x}_{k|k-1}^* = \mathbf{f}_k(\hat{\mathbf{x}}_{k-1|k-1}) = \hat{\mathbf{x}}_{k|k-1} \end{cases} \quad (1.63)$$

$$\quad (1.64)$$

Pour trouver la matrice de covariance de la loi jointe, une approximation de Gauss-Laplace est utilisée. Elle repose sur un développement au premier ordre de la quantité à l'intérieur du terme quadratique. Plus précisément, en notant :

$$J(\mathbf{x}_k, \mathbf{x}_{k-1}) = -\log p(\mathbf{x}_k, \mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) = \|\phi(\mathbf{x}_k, \mathbf{x}_{k-1})\|^2, \quad (1.65)$$

un développement à l'ordre 1 de la fonction ϕ aux points $\mathbf{x}_{k-1|k-1}^*$ et $\mathbf{x}_{k|k-1}^*$ nous fournit que :

$$p(\mathbf{x}_k, \mathbf{x}_{k-1}) \simeq \exp\left(-\frac{1}{2} \|\phi(\mathbf{x}_k, \mathbf{x}_{k-1}) + \mathbf{J}_\phi \boldsymbol{\delta}_x\|_{\mathbf{B}_k}^2\right) \quad (1.66)$$

$$\propto \exp\left(-\frac{1}{2} \boldsymbol{\delta}_x^\top \mathbf{J}_\phi^\top \mathbf{B}_k^{-1} \mathbf{J}_\phi \boldsymbol{\delta}_x\right) \quad (1.67)$$

avec $\mathbf{J}_\phi$ la matrice jacobienne de ϕ calculée en $\mathbf{x}_{k|k-1}^*$ et

$$\left\{ \boldsymbol{\delta}_x = [\mathbf{x}_k - \mathbf{x}_{k|k-1}^*, \mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^*] \right\}, \quad (1.68)$$

$$\left\{ \mathbf{B}_k = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k-1|k-1} \end{bmatrix} \right\}. \quad (1.69)$$

Par conséquent, la matrice de covariance de la loi jointe peut être approchée par $\mathbf{J}_\phi^\top \mathbf{B}_k^{-1} \mathbf{J}_\phi$. En récupérant le bloc diagonal haut de cette matrice, nous obtenons la covariance associée à

$p_{k|k-1}(\mathbf{x}_k)$ à partir de quelques manipulations matricielles élémentaires :

$$\mathbf{P}_{k|k-1} = \mathbf{J}_{\mathbf{f}_k} \mathbf{P}_{k-1|k-1} \mathbf{J}_{\mathbf{f}_k}^\top + \mathbf{Q}_k \quad (1.70)$$

où $\mathbf{J}_{\mathbf{f}_k}$ est la matrice jacobienne de $\mathbf{f}_k$ calculée en $\mathbf{x}_{k|k-1}^*$.

À l'étape de correction, les hypothèses de bruit gaussien combinées avec l'équation de mise à jour du filtre bayésien optimal (1.38) nous permettent d'écrire le logarithme de $p_{k|k}(\mathbf{x}_k)$:

$$\log p_{k|k}(\mathbf{x}_k) = C_2 + \frac{1}{2} \|\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k)\|_{\mathbf{R}_k}^2 + \frac{1}{2} \|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}}^2 \quad (1.71)$$

où $C_2 \in \mathbb{R}$ est une constante. Le mode est donc obtenu en résolvant le problème d'optimisation suivant :

$$\mathbf{x}_{k|k}^* = \underset{\tilde{\mathbf{x}}_k \in \mathbb{R}^p}{\operatorname{argmin}} - \log p_{k|k}(\tilde{\mathbf{x}}_k) \quad (1.72)$$

Contrairement à l'étape de prédiction, la solution n'est pas triviale et peut être obtenue en appliquant un algorithme d'optimisation. Comme le critère (1.71) est de type moindres carrés non linéaires, un algorithme de Gauss-Newton est généralement mis en oeuvre pour sa minimisation. Cet algorithme converge vers un point critique, et sous condition de convexité de la fonction, il correspond à son minimiseur. À chaque itération, nous pouvons montrer que l'itéré courant s'écrit [22] :

$$\mathbf{x}^{(l+1)} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^{(l)} \left(\mathbf{z}_k - (\mathbf{h}_k(\mathbf{x}^{(l)})) + \mathbf{J}_{\mathbf{h}_k}^{(l)} (\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}^{(l)}) \right) \quad (1.73)$$

où $\mathbf{J}_{\mathbf{h}_k}^{(l)}$ correspond à la matrice jacobienne de $\mathbf{h}_k$ calculée en $\mathbf{x}^{(l)}$, et $\mathbf{K}_k^{(l)}$ correspond à "l'équivalent" du gain de Kalman calculé à l'itération l :

$$\mathbf{K}_k^{(l)} = \mathbf{P}_{k|k-1} (\mathbf{J}_{\mathbf{h}_k}^{(l)})^\top (\mathbf{J}_{\mathbf{h}_k}^{(l)} \mathbf{P}_{k-1|k-1} (\mathbf{J}_{\mathbf{h}_k}^{(l)})^\top + \mathbf{R}_k)^{-1} \quad (1.74)$$

La matrice de covariance est alors obtenue de la même manière que précédemment : par une approximation de Gauss-Laplace autour du mode de la distribution corrigée.

1.4.1.4 Remarques sur le pistage d'une cible ponctuelle

Le pistage d'une cible ponctuelle est adapté à un scénario dans lequel une seule cible est observée par le système radar. En pratique, les méthodes présentées ne seront pas pertinentes dans notre application puisque nous nous focalisons sur un amas de débris qui peut être vu comme un ensemble de plusieurs cibles. Néanmoins, le formalisme des méthodes de filtrage et d'optimisation basées sur le *FKEI* nous servira de référence pour proposer des algorithmes adaptés à notre contexte.

1.4.2 Pistage d'une cible étendue

Comme brièvement décrit dans la section (1.2.2.2), une cible étendue est une cible donnant lieu à plusieurs mesures capteur à chaque instant. Dans le cas d'un système radar, elle génère des échos dans plusieurs cases de résolution.

Ainsi, lorsqu'une détection est réalisée sur plusieurs cases radars, cela signifie que plusieurs mesures provenant de différentes régions de l'objet vont être traitées. Les points de l'objet dont elles sont issues sont nommés réflecteurs : leur nombre et leur position dépendent en particulier de la surface équivalente radar des différentes parties de l'objet.

Dans ce contexte, pister uniquement le centroïde d'une cible n'est pas suffisant pour la décrire : un second jeu de paramètres doit être défini afin de caractériser l'étendue et la forme de l'objet.

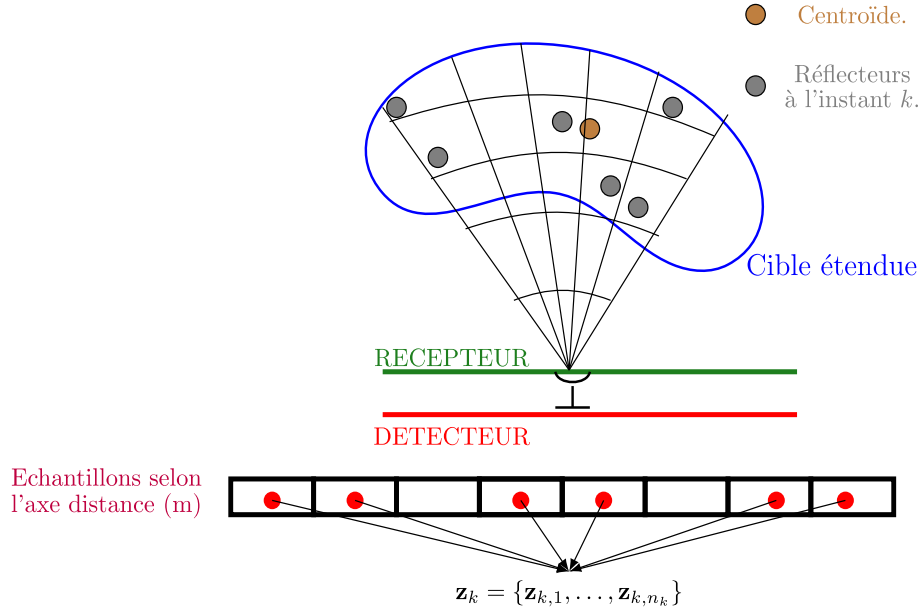


FIGURE 1.21 – Représentation simplifiée de la génération de mesures issues d'une cible étendue pour un système radar.

Ces paramètres peuvent caractériser son contour, dans un cas bidimensionnel, ou son volume dans un cas tridimensionnel. Si l'objet est supposé présenter une forme géométrique particulière alors ils peuvent donner ses dimensions.

Les méthode de pistage d'une cible étendue se divisent principalement en deux classes d'algorithmes [30] :

- la première s'appuie sur une représentation de l'objet sous forme d'une matrice aléatoire. Sous l'hypothèse qu'il est ellipsoïdal, il peut en effet être caractérisé à la fois par la position de son centroïde et par une matrice semi-définie positive représentant conjointement son orientation et son étendue [30][31][32] .
- La seconde classe modélise l'objet également par la position de son centroïde, mais son étendue est décrite de façon plus générique comme une forme de type convexe-étoilée [30][33][34] afin de traiter des cibles plus variées géométriquement. Le paramètre d'étendue est alors représenté par une fonction radiale correspondant à la distance entre le centroïde de l'objet et les points de son contour.

1.4.2.1 Modélisation par des matrices aléatoires

Cette approche est fondée sur une représentation de l'objet étendu par deux paramètres que sont le vecteur d'état du centroïde ainsi qu'une matrice aléatoire caractérisant son étendue [31][32]. Les réflecteurs obtenus peuvent alors se modéliser comme des points dispersés selon

une distribution gaussienne de matrice de covariance cette matrice d'étendue. Le support de cette loi caractérise ainsi la forme de l'objet par une ellipse.

Sous certaines hypothèses, les auteurs de [31] développent une variante du filtre de Kalman approchant la distribution jointe *a posteriori* de l'état du centroïde et de la matrice d'étendue par une distribution *Gaussienne-inverse Wishart* en mettant à profit la structure conjuguée de ce type de distribution. Les estimateurs de ces deux paramètres sont obtenus à partir de la moyenne jointe de cette distribution.

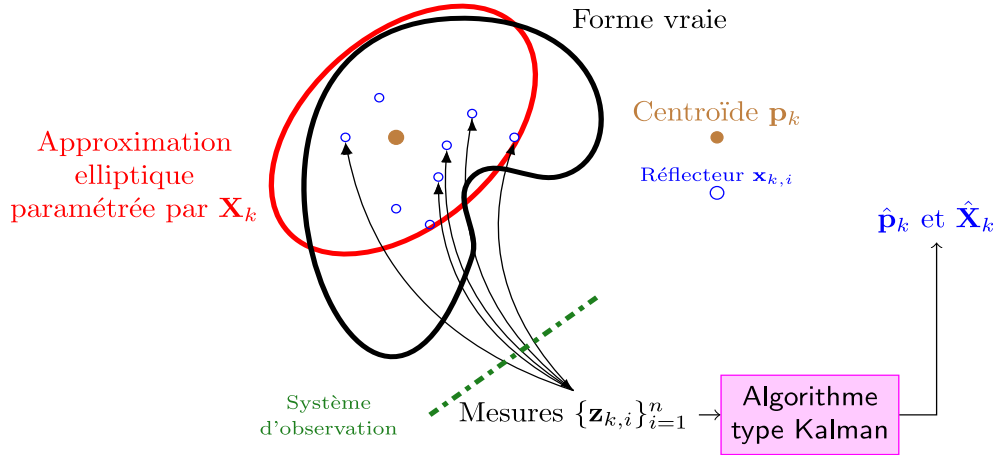


FIGURE 1.22 – Illustration de l'approche *matrice aléatoire*.

À chaque instant k , le centroïde est décrit par le vecteur d'état $\mathbf{p}_k \in \mathbb{R}^d$ contenant ses paramètres dynamiques, et l'étendue par une matrice symétrique définie positive $\mathbf{X}_k \in \mathbb{R}^{p \times p}$. Nous notons $\mathcal{Z}^k$ l'ensemble des réflecteurs fournis par le capteur de l'instant initial à l'instant k et $\mathbf{z}_k = \{\mathbf{z}_{k,i}\}_{i=1}^{n_k}$ l'ensemble des n_k réflecteurs à l'instant k .

De manière similaire au filtre de Kalman, l'objectif de l'approche est de déterminer une expression analytique de la distribution *a posteriori* $p(\mathbf{p}_k, \mathbf{X}_k | \mathcal{Z}^k)$ s'écrivant :

$$p(\mathbf{p}_k, \mathbf{X}_k | \mathcal{Z}^k) = p(\mathbf{p}_k | \mathbf{X}_k, \mathcal{Z}^k) p(\mathbf{X}_k | \mathcal{Z}^k) \quad (1.75)$$

Cette expression est fondée sur :

→ une estimation réursive gaussienne de $p(\mathbf{p}_k | \mathbf{X}_k, \mathcal{Z}^k)$:

$$p(\mathbf{p}_k | \mathbf{X}_k, \mathcal{Z}^k) = \mathcal{N}(\mathbf{p}_k; \hat{\mathbf{p}}_k, \mathbf{P}_{k|k} \odot \mathbf{X}_k) \quad (1.76)$$

où $\odot$ représente le produit de Kronecker [31] et $\mathbf{P}_{k|k}$ la matrice de covariance associée à une des composantes du repère de référence où est étudié le mouvement.

→ une approximation *inverse Wishart* de $p(\mathbf{X}_k | \mathcal{Z}^k)$:

$$p(\mathbf{X}_k | \mathcal{Z}^k) \simeq \mathcal{IW}(\mathbf{X}_k; \nu_{k|k}; \hat{\mathbf{X}}_{k|k}) \quad (1.77)$$

où $\nu_{k|k}$ et $\hat{\mathbf{X}}_{k|k}$ correspondent aux deux paramètres de la loi caractérisant l'étendue de l'objet.

La méthode est classiquement divisée en deux étapes de prédiction et de correction dont nous proposons de détailler brièvement les principes.

• Étape de prédiction :

L'étape de prédiction consiste à déterminer la distribution prédite du couple $\{\mathbf{p}_k, \mathbf{X}_k\}$ grâce à son modèle d'évolution défini par les distributions $p(\mathbf{X}_k|\mathbf{X}_{k-1})$ et $p(\mathbf{p}_k|\mathbf{p}_{k-1}, \mathbf{X}_k)$.

S'il est supposé que l'évolution de $\mathbf{X}_k$ n'a pas d'impact sur la loi de $\mathbf{p}_{k-1}$, alors $p(\mathbf{p}_{k-1}|\mathbf{X}_k, \mathcal{Z}^{k-1}) \simeq p(\mathbf{p}_{k-1}|\mathbf{X}_{k-1}, \mathcal{Z}^{k-1})$. Il est possible de montrer que les densités prédites s'écrivent sous la forme suivante :

$$\left\{ \begin{array}{l} p(\mathbf{p}_k|\mathbf{X}_k, \mathcal{Z}^{k-1}) = \int_{\mathbf{p}_{k-1} \in \mathbb{R}^d} p(\mathbf{p}_k|\mathbf{X}_k, \mathbf{p}_{k-1}) \underbrace{p(\mathbf{p}_{k-1}|\mathbf{X}_{k-1}, \mathcal{Z}^{k-1})}_{=\mathcal{N}(\mathbf{p}_{k-1}; \hat{\mathbf{p}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})} d_{\mathbb{R}^d}(\mathbf{p}_{k-1}) \end{array} \right. \quad (1.78)$$

$$\left\{ \begin{array}{l} p(\mathbf{X}_k|\mathcal{Z}^{k-1}) = \int_{\mathbf{X}_{k-1} \in \mathbb{R}^{p \times p}} p(\mathbf{X}_k|\mathbf{X}_{k-1}) \underbrace{p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1})}_{\simeq \mathcal{IW}(\mathbf{X}_k; \nu_{k-1|k-1}, \hat{\mathbf{X}}_{k-1|k-1})} d_{\mathbb{R}^{p \times p}}(\mathbf{X}_{k-1}) \end{array} \right. \quad (1.79)$$

Le vecteur d'état du centroïde est supposé suivre un modèle linéaire gaussien et l'évolution de l'étendue est modélisée par une loi de Wishart sous les formes suivantes :

$$\left\{ \begin{array}{l} \mathbf{p}_k = (\mathbf{F}_k \odot \mathbf{I}_p) \mathbf{p}_{k-1} + \mathbf{v}_k \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k \odot \mathbf{X}_k) \end{array} \right. \quad (1.80)$$

$$\left\{ \begin{array}{l} \mathbf{X}_k \sim \mathcal{W}\left(\delta_{k|k-1}, \frac{\mathbf{X}_{k-1}}{\delta_{k|k-1}}\right) \end{array} \right. \quad (1.81)$$

où $\mathcal{W}\left(\delta_{k|k-1}, \frac{\mathbf{X}_{k-1}}{\delta_{k|k-1}}\right)$ est la loi de Wishart de paramètre $\delta_{k|k-1}$ et $\frac{\mathbf{X}_{k-1}}{\delta_{k|k-1}}$. $\mathbf{v}_k$ est un bruit blanc et $\mathbf{Q}_k$ représente la matrice de covariance des composantes du bruit selon l'un des axes du repère de référence pour le mouvement.

L'étape de prédiction préservant le caractère Inverse Wishart-Gaussien de la loi jointe, les équations (1.78) et (1.79), impliquent alors que les distributions prédites s'écrivent sous la forme :

$$\left\{ \begin{array}{l} p(\mathbf{p}_k|\mathbf{X}_k, \mathcal{Z}^{k-1}) = \mathcal{N}(\mathbf{p}_k; \hat{\mathbf{p}}_{k|k-1}, \mathbf{P}_{k|k-1} \odot \mathbf{X}_k) \end{array} \right. \quad (1.82)$$

$$\left\{ \begin{array}{l} p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \simeq \mathcal{IW}(\mathbf{X}_k; \nu_{k|k-1}, \hat{\mathbf{X}}_{k|k-1}) \end{array} \right. \quad (1.83)$$

$\hat{\mathbf{p}}_{k|k-1}$ et $\mathbf{P}_{k|k-1}$ sont obtenus selon les équations de prédiction classiques du filtre de Kalman. $\nu_{k|k-1}$ et $\hat{\mathbf{X}}_{k|k-1}$ dépendent explicitement de $\nu_{k-1|k-1}$ et $\hat{\mathbf{X}}_{k-1|k-1}$ selon les formules décrites dans [31].

• Étape de correction :

A l'étape de correction, les données capteurs sont injectées à travers l'ensemble des mesures réflecteurs $\mathbf{z}_k$ issues de l'objet étendu afin d'en déduire une expression de $p(\mathbf{p}_k, \mathbf{X}_k|\mathcal{Z}^k)$.

Elles sont supposées suivre le modèle gaussien de dispersion suivant :

$$\mathbf{z}_{k,i} = \underbrace{(\mathbf{H}_k \odot \mathbf{I}_p)}_{\tilde{\mathbf{H}}_k} \mathbf{p}_k + \mathbf{s}_{k,i} \quad \mathbf{s}_{k,i} \sim \mathcal{N}(\mathbf{0}, \mathbf{X}_k) \quad (1.84)$$

où $\mathbf{H}_k$ est une matrice de projection conservant la partie position de $\mathbf{p}_k$ et $\mathbf{s}_{k,i}$ est supposé être un bruit blanc.

Selon ce modèle, il est possible de montrer que la vraisemblance jointe des mesures $\mathbf{x}$ peut se factoriser comme suit :

$$p(\mathbf{z}_k | \mathbf{X}_k, \mathbf{p}_k) \propto \mathcal{N}\left(\tilde{\mathbf{z}}_k; \tilde{\mathbf{H}}_k \mathbf{p}_k, \frac{\mathbf{X}_k}{n_k}\right) \mathcal{W}(\mathbf{Z}_k; n_k - 1, \mathbf{X}_k) \quad (1.85)$$

où $\tilde{\mathbf{z}}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbf{z}_{k,i}$ correspond à la moyenne des mesures courantes et $\mathbf{Z}_k$ correspond à leur matrice de covariance empirique :

$$\mathbf{Z}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} (\mathbf{z}_{k,i} - \tilde{\mathbf{z}}_k) (\mathbf{z}_{k,i} - \tilde{\mathbf{z}}_k)^\top \quad (1.86)$$

Cette écriture particulière permet d'obtenir une expression factorisée de la distribution *a posteriori*. En effet, en utilisant le fait qu'elle peut se décomposer comme suit :

$$p(\mathbf{p}_k, \mathbf{X}_k | \mathcal{Z}^k) \propto p(\mathbf{z}_k | \mathbf{p}_k, \mathbf{X}_k) p(\mathbf{p}_k, \mathbf{X}_k | \mathcal{Z}^{k-1}), \quad (1.87)$$

la substitution de (1.85) dans (1.87) permet d'obtenir la forme suivante :

$$p(\mathbf{p}_k, \mathbf{X}_k | \mathcal{Z}^k) \propto \mathcal{N}(\mathbf{p}_k; \hat{\mathbf{p}}_{k|k}, \mathbf{P}_{k|k}) \mathcal{IW}(\mathbf{X}_k; \nu_{k|k}; \hat{\mathbf{X}}_{k|k}) \quad (1.88)$$

où :

$$\begin{cases} \alpha_{k|k-1} = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \frac{1}{n_k} \end{cases} \quad (1.89)$$

$$\begin{cases} \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{s}_{k|k-1}^{-1} \end{cases} \quad (1.90)$$

$$\begin{cases} \hat{\mathbf{p}}_{k|k} = \hat{\mathbf{p}}_{k|k-1} + (\mathbf{K}_k \odot \mathbf{I}_p) (\tilde{\mathbf{z}}_k - \tilde{\mathbf{H}}_k \hat{\mathbf{p}}_{k|k-1}) \end{cases} \quad (1.91)$$

$$\begin{cases} \hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{N}_{k|k-1} + \mathbf{Z}_k \end{cases} \quad (1.92)$$

$$\begin{cases} \nu_{k|k} = \nu_{k|k-1} + n_k \end{cases} \quad (1.93)$$

$$\begin{cases} \mathbf{N}_{k|k-1} = \alpha_{k|k-1}^{-1} (\tilde{\mathbf{z}}_k - \tilde{\mathbf{H}}_k \hat{\mathbf{p}}_{k|k-1}) (\tilde{\mathbf{z}}_k - \tilde{\mathbf{H}}_k \hat{\mathbf{p}}_{k|k-1})^\top \end{cases} \quad (1.94)$$

La distribution *a posteriori* du couple $(\mathbf{p}_k, \mathbf{X}_k)$ étant une loi *inverse Wishart-Gaussienne*, un estimateur *a posteriori* est obtenu en calculant sa moyenne fournie par l'expression :

$$\begin{cases} \mathbb{E}(\mathbf{X}_k | \mathcal{Z}^k) = \frac{\hat{\mathbf{X}}_{k|k}}{\nu_{k|k} - p - 1} \end{cases} \quad (1.95)$$

$$\begin{cases} \mathbb{E}(\mathbf{p}_k | \mathcal{Z}^k) = \hat{\mathbf{p}}_{k|k} \end{cases} \quad (1.96)$$

Remarques sur l'approche matrice aléatoire

- Dans [35], une amélioration de l'approche de [31] est proposée afin de considérer la présence d'un bruit capteur dans le modèle d'observation des réflecteurs.
- Un des principaux défauts de cette méthode est qu'elle restreint la classe de cibles à reconstruire à avoir une forme elliptique ou du moins approchable par une ellipse. Cette hypothèse est assez restrictive, en particulier dans notre application dans laquelle les formes d'amas étudiés ne respectent pas ces conditions. Dans le cas d'objets avec une

forme plus complexe, cet algorithme admet ainsi des limites et aboutit à des formes estimées peu représentatives de la cible inconnue.

- Pour prendre en compte une forme plus irrégulière, une amélioration de cette méthode a été proposée dans [36] où l'objet peut être paramétrisé par un ensemble de plusieurs ellipses. Néanmoins, cette paramétrisation suppose de connaître ou d'estimer leur nombre et dépend des applications considérées.

1.4.2.2 Modélisation comme des formes convexes étoilées

Pour pouvoir pister des objets avec des formes plus diverses que des formes elliptiques, il est possible de considérer qu'ils respectent la propriété convexe-étoilée.

Une forme convexe-étoilée est une forme convexe par rapport à un certain point en son intérieur. Plus formellement, soit E un ensemble quelconque, et a un point de E . Cet ensemble est dit convexe étoilée si et seulement si :

$$\forall x \in E, \{(1-t)a + tx | t \in [0, 1]\} \subset E \quad (1.97)$$

À ce titre, une cible étendue sera convexe-étoilée si tout segment compris entre son centroïde et un point quelconque en sa surface ou en son volume est contenu dans l'objet. En particulier, ce genre représentation est adapté à des formes plus complexes comme celles des avions ou des voitures [37]. Les méthodes de pistage d'une cible étendue convexe-étoilée sont donc fondées sur une représentation analytique de cette dernière comportant :

- la position de son centroïde $\mathbf{p}_k$, et potentiellement ses composantes cinématiques (vitesse, accélération,...),
- une courbe paramétrée par un vecteur $\mathbf{a}_k$, représentant la distance radiale entre les points du contour et le centroïde. De par l'évolution de la forme de l'objet au cours du temps, $\mathbf{a}_k$ est un paramètre dynamique.

Il est également possible de considérer une modélisation non paramétrique des distances radiales de l'objet grâce au formalisme des processus gaussiens. Si l'incertitude jointe sur $\mathbf{a}_k$ et $\mathbf{p}_k$ est représentée par une distribution gaussienne, alors un filtre de Kalman étendu peut être mis en oeuvre pour en déduire une estimation à partir de l'ensemble des mesures des réflecteurs [30]. En effet, ces dernières peuvent être modélisées à travers une fonction non linéaire dépendant à la fois de $\mathbf{a}_k$, de $\mathbf{p}_k$ mais également de la position angulaire du réflecteur par rapport à un repère de référence attaché à l'objet comme illustré sur la figure 1.23.

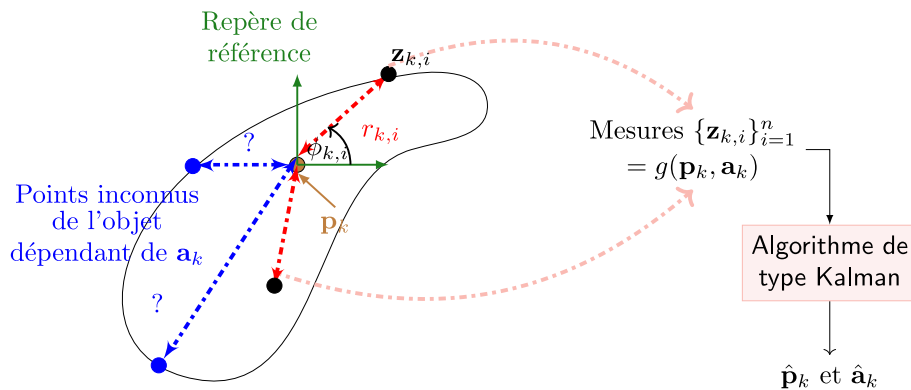


FIGURE 1.23 – Modélisation et pistage d'une forme convexe étoilée : $\phi_{k,i}$ correspond à l'angle d'orientation du réflecteur $\mathbf{z}_{k,i}$ et $r_{k,i}$ à la distance entre le centroïde $\mathbf{p}_k$ et $\mathbf{z}_{k,i}$.

Dans le cas bidimensionnel, chaque réflecteur sur le contour de l'objet peut être décrit selon le modèle d'observation suivant :

$$\mathbf{z}_{k,i} = \mathbf{p}_k + \mathbf{c}_{k,i} r_{k,i} + \mathbf{n}_{k,i} \quad \forall i \in \{1, \dots, n_k\} \quad (1.98)$$

où $\mathbf{n}_{k,i}$ est le bruit d'observation supposé gaussien de matrice de covariance $\mathbf{R}_{k,i}$ et $\mathbf{c}_{k,i}$ est le vecteur d'orientation unitaire s'écrivant :

$$\mathbf{c}_{k,i} = \frac{\mathbf{z}_{k,i} - \mathbf{p}_k}{\|\mathbf{z}_{k,i} - \mathbf{p}_k\|}. \quad (1.99)$$

$r_{k,i}$ correspond à la distance radiale entre la mesure i et $\mathbf{p}_k$ et dépend en particulier de l'angle d'orientation $\phi_{k,i}$ et du vecteur $\mathbf{a}_k$. Le modèle peut être rendu plus flexible dans le cas où des mesures sont situées au sein de l'objet, nous avons alors :

$$\mathbf{z}_{k,i} = \mathbf{p}_k + \varrho_{k,i} \mathbf{c}_{k,i} r_{k,i} + \mathbf{n}_{k,i} \quad \forall i \in \{1, \dots, n_k\} \quad (1.100)$$

où $\varrho_{k,i} \in [0, 1]$ est un paramètre aléatoire. Dans [38], il est montré que $\varrho_{k,i}$ peut être approché par une réalisation d'une distribution gaussienne.

Il reste à définir un modèle suffisamment générique pour la distance radiale. Différentes approches ont été présentées qui sont explicitées ci-après.

• Représentation fondée sur les hypersurfaces aléatoires

Dans cette approche initialement proposée dans [33], l'idée est de tirer parti de la 2π -périodicité des distances radiales pour les approcher par un développement en série de Fourier tronqué d'ordre noté ci-après N_F . Ainsi pour tout angle $\phi_{k,i}$:

$$r_{k,i} = \mathbf{m}(\phi_{k,i})^\top \mathbf{a}_k \quad (1.101)$$

où le vecteur de forme $\mathbf{a}_k$ contient les coefficients de Fourier et $\mathbf{m}(\phi_{k,i}) = \left[\frac{1}{2}, \cos(\phi_{k,i}), \sin(\phi_{k,i}), \dots, \cos(N_F \phi_{k,i}), \sin(N_F \phi_{k,i}) \right]^\top$. Par conséquent, le modèle de mesure peut s'écrire :

$$\mathbf{z}_{k,i} = \mathbf{p}_k + \mathbf{m}(\phi_{k,i})^\top \mathbf{a}_k \mathbf{c}_{k,i} + \mathbf{n}_{k,i} \quad \forall i \in \{1, \dots, n_k\} \quad (1.102)$$

et l'objectif est d'estimer la distribution *a posteriori* jointe de $[\mathbf{p}_k^\top, \mathbf{a}_k^\top]$. Il est à noter que l'ensemble des angles d'orientations sont des variables inconnues. Pour pallier ce problème, il existe plusieurs possibilités :

- elles peuvent être incorporées au problème d'estimation. Par conséquent, l'état à estimer est le vecteur $\mathbf{x}_k = \left[\mathbf{p}_k^\top, \mathbf{a}_k^\top, [\phi_{k,1}, \dots, \phi_{k,n_k}]^\top \right]^\top$. Ce problème devient néanmoins délicat en raison de son caractère mal-posé.
- Elles peuvent être prédites à chaque itération de l'algorithme à partir des mesures et du centroïde estimé.

Un défaut majeur de cette approche est sa sensibilité au nombre de coefficients considérés. En effet, selon le type d'objet à pister, celui-ci peut être difficile à déterminer et aboutit à des estimations de faible qualité. Dans [39], une généralisation de cette approche est proposée pour des cibles en trois dimensions mais limitée à des formes cylindriques.

• Représentation fondée sur les processus gaussiens

L'approche générique présentée dans le cas bidimensionnel dans [38] propose de modéliser les distances radiales au centroïde de l'objet à l'aide d'un processus gaussien. Ce type de modèle appartient à la classe des représentations dites bayésiennes non paramétriques qui présentent la particularité d'être définies sur des espaces de dimension infinie.

D'après la définition fournie par [28][40], un processus gaussien est un processus aléatoire $\{f(\mathbf{u})\}_{\mathbf{u} \in \mathbb{R}^d}$ tel que toute combinaison de points qui en est extraite est un vecteur gaussien. Il est caractérisé par deux paramètres :

- $\mu(\mathbf{u})$ est une fonction dans $\mathbb{R}^d$ dans $\mathbb{R}$ et définit sa fonction moyenne,
- $\kappa(\mathbf{u}, \mathbf{u}')$ définit son noyau où $\mathbf{u}'$ est une variable d'“entrée” quelconque. Il correspond à l'intercovariance entre $f(\mathbf{u})$ et $f(\mathbf{u}')$.

Il peut être noté sous la forme :

$$f \sim \mathcal{GP}(\mu(\mathbf{u}), \kappa(\mathbf{u}, \mathbf{u}')) \quad (1.103)$$

Connaissant N entrées $[\mathbf{u}_1, \dots, \mathbf{u}_N]^\top$ du processus, le vecteur $\mathbf{f} = [f(\mathbf{u}_1), \dots, f(\mathbf{u}_N)]^\top$ est alors un vecteur gaussien s'écrivant :

$$\mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K}) \quad (1.104)$$

où :

$$\begin{cases} \boldsymbol{\mu} = [\mu(\mathbf{u}_1), \dots, \mu(\mathbf{u}_N)]^\top & (1.105) \\ \mathbf{K}_{i,j} = \kappa(\mathbf{u}_i, \mathbf{u}_j) \quad \forall (i, j) \in \{1, \dots, N\}^2 & (1.106) \end{cases}$$

Pour le problème considéré, il est supposé que chaque distance radiale suit le modèle suivant :

$$r_{k,i} = f(\phi_{k,i}) + e_{k,i} \quad e_{k,i} \sim \mathcal{N}(0, \sigma_r^2) \quad (1.107)$$

où f est une réalisation d'un processus gaussien ayant pour entrée l'angle d'orientation $\phi_{k,i}$ et pour sortie une quantité représentant la distance entre le $i^{\text{ème}}$ réflecteur et le centroïde à une erreur près modélisée par le bruit $e_{k,i}$. Le processus se représente donc sous la forme :

$$f \sim \mathcal{GP}(r_o, \kappa(\phi_{k,i}, \phi_{k,j})) \quad \forall i \neq j \quad (1.108)$$

La moyenne r_o permet de quantifier la distance moyenne entre le centroïde et l'ensemble des points du contour de l'objet. Le noyau du processus $\kappa(\phi_{k,i}, \phi_{k,j})$ permet de caractériser la similarité entre la sortie $f(\phi_{k,i})$ et une sortie quelconque $f(\phi_{k,j})$.

Classiquement, il existe plusieurs noyaux pouvant être utilisés, chacun possédant leur particularité. Un choix judicieux est d'en considérer un permettant de prendre en compte la périodicité des angles d'orientations :

$$\kappa(\phi_{k,i}, \phi_{k,j}) = \sigma_f^2 \exp \left(- \frac{2 \sin^2 \left(\frac{|\phi_{k,i} - \phi_{k,j}|}{2} \right)}{l^2} \right) \quad (1.109)$$

Les paramètres de ce noyau ont chacun une signification propre :

- σ_f est un paramètre d'échelle, il permet de quantifier la dispersion des distances radiales de l'objet considéré.
- l est un paramètre de précision permettant de déterminer à partir de quel écart angulaire $f(\phi_{k,i})$ et $f(\phi_{k,j})$ deviennent faiblement corrélées.

Pour pister un objet étendu à partir de processus gaussiens, un ensemble de N_k angles test $\phi_k^t = [\phi_{k,1}^t, \dots, \phi_{k,N_k}^t]^\top$ est défini. L'objectif est de déterminer, conjointement avec les paramètres du centroïde, une estimation des valeurs du processus $\mathbf{f}_k = [f(\phi_{k,1}^t), \dots, f(\phi_{k,N_k}^t)]^\top$ et par suite, des distances radiales des points du contour correspondant aux angles considérés, à partir des mesures réflecteurs.

Il est possible d'y parvenir de manière analytique : dans le cas où le modèle d'évolution du centroïde est gaussien, un algorithme de type Kalman peut être mis en oeuvre dans lequel la distribution *a posteriori* des paramètres joints conserve une structure gaussienne.

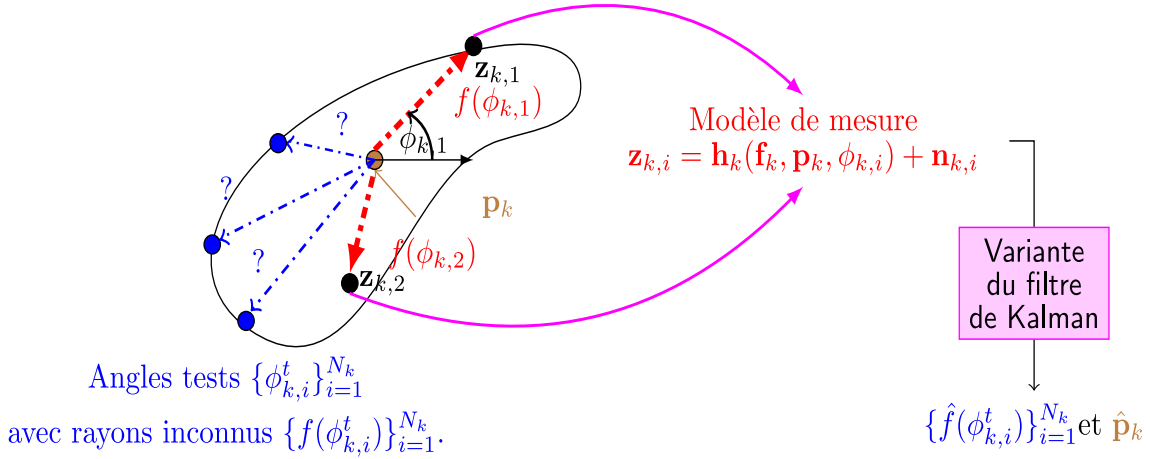


FIGURE 1.24 – Méthode de pistage d'un objet étendu fondée sur des processus gaussiens.

Formellement, le problème d'estimation consiste à déterminer une expression récursive de la distribution jointe $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ où $\mathbf{x}_k = [\mathbf{p}_k^\top, \mathbf{f}_k^\top]^\top$. Pour ce faire, il convient tout d'abord de réécrire le modèle d'observation. À partir du modèle du processus gaussien, il est possible de déterminer la "vraisemblance" de $r_{k,i}$ sachant l'ensemble des valeurs test $\mathbf{f}_k$. Un résultat classique permet d'exprimer la loi conditionnelle $p(r_{k,i} | \mathbf{f}_k)$. Il s'ensuit :

$$r_{k,i} = \left(\mathbf{k}^f(\phi_{k,i}) \right)^\top \mathbf{f}_k + e_{k,i}^r \quad e_{k,i}^r \sim \mathcal{N}(\mathbf{0}, \tilde{\sigma}_{k,i}^2) \quad (1.110)$$

Le vecteur $\mathbf{k}^f(\phi_{k,i})$ ainsi que la variance du modèle $\tilde{\sigma}_{k,i}^2$ sont obtenus en utilisant les propriétés de la moyenne et de la covariance d'une distribution conditionnelle gaussienne :

$$\begin{cases} \mathbf{k}^f(\phi_{k,i}) = \mathbf{K}(\phi_k^t, \phi_k^t)^{-1} \mathbf{k}(\phi_k^t, \phi_{k,i}) \end{cases} \quad (1.111)$$

$$\begin{cases} \tilde{\sigma}_{k,i}^2 = \kappa(\phi_{k,i}, \phi_{k,i}) + \sigma_r^2 - \mathbf{k}(\phi_k^t, \phi_{k,i})^\top \mathbf{K}(\phi_k^t, \phi_k^t)^{-1} \mathbf{k}(\phi_k^t, \phi_{k,i}) \end{cases} \quad (1.112)$$

avec :

$$\begin{cases} \mathbf{k}(\phi_k^t, \phi_{k,i}) = [\kappa(\phi_{k,1}^t, \phi_{k,i}), \dots, \kappa(\phi_{k,N_k}^t, \phi_{k,i})]^\top \end{cases} \quad (1.113)$$

$$\begin{cases} [\mathbf{K}(\phi_k^t, \phi_k^t)]_{i,j} = \kappa(\phi_{k,i}^t, \phi_{k,j}^t) \quad \forall (i,j) \in \{1, \dots, N_k\}^2 \end{cases} \quad (1.114)$$

Ainsi, en substituant l'expression (1.110) dans (1.98), il est obtenu le modèle d'observation gaussien suivant pour chaque réflecteur :

$$\mathbf{z}_{k,i} = \mathbf{h}_k(\mathbf{x}_k) + \tilde{\mathbf{e}}_{k,i} \quad \tilde{\mathbf{e}}_{k,i} \sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{R}}_{k,i}) \quad (1.115)$$

où $\mathbf{h}_k$ et $\tilde{\mathbf{e}}_{k,i}$ sont calculées de la manière suivante :

$$\begin{cases} \mathbf{h}_k(\mathbf{x}_k) = \mathbf{p}_k + \left(\mathbf{k}^f(\phi_{k,i}) \right)^\top \mathbf{f}_k \mathbf{c}_{k,i} & (1.116) \\ \tilde{\mathbf{e}}_{k,i} = e_{k,i}^T \mathbf{c}_{k,i} + \mathbf{n}_{k,i} & (1.117) \\ \tilde{\mathbf{R}}_{k,i} = \mathbf{R}_{k,i} + \tilde{\sigma}_{k,i}^2 \mathbf{c}_{k,i} \mathbf{c}_{k,i}^\top & (1.118) \end{cases}$$

En supposant de plus un modèle d'évolution gaussien sur $\mathbf{f}_k$ [38] ainsi qu'un modèle d'évolution sur $\mathbf{p}_k$, un filtre de Kalman étendu classique peut être mis en oeuvre afin de déterminer une approximation gaussienne de $p(\mathbf{x}_k | \mathbf{z}_{1:k})$.

Remarques sur l'approche processus gaussiens

- Tout comme pour l'approche *hypersurface aléatoire*, les angles d'orientations sont inconnus et doivent être prédits.
- Le formalisme de la méthode peut être également adapté afin de prendre en compte des mesures à l'intérieur de la forme. Les équations d'observations sont alors modifiées en intégrant le paramètre $\varrho_{k,i}$ de l'équation (1.100).
- Dans [41], cette approche est généralisée dans le cas d'objets en trois dimensions. Une paramétrisation en coordonnées sphériques est alors considérée et les angles d'orientations correspondent à des couples d'élévation et d'azimut $\{\phi_{k,i}, \theta_{k,i}\}_{i=1}^{n_k}$. Le processus gaussien admet donc deux entrées sous la forme $f(\gamma)$, avec $\gamma = [\phi, \theta]^\top$. Les fondements théoriques de la méthode restent les mêmes que dans le cas bidimensionnel mais le noyau du processus est fonction de deux vecteurs γ et γ' :

$$k(\gamma, \gamma') = \sigma_f^2 \exp \left(-\frac{d(\gamma, \gamma')}{2l^2} \right) \quad (1.119)$$

avec :

$$d(\gamma, \gamma') = \left(\arccos(\cos(\phi) \cos(\phi') \cos(\theta) \cos(\theta') + \cos(\phi) \cos(\phi') \sin(\theta) \sin(\theta') + \sin(\phi) \sin(\phi')) \right)^2 \quad (1.120)$$

1.5 Conclusions du chapitre

Ce chapitre nous a permis de présenter les différents outils instrumentaux et algorithmiques permettant de caractériser les débris spatiaux.

En particulier, leur modèle dynamique, leur modèle d'observation, ainsi que leurs méthodes de pistages associées ont été présentées. Ces concepts sont primordiaux et vont nous servir dans le cadre des développements théoriques et algorithmiques des prochains chapitres.

Dans le chapitre suivant, nous nous intéressons à présenter des concepts mathématiques tout aussi importants reliés à la théorie des groupes de Lie. En effet, de par la géométrie particulière

des amas de débris, une modélisation sur groupe de Lie de l'amas nous permettra de capturer de manière intrinsèque leur forme.

Chapitre 2

Généralités sur les groupes de Lie

Sommaire

2.1 Concepts fondamentaux	54
2.1.1 Définitions générales	54
2.1.2 Algèbre de Lie	55
2.1.3 Exemples de groupe de Lie	56
2.1.4 Opérateurs de non-commutativité	58
2.2 Dérivation et intégration	61
2.2.1 Dérivées d'ordre 1	61
2.2.2 Dérivées d'ordre 2	62
2.2.3 Intégration sur groupe de Lie	64
2.3 Statistiques et incertitudes	66
2.3.1 Définir une distribution sur variété	66
2.3.2 État de l'art des distributions sur variétés	67
2.3.3 Distribution gaussienne sur groupe de Lie	67
2.4 Algorithmes d'estimation	71
2.4.1 Algorithmes d'optimisation	71
2.4.2 Algorithmes de filtrage	73
2.5 Conclusions du chapitre	81

Ce chapitre est consacré à la présentation des groupes de Lie, qui sont des ensembles mathématiques qui joueront un rôle fondamental dans nos travaux. En effet, leur formalisme nous permettra de modéliser la géométrie d'un amas de débris spatiaux comme décrit dans la suite du manuscrit. Par conséquent, estimer sa forme pourra se reformuler comme un problème d'inférence sur groupe de Lie. Il apparaît ainsi nécessaire de détailler leurs principales propriétés sur lesquelles nous nous appuierons afin de développer de nouvelles méthodes de pistage dans les chapitres suivants.

Dans la première partie de ce chapitre, des fondements théoriques sont présentés : en particulier, la géométrie différentielle, l'analyse et l'intégration sont abordées.

Dans une deuxième partie, nous nous intéressons à définir des statistiques et incertitudes sur groupe de Lie : l'utilisation de distributions de probabilité sur groupe de Lie est en particulier motivée. Ensuite, les modèles statistiques introduits découleront naturellement sur une présentation des principaux algorithmes d'estimation sur groupe de Lie.

2.1 Concepts fondamentaux

2.1.1 Définitions générales

Un groupe de Lie G est défini comme un ensemble mathématique possédant les propriétés d'un groupe ainsi que d'une variété différentielle [42].

De par sa structure de groupe, G est muni d'une loi de composition interne associative. De plus, au sens de cette loi, il existe un élément neutre défini de sorte que chaque élément du groupe possède un inverse.

La structure de variété implique que G est un espace topologique localement homéomorphe à un espace vectoriel. La propriété de différentiabilité induit par ailleurs que l'opération de groupe et l'inversion sont différentiables. Autrement dit, il est possible de calculer la dérivée et l'intégrale de la composition de deux éléments ou de l'inverse d'un élément. À titre d'exemple, les groupes de Lie matriciels sont des groupes de Lie pour lesquels la loi de composition est l'opérateur multiplication de matrices et le neutre correspond à la matrice identité. Dans la suite de notre travail, nous nous focaliserons exclusivement sur ce type de groupes.

Comme G est une variété différentielle, il est possible de définir un espace tangent en tout point. Celui-ci est un espace vectoriel, noté $T_{\mathbf{X}}G$ en $\mathbf{X} \in G$, et dont la dimension définit la dimension de la variété. La structure de groupe devient alors particulièrement intéressante car elle permet de définir une application de l'espace tangent en un point $\mathbf{X}$ à l'espace tangent en un point $\mathbf{Y}$ [42][43]. Celle-ci, nommée *application tangente*, est définie de deux manières en raison, dans le cas général, de la non-commutativité du groupe. L'application tangente à gauche est définie à partir de l'opération de groupe à gauche comme suit :

$$\begin{aligned} L_{\mathbf{X}\mathbf{Y}}^L : T_{\mathbf{X}}G &\rightarrow T_{\mathbf{Y}}G \\ \mathbf{V} &\mapsto \mathbf{Y}\mathbf{X}^{-1}\mathbf{V} \end{aligned} \quad (2.1)$$

De manière analogue, l'application tangente à droite s'écrit :

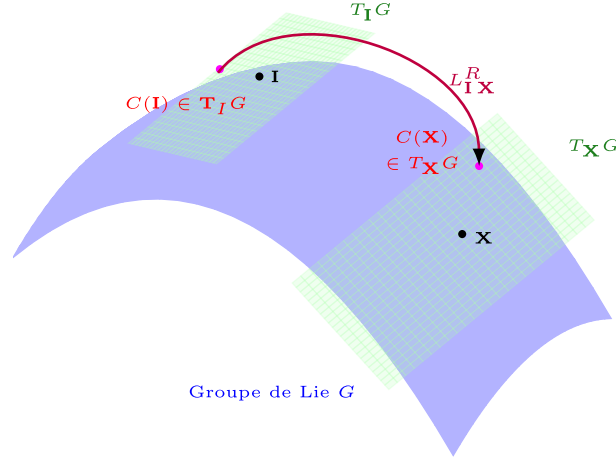
$$\begin{aligned} L_{\mathbf{X}\mathbf{Y}}^R : T_{\mathbf{X}}G &\rightarrow T_{\mathbf{Y}}G \\ \mathbf{V} &\mapsto \mathbf{V}\mathbf{X}^{-1}\mathbf{Y} \end{aligned} \quad (2.2)$$

Ces fonctions permettent de définir la notion d'invariance d'un champ de vecteurs. Notons T_G , le fibré tangent, qui correspond à l'ensemble tous les espaces tangents de G . Formellement, un champ de vecteur est une application différentiable $C(\cdot)$ de G dans le fibré tangent T_G , qui à chaque point de la variété associe un vecteur tangent.

Supposons un élément du champ de vecteurs en $\mathbf{X}$, noté $C(\mathbf{X})$, et un élément du champ de vecteurs en $\mathbf{I}$ noté $C(\mathbf{I})$, le champ de vecteurs C est dit invariant à gauche si et seulement si :

$$C(\mathbf{X}) = L_{\mathbf{I}\mathbf{X}}(C(\mathbf{I})) = \mathbf{X}C(\mathbf{I}) \quad (2.3)$$

Cette propriété d'invariance induit alors une correspondance entre l'espace tangent en $\mathbf{X}$ et l'espace tangent en $\mathbf{I}$. Il est par conséquent possible d'associer tout point de $T_{\mathbf{I}}G$ un point de $T_{\mathbf{X}}G$ via l'opération $L_{\mathbf{I}\mathbf{X}}^L$ ou $L_{\mathbf{I}\mathbf{X}}^R$.


 FIGURE 2.1 – Relation entre $T_I G$ et $T_X G$ à travers l'opérateur L_{IX}^R .

2.1.2 Algèbre de Lie

Comme vu dans la partie précédente, un champ de vecteurs C invariant à gauche dépend uniquement de sa valeur en l'identité $C(\mathbf{I})$. Ainsi, il est pertinent de s'intéresser aux propriétés de l'espace tangent à l'identité. Cet espace, que nous notons $\mathfrak{g}$, se nomme *algèbre de Lie*. Cela implique qu'il existe une application reliant tout point de $\mathfrak{g}$ à un point de G . Elle est obtenue en considérant une courbe paramétrée $\mathbf{X}(t)$ définie sur $\mathbb{R}$ et à valeurs dans G . La dérivée temporelle de cette courbe correspond physiquement à une "vitesse" donc à un élément sur son espace tangent. Il s'ensuit alors la relation suivante :

$$\frac{d\mathbf{X}(t)}{dt} = \mathfrak{a}\mathbf{X}(t) \quad \text{avec } \mathfrak{a}\mathbf{X}(t) \in T_{\mathbf{X}(t)}G \text{ et } \mathfrak{a} \in \mathfrak{g} \quad (2.4)$$

Sous condition que $\mathbf{X}(0) = \mathbf{I}$, la solution à cette équation différentielle s'écrit alors :

$$\mathbf{X}(t) = \text{Exp}_G(\mathfrak{a}t) \quad (2.5)$$

où Exp_G est appelé *exponentielle de groupe*. Cette application permet ainsi de relier un point de l'algèbre à un point du groupe. Dans notre cas, comme nous considérons des groupes de Lie matriciels, elle correspond directement à l'exponentielle de matrice classiquement définie par :

$$\text{Exp}_G(\mathfrak{a}) = \sum_{n=0}^{+\infty} \frac{\mathfrak{a}^n}{n!} \quad \forall \mathfrak{a} \in \mathfrak{g} \quad (2.6)$$

Nous voyons qu'il est possible d'associer à chaque point de $\mathfrak{g}$ un élément de G , néanmoins, relier inversement un point de G à un point de $\mathfrak{g}$ n'est pas toujours possible. En effet, cela viendrait à supposer que l'application $\text{Exp}_G(\cdot)$ est une bijection entre G et $\mathfrak{g}$ ce qui n'est généralement pas le cas. Dans le cas où nous nous plaçons sur un sous-ensemble ouvert A de G situé au voisinage de $\mathbf{I}$, il est tout de même possible de définir l'application Log_G comme sa réciproque locale par :

$$\text{Log}_G(\mathbf{A}) = \sum_{n=1}^{+\infty} (-1)^{n+1} \frac{(\mathbf{I} - \mathbf{A})^n}{n} \quad \forall \mathbf{A} \in A \quad (2.7)$$

Par conséquent, si $\mathbf{a}$ est un élément de $\mathfrak{g}$, alors $\text{Log}_G(\text{Exp}_G(\mathbf{a})) = \mathbf{a}$ si et seulement si $\mathbf{a}$ se situe au voisinage de l'élément $\text{Log}_G(\mathbf{I}) = \mathbf{0}$.

Comme $\mathfrak{g}$ est un espace vectoriel dont nous notons m la dimension, il est possible de définir un isomorphisme allant de $\mathbb{R}^m$ à $\mathfrak{g}$. Cet isomorphisme est noté $[\cdot]_G^\wedge$, et son isomorphisme inverse, allant de $\mathfrak{g}$ à $\mathbb{R}^m$, est noté $[\cdot]_G^\vee$. Si nous décomposons un élément $\mathbf{a} \in \mathbb{R}^m$ dans une base canonique $\{\mathbf{e}_i\}_{i=1}^m$, nous obtenons :

$$\mathbf{a} = \sum_{i=1}^m a_i \mathbf{e}_i \quad (2.8)$$

Par conséquent, l'image de $\mathbf{a}$ par $[\cdot]_G^\wedge$ notée $\mathbf{a} = [\mathbf{a}]_G^\wedge$ correspond à un élément de $\mathfrak{g}$ et peut s'écrire :

$$\mathbf{a} = \sum_{i=1}^m a_i \mathbf{E}_i \quad (2.9)$$

avec $\{\mathbf{E}_i\}_{i=1}^m$ une base canonique de $\mathfrak{g}$ telle que $\mathbf{E}_i = [\mathbf{e}_i]_G^\wedge$. Inversement, nous obtenons que $[\mathbf{a}]_G^\vee = \mathbf{a}$.

Ainsi, chaque élément $\mathbf{X}$ de G au voisinage de l'identité peut se réécrire au travers de ces deux nouvelles applications. En notant $\text{Exp}_G^\wedge(\cdot) = \text{Exp}_G([\cdot]_G^\wedge)$ et $\text{Log}_G^\vee(\cdot) = [\text{Log}_G(\cdot)]_G^\vee$, la relation entre $\mathbf{X}$ et son élément associé $\mathbf{a} \in \mathbb{R}^m$ peut s'écrire $\mathbf{X} = \text{Exp}_G^\wedge(\mathbf{a})$ et réciproquement $\mathbf{a} = \text{Log}_G^\vee(\mathbf{X})$.

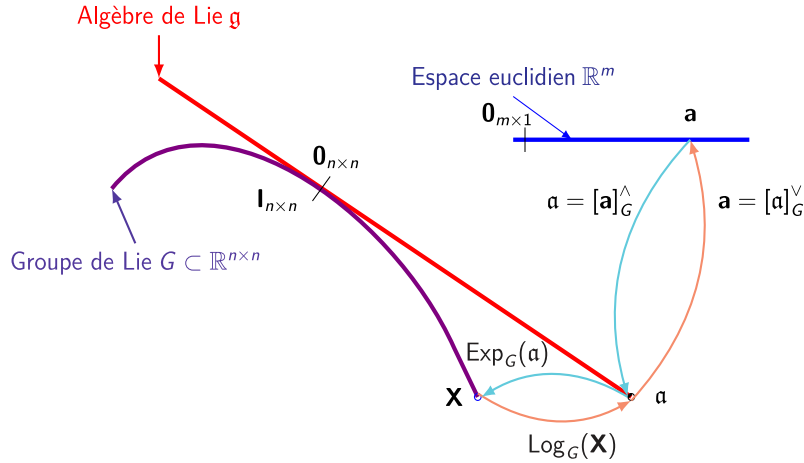


FIGURE 2.2 – Relation entre l'espace euclidien $\mathbb{R}^m$, l'algèbre de Lie $\mathfrak{g}$ et le groupe de Lie G .

2.1.3 Exemples de groupe de Lie

Les groupes de Lie sont des structures relativement abstraites mais nombre de groupes de Lie matriciels se retrouvent dans diverses applications réelles.

- Le groupe de Lie le plus trivial est $\mathbb{R}^n$ muni de la loi additive “+”. En effet, cet espace respecte toutes les conditions d'un groupe et d'une variété différentielle.
- Le groupe de Lie $SO(n)$ nommé *special orthogonal* est le groupe des matrices de rotation dans un espace de dimension n . Il s'écrit :

$$SO(n) = \{\mathbf{P} \in \mathbb{R}^{n \times n} | \mathbf{P} \mathbf{P}^\top = \mathbf{I}, |\mathbf{P}| = 1\}. \quad (2.10)$$

Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, il définit l'orientation d'un objet dans un repère et est particulièrement utilisé en robotique et en vision par ordinateur. Sa dimension est égale à $m = \frac{n(n-1)}{2}$

et son algèbre de Lie correspond à l'ensemble des matrices antisymétriques :

$$\mathfrak{so}(n) = \{[\mathbf{w}]_{\times} \mid \mathbf{w} \in \mathbb{R}^m\}, \quad (2.11)$$

où $\mathbf{w}$ est le vecteur de l'espace euclidien $\mathbb{R}^m$ isomorphe à l'élément de l'algèbre $[\mathbf{w}]_{\times}$ et $[\cdot]_{\times}$ est l'opérateur permettant de créer une matrice antisymétrique à partir d'un vecteur euclidien. Pour $m = 3$ et $\mathbf{w} = [w_1, w_2, w_3]^{\top}$, nous avons alors :

$$[\mathbf{w}]_{\times} = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}. \quad (2.12)$$

→ Le groupe de Lie $SE(n)$ correspond au produit semi-direct entre $SO(n)$ et $\mathbb{R}^n$:

$$SE(n) = \left\{ \mathbf{A} = \begin{bmatrix} \mathbf{P} & \mathbf{x} \\ \mathbf{0} & 1 \end{bmatrix} \mid \mathbf{P} \in SO(n), \mathbf{x} \in \mathbb{R}^n \right\}. \quad (2.13)$$

Si $n = 3$, il correspond à l'ensemble des rotations et des translations possibles pour un objet rigide en trois dimensions. Sa dimension de variété est égale à $\frac{n(n+1)}{2}$ et son algèbre de Lie se représente par l'ensemble des matrices suivantes :

$$\mathfrak{se}(n) = \left\{ \begin{bmatrix} [\mathbf{w}]_{\times} & \mathbf{u} \\ \mathbf{0} & 1 \end{bmatrix} \mid \mathbf{w} \in \mathbb{R}^{\frac{n(n-1)}{2}}, \mathbf{u} \in \mathbb{R}^n \right\}. \quad (2.14)$$

Physiquement parlant, les éléments de l'algèbre modélisent des vecteurs vitesses. En effet, $\mathbf{w}$ représente la vitesse de rotation instantanée associée à la rotation $\mathbf{P}$, et $\mathbf{u}$ la vitesse de translation instantanée associée à la position $\mathbf{x}$.

Il est intéressant de remarquer que $(\mathbb{R}^n, +)$ est un cas particulier de $SE(n)$, pour une matrice de rotation égale à l'identité. En effet, tout $\mathbf{x} \in \mathbb{R}^n$ peut se réécrire sous la forme matricielle suivante :

$$\mathbf{X} = \begin{bmatrix} \mathbf{I} & \mathbf{x} \\ \mathbf{0} & 1 \end{bmatrix}. \quad (2.15)$$

→ Le groupe $\mathbb{R}^{+*}$ muni de la loi multiplicative “ $\times$ ” est également un groupe de Lie possédant la propriété de commutativité. Dans ce cas, les fonctions exponentielle et logarithme sont en bijection puisqu'elles représentent celles classiquement connues dans le cas scalaire. Son algèbre de Lie s'écrit alors naturellement :

$$\mathfrak{R} = \{a \in \mathbb{R} \mid a = \log(b), b \in \mathbb{R}^{+*}\} \quad (2.16)$$

2.1.4 Opérateurs de non-commutativité

2.1.4.1 Adjoint de groupe

Généralement, un groupe de Lie n'est pas commutatif. Néanmoins, il existe un opérateur permettant de permuter deux éléments selon la loi de multiplication. Il est nommé opérateur adjoint et se présente comme une application définie sur $\mathfrak{g}$ et agissant sur $\mathbf{X} \in G$, à travers la formule suivante :

$$\begin{aligned} Ad_{\mathbf{X}}^G : \mathfrak{g} &\rightarrow \mathfrak{g} \\ \mathfrak{y} &\mapsto \mathbf{X} \mathfrak{y} \mathbf{X}^{-1} \end{aligned} \quad (2.17)$$

Cet opérateur admet les deux propriétés suivantes :

→ il a une structure d'homéomorphisme, ce qui signifie que :

$$Ad_{\mathbf{X}}^G(Ad_{\mathbf{Y}}^G(\mathfrak{y})) = Ad_{\mathbf{X}\mathbf{Y}}^G(\mathfrak{y}) \quad \forall \mathbf{X}, \mathbf{Y} \in G, \forall \mathfrak{y} \in \mathfrak{g} \quad (2.18)$$

→ Il est $\mathbb{R}$ -linéaire, ce qui implique que :

$$Ad_{\mathbf{X}}^G(\alpha \mathfrak{y} + \beta \mathfrak{z}) = \alpha Ad_{\mathbf{X}}^G(\mathfrak{y}) + \beta Ad_{\mathbf{X}}^G(\mathfrak{z}) \quad \forall \mathbf{X} \in G, \forall (\mathfrak{y}, \mathfrak{z}) \in \mathfrak{g}^2, \forall (\alpha, \beta) \in \mathbb{R}^2 \quad (2.19)$$

Un opérateur équivalent peut être défini sur $\mathbb{R}^m$. En mettant à profit la linéarité de l'opérateur $Ad_{\mathbf{X}}^G$ et des opérateurs $[\cdot]_G^\wedge$ et $[\cdot]_G^\vee$, il est possible de définir une application sous la forme :

$$\begin{aligned} \mathbb{R}^m &\rightarrow \mathbb{R}^m \\ \mathbf{a} &\mapsto [\mathbf{X}([\mathbf{a}]_G^\wedge) \mathbf{X}^{-1}]_G^\vee = Ad_G(\mathbf{X}) \mathbf{a} \end{aligned} \quad (2.20)$$

Selon cette définition, $Ad_G(\cdot)$ est une matrice appartenant à $\mathbb{R}^{m \times m}$. En se servant du fait que, $\mathbf{X} \text{Exp}_G^\wedge(\mathbf{a}) \mathbf{X}^{-1} = \text{Exp}_G^\wedge(\mathbf{X} \mathbf{a} \mathbf{X}^{-1})$, l'égalité dans (2.20) aboutit à :

$$\mathbf{X} \text{Exp}_G^\wedge(\mathbf{a}) = \text{Exp}_G^\wedge(Ad_G(\mathbf{X}) \mathbf{a}) \mathbf{X} \quad \forall \mathbf{X} \in G, \forall \mathbf{a} \in \mathbb{R}^m \quad (2.21)$$

et permet de représenter la non-commutativité directement sur le groupe.

Exemples d'adjoints :

- Pour un groupe de Lie commutatif, il est aisé de vérifier que $Ad_G(\mathbf{X})$ est la matrice identité, $\forall \mathbf{X} \in G$.
- Pour $G = SO(3)$, l'adjoint en un élément $\mathbf{R}$ se représente par lui-même, autrement dit : $Ad_{SO(3)}(\mathbf{R}) = \mathbf{R}$.
- Pour $G = SE(3)$ pour lequel chaque élément s'écrit :

$$\begin{bmatrix} \mathbf{P} & \mathbf{x} \\ \mathbf{0} & 1 \end{bmatrix}, \quad (2.22)$$

il se représente par la matrice :

$$Ad_G(\mathbf{X}) = \begin{bmatrix} \mathbf{P} & 0 \\ [\mathbf{x}]_{\times} \mathbf{P} & \mathbf{P} \end{bmatrix} \in \mathbb{R}^{6 \times 6}. \quad (2.23)$$

2.1.4.2 Adjoint sur l'algèbre

La non-commutativité du groupe implique une non-commutativité de l'algèbre. Par conséquent, cette propriété peut être caractérisée également sur cet espace à travers un autre opérateur adjoint défini en tout point $\mathfrak{y}$ de $\mathfrak{g}$ par :

$$\begin{aligned} ad_G^{\mathfrak{a}} : \mathfrak{g} &\rightarrow \mathfrak{g} \\ \mathfrak{y} &\mapsto \langle \mathfrak{a}, \mathfrak{y} \rangle \end{aligned} \quad (2.24)$$

où $\langle \cdot, \cdot \rangle$ définit le *crochet de Lie* qui vérifie :

$$\forall (\mathfrak{a}, \mathfrak{b}) \in \mathfrak{g}, \langle \mathfrak{a}, \mathfrak{b} \rangle = \mathfrak{a} \mathfrak{b} - \mathfrak{b} \mathfrak{a} \quad (2.25)$$

Tout comme l'opérateur adjoint de groupe, il peut être étendu sur l'espace vectoriel $\mathbb{R}^m$ par l'application :

$$\begin{aligned} \mathbb{R}^m &\rightarrow \mathbb{R}^m \\ \mathbf{b} &\rightarrow \langle [\mathbf{a}]_G^{\wedge}, [\mathbf{b}]_G^{\wedge} \rangle_G^{\vee} = ad_G(\mathbf{a}) \mathbf{b} \end{aligned} \quad (2.26)$$

où $ad_G(\cdot)$ définit une matrice dans $\mathbb{R}^{m \times m}$.

Pour un groupe de Lie commutatif, $\mathfrak{a} \mathfrak{b} = \mathfrak{b} \mathfrak{a}$ et la matrice $ad_G(\cdot) = \mathbf{0}$.

De manière générale, la forme de cet opérateur peut être déterminée en utilisant une base de $\mathbb{R}^m$. En effet, connaissant une base $\{\mathbf{e}_i\}_{i=1}^m$, nous pouvons écrire que :

$$\mathbf{a} = \sum_{i=1}^m a_i \mathbf{e}_i \quad \forall \mathbf{a} \in \mathbb{R}^m \quad (2.27)$$

L'opérateur ad_G étant linéaire, nous avons alors :

$$ad_G(\mathbf{a}) = \sum_{i=1}^m a_i ad_G(\mathbf{e}_i) \quad (2.28)$$

Ainsi, l'adjoint en tout point $\mathbf{a}$ de $\mathbb{R}^m$ peut être déterminé à partir de chaque élément de la base.

Pour certains groupes de Lie, le calcul de $ad_G(\mathbf{e}_i)$ est réalisable [43] :

→ Pour $G = SO(3)$, il est possible de montrer que $ad_G(\mathbf{e}_i) = \mathbf{E}_i \quad \forall i \in \{1, 2, 3\}$ où $\{\mathbf{E}_i\}_{i=1}^3$ est une base de $\mathfrak{so}(3)$.

→ Pour $G = SE(3)$, ces quantités peuvent s'écrire à partir de la même base $\{\mathbf{E}_i\}_{i=1}^3$:

$$ad_G(\mathbf{e}_1) = \begin{bmatrix} \mathbf{E}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 \end{bmatrix} \quad ad_G(\mathbf{e}_2) = \begin{bmatrix} \mathbf{E}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_2 \end{bmatrix} \quad ad_G(\mathbf{e}_3) = \begin{bmatrix} \mathbf{E}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_3 \end{bmatrix} \quad (2.29)$$

$$ad_G(\mathbf{e}_4) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E}_1 & \mathbf{0} \end{bmatrix} \quad ad_G(\mathbf{e}_5) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E}_2 & \mathbf{0} \end{bmatrix} \quad ad_G(\mathbf{e}_6) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E}_3 & \mathbf{0} \end{bmatrix} \quad . \quad (2.30)$$

Par ailleurs, il est intéressant de noter la relation entre $Ad_G()$ et $ad_G(.)$ donnée par la formule :

$$Ad_G(\text{Exp}_G^\wedge(\mathbf{a})) = \text{Exp}(ad_G(\mathbf{a})). \quad (2.31)$$

Elle s'avère utile selon si nous raisonnons dans le groupe de Lie ou l'espace euclidien qui lui est localement isomorphe.

2.1.4.3 Formules de Baker-Campbell-Hausdorff

Les groupes de Lie étant généralement non commutatifs, il s'ensuit que dans le cas général :

$$\text{Exp}_G^\wedge(\mathbf{a}) \text{Exp}_G^\wedge(\mathbf{b}) \neq \text{Exp}_G^\wedge(\mathbf{a} + \mathbf{b}). \quad (2.32)$$

Par conséquent, déterminer l'élément $\mathbf{z} \in \mathbb{R}^m$ vérifiant $\text{Exp}_G^\wedge(\mathbf{a}) \text{Exp}_G^\wedge(\mathbf{b}) = \text{Exp}_G^\wedge(\mathbf{z})$ n'est pas trivial. Les formules de Baker-Campbell-Hausdorff (BCH) fournissent une solution analytique à ce problème. Elles sont au nombre de deux et ont été établies dans [44]. La première formule s'écrit à l'ordre un :

$$\text{Log}_G^\vee(\text{Exp}_G^\wedge(-\mathbf{a}) \text{Exp}_G^\wedge(\mathbf{a} + \mathbf{b})) = \phi_G(-\mathbf{a}) \mathbf{b} + O(\|\mathbf{b}\|^2). \quad (2.33)$$

Dans cette formule apparaît une quantité nommée *matrice jacobienne à gauche de groupe* s'écrivant :

$$\phi_G(\mathbf{a}) = \sum_{n=0}^{+\infty} \frac{ad_G(\mathbf{a})^n}{(n+1)!} \quad (2.34)$$

Elle correspond à la dérivée dite à *gauche* de l'exponentielle de groupe dont nous détaillerons la définition dans la section suivante.

La seconde formule relie deux éléments $\mathbf{a}$ et $\mathbf{b}$ de la manière suivante :

$$\text{Log}_G^\vee(\text{Exp}_G^\wedge(\mathbf{a}) \text{Exp}_G^\wedge(\mathbf{b})) = \mathbf{b} + \psi_G(\mathbf{b}) \mathbf{a} + O(\|\mathbf{a}\|^2). \quad (2.35)$$

Elle fait apparaître l'inverse de la *matrice jacobienne à gauche du groupe* $\psi_G(.)$ et s'écrit à l'aide d'une série entière de l'opérateur adjoint pondéré par les nombres de Bernoulli $\{B_n\}_{n=0}^{+\infty}$ [45] :

$$\psi_G(\mathbf{a}) = \sum_{n=0}^{+\infty} \frac{B_n ad_G(\mathbf{a})^n}{n!}. \quad (2.36)$$

Il est intéressant de noter que lorsque G est un groupe de Lie commutatif, alors $\psi_G(\cdot)$ et $\phi_G(\cdot)$ sont égales à la matrice identité de par la nullité de $ad_G(\cdot)$. Par conséquent, les formules BCH s'écrivent : $\text{Exp}_G^\wedge(\mathbf{a}) \text{Exp}_G^\wedge(\mathbf{b}) = \text{Exp}_G^\wedge(\mathbf{a} + \mathbf{b})$.

2.2 Dérivation et intégration

2.2.1 Dérivées d'ordre 1

Nous nous intéressons dans cette partie à la généralisation sur groupe de Lie de la formule de dérivation classique sur un espace Euclidien. Nous savons que si $f : \mathbb{R}^m \rightarrow \mathbb{R}$, alors sa dérivée directionnelle classique est fournie par :

$$d_{\mathbf{h}}f(\mathbf{x}) = \left. \frac{\partial f(\mathbf{x} + t \mathbf{h})}{\partial t} \right|_{t=0} \quad \forall t \in \mathbb{R} \quad \forall \mathbf{h} \in \mathbb{R}^m. \quad (2.37)$$

Par conséquent, si nous supposons l'existence d'une fonction $g : G \rightarrow G'$, où G et G' sont deux groupes de Lie de dimensions respectives m et p , nous aimerions pouvoir définir quelle est la signification mathématique de différencier par rapport un élément du groupe G . Une première possibilité pour un groupe de Lie matriciel est d'en vectoriser chaque élément, et de dériver par rapport à chaque composante. Néanmoins, cette méthode n'est pas intrinsèque au sens où elle ne prend pas en compte la géométrie du groupe, ce qui peut poser des difficultés théoriques lorsque que nous cherchons à conserver les propriétés de celui-ci dans des algorithmes.

Une autre possibilité est alors de conserver sa géométrie en passant par sa représentation sur l'algèbre de Lie. En effet, l'incrément de dérivation $\mathbf{h}$ classiquement utilisé dans un cas euclidien, peut être directement défini dans le cas d'un groupe de Lie en passant par sa représentation euclidienne. L'opération $\mathbf{x} + \mathbf{h}$ peut alors être généralisée par une opération de groupe à travers la fonction exponentielle sous la forme " $\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h})$ " où " $\text{Exp}_G^\wedge(\mathbf{h}) \mathbf{X}$ " puisque cette opération peut être définie à droite ou à gauche dues aux propriétés de G' . De cette manière, la "dérivée à droite de g par rapport à $\mathbf{X}$ " se détermine par la formule [46] :

$$\frac{dg}{d\mathbf{X}}^R = \left. \frac{\partial \text{Log}_{G'}^\vee(g(\mathbf{X})^{-1} g(\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h})))}{\partial \mathbf{h}} \right|_{\mathbf{h}=0} \quad \mathbf{h} \in \mathbb{R}^m. \quad (2.38)$$

Cette dérivée est également connue sous le nom de *jacobienne à droite de g* . De même la dérivée à gauche s'écrit sous la forme :

$$\frac{dg}{d\mathbf{X}}^L = \left. \frac{\partial \text{Log}_{G'}^\vee(g(\text{Exp}_G^\wedge(\mathbf{h}) \mathbf{X}) g(\mathbf{X})^{-1})}{\partial \mathbf{h}} \right|_{\mathbf{h}=0} \quad \mathbf{h} \in \mathbb{R}^m \quad (2.39)$$

nommée *jacobienne à gauche de g* .

Lorsque g est une fonction à valeurs dans $\mathbb{R}^p$ ($p \geq 1$) muni de l'opération "+", ce qui sera généralement notre cas, l'opérateur $\text{Log}_{G'}^\vee(\cdot)$ devient l'identité. De plus, l'opération $g(\mathbf{X})^{-1} g(\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h}))$ se transforme en $g(\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h})) - g(\mathbf{X})$. Ce dernier terme s'éliminant par dérivation, nous alors faisons apparaître la définition des *dérivées de Lie* $\mathcal{L}_{g(\mathbf{X})}^L$ et $\mathcal{L}_{g(\mathbf{X})}^R$

respectivement à droite et à gauche de g en $\mathbf{X}$ [47] :

$$\mathcal{L}_{g(\mathbf{X})}^R = \left. \frac{\partial g(\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h}))}{\partial \mathbf{h}} \right|_{\mathbf{h}=\mathbf{0}} \quad \mathbf{h} \in \mathbb{R}^m, \quad (2.40)$$

$$\mathcal{L}_{g(\mathbf{X})}^L = \left. \frac{\partial g(\text{Exp}_G^\wedge(\mathbf{h}) \mathbf{X})}{\partial \mathbf{h}} \right|_{\mathbf{h}=\mathbf{0}} \quad \mathbf{h} \in \mathbb{R}^m. \quad (2.41)$$

2.2.1.1 Lien entre différentielle et dérivée

Considérons $f : G \rightarrow G'$ est une fonction différentiable et $\mathbf{X} \in G$. Soit $\boldsymbol{\eta}$ un champ de vecteurs et $\mathbf{X}(t)$ une courbe paramétrée telle que $\mathbf{X}(0) = \mathbf{X}$ et $\dot{\mathbf{X}}(0) = \boldsymbol{\eta}(\mathbf{X})$, alors l'application linéaire [48] :

$$\begin{aligned} T_{\mathbf{X}}G &\rightarrow T_{f(\mathbf{X})}G' \\ \boldsymbol{\eta}(\mathbf{X}) &\rightarrow df_{\mathbf{X}}(\boldsymbol{\eta}) = \left. \frac{df(\mathbf{X}(t))}{dt} \right|_{t=0} \end{aligned} \quad (2.42)$$

est appelée la différentielle de f en $\mathbf{X}$ et $df_{\mathbf{X}}(\boldsymbol{\eta})$ est un vecteur tangent à G' en $f(\mathbf{X})$. Si f est à valeurs réelles alors $df_{\mathbf{X}}(\boldsymbol{\eta})$ est un réel et nous notons $\boldsymbol{\eta}f$ la fonction :

$$\begin{aligned} G &\rightarrow \mathbb{R} \\ \mathbf{X} &\rightarrow (\boldsymbol{\eta}f)(\mathbf{X}) = df_{\mathbf{X}}(\boldsymbol{\eta}) \end{aligned} \quad (2.43)$$

que nous notons de façon simplifiée $\boldsymbol{\eta}f = df(\boldsymbol{\eta})$.

2.2.2 Dérivées d'ordre 2

Classiquement, sur un espace euclidien, une dérivée au second ordre d'une fonction à valeurs réelles est définie à partir d'une dérivation de la dérivée première.

Sur un groupe de Lie, la dérivée première d'une fonction à valeurs réelles, qui correspond à la dérivée de Lie, vit sur un espace tangent. Par conséquent, cette définition ne peut se généraliser directement car elle reviendrait à comparer la dérivée de la fonction en deux points différents donc sur deux espaces tangents distincts.

Pour pallier ce problème, il est nécessaire d'introduire la notion de différentiation selon un champ de vecteurs grâce à la définition de la *connexion affine*.

2.2.2.1 Connexion affine

Soit G un groupe de Lie et $C^\infty(G, T_G)$ l'union des champs de vecteurs sur G . Une connexion affine est une application bilinéaire ∇ de $C^\infty(G, T_G) \times C^\infty(G, T_G)$ dans $C^\infty(G, T_G)$ définie par [48][49][50] :

$$\begin{aligned} C^\infty(G, T_G) \times C^\infty(G, T_G) &\rightarrow C^\infty(G, T_G) \\ (\mathbf{a}, \mathbf{b}) &\mapsto \nabla_{\mathbf{a}} \mathbf{b} \end{aligned} \quad (2.44)$$

Cette application admet deux propriétés fondamentales pour toute fonction $f : G \rightarrow \mathbb{R}$ infiniment différentiable $\forall(\mathbf{a}, \mathbf{b}) \in C^\infty(G, T_G) \times C^\infty(G, T_G)$:

$$\begin{cases} \nabla_{f\mathbf{a}} \mathbf{b} = f \nabla_{\mathbf{a}} \mathbf{b} \\ \nabla_{\mathbf{a}}(f\mathbf{b}) = df(\mathbf{a})\mathbf{b} + f \nabla_{\mathbf{a}} \mathbf{b} \end{cases} \quad (2.45)$$

$$(2.46)$$

$df(\mathbf{a})$ correspond à la différentielle de f calculée en $\mathbf{a}$ au sens de la définition (2.43). Il est à noter que la seconde formule est connue sous le nom de *formule de Leibniz*.

Comme expliqué dans [48], il est possible de la mettre à profit pour obtenir une expression reliant une connexion affine et une dérivée d'ordre deux. Cette formule implique alors que :

$$df(\mathbf{a}\mathbf{b}) = df^2(\mathbf{a})\mathbf{b} + df(\nabla_{\mathbf{a}}\mathbf{b}) \quad (2.47)$$

→ Le terme à gauche de l'égalité correspond à la différentielle de f calculée en $\mathbf{a}\mathbf{b}$ et s'écrit d'après (2.43) :

$$df(\mathbf{a}\mathbf{b}) = (\mathbf{a}\mathbf{b}) f \quad (2.48)$$

→ Dans le premier terme à droite de l'égalité, nous faisons apparaître une quantité correspondant à la *dérivée covariante* de f au sens de la connexion ∇ et qui peut être vue comme une généralisation de la dérivée seconde.

→ Le second terme fait apparaître la différentielle de f appliquée à la connexion $\nabla_{\mathbf{a}}\mathbf{b}$ et peut s'écrire, d'après (2.43) :

$$df(\nabla_{\mathbf{a}}\mathbf{b}) = (\nabla_{\mathbf{a}}\mathbf{b}) f \quad (2.49)$$

Par conséquent, la dérivée covariante de f s'écrit explicitement en fonction de $\mathbf{a}$, $\mathbf{b}$ et de la connexion ∇ , sous la forme d'un opérateur $\tilde{\mathbf{H}}(\cdot)$

$$\begin{aligned} C^\infty(G, T_G) \times C^\infty(G, T_G) &\rightarrow C^\infty(G, T_G) \\ (\mathbf{a}, \mathbf{b}) &\mapsto \tilde{\mathbf{H}}(\mathbf{a}, \mathbf{b}) \end{aligned} \quad (2.50)$$

vérifiant :

$$\tilde{\mathbf{H}}(\mathbf{a}, \mathbf{b}) = (\mathbf{a}\mathbf{b} - \nabla_{\mathbf{a}}\mathbf{b}) f \quad (2.51)$$

2.2.2.2 Hessienne sur groupe de Lie

Comme G est un groupe de Lie, tout champ de vecteur est caractérisé par sa réalisation dans l'algèbre de Lie. Ainsi, il existe des connexions affines spécifiques admettant des propriétés inhérentes à sa structure de groupe plutôt qu'à sa structure de variété. Elle se nomment *connexions de Cartan-Schouten* [51][52]. Il peut en être définies trois qui ont la propriété commune d'être invariantes à gauche par l'action de groupe. De même, nous pouvons en définir trois qui sont invariantes par l'action à droite. Une connexion est invariante à gauche si elle vérifie la propriété suivante :

$$\mathbf{X} \nabla_{\mathbf{a}} \mathbf{b} = \nabla_{\mathbf{X}\mathbf{a}} \mathbf{b} \quad \forall \mathbf{X} \in G, \forall (\mathbf{a}, \mathbf{b}) \in \mathfrak{g}^2 \quad (2.52)$$

Nous nous focalisons sur la *connexion de Cartan-Schouten symétrique* définie pour deux éléments sur $\mathfrak{g}$ par :

$$\nabla_{\mathfrak{a}} \mathfrak{b} = \frac{1}{2} \langle \mathfrak{a}, \mathfrak{b} \rangle \quad \forall (\mathfrak{a}, \mathfrak{b}) \in \mathfrak{g}^2 \quad (2.53)$$

Selon cette connexion, l'opérateur $\tilde{\mathbf{H}}_{\mathbf{X}}(\cdot)$ peut alors s'écrire sur $\mathfrak{g} \times \mathfrak{g}$, d'après (2.51) :

$$\tilde{\mathbf{H}}(\mathfrak{a}, \mathfrak{b}) = (\mathfrak{a} \mathfrak{b} - \frac{1}{2} \langle \mathfrak{a}, \mathfrak{b} \rangle) f(\mathbf{X}) \quad (2.54)$$

$$= \frac{1}{2} (\mathfrak{a} \mathfrak{b} + \mathfrak{b} \mathfrak{a}) f(\mathbf{X}) \quad \forall \mathbf{X} \in G, \forall (\mathfrak{a}, \mathfrak{b}) \in \mathfrak{g}^2 \quad (2.55)$$

et il est montré [48] qu'il peut se réécrire sous la forme :

$$\tilde{\mathbf{H}}(\mathfrak{a}, \mathfrak{b}) = \mathfrak{a}^\top \mathbf{H}(\mathbf{X}) \mathfrak{b} \quad \forall \mathbf{X} \in G, \forall (\mathfrak{a}, \mathfrak{b}) \in \mathfrak{g}^2 \quad (2.56)$$

où :

$$\mathbf{H}(\mathbf{X}) = \frac{\partial}{\partial \mathbf{h}_1} \frac{\partial}{\partial \mathbf{h}_2} f(\mathbf{X} \text{Exp}_G^\wedge(\mathbf{h}_1) \text{Exp}_G^\wedge(\mathbf{h}_2)) \Big|_{\mathbf{h}_1=0, \mathbf{h}_2=0} \quad \forall \mathbf{X} \in G \quad \forall (\mathbf{h}_1, \mathbf{h}_2) \in (\mathbb{R}^m)^2 \quad (2.57)$$

où m est la dimension de G . Cet opérateur nous permet donc de définir une *matrice hessienne* sur le groupe en considérant une double dérivée à travers la fonction exponentielle et admettant la propriété de symétrie. En accord avec [47] et [48], il est alors possible de réaliser un développement de Taylor-Young sur groupe de Lie selon la formule suivante :

$$f(\mathbf{X} \text{Exp}_G^\wedge(\epsilon)) = f(\mathbf{X}) + \mathcal{L}_{f(\mathbf{X})}^R{}^\top \epsilon + \epsilon^\top \mathbf{H}(\mathbf{X}) \epsilon + O(\|\epsilon\|^2) \quad \forall \mathbf{X} \in G, \forall \epsilon \in \mathbb{R}^m \quad (2.58)$$

2.2.3 Intégration sur groupe de Lie

Dans les deux paragraphes précédents, nous avons pu définir des techniques permettant de différencier une fonction définie sur groupe de Lie à l'ordre un et deux. Comme nous souhaitons travailler dans nos travaux avec des distributions de probabilité, il paraît naturel de s'intéresser à la manière de les intégrer.

Un groupe de Lie est un ensemble mesurable, par conséquent, il est possible de définir une mesure de groupe $\mu(\cdot)$ permettant de réaliser une intégration à partir d'une fonction définie sur groupe de Lie à valeurs dans $\mathbb{R}$ sous la forme $\int_{\mathbf{X} \in G} f(\mathbf{X}) \mu(d\mathbf{X})$. S'il existe une paramétrisation globale de groupe notée $g : G \rightarrow \mathbb{R}^m$, qui associe à tout élément du groupe un élément sur $\mathbb{R}^m$, alors une intégrale sur groupe de Lie peut s'effectuer de la manière suivante :

$$\int_{\mathbf{X} \in G} f(\mathbf{X}) \mu(d\mathbf{X}) = c \int_{\mathbf{q} \in \mathbb{R}^m} f(g(\mathbf{q})) |\phi_G(\mathbf{q})| d\mathbf{q}, \quad (2.59)$$

où $c \in \mathbb{R}$ est une constante de normalisation et $d(\cdot)$ définit la mesure de Lebesgue sur $\mathbb{R}^m$. Dans certains cas particuliers, la jacobienne à gauche $\phi_G(\cdot)$ définie par (2.34) est connue : il est alors possible d'en déduire une expression analytique de cette intégrale. Néanmoins, elle est non triviale de manière générale.

Il existe des groupes de Lie particuliers dit unimodulaires vérifiant de bonnes propriétés d'intégration et particulièrement l'invariance par composition à droite ou à gauche et par inversion. Dans notre application, les groupes de Lie que nous utilisons respectent tous les

propriétés d'unimodularité. Par conséquent, nous proposons de définir les outils d'intégration sur ces groupes spécifiques.

2.2.3.1 Mesure de Haar et fonction modulaire

Nous recherchons une mesure respectant des propriétés d'invariance par multiplication *à droite* ou *à gauche*. Ainsi, elle généralise la mesure de Lebesgue sur espace euclidien, qui est invariante par translation.

Soit G un groupe de Lie et une fonction $f \in \mathcal{C}_c(G)$, où $\mathcal{C}_c(G)$ est l'ensemble des fonctions continues sur G à support compact, définissons de plus $\mathcal{P}(G)$ une tribu borélienne de G et $\mu_G(\cdot) : \mathcal{P}(G) \rightarrow \mathbb{R}^+$ une mesure borélienne sur G .

$\mu_G(\cdot)$ définit une *mesure de Haar à gauche* si la propriété suivante est vérifiée :

$$\int_{\mathbf{X} \in G} f(\mathbf{Y}\mathbf{X}) \mu_G(d\mathbf{X}) = \int_{\mathbf{X} \in G} f(\mathbf{X}) \mu_G(d\mathbf{X}). \quad (2.60)$$

Ceci implique qu'elle est invariante par multiplication *à gauche*. Elle permet alors de définir la forme linéaire suivante allant de $\mathcal{C}_c(G)$ dans $\mathbb{R}$ par :

$$f \rightarrow \int_{\mathbf{X} \in G} f(\mathbf{Y}\mathbf{X}\mathbf{Y}^{-1}) \mu_G(d\mathbf{X}) \quad (2.61)$$

et vérifiant également cette propriété d'invariance. Nous avons alors :

$$\int_{\mathbf{X} \in G} f(\mathbf{Y}\mathbf{X}\mathbf{Y}^{-1}) \mu_G(d\mathbf{X}) = \Delta(\mathbf{Y}) \int_{\mathbf{X} \in G} f(\mathbf{X}) \mu_G(d\mathbf{X}). \quad (2.62)$$

où l'application $\Delta : G \rightarrow \mathbb{R}^+$ est appelée *fonction modulaire* de G , et s'écrit par définition :

$$\mu_G(A\mathbf{Y}) = \Delta(\mathbf{Y}) \mu_G(A) \quad \forall A \in \mathcal{P}(G) \quad \forall \mathbf{Y} \in G. \quad (2.63)$$

Δ permet donc de définir la notion de groupe de Lie unimodulaire. En effet, c'est un groupe pour lequel la mesure de Haar devient bi-invariante (invariante *à gauche* et *à droite*) et $\Delta(\cdot)$ est égale à 1. Par conséquent, l'intégrale (2.62) s'écrit pour tout groupe de Lie unimodulaire :

$$\int_{\mathbf{X} \in G} f(\mathbf{Y}\mathbf{X}\mathbf{Y}^{-1}) \mu_G(d\mathbf{X}) = \int_{\mathbf{X} \in G} f(\mathbf{X}) \mu_G(d\mathbf{X}). \quad (2.64)$$

2.2.3.2 Intégration sur groupe de Lie unimodulaire

Les groupes de Lie unimodulaires présentent un intérêt lorsque nous cherchons à intégrer sur un groupe de Lie étant le produit direct de deux autres groupes de Lie.

Supposons G_1 et G_2 deux groupes de Lie et la fonction $f : G_1 \times G_2 \rightarrow \mathbb{R}$. Nous définissons l'intégrale de f sur $G = G_1 \times G_2$ par :

$$I = \int_{\mathbf{X} \in G_1 \times G_2} f(\mathbf{X}) \mu_G(d\mathbf{X}) \quad (2.65)$$

Dans un cadre générique, cette intégrale s'écrit [53] :

$$I = \int_{\mathbf{X}_1 \in G_1, \mathbf{X}_2 \in G_2} f(\mathbf{X}_1, \mathbf{X}_2) \Delta_{G_2}(\mathbf{X}_2) \mu_{G_1}(d\mathbf{X}_1) \mu_{G_2}(d\mathbf{X}_2) \quad (2.66)$$

Par définition, lorsque G_2 est unimodulaire, la mesure $\mu_{G_2}(\cdot)$ est bi-invariante et $\Delta_{G_2}(\cdot)$ est égale à 1. L'expression de l'intégrale se simplifie alors :

$$I = \int_{\mathbf{X}_1 \in G_1, \mathbf{X}_2 \in G_2} f(\mathbf{X}_1, \mathbf{X}_2) \mu_{G_1}(d\mathbf{X}_1) \mu_{G_2}(d\mathbf{X}_2) \quad (2.67)$$

et apparaît comme une intégrale séparable en $\mathbf{X}_1$ et $\mathbf{X}_2$, si f l'est également, ce qui est intéressant d'un point de vue calculatoire.

Par ailleurs, la propriété d'unimodularité permet de définir une formule d'intégration par parties sur groupe de Lie. Définissons $f_1 : G \rightarrow \mathbb{R}^p$ et $f_2 : G \rightarrow \mathbb{R}^p$ deux fonctions intégrables selon une mesure de Haar. Si les dérivées de Lie de f_1 et f_2 sont bornées par deux fonctions intégrables, il est possible de montrer que [47] :

$$\int_{\mathbf{X} \in G} f_1(\mathbf{X}) \mathcal{L}_{f_2(\mathbf{X})}^R \mu_G(d\mathbf{X}) = - \int_{\mathbf{X} \in G} \mathcal{L}_{f_1(\mathbf{X})}^R f_2(\mathbf{X}) \mu_G(d\mathbf{X}) \quad (2.68)$$

2.3 Statistiques et incertitudes

2.3.1 Définir une distribution sur variété

Pour décrire statistiquement une variable définie sur un groupe de Lie, par exemple à des fins de résoudre des problèmes d'estimation, il est nécessaire de lui associer une distribution de probabilité.

Formellement, elle peut être définie par une fonction $p(\mathbf{X})$ vérifiant :

$$\begin{cases} p(\mathbf{X}) > 0 \quad \forall \mathbf{X} \in G, \\ \int_{\mathbf{X} \in G} p(\mathbf{X}) \mu_G(d\mathbf{X}) = 1. \end{cases} \quad (2.69)$$

$$(2.70)$$

Une des difficultés dans la définition d'une distribution sur variétés est de parvenir à donner un sens à ses caractéristiques et en particulier à ses moments. En effet, les opérations usuelles utilisées en statistiques euclidiennes ne peuvent pas être directement généralisées sur variétés. Par conséquent, elles doivent être redéfinies en fonction des lois de composition et distances propres au groupe de Lie considéré. Il est par exemple possible d'obtenir une définition géométrique du moment d'ordre un à travers la moyenne de Fréchet [54]. Elle est définie comme étant le point $\bar{\mathbf{Y}}$ d'une variété $\mathcal{M}$ minimisant un critère dépendant d'une distance intrinsèque d :

$$\bar{\mathbf{Y}} = \operatorname{argmin}_{\mathbf{Y} \in \mathcal{M}} \int_{\mathbf{X} \in \mathcal{M}} d(\mathbf{Y}, \mathbf{X})^2 p(\mathbf{X}) \mu_{\mathcal{M}}(d\mathbf{X}) \quad (2.71)$$

où $\mu_{\mathcal{M}}$ est une mesure sur $\mathcal{M}$.

2.3.2 État de l'art des distributions sur variétés

Dans la littérature, il existe nombre de distributions définies sur des groupes de Lie et plus généralement sur des variétés différentielles. En effet, des lois de probabilités adaptées à la structure de ces espaces ont été construites pour répondre à des objectifs variés souvent dans des contextes d'estimation de paramètres.

Dans [55], une distribution gaussienne sur variété riemannienne est définie comme la distribution maximisant l'entropie d'une variable aléatoire. Selon ce principe, une distribution gaussienne riemannienne et admettant une forme analytique a été proposée sur la variété des matrices de covariance [56].

Il existe également des distributions établies dans un cadre non "gaussien". Sur les variétés de Stiefel : la distribution de Bingham-Langevin ainsi que la distribution de Bingham-Fisher ont récemment fait l'objet d'une attention particulière pour caractériser des variables respectant des propriétés d'orthogonalité [57][58] .

Sur groupe de Lie, nous pouvons citer la distribution de Von Mises-Fisher qui est largement utilisée sur les groupes $SO(2)$ et $SO(3)$, afin de modéliser un point aléatoirement reparté sur une sphère [59]. Sur $SO(2)$, il est également possible de définir une distribution gaussienne périodique sur l'angle de rotation associé à ce groupe [60].

2.3.3 Distribution gaussienne sur groupe de Lie

2.3.3.1 Définition générale

Les distributions décrites précédemment, soit ne sont pas adaptées aux groupes de Lie au sens où elles ne prennent pas en compte leurs propriétés algébriques, soit sont dédiées à des groupes de Lie particuliers. Idéalement, nous aimerions avoir accès à une distribution qui peut s'appliquer à tout groupe de Lie matriciel indifféremment de sa structure. Dans le cadre de ce manuscrit, nous nous intéressons à une densité particulière nommée *distribution gaussienne concentrée sur groupe de Lie*. Celle-ci, formalisée par [61][62], est pertinente à plusieurs points de vue :

- elle est définie comme la distribution à deux paramètres sur un groupe de Lie G de dimension m , réalisant le maximum d'entropie sous contraintes d'égalité sur ces derniers notés $\boldsymbol{\mu} \in G$ et $\mathbf{P} \in \mathbb{R}^{m \times m}$. Selon [47], nous pouvons l'obtenir en résolvant le problème d'optimisation suivant :

$$p = -\operatorname{argmin}_f \int_{\mathbf{X} \in G} \log f(\mathbf{X}) f(\mathbf{X}) \mu_G(d\mathbf{X}) \quad (2.72)$$

sous contraintes que :

$$\left\{ \int_{\mathbf{X} \in G} \operatorname{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{X}) p(\mathbf{X}) \mu_G(d\mathbf{X}) = \mathbf{0}_{m \times 1} \right. \quad (2.73)$$

$$\left. \mathbf{P} = \int_{\mathbf{X} \in G} \operatorname{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{X}) \operatorname{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{X})^\top p(\mathbf{X}) \mu_G(d\mathbf{X}). \right. \quad (2.74)$$

- Les deux paramètres $\boldsymbol{\mu}$ et $\mathbf{P}$ de la distribution peuvent s'apparenter à une moyenne et une covariance, bien que selon le groupe de Lie considéré, ils ne peuvent être rigoureusement considérés comme tels.

→ Elle admet des propriétés de symétrie et d'invariance par multiplication.

Il est possible de montrer que la *distribution gaussienne concentrée à gauche* s'écrit [63] :

$$p(\mathbf{Y}) = A^L(\mathbf{Y}, \boldsymbol{\mu}) \exp \left(-\frac{1}{2} \|\text{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{Y})\|_{\mathbf{P}}^2 \right), \quad \forall \mathbf{Y} \in G \quad (2.75)$$

avec :

$$A^L(\mathbf{Y}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^m |\phi_G(-\text{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{Y})) \mathbf{P} \phi_G(-\text{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{Y}))^\top|}}$$

L'opérateur $\|\cdot\|^2$ fait référence à la distance de Mahalanobis et $|\cdot|$ à l'opérateur déterminant d'une matrice.

De manière équivalente, nous pouvons définir une *distribution gaussienne concentrée à droite* sous la forme suivante :

$$p(\mathbf{Y}) = A^R(\mathbf{Y}, \boldsymbol{\mu}) \exp \left(-\frac{1}{2} \|\text{Log}_G^\vee(\mathbf{Y} \boldsymbol{\mu}^{-1})\|_{\mathbf{P}}^2 \right), \quad \forall \mathbf{Y} \in G \quad (2.76)$$

avec :

$$A^R(\mathbf{Y}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^m |\phi_G(-\text{Log}_G^\vee(\mathbf{Y} \boldsymbol{\mu}^{-1})) \mathbf{P} \phi_G(-\text{Log}_G^\vee(\mathbf{Y} \boldsymbol{\mu}^{-1}))^\top|}}$$

Au voisinage de $\boldsymbol{\mu}$, nous pouvons approcher $\phi_G(-\text{Log}_G^\vee(\mathbf{Y} \boldsymbol{\mu}^{-1})) \mathbf{P} \phi_G(-\text{Log}_G^\vee(\mathbf{Y} \boldsymbol{\mu}^{-1}))^\top$ par $\mathbf{P}$ et l'équation (2.75) devient :

$$p(\mathbf{Y}) \simeq \frac{1}{\sqrt{(2\pi)^m |\mathbf{P}|}} \exp \left(-\frac{1}{2} \|\text{Log}_G^\vee(\boldsymbol{\mu}^{-1} \mathbf{Y})\|_{\mathbf{P}}^2 \right), \quad \forall \mathbf{Y} \in G \quad (2.77)$$

Un échantillon de cette distribution peut être généré à partir d'un vecteur aléatoire gaussien appartenant à $\mathbb{R}^m$ où m est la dimension du groupe. Le principe est le suivant : soit un vecteur distribué selon une loi gaussienne de moyenne nulle et de matrice de covariance $\mathbf{P}$, sa projection sur le groupe grâce à la fonction $\text{Exp}_G^\wedge(\cdot)$ crée une matrice aléatoire $\mathbf{X}$ dont la distribution correspondante est de paramètres $\mathbf{I}$ et $\mathbf{P}$.

Par l'action de groupe, il est possible de multiplier $\mathbf{X}$ à gauche ou à droite par un élément noté $\boldsymbol{\mu}$, afin d'en déduire une nouvelle variable $\mathbf{Y}$ de nouveau paramètre de moyenne $\boldsymbol{\mu}$.

→ La multiplication à gauche va induire que $\mathbf{Y} = \boldsymbol{\mu} \text{Exp}_G^\wedge(\boldsymbol{\epsilon})$ suit une *distribution gaussienne concentrée à gauche* telle que $\mathbf{Y} \sim \mathcal{N}_G^L(\boldsymbol{\mu}, \mathbf{P})$,

→ celle à droite que $\mathbf{Y} = \text{Exp}_G^\wedge(\boldsymbol{\epsilon}) \boldsymbol{\mu}$ suit une *distribution gaussienne concentrée à droite* telle que $\mathbf{Y} \sim \mathcal{N}_G^R(\boldsymbol{\mu}, \mathbf{P})$.

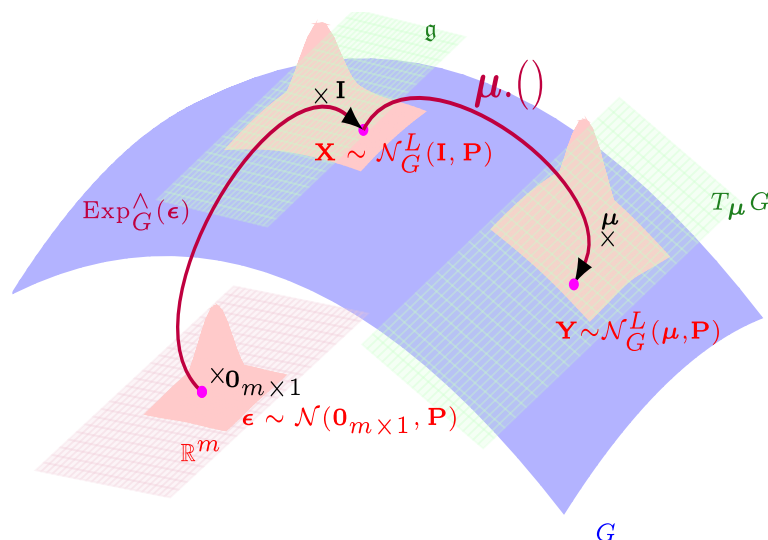


FIGURE 2.3 – Construction d'un échantillon $\mathbf{Y}$ selon une distribution gaussienne concentrée à gauche sur un groupe de Lie G à partir d'un échantillon gaussien sur $\mathbb{R}^m$.

2.3.3.2 Régions d'incertitude sur $SE(3)$

Une partie des développements présentés dans les chapitres suivants s'appuie sur les propriétés du groupe de Lie $SE(3)$, aussi étudions-nous particulièrement la loi gaussienne concentrée sur ce groupe. Nous nous intéressons plus particulièrement à un nuage de points généré selon une loi concentrée à droite car ce formalisme permet d'appréhender plus facilement l'influence des différents paramètres sur la forme du nuage. Supposons qu'un ensemble $\{\mathbf{Z}_i\}_{i=1}^N$ d'échantillons soient obtenus tels que :

$$\mathbf{Z}_i = \text{Exp}_{SE(3)}^{\wedge}(\boldsymbol{\epsilon}_i) \mathbf{M} \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \quad (2.78)$$

où :

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ \mathbf{0} & 1 \end{bmatrix} \quad (2.79)$$

et :

$$\text{Exp}_{SE(3)}^{\wedge}(\epsilon_i) = \begin{bmatrix} \mathbf{R}_i = \text{Exp}_{SO(3)}^{\wedge}(\mathbf{r}_i) & \mathbf{p}_i \\ \mathbf{0} & 1 \end{bmatrix} \quad (2.80)$$

Σ caractérise la variance en rotation et en translation sur l'espace vectoriel isomorphe à l'algèbre de Lie. $\mathbf{r}_i$ correspond à la partie rotationnelle de $\boldsymbol{\epsilon}_i$ et $\mathbf{p}_i$ à la projection sur le groupe de sa partie en translation.

Ce modèle permet d'identifier $\mathbf{M}$ comme le barycentre à *droite* du nuage de points $\{\mathbf{Z}_i\}_{i=1}^N$. Il est caractérisé par son orientation $\mathbf{R}$ et par sa position $\mathbf{p}$, et vérifie :

$$\sum_{i=1}^N \text{Log}_{SE(3)}^{\vee}(\mathbf{Z}_i \mathbf{M}^{-1}) \simeq \mathbf{0} \quad (2.81)$$

lorsque N tend vers l'infini.

En raison de l'action à droite de $\mathbf{M}$ sur chaque échantillon $\text{Exp}_{SE(3)}^\wedge(\boldsymbol{\epsilon}_i)$, la partie translation de $\mathbf{Z}_i$, notée $\mathbf{z}_i$ peut s'écrire de la manière suivante :

$$\mathbf{z}_i = \mathbf{R}_i \mathbf{p} + \mathbf{p}_i. \quad (2.82)$$

Chaque échantillon $\mathbf{z}_i$ est donc obtenu en pivotant la position du barycentre selon la rotation $\mathbf{R}_i$ puis en appliquant une translation de $\mathbf{p}_i$. Cet effet, dit de “bras de levier”, entraîne une dispersion spécifique des réalisations sous forme de “bananoïde”, comme déjà observée dans diverses études [64].

- Comme la rotation $SO(3)$ possède trois degrés de libertés, la courbure de la forme peut être définie selon trois axes. Elle est prépondérante selon un certain axe si la variance d'une des trois composantes rotationnelles de $\boldsymbol{\epsilon}_i$ est plus élevée que les deux autres.
- Son étendue dépend de l'ordre de grandeur de $\mathbf{p}$: en effet, plus ses coordonnées sont grandes, plus l'effet de bras de levier sera important.
- Sa largeur est caractérisée directement par le paramètre $\mathbf{p}_i$, puisque l'écart-type en translation sur l'algèbre de Lie permet de quantifier l'écart entre le centroïde $\mathbf{p}$ et les points du bord du nuage.

Sur les figures 2.4 et 2.5, sont représentées les réalisations des composantes en translation de $\{\mathbf{Z}_i\}_{i=1}^N$ avec $N = 1000$. Pour chacun de ces nuages, la position du barycentre est prise à $[10, 10, 10]$ m et la variance en translation fixée à 0.1^2m^2 . Sur la figure 2.4, la variance en rotation selon l'axe x est égale à 0.2^2rad^2 , alors que la variance en rotation selon l'axe x est supérieure sur la figure 2.5, égale à 0.5^2rad^2 .

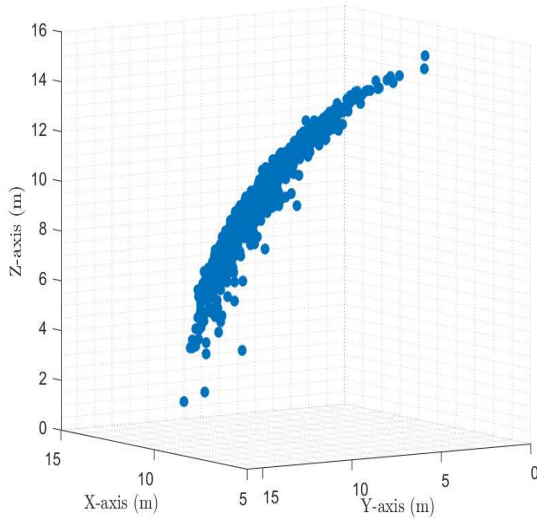


FIGURE 2.4 – Nuage de points observé pour $\mathbf{R}_i$ échantillonnée selon une distribution gaussienne concentrée sur $SO(3)$ de matrice de covariance $\boldsymbol{\Sigma}_R = \text{diag}([0.2^2, 0.1^2, 0.1^2])$.

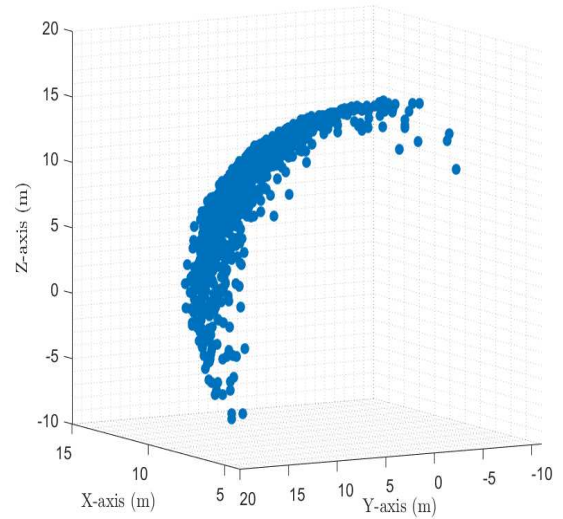


FIGURE 2.5 – Nuage de points observé pour $\mathbf{R}_i$ échantillonnée selon une distribution gaussienne concentrée sur $SO(3)$ de matrice de covariance $\boldsymbol{\Sigma}_R = \text{diag}([0.5^2, 0.1^2, 0.1^2])$.

2.4 Algorithmes d'estimation

2.4.1 Algorithmes d'optimisation

2.4.1.1 État de l'art

Un des aspects importants de notre travail est de pouvoir estimer des paramètres évoluant sur des groupes de Lie. Comme la résolution d'un problème d'estimation peut souvent être reformulée comme la minimisation d'un critère d'erreur, il apparaît naturel de s'intéresser à la manière d'optimiser des paramètres contraint par la structure de groupe.

De manière générale, les solutions analytiques sont difficiles à obtenir et il est souvent nécessaire de résoudre le problème en réalisant des approximations ou bien en utilisant des algorithmes itératifs.

Dans la littérature, nombre de travaux se sont intéressés à la manière de minimiser un critère sur groupe de Lie par l'utilisation de méthodes numériques [65]. L'objectif des algorithmes dédiés est de réaliser une mise à jour intrinsèque des paramètres tout en s'assurant que le critère à minimiser décroît bien au cours des itérations. Ils sont classiquement basés sur un calcul de direction de descente et construisent une suite d'itérés à partir d'une direction calculée sur l'espace tangent et reprojétée sur le groupe à travers une rétraction [49].

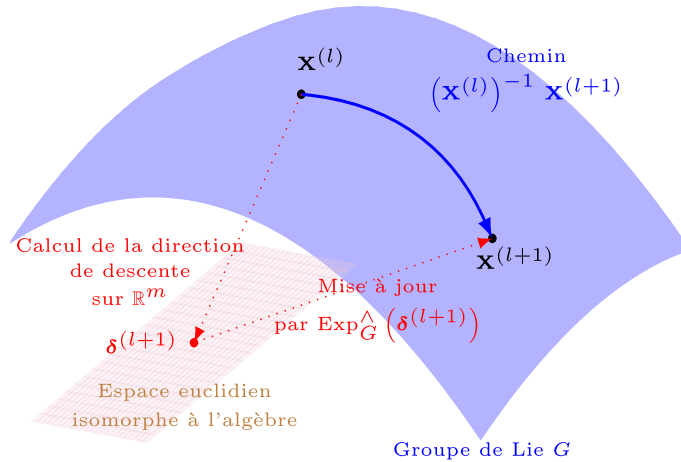


FIGURE 2.6 – Représentation géométrique de la mise à jour d'un paramètre par un algorithme d'optimisation sur groupe de Lie.

Une première possibilité est de mettre à profit la structure de variété riemannienne du groupe de Lie. La littérature sur ce sujet est riche [49]. Ainsi, des généralisations des algorithmes du gradient, de Newton et de Gauss-Newton ont été proposées dans [66][67][68]. Ces algorithmes trouvent initialement leur intérêt dans le domaine du traitement du son, des images texturées et des signaux biomédicaux où les paramètres inconnus présentent souvent des structures spécifiques. À titre d'exemple, nous pouvons citer l'espace des matrices de covariance pouvant caractériser un ensemble de données mesurées par un électroencéphalogramme [69].

Alternativement, il est possible de considérer des algorithmes dédiés aux groupes de Lie et qui prennent en compte leurs propriétés géométriques. Les premiers travaux d'optimisation dans ce contexte sont apparus dans la décennie 1990. Ils consistent à imposer que la solution recherchée vérifie les contraintes du groupe en utilisant des multiplicateurs de Lagrange [70].

Par la suite, des algorithmes de type Newton intrinsèque ont été développés, sur groupe de Lie compact dans un premier temps [65], puis pour des groupes non compacts dans un second temps [48][71]. Des approches de type Gauss-Newton ont plus récemment été proposées pour des familles de fonctions coût quadratique, dans un premier lieu, pour une application sur les groupes de Lie $SO(n)$ et $Sim(n)$ [72][73], puis sur des groupes de Lie plus génériques [63].

Le développement d'algorithmes sur groupe de Lie a trouvé sa motivation initialement dans le domaine de la séparation de sources [74]. En effet, les matrices à optimiser doivent souvent respecter des contraintes d'orthogonalité et appartiennent au groupe de Lie $SO(n)$ dans le cas où elles sont carrées. Plus récemment, les applications où des contraintes géométriques doivent être vérifiées sont devenues nombreuses. À titre d'exemple, nous pouvons citer les domaines de la vision par ordinateur, de la robotique ainsi que du médical. Dans ces exemples, il est d'intérêt de déterminer la pose d'une caméra ou d'un robot, ou encore le mouvement contraint d'une bactérie [75][76]. Ils peuvent être décrits comme des matrices appartenant aux groupes de Lie $SE(2)$ ou $SE(3)$ et estimées à partir de capteurs fournissant des mesures d'orientation et de translation.

Dans le cadre de nos travaux, nous désirons manipuler des groupes de Lie non compacts, notamment le groupe de Lie $SE(3)$. À cette fin, une attention particulière est accordée aux algorithmes de Newton et de Gauss-Newton. Les principes sont détaillés ci-dessous.

2.4.1.2 Algorithme de Newton sur groupe de Lie (LG-N)

Soit G un groupe de Lie non compact de dimension m et $J : G \rightarrow \mathbb{R}$ une fonction continue et différentiable à l'ordre deux, nous souhaitons déterminer le minimiseur de la fonction J :

$$\hat{\mathbf{X}} = \underset{\mathbf{X} \in G}{\operatorname{argmin}} J(\mathbf{X}) \quad (2.83)$$

Le processus itératif de Newton [48] consiste à construire la suite d'itérés $\{\mathbf{X}^{(l)}\}_{l \in \mathbb{N}}$ telle que :

$$\mathbf{X}^{(l+1)} = \mathbf{X}^{(l)} \operatorname{Exp}_G^\wedge(\boldsymbol{\delta}^{(l+1)}) \quad (2.84)$$

où $\boldsymbol{\delta}^{(l+1)} \in \mathbb{R}^m$ peut être calculée comme la quantité minimisant une approximation à l'ordre deux de la fonction J à l'itération l . Plus précisément,

$$\boldsymbol{\delta}^{(l+1)} = \underset{\boldsymbol{\delta} \in \mathbb{R}^m}{\operatorname{argmin}} \hat{J}(\mathbf{X}^{(l)} \operatorname{Exp}_G^\wedge(\boldsymbol{\delta})) \quad (2.85)$$

où $\hat{J}(\mathbf{X}^{(l)} \operatorname{Exp}_G^\wedge(\boldsymbol{\delta}))$ s'écrit :

$$\hat{J}(\mathbf{X}^{(l)} \operatorname{Exp}_G^\wedge(\boldsymbol{\delta})) = J(\mathbf{X}^{(l)}) + \left(\mathcal{L}_{J(\mathbf{X}^{(l)})}^R \right)^\top \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^\top \mathbf{H}(\mathbf{X}^{(l)}) \boldsymbol{\delta} \quad (2.86)$$

avec $\mathcal{L}_{J(\mathbf{X}^{(l)})}^R$ et $\mathbf{H}(\mathbf{X}^{(l)})$ correspondant respectivement aux dérivées d'ordre un et deux de J en $\mathbf{X}^{(l)}$ et dont les expressions ont été préalablement définies dans les équations (2.40) et (2.57). Si la hessienne est inversible, nous pouvons en déduire que la solution à ce problème de minimisation s'écrit sous la forme :

$$\boldsymbol{\delta}^{(l+1)} = -(\mathbf{H}(\mathbf{X}^{(l)}))^{-1} \mathcal{L}_{J(\mathbf{X}^{(l)})}^R. \quad (2.87)$$

Dans [48], il est prouvé que cette suite converge quadratiquement, sous certaines conditions [71], vers un point critique μ de J , vérifiant :

$$\mathcal{L}_{J(\mu)}^R = 0 \quad (2.88)$$

2.4.1.3 Algorithme de Gauss-Newton sur groupe de Lie (LG-GN)

L'algorithme de Gauss-Newton peut être vu comme un cas particulier de l'algorithme de Newton lorsque le critère à minimiser s'écrit sous la forme d'une somme de fonctions au carré. Cette méthode présente l'avantage de ne pas nécessiter le calcul de dérivées secondes.

Supposons une fonction $J : G \rightarrow \mathbb{R}$, continue et différentiable, définie selon la forme positive suivante :

$$J(\mathbf{X}) = \|\varepsilon(\mathbf{X})\|_{\Sigma}^2 \quad (2.89)$$

où G est de dimension m , $\varepsilon : G \rightarrow \mathbb{R}^d$ ($d \geq 1$) est une fonction continue et différentiable, et Σ est une matrice définie positive.

Nous cherchons à construire une suite d'itérés $\{\mathbf{X}^{(l)}\}_{l \in \mathbb{N}}$ sous la même forme que précédemment. Par un développement limité à l'ordre 1 de ε , la direction de descente $\delta^{(l+1)}$ est également calculée de manière optimale et correspond à la solution du problème d'optimisation :

$$\delta^{(l+1)} = \underset{\delta \in \mathbb{R}^m}{\operatorname{argmin}} \|\varepsilon(\mathbf{X}^{(l)}) + \mathbf{J}_{\varepsilon}^{(l)} \delta\|_{\Sigma}^2 \quad (2.90)$$

où $\mathbf{J}_{\varepsilon}^{(l)}$ correspond à la matrice jacobienne à droite de ε :

$$\mathbf{J}_{\varepsilon} = \left. \frac{\partial \varepsilon(\mathbf{X}^{(l)} \operatorname{Exp}_G^{\wedge}(\delta))}{\partial \delta} \right|_{\delta=0} \quad (2.91)$$

Nous pouvons alors montrer par les opérations de dérivations usuelles qu'elle s'écrit :

$$\delta^{(l+1)} = -(\mathbf{J}_{\varepsilon}^{\top(l)} \Sigma^{-1} \mathbf{J}_{\varepsilon}^{(l)})^{-1} \mathbf{J}_{\varepsilon}^{\top(l)} \Sigma^{-1} \varepsilon(\mathbf{X}^{(l)}). \quad (2.92)$$

Tout comme l'algorithme de Newton, cette méthode converge sous certaines conditions [72] vers un point critique de J .

2.4.2 Algorithmes de filtrage

2.4.2.1 État de l'art

Pour estimer des paramètres dynamiques à partir d'observations arrivant séquentiellement, il est nécessaire d'utiliser des algorithmes de filtrage récursif. Dans un cadre classique euclidien, ces algorithmes s'inscrivent généralement dans un formalisme bayésien en intégrant une information *a priori* sur les paramètres. Ils fournissent des intervalles de confiance sur les estimateurs obtenus. Lorsque les grandeurs d'intérêts évoluent sur groupe de Lie, il existe également dans la littérature des méthodes de filtrage permettant d'intégrer dynamiquement des contraintes *a priori* sur le groupe.

Comme pour les algorithmes d'optimisation, les premières méthodes d'estimation récursive ont été proposées sur variété riemannienne [48]. Elles ont été en particulier mises à profit dans des problèmes où des matrices de covariances sont inconnues et ont besoin d'être estimées récursivement tout en conservant leur caractère semi-défini positif. Ainsi, un filtre particulière sur variété riemannienne est développé dans [77] qui permet d'estimer un état évoluant sur espace euclidien conjointement avec les matrices de covariance des bruits de modèles. Il existe également des méthodes séquentielles permettant de prendre en compte les contraintes d'orthogonalité des paramètres dynamiques : à cette fin un filtre particulière sur variété de Stiefel est proposé dans [57].

Dans le cas spécifique des groupes de Lie, les travaux en filtrage sont plus nombreux en raison de leurs vastes applications, que ce soit en vision par ordinateur, pour estimer récursivement la pose d'une caméra ou d'un robot, ou en navigation inertielle, pour déterminer les orientations locales d'un mobile en déplacement. Des solutions algorithmiques généralisant les équations classiques du filtre de Kalman ont notamment été proposées. L'objectif de ces méthodes est ainsi d'approcher les modèles non-linéaires à l'aide d'un développement limité à l'ordre un sur groupe de Lie afin d'obtenir des expressions analytiques et aisées à mettre en oeuvre.

Dans l'article fondateur [78], un filtre de Kalman est présenté. Il repose sur un modèle d'état directement défini sur groupe de Lie. Ce filtre est intrinsèque au sens où, à chaque instant, les paramètres estimés appartiennent au groupe. Nommé *Invariant Kalman Filtering*, il permet d'estimer des paramètres dont les modèles d'évolution respectent des propriétés d'invariance et de symétrie. Il est à noter en particulier que les équations de ce filtre sont obtenues en minimisant une erreur d'estimation intrinsèque et ne sont pas établies dans un cadre bayésien. Ainsi, cet algorithme ne vise pas explicitement à calculer des distributions *a posteriori*. Il a en particulier prouvé son efficacité pour estimer les orientations locales d'un objet suivi par un système radar [79].

Il existe également des filtres fondés sur un formalisme probabiliste bayésien. En particulier, dans [63], un filtre de Kalman étendu ainsi qu'un filtre de Kalman étendu itéré sur groupe de Lie (*Extended Kalman Filtering* et *Iterated Extended Kalman Filtering*) sont développés à partir des distributions gaussiennes concentrées, de manière analogue aux *FKE* et *FKEI* standards présentés au chapitre 1. Dans ce cadre, un filtre d'information sur groupe de Lie est également décrit dans [80].

Des méthodes basées sur des transformées sans parfum ou sur des approches particulières ont finalement été proposées : en particulier, un filtre sans parfum sur groupe de Lie est décrit dans [81] et un filtre particulière intrinsèque est mis en oeuvre dans [82], dans le cadre spécifique du groupe de Lie $SE(3)$.

Dans la suite du paragraphe, nous proposons de détailler les approches analytiques fondées sur le filtre de Kalman, qui nous serviront de références pour la suite de nos travaux.

2.4.2.2 Filtre de Kalman invariant (IKF)

Soit G un groupe de Lie, et $\mathbf{X}(t) \in G \subset \mathbb{R}^{n \times n}$ la matrice contenant l'ensemble des paramètres à estimer à l'instant t qui est supposée satisfaire le modèle d'évolution suivant :

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{f}_t(\mathbf{X}(t)) + \mathbf{X}(t) \mathbf{W}(t) \quad (2.93)$$

où $\mathbf{f}_t : G \rightarrow T_{\mathbf{X}(t)}G$ est une fonction non linéaire de $\mathbf{X}(t)$ et $\mathbf{W}(t)$ est une matrice de bruit aléatoire appartenant à $T_{\mathbf{I}}G$. Il est également requis que la fonction $\mathbf{f}_t$ vérifie la propriété d'invariance suivante :

$$\mathbf{f}_t(\mathbf{A} \mathbf{B}) = \mathbf{A} \mathbf{f}_t(\mathbf{B}) - \mathbf{B} \mathbf{f}_t(\mathbf{A}) + \mathbf{f}_t(\mathbf{I}) \mathbf{B}, \quad \forall (\mathbf{A}, \mathbf{B}) \in G \times G. \quad (2.94)$$

Supposons qu'une observation euclidienne $\mathbf{y}_k \in \mathbb{R}^n$ soit disponible à l'instant de mesure t_k , sous la forme linéaire suivante :

$$\mathbf{y}_k = \mathbf{X}(t_k) \mathbf{d}_k + \mathbf{n}_k \quad (2.95)$$

où $\mathbf{d}_k \in \mathbb{R}^n$ est un vecteur connu représentant une entrée du modèle et $\mathbf{n}_k$ est un bruit blanc gaussien.

L'objectif de l'*IKF* est d'estimer récursivement $\mathbf{X}_k = \mathbf{X}(t_k)$ à partir des l'ensemble des observations de l'instant initial à l'instant courant $\{\mathbf{y}_1, \dots, \mathbf{y}_k\}$. Pour y parvenir, l'algorithme procède en deux étapes :

→ l'étape de prédiction consiste à déterminer, à chaque instant t , l'état $\hat{\mathbf{X}}(t)^-$ vérifiant l'équation différentielle suivante :

$$\frac{d\hat{\mathbf{X}}(t)^-}{dt} = \mathbf{f}_t(\hat{\mathbf{X}}(t)^-) \quad (2.96)$$

sous la condition initiale $\hat{\mathbf{X}}(t_{k-1}) = \hat{\mathbf{X}}_{k-1|k-1}$.

→ A l'instant de mesure t_k , l'étape de correction fournit l'état estimé $\hat{\mathbf{X}}_{k|k}$, en supposant qu'il s'écrit en fonction de l'état prédit $\hat{\mathbf{X}}_{k|k-1} = \hat{\mathbf{X}}(t_k)^-$ sous la forme :

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} \text{Exp}_G^\wedge \left(\mathbf{K}_k \left(\underbrace{\hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{y}_k - \mathbf{d}_k}_{\text{Vecteur d'innovation}} \right) \right) \quad (2.97)$$

avec $\mathbf{K}_k$ une matrice de gain pouvant être déterminée grâce à l'équation de Riccati [83]. Celle-ci est fondée sur l'étude de la propagation au fil des itérations des erreurs d'estimation dans l'algèbre de Lie. Soit la matrice d'erreur intrinsèque à gauche entre l'état réel et l'estimé recherché définie par :

$$\boldsymbol{\eta}_{k|k}^L = \mathbf{X}_k^{-1} \hat{\mathbf{X}}_{k|k}. \quad (2.98)$$

D'après l'équation précédente, il est possible de montrer que :

$$\boldsymbol{\eta}_{k|k}^L = \boldsymbol{\eta}_{k|k-1}^L \text{Exp}_G^\wedge \left(\mathbf{K}_k \left(\left(\boldsymbol{\eta}_{k|k-1}^L \right)^{-1} \mathbf{d}_k - \mathbf{d}_k + \hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{n}_k \right) \right) \quad (2.99)$$

où $\boldsymbol{\eta}_{k|k-1}^L$ est la matrice d'erreur de prédiction $\mathbf{X}_k^{-1} \hat{\mathbf{X}}_{k|k-1}$. En utilisant le fait que $\boldsymbol{\eta}_{k|k-1}^L$ est suffisamment petit, il est ensuite possible d'écrire que :

$$\boldsymbol{\eta}_{k|k-1}^L = \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_{k|k-1}) \quad (2.100)$$

Pareillement, en notant $\boldsymbol{\eta}_{k|k}^L = \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_{k|k})$, il s'ensuit que :

$$\boldsymbol{\epsilon}_{k|k} \simeq \boldsymbol{\epsilon}_{k|k-1} + \mathbf{K}_k \left(\left(\boldsymbol{\eta}_{k|k-1}^L \right)^{-1} \mathbf{d}_k - \mathbf{d}_k + \hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{n}_k \right). \quad (2.101)$$

Une linéarisation à l'ordre un de $\left(\boldsymbol{\eta}_{k|k-1}^L \right)^{-1}$ sous la forme :

$$\left(\boldsymbol{\eta}_{k|k-1}^L \right)^{-1} \simeq \mathbf{I} + [\boldsymbol{\epsilon}_{k|k-1}]_G^\wedge \quad (2.102)$$

fournit alors que :

$$\boldsymbol{\epsilon}_{k|k} \simeq \boldsymbol{\epsilon}_{k|k-1} + \mathbf{K}_k \left(-[\boldsymbol{\epsilon}_{k|k-1}]_G^\wedge \mathbf{d}_k + \hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{n}_k \right) \quad (2.103)$$

Par conséquent, l'équation de propagation de l'erreur s'écrit :

$$\boldsymbol{\epsilon}_{k|k} \simeq \boldsymbol{\epsilon}_{k|k-1} + \mathbf{K}_k \left(-[\boldsymbol{\epsilon}_{k|k-1}]_G^\wedge \mathbf{d}_k + \hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{n}_k \right) \quad (2.104)$$

et la matrice de covariance de l'erreur peut alors être calculée, conjointement avec le gain, selon l'équation de Riccati.

2.4.2.3 Filtre de Kalman sur groupe de Lie : approche bayésienne

Les filtres de Kalman sur groupe de Lie proposés dans [63] et [84] se fondent sur une approche bayésienne, de manière analogue au *FKE* classique. Plus précisément, l'objectif de ces méthodes est d'obtenir une expression récursive de la distribution *a posteriori* des paramètres d'intérêt en l'approchant par une distribution gaussienne concentrée. Elles s'appliquent à une large classe de modèles et sont adaptées pour des observations prenant leurs valeurs sur un groupe de Lie.

Tout comme dans le cas classique euclidien, nous pouvons en différencier deux principales :

- Le filtre de Kalman étendu sur groupe de Lie (*LG-FKE*), approchant la distribution *a posteriori* en linéarisant les modèles au premier ordre.
- Le filtre de Kalman étendu itéré (*LG-FKEI*) pour lequel les paramètres de la distribution gaussienne concentrée sont obtenus par résolution d'un problème d'optimisation sur groupe de Lie.

Contrairement à l'IKF, ces algorithmes sont formulés à temps discret.

• Filtre de Kalman étendu sur groupe de Lie

Considérons une matrice aléatoire $\mathbf{X}_k$ évoluant au cours du temps sur un groupe de Lie G de dimension m . De plus, considérons une observation $\mathbf{Z}_k$ appartenant à un groupe de Lie G' de dimension m' .

Le *LG-FKE* se fonde sur les modèles dynamiques et d'observation suivants :

$$\mathbf{X}_k = \mathbf{X}_{k-1} \text{Exp}_G^\wedge(\boldsymbol{\Omega}(\mathbf{X}_{k-1}) + \mathbf{v}_k), \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \quad (2.105)$$

$$\mathbf{Z}_k = \mathbf{h}_k(\mathbf{X}_k) \text{Exp}_{G'}^\wedge(\mathbf{n}_k), \quad \mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k) \quad (2.106)$$

où $\boldsymbol{\Omega} : G \rightarrow \mathbb{R}^m$ et $\mathbf{h}_k : G \rightarrow G'$ sont deux fonctions différentiables et potentiellement non linéaires.

L'objectif est d'approcher récursivement la distribution *a posteriori* $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$ à chaque instant k par une distribution gaussienne concentrée sous la forme :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \mathcal{N}_G^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k|k}, \mathbf{P}_{k|k}) \quad (2.107)$$

Par conséquent, ce problème revient à déterminer récursivement les paramètres $\{\hat{\mathbf{X}}_{k|k}, \mathbf{P}_{k|k}\}$. Il est à noter qu'il est également possible de décrire cet algorithme par l'utilisation d'une distribution gaussienne à *droite*. Par souci de cohérence, nous garderons par la suite le formalisme à *gauche*.

Au même titre que le *FKE*, le *LG-FKE* procède classiquement en deux étapes :

→ l'étape de prédiction consiste à approcher la distribution prédite par la distribution $\mathcal{N}_G^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k|k-1}, \mathbf{P}_{k|k-1})$ construite à partir de la distribution précédente $\mathcal{N}_G^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ et du modèle d'évolution (2.105).

Le moment du premier ordre $\hat{\mathbf{X}}_{k|k-1}$ est déterminé en propageant l'état précédent selon le modèle d'évolution non bruité :

$$\hat{\mathbf{X}}_{k|k-1} = \hat{\mathbf{X}}_{k-1|k-1} \text{Exp}_G^\wedge(\boldsymbol{\Omega}(\hat{\mathbf{X}}_{k-1|k-1})) \quad (2.108)$$

Le moment du second ordre $\mathbf{P}_{k|k-1}$ est mis à jour en mettant à profit la relation liant les erreurs précédente et prédite, projetées sur l'espace vectoriel isomorphe à l'algèbre. Ces deux erreurs s'écrivent respectivement :

$$\begin{cases} \boldsymbol{\epsilon}_{k-1|k-1} = \text{Log}_G^\vee(\hat{\mathbf{X}}_{k-1|k-1}^{-1} \mathbf{X}_k) \\ \boldsymbol{\epsilon}_{k|k-1} = \text{Log}_G^\vee(\hat{\mathbf{X}}_{k|k-1}^{-1} \mathbf{X}_k) \end{cases} \quad (2.109)$$

$$(2.110)$$

En utilisant la première formule BCH (2.33), il est montré qu'elles sont liées par la formule suivante :

$$\boldsymbol{\epsilon}_{k|k-1} = \mathbf{F}_k \boldsymbol{\epsilon}_{k-1|k-1} + \phi_G(-\boldsymbol{\Omega}(\hat{\mathbf{X}}_{k-1|k-1})) \mathbf{v}_k + O(\|\mathbf{v}_k\|^2, \|\boldsymbol{\epsilon}_{k-1|k-1}\|^2) \quad (2.111)$$

où :

$$\begin{cases} \mathbf{F}_k = \text{Ad}_G(\text{Exp}_G^\wedge(\boldsymbol{\Omega}(-\hat{\mathbf{X}}_{k-1|k-1}))) + \phi_G(\boldsymbol{\Omega}(-\hat{\mathbf{X}}_{k-1|k-1})) \mathbf{C}_k, \\ \mathbf{C}_k = \left. \frac{\partial \boldsymbol{\Omega}(\hat{\mathbf{X}}_{k-1|k-1} \text{Exp}_G^\wedge(\boldsymbol{\delta}))}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta}=\mathbf{0}} \end{cases} \quad (2.112)$$

$$(2.113)$$

Comme $\mathbf{P}_{k|k-1}$ correspond à la matrice de covariance de $\boldsymbol{\epsilon}_{k|k-1}$, nous obtenons, en négligeant les termes d'ordre supérieur à deux :

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \phi_G(-\boldsymbol{\Omega}(\hat{\mathbf{X}}_{k-1|k-1})) \mathbf{Q}_k \phi_G(-\boldsymbol{\Omega}(\hat{\mathbf{X}}_{k-1|k-1}))^\top \quad (2.114)$$

→ L'étape de mise à jour consiste à approcher la distribution *a posteriori* à partir de l'approximation de la distribution prédite. Pour y parvenir, une mise à jour de l'erreur sur l'algèbre de Lie est obtenue par linéarisation à l'ordre un du modèle d'observation.

Après projection sur l'espace euclidien localement isomorphe à l'algèbre de Lie de (2.106), il vient que :

$$\tilde{\mathbf{z}}_k = \mathbf{H}_k \boldsymbol{\epsilon}_{k|k-1} + \mathbf{n}_k \quad (2.115)$$

avec :

$$\left\{ \begin{array}{l} \mathbf{H}_k = - \frac{\partial \text{Log}_G^\vee(\mathbf{h}_k(\hat{\mathbf{X}}_{k|k-1} \text{Exp}_G^\wedge(\boldsymbol{\delta}))^{-1} \mathbf{Z}_k)}{\partial \boldsymbol{\delta}} \Big|_{\boldsymbol{\delta}=\mathbf{0}} \\ \tilde{\mathbf{z}}_k = \text{Log}_G^\vee(\mathbf{h}_k(\hat{\mathbf{X}}_{k|k-1})^{-1} \mathbf{Z}_k) \end{array} \right. \quad (2.116)$$

En mettant à profit le fait que $p(\tilde{\mathbf{z}}_k | \boldsymbol{\epsilon}_{k|k-1}, \mathbf{Z}_1, \dots, \mathbf{Z}_{k-1}) = \mathcal{N}(\tilde{\mathbf{z}}_k; \mathbf{H}_k \boldsymbol{\epsilon}_{k|k-1}, \mathbf{R}_k)$ et que $p(\boldsymbol{\epsilon}_{k|k-1} | \mathbf{Z}_1, \dots, \mathbf{Z}_{k-1}) = \mathcal{N}(\boldsymbol{\epsilon}_{k|k-1}; \mathbf{0}, \mathbf{P}_{k|k-1})$, nous pouvons en déduire que :

$$p(\boldsymbol{\epsilon}_{k|k-1} | \mathbf{Z}_1, \dots, \mathbf{Z}_k) = \mathcal{N}(\boldsymbol{\epsilon}_{k|k-1}; \mathbf{m}_{k|k}, \mathbf{P}_{k|k}^-) \quad (2.118)$$

avec :

$$\left\{ \begin{array}{l} \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\ \mathbf{m}_{k|k} = \mathbf{K}_k \tilde{\mathbf{z}}_k \\ \mathbf{P}_{k|k}^- = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \end{array} \right. \quad (2.119)$$

$$\mathbf{m}_{k|k} = \mathbf{K}_k \tilde{\mathbf{z}}_k \quad (2.120)$$

$$\mathbf{P}_{k|k}^- = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \quad (2.121)$$

Afin d'obtenir l'approximation gaussienne concentrée désirée, il est nécessaire de reparamétriser l'erreur afin qu'elle soit centrée. En effet, à cette étape du raisonnement, l'état vrai et l'état estimé sont reliés comme suit :

$$\mathbf{X}_k = \hat{\mathbf{X}}_{k|k-1} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_{k|k-1}) \quad (2.122)$$

où $\boldsymbol{\epsilon}_{k|k-1}$ est à moyenne non nulle. En se servant du fait que, $\boldsymbol{\epsilon}_{k|k-1} = \mathbf{m}_{k|k} + \boldsymbol{\epsilon}_{k|k}$ où $\boldsymbol{\epsilon}_{k|k}$ est un vecteur aléatoire gaussien centré, il est possible d'écrire que :

$$\mathbf{X}_k = \hat{\mathbf{X}}_{k|k} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_{k|k}) \quad \text{avec} \quad \boldsymbol{\epsilon}_{k|k} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{k|k}) \quad (2.123)$$

où :

$$\left\{ \begin{array}{l} \hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} \text{Exp}_G^\wedge(\mathbf{m}_{k|k}), \\ \mathbf{P}_{k|k} = \phi_G(-\mathbf{m}_{k|k}) \mathbf{P}_{k|k}^- \phi_G(-\mathbf{m}_{k|k})^\top. \end{array} \right. \quad (2.124)$$

$$\mathbf{P}_{k|k} = \phi_G(-\mathbf{m}_{k|k}) \mathbf{P}_{k|k}^- \phi_G(-\mathbf{m}_{k|k})^\top. \quad (2.125)$$

Comme $\boldsymbol{\epsilon}_{k|k}$ est centrée, nous en déduisons que :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \mathcal{N}_G^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k|k}, \mathbf{P}_{k|k}). \quad (2.126)$$

• Filtre de Kalman étendu itéré

Tout comme le *LG-FKE*, cet algorithme se fonde sur une approximation gaussienne concentrée de la distribution *a posteriori*. Néanmoins, la linéarisation des modèles s'effectue de manière itérative en reformulant les deux étapes de l'algorithme comme deux problèmes d'optimisation selon le même principe que le *FKEI* standard détaillé dans le chapitre 1.

Ainsi, à chaque étape, un critère d'optimisation est défini et minimisé, revenant à déterminer le maximiseur des distributions *a posteriori*. En mettant à profit que les critères s'écrivent comme la norme d'une fonction vectorielle, il est possible de les optimiser en appliquant un algorithme de Gauss-Newton, tel que décrit dans la sous-section (2.4.1). La solution obtenue peut alors être utilisée afin d'approcher le moment d'ordre deux selon une méthode de Gauss-Laplace sur groupe de Lie.

Pour développer les équations de l'algorithme, le modèle d'évolution est supposé s'écrire sous la forme suivante :

$$\mathbf{X}_k = \mathbf{f}_k(\mathbf{X}_{k-1}) \text{Exp}_G^\wedge(\mathbf{v}_k) \quad \text{avec } \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \quad (2.127)$$

et le modèle d'observation vérifie l'équation (2.106).

L'algorithme comporte à chaque instant les deux étapes classiques d'un filtre.

→ L'objectif de l'étape de prédiction est d'approcher la distribution prédite $p(\mathbf{X}_k | \mathbf{Z}_{1:k-1})$, par une loi gaussienne concentrée notée $\mathcal{N}_{G'}^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k|k-1}, \mathbf{P}_{k|k-1})$.

Suivant [63] et [85], cela est équivalent à approcher $p(\mathbf{X}_k, \mathbf{X}_{k-1} | \mathbf{Z}_{1:k-1})$ par une distribution gaussienne concentrée puis à marginaliser selon $\mathbf{X}_{k-1}$. Le moment d'ordre un de cette distribution jointe, assimilé à son maximiseur, est calculé en minimisant le critère suivant :

$$\begin{aligned} \{\mathbf{X}_{k|k-1}^*, \mathbf{X}_{k-1|k-1}^*\} = \underset{\mathbf{X}, \tilde{\mathbf{X}}}{\operatorname{argmin}} \quad & \left\| \text{Log}_G^\vee \left(\mathbf{f}_k(\tilde{\mathbf{X}})^{-1} \mathbf{X} \right) \right\|_{\mathbf{Q}_k}^2 \\ & + \left\| \text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k-1|k-1} \right)^{-1} \tilde{\mathbf{X}} \right) \right\|_{\mathbf{P}_{k-1|k-1}}^2, \end{aligned} \quad (2.128)$$

où $\hat{\mathbf{X}}_{k-1|k-1}$ et $\mathbf{P}_{k-1|k-1}$ sont les paramètres de la distribution *a posteriori* estimée à l'instant précédent $\hat{p}(\mathbf{X}_{k-1} | \mathbf{z}_{1:k-1}) = \mathcal{N}_G^L(\mathbf{X}_{k-1}; \hat{\mathbf{X}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$.

La solution $\mathbf{X}_{k-1|k-1}^*$ est triviale et est égale à l'état estimé précédent $\hat{\mathbf{X}}_{k-1|k-1}$.

Concernant $\mathbf{X}_{k|k-1}^*$, il peut être obtenu en propageant l'état estimé précédent à travers l'équation dynamique privée de bruit :

$$\mathbf{X}_{k|k-1}^* = \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}). \quad (2.129)$$

Il est utilisé comme premier paramètre de la distribution prédite : $\hat{\mathbf{X}}_{k|k-1} = \mathbf{X}_{k|k-1}^*$.

Une approximation de Gauss-Laplace sur groupe de Lie de la distribution jointe, comme expliqué en annexe (B.3.2.1), permet d'obtenir une approximation de sa matrice de covariance.

Il est alors possible de montrer qu'elle s'écrit sous la forme suivante :

$$\mathbf{P}_k^* = \left(\begin{bmatrix} \mathbf{I} & \mathbf{F}_k \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^\top \Sigma_p^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{F}_k \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right)^{-1} \quad (2.130)$$

où $\mathbf{F}_k$ correspond à la matrice jacobienne à droite de $\mathbf{f}_k$ calculée en $\hat{\mathbf{X}}_{k-1|k-1}$ et obtenue selon l'équation (2.38) :

$$\mathbf{F}_k = \left. \frac{\partial (\text{Log}_G^\vee(\mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}))^{-1} \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1} \text{Exp}_G^\wedge(\boldsymbol{\delta})))}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta}=\mathbf{0}} \quad (2.131)$$

et

$$\boldsymbol{\Sigma}_p = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k-1|k-1} \end{bmatrix}. \quad (2.132)$$

La matrice de covariance de la distribution prédite est alors obtenue en conservant le bloc gauche-haut de $\mathbf{P}_k^*$. Nous en déduisons :

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k \quad (2.133)$$

→ L'étape de correction permet de déterminer une approximation de la distribution *a posteriori* fournie par :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \mathcal{N}_G^L(\mathbf{X}_k; \hat{\mathbf{X}}_{k|k}, \mathbf{P}_{k|k}) \quad (2.134)$$

$\hat{\mathbf{X}}_{k|k}$ est à nouveau obtenu en calculant le maximiseur de $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$, puis une approximation de Gauss-Laplace est effectuée pour en déduire $\mathbf{P}_{k|k}$.

Le critère d'optimisation est déterminé en utilisant la règle de Bayes. En effet, nous avons :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) = \frac{p(\mathbf{Z}_k | \mathbf{X}_k) p(\mathbf{X}_k | \mathbf{Z}_{1:k-1})}{p(\mathbf{Z}_k | \mathbf{Z}_{1:k-1})}. \quad (2.135)$$

L'anti-logarithme de $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$ peut alors s'écrire :

$$-2 \log p(\mathbf{X}_k | \mathbf{Z}_{1:k}) = C + \left\| \text{Log}_{G'}^\vee \left((\mathbf{h}_k(\mathbf{X}_k))^{-1} \mathbf{Z}_k \right) \right\|_{\mathbf{R}_k}^2 + \left\| \text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1} \right)^{-1} \mathbf{X}_k \right) \right\|_{\mathbf{P}_{k|k-1}}^2, \quad (2.136)$$

où $C \in \mathbb{R}$ est une constante.

Par conséquent, le paramètre $\hat{\mathbf{X}}_{k|k}$ est obtenu en résolvant le problème d'optimisation suivant :

$$\hat{\mathbf{X}}_{k|k} = \underset{\mathbf{X} \in G}{\text{argmin}} \left\| \text{Log}_{G'}^\vee \left((\mathbf{h}_k(\mathbf{X}))^{-1} \mathbf{Z}_k \right) \right\|_{\mathbf{R}_k}^2 + \left\| \text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1} \right)^{-1} \mathbf{X} \right) \right\|_{\mathbf{P}_{k|k-1}}^2, \quad (2.137)$$

grâce à l'algorithme de Gauss-Newton sur groupe de Lie. Comme expliqué dans la partie (2.4.1.3), à chaque itération l de l'algorithme, la mise à jour de l'état $\mathbf{X}_k^{(l)}$ s'écrit :

$$\mathbf{X}_k^{(l+1)} = \mathbf{X}_k^{(l)} \text{Exp}_G^\wedge(\boldsymbol{\delta}^{(l+1)}) \quad (2.138)$$

avec $\boldsymbol{\delta}^{(l+1)}$ calculée selon l'équation (2.92) :

$$\boldsymbol{\delta}^{(l+1)} = - \left(\mathbf{J}^{(l)\top} \boldsymbol{\Sigma}_c^{-1} \mathbf{J}^{(l)} \right)^{-1} \mathbf{J}^{(l)\top} \boldsymbol{\Sigma}_c^{-1} \mathbf{e} \left(\mathbf{X}_k^{(l)} \right) \quad (2.139)$$

dans laquelle :

$$\begin{cases} \Sigma_c = \begin{bmatrix} \mathbf{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k|k-1} \end{bmatrix}, \\ \mathbf{J}^{(l)} = \left[\frac{\partial \left(\text{Log}_{G'}^\vee \left(\mathbf{h}_k(\mathbf{X}_k^{(l)}) \text{Exp}_G^\wedge(\boldsymbol{\delta}) \right)^{-1} \mathbf{Z}_k \right)}{\partial \boldsymbol{\delta}} \right]_{\boldsymbol{\delta}=\mathbf{0}}^\top, \frac{\partial \left(\text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1} \right)^{-1} \mathbf{X}_k^{(l)} \text{Exp}_G^\wedge(\boldsymbol{\delta}) \right)}{\partial \boldsymbol{\delta}} \right) \Big|_{\boldsymbol{\delta}=\mathbf{0}}^\top \right]^\top, \\ \mathbf{e}(\mathbf{X}_k^{(l)}) = \left[\text{Log}_{G'}^\vee \left(\left(\mathbf{h}_k(\mathbf{X}_k^{(l)}) \right)^{-1} \mathbf{Z}_k \right)^\top, \text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1} \right)^{-1} \mathbf{X}_k^{(l)} \right)^\top \right]^\top. \end{cases}$$

Il est intéressant de remarquer que chaque itération de cette procédure d'optimisation revient à réaliser une mise à jour de l'état à la façon d'un filtre de Kalman étendu mais en utilisant à chaque fois un nouveau point de linéarisation, raffinant ainsi l'approximation de l'équation d'observation. Par conséquent, $\hat{\mathbf{X}}_{k|k}$ est estimée par la valeur de $\mathbf{X}_k^{(L)}$ où L correspond à la dernière itération de l'algorithme de Gauss-Newton.

En réalisant un développement limité à l'ordre un de $\mathbf{e}(\mathbf{X}_k)$ en $\hat{\mathbf{X}}_{k|k}$, il est obtenu l'approximation suivante de la distribution *a posteriori* :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \propto \exp \left(-\frac{1}{2} \left\| \text{Log}_G^\vee \left(\left(\hat{\mathbf{X}}_{k|k} \right)^{-1} \mathbf{X}_k \right) \right\|_{2(\mathbf{H}^{(L)})^{-1}}^2 \right),$$

avec $\mathbf{H}^{(L)} = (\mathbf{J}^{(L)\top} \Sigma_c^{-1} \mathbf{J}^{(L)})$. Par identification avec l'expression d'une distribution gaussienne concentrée, le paramètre de covariance est alors donné par :

$$\mathbf{P}_{k|k} = 2(\mathbf{H}^{(L)})^{-1}.$$

Par conséquent, l'approximation gaussienne concentrée s'écrit finalement :

$$p(\mathbf{X}_k | \mathbf{Z}_{1:k}) \simeq \mathcal{N}_G^L(\mathbf{X}_k; \mathbf{X}_k^{(L)}, 2(\mathbf{H}^{(L)})^{-1}). \quad (2.140)$$

2.5 Conclusions du chapitre

Dans ce chapitre, les principales notions théoriques et algorithmiques sur les groupes de Lie nécessaires pour introduire les contributions de ce travail de thèse ont été présentées. Nous nous sommes ainsi attachés à présenter les différents concepts avec le niveau suffisant de détails pour pouvoir les manipuler rigoureusement sans trop alourdir la lecture.

En effet, les notions introduites sont d'intérêt majeur car elles vont nous servir de références dans le chapitre suivant, pour proposer un modèle ainsi que deux algorithmes sur groupe de Lie permettant de caractériser et de pister un amas de débris spatiaux.

Par ailleurs, les éléments d'intégration et de dérivation vont nous permettre par la suite de proposer une borne de Cramér-Rao bayésienne sur groupe de Lie.

Chapitre 3

Modèles et algorithmes sur groupe de Lie pour le pistage d'un amas de débris spatiaux

Sommaire

3.1	Modélisation de l'amas comme une cible étendue sur groupe de Lie	83
3.1.1	Démarche proposée	83
3.1.2	Modèle de dispersion des réflecteurs	84
3.1.3	Modèle d'observation	85
3.1.4	Illustrations du modèle	86
3.1.5	Problème d'estimation considéré	87
3.2	Estimation statique des paramètres de l'amas	89
3.2.1	Estimateur <i>a posteriori</i>	89
3.2.2	Calcul de l'estimateur par un algorithme d'optimisation	92
3.3	Algorithme de pistage de l'amas	95
3.3.1	Modèles d'évolution	96
3.3.2	Description de l'algorithme de pistage	98
3.4	Validation expérimentale des algorithmes	103
3.4.1	Définitions des métriques d'estimation	103
3.4.2	Mise en oeuvre de l'algorithme statique	107
3.4.3	Mise en oeuvre de l'algorithme de pistage	115
3.5	Conclusions du chapitre	124

Dans ce chapitre, nous nous appuyons sur les concepts introduits dans le chapitre 2 pour proposer une modélisation des amas de débris spatiaux adaptée à leur dispersion observée en chapitre 1. Plus précisément, nous les considérons comme des objets étendus de forme incurvée qui peuvent être efficacement décrits par des paramètres définis sur des groupes de Lie. À partir de cette représentation, deux algorithmes sont développés pour estimer dans un contexte statique puis dynamique les caractéristiques d'un amas.

- Le premier algorithme que nous proposons est de type statique : son objectif est d'estimer à un instant donné la forme et la position du centroïde de l'amas. Il est fondé sur une méthode d'optimisation de type Newton sur groupe de Lie.
- Le deuxième met à profit des modèles d'évolution de ces paramètres au cours du temps pour réaliser un pistage de l'amas. La mise à jour récursive des estimations peut être décrite comme une variante de l'algorithme *LG-FKEI*.

Le chapitre est organisé de la manière suivante :

- premièrement, en s'inspirant des méthodes de pistage d'objets étendus de l'état de l'art, nous proposons de définir des modèles de dispersion de l'amas et d'observation, prenant en compte sa configuration géométrique à travers une représentation sur le groupe de Lie $SE(3)$.
- Ensuite, les algorithmes d'estimation et de pistage proposés et fondés sur ce modèle sont présentés et détaillés.
- Dans une dernière partie, ces algorithmes sont testés et validés sur deux scénarios permettant de simuler l'évolution d'un amas de débris. Le premier est directement fondé sur les modèles proposés alors que le second repose sur un modèle physique et permet de s'assurer de la pertinence de ces derniers.

3.1 Modélisation de l'amas comme une cible étendue sur groupe de Lie

3.1.1 Démarche proposée

Pour pister un amas de débris, nous proposons d'adapter les méthodes de cibles étendues de façon à prendre en compte sa forme géométrique. En effet, dans notre contexte d'application, les différents morceaux de débris sont proches entre eux et de nombre importants, par conséquent, une approche multi-cibles, par exemple fondée sur les ensembles finis aléatoires [86], ne serait pas pertinente. Bien qu'il soit ellipsoïdal à l'issue de sa formation, il tend à s'incurver progressivement au cours du temps. La figure 3.1 illustre ce phénomène mis en évidence dans le chapitre 1. Ainsi, de manière similaire à l'approche fondée sur les *matrices aléatoires*, nous supposons que l'étendue de l'amas est paramétrisée par une matrice de covariance représentant la dispersion des réflecteurs en son sein.

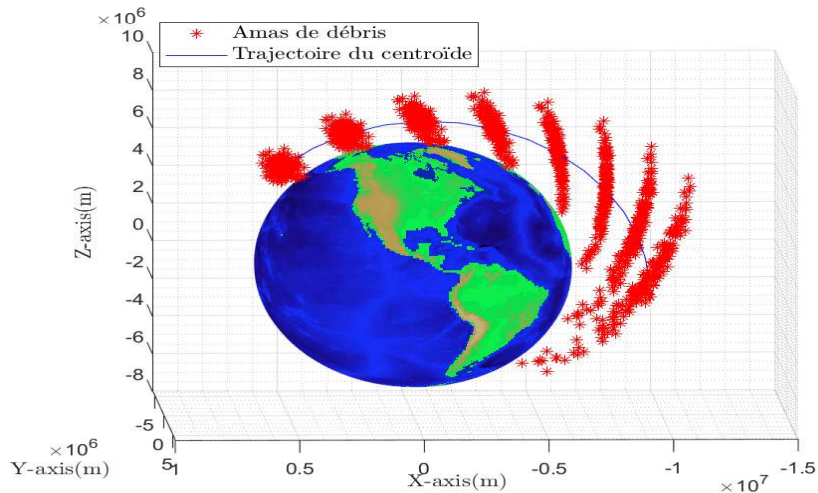


FIGURE 3.1 – Trajectoire d'un amas de débris se courbant sous l'effet de la force de gravitation. L'amas est représenté toutes les 2000 secondes.

Néanmoins, contrairement à cette méthode, modéliser la distribution des débris par une loi gaussienne, ou un mélange de lois gaussienne n'apparaît pas une approche pertinente. Dans

le premier cas, la courbure de l'amas ne serait pas prise en compte, et dans le second cas, le nombre de paramètres à estimer serait conséquent et se poserait la question du choix de l'ordre.

Une contribution de cette thèse est de proposer un modèle aléatoire de l'amas intrinsèquement adapté à sa forme et dépendant d'un nombre réduit de paramètres. Il est fondé sur l'observation que la répartition spatiale de débris est similaire à celle d'un ensemble d'échantillons statistiques obtenus sur le groupe de Lie $SE(3)$.

Contrairement aux approches classiques, nous proposons ainsi de définir un modèle de cible étendue de type *matrice aléatoire* sur $SE(3)$ permettant de prendre en compte :

- la géométrie de la forme, en particulier par la définition de son étendue à travers une matrice de covariance intrinsèque sur $SE(3)$ et d'un centroïde géométrique.
- la répartition des observations capteurs dans cette forme spécifique.

3.1.2 Modèle de dispersion des réflecteurs

Afin de modéliser la dispersion des réflecteurs au sein de l'amas, nous proposons de nous appuyer sur le formalisme des distributions gaussiennes concentrées. En effet, comme observé dans le chapitre 2, les échantillons issus de ce type de distribution forment, après projection sur $\mathbb{R}^3$, un nuage de points courbé et de forme similaire à celles observées sur la figure 3.1.

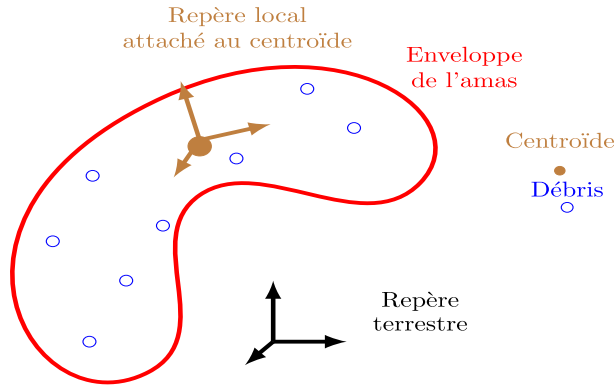


FIGURE 3.2 – Modélisation géométrique de l'amas de débris.

À un instant k donné, les objets à l'intérieur de l'amas ne sont pas tous détectés par le capteur radar : seuls n_k d'entre eux réfléchissent fortement le signal radar. L'idée clé de notre approche est que ces derniers, appelés réflecteurs, peuvent être obtenus de la manière suivante :

$$\mathbf{E}_{k,i} \sim \mathcal{N}_{SE(3)}^L(\mathbf{I}, \mathbf{S}_k), \quad \forall i \in \{1, \dots, n_k\} \text{ soit} \quad (3.1)$$

$$\mathbf{E}_{k,i} = \text{Exp}_{SE(3)}^{\wedge}(\epsilon_{k,i}) \quad \text{avec} \quad \epsilon_{k,i} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_k) \quad (3.2)$$

où la matrice de covariance $\mathbf{S}_k$ est une matrice de forme qui détermine la courbure et l'étendue de la dispersion des débris. Elle joue un rôle crucial dans la caractérisation de l'amas.

Les positions des réflecteurs dans un repère local B attaché à l'amas correspondent alors à la partie translation des $\{\mathbf{E}_{k,i}\}_{i=1}^{n_k}$: nous les notons $\{\mathbf{p}_{k,i,B}\}_{i=1}^{n_k}$

Afin de prendre en compte la dynamique de l'amas, ce dernier doit être modélisé dans un

repère terrestre T . En appliquant les formules de changement de repère, nous obtenons :

$$\mathbf{p}_{k,i,T} = \mathbf{R}_{TB,k} \mathbf{p}_{k,i,B} + \mathbf{p}_k \quad \forall i \in \{1, \dots, n_k\} \quad (3.3)$$

où le couple $\{\mathbf{R}_{TB,k}, \mathbf{p}_k\}$ définit les paramètres du centroïde. Ainsi, $\mathbf{R}_{TB,k}$ est la matrice de rotation entre le repère terrestre et le repère local et $\mathbf{p}_k$ à sa position dans le repère T .

Ce modèle peut alors se réécrire de manière compacte sur le groupe de Lie $SE(3)$ sous la forme :

$$\mathbf{Z}_{k,i} = \mathbf{M}_k \mathbf{E}_{k,i} \quad \forall i \in \{1, \dots, n_k\} \quad (3.4)$$

et :

$$\left\{ \begin{array}{l} \mathbf{M}_k = \begin{bmatrix} \mathbf{R}_{TB,k} & \mathbf{p}_k \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{Z}_{k,i} = \begin{bmatrix} \mathbf{R}_{k,i} & \mathbf{p}_{k,i,T} \\ \mathbf{0} & 1 \end{bmatrix} \end{array} \right. \quad (3.5)$$

$$(3.6)$$

où $\mathbf{R}_{k,i}$ correspond à la matrice de rotation du $i^{\text{ième}}$ réflecteur. Il convient de noter qu'elle n'est pas directement observable et n'a pas nécessairement de sens physique. Il s'agit ici d'une variable latente qui nous sert à construire notre modèle.

3.1.3 Modèle d'observation

Le système d'observation, un radar dans notre étude, fournit uniquement des mesures relatives à la position des débris. En prenant en compte l'incertitude sur chaque mesure, le modèle d'observation des réflecteurs peut alors s'écrire :

$$\mathbf{z}_{k,i} = \mathbf{\Pi}(\mathbf{Z}_{k,i}) + \mathbf{n}_{k,i} \quad (3.7)$$

où $\mathbf{\Pi}$ est une fonction allant de $SE(3)$ à $\mathbb{R}^3$ et $\mathbf{n}_{k,i}$ est le bruit de mesure associé au $i^{\text{ième}}$ réflecteur : il est supposé centré gaussien et de matrice de covariance $\mathbf{U}_k$.

- Dans le cas où nous supposons que les mesures $\{\mathbf{z}_{k,i}\}_{i=1}^{n_k}$ sont en coordonnées cartésiennes, alors $\mathbf{\Pi}$ est une fonction permettant de conserver la partie translation de $\mathbf{Z}_{k,i}$ égale à $\mathbf{p}_{k,i,T}$.
- Si des mesures radar de distance et d'angles sont disponibles, alors $\mathbf{\Pi}$ s'écrit comme la composition de deux transformations permettant de successivement :
 - convertir les coordonnées des réflecteurs dans le repère local du radar à travers une matrice de changement de repère,
 - transformer les positions cartésiennes en coordonnées sphériques telles que définies par les équations (1.24), (1.25) et (1.26).

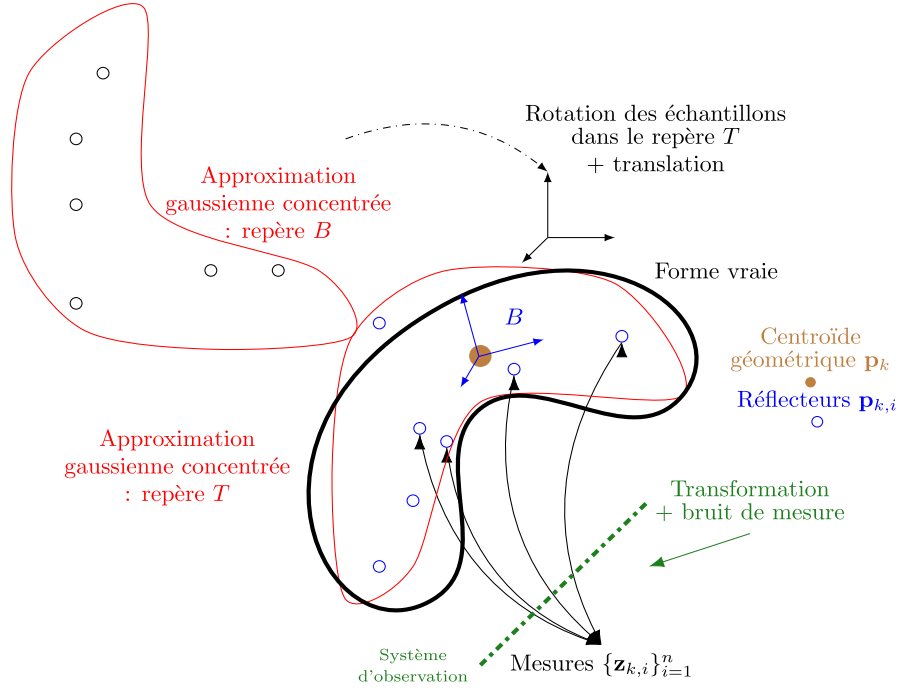


FIGURE 3.3 – Représentation géométrique du modèle proposé : la multiplication à gauche permet de transporter les réflecteurs sur une forme courbée approchant la forme vraie.

3.1.4 Illustrations du modèle

Le modèle proposé est illustré à travers la génération de deux ensembles d'échantillons.

Premièrement, nous avons simulé un ensemble d'échantillons $\{\mathbf{Z}_i\}_{i=1}^N$, avec $N = 2000$, selon notre modèle en considérant une multiplication à gauche égale à l'identité. Ensuite, nous avons pu isoler les échantillons en positions en conservant les composantes de translation de $SE(3)$, puis les représenter en rouge sur la figure (3.4). Nous observons qu'ils sont géométriquement centrés en $\mathbf{0}$, ce qui est consistant avec le modèle puisque ce point correspond au centroïde géométrique du nuage de points ainsi généré.

Nous avons également obtenu de nouveaux échantillons $\{\tilde{\mathbf{Z}}_i\}_{i=1}^N$ en multipliant chaque $\mathbf{Z}_i$ obtenu précédemment par la matrice :

$$\mathbf{M} = \begin{bmatrix} \text{Exp}_{SO(3)}^{\wedge} \left(\left[\frac{\pi}{4}, 0, 0 \right] \right) & \begin{bmatrix} 500 & 500 & 500 \end{bmatrix}^{\top} \\ \mathbf{0} & 1 \end{bmatrix} \quad (3.8)$$

Cette opération implique que chaque échantillon rouge sur la figure (3.4) subit une rotation d'angle $\frac{\pi}{4}$ selon l'axe des z suivie d'une translation de 500 m selon chaque axe. L'ensemble résultant de ces transformations est représenté par les échantillons bleus également sur la figure (3.4).

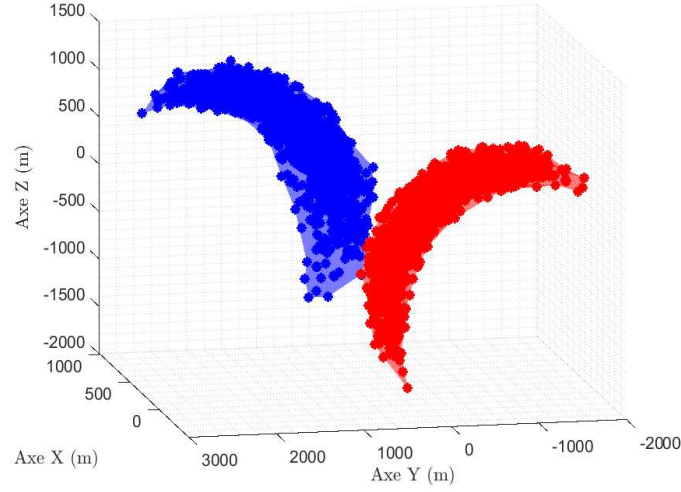


FIGURE 3.4 – Exemples de formes générées selon le modèle (3.7).

Dans un second temps, nous avons généré un nuage de points présentant les caractéristiques d'un amas de débris spatiaux. À ce titre, nous avons considéré une matrice de covariance de forme fournissant une étendue spatiale, réaliste et cohérente avec celle observée sur la figure d'expansion de l'amas. Par conséquent, le paramètre d'étendue de la matrice de covariance $\mathbf{S}_k$ a été fixé à un ordre proche de 10^6m . Le modèle d'observation (3.7) nous a ainsi permis de générer 10 réflecteurs bruités avec une matrice de covariance de bruit $\mathbf{U}_k$ prise égale à $10^3 \mathbf{I}_3$. Ils sont représentés sur la figure 3.5 de même que l'enveloppe de l'amas.

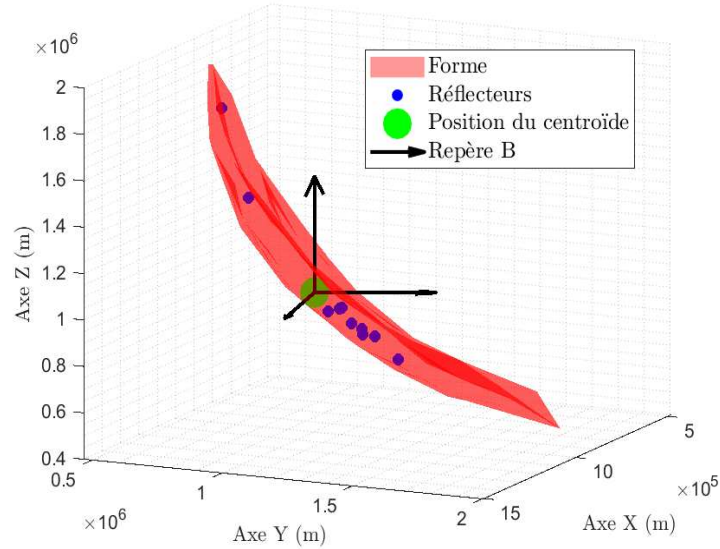


FIGURE 3.5 – Dispersion des réflecteurs générés selon les modèles (3.4) et (3.7).

3.1.5 Problème d'estimation considéré

Selon notre modèle, la dispersion des réflecteurs est donc décrite par deux paramètres inconnus :

→ la matrice :

$$\mathbf{M}_k = \begin{bmatrix} \mathbf{R}_{TB,k} & \mathbf{p}_k \\ \mathbf{0} & 1 \end{bmatrix} \in SE(3) \quad (3.9)$$

décrivant le centroïde inconnu.

→ la matrice $\mathbf{S}_k$ paramétrisant la forme inconnue et appartenant à l'ensemble des matrices symétriques définies positives.

Pour caractériser l'amas de débris, nous définissons donc un problème inverse où les inconnus sont concaténés dans la matrice $\mathbf{A}_k = \text{blkdiag}(\mathbf{M}_k, \mathbf{S}_k)$.

Nous constatons premièrement que l'espace des matrices symétriques définies positives n'est pas un groupe de Lie mais admet néanmoins une structure naturelle de variété riemannienne. Ainsi, il pourrait être possible de réaliser l'estimation de $\mathbf{S}_k$ conjointement avec celle $\mathbf{M}_k$ en utilisant des algorithmes d'estimation dédiés puisque $SE(3)$ est une variété riemannienne. Néanmoins, cette approche n'est pas consistante avec le modèle proposé : en particulier la forme caractéristique que nous pouvons générer est due à la géométrie de $SE(3)$ en tant que groupe et non en tant que variété.

Pour pallier ce problème, nous proposons de “contraindre” $\mathbf{S}_k$ à s'écrire comme un ensemble d'éléments sur groupe de Lie afin de proposer un formalisme unifié d'estimation.

Une solution est alors de décomposer $\mathbf{S}_k$ sur une base de vecteurs propres selon la forme :

$$\mathbf{S}_k = \mathbf{P}_k^\top \mathbf{D}_k \mathbf{P}_k. \quad (3.10)$$

Selon cette décomposition, $\mathbf{P}_k$ est une matrice vérifiant les propriétés suivantes :

$$\begin{cases} |\mathbf{P}_k| = 1 \end{cases} \quad (3.11)$$

$$\begin{cases} \mathbf{P}_k \mathbf{P}_k^\top = \mathbf{I} \end{cases} \quad (3.12)$$

et est donc un élément du groupe de Lie $SO(6)$.

Concernant $\mathbf{D}_k$, ses propriétés impliquent qu'elle vit sur l'ensemble des matrices diagonales à valeurs positives $D_6(\mathbb{R}^{+*})$:

$$D_6(\mathbb{R}^{+*}) = \left\{ \mathbf{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_6 \end{bmatrix} \middle| \forall i \in \{1, \dots, 6\}, d_i > 0 \right\} \quad (3.13)$$

Cet ensemble est en bijection avec l'espace produit $(\mathbb{R}^{+*})^6 = \mathbb{R}^{+*} \times \dots \times \mathbb{R}^{+*}$ qui est un groupe de Lie comme nous l'avons énoncé dans le chapitre 2. Par conséquent, $D_6(\mathbb{R}^{+*})$ en est également un et il admet en particulier la propriété de commutativité.

Ainsi, le problème d'estimation de $\mathbf{A}_k$ peut se réécrire globalement comme un problème d'inférence sur le groupe de Lie produit $G = SE(3) \times SO(6) \times D_6(\mathbb{R}^{+*})$ où la variable à estimer s'écrit :

$$\mathbf{X}_k = \text{blkdiag}(\mathbf{M}_k, \mathbf{P}_k, \mathbf{D}_k) \quad (3.14)$$

3.2 Estimation statique des paramètres de l'amas

Dans un premier temps, nous nous intéressons à estimer les paramètres d'un amas, c'est-à-dire sa position et sa forme, à un instant donné, à partir des mesures capteur. À cette fin, nous définissons un algorithme d'optimisation.

Le choix de travailler premièrement dans un contexte statique est lié à deux arguments :

- il permet d'introduire des concepts et résultats qui seront réinvestis dans l'algorithme de pistage,
- de plus, il présente un intérêt théorique à travers la borne de Cramér-Rao sur groupe de Lie que nous proposerons dans un contexte statique dans le chapitre 4.

Par conséquent, dans la section explicative de cet algorithme, l'indice k est omis de sorte que l'ensemble des variables $\{\mathbf{S}_k, \mathbf{X}_k, \mathbf{M}_k, \mathbf{P}_k, \mathbf{D}_k, \mathbf{U}_k, \mathbf{Z}_{k,i}, \mathbf{z}_{k,i}, n_k\}$ est remplacé par l'ensemble $\{\mathbf{S}, \mathbf{X}, \mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{U}, \mathbf{Z}_i, \mathbf{z}_i, n\}$.

Pour estimer l'ensemble des paramètres de l'amas, nous mettons en oeuvre un algorithme de type Newton - Gauss-Newton qui est une variante des algorithmes proposés par [72], dont nous détaillons le principe ci-dessous.

3.2.1 Estimateur *a posteriori*

Comme les variables paramétrisant l'amas sont définies sur des groupes de Lie, nous proposons de réaliser leur estimation dans un contexte bayésien permettant d'intégrer une contrainte *a priori* sur chacune d'entre elles.

Notre objectif est ainsi de déterminer la distribution *a posteriori* de $\mathbf{X}$ dont nous aimerions extraire un estimateur. Déterminer l'espérance de cette distribution est un problème délicat puisque cette quantité n'est pas définie de manière triviale comme sur un espace euclidien. Comme nous l'avons vu dans le chapitre 2, elle est fondée sur une distance sur $SE(3)$. Alternativement, nous proposons de chercher un estimateur de type maximum *a posteriori* :

$$\hat{\mathbf{X}} = \underset{\mathbf{X} \in G}{\operatorname{argmax}} p(\mathbf{X}|\mathbf{z}), \quad (3.15)$$

où $\mathbf{z} = \{\mathbf{z}_i\}_{i=1}^n$. Dans le modèle proposé, un aspect important à prendre à compte est la présence des variables $\mathbf{Z} = \{\mathbf{Z}_i\}_{i=1}^n$ appartenant à $\underbrace{SE(3) \times \dots \times SE(3)}_{n \text{ produits direct de } SE(3)} = SE(3)^n$ qui in-

terviennent dans le modèle direct. Elles n'ont pas de sens physique et servent uniquement à définir proprement la représentation sur groupe de Lie. Ce sont donc des variables de nuisances car elles "interfèrent" entre les observations et les inconnues. Par conséquent, obtenir la distribution de $p(\mathbf{X}|\mathbf{z})$ nécessite le calcul de la vraisemblance marginalisée $p(\mathbf{z}|\mathbf{X})$ de la manière suivante :

$$p(\mathbf{z}|\mathbf{X}) = \int_{\mathbf{Z} \in SE(3)^n} p(\mathbf{z}|\mathbf{Z}) p(\mathbf{Z}|\mathbf{X}) \mu_{SE(3)^n}(d\mathbf{Z}) \quad (3.16)$$

où $\mu_{SE(3)^n}$ est une mesure de Haar sur $SE(3)^n$.

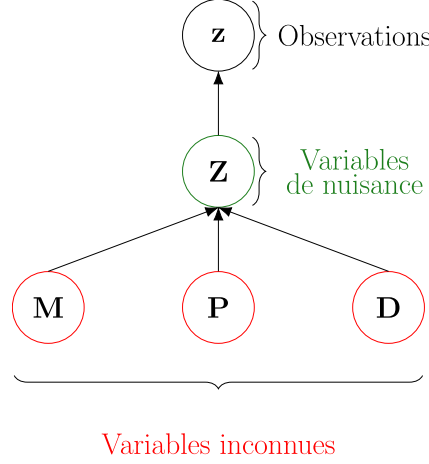


FIGURE 3.6 – Représentation hiérarchique du problème d'estimation

L'intégrale (3.16) n'est pas calculable analytiquement et difficilement approchable par une distribution paramétrique. Pour pallier ce problème, nous proposons d'estimer ces variables au même titre que nos paramètres d'intérêts et de considérer l'ensemble des variables inconnues sous la forme $\mathbf{X}^Z = \text{blkdiag}(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z})$ appartenant au groupe de Lie produit $G^Z = SE(3) \times SO(6) \times D_6(\mathbb{R}^{+*}) \times SE(3)^n$ de dimension m_Z . Plus précisément, nous cherchons à déterminer un maximiseur de la distribution *a posteriori* $p(\mathbf{X}^Z | \mathbf{z})$ noté :

$$\hat{\mathbf{X}}^Z = \underset{\mathbf{X}^Z \in G^Z}{\operatorname{argmax}} p(\mathbf{X}^Z | \mathbf{z}), \quad (3.17)$$

L'expression peut alors s'écrire de manière plus détaillée, en utilisant la règle de Bayes. En effet, par indépendance conditionnelle des $\{\mathbf{z}_i\}_{i=1}^n$ et des $\{\mathbf{Z}_i\}_{i=1}^n$ et selon le modèle hiérarchique illustré en 3.6, nous pouvons montrer que :

$$p(\mathbf{X}^Z | \mathbf{z}) \propto \left(\prod_{i=1}^n p(\mathbf{z}_i | \mathbf{Z}_i) p(\mathbf{Z}_i | \mathbf{M}, \mathbf{P}, \mathbf{D}) \right) \underbrace{p(\mathbf{M}, \mathbf{P}, \mathbf{D})}_{\text{termes } a \text{ priori}}. \quad (3.18)$$

Détaillons à présent chacun de ces termes.

3.2.1.1 Expression de la vraisemblance

D'après le modèle (3.4), les variables $\{\mathbf{Z}_i\}_{i=1}^n$, conditionnellement à $\mathbf{M}$ et $\mathbf{S}$, sont distribuées sous une loi gaussienne concentrée à gauche, soit :

$$p(\mathbf{Z}_i | \mathbf{M}, \mathbf{S}) = \mathcal{N}_{SE(3)}^L(\mathbf{Z}_i; \mathbf{M}, \mathbf{S}). \quad (3.19)$$

En mettant à profit l'équation (3.10), cette distribution peut se réécrire, au voisinage du maximum *a posteriori*, en fonction de $\mathbf{M}$, $\mathbf{P}$ et $\mathbf{D}$:

$$p(\mathbf{Z}_i | \mathbf{M}, \mathbf{P}, \mathbf{D}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{D}|}} \exp \left(-\frac{1}{2} \|\text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)\|_{\mathbf{P}^\top \mathbf{D} \mathbf{P}}^2 \right), \quad (3.20)$$

où $m = 6$. Les $\{\mathbf{Z}_i\}_{i=1}^n$ étant supposés mutuellement indépendants conditionnellement aux variables inconnues, la distribution de $\mathbf{Z}|\mathbf{M}, \mathbf{P}, \mathbf{D}$ s'écrit :

$$p(\mathbf{Z}|\mathbf{M}, \mathbf{P}, \mathbf{D}) = \prod_{i=1}^n p(\mathbf{Z}_i|\mathbf{M}, \mathbf{P}, \mathbf{D}) \quad (3.21)$$

$$= \frac{1}{\sqrt{(2\pi)^{mn} \left(\prod_{j=1}^6 d_j \right)^n}} \exp \left(-\frac{1}{2} \sum_{i=1}^n \|\text{Log}_{SE(3)}^{\vee}(\mathbf{M}^{-1} \mathbf{Z}_i)\|_{\mathbf{P}^\top \mathbf{D} \mathbf{P}}^2 \right), \quad (3.22)$$

où les $\{d_j\}_{j=1}^6$ correspondent aux composantes diagonales de $\mathbf{D}$. Ce sont donc des variables inconnues qui sont à estimer.

Concernant la vraisemblance des mesures $\{\mathbf{z}_i\}_{i=1}^n$, nous en déduisons d'après l'équation (3.7) qu'il s'agit d'une loi gaussienne multivariée sur $\mathbb{R}^3$, soit :

$$p(\mathbf{z}_i|\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z}_i) = \mathcal{N}(\mathbf{z}_i; \Pi(\mathbf{Z}_i), \mathbf{U}) \quad (3.23)$$

$$\propto \exp \left(-\frac{1}{2} \|\mathbf{z}_i - \Pi(\mathbf{Z}_i)\|_{\mathbf{U}}^2 \right). \quad (3.24)$$

De la même façon, comme les $\mathbf{z} = \{\mathbf{z}_i\}_{i=1}^n$ sont mutuellement indépendants conditionnellement à $\mathbf{Z}$, il vient que :

$$p(\mathbf{z}|\mathbf{Z}) = \prod_{i=1}^n p(\mathbf{z}_i|\mathbf{Z}_i) \quad (3.25)$$

$$\propto \exp \left(-\frac{1}{2} \sum_{i=1}^n \|\mathbf{z}_i - \Pi(\mathbf{Z}_i)\|_{\mathbf{U}}^2 \right). \quad (3.26)$$

3.2.1.2 Expression des *a priori*

Il reste à définir les lois *a priori* des paramètres inconnus qui sont tous définis sur des groupes de Lie. Nous choisissons de façon générique de les représenter par des lois gaussiennes concentrées sur leurs ensembles de définition respectifs. Ainsi, la méthode d'estimation développée dans le cas statique se transposera simplement au cas dynamique.

Nous supposons que $\mathbf{M}$, $\mathbf{P}$ et $\mathbf{D}$ sont indépendants, par conséquent :

$$p(\mathbf{M}, \mathbf{P}, \mathbf{D}) = p(\mathbf{M}) p(\mathbf{P}) p(\mathbf{D}). \quad (3.27)$$

Nous choisissons donc une loi *a priori* gaussienne concentrée sur $SO(6)$ pour $\mathbf{P}$ ainsi qu'une loi gaussienne concentrée sur $D_6(\mathbb{R}^{+*})$ pour $\mathbf{D}$:

$$\begin{cases} p(\mathbf{P}) = \mathcal{N}_{SO(6)}^L(\mathbf{P}; \boldsymbol{\mu}_P, \mathbf{Q}_P) \end{cases} \quad (3.28)$$

$$\begin{cases} p(\mathbf{D}) = \mathcal{N}_{D_6(\mathbb{R}^{+*})}^L(\mathbf{D}; \boldsymbol{\mu}_D, \mathbf{Q}_D). \end{cases} \quad (3.29)$$

Il est à noter que la distribution gaussienne concentrée sur $D_6(\mathbb{R}^{+*})$ peut être vue comme une généralisation de la loi log-normale sur $((\mathbb{R}^{+*}), *)$.

Concernant le paramètre $\mathbf{M}$, nous décrivons son incertitude *a priori* également par une distribution gaussienne concentrée sur $SE(3)$:

$$p(\mathbf{M}) = \mathcal{N}_{SE(3)}^L(\mathbf{M}; \boldsymbol{\mu}_M, \mathbf{Q}_M). \quad (3.30)$$

3.2.2 Calcul de l'estimateur par un algorithme d'optimisation

Afin de maximiser $p(\mathbf{X}^Z | \mathbf{z})$, nous proposons de nous intéresser à la minimisation de son anti-logarithme $-\log p(\mathbf{X}^Z | \mathbf{z})$, de la même manière que illustré dans [28]. Par conséquent, nous définissons la fonction de minimisation suivante :

$$J(\mathbf{X}^Z) = -2 \log p(\mathbf{X}^Z | \mathbf{z}) \quad (3.31)$$

et nous cherchons désormais à résoudre le problème d'optimisation suivant :

$$\hat{\mathbf{X}}^Z = \underset{\mathbf{X}^Z \in G^Z}{\operatorname{argmin}} J(\mathbf{X}^Z). \quad (3.32)$$

Nous constatons alors que le critère J admet une structure particulière.

- Premièrement, nous constatons que le logarithme de la distribution (3.20) peut s'écrire comme un terme quadratique dépendant de $\mathbf{M}$, $\mathbf{P}$ et $\mathbf{D}$. En effet, nous avons :

$$\|\operatorname{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)\|_{\mathbf{P}^\top \mathbf{D} \mathbf{P}}^2 = \|\mathbf{D}^{-1/2} \mathbf{P} \operatorname{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)\|_2^2 \quad (3.33)$$

Ainsi, cette écriture se prête bien à la mise en place d'un algorithme d'optimisation de type Gauss-Newton.

- De plus, de par la présence des composantes diagonales de $\mathbf{D}$ en facteur dans (3.22), la log-vraisemblance des mesures ne peut s'écrire comme une norme vectorielle au carré contrairement aux autres termes intervenant dans J .

Par conséquent, nous pouvons le décomposer de la manière suivante :

$$J(\mathbf{X}^Z) = \underbrace{J_{nq}(\mathbf{D})}_{\text{Partie non quadratique}} + \underbrace{J_q(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z})}_{\text{Partie norme vectorielle}}, \quad (3.34)$$

où :

$$J_{nq}(\mathbf{D}) = n \sum_{j=1}^6 \log(d_j) \quad (3.35)$$

$$\begin{aligned} J_q(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z}) = & \sum_{i=1}^n \left(\|\mathbf{z}_i - \Pi(\mathbf{Z}_i)\|_{\mathbf{U}}^2 + \|\mathbf{D}^{-1/2} \mathbf{P} \operatorname{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)\|_2^2 \right) \\ & + \|\operatorname{Log}_{SE(3)}^\vee(\boldsymbol{\mu}_M^{-1} \mathbf{M})\|_{\mathbf{Q}_M}^2 + \|\operatorname{Log}_{SO(6)}^\vee(\boldsymbol{\mu}_P^{-1} \mathbf{P})\|_{\mathbf{Q}_P}^2 + \|\operatorname{Log}_{D_6(\mathbb{R}^{+*})}^\vee(\boldsymbol{\mu}_D^{-1} \mathbf{D})\|_{\mathbf{Q}_D}^2 \end{aligned} \quad (3.36)$$

Par sa structure de norme vectorielle, J_q peut se réécrire en faisant intervenir une fonction différentiable $\varepsilon : G^Z \rightarrow \mathbb{R}^{m_Z}$ telle que :

$$J_q(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z}) = \|\varepsilon(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z})\|_{\boldsymbol{\Sigma}_q}^2 \quad (3.37)$$

avec :

$$\varepsilon(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z}) = \left[\Phi(\mathbf{Z})^\top, \Psi(\mathbf{X}, \mathbf{Z})^\top, \text{Log}_G^\vee \left((\mathbf{X}_P)^{-1} \mathbf{X} \right)^\top \right]^\top. \quad (3.38)$$

Détaillons les différents composantes de cette fonction.

Les fonctions Φ et Ψ s'écrivent :

$$\Phi(\mathbf{Z}) = \left[\phi(\mathbf{z}_1, \mathbf{Z}_1)^\top, \dots, \phi(\mathbf{z}_n, \mathbf{Z}_n)^\top \right]^\top, \quad (3.39)$$

avec $\phi(\mathbf{z}_i, \mathbf{Z}_i) = \mathbf{z}_i - \Pi(\mathbf{Z}_i)$, $\forall i \in \{1, \dots, n\}$.

$$\Psi(\mathbf{X}, \mathbf{Z}) = \left[\psi(\mathbf{X}, \mathbf{Z}_1)^\top, \dots, \psi(\mathbf{X}, \mathbf{Z}_n)^\top \right]^\top, \quad (3.40)$$

Finalement ; $\psi(\mathbf{X}, \mathbf{Z}_i) = \mathbf{D}^{-1/2} \mathbf{P} \text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)$, $\forall i \in \{1, \dots, n\}$.

$$\mathbf{X}_P = \text{blkdiag}(\boldsymbol{\mu}_M, \boldsymbol{\mu}_P, \boldsymbol{\mu}_D) \quad (3.41)$$

$$\boldsymbol{\Sigma}_q = \text{blkdiag}(\mathbf{I}_{n \times n} \odot \mathbf{U}, \mathbf{I}_{6n \times 6n}, \mathbf{Q}_M, \mathbf{Q}_P, \mathbf{Q}_D). \quad (3.42)$$

où $\text{blkdiag}(\cdot)$ correspond à l'opérateur prenant en entrée un ensemble de matrices et retournant en sortie une matrice bloc diagonale. Ainsi, le critère se réécrit de manière plus compacte :

$$J(\mathbf{X}^Z) = n \sum_{j=1}^6 \log(d_j) + \|\varepsilon(\mathbf{M}, \mathbf{P}, \mathbf{D}, \mathbf{Z})\|_{\boldsymbol{\Sigma}_q}^2. \quad (3.43)$$

Cette fonction est différentiable mais néanmoins difficile à minimiser analytiquement. Nous proposons donc mettre en place une procédure numérique afin de pouvoir approcher son minimiseur ou du moins un point critique. À cette fin, un algorithme d'optimisation intrinsèque sur groupe de Lie est considéré.

La structure semi-norme vectorielle du critère implique qu'il n'est pas possible d'appliquer directement un algorithme *LG-GN*. Un algorithme *LG-N* pourrait être employé mais il n'assurait pas que la matrice hessienne conserve son caractère défini ou semi-défini positif.

Nous proposons alors d'utiliser la structure particulière du critère afin de mettre en oeuvre un algorithme de Newton modifié et adapté à celle-ci.

De manière similaire au *LG-N* classique défini dans le chapitre 2, nous cherchons à obtenir une suite d'itérés $\{\mathbf{X}^{Z(l)}\}_{l=1}^N$ sous la forme :

$$\mathbf{X}^{Z(l+1)} = \mathbf{X}^{Z(l)} \exp_{G^Z}^\wedge(\boldsymbol{\delta}_Z^{(l+1)}) \quad (3.44)$$

où à chaque itération l , une direction de descente optimale $\boldsymbol{\delta}_Z^{(l+1)}$ peut être obtenue classiquement en minimisant une approximation à l'ordre 2 de $J(\mathbf{X}^{Z(l)} \exp_{G^Z}^\wedge(\boldsymbol{\delta}))$. Dans notre cas, nous allons déduire une direction de descente optimale adaptée dont nous détaillons maintenant le calcul.

• Calcul de la direction de descente optimale

Comme nous venons de voir, nous souhaitons déterminer une approximation à l'ordre deux du critère afin de calculer une direction de descente optimale. D'après [47], toute fonction deux

fois différentiable à valeurs dans $\mathbb{R}$ admet un développement limité de Taylor-Young à l'ordre 2 et il est possible d'obtenir une approximation de J_{nq} notée $\tilde{J}_{nq}$ sous la forme :

$$\tilde{J}_{nq}(\mathbf{X}^{Z(l)} \text{Exp}_G^\wedge(\boldsymbol{\delta})) = J_{nq}(\mathbf{X}^{Z(l)}) + \left(\mathcal{L}_{J_{nq}(\mathbf{X}^{Z(l)})}^R \right)^\top \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^\top \mathbf{H}_{nq}(\mathbf{X}^{Z(l)}) \boldsymbol{\delta} \quad (3.45)$$

où $\mathcal{L}_{J_{nq}(\mathbf{X}^{Z(l)})}^R$ et $\mathbf{H}_{nq}(\mathbf{X}^{Z(l)})$ correspondent respectivement à la dérivée de Lie à droite et à la hessienne sur groupe de Lie de J_{nq} calculée en $\mathbf{X}^{Z(l)}$ selon l'équation (2.51).

Ainsi, nous proposons de chercher la descente $\boldsymbol{\delta}_Z^{(l+1)} \in \mathbb{R}^{m_Z}$ minimisant la somme :

- d'un développement direct à l'ordre 2 de $J_{nq}(\mathbf{X}^{Z(l)} \text{exp}_{G^Z}^\wedge(\boldsymbol{\delta}))$ selon l'équation (3.45).
- d'une approximation à l'ordre 2 de J_q , notée J_{nq} , obtenue par développement de Taylor à l'ordre 1 de $\boldsymbol{\varepsilon}(\mathbf{X}^{Z(l)} \text{exp}_{G^Z}^\wedge(\boldsymbol{\delta}))$ à la manière de la méthode de Gauss-Newton :

$$\tilde{J}_q(\mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta})) = \|\boldsymbol{\varepsilon}(\mathbf{X}^{Z(l)}) + \mathbf{J}_\varepsilon^{(l)\top} \boldsymbol{\delta}\|_{\Sigma_q}^2 \quad (3.46)$$

où $\mathbf{J}_\varepsilon^{(l)}$ correspond à la matrice jacobienne sur groupe de Lie de $\boldsymbol{\varepsilon}$ calculée en $\mathbf{X}^{Z(l)}$:

$$\mathbf{J}_\varepsilon^{(l)} = \left. \frac{\partial \boldsymbol{\varepsilon}(\mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta}))}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta}=\mathbf{0}} \quad (3.47)$$

La somme de ces deux approximations aboutit alors à l'estimation suivante de J :

$$\tilde{J}(\mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta})) = J(\mathbf{X}^{Z(l)}) + \left(\mathcal{L}_{J(\mathbf{X}^{Z(l)})}^R \right)^\top \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^\top \mathbf{H}_J^{(l)} \boldsymbol{\delta}. \quad (3.48)$$

avec :

$$\begin{cases} \mathbf{H}_J^{(l)} = \mathbf{H}_{nq}(\mathbf{X}^{Z(l)}) + 2 \mathbf{J}_\varepsilon^{(l)\top} \Sigma_q^{-1} \mathbf{J}_\varepsilon^{(l)}, \\ \mathcal{L}_{J(\mathbf{X}^{Z(l)})}^R = \mathcal{L}_{J_{nq}(\mathbf{X}^{Z(l)})}^R + 2 \mathbf{J}_\varepsilon^{(l)\top} \Sigma_q^{-1} \boldsymbol{\varepsilon}(\mathbf{X}^{Z(l)}). \end{cases} \quad (3.49)$$

$$(3.50)$$

La direction de descente optimale est obtenue en minimisant la quantité (3.48) par rapport à $\boldsymbol{\delta}$. De par sa forme quadratique, cela est équivalent à déterminer un point critique annulant sa dérivée. Si $\mathbf{H}_J^{(l)}$ est inversible, nous en déduisons alors que :

$$\boldsymbol{\delta}_Z^{(l+1)} = - \left(\mathbf{H}_J^{(l)} \right)^{-1} \left(\mathcal{L}_{J(\mathbf{X}^{Z(l)})}^R \right). \quad (3.51)$$

Algorithme 1 : Algo-Newton Pseudo-code de l'algorithme de Newton proposé.

Entrées : Paramètre initial $\mathbf{X}^{Z(0)} = \text{blkdiag}(\mathbf{M}^{(0)}, \mathbf{P}^{(0)}, \mathbf{D}^{(0)}, \mathbf{Z}^{(0)})$, mesures $\mathbf{z}$, critère $J(\mathbf{z}, \cdot)$, nombre maximum d'itérations L .

tant que $1 \leq l < L$ **faire**

- Approximation du critère :

$$\tilde{J} \leftarrow \tilde{J}_{nq} + \tilde{J}_q$$

- Calcul de la direction de descente optimale :

$$\boldsymbol{\delta}_Z^{(l+1)} \leftarrow \underset{\boldsymbol{\delta} \in \mathbb{R}^{m_Z}}{\text{argmin}} \left(\tilde{J}_q(\cdot, \mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta})) + \tilde{J}_{nq}(\cdot, \mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta})) \right)$$

- Mise à jour :

$$\mathbf{X}^{Z(l+1)} \leftarrow \mathbf{X}^{Z(l)} \text{Exp}_{G^Z}^\wedge(\boldsymbol{\delta}_Z^{(l+1)});$$

- $l \leftarrow l + 1$;

retourner $\mathbf{X}^{(L)}$

3.2.2.1 Remarques sur l'approche

- La formule (3.49) assure la semi-définie positivité de $\mathbf{H}_J^{(l)}$ puisque, comme détaillé dans l'annexe (C), $\mathbf{H}_{nq}^{(l)}$ est une matrice nulle. Ainsi, $\boldsymbol{\delta}_Z^{(l+1)}$ est garantie d'être une direction de descente.
- En pratique, le nombre d'itérations sera assez faible : une stabilité du critère d'optimisation s'observe aux alentours de 10 itérations pour les scénarios de test considérés en partie (3.4).
- La méthode proposée ne garantit pas qu'il y ait convergence vers un minimum global. En effet, le critère à optimiser n'est pas convexe et peut présenter plusieurs minimums locaux.

3.3 Algorithme de pistage de l'amas

Afin de suivre l'évolution temporelle de la forme et du centroïde de l'amas, nous proposons maintenant de mettre en place un algorithme de pistage permettant d'estimer récursivement les paramètres $\mathbf{M}_k$, $\mathbf{P}_k$ et $\mathbf{D}_k$ grâce à des mesures réflecteurs $\mathbf{z}_{1:k} = [\{\mathbf{z}_{1,i}\}_{i=1}^{n_1}, \dots, \{\mathbf{z}_{k,i}\}_{i=1}^{n_k}]$ obtenues suivant les modèles (3.4) et (3.7).

De manière analogue à l'approche de [31], nous considérons une variante du filtre de Kalman où la distribution jointe *a posteriori* de l'état du centroïde de la matrice d'étendue est supposée prendre une forme paramétrique : nous proposons plus précisément d'approcher cette distribution par une distribution gaussienne concentrée sur le groupe de Lie produit associé à l'ensemble de tous les paramètres inconnus.

Afin d'identifier les moments de celle-ci, nous mettons en oeuvre un algorithme *LG-IEKF* admettant trois différences fondamentales par rapport à l'algorithme proposé dans [63] :

- Il considère à chaque instant plusieurs mesures liées à la dispersion des réflecteurs.
- Il prend en compte la présence de variables de nuisances.
- L'étape de correction est fondée sur l'optimisation d'un critère modifié non quadratique telle que décrite dans la section (3.2). La matrice de covariance de la loi *a posteriori*

d'intérêt est par conséquent déterminée par une approximation de Laplace et non de Gauss-Laplace.

Afin de décrire les différentes étapes de l'algorithme, nous devons d'abord définir les modèles d'évolution régissant la dynamique des différents paramètres à pister. Comme la position du centroïde de l'amas change au cours du temps, sa trajectoire dépend de sa vitesse, en accord avec le modèle d'évolution décrit dans le chapitre 1. Par conséquent, cette dernière doit être intégrée au problème d'estimation.

3.3.1 Modèles d'évolution

3.3.1.1 Modèle d'évolution du centroïde

Nous avons vu dans le chapitre 2 en section (1.3) que la trajectoire d'un débris, dont l'altitude est suffisamment élevée, est déterminée principalement par la force gravitationnelle. Nous supposons que c'est également le cas pour le centroïde de l'amas.

Si nous notons $\mathbf{v}_k$ sa vitesse à l'instant k , $\mathbf{p}_k$ étant sa position, son évolution dynamique peut se réécrire, grâce à (1.36) selon les deux équations aux différences suivantes :

$$\mathbf{p}_k = \underbrace{\mathbf{p}_{k-1} + T \mathbf{f}_p(\mathbf{v}_{k-1})}_{\tilde{\mathbf{f}}_p(\mathbf{p}_{k-1}, \mathbf{v}_{k-1})} + \mathbf{w}_{p,k}, \quad (3.52)$$

$$\mathbf{v}_k = \underbrace{\mathbf{v}_{k-1} + T \mathbf{f}_v(\mathbf{p}_{k-1}, \mathbf{v}_{k-1})}_{\tilde{\mathbf{f}}_v(\mathbf{p}_{k-1}, \mathbf{v}_{k-1})} + \mathbf{w}_{v,k}, \quad (3.53)$$

avec :

$$\mathbf{f}_p(\mathbf{v}_{k-1}) = \mathbf{v}_{k-1}, \quad (3.54)$$

$$\mathbf{f}_v(\mathbf{p}_{k-1}, \mathbf{v}_{k-1}) = -\boldsymbol{\Omega} \wedge \boldsymbol{\Omega} \wedge \mathbf{p}_{k-1} - 2 \boldsymbol{\Omega} \wedge \mathbf{v}_{k-1} - \frac{G M}{\|\mathbf{p}_{k-1}\|^3} \mathbf{p}_{k-1}. \quad (3.55)$$

Nous rappelons, dans ces deux formules, que T correspond à la période de discrétisation du modèle continu, $\boldsymbol{\Omega}$ au vecteur vitesse de rotation de la Terre, M à sa masse et G à la constante gravitationnelle.

Les vecteurs $\mathbf{w}_{p,k}$ and $\mathbf{w}_{v,k}$ représentent respectivement les bruits de position et de vitesse. Ils sont supposés centrés et gaussiens. Ils modélisent, non seulement, les perturbations physiques dues aux forces agissant sur l'amas et qui n'admettent pas d'expressions analytiques (vents solaires, attraction lunaire, ...) mais également les erreurs introduites par la discrétisation du modèle.

De par la corrélation entre la position et la vitesse, il apparaît nécessaire d'introduire cette dernière dans le problème d'estimation. Par conséquent, la matrice inconnue $\mathbf{M}_k$ doit être augmentée avec le vecteur vitesse $\mathbf{v}_k \in \mathbb{R}^3$. Pour obtenir une représentation unifiée sur groupe de Lie, nous nous servons du fait que $\mathbb{R}^3$ peut être considéré comme un cas particulier du groupe $SE(3)$ lorsque la matrice de rotation est égale à la matrice identité. Nous considérons donc un état augmenté du centroïde sur le groupe $SE(3) \times \mathbb{R}^3$ prenant la forme :

$$\mathbf{M}_k^v = \left[\begin{array}{c|cc} \mathbf{M}_k & \mathbf{0}_{6 \times 6} & \\ \hline \mathbf{0}_{4 \times 4} & \mathbf{I}_{3 \times 3} & \mathbf{v}_k \\ & \mathbf{0}_{1 \times 3} & 1 \end{array} \right],$$

où $\mathbf{M}_k$ est la matrice appartenant à $SE(3)$ et définie par l'équation (3.5). Pour modéliser la dynamique de la matrice d'orientation $\mathbf{R}_{TB,k}$ incluse dans $\mathbf{M}_k$, nous optons pour une marche aléatoire sur le groupe $SO(3)$ permettant de prendre en compte le fait qu'elle n'est pas constante au cours du temps mais varie lentement. Ce type de modèle d'évolution généralise les marches aléatoires classiques euclidiennes et a déjà été utilisé dans diverses études. $\mathbf{R}_{TB,k}$ vérifie alors :

$$\mathbf{R}_{TB,k} = \mathbf{R}_{TB,k-1} \text{Exp}_{SO(3)}^\wedge(\mathbf{w}_{R,k}) \quad (3.56)$$

où $\mathbf{w}_{R,k}$ est un vecteur gaussien centré.

De cette manière, en concaténant les équations (3.52), (3.53) et (3.56), le modèle d'évolution global peut se réécrire sous la forme compacte suivante :

$$\mathbf{M}_k^v = \mathbf{f}_k^v(\mathbf{M}_{k-1}^v) \text{Exp}_G^\wedge(\mathbf{w}_k), \quad (3.57)$$

avec :

$$\mathbf{f}_k^v(\mathbf{M}_{k-1}^v) = \begin{bmatrix} \mathbf{R}_{TB,k-1} \tilde{\mathbf{f}}_p(\mathbf{p}_{k-1}, \mathbf{v}_{k-1}) & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{1 \times 3} & 1 & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \tilde{\mathbf{f}}_v(\mathbf{p}_{k-1}, \mathbf{v}_{k-1}) \\ \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}$$

et $\mathbf{w}_k = [\mathbf{w}_{R,k}^T, \mathbf{B}(\mathbf{w}_{R,k})\mathbf{w}_{p,k}^T, \mathbf{w}_{v,k}^T]^T$ où $\mathbf{B}(\mathbf{w}_{R,k})$ est une matrice dépendant de $\mathbf{w}_{R,k}$ s'écrivant [87] :

$$\mathbf{B}(\mathbf{w}_{R,k}) = \mathbf{I} + \frac{(1 - \cos(\|\mathbf{w}_{R,k}\|)) [\mathbf{w}_{R,k}]_\times}{\|\mathbf{w}_{R,k}\|^2} + \frac{(\|\mathbf{w}_{R,k}\| - \sin(\|\mathbf{w}_{R,k}\|)) [\mathbf{w}_{R,k}]_\times^2}{\|\mathbf{w}_{R,k}\|^3}. \quad (3.58)$$

Si la matrice de covariance de $\mathbf{w}_{R,k}$ prend de faibles valeurs, ce qui est le cas pour des variables angulaires, alors $\mathbf{B}(\mathbf{w}_{R,k})$ peut être raisonnablement approchée par la matrice identité. Il s'ensuit que $\mathbf{w}_k \simeq [\mathbf{w}_{R,k}^T, \mathbf{w}_{p,k}^T, \mathbf{w}_{v,k}^T]^T$: ce vecteur est gaussien, centré et de matrice de covariance notée $\mathbf{W}_k$.

Il est à noter qu'il aurait été également possible de définir un état augmenté sur le groupe $SO(3) \times \mathbb{R}^6$ en considérant d'une part $\mathbf{R}_{TB,k}$ sur $SO(3)$ et d'autre part le couple $(\mathbf{p}_k, \mathbf{v}_k)$ sur $\mathbb{R}^6$. Dans ce cas, le modèle d'observation de l'amas aurait dû être légèrement modifié, en introduisant une fonction permettant de transformer un élément de $SO(3) \times \mathbb{R}^3$ en un élément de $SE(3)$.

3.3.1.2 Modèle d'évolution de la matrice d'étendue

Pour garantir que $\mathbf{S}_k$ présente les propriétés d'une matrice de covariance, une solution est d'assujétir $\mathbf{D}_k$ et $\mathbf{P}_k$ à évoluer sur leurs groupes de Lie de définition. À nouveau, nous considérons des marches aléatoires. Elles prennent la forme :

$$\mathbf{D}_k = \mathbf{D}_{k-1} \text{Exp}_{D_6(\mathbb{R}^{+*})}^\wedge(\boldsymbol{\epsilon}_{\mathbf{D}_k}), \quad (3.59)$$

$$\mathbf{P}_k = \mathbf{P}_{k-1} \text{Exp}_{SO(6)}^\wedge(\boldsymbol{\epsilon}_{\mathbf{P}_k}), \quad (3.60)$$

avec $\boldsymbol{\epsilon}_{\mathbf{D}_k}$ et $\boldsymbol{\epsilon}_{\mathbf{P}_k}$ deux bruits d'évolution caractérisant les deux marches aléatoires. Ils sont supposés centrés et gaussiens.

3.3.1.3 Modèle d'état augmenté

Du fait de la prise en compte de la vitesse, l'état à estimer est différent de l'approche statique et s'écrit $\mathbf{X}_k^v = \text{blkdiag}(\mathbf{M}_k^v, \mathbf{P}_k, \mathbf{D}_k)$. Il appartient au groupe $G_X = SE(3) \times \mathbb{R}^3 \times SO(6) \times D_6(\mathbb{R}^{+*})$ de dimension m_X .

Pour faciliter la présentation de notre algorithme de pistage, il paraît judicieux de réécrire un modèle augmenté concaténant les modèles d'évolution de chacune des variables. Nous avons alors :

$$\mathbf{X}_k^v = \mathbf{f}_k(\mathbf{X}_{k-1}^v) \text{Exp}_{G_X}^\wedge(\mathbf{w}_k^v) \text{ avec } \mathbf{w}_k^v \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_k^v) \quad (3.61)$$

où $\mathbf{f}_k(\mathbf{X}_{k-1}^v) = \text{blkdiag}([\mathbf{f}_k^v(\mathbf{M}_{k-1}^v), \mathbf{P}_{k-1}, \mathbf{D}_{k-1}])$ et $\mathbf{w}_k^v$ regroupe l'ensemble des bruits d'état.

3.3.2 Description de l'algorithme de pistage

Dans cette sous-section, nous présentons les principes théoriques de l'algorithme de pistage proposé. L'objectif visé est d'approcher la distribution de l'état augmenté $\mathbf{X}_k^v$ sous la forme :

$$p(\mathbf{X}_k^v | \mathbf{z}_{1:k}) \simeq \mathcal{N}_{G_X}^L(\mathbf{X}_k^v; \hat{\mathbf{X}}_{k|k}^v, \boldsymbol{\Sigma}_{k|k}^v). \quad (3.62)$$

Pour y parvenir, nous procédons classiquement en deux étapes :

- l'étape de prédiction, qui est menée de manière équivalente à celle du *LG-FKEI*,
- l'étape de correction, dont nous proposons une formulation spécifique due à la forme du critère à optimiser et prenant en compte la présence des variables latentes $\mathbf{Z}_k = \{\mathbf{Z}_{k,i}\}_{i=1}^{n_k}$. Enfin, la matrice de covariance est approchée autour de l'optimum grâce à une méthode de Laplace sur groupe de Lie.

3.3.2.1 Étape de prédiction

En accord avec le modèle d'état augmenté (3.61), l'étape de prédiction considérée est exactement la même que celle de l'algorithme *LG-FKEI* classique, et nous approchons la distribution prédite sous la forme :

$$p(\mathbf{X}_k^v | \mathbf{z}_{1:k-1}) \simeq \mathcal{N}_{G_X}^L(\mathbf{X}_k^v; \hat{\mathbf{X}}_{k|k-1}^v, \boldsymbol{\Sigma}_{k|k-1}^v). \quad (3.63)$$

Comme expliqué dans le chapitre 2 à travers l'équation (2.127), les paramètres de la distribution jointe sont calculés en minimisant le critère suivant :

$$\begin{aligned} \{\mathbf{X}_{k|k-1}^{v,*}, \mathbf{X}_{k-1}^{v,*}\} = \underset{\mathbf{X}, \tilde{\mathbf{X}}}{\operatorname{argmin}} \quad & \left\| \operatorname{Log}_{G_X}^{\vee} \left(\mathbf{f}_k(\tilde{\mathbf{X}})^{-1} \mathbf{X} \right) \right\|_{\mathbf{W}_k^v}^2 \\ & + \left\| \operatorname{Log}_{G_X}^{\vee} \left(\left(\hat{\mathbf{X}}_{k-1|k-1}^v \right)^{-1} \tilde{\mathbf{X}} \right) \right\|_{\Sigma_{k-1|k-1}^v}^2, \end{aligned} \quad (3.64)$$

où $\hat{\mathbf{X}}_{k-1|k-1}^v$ et $\Sigma_{k-1|k-1}^v$ sont les paramètres de la distribution *a posteriori* à l'instant précédent. Celle-ci vérifie :

$$p(\mathbf{X}_{k-1}^v | \mathbf{z}_{1:k-1}) \simeq \mathcal{N}_{G_X}^L(\mathbf{X}_{k-1}^v; \hat{\mathbf{X}}_{k-1|k-1}^v, \Sigma_{k-1|k-1}^v). \quad (3.65)$$

En accord avec les développements du chapitre 2, il vient que :

$$\begin{cases} \hat{\mathbf{X}}_{k|k-1}^v = \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}^v) = \mathbf{X}_{k|k-1}^{v,*} \\ \Sigma_{k|k-1}^v = \mathbf{F}_k \Sigma_{k-1|k-1}^v \mathbf{F}_k^\top + \mathbf{W}_k^v \end{cases} \quad (3.66)$$

$$\quad (3.67)$$

avec :

$$\mathbf{F}_k = \left. \frac{\partial (\operatorname{Log}_{G_X}^{\vee}(\mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}^v)^{-1} \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}^v) \operatorname{Exp}_{G_X}^{\wedge}(\boldsymbol{\delta})))}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta}=\mathbf{0}} \quad (3.68)$$

3.3.2.2 Étape de correction

À l'instar du *LG-FKEI*, l'étape de correction est fondée sur une approximation de la distribution *a posteriori* sous la forme gaussienne concentrée décrite par l'équation (3.62).

De manière similaire à l'approche statique, la spécificité de notre algorithme est que les variables de nuisances $\mathbf{Z}_k = \{\mathbf{Z}_{k,i}\}_{i=1}^{n_k}$ sont intégrées dans le problème d'estimation et nous cherchons à estimer le nouvel état :

$$\mathbf{X}_k^{v,Z} = \operatorname{blkdiag}(\mathbf{M}_k^v, \mathbf{P}_k, \mathbf{D}_k, \mathbf{Z}_k) \quad (3.69)$$

appartenant au groupe $G_X^Z = G_X \times SE(3)^{n_k}$. Nous proposons alors d'estimer la distribution *a posteriori* de $\mathbf{X}_k^{v,Z}$:

$$p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k}) \simeq \mathcal{N}_{G_X^Z}^L(\mathbf{X}_k^{v,Z}; \hat{\mathbf{X}}_{k|k}^{v,Z}, \Sigma_{k|k}^{v,Z}). \quad (3.70)$$

où $\mathbf{X}_{k|k}^{v,Z}$ est obtenu par résolution du problème d'optimisation défini par :

$$\hat{\mathbf{X}}_{k|k}^{v,Z} = \underset{\mathbf{X}_k^{v,Z} \in G_X^Z}{\operatorname{argmin}} J^{(k)}(\mathbf{X}_k^{v,Z}) \quad (3.71)$$

où $J^{(k)}(\mathbf{X}_k^{v,Z}) = -2 \log p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k})$.

La règle de Bayes appliquée au modèle hiérarchique définie sur la figure 3.6, fournit de manière analogue au cas statique :

$$p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k}) \propto p(\mathbf{z}_k | \mathbf{Z}_k) p(\mathbf{Z}_k | \mathbf{X}_k^{v,Z}) p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k-1}), \quad (3.72)$$

où la différence fondamentale dans ce cas est que les variables $\mathbf{M}_k^v$, $\mathbf{P}_k$ et $\mathbf{D}_k$ sont corrélées *a posteriori* de sorte que leur loi jointe de prédiction n'est pas séparable.

Le critère d'optimisation s'écrit alors de la manière suivante :

$$J^{(k)}(\mathbf{X}_k^{v,Z}) = n_k \sum_{i=1}^6 \log(d_{k,i}) + \sum_{i=1}^{n_k} \left(\|\phi(\mathbf{z}_{k,i}, \mathbf{Z}_{k,i})\|_{\mathbf{U}_k}^2 + \|\psi(\mathbf{X}_k^v, \mathbf{Z}_{k,i})\|_2^2 \right) + \left\| \text{Log}_{G_X}^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1}^v \right)^{-1} \mathbf{X}_k^v \right) \right\|_{\Sigma_{k|k-1}^v}^2 + C, \quad (3.73)$$

où $C \in \mathbb{R}$ est une constante, les $\{d_{k,i}\}_{i=1}^6$ correspondent aux composantes diagonales de $\mathbf{D}_k$ et $\phi : SE(3) \rightarrow \mathbb{R}^3$ et $\psi : G_X^Z \rightarrow \mathbb{R}^6$ sont deux fonctions différentiables. $J^{(k)}$ peut donc se réécrire sous la forme :

$$J^{(k)}(\mathbf{X}_k^{v,Z}) = J_{nq}^{(k)}(\mathbf{D}_k) + J_q^{(k)}(\mathbf{X}_k^{v,Z}) \quad (3.74)$$

où $J_q^{(k)}$ s'écrit comme une norme vectorielle au carré $\left\| \boldsymbol{\varepsilon}^{(k)}(\mathbf{X}_k^{v,Z}) \right\|_{\Sigma_{k|k-1}^a}^2$ avec :

$$\boldsymbol{\varepsilon}^{(k)}(\mathbf{X}_k^{v,Z}) = \left[\Phi(\mathbf{Z}_k)^\top, \Psi(\mathbf{X}_k^v, \mathbf{Z}_k)^\top, \text{Log}_{G_X}^\vee \left(\left(\hat{\mathbf{X}}_{k|k-1}^v \right)^{-1} \mathbf{X}_k^v \right)^\top \right]^\top \quad (3.75)$$

$$\Sigma_{k|k-1}^a = \text{blkdiag}(\mathbf{U}_k \otimes \mathbf{I}_{n_k \times n_k}, \mathbf{I}_{6 \times 6}, \Sigma_{k|k-1}^v) \quad (3.76)$$

et Φ et Ψ sont les fonctions définies par les équations (3.39) et (3.40).

Pour résoudre le problème d'optimisation, nous déployons l'algorithme de Newton proposé dans la section précédente fondé la direction de descente optimale fournie par l'équation (3.51).

En effet, étant donné que le critère à optimiser est similaire à celui de l'équation (3.43), les jacobienues et les hessiennes entrant en jeu ont les mêmes expressions. Elles sont détaillées dans l'annexe B.

Nous pouvons alors en déduire, au bout de L itérations de l'algorithme, une solution $\mathbf{X}_k^{v,Z(L)}$ proche d'un point critique du critère et vérifiant :

$$\mathcal{L}_{J(\mathbf{X}_k^{v,Z(L)})}^R \simeq \mathbf{0} \quad (3.77)$$

Ce point critique est utilisé en tant que premier paramètre de l'approximation gaussienne concentrée de sorte que $\hat{\mathbf{X}}_{k|k}^{v,Z} = \mathbf{X}_k^{v,Z(L)}$. L'algorithme d'optimisation étant intégré dans l'étape de correction du filtre, il est important de souligner qu'il bénéficie de sa convergence. Par conséquent, peu d'itérations sont nécessaires.

Pour identifier la covariance, nous décrivons maintenant l'approximation de Laplace.

3.3.2.3 Approximation de Laplace

Contrairement à l'algorithme *LG-FKEI* classique, il n'est pas possible d'effectuer une approximation de Gauss-Laplace pour en déduire la matrice de covariance estimée $\Sigma_{k|k}^v$.

Nous proposons d'utiliser une approximation de Laplace qui est fondée sur un développement

de Taylor-Young à l'ordre 2 du critère $J^{(k)}$.

Comme présenté dans la partie (3.2.2), il peut s'approcher en $\hat{\mathbf{X}}_{k|k}^{v,Z}$ sous la forme :

$$J^{(k)}(\mathbf{X}_k^{v,Z}) \simeq J^{(k)}(\hat{\mathbf{X}}_{k|k}^{v,Z}) + \left(\mathcal{L}_{J(\hat{\mathbf{X}}_{k|k}^{v,Z})}^R \right)^\top \text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) + \frac{1}{2} \left(\text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) \right)^\top \mathbf{H}(\hat{\mathbf{X}}_{k|k}^{v,Z}) \text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) \quad (3.78)$$

où $\mathcal{L}_{J(\hat{\mathbf{X}}_{k|k}^{v,Z})}^R$ et $\mathbf{H}(\hat{\mathbf{X}}_{k|k}^{v,Z})$ correspondent respectivement à la dérivée de Lie et à la hessienne sur groupe de Lie de $J^{(k)}$ calculées en $\hat{\mathbf{X}}_{k|k}^{v,Z}$.

Puisque $\mathcal{L}_{J(\hat{\mathbf{X}}_{k|k}^{v,Z})}^R \simeq \mathbf{0}$, l'équation précédente peut se réécrire :

$$J^{(k)}(\mathbf{X}_k^{v,Z}) \simeq J^{(k)}(\hat{\mathbf{X}}_{k|k}^{v,Z}) + \frac{1}{2} \left(\text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) \right)^\top \mathbf{H}(\hat{\mathbf{X}}_{k|k}^{v,Z}) \text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) \quad (3.79)$$

En se servant du fait que $J^{(k)}(\mathbf{X}_k^{v,Z}) = -2 \log p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k})$, il s'ensuit une approximation gaussienne concentrée de la distribution *a posteriori* sous la forme :

$$p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k}) \propto \exp \left(-\frac{1}{2} \left\| \text{Log}_{G_X^Z}^\vee \left(\left(\hat{\mathbf{X}}_{k|k}^{v,Z} \right)^{-1} \mathbf{X}_k^{v,Z} \right) \right\|_{2(\mathbf{H}(\hat{\mathbf{X}}_{k|k}^{v,Z}))^{-1}}^2 \right).$$

Par conséquent :

$$p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k}) \simeq \mathcal{N}_{G_X^Z}^L(\mathbf{X}_k^{v,Z}; \hat{\mathbf{X}}_{k|k}^{v,Z}, \Sigma_{k|k}^{v,Z}), \quad (3.80)$$

avec par identification :

$$\Sigma_{k|k}^{v,Z} = 2 \left(\mathbf{H}(\hat{\mathbf{X}}_{k|k}^{v,Z}) \right)^{-1}. \quad (3.81)$$

Pour obtenir la distribution *a posteriori* de $\mathbf{X}_k^v$, il est nécessaire d'extraire les variables latentes de $p(\mathbf{X}_k^{v,Z} | \mathbf{z}_{1:k})$.

Pour y parvenir, nous devons :

- conserver uniquement la partie correspondant aux composantes $\mathbf{X}_k^v$ dans la moyenne et la covariance de $\mathbf{X}_k^{v,Z}$, notées respectivement $\hat{\mathbf{X}}_{k|k}^v$ et $\Sigma_{k|k}^v$.
- approcher la distribution de $\mathbf{X}_k^v$ par :

$$p(\mathbf{X}_k^v | \mathbf{z}_{1:k}) \simeq \mathcal{N}_{G_X}^L(\mathbf{X}_k^v; \hat{\mathbf{X}}_{k|k}^v, \Sigma_{k|k}^v) \quad (3.82)$$

Algorithme 2 : Pseudo-code de l'algorithme de pistage proposé.

Entrées : $\left[\hat{\mathbf{X}}_{0|0}^v, \Sigma_{0|0}^v, \mathbf{W}_k^v, \{\{\mathbf{z}_{k,i}\}_{i=1}^{n_k}\}_{k=1}^N (N : \text{nombre d'instants}) \right]$

$k = 1$

tant que $k \leq N$ **faire**

 → **Étape de prédiction**

 • Propagation de l'état (centroïde + forme) :

$$\hat{\mathbf{X}}_{k|k-1}^v \leftarrow \mathbf{f}_k(\hat{\mathbf{X}}_{k-1|k-1}^v)$$

 • Propagation de la covariance associée :

$$\Sigma_{k|k-1}^v \leftarrow \mathbf{F}_k \Sigma_{k-1|k-1}^v \mathbf{F}_k^\top + \mathbf{W}_k^v$$

 → **Étape de correction**

 • Intégration des variables latentes dans les inconnus :

$$\mathbf{X}_k^{v,Z} = \text{blkdiag}(\mathbf{X}_k^v, \mathbf{Z}_k)$$

 • Mise à jour des paramètres selon l'algorithme (1) :

$$\left[\hat{\mathbf{X}}_{k|k}^{v,Z}, \mathbf{H}_J^{(L)} \right] \leftarrow \text{Algo-Newton}(\{\mathbf{z}_{k,i}\}_{i=1}^{n_k}, \hat{\mathbf{X}}_{k|k-1}^v)$$

 • Approximation de Laplace :

$$\Sigma_k^{v,Z} \leftarrow 2 \left(\mathbf{H}_J^{(L)} \right)^{-1}$$

 • Extraction des composantes de l'état :

$$[s_1, s_1] = \text{size}(\mathbf{X}_k^v)$$

$$\hat{\mathbf{X}}_{k|k}^v \leftarrow \hat{\mathbf{X}}_{k|k}^{v,Z} [1 : s_1, 1 : s_1] ;$$

$$\Sigma_{k|k}^v \leftarrow \Sigma_{k|k}^{v,Z} [1 : m_X, 1 : m_X] ;$$

 → $k = k + 1$

retourner $\{\mathbf{X}_{k|k}^v\}_{k=1}^N, \{\Sigma_{k|k}^v\}_{k=1}^N$

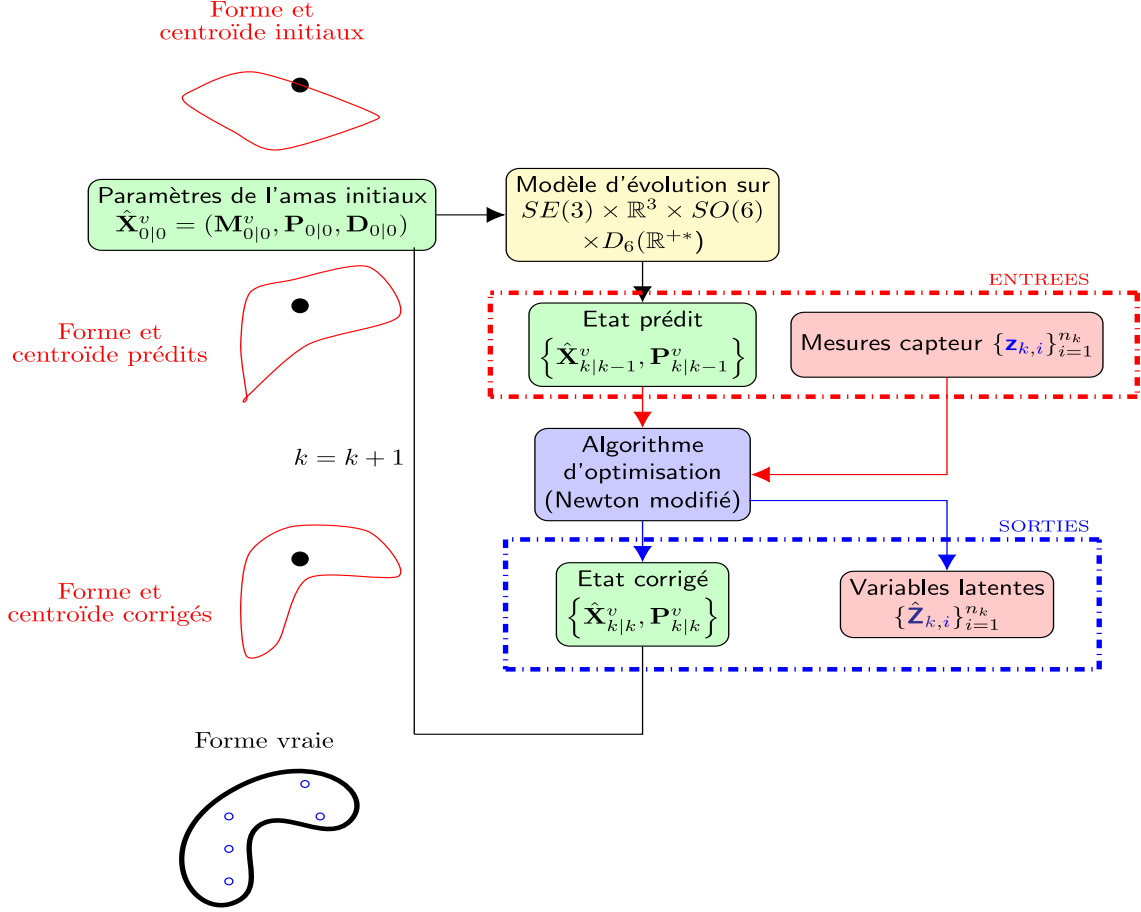


FIGURE 3.7 – Architecture de l'algorithme proposé.

3.4 Validation expérimentale des algorithmes

Dans cette section, nous mettons en oeuvre les deux algorithmes proposés afin d'évaluer leur performances. Chacune des deux méthodes est testée sur deux scénarios.

Elles sont premièrement appliquées à un scénario où les caractéristiques de l'amas ont été obtenues directement par simulation du modèle proposé sur groupe de Lie. Ensuite, il est considéré un scénario où la géométrie de l'amas résulte directement de son modèle physique d'évolution. Ce cas de figure permet d'étudier la pertinence de la paramétrisation sur groupe de Lie considérée.

Pour valider l'intérêt de l'approche proposée, elle est comparée à une méthode de l'état de l'art [38] dédiée au pistage d'une cible étendue quelconque. Un algorithme fondé sur les processus gaussiens est considéré car ils sont plus génériques que les modèles à base de matrices aléatoires.

3.4.1 Définitions des métriques d'estimation

Dans cette première partie, nous proposons de mettre en place les différentes métriques utilisées afin d'évaluer quantitativement les performances des algorithmes. En particulier, nous introduisons la notion de métrique intrinsèque sur groupe de Lie et de distance de Hausdorff.

3.4.1.1 Métrique intrinsèque sur groupe de Lie

Classiquement, il est possible d'évaluer les performances d'un algorithme d'estimation en utilisant des métriques usuelles telles que des normes euclidiennes. En effet, si $\hat{\mathbf{x}}$ est un estimateur quelconque d'un paramètre $\mathbf{x} \in \mathbb{R}^m$, alors l'erreur commise $\mathbf{e}$ peut s'écrire sous une forme additive :

$$\hat{\mathbf{x}} = \mathbf{x} + \mathbf{e}, \quad (3.83)$$

et :

$$\|\mathbf{e}\| = \|\mathbf{x} - \hat{\mathbf{x}}\|. \quad (3.84)$$

Dans le cadre de notre approche, les paramètres évoluent sur groupe de Lie : par conséquent, il semble nécessaire de définir une métrique prenant en compte la géométrie de leurs ensembles de définition.

Supposons un groupe de Lie matriciel G de dimension m , alors la relation entre une matrice inconnue $\mathbf{X} \in G$ et un estimateur $\hat{\mathbf{X}}$ peut s'écrire :

$$\hat{\mathbf{X}} = \mathbf{X} \text{Exp}_G^\wedge(\mathbf{e}) \quad (3.85)$$

où le terme d'erreur sur l'espace vectoriel isomorphe à l'algèbre de Lie, $\mathbf{e} \in \mathbb{R}^m$ vérifie, à condition d'être suffisamment petit :

$$\mathbf{e} = \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}). \quad (3.86)$$

Par conséquent, une possibilité d'évaluation de l'erreur intrinsèque entre $\mathbf{X}$ et $\hat{\mathbf{X}}$ est de considérer la norme suivante :

$$E = \|\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})\| \quad (3.87)$$

Il est également possible de considérer que le paramètre et son estimé s'écrivent selon une relation d'erreur obtenue *à droite* :

$$\hat{\mathbf{X}} = \text{Exp}_G^\wedge(\mathbf{e}') \mathbf{X} \quad (3.88)$$

et l'erreur intrinsèque *à droite* peut être définie par :

$$E' = \|\text{Log}_G^\vee(\hat{\mathbf{X}} \mathbf{X}^{-1})\|. \quad (3.89)$$

Dans le cas du groupe de Lie $G = SO(3)$, il est possible de montrer que (3.87) correspond à la distance géodésique entre $\mathbf{X}$ et $\hat{\mathbf{X}}$ [55]. Toutefois, elle ne respecte pas généralement la définition d'une distance, particulièrement pour le groupe de Lie $SE(3)$. Malgré cela, elle admet un sens physique et peut se comprendre comme la "différence" entre deux éléments d'un groupe reprojctée sur son espace euclidien associé. De plus, elle est cohérente avec les critères d'erreurs considérés dans nos algorithmes, hérités de l'expression de la loi gaussienne concentrée.

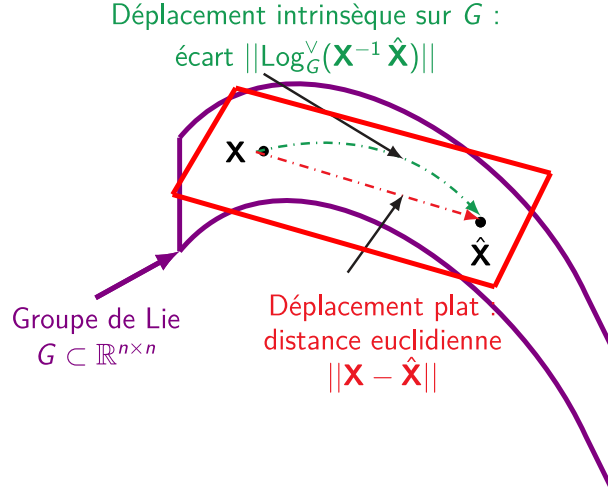


FIGURE 3.8 – Représentation géométrique de l'écart intrinsèque.

Dans le cadre de nos évaluations, nous proposons alors d'utiliser cette métrique d'erreur et de considérer en particulier l'erreur *à gauche* définie par l'équation (3.87). À partir de celle-ci, nous pouvons définir une *Erreur quadratique moyenne* (RMSE) intrinsèque *a posteriori* sur groupe de Lie et permettant, de manière analogue au cas euclidien, de mesurer la qualité de $\hat{\mathbf{X}}$. Elle peut s'écrire :

$$E = \mathbb{E}_{p(\mathbf{Z}, \mathbf{X})} \left(\|\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}})\|^2 \right) \quad (3.90)$$

$$= \int_{\mathbf{X} \in G} \int_{\mathbf{Z} \in G'} \|\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}})\|^2 p(\mathbf{Z}, \mathbf{X}) \mu_S(d\mathbf{Z}, d\mathbf{X}) \quad (3.91)$$

où $\mathbf{Z}$ désigne les observations qui sont définies sur un groupe de Lie G' . $\mathbb{E}$ est l'opérateur espérance classique, S est le groupe de Lie produit sur lequel évolue le couple $(\mathbf{Z}, \mathbf{X})$ et μ_S est une mesure de Haar sur S .

Elle peut donc être approchée par Monte-Carlo sous la forme :

$$E \simeq \sqrt{\frac{1}{N_r} \sum_{i=1}^{N_r} \|\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}}^{(i)})\|_2^2} \quad (3.92)$$

où chaque estimateur $\hat{\mathbf{X}}^{(i)}$, avec $i = 1, \dots, N_r$, est calculé à partir de réalisations différentes des bruits de modèles.

3.4.1.2 Métrique de comparaison de forme

Lorsque nous cherchons à évaluer nos algorithmes, nous avons besoin en particulier de déterminer les performances d'estimation de l'étendue de l'amas.

Comme nous venons de voir précédemment, l'équation (3.87) permet de calculer une erreur intrinsèque. Dans le cadre de notre approche, le couple $(\mathbf{P}_k, \mathbf{D}_k)$ caractérise directement la géométrie de l'amas. Soient $\hat{\mathbf{D}}_{k|k}$ et $\hat{\mathbf{P}}_{k|k}$ des estimateurs respectivement de $\mathbf{D}_k$ et $\mathbf{P}_k$, les

erreurs intrinsèques sur $\mathbf{P}_k$ et $\mathbf{D}_k$ définies par :

$$E_{\mathbf{P}_k} = \|\text{Log}_{SO(6)}^{\vee}(\mathbf{P}_k^{-1}\hat{\mathbf{P}}_{k|k})\| \quad (3.93)$$

$$E_{\mathbf{D}_k} = \|\text{Log}_{D_6(\mathbb{R}^{+*})}^{\vee}(\mathbf{D}_k^{-1}\hat{\mathbf{D}}_{k|k})\| \quad (3.94)$$

sont de bons indicateurs de la qualité de la reconstruction de la forme.

Cette erreur est uniquement interprétable au sens de notre modèle sur groupe de Lie. Par conséquent, si nous désirons comparer les performances de notre approche avec notre approche avec une méthode présentant une modélisation différente fondée sur un autre jeu de paramètres, alors ces métriques d'erreur ne peuvent pas être utilisées.

Pour pallier ce problème, nous proposons d'utiliser une distance générique permettant directement de comparer la forme de deux nuages de points et permettant de quantifier les erreurs de reconstruction en comparant les positions des différents points des nuages.

• Distance de Hausdorff :

La distance de Hausdorff est une distance permettant de quantifier la similarité entre deux nuages de points. Elle est régulièrement utilisée dans la littérature afin de mesurer le degré de ressemblance entre deux images en comparant les valeurs de niveaux de gris des pixels [88]. Formellement, elle peut être vue comme une quantité calculant l'“écart” entre deux ensembles S_1 et S_2 :

$$D^H(S_1, S_2) = \max(\{d(S_1, \mathbf{b}) | \mathbf{b} \in S_2\}) \quad (3.95)$$

où $d(A, b) = \min(\{d(a, b) | a \in A\})$ et d est une distance classique.

Lorsque S_1 et S_2 sont inclus dans $\mathbb{R}^3$, il est possible d'utiliser la distance euclidienne telle que $\forall(\mathbf{a}, \mathbf{b}) \in (\mathbb{R}^3)^2, d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|_2$.

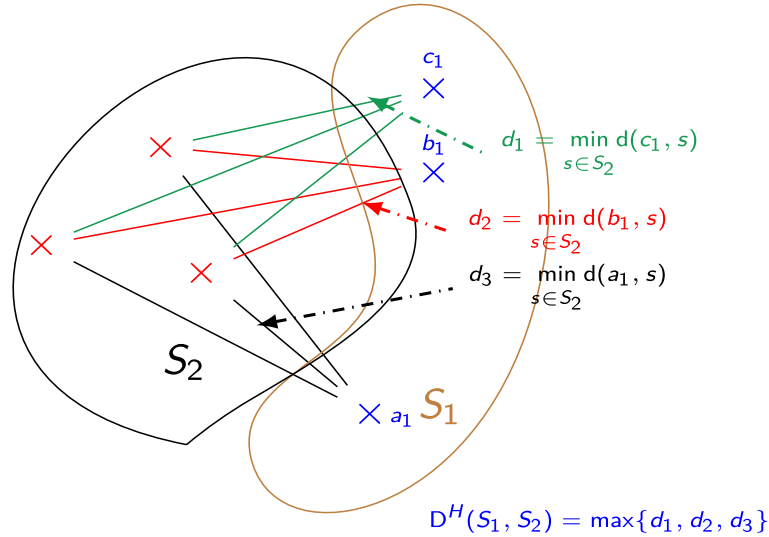


FIGURE 3.9 – Principe du calcul de la distance de Hausdorff dans le cas où deux ensembles S_1 et S_2 sont constitués de 3 points.

• Utilisation pratique : protocole de comparaison

Afin de comparer les reconstructions des formes obtenues, nous proposons d'utiliser cette distance de sorte à :

- calculer un écart entre le nuage des points prédit selon notre approche avec un nuage de points généré selon le vrai modèle.
- calculer l'écart entre un nuage des points prédit selon une approche alternative avec le nuage de points vrai.

Selon notre méthode, la forme est paramétrisée par la matrice de covariance $\mathbf{S}_k$ et les points prédits sont obtenus en :

1. générant N_p éléments prédits selon le modèle (3.4) en utilisant la matrice estimée $\hat{\mathbf{S}}_k$ comme matrice de covariance et l'état estimé $\hat{\mathbf{M}}_k$ comme centroïde,
2. conservant la partie position des éléments obtenus pour obtenir un ensemble de points $\{\hat{\mathbf{z}}_{GdL,k,i}\}_{i=1}^{N_p}$ représentant géométriquement la forme.

Comme notre approche permet de générer des points prédits à l'intérieur de l'amas, la méthode à laquelle nous souhaitons nous comparer, celle fondée sur les processus gaussiens, doit également prendre en compte cet aspect.

En reprenant les notations du chapitre 1, nous pouvons à partir de chaque estimation du vecteur des distances radiales $\hat{\mathbf{f}}_k = [\hat{f}_{k,1}, \dots, \hat{f}_{k,N_t}]$, générer N_p points prédits selon le modèle suivant :

$$\hat{\mathbf{z}}_{GP,k,j,l} = \hat{\mathbf{p}}_k + \varrho_{k,l} \hat{f}_{k,j} \mathbf{c}_{k,j} \quad \forall l \in \{1, \dots, N_l\} \quad \forall j \in \{1, \dots, N_t\} \text{ avec } \varrho_{k,l} \sim \mathcal{U}[0, 1] \quad (3.96)$$

Nous obtenons ainsi N_l points échantillonnés uniformément sur le segment $[\hat{\mathbf{p}}_k, \hat{\mathbf{p}}_k + \hat{f}_{k,j} \mathbf{c}_{k,j}]$ de sorte que $N_l N_t = N_p$.

La métrique d considérée est la distance euclidienne classique.

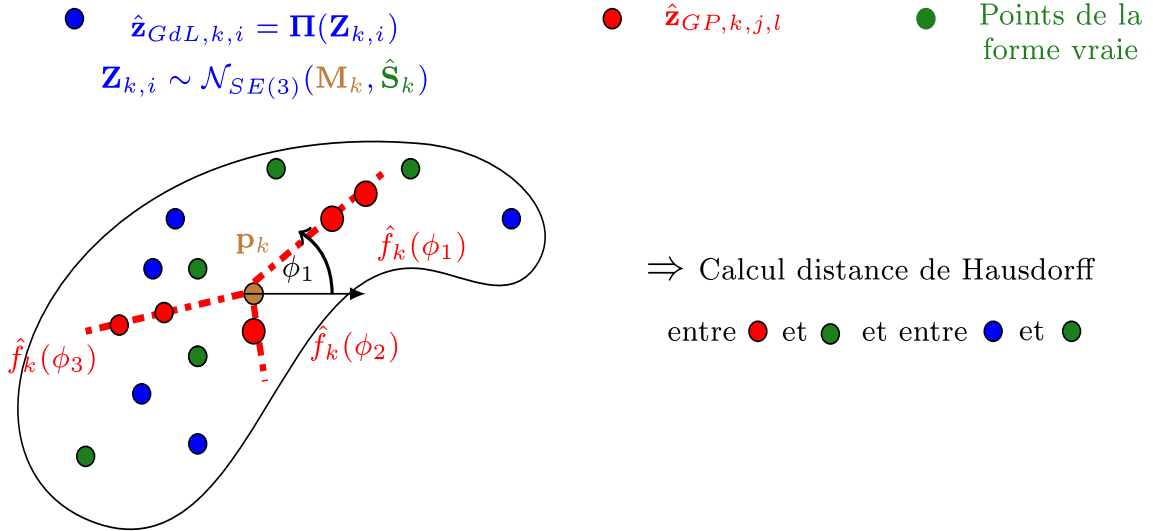


FIGURE 3.10 – Comparaison des points prédits aux points issus de la forme vraie.

3.4.2 Mise en oeuvre de l'algorithme statique

Dans cette partie, nous proposons de tester les performances de notre méthode à travers des simulations effectuées sur des données simulées selon notre modèle puis selon le modèle physique.

3.4.2.1 Premier scénario

• Simulation des données

Nous considérons dans ce protocole un ensemble de mesures réflecteurs générées de manière hiérarchique à partir du modèle *a priori* et du modèle d'observation proposé.

La matrice de forme vraie $\mathbf{S}$ est alors simulée :

- en échantillonnant $\mathbf{D}$ selon $\mathcal{N}_{SO(6)}^L(\boldsymbol{\mu}_D, \boldsymbol{\Sigma}_D)$
- en échantillonnant $\mathbf{P}$ selon $\mathcal{N}_{D_6(\mathbb{R}^{*+})}^L(\boldsymbol{\mu}_P, \boldsymbol{\Sigma}_P)$
- puis en calculant $\mathbf{S} = \mathbf{P}^\top \mathbf{D} \mathbf{P}$.

De la même manière, $\mathbf{M}$ est échantillonné directement selon $\mathcal{N}_{SE(3)}^L(\boldsymbol{\mu}_M, \boldsymbol{\Sigma}_M)$.

Les réflecteurs $\{\mathbf{Z}_i\}_{i=1}^n$ sont obtenus en substituant $\mathbf{M}$ et $\mathbf{P}$ et $\mathbf{D}$ dans le modèle(3.4).

De plus, nous supposons que les observations sont directement les mesures de position des réflecteurs. Par conséquent, la fonction $\mathbf{\Pi}$ récupère directement la partie translation de chaque $\mathbf{Z}_i$.

Enfin, la matrice de covariance du bruit de mesure est déterminée de sorte que ses valeurs correspondent aux incertitudes obtenues pour des mesures radars de distance et d'angles d'arrivée fournies par les équations (1.24), (1.25) et (1.26). Elle peut par exemple être calculée à partir d'une transformée sans parfum permettant de déterminer l'incertitude propagée à travers la non-linéarité de ces mesures radars.

• Réglages de l'algorithme

Les distributions *a priori* gaussiennes concentrées sur $\mathbf{P}$ et $\mathbf{D}$ définies selon les équations (3.28) et (3.29) doivent générer des couples $(\mathbf{P}, \mathbf{D})$ donnant lieu à des nuages de points suffisamment courbés.

- Les paramètres $\boldsymbol{\mu}_D$ et $\boldsymbol{\mu}_P$ peuvent être vus comme ceux caractérisant une forme “moyenne” *a priori*.
- Leurs incertitudes respectives $\mathbf{Q}_D$ et $\mathbf{Q}_P$ autorisent des formes “banane” plus ou moins éloignées de cette forme moyenne.

Lorsque nous construisons ce type *a priori*, il est à noter que l'écart entre les valeurs propres de la partie orientation de $\mathbf{D}$ et celles de sa partie position est significatif. En particulier, nous constatons qu'elles peuvent être très petites en orientation (de l'ordre de 10^{-7} rad^2) et très grande en position (de l'ordre de 10^6 m^2), dû à l'effet de bras de levier. Par conséquent, nous construisons un couple de paramètres $(\mathbf{Q}_D, \mathbf{Q}_P)$ prenant en compte cet aspect : les valeurs d'incertitude en orientation sont prises de l'ordre de grandeur de 10^{-7} rad^2 et celles en position de l'ordre de 10^6 m^2 .

L'*a priori* sur le centroïde $\mathbf{M}$ peut être divisé en un *a priori* sur son orientation et sur sa position. Ils sont conjointement distribués selon une loi gaussienne concentrée sur $SE(3)$. La composante rotationnelle de $\boldsymbol{\mu}_M$ est prise égale au vecteur nul afin d'obtenir une rotation moyenne égale à l'identité. Sa partie en translation est obtenue à partir d'une valeur *a priori* correspondant à une altitude de 10^6 m , dans le repère *ECEF*, et représentant une valeur physique cohérente pour une position d'un amas en orbite basse. Leurs incertitudes sont quantifiées par une matrice de covariance diagonale de valeurs 10^{-3} rad^2 et 10^4 m^2 respectivement sur les trois composantes en orientation et en translation.

• Initialisation

Afin de mettre en oeuvre l'algorithme proposé, nous devons choisir judicieusement les paramètres initiaux $\mathbf{M}_o$, $\mathbf{P}_o$ et $\mathbf{D}_o$.

- Nous construisons une matrice $\mathbf{S}_o$ dont la partie orientation est déterminée en considérant une erreur aléatoire sur la partie orientation des matrices vraies $\mathbf{P}$ et $\mathbf{D}$. La partie position est initialisée à partir de la matrice de covariance empirique de l'ensemble des réflecteurs. Une décomposition classique en valeurs propres et vecteur propres permet d'obtenir les deux initialisations $\mathbf{P}_o$ et $\mathbf{D}_o$.
- Concernant $\mathbf{M}_o$, sa position initiale peut être obtenue à partir de la moyenne empirique des mesures réflecteurs et son orientation initiale peut être déterminée par les vecteurs propres d'une analyse en composante principales sur le nuage de points généré par la matrice de forme *a priori*.
- Les matrices initiales $\mathbf{Z}_o$ sont déterminées en considérant que chacune d'entre elles possède une matrice de rotation égale à la rotation *a priori* du centroïde et que la partie position correspond à la position bruitée du réflecteur associé.

Les principaux paramètres de simulation sont précisés dans le tableau (3.1).

Nombre de paramètres inconnus	$m_Z = 27$
Nombre de réflecteurs	$n = 10$
Nombre d'itérations	$n_{it} = 12$
Nombre de simulations de Monte-Carlo	$N_r = 50$
Covariances <i>a priori</i>	$\mathbf{Q}_D = \text{diag}([10^{-7}, 10^{-7}, 10^{-7}, 10^6, 10^6, 10^6])$ $\mathbf{Q}_P = 10^{-4} \text{diag}([1, 1, 1, 1, 1, 1])$

TABLEAU 3.1 – Paramètres de simulation de l'algorithme.

• Résultats de simulation obtenus

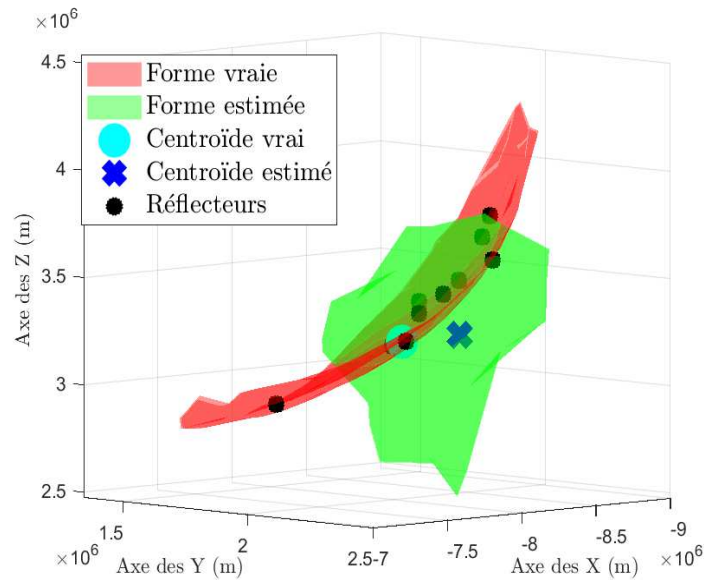


FIGURE 3.11 – Forme et centroïde reconstruits à la première itération.

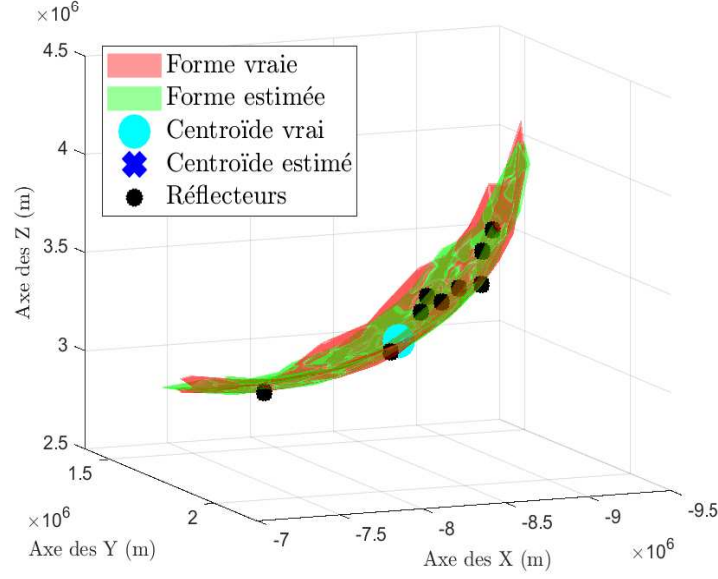


FIGURE 3.12 – Forme et centroïde reconstruits à la dernière itération.

Le comportement de l'algorithme peut d'abord être vérifié d'un point de vue visuel : des points prédits ainsi que des points vrais sont simulés à partir des modèles (3.4) et (3.7) dépourvus du bruit capteur, en utilisant respectivement la matrice vraie $\mathbf{S}$ et la matrice estimée $\hat{\mathbf{S}}$. Ils sont ensuite triangulés afin d'obtenir une représentation compacte des formes vraie et estimée.

- Sur la figure 3.11, nous affichons la forme théorique superposée à la forme initiale obtenue à la première itération de l'algorithme. Nous représentons également les centroïdes vrai et initial. Nous constatons que la forme initiale est de type ellipsoïdale puisque la matrice de covariance de forme est initialisée à partir de mesures euclidiennes.
- Sur la figure 3.12, nous représentons les mêmes formes à la dernière itération. Nous constatons alors que l'estimation est visuellement de très bonne qualité, puisque la forme prédite ainsi que le centroïde se superposent presque parfaitement à la forme et au centroïde théorique. La géométrie est en particulier très bien reconstruite.

Pour valider qualitativement ces observations, nous pouvons évaluer les performances de l'algorithme en :

- traçant le critère d'optimisation au cours des itérations,
 - calculant les erreurs intrinsèques sur $\mathbf{P}$, $\mathbf{D}$ et $\mathbf{M}$.
- Sur la figure 3.13, nous avons représenté le critère d'optimisation défini par l'équation (3.43) au fil des itérations. Nous constatons sa décroissance puis sa convergence au bout de 5 itérations ce qui confirme le bon comportement de notre algorithme.
 - Sur les figures 3.14 et 3.15, nous avons tracé les erreurs moyennes, pour chaque $\mathbf{X} \in \{\mathbf{M}, \mathbf{P}, \mathbf{D}\}$, grâce à la RMSE définie selon l'équation (3.92), et ceci pour N_r réalisations à chaque itération de l'algorithme. Nous observons que ces erreurs obéissent au même comportement que le critère d'optimisation : elles décroissent puis convergent vers une valeur fixe au bout de 6 à 7 itérations. Par conséquent, elles permettent de confirmer la stabilité et la convergence de l'algorithme.

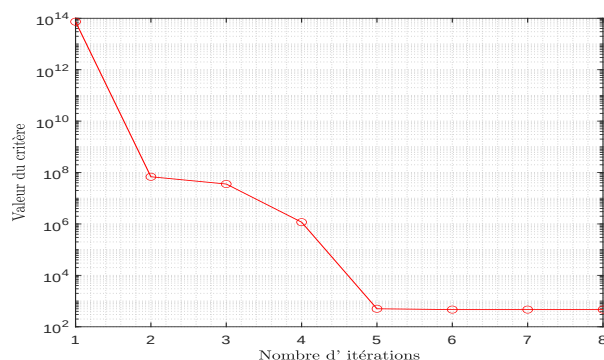


FIGURE 3.13 – Évolution du critère d'optimisation au cours des 8 premières itérations.

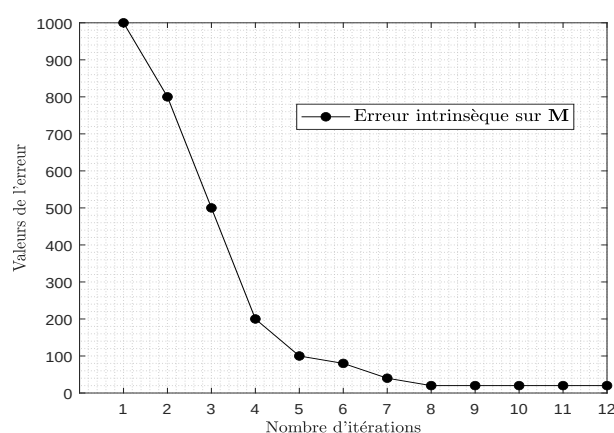


FIGURE 3.14 – Évolution de la RMSE intrinsèque sur M_k au cours des itérations.

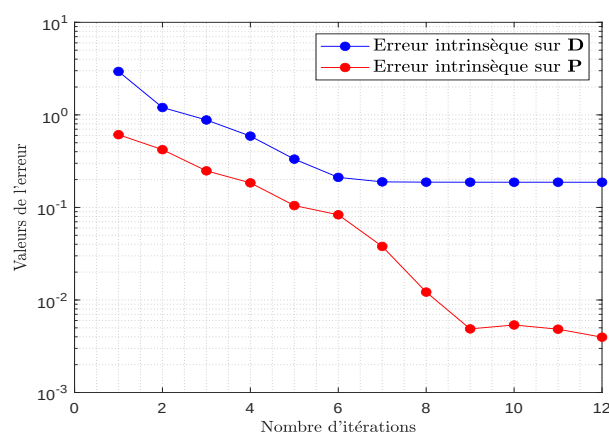


FIGURE 3.15 – Évolution de la RMSE intrinsèque sur D_k et P_k au cours des itérations.

3.4.2.2 Simulation selon le modèle physique

À présent, nous considérons que l'amas que nous cherchons à reconstruire est entièrement déterminé par le modèle de gravitation et non plus par simulation des modèles proposés. Pour

y parvenir, nous générons un amas de débris que nous faisons évoluer au cours du temps selon le modèle (1.3.2) du chapitre 1, puis nous récupérons ensuite son enveloppe.

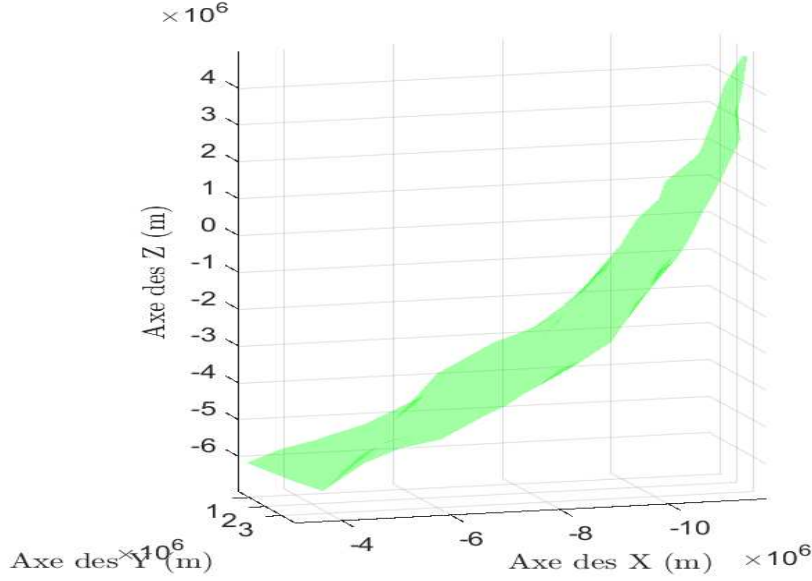


FIGURE 3.16 – Enveloppe de l'amas considéré pour le test de l'algorithme.

Les réflecteurs de l'amas sont obtenus en échantillonnant aléatoirement sur l'ensemble des points du nuage. Nous considérons que peu de mesures sont disponibles et nous fixons leur nombre à 10.

Dans le cadre de ce scénario, nous supposons que les mesures correspondent à des mesures radars d'angles d'arrivées et de distances associées à chacun des réflecteurs. Cette non-linéarité est prise en compte dans la définition de Π qui s'écrit sous la forme d'un opérateur :

$$\begin{aligned} \Pi : SE(3) &\rightarrow \mathbb{R}^3 \\ \mathbf{X} &\mapsto \mathbf{h}_{rad}(\mathbf{P}_{rad}(\mathbf{A}\mathbf{X}\mathbf{B} - \mathbf{p}_{rad})) \end{aligned} \quad (3.97)$$

où $\mathbf{P}_{rad}$ correspond à la matrice de passage du repère terrestre au repère radar.

$\mathbf{p}_{rad}$ est la position du radar dans le repère terrestre et $\mathbf{h}_{rad}$ est la fonction qui permet de transformer des coordonnées cartésiennes en coordonnées sphérique dont chaque composante s'écrit grâce aux équations (1.24), (1.25) et (1.26).

Enfin, $\mathbf{A} \in \mathbb{R}^{3 \times 4}$ et $\mathbf{B} \in \mathbb{R}^{4 \times 1}$ sont deux matrices permettant de prélever la partie translation d'un élément de $SE(3)$. L'incertitude sur chaque mesure est alors déterminée selon les formules de Woodward définies à travers les formules (1.13), (1.22) et (1.23) et leurs valeurs sont données dans le tableau (3.2).

Incertitude en distance	$\sigma_r = 150 \text{ m}$
Incertitude en élévation	$\sigma_\theta = 0.8^\circ$
Incertitude en azimuth	$\sigma_\phi = 0.3^\circ$

TABEAU 3.2 – Incertitudes de mesures.

• Intégration d'un *a priori* de forme

Nous considérons, comme précédemment, un *a priori* de sorte que les valeurs propres associées à la matrice de forme *a priori* ont des valeurs très petites en orientation et très grandes en translation.

De plus, nous constatons que l'étendue de l'amas est de l'ordre de $10^6 m^3$ de manière similaire à la forme simulée précédemment. Ainsi, les valeurs des incertitudes associées aux covariances $\mathbf{Q}_D$ et $\mathbf{Q}_P$ sont calculées pareillement.

• Résultats de simulation obtenus

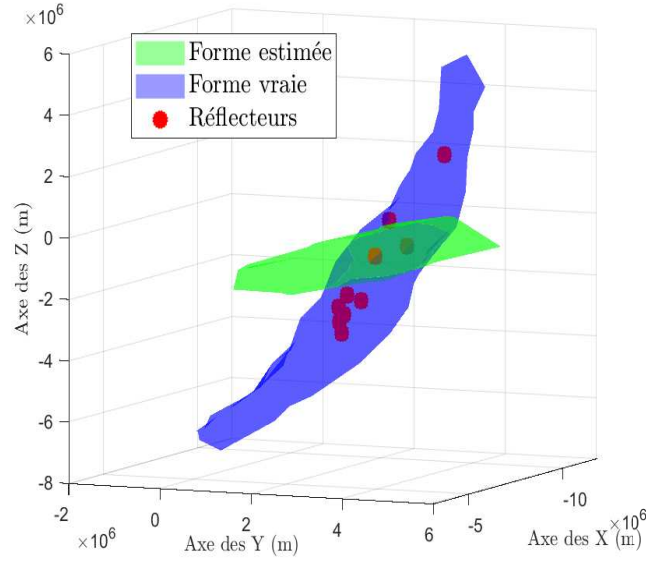


FIGURE 3.17 – Forme estimée à l'itération initiale superposée à la forme de l'amas et aux réflecteurs.

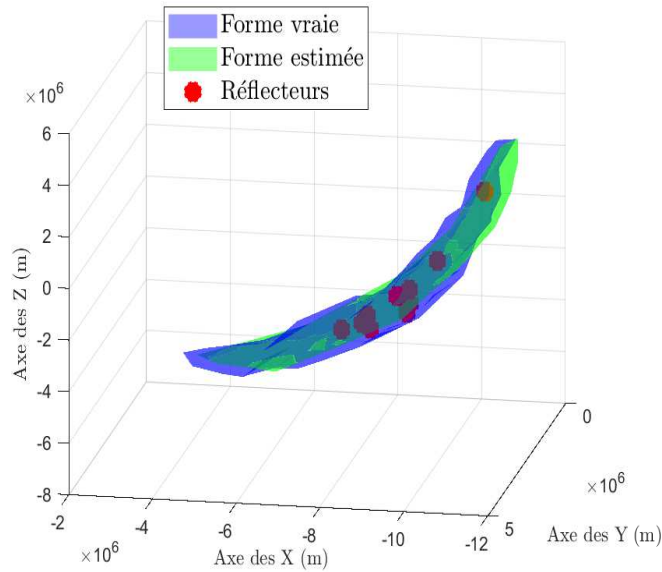


FIGURE 3.18 – Forme reconstruite selon un premier angle de vue.

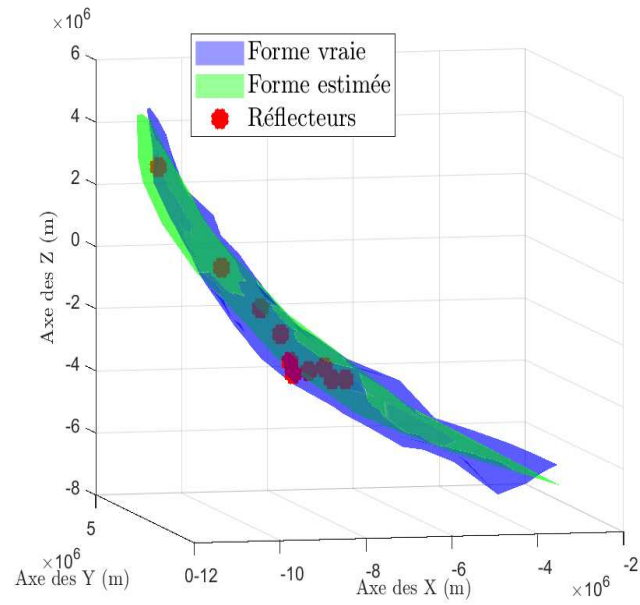


FIGURE 3.19 – Forme reconstruite selon un second angle de vue.

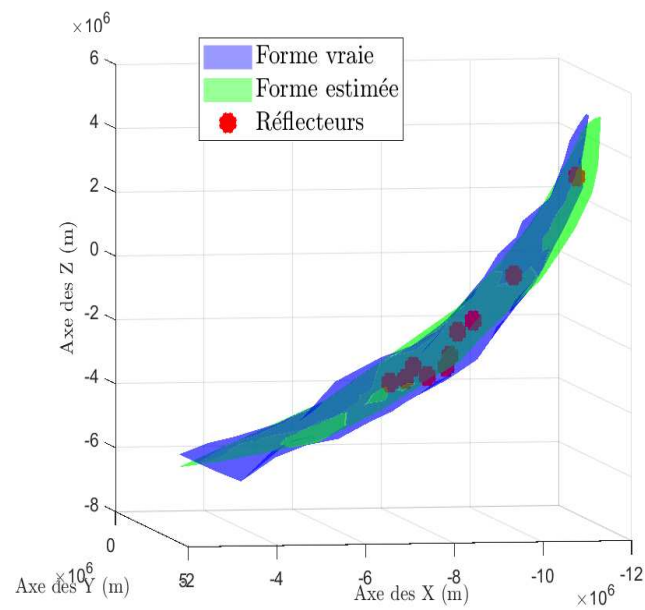


FIGURE 3.20 – Forme reconstruite selon un troisième angle de vue.

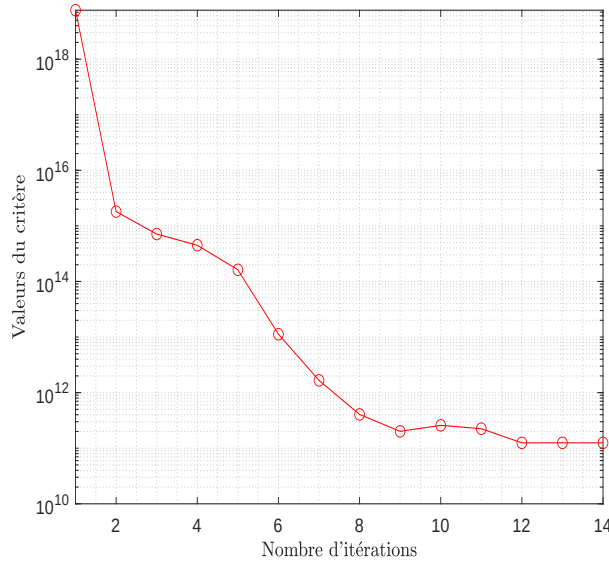


FIGURE 3.21 – Évolution du critère d'optimisation au cours des itérations.

Sur la figure 3.17, nous représentons la forme initiale estimée à partir de la matrice de covariance des réflecteurs. Les figures 3.18, 3.19 et 3.20 montrent l'estimation de la forme obtenue sous trois vues différentes afin de vérifier la bonne reconstruction en termes d'orientation, d'étendue et de courbure. Nous constatons que l'algorithme est en capacité de retrouver la géométrie de l'amas et particulièrement la courbure grâce à la combinaison de l'information des mesures avec l'information *a priori* de forme.

La convergence de l'algorithme peut être validée en traçant le critère d'optimisation en figure 3.21. De par les fortes non-linéarités introduites par la jacobienne des mesures radar combinées avec celles de l'opérateur $\text{Exp}^{\wedge}(\cdot)$, le critère a tendance à décroître assez doucement pour converger au bout d'une dizaine d'itérations (précisément 12). Néanmoins, son allure globale est cohérente en termes de décroissance et de convergence.

3.4.3 Mise en oeuvre de l'algorithme de pistage

À travers cette section, nous nous intéressons désormais à simuler et tester l'algorithme de pistage décrit par le pseudo-code 2.

À l'instar du cas statique, deux scénarios sont ainsi considérés. Dans une première partie, nous supposons que l'amas est généré selon les modèles d'évolution. Dans une seconde partie, les débris constituant l'amas sont supposés évoluer indépendamment les uns des autres selon le modèle de gravitation.

3.4.3.1 Premier scénario

• Protocole de simulation

Dans cette série d'expériences, nous simulons l'évolution de l'amas de débris selon le modèle d'état proposé par l'équation (3.61). À cet effet, nous supposons une durée de simulation de 15s. Il est à noter que cette valeur est en réalité trop petite pour que l'amas puisse se courber

à partir de son modèle dynamique. nous la fixons ainsi à des fins de validation de l'algorithme mais ajustons les hyperparamètres de façon à générer des bananaïdes.

- Les paramètres de forme $\mathbf{D}_k$ et $\mathbf{P}_k$ sont déterminés par les équations aux différences (3.59) avec une étendue spatiale initiale de 10^4 m^3 et une distorsion réaliste générée selon l'axe des abscisses. La matrice initiale $\mathbf{S}_o = \mathbf{P}_o^\top \mathbf{D}_o \mathbf{P}_o$ est obtenue selon le processus décrit dans la sous-section (3.4.2.1).
- La trajectoire du centroïde est gouvernée par les équations (3.52) et (3.53). Concernant sa matrice d'orientation, elle est supposée invariante au cours du temps et générée selon (1.24).
- Les écarts-types des bruits de modèle sont résumés dans le tableau (3.3).

Position	$\sigma_p = 1 \text{ m}$
Vitesse	$\sigma_v = 10^{-1} \text{ m s}^{-1}$
Orientation	$\sigma_R = 10^{-3} \text{ rad}$
Forme $\mathbf{D}_k$	$\sigma_{D_k} = [10^{-3}, 10^{-3}, 10^{-3}, 1, 1, 1]$
Forme $\mathbf{P}_k$	$\sigma_{P_k} = 10^{-4} \mathbf{1}_{15 \times 1}$

TABEAU 3.3 – Écarts-types des bruits des modèles d'état.

À l'instant initial, le centroïde est supposé être localisé à une altitude de 10^5 m avec une vitesse initiale de $7 \cdot 10^3 \text{ m s}^{-1}$. Cet ordre de grandeur est consistant avec le rayon orbital considéré. Les mesures réflecteurs sont supposées être disponibles chaque seconde et le nombre n_k de réflecteurs à l'instant k est supposé distribué selon une loi de Poisson de moyenne 10. Quant aux mesures, elles sont calculées selon les équations (3.4) et (3.7) avec un écart-type de bruit égal à 25 m, obtenu en projetant les incertitudes en angle et distance radar fournies par les formules de Woodward sur des coordonnées cartésiennes.

Erreur initiale sur $\mathbf{D}_0$	$\text{Exp}_{D_6 \mathbb{R}^{+*}}^\wedge([0.5, 0.5, 0.5, 500, 500, 500])$
Erreur initiale sur $\mathbf{P}_0$	$\text{Exp}_{SO(6)}^\wedge(10^{-2} \mathbf{1}_{15 \times 1})$
Erreur initiale sur $\mathbf{M}_0$	$\text{Exp}_{SE(3)}^\wedge([10^{-3}, 10^{-3}, 10^{-3}, 350, 350, 350])$
Nombre d'itérations de l'algorithme de Newton	$n_{it} = 5$

TABEAU 3.4 – Paramètres de l'algorithme.

Finalement, l'algorithme est initialisé comme indiqué dans le tableau (3.4). Il est à noter que puisque les mesures sont dix fois moins fréquentes que la période d'échantillonnage du modèle d'état, l'étape de prédiction de l'algorithme est répétée dix fois.

• Remarques sur le coût calculatoire

Le nombre de paramètres intrinsèques à estimer à chaque instant est égal à 30 qui est la dimension de G_X . De plus, à l'issue de l'étape de prédiction, nous devons prendre en compte la présence de n_k variables cachées, chacune de dimension 6. Nous nous retrouvons ainsi avec $6n_k + 30$ inconnus, aboutissant à un problème d'estimation en grande dimension.

D'un point de vue numérique, nous mettons à profit le faible nombre de mesures ainsi que la parcimonie des matrices jacobiniennes afin de réduire le coût calculatoire en particulier en termes d'inversion matricielle. L'étape de mise à jour est ainsi fondée sur un algorithme de Newton efficace et rapide de l'ordre de 0.08s.

• Performance et stabilité

Pour évaluer les performances d'estimation, nous avons visualisé au cours du temps l'évolution de la forme et de la position du centroïde reconstruits.

Sur la figure 3.24, nous observons que :

- l'algorithme arrive à recouvrer rapidement la position du centroïde (au bout de 5 secondes),
- les formes de l'amas estimée et véritable coïncident également assez rapidement. Plus précisément, la courbure de la forme est retrouvée au bout de 5 secondes.

Nous validons ces résultats numériquement en traçant les erreurs intrinsèques moyennes pour $\mathbf{M}_k$, $\mathbf{D}_k$ et $\mathbf{P}_k$. Elles sont représentées sur les figures 3.22 et 3.23 et calculées pour 50 réalisations de Monte-Carlo. Nous constatons alors que les erreurs d'estimation se stabilisent en peu d'itérations. Les valeurs finales pour la matrice de forme ainsi que pour les paramètres du centroïde, en particulier, sont très proches de zéro. Ces résultats corroborent les observations précédentes.

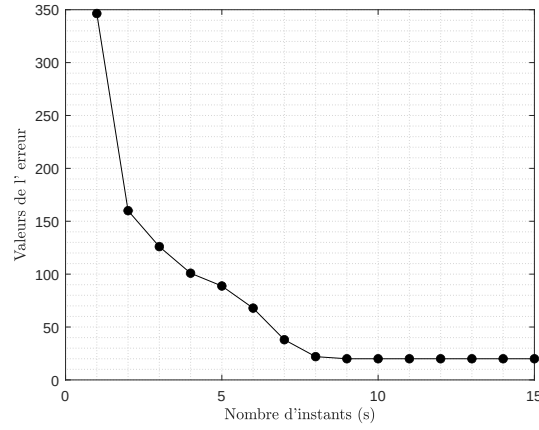


FIGURE 3.22 – Évolution de l'erreur d'estimation intrinsèque sur $\mathbf{M}_k$ au cours du temps.

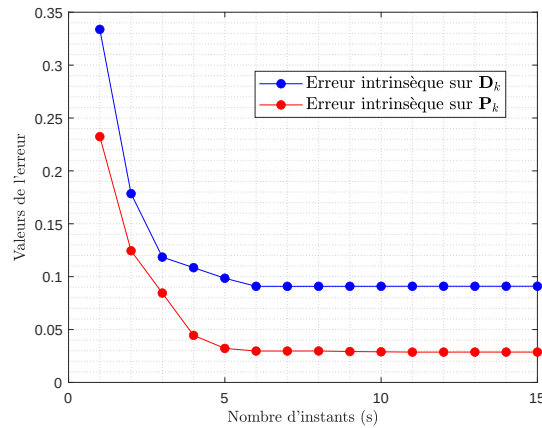


FIGURE 3.23 – Évolution de l'erreur d'estimation intrinsèque sur $\mathbf{P}_k$ et $\mathbf{D}_k$ au cours du temps.

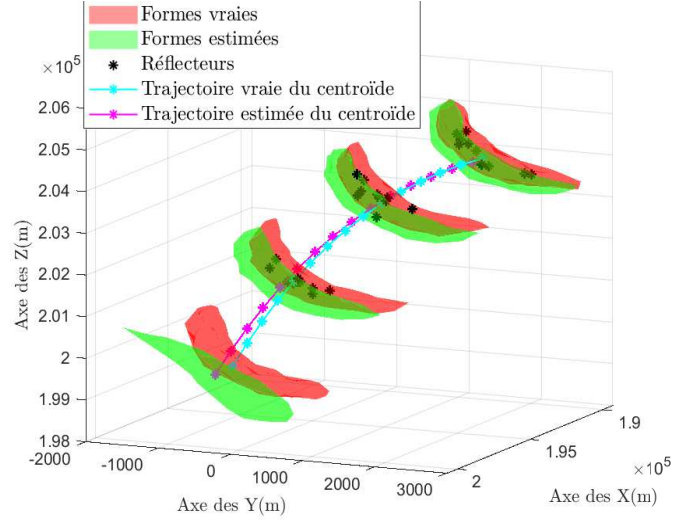


FIGURE 3.24 – Évolution des formes et de la trajectoire du centroïde estimées superposées aux formes et trajectoires du centroïde vraies.

• Robustesse et comparaison

Dans cette partie, nous proposons d'explorer la robustesse de notre approche et de la comparer avec les performances obtenues avec l'approche fondée sur les processus gaussiens. En particulier, nous nous intéressons à étudier l'influence du degré de distorsion ainsi que de l'étendue de l'amas. À cette fin, nous introduisons deux paramètres, notés s (m) et α (rad). Ils permettent de construire une matrice de forme $\mathbf{S}_k$ de sorte que l'ensemble des points générés selon (3.4) se dispersent plus ou moins loin avec une courbure variée.

- Le premier paramètre s contrôle l'étendue de l'amas et correspond à l'ordre de grandeur utilisé afin de générer le bras de levier.
- α est un paramètre angulaire contrôlant la distorsion de la forme à travers les composantes en rotation de $\mathbf{D}_k$.

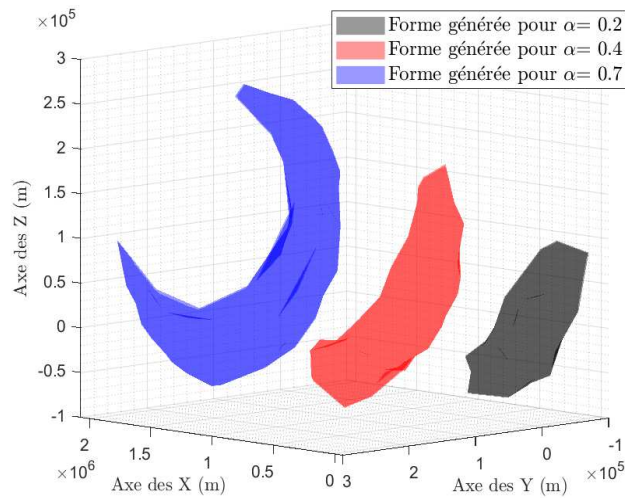


FIGURE 3.25 – Formes générées selon notre modèle pour différentes valeurs de α et $s = 10^5$ m.

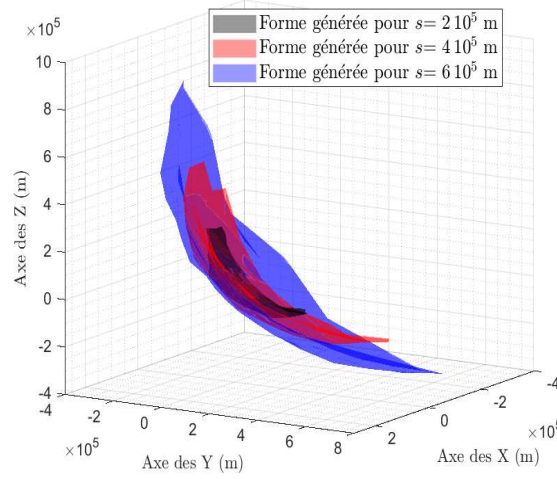


FIGURE 3.26 – Formes générées selon notre modèle pour différentes valeurs de s et $\alpha = 0.3 \text{ rad}$.

Pour étudier leur impact, nous faisons varier ces paramètres.

- Pour chacun d'eux, nous simulons un amas initial que nous faisons ensuite évoluer selon (3.61). Le même modèle d'évolution que pour notre méthode a été utilisé pour la cinématique du centroïde. Hormis la forme originelle, les paramètres de simulation sont réglés comme dans le scénario (3.4.3.1).
- Ensuite, notre méthode et celle fondée sur les processus gaussiens sont appliquées pour pister l'amas.
- Les distances de Hausdorff sont alors calculées pour les deux approches selon le protocole décrit dans la section (3.4.1.2). Elles sont moyennées à partir de 50 simulations de Monte-Carlo et sur toute la trajectoire de l'amas : nous les avons représentées sur les figures 3.27 et 3.28.

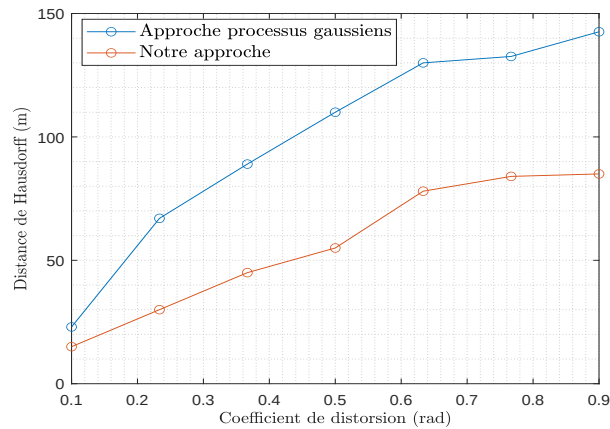


FIGURE 3.27 – Distance de Hausdorff moyenne en fonction de α pour notre approche et l'approche fondée sur les processus gaussiens pour $s = 10^2 \text{ m}$.

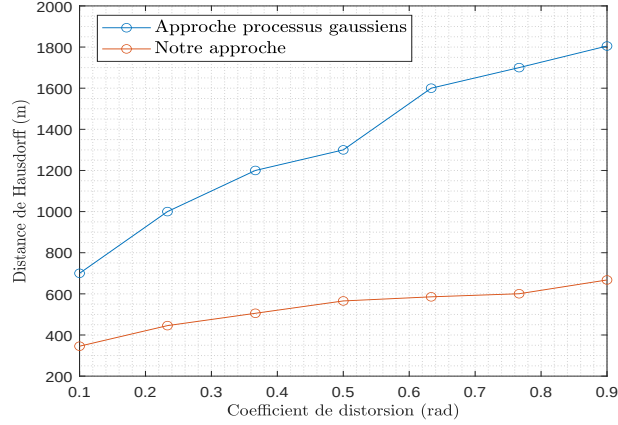


FIGURE 3.28 – Distance de Hausdorff moyenne en fonction de α pour notre approche et l'approche fondée sur les processus gaussiens pour $s = 10^4$ m.

Sur la figure 3.27, la distance de Hausdorff augmente plus rapidement pour la méthode à base de processus gaussiens quand le paramètre de distorsion devient plus grand. Ce comportement est d'autant plus marqué lorsque le paramètre s prend une grande valeur comme observée sur la figure 3.28. Les causes de ces observations sont les suivantes :

- premièrement, la modélisation par des processus gaussiens suppose que l'objet étendu à pister est convexe étoilé. Cette condition n'est en réalité plus respectée lorsque la courbure devient trop importante.
- Deuxièmement, notre approche introduit une information *a priori* sur la géométrie de la forme qui améliore significativement sa reconstruction quand un faible nombre de mesures (10 en moyenne dans ces scénarios) est disponible.

• Remarques

Pour la mise en oeuvre de l'approche fondée sur les processus gaussiens, nous avons supposé que le vecteur des distances radiales $\mathbf{f}_k$ évolue selon une marche aléatoire, de la même manière que les paramètres de la matrice d'étendue du modèle sur groupe de Lie :

$$\mathbf{f}_k = \mathbf{f}_{k-1} + \mathbf{n}_k \quad \mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{\mathbf{f}_k}) \quad (3.98)$$

où la matrice de covariance du bruit d'état $\mathbf{Q}^{\mathbf{f}_k}$ dépend du noyau du processus κ défini selon (1.119). De cette façon, en comparaison de la représentation sur groupe de Lie, nous n'introduisons pas d'informations supplémentaires sur la dynamique du contour de l'amas. Les seuls hyperparamètres jouant un rôle clé dans la précision de l'estimation sont σ_f^2 et l de l'équation (1.119). Afin de s'assurer que cette méthode n'est pas désavantagée, ils ont été estimés de manière à maximiser la vraisemblance par rapport aux mesures initiales comme expliqué dans [28]. Pour cela, une méthode classique de descente de gradient est appliquée.

3.4.3.2 Second scénario : simulation selon le modèle gravitationnel

• Protocole de simulation

À présent, nous considérons que l'évolution de l'amas est directement gouvernée par le modèle de gravitation : par conséquent, chaque débris suit le modèle d'évolution défini par l'équation (1.36). À cet effet, les positions et les vitesses initiales de ces derniers sont échantillonnées selon une distribution gaussienne multivariée dont la covariance détermine la forme ellipsoïdale du nuage initial. Ainsi, le centroïde euclidien initial de l'amas est confondu avec son centroïde géométrique. Les valeurs des paramètres permettant de simuler l'amas initial sont présentées dans le tableau (3.5).

Altitude	10^6m
Longitude	-0.6°
Latitude	44°
Vitesse	$8 \cdot 10^3 \text{m s}^{-1}$
Nombre de débris simulés	2000

TABLEAU 3.5 – Paramètres physiques de l'amas initial.

Chacun des débris évolue indépendamment selon les équations (3.52) et (3.53) où les écarts-types des bruit de modèles sont fixés aux mêmes valeurs que pour le scénario précédent.

Le temps de simulation est pris relativement grand, égal à $t = 3000 \text{ s}$, de sorte à pouvoir étudier le comportement à long terme des approches à comparer. À chaque instant de mesure, le nombre de réflecteurs détectés par le capteur radar est choisi comme une réalisation d'une loi de Poisson de moyenne 10. Ils sont tirés aléatoirement parmi les débris dans un rayon de 100 km autour du centroïde et nous supposons qu'ils sont disponibles avec une période fixée à 1 s. L'écart-type du bruit de mesure est pris égal à 25 m.

• Paramètres de l'algorithme

Les paramètres initiaux de l'algorithme sont obtenus à partir des caractéristiques de la forme initiale de l'amas :

- premièrement, la matrice de forme est construite comme la matrice de covariance des mesures et la position du centroïde est obtenue à partir de leur moyenne.
- Ensuite, la matrice d'orientation du centroïde est initialisée aléatoirement en considérant une réalisation d'une distribution gaussienne concentrée sur $SO(3)$.
- Enfin, les orientations des mesures sont initialisées en introduisant une erreur sur cette matrice d'orientation afin de générer des orientations dispersées autour d'une orientation moyenne pour chaque réflecteur.

• Analyse comparative

Sur les figures 3.29 et 3.30, les formes estimées et vraies de l'amas sont représentées pour notre méthode ainsi que pour l'approche à base de processus gaussiens, respectivement à cinq instants : $t = 1000, 1800, 2200, 2600$, et 3000 s . Nous pouvons remarquer que :

- lorsque sa forme reste relativement ellipsoïdale, globalement jusqu'à $t = 1800 \text{ s}$, les deux méthodes ont des performances similaires.
- Néanmoins, quand le degré de distorsion et l'étendue de l'amas augmentent, entre 1800 s et 3000 s , notre modélisation permet de mieux capturer son enveloppe contrairement à l'approche par processus gaussiens. En effet, cette dernière tend à sous-estimer l'étendue et aplatir la forme : elle échoue à reproduire sa courbure.

Pour quantifier les performances de reconstruction, nous calculons les distances de Hausdorff moyennes entre les amas vrais et estimés tout au long de la trajectoire. Pour y parvenir, 50

simulations de Monte-Carlo sont une nouvelle fois utilisées. Il est à noter que cette métrique prend en compte non seulement l'erreur en forme mais également en position du centroïde. Les erreurs obtenues sont tracées sur la figure 3.31. En accord avec l'observation visuelle, elles sont plus élevées pour l'approche fondée sur les processus gaussiens et augmentent au cours du temps, au fur et à mesure que la courbure de l'amas devient plus prononcée et que son étendue spatiale croît. Nous remarquons en particulier la consistance par rapport aux résultats établis dans la sous-section (3.4.3.1) sur l'influence de la courbure et de l'étendue.

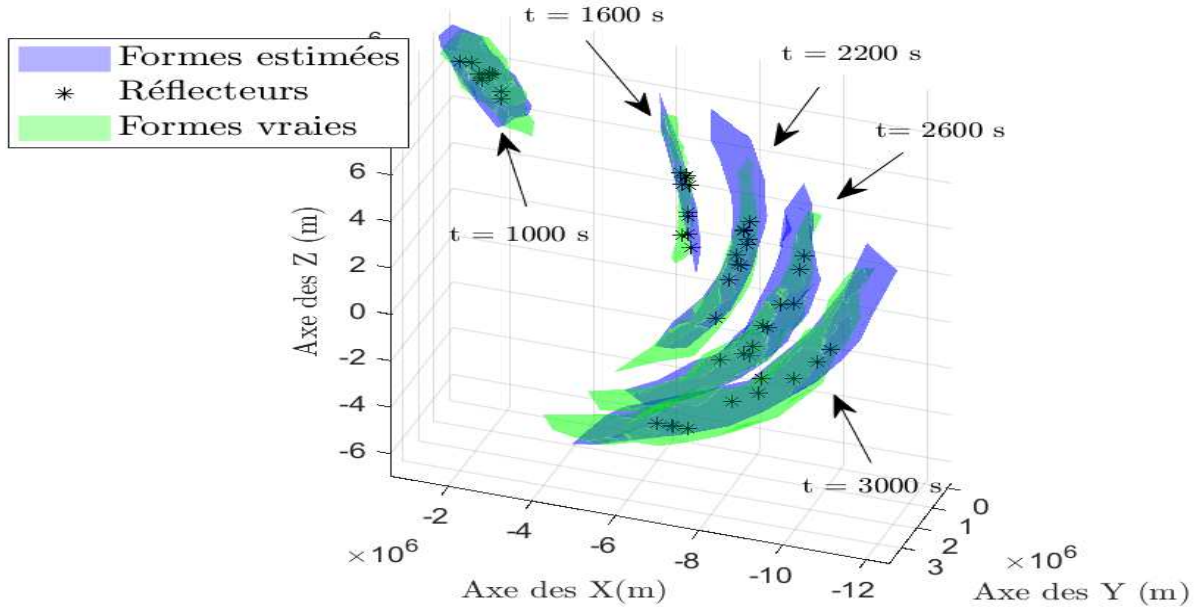


FIGURE 3.29 – Formes estimées superposées aux formes vraies pour différents instants : notre approche.

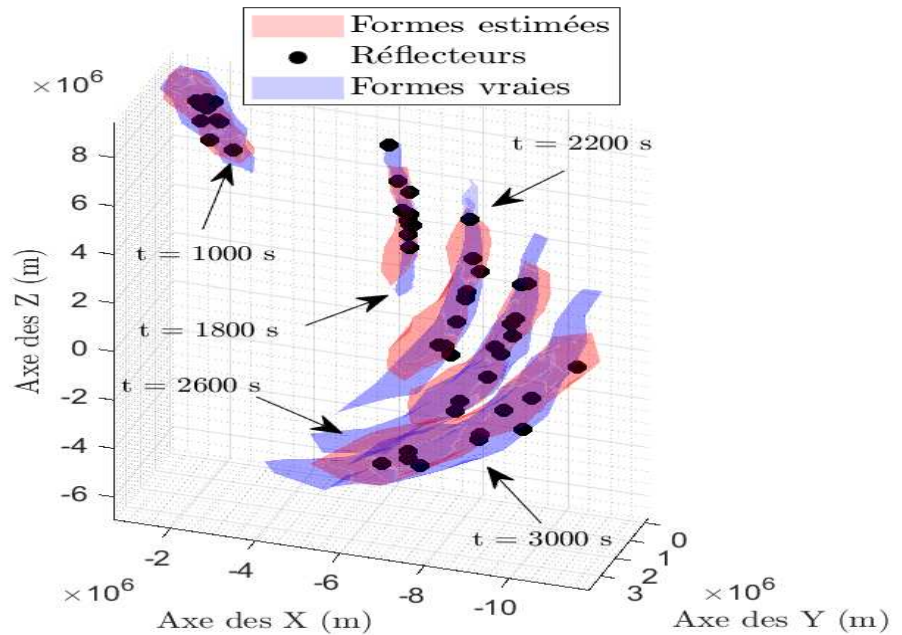


FIGURE 3.30 – Formes estimées superposées aux formes vraies aux mêmes instants : approche processus gaussiens.

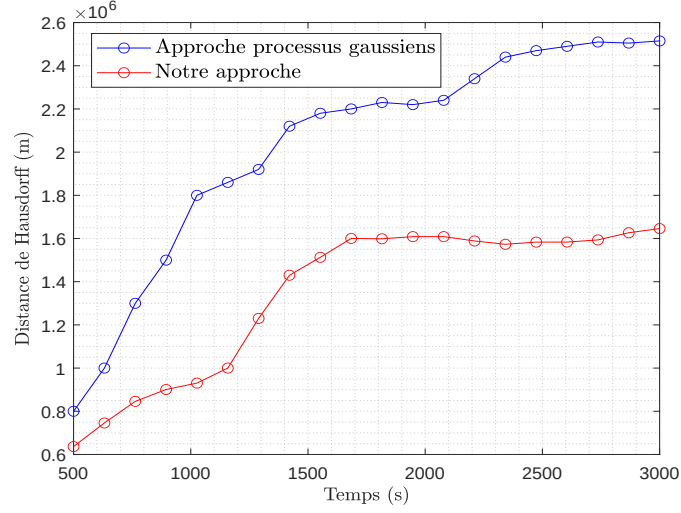


FIGURE 3.31 – Représentation des distances de Hausdorff moyennées en fonction du temps.

Comme indiqué dans la section (3.4.3.1), les hyperparamètres de la méthode fondée sur les processus gaussiens ont été obtenus en maximisant la vraisemblance des mesures. Cependant, ce choix ne garantit pas de fournir une distance de Hausdorff minimale entre les formes vraies et estimées. Pour étudier leur impact, nous avons mise en oeuvre la méthode pour différents couples $\{\sigma_f, l\}$ tout en conservant les mêmes paramètres de simulation de ce scénario. Les distances de Hausdorff moyennées sur le temps et sur 50 réalisations de Monte-Carlo correspondant à différentes trajectoires de l'amas sont représentées sur la figure 3.32.

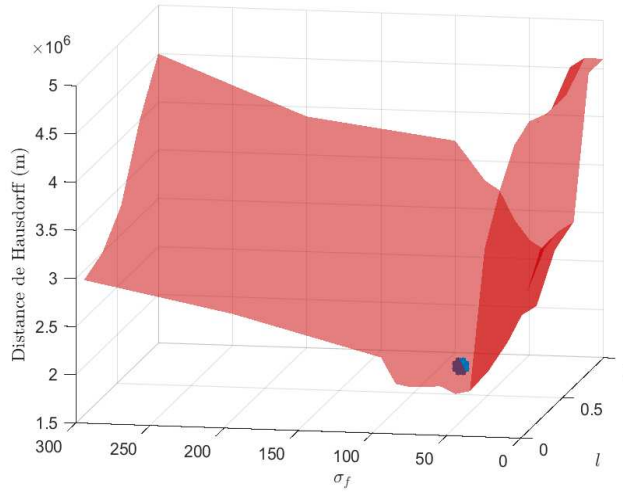


FIGURE 3.32 – Distances de Hausdorff obtenues pour différentes valeurs de σ_f comprises entre 1 et 300 et l entre 0 et 1. Le point bleu correspond à la distance de Hausdorff obtenue en déterminant les paramètres au sens du maximum de vraisemblance.

Nous observons, d'une part, qu'elles admettent un minimum égal à $1.92 \cdot 10^6$ m pour un couple $(\sigma_f^*, l^*) = (50, 0.11)$, et d'autre part, que la distance de Hausdorff associée au maximum de vraisemblance est proche de ce minimum et vaut $2.01 \cdot 10^6$ m. Cela valide la pertinence de ce

dernier. Enfin, quelles que soient les valeurs des hyperparamètres, notre méthode surpasse l'approche concurrente avec une distance de Hausdorff égale à $1.34 \cdot 10^6$ m.

- Limites de l'approche

Cette approche donne de bons résultats sur nos scénarios considérés, mais reste néanmoins sensible aux paramètres *a priori* considérés dans l'algorithme. En effet, selon l'amas considéré, le réglage des hyperparamètres peut entraîner des problèmes numériques. Ainsi, cela nécessite de définir des paramètres ayant des ordres de grandeurs parfois très faibles (de l'ordre de 10^{-7}). En particulier, si l'ordre de grandeur de l'étendue de la forme à estimer est très grand, ces paramètres peuvent prendre des valeurs très proches de zéro et poser des problèmes de conditionnement des matrices. Ils apparaissent notamment pour inverser la matrice hessienne. Lorsque la dispersion de l'amas devient très grande, un pistage de type cible étendue ne serait par conséquent pas forcément judicieux et un pistage multi-cibles pourrait être considéré afin d'estimer individuellement les trajectoires de chaque débris de l'amas.

3.5 Conclusions du chapitre

Dans ce chapitre, les premières contributions, en termes de modélisation et algorithmie ont été présentées. Nous avons proposé une nouvelle représentation d'état de type *matrice aléatoire* sur $SE(3)$ permettant de prendre en compte la configuration géométrique particulière de l'amas de débris. Selon ce modèle, il est ainsi décrit par une matrice de covariance modélisant sa forme et un état sur $SE(3)$ modélisant son centroïde géométrique.

Ensuite, deux algorithmes d'estimation ont été détaillés afin de déterminer une estimation de ces paramètres.

Le premier est un algorithme hors-ligne, réalisant l'estimation de manière statique et inspiré d'un algorithme de Newton. Il est adapté à la structure de notre modèle d'observation.

Le second est un algorithme de pistage récursif, fondé sur notre modèle sur groupe de Lie. Il se présente comme une variante du *LG-FKEI* permettant de prendre en compte l'évolution de l'amas au cours du temps et dans lequel la mise à jour des paramètres à chaque instant est obtenue grâce à l'algorithme 1.

Nos analyses expérimentales permettent d'observer leur bon comportement ainsi que des performances satisfaisantes en particulier en terme de reconstruction de l'étendue de l'objet.

La comparaison avec une méthode de l'état de l'art nous a permis de justifier que cette approche est plus adaptée géométriquement pour estimer les caractéristiques de ce type d'amas. Il peut être délicat d'interpréter l'ordre de grandeur des erreurs obtenues. À cet effet, nous nous intéressons dans le chapitre suivant au calcul de bornes sur l'erreur d'estimation pour des problèmes formulés sur groupe de Lie.

Chapitre 4

Mise en place d’une borne bayésienne d’estimation sur groupe de Lie

Sommaire

4.1	État de l’art des bornes de performance	126
4.1.1	État de l’art des bornes euclidiennes	126
4.1.2	État de l’art des bornes de performances sur variétés	127
4.2	Proposition d’une borne de Cramér-Rao sur groupe de Lie	128
4.2.1	Motivations	128
4.2.2	Expression générale	129
4.2.3	Hypothèses sur $\hat{\mathbf{X}}$ et expression approchée de la borne	133
4.2.4	Expression analytique	134
4.3	Mise en oeuvre sur le modèle d’observation proposé	138
4.3.1	Expression de $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$	139
4.3.2	Expression de $\mathbb{E}(\mathbf{H}_{\mathbf{M}}(\mathbf{M}))$	142
4.3.3	Simulations	143
4.4	Conclusion du chapitre	146

L’objectif de ce chapitre est de s’intéresser à la manière de quantifier les erreurs d’estimation d’un algorithme estimant des paramètres évoluant sur groupe de Lie. Dans le cadre de nos expériences, nous avons été capables d’évaluer des erreurs moyennes intrinsèques d’estimation. Nous avons constaté que ces erreurs ont un bon comportement car elles convergent vers une faible valeur proche de zéro. Néanmoins, la question que nous pouvons légitimement nous poser est “quelle est la plus petite valeur atteignable de l’erreur selon les modèles utilisés?”. Elle admet une réponse dans le cadre de l’estimation sur espace euclidien à travers les bornes de performance et en particulier la *borne de Cramér-Rao* et la *borne de Cramér-Rao bayésienne*. Sur groupe de Lie, qui plus est dans un cadre bayésien, il n’est pas trivial de pouvoir calculer une borne d’erreur à partir d’une erreur d’estimation de type intrinsèque.

Notre contribution dans ce chapitre est de proposer une nouvelle borne d’erreur bayésienne sur groupe de Lie minorant l’erreur moyenne d’estimation intrinsèque définie par (3.92).

Dans la première partie du chapitre, un rappel sur les bornes de performances sur espace euclidien est effectué.

Ensuite, les bornes de performances déjà existantes sur variétés riemanniennes et sur groupe de Lie sont rappelées.

Enfin, la dernière partie du chapitre est consacrée à la présentation et aux détails théoriques de la borne proposée. Elle est validée sur un scénario d’estimation du centroïde géométrique d’un amas de débris.

4.1 État de l'art des bornes de performance

4.1.1 État de l'art des bornes euclidiennes

Classiquement, lorsque nous cherchons à estimer un paramètre euclidien, il est pertinent de déterminer l'erreur d'estimation moyenne minimale commise selon un certain modèle d'observation de ce paramètre. Les bornes de performance d'un estimateur sont naturellement apparues dans des contextes d'inférence dans des applications diverses [89] au milieu du 20^{ème} siècle. La première borne proposée a été développée par C. Rao et H. Cramér sous le nom bien connu de *borne de Cramér-Rao* [90]. Dans un contexte fréquentiste où le paramètre inconnu n'est pas probabilisé, elle fournit un minorant sur l'erreur d'estimation moyenne quadratique d'un paramètre dans le cas où son estimateur est non biaisé. Une expression généralisée peut en être obtenue grâce à la *borne de Battacharya* [91]. Dans un contexte bayésien, il est nécessaire de prendre en plus en compte l'information *a priori* introduite sur la paramètre inconnu. une famille de bornes spécifiques, nommée *Ziv-Zakai* [92], a plus récemment été développée permettant de définir des erreurs minimales atteignables sur des problèmes de détection en traitement du signal. Une deuxième famille de bornes bayésiennes fondée sur le principe d'*inégalité sur la covariance*, nommée *Wess-Weinstein*, existe également et a abouti à la proposition de la *borne de Wess-Weinstein* ainsi que de la *borne de Cramér-Rao bayésienne* [93][94].

Selon les contextes d'applications, des améliorations hybrides de ces bornes ont été également développées, dans le cas où le paramètre inconnu présente une partie aléatoire et une autre déterministe [95]. Dans la suite de cette section, nous nous intéressons plus précisément à rappeler la borne de Cramér-Rao que nous nous attachons à généraliser sur groupe de Lie.

4.1.1.1 Borne de Cramér-Rao fréquentiste

Supposons qu'un paramètre $\boldsymbol{\theta} \in \mathbb{R}^p$ soit inconnu et qu'un ensemble d'observations $\{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ distribuées selon la vraisemblance $p(\mathbf{z}|\boldsymbol{\theta})$ soit disponible. Notons de plus $\hat{\boldsymbol{\theta}}$, un estimateur fréquentiste et non biaisé de $\boldsymbol{\theta}$ vérifiant la propriété :

$$\int_{\boldsymbol{\theta} \in \mathbb{R}^p} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) p(\mathbf{z}|\boldsymbol{\theta}) d\mathbf{z} = \mathbf{0}. \quad (4.1)$$

L'erreur quadratique moyenne s'écrivant :

$$E = \mathbb{E}_{p(\mathbf{z}|\boldsymbol{\theta})} (||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||^2), \quad (4.2)$$

vérifie :

$$E > \text{tr}(\mathbf{J}^{-1}), \quad (4.3)$$

où :

$$\mathbf{J} = -\mathbb{E}_{p(\mathbf{z}|\boldsymbol{\theta})} \left(\frac{\partial^2 \log p(\mathbf{z}|\boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}} \right). \quad (4.4)$$

correspond à la matrice d'information de Fisher de $\boldsymbol{\theta}$. tr correspond à l'opérateur trace de matrice et $\text{tr}(\mathbf{J}^{-1})$ est appelée la borne de Cramér-Rao paramétrique.

4.1.1.2 Borne de Cramér-Rao bayésienne

Supposons désormais que le paramètre $\boldsymbol{\theta}$ est aléatoire et distribué selon une loi *a priori* $p(\boldsymbol{\theta})$. Pour tout estimateur *a posteriori* $\hat{\boldsymbol{\theta}}_p$, l'erreur moyenne quadratique jointe s'écrit :

$$E_b = \mathbb{E}_{p(\mathbf{z}, \boldsymbol{\theta})} \left(\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_p\|^2 \right) \quad (4.5)$$

Il est démontré dans [20] qu'elle vérifie l'inégalité suivante :

$$E_b \geq \text{tr} \left(\mathbb{E}_{p(\mathbf{z}|\boldsymbol{\theta})} \left(\nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta})^\top \right) \right) \quad (4.6)$$

où $\nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta})$ correspond au gradient de $\log p(\mathbf{z}, \boldsymbol{\theta})$ en $\boldsymbol{\theta}$. Si $p(\boldsymbol{\theta})$ converge suffisamment vite vers 0 lorsque $\boldsymbol{\theta}$ prend de grandes valeurs, alors :

$$E_b \geq \text{tr}(\mathbf{J}_b^{-1}) \quad (4.7)$$

où $\mathbf{J}_b$ correspond à la matrice d'information de Fisher jointe s'écrivant :

$$\mathbf{J}_b = -\mathbb{E}_{p(\mathbf{z}, \boldsymbol{\theta})} \left(\frac{\partial^2 \log p(\mathbf{z}, \boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}} \right) \quad (4.8)$$

- Lorsque l'estimateur considéré est la moyenne *a posteriori*, la borne est atteinte car il est optimal au sens de la minimisation de l'erreur (4.5).
- Il est également proposé dans [20] une borne de Cramér-Rao récursive prenant en compte des modèles dynamiques sur les paramètres inconnus. Dans le cas où les modèles sur les paramètres et les observations sont linéaires gaussiens, l'expression récursive est la trace de la matrice de covariance calculée récursivement par le filtre de Kalman. De plus, elle est atteinte.

4.1.2 État de l'art des bornes de performances sur variétés

Comme décrit dans le chapitre 2, estimer un paramètre évoluant sur variétés est devenu crucial dans diverses applications en traitement du signal, en robotique ou dans le domaine biomédical [49][55][66][96]. Pour le résoudre, plusieurs algorithmes dédiés prenant en compte les contraintes géométriques de ces ensembles de définition ont été développés sur variétés riemanniennes et sur groupe de Lie [70]. Pour évaluer leurs performances, il est capital de pouvoir déterminer, comme dans le cas euclidien, l'erreur d'estimation moyenne minimale théorique pouvant être atteinte. Il se pose alors naturellement la question de déterminer des bornes de performances d'estimation sur des structures de types variétés.

Afin d'y parvenir, plusieurs difficultés doivent être traitées :

- premièrement, la métrique euclidienne classique n'est pas adaptée pour le calcul d'erreurs d'estimation sur variétés. Par conséquent, des métriques intrinsèques doivent être considérées.
- Deuxièmement, les calculs à effectuer aboutissent à des intégrales de densité de probabilité qui ne sont pas définies selon la mesure de Lebesgue.
- Enfin, de manière analogue au cas euclidien, il est nécessaire de définir une dérivée du second ordre, qui dépend de la connexion affine utilisée.

Dans la littérature, des généralisations des bornes de performance fréquentistes ont été proposées et mises en oeuvre pour des estimateurs supposés non biaisés. À ce titre, une borne de Cramér-Rao paramétrique sur variété riemannienne est développée dans [97] qui admet une expression analytique dans le cadre de variétés d'intérêt comme l'ensemble des matrices de covariance ou les variétés de Grassmann [98]. Une modification de cette borne est par ailleurs introduite en particulier dans le cadre des distributions elliptiques. Ces distributions sont décrites par des matrices de dispersion inconnues et contraintes à évoluer sur l'espace des matrices symétrique et semi-définies positives [99]. À cet effet, des métriques propres aux variétés des covariances peuvent être utilisées comme la métrique log-euclidienne ou la métrique de Rao-Fisher, cette dernière définissant une distance géodésique [100].

Dans le cadre spécifique des groupes de Lie, des bornes fréquentistes ont également été proposées dans la littérature. Comme il n'est pas trivial de définir une distance sur un groupe de Lie, une métrique d'erreur intrinsèque est définie directement sur l'algèbre de Lie [101]. Dans un cadre générique, une borne de Cramér-Rao paramétrique est mise en place dans [101] qui fournit une expression analytique pouvant s'appliquer sur tout groupe de Lie matriciel. En particulier, cette borne est montrée être un cas particulier de la borne de [97] lorsque la métrique considérée sur le groupe est bi-invariante, ce qui est le cas du groupe $SO(3)$. Dans [102], une borne de Cramér-Rao est déterminée plus spécifiquement pour des observations issues de distributions symétriques et concentrées obtenues sur groupe de Lie.

Dans un cadre bayésien et particulièrement pour le cas du groupe $SO(n)$, une métrique intrinsèque de Hilbert-Schmidt permettant de quantifier les erreurs est mise en place et aboutit à la définition d'une *borne de Hilbert-Schmidt* bayésienne sur $SO(n)$ [103].

À notre connaissance, il n'existe pas dans la littérature de borne de type Cramér-Rao générique sur groupe de Lie se plaçant dans un contexte bayésien.

4.2 Proposition d'une borne de Cramér-Rao sur groupe de Lie

4.2.1 Motivations

Les bornes de Cramér-Rao fréquentistes proposées par [101] et [102] ne sont pas adaptées si :

- les observations vivent sur groupe de Lie. En effet, dans le cas de [101], les observations sont supposées prendre leurs valeurs sur un espace euclidien, ce qui n'est pas forcément le cas dans diverses applications en particulier lorsque des mesures de rotations et de translations sont disponibles et vivent sur $SE(3)$.
- Le contexte est bayésien et le paramètre inconnu est aléatoire.
- Le groupe de Lie considéré est non compact. En effet, les bornes proposées sont établies dans un cadre compact et ne sont donc pas applicables à des groupes comme $SE(3)$.

Pour prendre en compte ces trois aspects, nous proposons de mettre en place une borne de Cramér-Rao :

- dans un contexte bayésien,
- applicable à tout groupe de Lie et considérant des observations sur groupe de Lie.

La borne que nous proposons est obtenue en s'inspirant de la démonstration de la borne bayésienne euclidienne établie dans [104].

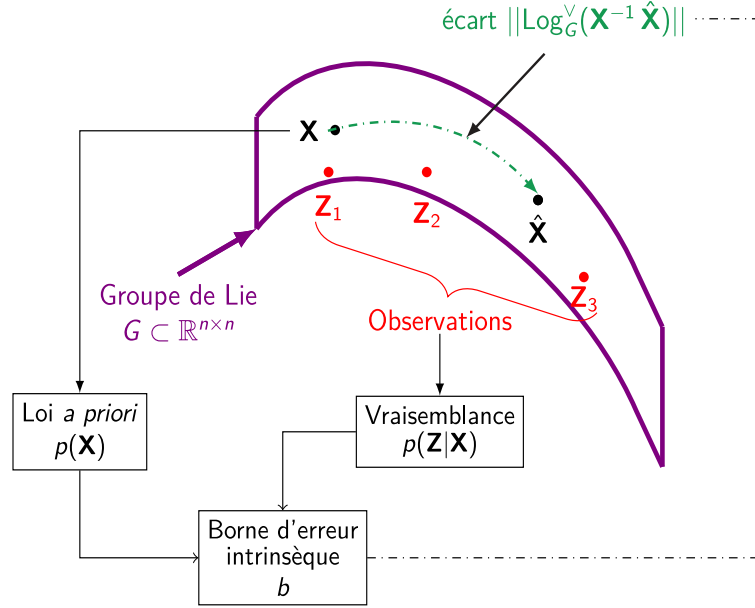


FIGURE 4.1 – Construction d'une borne bayésienne sur groupe de Lie dans le cas où les observations et le paramètre inconnu vivent sur le même groupe de Lie.

Une des difficultés est le calcul d'intégrales de densités de probabilité que nous pouvons définir par rapport à une mesure de groupe, comme la mesure de Haar.

Pour la pallier, nous proposons de considérer le développement de la borne dans un cadre matriciel et unimodulaire. Ce choix se justifie par le fait que la plupart des groupes de Lie physiques respectent cette propriété (contrairement à celle de compacité) et que cette dernière permet d'effectuer des calculs d'intégration plus avancés. En particulier, il est possible de faire apparaître des quantités intuitives en utilisant les propriétés d'invariance de mesure et une intégration par parties sur groupe de Lie, comme décrit dans la section (2.2.3.2).

4.2.2 Expression générale

Afin de développer notre borne, nous considérons une matrice aléatoire inconnue $\mathbf{X}$ appartenant à un groupe de Lie G unimodulaire. Elle est distribuée *a priori* selon la densité dénommée $p(\mathbf{X})$. De plus, nous supposons l'existence d'un ensemble d'observations indépendantes $\mathbf{Z} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_n\}$, où chaque $\mathbf{Z}_i$ appartient à un groupe de Lie unimodulaire noté G' , distribué selon la vraisemblance $p(\mathbf{Z}|\mathbf{X})$. Par ailleurs, nous considérons $\hat{\mathbf{X}}$ un estimateur *a posteriori* de $\mathbf{X}$.

Théorème 1. Notons E_{int} l'erreur intrinsèque moyenne entre $\mathbf{X}$ et $\hat{\mathbf{X}}$ dont nous rappelons qu'elle s'écrit :

$$E_{\text{int}} = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} \|\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}})\|^2 p(\mathbf{Z}, \mathbf{X}) \mu_{G'^n \times G}(d\mathbf{Z}, d\mathbf{X}) \quad (4.9)$$

Elle vérifie l'inégalité suivante :

$$E_{\text{int}} \geq \text{tr} \left(\mathbb{E}_{p(\mathbf{Z}, \mathbf{X})} \left(\psi_G \left(\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right) \right) \mathbf{P} \mathbb{E}_{p(\mathbf{Z}, \mathbf{X})} \left(\psi_G \left(\text{Log}_G^V(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right)^\top \right) \right) \quad (4.10)$$

où :

- $\psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right)$ correspond à la matrice jacobienne à gauche inverse de G calculée en $\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})$ dont l'expression a été donnée par l'équation (2.36).
- $\mathbf{P}$ correspond à l'inverse de l'espérance de la hessienne sur groupe de Lie de $-\log p(\mathbf{Z}, \mathbf{X})$ définie par l'équation (2.51). Elle peut être vue comme une matrice d'information de Fisher sur groupe de Lie et s'écrit :

$$\mathbf{P} = -\mathbb{E}_{p(\mathbf{Z}, \mathbf{X})} \left(\frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \Big|_{\epsilon_1=0, \epsilon_2=0} \right)^{-1}. \quad (4.11)$$

4.2.2.1 Démonstration

Pour démontrer cette inégalité, nous nous inspirons de la démonstration de la borne de Cramér-Rao *a posteriori* classique. Nous définissons dans un premier temps le vecteur aléatoire suivant :

$$\epsilon = \begin{bmatrix} \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \\ \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} \end{bmatrix} \quad (4.12)$$

Nous calculons ensuite sa matrice de corrélation afin de construire une matrice dont l'un des blocs fait apparaître la matrice de l'erreur d'estimation moyenne intrinsèque. En effet, si nous considérons la matrice :

$$\mathbf{R} = \mathbb{E} \left(\epsilon \epsilon^\top \right), \quad (4.13)$$

nous voyons qu'elle peut s'écrire par bloc sous la forme :

$$\mathbf{R} = \begin{bmatrix} \mathbb{E}[\delta_{\hat{\mathbf{X}}} \delta_{\hat{\mathbf{X}}}^\top] & \mathbb{E}(\delta_{\hat{\mathbf{X}}} \nabla_{\mathbf{X}} Lp^\top) \\ \mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top) & \mathbb{E}(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top) \end{bmatrix} \quad (4.14)$$

où :

$$\begin{cases} \delta_{\hat{\mathbf{X}}} = \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \end{cases} \quad (4.15)$$

$$\begin{cases} \nabla_{\mathbf{X}} Lp = \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} \end{cases} \quad (4.16)$$

Comme $\mathbf{R}$ est semi-définie positive, son complément de Schur l'est également. Par conséquent :

$$\text{tr} \left(\mathbb{E} \left(\delta_{\hat{\mathbf{X}}} \delta_{\hat{\mathbf{X}}}^\top \right) \right) \geq \text{tr} \left(\mathbb{E} \left(\delta_{\hat{\mathbf{X}}} \nabla_{\mathbf{X}} Lp^\top \right) \mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right)^{-1} \mathbb{E} \left(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top \right) \right) \quad (4.17)$$

Cette formule nous permet d'obtenir une inégalité sur la matrice d'erreur $\mathbb{E} \left(\delta_{\hat{\mathbf{X}}} \delta_{\hat{\mathbf{X}}}^\top \right)$. Néanmoins, le terme de droite n'admet généralement pas d'expression analytique et (4.17) n'est donc pas directement applicable à un problème d'estimation.

Ainsi, il semble judicieux de l'explicitier de sorte à obtenir une expression analytique faisant intervenir des quantités que nous sommes capables d'évaluer. Nous constatons que cette borne fait apparaître deux calculs d'intégrale : $\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top \right)$ et $\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right)$. Par conséquent,

nous allons nous intéresser au calcul explicite de chacune d'entre elles, en utilisant les propriétés d'invariance d'intégrale sur groupe de Lie unimodulaire rappelées dans la sous-section (2.2.3.1) du chapitre 2. En particulier, ces propriétés vont nous permettre de mettre à profit la formule d'intégration par parties sur groupe de Lie.

1) Calcul de $\mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top)$:

Par construction et par propriétés d'unimodularité, nous savons que :

$$\mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top) = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})^\top p(\mathbf{Z}, \mathbf{X}) \Delta_G(\mathbf{X}) \mu_{G'^n}(d\mathbf{Z}) \mu_{G'}(d\mathbf{X}) \quad (4.18)$$

Comme $\mathbf{X}$ est unimodulaire, $\Delta_G(\mathbf{X}) = 1$ et d'après l'équation (2.64) :

$$\mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top) = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})^\top p(\mathbf{Z}, \mathbf{X}) \mu_{G'^n}(d\mathbf{Z}) \mu_G(d\mathbf{X}) \quad (4.19)$$

Une manière de s'affranchir de la log-vraisemblance est de remarquer que :

$$\frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} = \frac{\partial p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))^{-1} \Big|_{\epsilon_{\mathbf{X}}=0} \quad (4.20)$$

$$= \frac{\partial p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} p(\mathbf{Z}, \mathbf{X})^{-1} \quad (4.21)$$

En substituant (4.21) dans (4.19), nous obtenons :

$$\mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top) = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} \frac{\partial p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}))}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0} \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})^\top \mu_G(d\mathbf{X}) \mu_{G'^n}(d\mathbf{Z}) \quad (4.22)$$

En supposant que p est une densité de probabilité bornée, nous pouvons effectuer une intégration par parties sur groupe de Lie, selon le procédé de la démonstration de l'équation (12.14) de [47] et de (2.68). Nous obtenons alors :

$$\mathbb{E}(\nabla_{\mathbf{X}} Lp \delta_{\hat{\mathbf{X}}}^\top) = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} p(\mathbf{Z}, \mathbf{X}) \frac{\partial \text{Log}_G^\vee(\text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}) \mathbf{X}^{-1} \hat{\mathbf{X}})}{\partial \epsilon_{\mathbf{X}}} \Big|_{\epsilon_{\mathbf{X}}=0}^\top \mu_G(d\mathbf{X}) \mu_{G'^n}(d\mathbf{Z}) \quad (4.23)$$

À ce stade du calcul, nous ne pouvons pas faire plus de simplifications dans le cas général. Néanmoins, nous nous plaçons sous l'hypothèse où la distribution *a posteriori* $p(\mathbf{X}|\mathbf{Z})$ est relativement peu étalée, ce qui est vérifié notamment si le nombre n d'observations est grand. Dans ce cas, l'estimateur $\hat{\mathbf{X}}$, par définition obtenu à partir de $p(\mathbf{X}|\mathbf{Z})$, se trouve à une "distance" raisonnable des réalisations correspondant à une valeur non négligeable de $p(\mathbf{X}|\mathbf{Z})$.

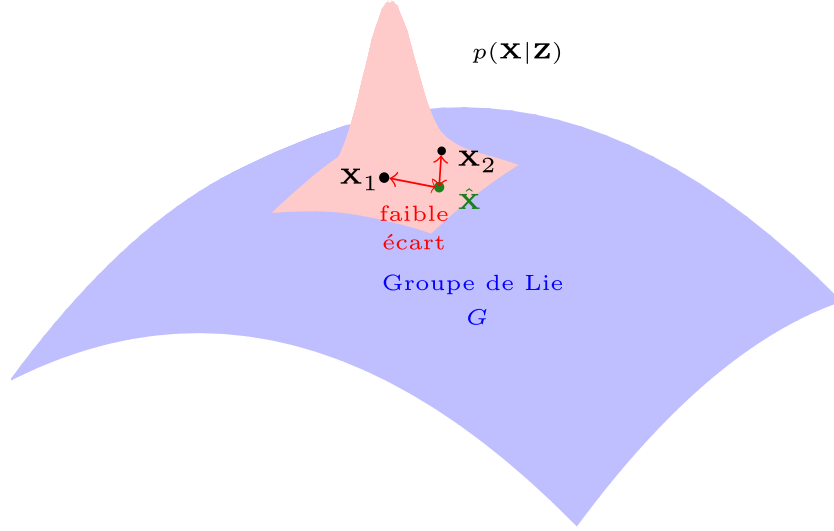


FIGURE 4.2 – Représentation du support de la distribution *a posteriori* $p(\mathbf{X}|\mathbf{Z})$ sur G : l'estimateur $\hat{\mathbf{X}}$ est représenté en vert. $\mathbf{X}_1$ et $\mathbf{X}_2$ correspondant à des réalisations selon $p(\mathbf{X}|\mathbf{Z})$.

Par conséquent, $\forall \mathbf{X}$ de densité de probabilité $p(\mathbf{X}|\mathbf{Z})$ non négligeable, $\mathbf{X}^{-1} \hat{\mathbf{X}}$ peut s'écrire :

$$\mathbf{X}^{-1} \hat{\mathbf{X}} = \text{Exp}_G^\wedge(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \quad (4.24)$$

Ainsi, en utilisant la seconde formule BCH (2.35), nous obtenons :

$$\text{Log}_G^\vee(\text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}) \mathbf{X}^{-1} \hat{\mathbf{X}}) = \text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) + \psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right) \epsilon_{\mathbf{X}} + O(\|\epsilon_{\mathbf{X}}\|^2)$$

Soit en dérivant par rapport à ϵ :

$$\left. \frac{\partial \text{Log}_G^\vee(\text{Exp}_G^\wedge(\epsilon_{\mathbf{X}}) \mathbf{X}^{-1} \hat{\mathbf{X}})}{\partial \epsilon_{\mathbf{X}}} \right|_{\epsilon_{\mathbf{X}}=0} = \psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right) \quad (4.25)$$

En remplaçant cette équation dans (4.23), il vient que :

$$\begin{aligned} \mathbb{E} \left(\nabla_{\mathbf{X}} L p \delta_{\hat{\mathbf{X}}}^\top \right) &= \int_{\mathbf{X} \in G} \int_{\mathbf{Z} \in G^n} p(\mathbf{Z}, \mathbf{X}) \psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right)^\top \mu_{G^n}(d\mathbf{Z}) \mu_G(d\mathbf{X}) \\ &= \boxed{\mathbb{E} \left(\psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right)^\top \right)} \end{aligned} \quad (4.26)$$

Par transposition, nous obtenons naturellement que :

$$\mathbb{E} \left(\delta_{\hat{\mathbf{X}}} \nabla_{\mathbf{X}} L p^\top \right) = \boxed{\mathbb{E} \left(\psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right) \right)} \quad (4.27)$$

2) Calcul de $\mathbb{E}(\nabla_{\mathbf{X}} L p \nabla_{\mathbf{X}} L p^\top)$

Cette quantité s'écrit sous la forme d'une intégrale :

$$\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right) = \int_{\mathbf{Z} \in G'^n} \int_{\mathbf{X} \in G} \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_1))}{\partial \epsilon_1} \bigg|_{\epsilon_1=0} \frac{\partial \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_2))}{\partial \epsilon_2} \bigg|_{\epsilon_2=0}^\top p(\mathbf{Z}, \mathbf{X}) \mu_G(d\mathbf{X}) \mu_{G'^n}(d\mathbf{Z}) \quad (4.28)$$

Nous pouvons développer cette expression en appliquant (4.21) par exemple au terme dépendant de ϵ_2 puis en permutant les intégrales et les dérivées selon l'équation (12.14) de [47] :

$$\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \left[\int_{\mathbf{Z} \in G'^n} \int_{\tilde{\mathbf{X}} \in G} \log p(\mathbf{Z}, \tilde{\mathbf{X}} \text{Exp}_G^\wedge(\epsilon_1)) p(\mathbf{Z}, \tilde{\mathbf{X}} \text{Exp}_G^\wedge(\epsilon_2)) \mu_G(d\tilde{\mathbf{X}}) \mu_{G'^n}(d\mathbf{Z}) \right] \bigg|_{\epsilon_1=0, \epsilon_2=0} \quad (4.29)$$

Il est à noter que cette inversion est possible si p est supposée bornée et tend vers zéro lorsque $\mathbf{X}$ prend de grandes valeurs.

Par le changement de variable $\tilde{\mathbf{X}} = \mathbf{X} \text{Exp}_G^\wedge(\epsilon_2)$ et par invariance de $\mu_G(\cdot)$, nous obtenons :

$$\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \left[\int_{\mathbf{Z} \in G'^n} \int_{\tilde{\mathbf{X}} \in G} \log p(\mathbf{Z}, \tilde{\mathbf{X}} \text{Exp}_G^\wedge(-\epsilon_2) \text{Exp}_G^\wedge(\epsilon_1)) p(\mathbf{Z}, \tilde{\mathbf{X}}) \mu_G(d\tilde{\mathbf{X}}) \mu_{G'^n}(d\mathbf{Z}) \right] \bigg|_{\epsilon_1=0, \epsilon_2=0} \quad (4.30)$$

Nous avons alors :

$$\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right) = - \int_{\mathbf{Z} \in G'^n} \int_{\tilde{\mathbf{X}} \in G} \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z}, \tilde{\mathbf{X}} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \bigg|_{\epsilon_1=0, \epsilon_2=0} p(\mathbf{Z}, \tilde{\mathbf{X}}) \mu_G(d\tilde{\mathbf{X}}) \mu_{G'^n}(d\mathbf{Z}) \quad (4.31)$$

Soit :

$$\mathbb{E} \left(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top \right) = - \mathbb{E} \left(\frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \bigg|_{\epsilon_1=0, \epsilon_2=0} \right) \quad (4.32)$$

Nous aboutissons donc à une expression de l'intégrale, dépendant uniquement de la hessienne de la log-vraisemblance par rapport à $\mathbf{X}$, soit :

$$\mathbb{E}(\nabla_{\mathbf{X}} Lp \nabla_{\mathbf{X}} Lp^\top) = -\mathbb{E}(\mathbf{H}(\mathbf{X})) \quad (4.33)$$

avec :

$$\mathbf{H}(\mathbf{X}) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z}, \mathbf{X} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \bigg|_{\epsilon_1=0, \epsilon_2=0} \quad (4.34)$$

En remplaçant les équations (4.26), (4.27) et (4.33) dans (4.17), nous obtenons alors l'équation (4.10).

4.2.3 Hypothèses sur $\hat{\mathbf{X}}$ et expression approchée de la borne

À travers l'équation (4.10), nous obtenons une inégalité sur la RMSE intrinsèque d'estimation. Cependant, elle n'est pas directement exploitable pour deux raisons :

- La matrice jacobienne inverse $\psi_G(\cdot)$ n'admet pas d'expression analytique dans un cadre général.
- L'erreur intrinsèque intervient dans les deux membres de l'inégalité.

Pour tenter de réduire dans un premier temps la difficulté du problème et obtenir une expression analytique, nous proposons d'effectuer une approximation raisonnable et réaliste de la matrice jacobienne inverse.

En utilisant de nouveau le fait que $p(\mathbf{X}|\mathbf{Z})$ est peu étalée, ψ_G peut être approchée par un développement limité à l'ordre 1. Ainsi, nous avons :

$$\psi_G \left(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}) \right) \simeq \mathbf{I} - \frac{1}{2} \text{ad}_G(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \quad (4.35)$$

En substituant cette formule dans l'expression de la borne, nous obtenons que le terme à droite de l'équation (4.10), noté e_{LG} s'écrit :

$$e_{LG} \simeq \underbrace{\text{tr} \left(\mathbf{P} - \frac{1}{2} \mathbf{A} + \frac{1}{4} \mathbf{B} \right)}_{\text{fonction dépendant de } E_{int}} \quad (4.36)$$

où :

$$\begin{cases} \mathbf{A} = \mathbf{P} \mathbb{E} \left(\text{ad}_G(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \right)^\top + \mathbb{E} \left(\text{ad}_G(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \right) \mathbf{P} \end{cases} \quad (4.37)$$

$$\begin{cases} \mathbf{B} = \mathbb{E} \left(\text{ad}_G(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \right) \mathbf{P} \mathbb{E} \left(\text{ad}_G(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \right)^\top \end{cases} \quad (4.38)$$

En raison de la semi-définie positivité de $\mathbf{P}$, le terme $\mathbf{B}$ l'est également et sa trace est donc positive. Par conséquent, E_{int} est bornée par la quantité e_{LG}^1 s'écrivant :

$$e_{LG}^1 = \text{tr} \left(\mathbf{P} - \frac{1}{2} \mathbf{A} \right). \quad (4.39)$$

Nous proposons d'utiliser l'inégalité obtenue afin d'en déduire une formule analytique dans le cas de groupes de Lie d'intérêt, en particulier les groupes $SO(3)$ et $SE(3)$, qui interviennent dans notre application.

4.2.4 Expression analytique

Résultat préliminaire :

Soit $\mathbf{A}$ une matrice symétrique de taille $n \times n$ et $\mathbf{B}$ une matrice antisymétrique, alors :

$$\text{tr}(\mathbf{A} \mathbf{B}) = \text{tr}(\mathbf{B} \mathbf{A}) = 0 \quad (4.40)$$

4.2.4.1 Cas de $SO(3)$

Dans le cas du groupe $SO(3)$, la matrice $\text{ad}_G(\cdot)$ est antisymétrique, par conséquent :

$$\text{ad}_G(\mathbf{a}) = -\text{ad}_G(\mathbf{a})^\top \quad \forall \mathbf{a} \in \mathbb{R}^3 \quad (4.41)$$

Ainsi, la trace de $\mathbf{A}$ est nulle puisque la trace d'une matrice antisymétrique multipliée par une matrice symétrique, qui est une propriété de $\mathbf{P}$, est de trace nulle d'après (4.40).

Par conséquent, l'inégalité de Cramér-Rao proposée sur $SO(3)$ s'écrit alors directement sous la forme :

$$E_{int} \geq \text{tr}(\mathbf{P}) = -\mathbb{E}(\text{tr}(\mathbf{H})) \quad (4.42)$$

Cette borne est obtenue en négligeant les termes de courbure dans (4.35). Cela revient à considérer que $SO(3)$ est localement commutatif au voisinage de l'identité et donc en $\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})$.

4.2.4.2 Cas de $SE(3)$

Pour déterminer une borne spécifique sur $SE(3)$, nous proposons premièrement de mettre à profit l'expression de son adjoint en $\text{Log}_{SE(3)}^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})$, qui admet la structure suivante :

$$ad_{SE(3)}(\mathbf{a}) = \left[\begin{array}{c|c} \mathbf{A}_1 & \mathbf{0}_{6 \times 6} \\ \hline \mathbf{A}_2 & \mathbf{A}_1 \end{array} \right]. \quad (4.43)$$

où $\mathbf{a} = \text{Log}_{SE(3)}^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})$ et

$$\begin{cases} \mathbf{A}_1 = [\mathbf{a}_R]_\times \\ \mathbf{A}_2 = [\mathbf{a}_p]_\times \end{cases} \quad (4.44)$$

$$(4.45)$$

avec $\mathbf{a}_R \in \mathbb{R}^3$ qui correspond au sous-vecteur des composantes rotationnelles de $\mathbf{a}$ et $\mathbf{a}_p$ à celui des composantes en translation.

Comme l'adjoint s'écrit sous forme d'une représentation par blocs, nous décomposons la matrice $\mathbf{P}$ sous la même forme :

$$\mathbf{P} = \left[\begin{array}{c|c} \mathbf{P}_1 & \mathbf{P}_2 \\ \hline \mathbf{P}_3 & \mathbf{P}_4 \end{array} \right]. \quad (4.46)$$

En raison de la symétrie de $\mathbf{P}$, $\mathbf{P}_1$ et $\mathbf{P}_4$ sont deux matrices également symétriques appartenant à $\mathbb{R}^{3 \times 3}$ et $\mathbf{P}_2$ et $\mathbf{P}_3$ sont deux matrices définies telles que $\mathbf{P}_2^\top = \mathbf{P}_3$.

• Cas d'une matrice $\mathbf{P}$ diagonale

Si la matrice $\mathbf{P}$ est diagonale, alors nous montrons aisément que :

$$\text{tr}(\mathbb{E}(\mathbf{A})) = 0 \quad (4.47)$$

En effet, la matrice $\mathbf{A}$ est de trace nulle car elle est calculée en effectuant le produit d'une matrice antisymétrique, de trace nulle, par une matrice diagonale.

Par conséquent, la borne d'erreur s'écrit plus simplement dans ce cas :

$$b_{SE(3)}^1 = \text{tr}(\mathbf{P}). \quad (4.48)$$

L'expression est alors la même que dans le cas de $SO(3)$.

• Cas d'une matrice $\mathbf{P}$ quelconque

Dans cette partie, nous réarrangeons l'inégalité obtenue pour E_{int} et mettons à profit les propriétés de $SE(3)$ pour en déduire une borne explicite.

Théorème 2. *Si $\mathbf{P}$ est une matrice non diagonale, une borne de Cramér-Rao intrinsèque sur $SE(3)$ s'écrit sous la forme :*

$$b_{SE(3)}^1 = \left(-\frac{\sqrt{2}\alpha}{2} + \sqrt{\frac{\alpha^2}{2} + \text{tr}(\mathbf{P})} \right)^2 \quad (4.49)$$

où $\alpha = \text{tr}(\mathbf{P}_3 \mathbf{P}_3^\top)$.

Démonstration :

Afin de démontrer ce théorème, nous transformons l'inégalité (4.10) s'écrivant sous la forme :

$$E_{int} \geq g(E_{int}) \quad (4.50)$$

où $g : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction continue, en une inégalité du type :

$$f(E_{int}) \geq 0 \quad (4.51)$$

où $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction continue dont nous cherchons à déterminer une expression explicite. L'étude des propriétés de f nous permet alors d'extraire la borne de performance recherchée.

Pour y parvenir, nous détaillons premièrement le calcul de $ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \mathbf{P}$ et $\mathbf{P} ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}))^\top$ afin d'obtenir une expression explicite de $g(E_{int})$.

En utilisant les équations (4.45) et (4.46) nous obtenons que :

$$\begin{cases} ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \mathbf{P} = \left[\begin{array}{c|c} \mathbf{A}_1 \mathbf{P}_1 & \mathbf{A}_1 \mathbf{P}_2 \\ \hline \mathbf{A}_2 \mathbf{P}_1 + \mathbf{A}_1 \mathbf{P}_3 & \mathbf{A}_2 \mathbf{P}_2 + \mathbf{A}_1 \mathbf{P}_4 \end{array} \right] \\ \mathbf{P} ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}))^\top = \left[\begin{array}{c|c} \mathbf{P}_1 \mathbf{A}_1^\top & \mathbf{P}_1 \mathbf{A}_2^\top + \mathbf{P}_2 \mathbf{A}_1^\top \\ \hline \mathbf{P}_3 \mathbf{A}_1^\top & \mathbf{P}_3 \mathbf{A}_2^\top + \mathbf{P}_4 \mathbf{A}_1^\top \end{array} \right] \end{cases} \quad (4.52)$$

$$(4.53)$$

La matrice $\mathbf{A}_s = ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}})) \mathbf{P} + \mathbf{P} ad_{SE(3)}(\text{Log}_G^\vee(\mathbf{X}^{-1} \hat{\mathbf{X}}))^\top$ s'écrit alors :

$$\mathbf{A}_s = \left[\begin{array}{c|c} \mathbf{A}_1 \mathbf{P}_1 + \mathbf{P}_1 \mathbf{A}_1^\top & \mathbf{P}_1 \mathbf{A}_2 + \mathbf{P}_2 \mathbf{A}_1^\top + \mathbf{A}_1 \mathbf{P}_2 \\ \hline \mathbf{A}_2 \mathbf{P}_1 + \mathbf{P}_3 \mathbf{A}_1^\top + \mathbf{A}_1 \mathbf{P}_3 & \mathbf{A}_2 \mathbf{P}_2 + \mathbf{A}_1 \mathbf{P}_4 + \mathbf{P}_3 \mathbf{A}_2^\top + \mathbf{P}_4 \mathbf{A}_1^\top \end{array} \right] \quad (4.54)$$

En récupérant la trace de cette matrice, nous obtenons que :

$$\text{tr}(\mathbf{A}_s) = \text{tr}(\mathbf{A}_1 \mathbf{P}_1 + \mathbf{P}_1 \mathbf{A}_1^\top + \mathbf{A}_2 \mathbf{P}_2 + \mathbf{A}_1 \mathbf{P}_4 + \mathbf{P}_3 \mathbf{A}_2^\top + \mathbf{P}_4 \mathbf{A}_1^\top) \quad (4.55)$$

Cette expression peut se simplifier, en effet :

→ $\mathbf{A}_1 = -\mathbf{A}_1^\top$, par conséquent $\text{tr}(\mathbf{A}_1 \mathbf{P}_1 + \mathbf{P}_1 \mathbf{A}_1^\top) = \text{tr}(\mathbf{A}_1 \mathbf{P}_4 + \mathbf{P}_4 \mathbf{A}_1^\top) = 0$, d'après (4.40) puisque $\mathbf{P}_1$ et $\mathbf{P}_4$ sont symétriques.

→ $\mathbf{P}_2 = \mathbf{P}_3^\top$, par conséquent $\text{tr}(\mathbf{A}_2 \mathbf{P}_2 + \mathbf{P}_3 \mathbf{A}_2^\top) = 2 \text{tr}(\mathbf{A}_2 \mathbf{P}_3^\top)$.

Nous en déduisons finalement que :

$$\text{tr}(\mathbb{E}(\mathbf{A}_s)) = 2 \text{tr}(\mathbb{E}(\mathbf{A}_2) \mathbf{P}_3^\top) \quad (4.56)$$

Comme $\mathbf{A} = \mathbb{E}(\mathbf{A}_s)$, nous obtenons que le membre de droite de l'inégalité (4.50) s'écrit :

$$g(E_{int}) = \text{tr}(\mathbf{P} - \mathbb{E}(\mathbf{A}_2) \mathbf{P}_3^\top) \quad (4.57)$$

Afin d'obtenir une borne explicite, nous cherchons à faire disparaître du terme de droite l'erreur intrinsèque E_{int} .

Pour y parvenir, nous mettons à profit l'inégalité de Cauchy-Schwartz sur l'espace des matrices carrées de taille 6×6 . Nous avons alors :

$$\left| \text{tr}(\mathbf{A}_2 \mathbf{P}_3^\top) \right| \leq \sqrt{\text{tr}(\mathbf{A}_2 \mathbf{A}_2^\top)} \sqrt{\text{tr}(\mathbf{P}_3 \mathbf{P}_3^\top)} \quad (4.58)$$

Par monotonie de l'opérateur $\mathbb{E}$, nous obtenons que :

$$-\mathbb{E}\left(\sqrt{\text{tr}(\mathbf{A}_2 \mathbf{A}_2^\top)}\right) \sqrt{\text{tr}(\mathbf{P}_3 \mathbf{P}_3^\top)} \leq -\text{tr}(\mathbb{E}(\mathbf{A}_2) \mathbf{P}_3^\top) \quad (4.59)$$

Ainsi, le terme à gauche fait apparaître l'erreur moyenne sur les composantes en translation de $\mathbf{a}$. En effet, il est possible de voir que $\text{tr}(\mathbf{A}_2 \mathbf{A}_2^\top) = 2 \|\mathbf{a}_p\|^2$. Par conséquent :

$$-\sqrt{2} \mathbb{E}(\|\mathbf{a}_p\|) \sqrt{\text{tr}(\mathbf{P}_3 \mathbf{P}_3^\top)} \leq -\text{tr}(\mathbb{E}(\mathbf{A}_2) \mathbf{P}_3^\top) \quad (4.60)$$

Pour faire apparaître la quantité $E_{int} = \mathbb{E}(\|\mathbf{a}\|^2)$, nous nous servons de deux choses :

1. La quantité $-\mathbb{E}(\|\mathbf{a}_p\|)$ peut être minorée en utilisant l'inégalité de Cauchy-Schwartz sur l'espace des variables aléatoires réelles :

$$\mathbb{E}(\|\mathbf{a}_p\|) \leq \sqrt{\mathbb{E}(\|\mathbf{a}_p\|^2)} \quad (4.61)$$

2. La quantité $-\mathbb{E}(\|\mathbf{a}_p\|^2)$ peut être également minorée par puisque $\mathbb{E}(\|\mathbf{a}\|^2) \geq \mathbb{E}(\|\mathbf{a}_p\|^2)$. Nous obtenons finalement l'inégalité suivante :

$$-\sqrt{2 \mathbb{E}(\|\mathbf{a}\|^2)} \alpha \leq -\sqrt{2 \mathbb{E}(\|\mathbf{a}_p\|^2)} \alpha \leq -\sqrt{2} \alpha \mathbb{E}(\|\mathbf{a}_p\|) \leq -\text{tr}(\mathbb{E}(\mathbf{A}_2) \mathbf{P}_3^\top) \alpha \quad (4.62)$$

où :

$$\alpha = \sqrt{\text{tr}(\mathbf{P}_3 \mathbf{P}_3^\top)}. \quad (4.63)$$

Ceci nous permet, d'après l'équation (4.57), d'en déduire une inéquation sur l'erreur sous la forme :

$$\mathbb{E}(\|\mathbf{a}\|^2) \geq \text{tr}(\mathbf{P}) - \sqrt{2} \alpha \sqrt{\mathbb{E}(\|\mathbf{a}\|^2)} \quad (4.64)$$

Par conséquent, nous constatons que $\sqrt{E_{int}}$ vérifie l'inégalité $h(\sqrt{E_{int}}) > 0$ avec h une fonction quadratique et convexe s'écrivant :

$$h(x) = x^2 + \sqrt{2}\alpha x - \text{tr}(\mathbf{P}) \quad (4.65)$$

Par définition, son discriminant s'écrit alors :

$$\Delta = 2\alpha^2 + 4\text{tr}(\mathbf{P}) \quad (4.66)$$

Dès lors que $\mathbf{P}$ est une matrice non nulle, ce discriminant est strictement positif et h admet donc deux racines :

$$\begin{cases} x_1 = -\frac{\sqrt{2}\alpha}{2} - \sqrt{\frac{\alpha^2}{2} + \text{tr}(\mathbf{P})} \\ x_2 = -\frac{\sqrt{2}\alpha}{2} + \sqrt{\frac{\alpha^2}{2} + \text{tr}(\mathbf{P})} \end{cases} \quad (4.67)$$

$$\quad (4.68)$$

Par définition, x_1 est négative, puisqu'elle est la somme de deux termes négatifs. x_2 est positive puisque $\frac{\alpha^2}{2} + \text{tr}(\mathbf{P}) \geq \frac{\alpha^2}{2}$.

Il s'ensuit ainsi que $h(x) > 0 \forall x \in]-\infty, x_1] \cup [x_2, +\infty[$ car h est une fonction convexe. Comme $\sqrt{E_{int}}$ est positive, elle appartient à l'intervalle $[x_2, +\infty[$. Notre borne sur E_{int} est donc directement x_2^2 et nous en déduisons :

$$E_{int} \geq \left(-\frac{\sqrt{2}\alpha}{2} + \sqrt{\frac{\alpha^2}{2} + \text{tr}(\mathbf{P})} \right)^2 \quad (4.69)$$

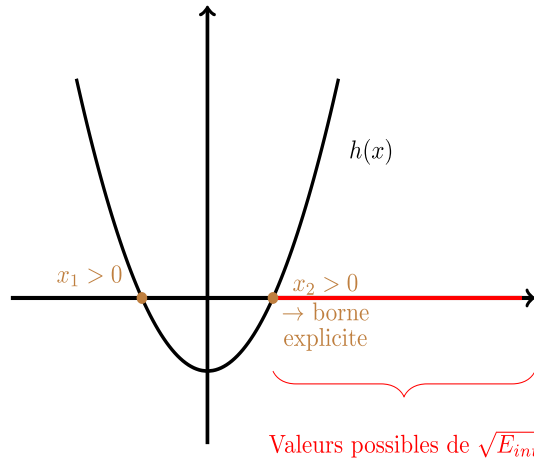


FIGURE 4.3 – Polynôme fonction de $\sqrt{E_{int}}$: la borne sur $SE(3)$ est fournie par x_2^2 .

4.3 Mise en oeuvre sur le modèle d'observation proposé

Dans cette partie, nous proposons de mettre en oeuvre la borne de Cramér-Rao pour un problème d'inférence sur $SE(3)$ dans lequel la vraisemblance et la loi *a priori* sont des distributions gaussiennes concentrées.

Plus précisément, nous nous plaçons dans le cadre du modèle développé dans le chapitre 3 mais en supposant que les mesures sur $SE(3)$ sont directement disponibles. Un ensemble d'observations mutuellement indépendantes $\mathbf{Z} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_n\}$ dispersé selon une forme courbée est disponible et appartient au groupe de Lie $SE(3)^n$. Nous rappelons qu'elles se représentent selon le modèle d'observation suivant :

$$\mathbf{Z}_i = \mathbf{M} \text{Exp}_{SE(3)}^\wedge(\mathbf{e}_i) \quad \mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{S}) \quad \forall i \in \{1, \dots, n\} \quad (4.70)$$

où $\mathbf{M}$ est supposé inconnu et aléatoire, défini sur $SE(3)$. Il représente le centroïde géométrique de la forme. Nous rappelons que le modèle *a priori* de $\mathbf{M}$ peut s'écrire sous la forme :

$$\mathbf{M} = \mathbf{M}_o \text{Exp}_{SE(3)}^\wedge(\mathbf{e}_M) \quad \mathbf{e}_M \sim \mathcal{N}(\mathbf{0}, \Sigma_M) \quad (4.71)$$

Nous cherchons donc à en déduire l'erreur d'estimation minimale atteignable sur le paramètre $\mathbf{M}$.

Afin d'obtenir une expression analytique de la borne, l'unique quantité à calculer est la hessienne sur groupe de Lie $\mathbf{H}(\mathbf{M})$ permettant d'en déduire la matrice d'information de Fisher sur groupe de Lie définie par l'équation (4.34). En mettant à profit la règle de Bayes, cette matrice peut se décomposer en deux termes :

$$\mathbf{H}(\mathbf{M}) = \underbrace{\mathbf{H}_{\mathbf{Z}}(\mathbf{M})}_{\text{Terme de vraisemblance}} + \underbrace{\mathbf{H}_{\mathbf{M}}(\mathbf{M})}_{\text{Terme } a \text{ priori}} \quad (4.72)$$

où :

$$\left\{ \begin{array}{l} \mathbf{H}_{\mathbf{Z}}(\mathbf{M}) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z} | \mathbf{M} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \Big|_{\epsilon_1=0, \epsilon_2=0} \\ \mathbf{H}_{\mathbf{M}}(\mathbf{M}) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{M} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \Big|_{\epsilon_1=0, \epsilon_2=0} \end{array} \right. \quad (4.73)$$

$$\left\{ \begin{array}{l} \mathbf{H}_{\mathbf{Z}}(\mathbf{M}) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{Z} | \mathbf{M} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \Big|_{\epsilon_1=0, \epsilon_2=0} \\ \mathbf{H}_{\mathbf{M}}(\mathbf{M}) = \frac{\partial}{\partial \epsilon_1} \frac{\partial}{\partial \epsilon_2} \log p(\mathbf{M} \text{Exp}_G^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) \Big|_{\epsilon_1=0, \epsilon_2=0} \end{array} \right. \quad (4.74)$$

La matrice d'information de Fisher s'écrit à l'aide de ces deux variables sous la forme :

$$\mathbf{P} = -(\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M})) + \mathbb{E}(\mathbf{H}_{\mathbf{M}}(\mathbf{M})))^{-1} \quad (4.75)$$

et la borne donnée par l'équation (4.49).

4.3.1 Expression de $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$:

Comme chaque $\mathbf{Z}_i$ est distribuée selon une gaussienne concentrée de paramètres $\mathbf{M}$ et $\mathbf{S}$, de par leur indépendance, nous pouvons écrire :

$$\log p(\mathbf{Z} | \mathbf{M}) = C_1 - \frac{1}{2} \sum_{i=1}^n \|\text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)\|_{\mathbf{S}}^2. \quad (4.76)$$

Nous avons alors, $\forall (\epsilon_1, \epsilon_2) \in \mathbb{R}^6$:

$$\log p(\mathbf{Z} | \mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2)) = C_1 - \frac{1}{2} \sum_{i=1}^n \left\| \text{Log}_{SE(3)}^\vee((\mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_G^\wedge(\epsilon_2))^{-1} \mathbf{Z}_i) \right\|_{\mathbf{S}}^2 \quad (4.77)$$

Comme ϵ_1 et ϵ_2 sont deux incréments qui vont tendre vers zéro, nous pouvons écrire que $\text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2) = \text{Exp}_G^\wedge(\epsilon_1 + \epsilon_2)$. Par conséquent, l'équation précédente prend la forme suivante :

$$\log p(\mathbf{Z} | \mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2)) = C_1 - \frac{1}{2} \sum_{i=1}^n \left\| \text{Log}_{SE(3)}^\vee(\text{Exp}_{SE(3)}^\wedge(-(\epsilon_1 + \epsilon_2)) \mathbf{M}^{-1} \mathbf{Z}_i) \right\|_{\mathbf{S}}^2 \quad (4.78)$$

Nous avons alors, en accord avec la seconde formule BCH :

$$\text{Log}_{SE(3)}^\vee(\text{Exp}_{SE(3)}^\wedge(-(\epsilon_1 + \epsilon_2)) \mathbf{M}^{-1} \mathbf{Z}_i) = \text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i) - \psi_{SE(3)}(\delta_{MZi}) (\epsilon_1 + \epsilon_2) + O(\|\epsilon_1 + \epsilon_2\|^2) \quad (4.79)$$

avec : $\delta_{MZi} = \text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)$. Par conséquent, en différenciant deux fois (4.78) par rapport à ϵ_1 et ϵ_2 en accord avec l'équation précédente, nous obtenons :

$$\mathbf{H}_{\mathbf{Z}}(\mathbf{M}) = - \sum_{i=1}^n \psi_{SE(3)}(\delta_{MZi})^\top \mathbf{S}^{-1} \psi_{SE(3)}(\delta_{MZi}) \quad (4.80)$$

D'où :

$$\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M})) = - \sum_{i=1}^n \mathbb{E} \left(\psi_{SE(3)}(\delta_{MZi})^\top \mathbf{S}^{-1} \psi_{SE(3)}(\delta_{MZi}) \right)$$

Posons $\mathbf{f}_{\psi_{SE(3)}}(.)$ la fonction $\psi_{SE(3)}(.)^\top \mathbf{S}^{-1} \psi_{SE(3)}(.)$. Par définition de $\mathbb{E}$, nous avons :

$$\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M})) = - \sum_{i=1}^n \int_{\mathbf{Z} \in G^n} \mathbf{f}_{\psi_{SE(3)}}(\text{Log}_G^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)) p(\mathbf{Z} | \mathbf{M}) \mu_{SE(3)}^n(d\mathbf{Z}) \quad (4.81)$$

En posant le changement de variable $\mathbf{a}_i = \text{Log}_{SE(3)}^\vee(\mathbf{M}^{-1} \mathbf{Z}_i)$, il en vient que :

$$\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M})) = - \sum_{i=1}^n \int_{\mathbf{a}_i \in \mathbb{R}^p} \mathbf{f}_{\psi_{SE(3)}}(\mathbf{a}_i) \mathcal{N}(\mathbf{0}_{p \times 1}, \mathbf{S}) \mu_{\mathbb{R}^p}(d\mathbf{a}_i) \quad (4.82)$$

$$= - \sum_{i=1}^n \mathbb{E}(\mathbf{f}_{\psi_{SE(3)}}(\mathbf{a}_i)) \quad (4.83)$$

Ainsi, $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$ correspond directement à l'espérance de $\mathbf{f}_{\psi_{SE(3)}}(\mathbf{a}_i)$ sous une distribution gaussienne euclidienne centrée.

Néanmoins, cette expression ne peut pas être développée, en raison de la non-linéarité de

$\psi_{SE(3)}(\cdot)$. Pour pallier ce problème, nous utilisons le fait que $\psi_{SE(3)}(\cdot)$ peut être raisonnablement approchée à l'ordre 1. Plus précisément, cette approximation s'écrit :

$$\hat{\psi}_{SE(3)}(\delta_{MZi}) = \mathbf{I}_{6 \times 6} - \frac{1}{2} ad_{SE(3)}(\delta_{MZi}) \quad (4.84)$$

Et $\mathbb{E}(\mathbf{H}_Z(\mathbf{M}))$ peut alors être approchée par :

$$\hat{\mathbb{E}}(\mathbf{H}_Z(\mathbf{M})) = -n \mathbb{E} \left(\hat{\psi}_G(\delta_{MZi})^\top \mathbf{S}^{-1} \hat{\psi}_G(\delta_{MZi}) \right) \quad (4.85)$$

Ainsi, cette approximation nous amène à calculer deux quantités : $\mathbb{E}(ad_{SE(3)}(\mathbf{a}_i))$ et $\mathbb{E}(ad_{SE(3)}(\mathbf{a}_i)^\top ad_{SE(3)}(\mathbf{a}_i))$. Pour y parvenir, nous décomposons $\mathbf{a}$ dans sa base canonique et naturelle :

$$\mathbf{a}_i = \sum_{j=1}^6 a_{i,j} \mathbf{e}_j \quad (4.86)$$

où $\{\mathbf{e}_i\}_{i=1}^6$ est la base canonique de $\mathbb{R}^6$. Par linéarité de l'opérateur $ad_{SE(3)}(\cdot)$, nous avons ainsi :

$$ad_{SE(3)}(\mathbf{a}_i) = \sum_{j=1}^6 a_{i,j} ad_{SE(3)}(\mathbf{e}_j) \quad (4.87)$$

Par définition de $\mathbf{a}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{S})$, $\mathbb{E}(\mathbf{a}_i) = \mathbf{0}$, d'où, par passage à l'espérance, nous obtenons :

$$\mathbb{E}(ad_{SE(3)}(\mathbf{a}_i)) = \mathbf{0}_{6 \times 6} \quad (4.88)$$

Pour obtenir une expression de $\mathbb{E}(ad_{SE(3)}(\mathbf{a}_i)^\top ad_{SE(3)}(\mathbf{a}_i))$, nous nous servons du fait que :

$$ad_{SE(3)}(\mathbf{a}_i)^\top ad_{SE(3)}(\mathbf{a}_i) = \sum_{j=1}^6 a_{i,j} ad_{SE(3)}(\mathbf{e}_j)^\top \sum_{k=1}^6 a_{i,k} ad_{SE(3)}(\mathbf{e}_k) \quad (4.89)$$

Par passage à l'espérance, les coefficients de $\mathbf{S}$, $\mathbb{E}(a_{i,j} a_{i,k}) = (\mathbf{S})_{j,k} \forall (j, k) \in \{1, \dots, 6\}^2$ vont ainsi intervenir, soit :

$$\mathbb{E}(ad_{SE(3)}(\mathbf{a}_i)^2) = \sum_{j=1}^6 \sum_{k=1}^6 (\mathbf{S})_{j,k} ad_{SE(3)}(\mathbf{e}_k)^\top ad_{SE(3)}(\mathbf{e}_k) \quad (4.90)$$

Nous aboutissons alors à :

$$\hat{\mathbb{E}}(\mathbf{H}_Z(\mathbf{M})) = - \left(n \mathbf{S}^{-1} + \frac{n}{4} \sum_{j=1}^6 \sum_{k=1}^6 (\mathbf{A}_Z)_{j,k} \right) \quad (4.91)$$

en rassemblant les équations (4.88) et (4.90), en développant l'équation, puis en conservant uniquement les termes d'ordre 2. $\mathbf{A}_Z$ est un tenseur de dimension $6 \times 6 \times 6 \times 6$ tel que, $\forall (j, k) \in \{1, \dots, 6\}^2$:

$$(\mathbf{A}_Z)_{j,k} = (\mathbf{S})_{j,k} ad_{SE(3)}(\mathbf{e}_j)^\top \mathbf{S}^{-1} ad_{SE(3)}(\mathbf{e}_k) \quad (4.92)$$

4.3.2 Expression de $\mathbb{E}(\mathbf{H}_M(\mathbf{M}))$

De manière analogue, $\forall(\epsilon_1, \epsilon_2) \in (\mathbb{R}^6)^2$, nous pouvons écrire le logarithme de la loi *a priori* en $\mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2)$:

$$\log p(\mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2)) = C_1 - \frac{1}{2} \left\| \text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2)) \right\|^2 \quad (4.93)$$

Comme précédemment, le terme à l'intérieur de la norme peut être développé, et cela en supposant que $\text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2) = \text{Exp}_{SE(3)}^\wedge(\epsilon_1 + \epsilon_2)$. Néanmoins, cela nécessite de faire apparaître l'opérateur $Ad_{SE(3)}(\cdot)$ de sorte à pouvoir appliquer la même formule BCH. Plus précisément, nous avons :

$$\begin{aligned} \text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M} \text{Exp}_{SE(3)}^\wedge(\epsilon_1) \text{Exp}_{SE(3)}^\wedge(\epsilon_2)) &= \\ \text{Log}_{SE(3)}^\vee(\text{Exp}_{SE(3)}^\wedge(Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M})(\epsilon_1 + \epsilon_2)) \mathbf{M}_o^{-1} \mathbf{M}) &= \\ = \boldsymbol{\delta}_o + \boldsymbol{\psi}_{SE(3)}(\boldsymbol{\delta}_o) Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M})(\epsilon_1 + \epsilon_2) + O(\|\epsilon_1 + \epsilon_2\|^2) \end{aligned} \quad (4.94)$$

où $\boldsymbol{\delta}_o = \text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M})$. Le terme précédent peut être intégré dans l'équation (4.93). En différenciant deux fois, nous aboutissons à :

$$\mathbf{H}_M(\mathbf{M}) = -Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M})^\top \boldsymbol{\psi}_{SE(3)}(\boldsymbol{\delta}_o)^\top \boldsymbol{\Sigma}_M^{-1} \boldsymbol{\psi}_{SE(3)}(\boldsymbol{\delta}_o) Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M}) \quad (4.95)$$

Nous avons alors, après quelques simplifications décrites en annexe D et en conservant uniquement les termes d'ordre 2 :

$$\mathbb{E}(\mathbf{H}_M(\mathbf{M})) = -\mathbb{E} \left(\boldsymbol{\Sigma}_M^{-1} - \frac{1}{2} ad_{SE(3)}(\boldsymbol{\delta}_o)^{2\top} \boldsymbol{\Sigma}_M^{-1} - \boldsymbol{\Sigma}_M^{-1} \frac{1}{2} ad_{SE(3)}(\boldsymbol{\delta}_o)^2 \right) \quad (4.96)$$

$$= - \left(\boldsymbol{\Sigma}_M^{-1} - \frac{1}{2} \sum_{j=1}^6 \sum_{k=1}^6 (\mathbf{A}_{M1} + \mathbf{A}_{M2})_{j,k} \right) \quad (4.97)$$

où $\mathbf{A}_{M1}$, $\mathbf{A}_{M2}$ sont des tenseurs de dimension $6 \times 6 \times 6 \times 6$ tels que :

$$(\mathbf{A}_{M1})_{j,k} = (\boldsymbol{\Sigma}_M)_{j,k} ad_{SE(3)}(\mathbf{e}_j)^\top ad_{SE(3)}(\mathbf{e}_k)^\top \boldsymbol{\Sigma}_M^{-1} \forall (j, k) \in \{1, \dots, 6\}^2 \quad (4.98)$$

$$(\mathbf{A}_{M2})_{j,k} = \boldsymbol{\Sigma}_M^{-1} (\boldsymbol{\Sigma}_M)_{j,k} ad_{SE(3)}(\mathbf{e}_j) ad_{SE(3)}(\mathbf{e}_k) \forall (j, k) \in \{1, \dots, 6\}^2 \quad (4.99)$$

Finalement, en rassemblant (4.91) et (4.97), nous obtenons une expression globale de $\mathbf{P}$ sous la forme :

$$\mathbf{P} = \left[n \mathbf{S}^{-1} + \frac{n}{4} \sum_{j=1}^6 \sum_{k=1}^6 (\mathbf{A}_Z)_{j,k} + \boldsymbol{\Sigma}_M^{-1} - \frac{1}{2} \sum_{j=1}^6 \sum_{k=1}^6 (\mathbf{A}_{M1} + \mathbf{A}_{M2})_{j,k} \right]^{-1} \quad (4.100)$$

et la borne peut alors être obtenue selon l'équation (4.49).

4.3.2.1 Remarques sur la borne

- Cette borne est obtenue de manière approchée mais permet de considérer des termes non négligeables inhérents à la structure géométrique de $SE(3)$.
- La matrice d'information de Fisher résultante est approchée, par conséquent, la propriété de semi-définie positivité peut ne pas être respectée. Néanmoins, en raison de l'ordre de grandeur des valeurs des tenseurs $\mathbf{A}_Z$, $\mathbf{A}_{M1}$ et $\mathbf{A}_{M2}$ plus faible que celui des valeurs de $\mathbf{S}$ et Σ_M , cette propriété sera valable numériquement.

4.3.3 Simulations

Afin d'illustrer numériquement la borne établie dans le cas de $SE(3)$, nous proposons de considérer une version simplifiée du problème d'estimation que nous avons établi dans le chapitre 3. Plus précisément, la matrice d'étendue $\mathbf{S}$ est supposée connue et les orientations de chaque réflecteur sont disponibles. Par conséquent, le problème d'estimation peut s'écrire sous la forme d'un problème d'optimisation quadratique et peut classiquement être résolu par un algorithme *LG-GN*.

Notre objectif est double : comparer d'une part l'erreur moyenne d'estimation de notre approche avec la borne théorique et d'autre part, étudier le comportement de cette borne selon les caractéristiques des données.

L'erreur moyenne d'estimation intrinsèque est approchée par une méthode de Monte-Carlo. Les estimés sont obtenus à partir de différentes réalisations de l'état et des observations simulées selon les équations (4.70) et (4.71). Afin d'approcher statistiquement le mieux possible l'espérance théorique, un grand nombre de réalisations est considéré. La borne proposée est, quant à elle, calculée numériquement selon la formule décrite par (4.100), qui repose sur une approximation à l'ordre 1 de la jacobienne inverse, conformément à (4.35).

Les valeurs des paramètres de simulation sont précisés dans le tableau (4.1), les variables $\mathbf{S}$ et Σ_M sont par ailleurs testées pour différentes valeurs en fonction des simulations considérées. Les réglages algorithmiques pour le test de la méthode de Gauss-Newton sont donnés dans le tableau (4.2).

Nombre d'observations	$n = 10$ à 100
Nombre de simulations Monte Carlo	$Nr = 500$
Moyenne <i>a priori</i> $\mathbf{M}_o$	$\text{Exp}_{SE(3)}^\wedge([0.1, 0.1, 0.1, 100, 100, 100])$

TABLEAU 4.1 – Paramètres de simulation.

Erreur d'initialisation sur $\mathbf{M}$	$e_M = \exp_{SE(3)}^\wedge([0.1; 0.1; 0.1; 10; 10; 10])$
Nombre d'itérations LG-GN	$n_b = 10$

TABLEAU 4.2 – Réglages algorithmiques.

• Influence du nombre d'observations

Dans cette série d'expériences, nous fixons $\mathbf{S}$ et Σ_M et faisons varier le nombre d'observations. $\mathbf{S}$ est simulée de sorte à obtenir des réflecteurs dispersés sur une étendue de l'ordre de 10^5 m.

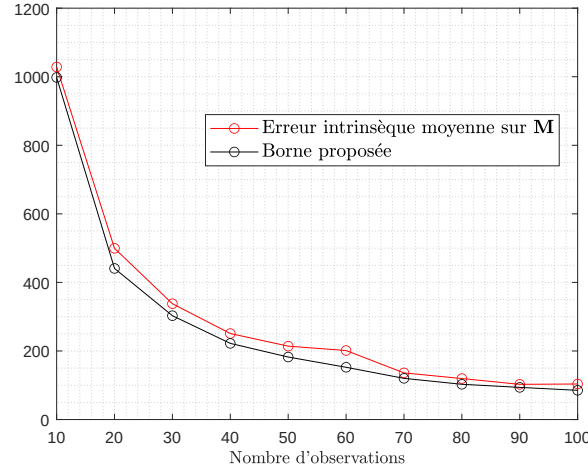


FIGURE 4.4 – Évolution de la borne proposée superposée à l'erreur moyenne quadratique pour $Nr = 500$ réalisations et $\Sigma_M = \text{diag}([0.1, 0.1, 0.1, 100, 100, 100])$.

Nous observons premièrement que la borne admet un comportement cohérent puisqu'elle minore l'erreur quelque soit le nombre d'observations considéré. De plus, cette borne admet un comportement asymptotique similaire à celui de l'erreur, en se stabilisant pour un grand nombre d'observations.

• Influence des paramètres *a priori*

Nous nous intéressons également au comportement de la borne lorsque l'*a priori* sur les composantes de translation de $\mathbf{M}_o$ évolue. Comme précédemment, $\mathbf{S}$ est fixée avec une étendue de l'ordre de 10^5 m.

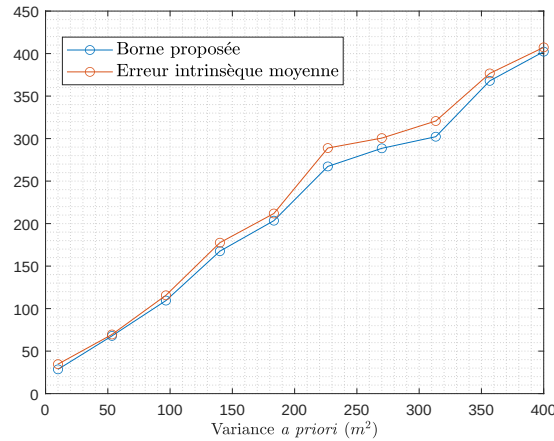


FIGURE 4.5 – Évolution de la borne proposée superposée à l'erreur moyenne quadratique en fonction de la variance *a priori* pour $n = 50$ et avec une covariance en rotation *a priori* égale à $\text{diag}([0.1, 0.1, 0.1])$.

Lorsque la dispersion *a priori* prend de grandes valeurs, l'erreur d'estimation a tendance à logiquement se dégrader. Nous observons que la borne conserve une cohérence en termes d'évolution. En effet, elle garde la même allure que l'erreur tout en la minorant.

• Influence de la distorsion de la forme

Afin d'étudier l'influence des paramètres sur notre borne, nous nous intéressons désormais à modifier la matrice de forme en faisant varier son paramètre de distorsion, décrit dans le chapitre 3. Ce paramètre agissant sur la courbure de la forme, il traduit les propriétés géométriques de $SE(3)$. Par conséquent, nous proposons d'étudier son influence en comparant notre borne $b_{SE(3)}^1$ à la borne obtenue dans le cas où les termes d'ordre supérieur à 1 sont négligés c'est à dire lorsque la jacobienne inverse est approchée par l'identité. La borne à l'ordre 0 s'écrit alors :

$$b_{SE(3)}^0 = \text{tr} (n \mathbf{S}^{-1} + \Sigma_M^{-1})^{-1} \quad (4.101)$$

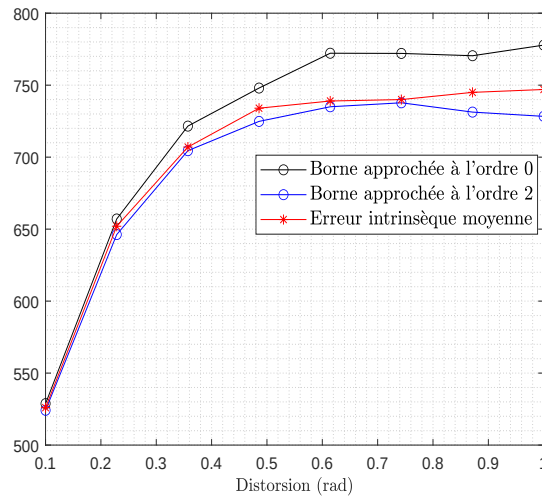


FIGURE 4.6 – Évolution de la borne proposée en fonction de la distorsion avec $n = 50$, $Nr = 500$ et $\Sigma_M = \text{diag}([0.1, 0.1, 0.1, 100, 100, 100])$ superposée à l'erreur moyenne quadratique.

- Premièrement, nous constatons que les deux bornes sont très proches entre elles lorsque le degré de distorsion est proche de zéro. Cette situation correspond à une forme avec une très faible courbure.
- Au contraire, lorsqu'il prend de grandes valeurs, l'écart entre les deux bornes augmente. Ainsi, lorsque la forme devient très distordue, l'approximation à l'ordre 0 n'est plus valide et la borne n'est plus conservative. Par conséquent, comme observée sur la figure 4.6, l'erreur moyenne devient plus petite qu'elle.

Nous pouvons en déduire que notre borne est bien plus adaptée pour des paramètres appartenant à $SE(3)$, en particulier lorsque les termes d'ordre supérieur à 0 dans les expressions développés des opérateurs de groupe de Lie sont non négligeables. Comme $b_{SE(3)}^0$ ne considère pas la structure géométrique de $SE(3)$, elle devient alors faussée et s'éloigne de $b_{SE(3)}^1$.

4.4 Conclusion du chapitre

Dans ce chapitre, nous avons proposé une nouvelle contribution théorique qui est une borne de Cramér-Rao bayésienne sur groupe de Lie.

Dans un premier temps, une inégalité sur l’erreur intrinsèque moyenne d’estimation a été établie dans un cadre générique de groupes de Lie matriciels unimodulaires. Ensuite, nous nous sommes intéressés en particulier aux groupes de Lie $SO(3)$ et $SE(3)$ afin d’en déduire une expression spécifique de la borne. Nous avons ensuite appliqué cette borne à notre modèle d’observation d’amas développé dans le chapitre 3. Cette mise en oeuvre numérique nous a ainsi permis de valider la cohérence et les propriétés de la borne proposée.

Une perspective directe de ce travail serait de considérer dans le problème d’estimation, premièrement, la présence des variables de nuisances et, deuxièmement, l’intégration des paramètres de forme de l’amas. Nous pourrions ensuite en déduire une borne d’erreur sur l’ensemble des paramètres modélisant l’amas. Cependant, prendre en compte ces variables supplémentaires suppose de considérer un modèle *a priori* hiérarchique qui complique l’expression de la borne.

Conclusions et perspectives

Conclusions

Cette thèse porte sur le pistage de débris spatiaux qui sont des objets artificiels et non fonctionnels en orbite autour de la Terre, et dont le nombre ne cesse d'augmenter. Dans ce cadre, nous nous sommes intéressés à mettre en place un modèle ainsi que des algorithmes pour estimer conjointement la position du centroïde et la forme d'un d'amas de débris spatiaux évoluant en orbite basse autour de la Terre.

Une première contribution a été de traiter l'amas comme une cible étendue. Si un système radar est utilisé pour la surveillance spatiale, il occupe en effet plusieurs cellules de résolution. Afin de déterminer ses caractéristiques, il a donc été nécessaire de définir des paramètres décrivant non seulement son centroïde mais également sa forme. Nous avons alors proposé de modéliser celle-ci à partir d'une matrice aléatoire en s'inspirant de l'approche développée par [31].

Afin de prendre en compte la géométrie de l'amas dont l'évolution est gouvernée par la force de gravitation selon le modèle classique képlerien, nous avons proposé de reformuler le problème d'estimation des caractéristiques de l'amas comme un problème d'inférence sur groupe de Lie. En effet, les débris tendent à se disperser sous forme de bananoïde, ainsi que nous avons pu l'observer en simulation, de manière similaire à des échantillons distribués selon des lois gaussiennes sur concentrées sur $SE(3)$. Les groupes de Lie sont des structures mathématiques élégantes sur lesquelles les calculs classiques d'intégration, de dérivation et statistiques peuvent être mis en place tout en considérant leurs propriétés géométriques.

Une seconde contribution de ce travail a été de définir un modèle d'état et un modèle d'observation donnant lieu à des mesures radar dont la répartition spatiale épouse la forme de l'amas. Pour y parvenir, le formalisme des distributions gaussiennes concentrées sur $SE(3)$ nous a permis de définir un modèle intrinsèque sur ce groupe de Lie et générant des mesures se dispersant selon la manière souhaitée.

Par la suite, un algorithme d'estimation, dans un contexte statique, a premièrement été mis en place. Il est fondé sur l'algorithme de Newton sur groupe de Lie dans lequel une contribution a été de modifier le calcul de descente afin d'obtenir une mise à jour efficace numériquement et adaptée à la structure du critère à optimiser. Plus précisément, notre approche assure que la matrice hessienne intervenant dans le calcul de l'incrément à chaque itération est semi-définie positive.

Ensuite, une version récursive a été développée, fondée sur un filtre de Kalman étendu itéré sur groupe de Lie. À l'instar de l'algorithme statique, une contribution a été de prendre en compte la présence de variables cachées, les propriétés du critère à optimiser, ainsi que la présence de plusieurs mesures, issues de différentes parties de la cible étendue, à chaque étape de l'algorithme. En particulier, une approximation de Laplace sur groupe de Lie a été mise en oeuvre afin de prendre en compte la structure non quadratique du critère d'optimisation à l'étape de correction.

Ces algorithmes ainsi que le modèle ont été validés numériquement à travers des simulations

selon des scénarios réalistes quant à l'évolution physique des débris. Les performances de l'algorithme de pistage ont en particulier été testées et comparées à une méthode de l'état de l'art dédiée aux cibles étendues basée sur les processus gaussiens, permettant de pister des objets de forme quelconque. Il s'est révélé plus efficace en terme de reconstruction de la géométrie de l'amas.

Une dernière contribution de ce manuscrit a été de proposer une borne de performance bayésienne sur groupe de Lie. Cette borne est à notre connaissance la première établie dans un cadre bayésien et considérant la présence d'observations sur groupe de Lie. Elle a d'abord été étudiée pour les groupes de Lie matriciels unimodulaires. Ensuite, elle a été explicitée dans le cas des groupes de Lie que nous manipulons dans notre application physique, plus précisément $SO(3)$ et $SE(3)$. Finalement, elle a été déterminée de manière analytique et mise en oeuvre dans le cadre de notre modèle d'observation d'un amas de débris.

Perspectives

Les perspectives de ce travail sont multiples que ce soit d'un point de vue théorique ou algorithmique.

Premièrement, l'algorithme de pistage proposé pourrait être robustifié en prenant en compte la présence de données aberrantes. Un critère de décision fondée sur un fenêtrage statistique intrinsèque sur groupe de Lie pourrait ainsi être mis en place. La difficulté serait de calculer des frontières de décision relatives à des probabilités de fausse alarme prédéfinies.

Par ailleurs, un amas occupe plusieurs cases de résolution non nécessairement contiguës du système radar. Pour aider à sa détection, il serait possible d'appliquer une approche de type "track before detect" [105]. Cette méthode permet classiquement d'améliorer les performances de détection radar à partir des résultats du pistage. En accord avec notre modèle d'observation, un modèle statistique sur groupe de Lie pourrait être ainsi proposé mais se poserait néanmoins la question de la complexité calculatoire.

De plus, lorsque l'amas s'avère trop étendu et que les différents débris sont suffisamment éloignés les uns des autres, un pistage multi-cibles pourrait être mis en oeuvre. Par conséquent, nous pourrions nous demander comment déterminer un seuil à partir duquel l'algorithme de pistage d'une cible étendue ne serait plus pertinent et un pistage de type multi-cibles devrait être considéré. Le choix du seuil serait à nouveau difficile à définir puisqu'il reposerait sur une statistique de test définie sur un groupe de Lie et dont la distribution ne pourrait pas aisément être explicitée. Ensuite, dans ce cadre multi-cibles, des méthodes fondées sur les ensembles finis aléatoires pourraient être mises en places puisqu'elles ont montré leur efficacité dans diverses applications, à travers le filtre *probability density hypothesis* et le filtre *multi-Bernoulli* [106][86]. Les paramètres fournis par la sortie de l'algorithme sur groupe de Lie devraient être injectés dans celui fondé sur les ensembles finis aléatoires. Dans ce contexte, un point délicat est l'initialisation de cette méthode puisqu'elle est fondée sur une paramétrisation différente. Dans le cas où plusieurs amas de débris sont observés simultanément, un pistage étendu multi-cibles pourrait également être mis en place en combinant le formalisme des groupes de Lie et des ensembles finis aléatoires. L'ensemble des paramètres d'un amas appartenant au groupe de Lie $SE(3) \times SO(6) \times D_6(\mathbb{R}^{+*})$ pourrait alors être vu comme un élément d'un ensemble fini aléatoire. Cependant, les mises en oeuvre des algorithmes de type ensembles finis aléatoires sont fondées sur des approximations par mélange de gaussiennes ou mesures empiriques qu'il conviendrait de généraliser.

D'autres perspectives concernent la borne de Cramér-Rao sur groupe de Lie que nous avons

développée. Dans ce manuscrit, la borne a été appliquée à un problème d'estimation simplifié avec un état appartenant à $SE(3)$. Afin de la rendre plus réaliste et cohérente avec notre problème de pistage de débris, il serait pertinent de la déterminer :

- dans le cas où la matrice de forme est supposée inconnue. De par la structure du modèle proposé, il serait néanmoins difficile d'obtenir une expression analytique de la borne associée à l'état $SE(3) \times SO(6) \times D_6(\mathbb{R}^{+*})$. De plus, la gestion des variables latentes devrait être aussi prise en compte dans les développements calculatoires. Une procédure de type Monte-Carlo pourrait alors être utilisée pour obtenir une approximation numérique mais nécessiterait un coût calculatoire assez conséquent.
- dans un cadre récursif : une borne de Cramér-Rao en ligne pourrait alors être mise en place afin de la comparer aux performances obtenues par l'algorithme de pistage de l'amas. De la même manière que pour notre borne, elle pourrait être développée en utilisant la démonstration récursive de la borne de Cramér-Rao euclidienne *a posteriori* en ligne proposée par [94] et [104].

À travers notre modèle, nous pourrions également proposer des généralisations des bornes d'erreurs classiques. Ainsi, il serait possible de mettre en place une borne de *Ziv-Zakai* ou une borne de *Wess-Weinstein* sur groupe de Lie. Les propriétés conservatives des ces bornes pourraient être préservées en considérant des métriques d'erreurs différentes comme une pseudo-métrique bi-invariante [107]. Néanmoins, les difficultés calculatoires en particulier en termes d'intégration et de dérivation, seraient toujours présentes.

Annexes

Annexe A

Résolution et discrétisation d'une équation différentielle

Sommaire

A.1 Cas classique linéaire	151
A.2 Cas non linéaire	152

A.1 Cas classique linéaire

Considérons un vecteur d'état $\mathbf{x}(t)$ (avec $t \in \mathbb{R}$) modélisé par une équation différentielle linéaire matricielle :

$$\dot{\mathbf{x}}(t) = \mathbf{F} \mathbf{x}(t) + \mathbf{n}(t) \quad (\text{A.1})$$

avec :

- $\mathbf{F}$ la matrice d'état.
- $\mathbf{n}(t)$ un terme additif pouvant être un bruit .

La solution générale s'écrit alors :

$$\mathbf{x}(t) = e^{\mathbf{F}t} \boldsymbol{\alpha}(t) \quad (\text{A.2})$$

où $\boldsymbol{\alpha}(t)$ est une fonction du temps à déterminer.

a - Résolution :

La méthode de la variation de la constante nous permet de déterminer cette fonction. Nous en déduisons alors que la solution sous la forme :

$$\mathbf{x}(t) = e^{\mathbf{F}t} \left(\boldsymbol{\alpha}(t_0) + \int_{t_0}^t e^{-\mathbf{F}t'} \mathbf{n}(t') dt' \right) \quad (\text{A.3})$$

Où t_0 est un instant quelconque.

b - Discrétisation :

En discrétisant l'équation (A.3) entre les instants t_{k-1} et t_k , nous obtenons :

$$\mathbf{x}(t_k) = e^{\mathbf{F} t_k} \left(\boldsymbol{\alpha}(t_{k-1}) + \int_{t_{k-1}}^{t_k} e^{-\mathbf{F} t} \mathbf{n}(t) dt \right) \quad (\text{A.4})$$

que l'on peut écrire également :

$$\boxed{\mathbf{x}(t_k) = e^{\mathbf{F} T} \mathbf{x}(t_{k-1}) + \int_{t_{k-1}}^{t_k} e^{\mathbf{F}(t_k-t)} \mathbf{n}(t) dt} \quad (\text{A.5})$$

où $T = t_k - t_{k-1}$ est la période d'échantillonnage. Un développement limité à l'ordre 1 de la matrice exponentielle nous permet d'approcher la solution par :

$$\mathbf{x}(t_k) \cong (\mathbf{I} + \mathbf{F} T) \mathbf{x}(t_{k-1}) + \int_{t_{k-1}}^{t_k} (I + \mathbf{F}(t_k - t)) \mathbf{n}(t) dt \quad (\text{A.6})$$

Finalement, l'équation discrète approchée à l'ordre 1 s'écrit :

$$\boxed{\mathbf{x}_k = \tilde{\mathbf{F}} \mathbf{x}_{k-1} + \mathbf{n}_k} \quad (\text{A.7})$$

avec

$$\begin{cases} \tilde{\mathbf{F}} = & \mathbf{F} + \mathbf{I} T \\ \mathbf{n}_k = & \int_{t_{k-1}}^{t_k} (I + \mathbf{F}(t_k - t)) \mathbf{n}(t) dt \end{cases} \quad (\text{A.8})$$

A.2 Cas non linéaire

Considérons maintenant le problème où le vecteur d'état $\mathbf{x}(t)$ est régi par une équation différentielle non linéaire du type :

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{n}(t) \quad (\text{A.9})$$

où $\mathbf{f}$ est une fonction non linéaire de $\mathbf{x}$.

1ière méthode :

La discrétisation et la linéarisation sont réalisées en prenant l'intégrale de l'équation (A.9), entre les deux instants discrets t_{k-1} et t_k . Autrement dit, nous obtenons :

$$\int_{t_{k-1}}^{t_k} \dot{\mathbf{x}}(t) dt = \int_{t_{k-1}}^{t_k} (\mathbf{f}(\mathbf{x}(t)) + \mathbf{n}(t)) dt \quad (\text{A.10})$$

Pour calculer l'intégrale de droite dans (A.10), nous devons alors approcher $\mathbf{f} \circ \mathbf{x}$ et $\mathbf{n}$ par rapport à la variable t .

- Approximation à l'ordre 0 :

Nous approchons $\mathbf{f} \circ \mathbf{x}$ et $\mathbf{n}$ par leur valeur en t_{k-1} soient $\mathbf{f}(\mathbf{x}(t_{k-1}))$ et $\mathbf{n}(t_{k-1})$. L'équation (A.10) devient alors :

$$\boxed{\mathbf{x}(t_k) = \mathbf{x}(t_{k-1}) + \mathbf{f}(\mathbf{x}(t_{k-1})) T + \mathbf{n}(t_k) T} \quad (\text{A.11})$$

• Approximation à l'ordre 1 :

Nous linéarisons à l'ordre 1 $\mathbf{f} \circ \mathbf{x}$ en considérant comme point de linéarisation t_{k-1} , soit :

$$\mathbf{f}(\mathbf{x}(t)) = \mathbf{f}(\mathbf{x}(t_{k-1})) + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \dot{\mathbf{x}}(t_{k-1}) (t - t_{k-1})$$

où $\nabla \mathbf{f}_{\mathbf{x}(t_{k-1})}$ correspond à la jacobienne de $\mathbf{f}$ calculée en $\mathbf{x}(t_{k-1})$.

En utilisant l'équation (A.10), nous obtenons alors :

$$\mathbf{x}(t_k) = \mathbf{x}(t_{k-1}) + \mathbf{f}(\mathbf{x}(t_{k-1})) T + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \mathbf{x}(t_{k-1}) \frac{T^2}{2} + \mathbf{n}(t_k) T \quad (\text{A.12})$$

$$= \boxed{(\nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \frac{T^2}{2} + \mathbf{I}) \mathbf{x}(t_{k-1}) + \mathbf{f}(\mathbf{x}(t_{k-1})) T + \mathbf{n}(t_k) T} \quad (\text{A.13})$$

2ième méthode :

La deuxième méthode consiste à faire directement une linéarisation de l'équation (A.9) puis de résoudre et discrétiser l'équation différentielle linéarisée selon la méthode de la partie I. Autrement dit, l'équation (A.9) devient :

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t_{k-1})) + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} (\mathbf{x}(t) - \mathbf{x}(t_{k-1})) + \mathbf{n}(t) \quad (\text{A.14})$$

soit :

$$\dot{\mathbf{x}}(t) = \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \mathbf{x}(t) + \mathbf{f}(\mathbf{x}(t_{k-1})) - \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \mathbf{x}(t_{k-1}) + \mathbf{n}(t) \quad (\text{A.15})$$

D'après (A.6), $\mathbf{x}(t)$ s'écrit après discrétisation , à $t = t_k$:

$$\mathbf{x}(t_k) = e^{\nabla \mathbf{f}_{\mathbf{x}(t-1)} T} \mathbf{x}(t_{k-1}) + \int_{t_{k-1}}^{t_k} e^{\nabla \mathbf{f}_{\mathbf{x}(t-1)} (t_k - t)} (\mathbf{f}(\mathbf{x}(t-1)) - \nabla \mathbf{f}_{\mathbf{x}(t-1)} \mathbf{x}(t-1) + \mathbf{n}(t)) dt \quad (\text{A.16})$$

Comme précédemment, nous pouvons approcher la matrice exponentielle :

- Par un développement limité à l'ordre 2 dans le premier terme soit : $e^{\nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} T} = \mathbf{I} + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} T + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \frac{T^2}{2}$.
- Par un développement limité à l'ordre 1 dans le second terme soit : $e^{\nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} (t_k - T)} = \mathbf{I} + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} T$.

Ces deux approximations nous permettent d'obtenir l'équation suivante :

$$\mathbf{x}(t_k) = \left(\mathbf{I} + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} T + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \frac{T^2}{2} \right) \mathbf{x}(t_{k-1}) \quad (\text{A.17})$$

$$+ \int_{t_{k-1}}^{t_k} (\mathbf{I} + \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} T) (\mathbf{f}(\mathbf{x}(t_{k-1})) - \nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \mathbf{x}(t_{k-1})) + \mathbf{n}(t) dt \quad (\text{A.18})$$

Soit après intégration et simplification :

$$\boxed{\mathbf{x}(t_k) = \left(\nabla \mathbf{f}_{\mathbf{x}(t_{k-1})} \frac{T^2}{2} + \mathbf{I} \right) \mathbf{x}(t_{k-1}) + \mathbf{f}(\mathbf{x}(t_{k-1})) T + \mathbf{n}(t_k) T} \quad (\text{A.19})$$

Annexe B

Calculs usuels sur groupe de Lie

Sommaire

B.1	Calculs usuels de dérivée de Lie	155
B.2	Calcul de hessienne sur groupe de Lie	158
B.3	Approximation de Gauss-Laplace et Laplace	158
B.3.1	Rappels dans le cas euclidien	158
B.3.2	Sur groupe de Lie	159

B.1 Calculs usuels de dérivée de Lie

Dans cette section, nous illustrons les définitions de dérivées proposées en chapitre 2 en les appliquant sur des fonctions usuelles définies sur groupe de Lie.

- Fonction à valeurs dans un groupe :

Considérons la fonction f définie par :

$$\begin{aligned} f &: \mathbb{R}^m \rightarrow G' \\ \tau &\mapsto \text{Exp}_{G'}^{\wedge}(\tau) \end{aligned} \tag{B.1}$$

Sa dérivée à *droite* est définie par :

$$\left. \frac{\partial \text{Log}_{G'}^{\vee}(\text{Exp}_{G'}^{\wedge}(\mathbf{h})^{-1} \text{Exp}_{G'}^{\wedge}(\tau + \mathbf{h}))}{\partial \mathbf{h}} \right|_{\mathbf{h}=\mathbf{0}} \tag{B.2}$$

correspondant à la dérivée à *droite* de l'exponentielle de groupe. De manière analogue, sa dérivée à *gauche* est définie par :

$$\left. \frac{\partial \text{Log}_{G'}^{\vee}(\text{Exp}_{G'}^{\wedge}(\tau + \mathbf{h}) \text{Exp}_{G'}^{\wedge}(\mathbf{h})^{-1})}{\partial \mathbf{h}} \right|_{\mathbf{h}=\mathbf{0}} \tag{B.3}$$

correspondant à la définition de $\phi_{G'}(\tau)$ introduite dans le paragraphe précédent à travers la première formule BCH.

- Fonction à valeurs dans $\mathbb{R}$:

Supposons désormais une fonction f définie sur un sous-ensemble A d'un groupe de Lie matriciel G au voisinage de l'identité, à valeurs dans $\mathbb{R}$ et représentant la norme au carré du logarithme de l'inverse :

$$\begin{aligned} f &: A \rightarrow \mathbb{R} \\ \mathbf{X} &\mapsto \|\text{Log}_G^\vee(\mathbf{X}^{-1})\|^2 \end{aligned} \quad (\text{B.4})$$

Nous nous intéressons au calcul de sa dérivée de Lie à droite. D'après l'équation (2.40), et en utilisant le fait que $(\mathbf{X} \text{Exp}_G^\wedge(\boldsymbol{\delta}_\mathbf{X}))^{-1} = \text{Exp}_G^\wedge(-\boldsymbol{\delta}_\mathbf{X}) \mathbf{X}^{-1}$, il est possible d'écrire que :

$$\mathcal{L}_{f(\mathbf{X})}^R = \left. \frac{\partial f(\mathbf{X} \text{Exp}_{SE(3)}^\wedge(\boldsymbol{\delta}_\mathbf{X}))}{\partial \boldsymbol{\delta}_\mathbf{X}} \right|_{\boldsymbol{\delta}_\mathbf{X}=\mathbf{0}} \quad (\text{B.5})$$

En substituant f par son expression, nous obtenons alors :

$$\mathcal{L}_{f(\mathbf{X})}^R = \left. \frac{\partial \|\text{Log}_G^\vee(\text{Exp}_G^\wedge(-\boldsymbol{\delta}_\mathbf{X}) \mathbf{X}^{-1})\|^2}{\partial \boldsymbol{\delta}_\mathbf{X}} \right|_{\boldsymbol{\delta}_\mathbf{X}=\mathbf{0}} \quad (\text{B.6})$$

Une manière de développer cette expression est d'utiliser une des formules BCH. En effet, comme A est un voisinage de l'identité, nous pouvons écrire que $\mathbf{X} = \text{Exp}_G^\wedge(\mathbf{a})$ avec $\mathbf{a} \in \mathbb{R}^m$ (m étant la dimension de G), alors la seconde formule BCH (2.35) nous fournit :

$$\text{Log}_G^\vee(\text{Exp}_G^\wedge(-\boldsymbol{\delta}_\mathbf{X}) \mathbf{X}^{-1}) = \text{Log}_G^\vee(\mathbf{X}^{-1}) - \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1})) \boldsymbol{\delta}_\mathbf{X}, \quad (\text{B.7})$$

où $\boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))$ est la jacobienne à gauche inverse du groupe calculée en $\text{Log}_G^\vee(\mathbf{X}^{-1})$. Ainsi, cette expression prise au carré peut être développée :

$$\begin{aligned} \|\text{Log}_G^\vee(\text{Exp}_G^\wedge(-\boldsymbol{\delta}_\mathbf{X}) \mathbf{X}^{-1})\|^2 &= \|\text{Log}_G^\vee(\mathbf{X}^{-1})\|^2 - 2 \text{Log}_G^\vee(\mathbf{X}^{-1})^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1})) \boldsymbol{\delta}_\mathbf{X} \\ &\quad + \boldsymbol{\delta}_\mathbf{X}^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1})) \boldsymbol{\delta}_\mathbf{X}. \end{aligned} \quad (\text{B.8})$$

La dérivée de Lie s'obtient alors en dérivant cette expression quadratique par rapport à $\boldsymbol{\delta}_\mathbf{X}$ et en la calculant en $\boldsymbol{\delta}_\mathbf{X} = \mathbf{0}$:

$$\mathcal{L}_{f(\mathbf{X})}^R = -2 \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))^\top \text{Log}_G^\vee(\mathbf{X}^{-1}). \quad (\text{B.9})$$

- Fonction à valeurs dans $\mathbb{R}^p$:

Considérons un ultime exemple où la fonction f est définie sur $G = SE(3)$ de la manière suivante :

$$\begin{aligned} f &: SE(3) \rightarrow \mathbb{R}^6 \\ \mathbf{X} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ \mathbf{0} & 1 \end{bmatrix} &\mapsto \mathbf{R} \mathbf{p} \end{aligned} \quad (\text{B.10})$$

Sa dérivée de Lie à droite peut être calculée selon la formule précédente :

$$\mathcal{L}_{g(X)}^R = \left. \frac{\partial f(\mathbf{X} \text{Exp}_{SE(3)}^\wedge(\boldsymbol{\delta}_X))}{\partial \boldsymbol{\delta}_X} \right|_{\boldsymbol{\delta}_X=\mathbf{0}} \quad (\text{B.11})$$

Comme $SE(3)$ est en produit semi-direct avec $SO(3)$ et $\mathbb{R}^3$, l'incrément de dérivation $\boldsymbol{\delta}_X$ peut être séparé en un incrément associé à la dérivation selon $\mathbf{R}$, et un incrément associé à celle de $\mathbf{p}$. Par conséquent, nous devons dériver partiellement selon chacun des deux termes. La fonction étant à valeurs dans $\mathbb{R}^6$, nous allons calculer des quantités correspondant à des matrices jacobiennes de dimension 6×6 . En remplaçant f par son expression, la dérivée de Lie selon $\mathbf{R}$ s'écrit alors :

$$\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{R}} = \left. \frac{\partial \mathbf{R} \text{Exp}_{SO(3)}^\wedge(\boldsymbol{\delta}_R) \mathbf{p}}{\partial \boldsymbol{\delta}_R} \right|_{\boldsymbol{\delta}_R=\mathbf{0}} \quad (\text{B.12})$$

et la dérivée de Lie selon $\mathbf{p}$:

$$\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{p}} = \left. \frac{\partial \mathbf{R} \mathbf{p} \text{Exp}_{\mathbb{R}^3}^\wedge(\boldsymbol{\delta}_p)}{\partial \boldsymbol{\delta}_p} \right|_{\boldsymbol{\delta}_p=\mathbf{0}} \quad (\text{B.13})$$

Le calcul de $\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{R}}$ nous amène à :

$$\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{R}} = \left. \frac{\partial \mathbf{R} \text{Exp}_{SO(3)}^\wedge(\boldsymbol{\delta}_R) \mathbf{p}}{\partial \boldsymbol{\delta}_R} \right|_{\boldsymbol{\delta}_R=\mathbf{0}} \quad (\text{B.14})$$

$$= \left. \frac{\mathbf{R} \partial \text{Exp}_{SO(3)}^\wedge(\boldsymbol{\delta}_R) \mathbf{p}}{\partial \boldsymbol{\delta}_R} \right|_{\boldsymbol{\delta}_R=\mathbf{0}} \quad (\text{B.15})$$

$$= [\mathbf{R} \mathbf{E}_1 \mathbf{p}, \dots, \mathbf{R} \mathbf{E}_p \mathbf{p}] \quad (\text{B.16})$$

où $\{\mathbf{E}_i\}_{i=1}^p$ est une base de $\mathfrak{so}(3)$. Concernant $\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{p}}$, nous avons :

$$\left[\mathcal{L}_{g(X)}^R \right]_{\mathbf{p}} = \left. \frac{\partial \mathbf{R} \mathbf{p} \text{Exp}_{\mathbb{R}^3}^\wedge(\boldsymbol{\delta}_p)}{\partial \boldsymbol{\delta}_p} \right|_{\boldsymbol{\delta}_p=\mathbf{0}} \quad (\text{B.17})$$

$$= \left. \frac{\partial \mathbf{R} (\mathbf{p} + \boldsymbol{\delta}_p)}{\partial \boldsymbol{\delta}_p} \right|_{\boldsymbol{\delta}_p=\mathbf{0}} \quad (\text{B.18})$$

$$= \mathbf{R} \quad (\text{B.19})$$

Par conséquent, la dérivée de Lie de f est la matrice jacobienne suivante :

$$\begin{bmatrix} [\mathbf{R} \mathbf{E}_1 \mathbf{p}, \dots, \mathbf{R} \mathbf{E}_p \mathbf{p}] & \mathbf{R} \end{bmatrix} \quad (\text{B.20})$$

B.2 Calcul de hessienne sur groupe de Lie

À titre d'exemple, nous pouvons reprendre l'exemple de la fonction f de la sous-section précédente définie par :

$$\begin{aligned} f &: A \rightarrow \mathbb{R} \\ \mathbf{X} &\mapsto \|\text{Log}_G^\vee(\mathbf{X}^{-1})\|^2 \end{aligned} \quad (\text{B.21})$$

Le calcul de hessienne peut être effectuée en utilisant la formule (2.57) :

$$\mathbf{H}(\mathbf{X}) = \frac{\partial \|\text{Log}_G^\vee((\mathbf{X} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1) \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_2))^{-1})\|^2}{\partial \boldsymbol{\epsilon}_1 \partial \boldsymbol{\epsilon}_2} \Big|_{\boldsymbol{\epsilon}_1=\mathbf{0}, \boldsymbol{\epsilon}_2=\mathbf{0}} \quad (\text{B.22})$$

Comme $\boldsymbol{\epsilon}_1$ et $\boldsymbol{\epsilon}_2$ ont des valeurs proches de zéro, nous pouvons écrire que $\text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2) = \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1) \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_2)$ et le calcul de la hessienne se ramène à :

$$\mathbf{H}(\mathbf{X}) = \frac{\partial \|\text{Log}_G^\vee((\mathbf{X} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2))^{-1})\|^2}{\partial \boldsymbol{\epsilon}_1 \partial \boldsymbol{\epsilon}_2} \Big|_{\boldsymbol{\epsilon}_1=\mathbf{0}, \boldsymbol{\epsilon}_2=\mathbf{0}} \quad (\text{B.23})$$

De plus, en utilisant la seconde formule BCH, nous pouvons écrire de manière développée le terme à l'intérieur du carré :

$$\text{Log}_G^\vee((\mathbf{X} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2))^{-1}) = \text{Log}_G^\vee(\mathbf{X}^{-1}) - \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2) \quad (\text{B.24})$$

Par conséquent, en prenant la norme au carré de cette expression nous obtenons que :

$$\begin{aligned} \|\text{Log}_G^\vee((\mathbf{X} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2))^{-1})\|^2 &= \|\text{Log}_G^\vee(\mathbf{X}^{-1})\|^2 - 2 \text{Log}_G^\vee(\mathbf{X}^{-1})^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1})) \\ &\quad (\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2) + (\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2)^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2) \end{aligned} \quad (\text{B.25})$$

Une dérivation successive de cette quantité par rapport à $\boldsymbol{\epsilon}_1$ et $\boldsymbol{\epsilon}_2$ et calculée pour $\boldsymbol{\epsilon}_1 = \boldsymbol{\epsilon}_2 = \mathbf{0}$ aboutit à une formule explicite de la Hessienne :

$$\mathbf{H}(\mathbf{X}) = 2 \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1}))^\top \boldsymbol{\psi}_G(\text{Log}_G^\vee(\mathbf{X}^{-1})) \quad (\text{B.26})$$

B.3 Approximation de Gauss-Laplace et Laplace

B.3.1 Rappels dans le cas euclidien

B.3.1.1 Approximation de Gauss-Laplace

Considérons une densité de probabilité $p : \mathbb{R}^m \rightarrow \mathbb{R}$ de forme inconnue pouvant s'écrire de la manière suivante :

$$p(\mathbf{x}) \simeq \alpha \exp(-\|\boldsymbol{\phi}(\mathbf{x})\|^2) \quad (\text{B.27})$$

où $\boldsymbol{\phi} : \mathbb{R}^p \rightarrow \mathbb{R}^m$ est une fonction continue et différentiable et α une constante dans $\mathbb{R}^{+*}$. L'approximation de Gauss-Laplace consiste à effectuer un développement de Taylor-Young à

l'ordre 1 de ϕ afin d'obtenir une approximation gaussienne de p . Plus précisément nous avons :

$$\phi(\mathbf{x}) \simeq \phi(\hat{\mathbf{x}}) + \mathbf{J}_{\hat{\mathbf{x}}}(\mathbf{x} - \hat{\mathbf{x}}) \quad (\text{B.28})$$

où $\mathbf{J}_{\hat{\mathbf{x}}}$ est la jacobienne de ϕ calculée en $\hat{\mathbf{x}}$. Si $\hat{\mathbf{x}}$ est un point critique de $\|\phi\|^2$, alors p peut être approchée par :

$$p(\mathbf{x}) \simeq \alpha \exp(-\|\phi(\hat{\mathbf{x}}) + \mathbf{J}_{\hat{\mathbf{x}}}(\mathbf{x} - \hat{\mathbf{x}})\|^2) \quad (\text{B.29})$$

Comme $\hat{\mathbf{x}}$ est un point critique les termes en $\mathbf{x} - \hat{\mathbf{x}}$, d'où :

$$p(\mathbf{x}) \simeq \alpha' \exp\left(-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^{\top} \mathbf{P}^{-1}(\mathbf{x} - \hat{\mathbf{x}})\right) \quad (\text{B.30})$$

où α' est une constante positive et $\mathbf{P} = \frac{1}{2}(\mathbf{J}_{\hat{\mathbf{x}}}^{\top} \mathbf{J}_{\hat{\mathbf{x}}})^{-1}$. Par conséquent, $p(\mathbf{x})$ peut s'écrire sous la forme :

$$p(\mathbf{x}) \simeq \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P}) \quad (\text{B.31})$$

B.3.1.2 Approximation de Laplace

Considérons désormais une densité de probabilité plus générique s'écrivant sous la forme :

$$p(\mathbf{x}) = \alpha \exp(-J(\mathbf{x})) \quad (\text{B.32})$$

où J est une fonction continue et deux fois différentiable et $\alpha \in \mathbb{R}$. L'approximation de Laplace consiste à approcher à l'ordre J à partir d'un point $\hat{\mathbf{x}}$:

$$J(\mathbf{x}) = J(\hat{\mathbf{x}}) + \nabla J^{\top}(\hat{\mathbf{x}} - \mathbf{x}) + \frac{1}{2}(\hat{\mathbf{x}} - \mathbf{x})^{\top} \mathbf{H}(\hat{\mathbf{x}} - \mathbf{x}) \quad (\text{B.33})$$

o ∇J et $\mathbf{H}$ sont respectivement le gradient et la hessienne de J calculée en $\hat{\mathbf{x}}$.

Si $\hat{\mathbf{x}}$ est un point critique de J , son gradient est nul, et p peut s'approcher par :

$$p(\mathbf{x}) = \alpha' \exp\left(-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^{\top} \mathbf{H}(\mathbf{x} - \hat{\mathbf{x}})\right) \quad (\text{B.34})$$

où $\alpha' > 0$ est une constante soit :

$$p(\mathbf{x}) \simeq \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{H}^{-1}) \quad (\text{B.35})$$

B.3.2 Sur groupe de Lie

B.3.2.1 Approximation de Gauss-Laplace

Considérons désormais une distribution de probabilité définie sur un groupe de Lie G s'écrivant :

$$p(\mathbf{X}) = \alpha \exp(-\|\phi(\mathbf{X})\|^2) \quad (\text{B.36})$$

où $\phi : G \rightarrow \mathbb{R}^m$ est une fonction différentiable.

Comme dans le cas euclidien, l'idée est d'approcher par un développement limité à l'ordre 1 la fonction ϕ :

$$\phi(\mathbf{X}) \simeq \phi(\hat{\mathbf{X}}) + \mathbf{J}_\phi \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X}) \quad (\text{B.37})$$

où $\mathbf{J}_\phi$ est la jacobienne sur groupe de Lie de ϕ en $\hat{\mathbf{X}}$:

$$\mathbf{J}_\phi = \left. \frac{\partial \phi(\hat{\mathbf{X}} \text{Exp}_G^\wedge(\boldsymbol{\delta}))}{\partial \boldsymbol{\delta}} \right|_{\boldsymbol{\delta}=\mathbf{0}} \quad (\text{B.38})$$

De manière analogue au cas euclidien, si $\hat{\mathbf{X}}$ est un point critique de $\|\phi(\mathbf{X})\|^2$, alors cette dernière peut être approchée par un développement limité à l'ordre 2 par :

$$\|\phi(\mathbf{X})\|^2 = \|\phi(\hat{\mathbf{X}})\|^2 + \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})^\top \mathbf{J}_\phi^\top \mathbf{J}_\phi \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X}) \quad (\text{B.39})$$

et $p(\mathbf{X})$ approchée par :

$$\hat{p}(\mathbf{X}) = \alpha \exp \left(-\|\phi(\hat{\mathbf{X}}) + \mathbf{J}_\phi \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})\|^2 \right) \quad (\text{B.40})$$

$$= \alpha' \exp \left(-\text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})^\top \mathbf{J}_\phi^\top \mathbf{J}_\phi \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X}) \right) \quad (\text{B.41})$$

avec $\alpha' = \alpha \exp \left(-\|\phi(\hat{\mathbf{X}})\|^2 \right)$; Par conséquent, p peut être approchée par une distribution gaussienne concentrée à gauche sous la forme :

$$\hat{p}(\mathbf{X}) = \mathcal{N}_G^L(\mathbf{X}; \hat{\mathbf{X}}, \mathbf{P}) \quad (\text{B.42})$$

avec $\mathbf{P} = \frac{1}{2} \left(\mathbf{J}_\phi^\top \mathbf{J}_\phi \right)^{-1}$.

B.3.2.2 Approximation de Laplace

Considérons désormais que $\mathbf{X} \in G$ de dimension m suit une densité de probabilité plus générique s'écrivant sous la forme :

$$p(\mathbf{X}) = \alpha \exp(-J(\mathbf{X})) \quad (\text{B.43})$$

où J est une fonction continue et deux fois différentiable. L'approximation de Laplace consiste à développer à l'ordre 2 J à partir d'un point $\hat{\mathbf{X}}$:

$$J(\mathbf{X}) = J(\hat{\mathbf{X}}) + \left(\mathcal{L}_{J(\hat{\mathbf{X}})}^R \right)^\top \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X}) + \frac{1}{2} \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})^\top \mathbf{H} \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X}) \quad (\text{B.44})$$

où :

$$\left\{ \begin{array}{l} \mathcal{L}_{J(\hat{\mathbf{X}})}^R = \left. \frac{\partial J(\hat{\mathbf{X}} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}))}{\partial \boldsymbol{\epsilon}} \right|_{\boldsymbol{\epsilon}=\mathbf{0}} \quad \text{avec } \boldsymbol{\epsilon} \in \mathbb{R}^m \\ \mathbf{H} = \left. \frac{\partial}{\partial \boldsymbol{\epsilon}_1} \frac{\partial}{\partial \boldsymbol{\epsilon}_2} J(\hat{\mathbf{X}} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1) \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_2)) \right|_{\boldsymbol{\epsilon}_1=\mathbf{0}, \boldsymbol{\epsilon}_2=\mathbf{0}} \quad \forall (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) \in (\mathbb{R}^m)^2. \end{array} \right. \quad (\text{B.45})$$

$$\left\{ \begin{array}{l} \mathcal{L}_{J(\hat{\mathbf{X}})}^R = \left. \frac{\partial J(\hat{\mathbf{X}} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}))}{\partial \boldsymbol{\epsilon}} \right|_{\boldsymbol{\epsilon}=\mathbf{0}} \quad \text{avec } \boldsymbol{\epsilon} \in \mathbb{R}^m \\ \mathbf{H} = \left. \frac{\partial}{\partial \boldsymbol{\epsilon}_1} \frac{\partial}{\partial \boldsymbol{\epsilon}_2} J(\hat{\mathbf{X}} \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_1) \text{Exp}_G^\wedge(\boldsymbol{\epsilon}_2)) \right|_{\boldsymbol{\epsilon}_1=\mathbf{0}, \boldsymbol{\epsilon}_2=\mathbf{0}} \quad \forall (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) \in (\mathbb{R}^m)^2. \end{array} \right. \quad (\text{B.46})$$

Si $\hat{\mathbf{X}}$ est un point critique de J , sa dérivée de Lie est nulle, et p peut s'approcher par :

$$p(\mathbf{X}) = \alpha' \exp^{-\frac{1}{2} \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})^\top \mathbf{H} \text{Log}_G^\vee(\hat{\mathbf{X}}^{-1} \mathbf{X})} \quad (\text{B.47})$$

soit :

$$p(\mathbf{X}) \simeq \mathcal{N}_G^L(\mathbf{X}; \hat{\mathbf{X}}, \mathbf{H}^{-1}) \quad (\text{B.48})$$

Annexe C

Calculs des dérivées du critère d'optimisation

Sommaire

C.1	Calcul de la jacobienne de $\varepsilon(\mathbf{X}_k^{v,Z})$	162
C.2	Calcul de la dérivée de Lie et de la hessienne de $J_{nq}(\mathbf{X}_k^{v,Z})$	165

C.1 Calcul de la jacobienne de $\varepsilon(\mathbf{X}_k^{v,Z})$

La jacobienne globale $\mathbf{J}_\varepsilon$ est construite à partir des matrices jacobienues obtenues selon chaque paramètre.

• Jacobienne de $\phi(\cdot, \mathbf{Z}_{k,i})$

Nous rappelons que $\mathbf{Z}_{k,i} \in SE(3)$. Par définition de la jacobienne sur groupe de Lie, nous avons :

$$\mathbf{J}_{\phi_{\mathbf{Z},i}} = \left. \frac{\partial \phi(\cdot, \mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\delta \mathbf{Z}))}{\partial \delta \mathbf{Z}} \right|_{\delta \mathbf{Z}=\mathbf{0}} \quad (\text{C.1})$$

Il s'ensuit :

$$\mathbf{J}_{\phi_{\mathbf{Z},i}} = \left. \frac{\partial (\mathbf{z}_{k,i} - \Pi(\mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\delta \mathbf{Z})))}{\partial \delta \mathbf{Z}} \right|_{\delta \mathbf{Z}=\mathbf{0}} \quad (\text{C.2})$$

$$= \left. \frac{\partial (-\mathbf{A} \mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\delta \mathbf{Z}) \mathbf{B})}{\partial \delta \mathbf{Z}} \right|_{\delta \mathbf{Z}=\mathbf{0}} \quad (\text{C.3})$$

où :

$$\begin{cases} \mathbf{A} = [\mathbf{I}_3, \mathbf{0}_{3 \times 1}] \\ \mathbf{B} = [\mathbf{0}_{1 \times 3}, 1]^\top. \end{cases}$$

Pour obtenir une expression analytique, nous introduisons $\{\mathbf{G}_l\}_{l=1}^p$, une base de vecteurs de $\mathfrak{se}(3)$ (avec $p = 6$). La dérivée de $\mathbf{A} \mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\delta \mathbf{Z}) \mathbf{B}$ en $\delta \mathbf{Z}$ lorsqu'il tend vers $\mathbf{0}$ est alors donnée par :

$$\left[\mathbf{A} \mathbf{Z}_{k,i} \mathbf{G}_1 \mathbf{B} \mid \dots \mid \mathbf{A} \mathbf{Z}_{k,i} \mathbf{G}_p \mathbf{B} \right]. \quad (\text{C.4})$$

Ainsi, cela implique :

$$\mathbf{J}_{\phi_{\mathbf{Z},i}} = - \left[\mathbf{A} \mathbf{Z}_{k,i} \mathbf{G}_1 \mathbf{B} \mid \dots \mid \mathbf{A} \mathbf{Z}_{k,i} \mathbf{G}_p \mathbf{B} \right]. \quad (\text{C.5})$$

• **Jacobienne de $\psi(\mathbf{X}_k^v, \mathbf{Z}_{k,i})$ selon $\mathbf{M}_k^v$, $\mathbf{P}_k$ et $\mathbf{D}_k$**

Elle est donnée par :

$$\mathbf{J}_{\psi_{\mathbf{X},i}} = \left[\mathbf{J}_{\mathbf{M}^v,i}, \mathbf{J}_{\mathbf{P},i}, \mathbf{J}_{\mathbf{D},i} \right] \quad (\text{C.6})$$

où $\mathbf{J}_{\mathbf{M}^v,i}$, $\mathbf{J}_{\mathbf{P},i}$ et $\mathbf{J}_{\mathbf{D},i}$ sont les matrices jacobienues respectivement selon $\mathbf{M}_k^v$, $\mathbf{P}_k$, et $\mathbf{D}_k$.

1) *Jacobienne selon $\mathbf{M}_k^v$:*

Premièrement, la jacobienne selon $\mathbf{M}_k$, notée $\mathbf{J}_{\mathbf{M},i}$ est calculée. Ensuite, elle est complétée par des valeurs nulles afin d'obtenir $\mathbf{J}_{\mathbf{M}^v,i}$.

$$\begin{aligned} \mathbf{J}_{\mathbf{M},i} &= \left. \frac{\partial \psi(\mathbf{M}_k \text{Exp}_{SE(3)}^\wedge(\delta \mathbf{M}), \mathbf{P}_k, \mathbf{D}_k, \mathbf{Z}_{k,i})}{\partial \delta \mathbf{M}} \right|_{\delta \mathbf{M}=\mathbf{0}} \\ &= \left. \frac{\partial (\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\text{Exp}_{SE(3)}^\wedge(-\delta \mathbf{M}) \mathbf{M}_k^{-1} \mathbf{Z}_{k,i}))}{\partial \delta \mathbf{M}} \right|_{\delta \mathbf{M}=\mathbf{0}} \\ &\simeq \left. \frac{-\partial \left(\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) + \mathbf{D}_k^{-1/2} \mathbf{P}_k \delta \mathbf{M} \right)}{\partial \delta \mathbf{M}} \right|_{\delta \mathbf{M}=\mathbf{0}} \\ &= -\mathbf{D}_k^{-1/2} \mathbf{P}_k. \end{aligned} \quad (\text{C.7})$$

Ainsi :

$$\mathbf{J}_{\mathbf{M}^v,i} = \left[-\mathbf{D}_k^{-1/2} \mathbf{P}_k, \mathbf{0}_{6 \times 3} \right]. \quad (\text{C.8})$$

2) *Jacobienne selon $\mathbf{P}_k$:*

$$\begin{aligned} \mathbf{J}_{\mathbf{P},i} &= \left. \frac{\partial \psi(\mathbf{X}_k, \mathbf{P}_k \text{Exp}_{SO(6)}^\wedge(\delta \mathbf{P}), \mathbf{D}_k, \mathbf{Z}_{k,i})}{\partial \delta \mathbf{P}} \right|_{\delta \mathbf{P}=\mathbf{0}} \\ &= \left. \frac{\partial \left(\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Exp}_{SO(6)}^\wedge(\delta \mathbf{P}) \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) \right)}{\partial \delta \mathbf{P}} \right|_{\delta \mathbf{P}=\mathbf{0}} \end{aligned} \quad (\text{C.9})$$

Nous considérons $\{\mathbf{E}_l\}_{l=1}^g$ une base de vecteurs de $\mathfrak{so}(6)$ (avec $g = 15$). Ainsi, nous avons :

$$\mathbf{J}_{\mathbf{P},i} = [\mathbf{D}_k^{-1/2} \mathbf{P}_k \mathbf{E}_1 \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) | \dots \dots | \mathbf{D}_k^{-1/2} \mathbf{P}_k \mathbf{E}_g \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i})] \quad (\text{C.10})$$

3) *Jacobienne selon $\mathbf{D}_k$*

$$\begin{aligned} \mathbf{J}_{\mathbf{D},i} &= \left. \frac{\partial \psi(\mathbf{X}_k, \mathbf{P}_k, \mathbf{D}_k \text{Exp}_{D_6(\mathbb{R}^{+*})}^\wedge(\boldsymbol{\delta}_{\mathbf{D}}), \mathbf{Z}_{k,i})}{\partial \boldsymbol{\delta}_{\mathbf{D}}} \right|_{\boldsymbol{\delta}_{\mathbf{D}}=\mathbf{0}} \\ &= \left. \frac{\partial \left(\text{Exp}_{D_6(\mathbb{R}^{+*})}^\wedge(-\frac{1}{2} \boldsymbol{\delta}_{\mathbf{D}}) \mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) \right)}{\partial \boldsymbol{\delta}_{\mathbf{D}}} \right|_{\boldsymbol{\delta}_{\mathbf{D}}=\mathbf{0}} \end{aligned} \quad (\text{C.11})$$

Notons que $\left. \frac{\partial \left(\text{Exp}_{D_6(\mathbb{R}^{+*})}^\wedge(-\frac{1}{2} \boldsymbol{\delta}_{\mathbf{D}}) \right)}{\partial (\boldsymbol{\delta}_{\mathbf{D}})_l} \right|_{(\boldsymbol{\delta}_{\mathbf{D}})_l=0} = -\frac{1}{2} \Upsilon_{l,l}$ où $\Upsilon_{l,l}$ est une matrice de taille 6×6

où toutes les entrées valent zéro à l'exception de l'élément l de la diagonale. Par conséquent, nous avons :

$$\mathbf{J}_{\mathbf{D},i} = -\frac{1}{2} \text{diag} \left(\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) \right). \quad (\text{C.12})$$

• **Jacobienne de $\psi(\mathbf{X}_k^v, \mathbf{Z}_{k,i})$ selon $\mathbf{Z}_{k,i}$**

$$\begin{aligned} \mathbf{J}_{\psi_{\mathbf{Z},i}} &= \left. \frac{\partial \psi(\mathbf{X}_k, \mathbf{P}_k, \mathbf{D}_k, \mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\boldsymbol{\delta}_{\mathbf{Z}}))}{\partial \boldsymbol{\delta}_{\mathbf{Z}}} \right|_{\boldsymbol{\delta}_{\mathbf{Z}}=\mathbf{0}} \\ &= \left. \frac{\partial \left(\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i} \text{Exp}_{SE(3)}^\wedge(\boldsymbol{\delta}_{\mathbf{Z}})) \right)}{\partial \boldsymbol{\delta}_{\mathbf{Z}}} \right|_{\boldsymbol{\delta}_{\mathbf{Z}}=\mathbf{0}} \\ &\simeq \left. \frac{\partial \left(\mathbf{D}_k^{-1/2} \mathbf{P}_k \text{Log}_{SE(3)}^\vee(\mathbf{M}_k^{-1} \mathbf{Z}_{k,i}) + \mathbf{D}_k^{-1/2} \mathbf{P}_k \boldsymbol{\delta}_{\mathbf{Z}} \right)}{\partial \boldsymbol{\delta}_{\mathbf{Z}}} \right|_{\boldsymbol{\delta}_{\mathbf{Z}}=\mathbf{0}} \\ &= \mathbf{D}_k^{-1/2} \mathbf{P}_k. \end{aligned} \quad (\text{C.13})$$

• **Jacobienne de $\text{Log}_{G_x}^{\vee} \left(\left(\hat{\mathbf{X}}_{k|k-1}^v \right)^{-1} \mathbf{X}_k^v \right)$**

$$\begin{aligned}
 \mathbf{J}_p &= \left. \frac{\partial \text{Log}_{G'}^{\vee} \left(\left(\hat{\mathbf{X}}_{k|k-1}^v \right)^{-1} \mathbf{X}_k^v \text{Exp}_{G'}^{\wedge}(\delta_{\mathbf{X}}) \right)}{\partial \delta_{\mathbf{X}}} \right|_{\delta_{\mathbf{X}}=0} \\
 &\simeq \left. \frac{\partial \left(\text{Log}_{G_X}^{\vee} \left(\left(\hat{\mathbf{X}}_{k|k-1}^v \right)^{-1} \mathbf{X}_k^v \right) + \text{Log}_{G_X}^{\vee} (\text{Exp}_{G'}^{\wedge}(\delta_{\mathbf{X}})) \right)}{\partial \delta_{\mathbf{X}}} \right|_{\delta_{\mathbf{X}}=0} \\
 &= \mathbf{I}_{30 \times 30}
 \end{aligned} \tag{C.14}$$

• **Jacobienne globale**

Par concaténant les jacobienues calculées, nous obtenons que la jacobienne globale $\mathbf{J}_{\epsilon}$ de dimension $(9n_k + 30) \times (6n_k + 30)$ s'écrit :

$$\mathbf{J}_{\epsilon} = \begin{bmatrix} \mathbf{0}_{3n_k \times 30} & \begin{bmatrix} \mathbf{J}_{\phi_{\mathbf{Z},1}} & & \\ & \ddots & \\ & & \mathbf{J}_{\phi_{\mathbf{Z},n_k}} \end{bmatrix} \\ \begin{bmatrix} \mathbf{J}_{\psi_{\mathbf{X},1}} \\ \vdots \\ \mathbf{J}_{\psi_{\mathbf{X},n_k}} \end{bmatrix} & \begin{bmatrix} \mathbf{J}_{\psi_{\mathbf{Z},1}} & & \\ & \ddots & \\ & & \mathbf{J}_{\psi_{\mathbf{Z},n_k}} \end{bmatrix} \\ \mathbf{I}_{30 \times 30} & \mathbf{0}_{30 \times 6n_k} \end{bmatrix} \tag{C.15}$$

C.2 Calcul de la dérivée de Lie et de la hessienne de $J_{nq}(\mathbf{X}_k^{v,Z})$

• **Calcul de la dérivée de Lie**

$$J_{nq}^{(k)}(\mathbf{D}_k) = n_k \sum_{j=1}^n \log(d_{k,j}) \tag{C.16}$$

La dérivée de Lie à droite de $J_{nq}(\mathbf{D}_k)$ peut naturellement être calculée de la manière suivante :

$$\left[\mathcal{L}^R(J_{nq}^{(k)}(\mathbf{D}_k)) \right]_j = \left. \frac{\partial \left(n \sum_{j=1}^6 \log(d_{k,j} \exp(\tilde{d}_i)) \right)}{\partial \tilde{d}_i} \right|_{\tilde{d}_i=0} \tag{C.17}$$

Comme $\log$ et $\exp$ correspondent respectivement aux fonctions logarithmes et exponentielles sur $]0, +\infty[$ et sur $\mathbb{R}$, nous obtenons :

$$n_k \sum_{j=1}^6 \log(d_{k,j} \exp(\tilde{d}_j)) = n_k \sum_{j=1}^6 \log(d_{k,j}) + n_k \sum_{j=1}^6 \tilde{d}_j \quad \forall j \in \{1, \dots, 6\} \quad (\text{C.18})$$

Ceci implique alors que :

$$\left. \frac{\partial \left(n_k \sum_{j=1}^6 \log(d_{k,j} \exp(\tilde{d}_j)) \right)}{\partial \tilde{d}_j} \right|_{\tilde{d}_i=0} = n_k \quad \forall i \in \{1, \dots, 6\} \quad (\text{C.19})$$

et :

$$\mathcal{L}^R(J_{nq}^{(k)}(\mathbf{D}_k)) = [n_k, \dots, n_k]^\top \quad (\text{C.20})$$

• Calcul de la hessienne

Par définition de la hessienne sur groupe de Lie, chaque entrée $\mathbf{H}_{nq,i,l}$ la matrice $\mathbf{H}_{nq}$ la définissant s'écrit :

$$\mathbf{H}_{nq,i,l} = \left. \frac{\partial \left(n_k \sum_{j=1}^6 \log(d_{k,j} \exp(\tilde{d}_i) \exp(\tilde{d}_l)) \right)}{\partial \tilde{d}_i \partial \tilde{d}_l} \right|_{\tilde{d}_i=0, \tilde{d}_l=0} \quad (\text{C.21})$$

Or :

$$\log(d_{k,j} \exp(\tilde{d}_i) \exp(\tilde{d}_l)) = \log(d_{k,j}) + (\tilde{d}_i + \tilde{d}_l) \quad \forall (i, l) \in \{1, \dots, 6\}^2 \quad (\text{C.22})$$

Par conséquent, une dérivation successive selon $\tilde{d}_i$ et $\tilde{d}_k$ fournit :

$$\mathbf{H}_{i,k} = 0. \quad (\text{C.23})$$

Annexe D

Compléments sur le calcul de la borne de Cramér-Rao

Sommaire

D.1	Rappel de la démonstration de la borne de Cramér-Rao <i>a posteriori</i> (4.7)	167
D.2	Détails du calcul de $\mathbb{E}(\mathbf{H}_{\mathbf{Z}}(\mathbf{M}))$	168
D.3	Détails du calcul de $\mathbb{E}(\mathbf{H}_{\mathbf{M}}(\mathbf{M}))$	169

D.1 Rappel de la démonstration de la borne de Cramér-Rao *a posteriori* (4.7)

Soit $\boldsymbol{\theta} \in \mathbb{R}^m$ un vecteur aléatoire de loi $p(\boldsymbol{\theta})$ et $\mathbf{z}$ une observation distribuée selon $p(\mathbf{z}|\boldsymbol{\theta})$.
Considérons le vecteur aléatoire :

$$\boldsymbol{\phi} = \begin{bmatrix} \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \\ \nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta}) \end{bmatrix} \quad (\text{D.1})$$

dont la matrice de corrélation s'écrit :

$$\mathbb{E}(\boldsymbol{\phi} \boldsymbol{\phi}^\top) = \begin{bmatrix} \mathbf{C} & \mathbf{M} \\ \mathbf{M}^\top & \mathbf{P}^{-1} \end{bmatrix} \quad (\text{D.2})$$

où :

$$\begin{cases} \mathbf{M} = \mathbb{E}(\nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta})^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})) & (\text{D.3}) \\ \mathbf{C} = \mathbb{E}((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top) & (\text{D.4}) \\ \mathbf{P}^{-1} = \mathbb{E}(\nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta}) (\nabla_{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta})^\top)) & (\text{D.5}) \end{cases}$$

En supposant que $p(\boldsymbol{\theta})$ tend vers zéro lorsque $\boldsymbol{\theta}$ tend vers l'infini, il en est déduit par intégration par parties que :

$$\mathbf{M} = \mathbf{M}_1 - \int \mathbf{I} p(\mathbf{z}, \boldsymbol{\theta}) d\mathbf{z} d\boldsymbol{\theta} \quad (\text{D.6})$$

$$= -\mathbb{E}(\mathbf{I}) \quad (\text{D.7})$$

$$= -\mathbf{I} \quad (\text{D.8})$$

où :

$$[\mathbf{M}_1]_{i,j} = \left[\left[\int p(\mathbf{z}, \boldsymbol{\theta}) (\hat{\theta}_j - \theta_j) d\mathbf{z} d\boldsymbol{\theta}_{-i} \right]_{\theta_i}^{\theta_i} \right]_{i,j} = 0 \quad (\text{D.9})$$

et :

$$\boldsymbol{\theta}_{-i} = [\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n]^\top \quad (\text{D.10})$$

En utilisant le même argument, il est obtenu que :

$$\mathbf{P}^{-1} = - \int p(\mathbf{z}, \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \log p(\mathbf{z}, \boldsymbol{\theta}) d\mathbf{z} d\boldsymbol{\theta} \quad (\text{D.11})$$

$$= \mathbb{E} \left(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\mathbf{z}, \boldsymbol{\theta}) \right) \quad (\text{D.12})$$

Comme $\mathbb{E}(\boldsymbol{\phi} \boldsymbol{\phi}^\top)$ est une matrice semi-définie positive, son complément de Schur l'est également. Par conséquent :

$$\text{tr}(\mathbf{C} - \mathbf{M}^\top \mathbf{P}^{-1} \mathbf{M}) \geq 0 \quad (\text{D.13})$$

Comme $\mathbf{M} = \mathbf{I}$ et $\text{tr}(\mathbf{C}) = \mathbb{E}_{p(\mathbf{z}, \boldsymbol{\theta})}(\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_p\|^2)$, nous obtenons ainsi l'inégalité (4.7).

D.2 Détails du calcul de $\mathbb{E}(\mathbf{H}_Z(\mathbf{M}))$

D'après l'équation (4.85), nous savons que :

$$\mathbf{H}_Z(\mathbf{M}) = -\mathbb{E} \left(\hat{\boldsymbol{\psi}}_G(\boldsymbol{\delta}_{MZi})^\top \mathbf{S}^{-1} \hat{\boldsymbol{\psi}}_G(\boldsymbol{\delta}_{MZi}) \right) n \quad (\text{D.14})$$

Le développement limité à l'ordre 1 de $\boldsymbol{\psi}_{SE(3)}(\boldsymbol{\delta}_{MZi})$ est fourni par :

$$\boldsymbol{\psi}_{SE(3)}(\boldsymbol{\delta}_{MZi}) = \mathbf{I} - \frac{1}{2} ad_{SE(3)}(\boldsymbol{\delta}_{MZi}) \quad (\text{D.15})$$

Par conséquent, $\mathbf{H}_Z(\mathbf{M})$ peut s'écrire :

$$\mathbf{H}_Z(\mathbf{M}) \simeq \left(\mathbf{I} - \frac{1}{2} ad_{SE(3)}(\boldsymbol{\delta}_{MZi}) \right)^\top \mathbf{S}^{-1} \left(\mathbf{I} - \frac{1}{2} ad_{SE(3)}(\boldsymbol{\delta}_{MZi}) \right) \quad (\text{D.16})$$

$$\simeq \mathbf{S}^{-1} - \frac{1}{2} \left(ad_{SE(3)}(\boldsymbol{\delta}_{MZi})^\top \mathbf{S}^{-1} + \mathbf{S}^{-1} ad_{SE(3)}(\boldsymbol{\delta}_{MZi}) \right) \quad (\text{D.17})$$

$$+ \frac{1}{4} ad_{SE(3)}(\boldsymbol{\delta}_{MZi})^\top \mathbf{S}^{-1} ad_{SE(3)}(\boldsymbol{\delta}_{MZi}) \quad (\text{D.18})$$

En prenant l'espérance de cette fonction, il en vient que les termes en $ad_{SE(3)}(\delta_{MZi})$ s'annulent et nous obtenons alors l'expression (4.91).

D.3 Détails du calcul de $\mathbb{E}(\mathbf{H}_M(\mathbf{M}))$

D'après l'équation (4.95), nous savons que :

$$\mathbf{H}_M(\mathbf{M}) = -Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M})^\top \boldsymbol{\psi}_{SE(3)}(\delta_o)^\top \boldsymbol{\Sigma}_M^{-1} \boldsymbol{\psi}_{SE(3)}(\delta_o) Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M}) \quad (\text{D.19})$$

Nous savons d'après [43] que :

$$Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M}) = \expm \left(ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \right). \quad (\text{D.20})$$

Par conséquent, $Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M})$ peut être approchée à l'ordre 1 par :

$$Ad_{SE(3)}(\mathbf{M}_o^{-1} \mathbf{M}) \simeq \mathbf{I} + ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \quad (\text{D.21})$$

Comme nous savons que :

$$\boldsymbol{\psi}_{SE(3)}(\delta_o) = \mathbf{I} - \frac{1}{2} ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \quad (\text{D.22})$$

nous en déduisons que $\mathbf{H}_M(\mathbf{M})$ peut être approchée sous la forme :

$$\begin{aligned} \mathbf{H}_M(\mathbf{M}) \simeq & - \left(\mathbf{I} + ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \right)^\top \left(\mathbf{I} - \frac{1}{2} ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \right)^\top \\ & \boldsymbol{\Sigma}_M^{-1} \left(\mathbf{I} + ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \right) \left(\mathbf{I} - \frac{1}{2} ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right) \right) \end{aligned} \quad (\text{D.23})$$

En développant cette formule, nous obtenons :

$$\begin{aligned} \mathbf{H}_M(\mathbf{M}) \simeq & \mathbf{I} - \frac{1}{2} ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right)^2 \boldsymbol{\Sigma}_M^{-1} - \\ & \frac{1}{2} \left(ad_{SE(3)} \left(\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M}) \right)^\top \right)^2 \boldsymbol{\Sigma}_M^{-1} + \text{terme d'ordre 4} \end{aligned} \quad (\text{D.24})$$

En décomposant $\text{Log}_{SE(3)}^\vee(\mathbf{M}_o^{-1} \mathbf{M})$ dans une base de $\mathbb{R}^6$ puis en passant à l'espérance, nous obtenons alors l'expression (4.97).

Bibliographie

- [1] D. KESSLER : Sources of orbital debris and the projected environment for future space-craft. *Advances in Space Research*, 1981.
- [2] H. KLINKRAD : *Space Debris*. Wiley, 2010.
- [3] J. MCDOWELL : Rapport technique, Orbital Stop Press, 2009.
- [4] E. PATE-CORNELL et M. SACHON : Risks of particle hits during space walks in low earth orbit. *IEEE Transactions on Aerospace and Electronic Systems*, 2001.
- [5] E. FLETCHER : Space situational awareness : An analysis of threats, consequences and future solutions for Europe's security, safety and prosperity. Rapport technique, European Space Agency, 2010.
- [6] J. E. REED : The AN/FPS-85 radar system. *Proceedings of the IEEE*, 1969, 1969.
- [7] P. CHORMAN : Cobra dane space surveillance capabilities. 4th European Conference on SpaceDebris (ESA SP-587), 2005.
- [8] T. MICHAL, J. P. EGLIZEAUD et J. BOUCHARD : Graves : the new french system for space surveillance. *In Proceedings of the 4th European Conference on SpaceDebris (ESA SP-587)*, 2004.
- [9] J. DARRICAU : *Physique et théorie du radar*. 1993.
- [10] S. S. BLACKMAN et R. POPOLI : *Design and Analysis of Modern Tracking Systems*. 1999.
- [11] M. A. RICHARDS, J.A. SCHEER et W.A. HOLM : *Principles of Modern Radar : Basic Principles*. chapitre 8, pages 19–28. Scitech publishing, 2010.
- [12] J. PETITJEAN : *Contributions au traitement spatio-temporel fondé sur un modèle autorégressif vectoriel des interférences pour améliorer la détection de petites cibles lentes dans un environnement de fouillis hétérogène Gaussien et non Gaussien*. Thèse de doctorat, Université de Bordeaux, 2010.
- [13] B. MAHAFAZ : *Radar systems analysis and desing using Matlab*. 2000.
- [14] J. E. PRUSSING et B. A. CONWAY : *Orbital Mechanics*. Oxford University Press, 1993.
- [15] J.T. HORWOOD et A.B. POORE : Orbital state uncertainty realism. Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, 2011.
- [16] C. YANEZ, M. GUPTA, V. MORAND et J.C. DOLADO : On the Gaussianity validity time for orbital uncertainty propagation. Darmstadt, Germany, Janvier 2019. *International Conference on Information Fusion*.
- [17] S. ARULAMPALAM, S. MASKELL, N. GORDON et T. CLAPP : A tutorial on particle filters for online non-linear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):139–1188, Août 2002.
- [18] R. KALMAN : A new approach to linear filtering and prediction problems. *ASME-Journal of Basic Engineering*, 82:35–45, Août 1960.
- [19] J.J. YIN, J.Q. ZHANG et J. ZHAO : CDF-KF algorithm for conditionally linear gaussian state space models. *In 2nd International Conference on Advanced Computer Control (ICACC)*, 2010.

- [20] N. de Freitas A. DOUCET et N. GORDON : *Sequential Monte Carlo methods in practice*. Springer-Verlag,, 2001.
- [21] H. M. MENEGAZ, J. Y. ISHIHARA, G. A. BORGES et A. N. VARGAS : A new method for the nonlinear transformation of means and covariances in nonlinear filters. *IEEE Transactions on Automatic Control*, 60:2583–2598, Octobre 2015.
- [22] B.M. BELL et F.W. CATHEY : The iterated Kalman filter update as a Gauss-Newton method. *IEEE Transactions on Automatic Control*, 38(2):294–297, 1993.
- [23] S. JULIER et J. UHLMANN : Unscented filtering and nonlinear estimation. *IEEE Transactions on Automatic Control*, 45:477–482, Août 2000.
- [24] D.J. JWO, C.H. TSENG, J.C. LIU et H.D. LEE : Unscented Kalman filtering for single camera based motion and shape estimation. *Sensors*, 11, Juillet 2011.
- [25] B. D. RIPLEY : *Stochastic simulation*. Wiley and Sons, 1987.
- [26] A. DOUCET, S. GODSILL et C. ANDRIEU : On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3), 2000.
- [27] B.M. BELL et F.W. CATHEY : Incremental least squares methods and the extended Kalman filter. *SIAM Journal on Optimization*, 6(3):807–822, 1996.
- [28] C. BISHOP : *Pattern recognition and machine learning*. Springer, 2006.
- [29] J. NOCEDAL et S. J. WRIGHT : *Numerical optimization*. Springer, 2006.
- [30] K. GRANSTROM, M. BAUM et S. REUTER : Extended object tracking : Introduction, overview and applications. *Journal of Advances in Information Fusion*, 12(2):139–174, Décembre 2016.
- [31] J. W. KOCH et M. FELDMANN : Cluster tracking under kinematical constraints using random matrices. *Robotics and Autonomous Systems*, 2009.
- [32] J. W. KOCH et M. FELDMANN : Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 102:1042–1059, Juillet 2008.
- [33] M. BAUM et U.D. HANEBECK : Extended object tracking with random hypersurface models. *IEEE Transactions on Aerospace and Electronic Systems*, 50(1):1042–1059, Janvier 2014.
- [34] M. BAUM et U.D. HANEBECK : Shape tracking of extended objects and group targets with star-convex rhms. Chicago, USA, Janvier 2011. *International Conference on Information Fusion*.
- [35] M. FELDMANN, D. FRÄNKEN et W. KOCH : Tracking of extended objects and group targets using random matrices. *IEEE Transactions on Signal Processing*, 59, 2011.
- [36] J. LIAN et X. R. LI : Tracking of extended object or target group using random matrix — part II : Irregular object.
- [37] K. THORMANN, Marcus BAUM et Jens HONER : Extended target tracking using Gaussian processes with high-resolution automotive radar. *In International Conference on Information Fusion*, 2018.
- [38] N. WAHLSTRÖM et E. ÖZKAN : Extended target tracking using Gaussian processes. *IEEE Transactions on Signal Processing*, 63:4165–4178, Juillet 2015.
- [39] F. FAION, M. BAUM et U. D. HANEBECK : Tracking 3D shapes in noisy point clouds with random hypersurface models. *In International Conference on Information Fusion*, 2012.
- [40] C. RASMUSSEN et C. WILLIAMS : *Gaussian processes for Machine Learning*. MIT Press, 2006.

- [41] M. KUMRU et E. ÖZKAN : 3D extended object tracking using recursive Gaussian processes. *In International Conference on Information Fusion*, 2018.
- [42] A. BAKER : *Matrix Groups : an Introduction to Lie Group Theory*. Springer, 2003.
- [43] J. M. SELIG : *Geometrical Methods in Robotics*. Springer, 1996.
- [44] W. MILLER : *Symmetry Groups and their applications*. Academic Press, 1972.
- [45] M. ABRAMOWITZ et I. A. STEGUN : “§23.1 : Bernoulli and Euler Polynomials and the Euler-Maclaurin Formula”, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover publications, 1972.
- [46] J. SOLA, J. DERAY et D. ATCHUTHAN : A micro Lie theory for state estimation in robotics. 2018.
- [47] G. S. CHIRIKJIAN : *Stochastic Models, Information Theory and Lie Groups, vol 2*. pages 56–57. Birkhäuser, 2012.
- [48] J. H. MANTON et R. MAHONY : The geometry of the Newton method on non-compact Lie group. *Journal of Global Optimization*, 23:309–327, 2002.
- [49] P. A. ABSIL, R. MAHONY et R. SEPULCHRE : *Optimization algorithm on matrix manifolds*. Princeton, 2008.
- [50] J. M. LEE : *Introduction to smooth manifolds*. Springer, 2003.
- [51] X. PENNEC : Bi-invariant means on Lie groups with Cartan-Schouten connections. *In Geometric Science of Information*, 2013.
- [52] X. PENNEC, S. SOMMER et T. FLETCHER : *Riemannian Geometric Statistics in Medical Image Analysis 1st Edition*. Academic Press, 2019.
- [53] J. FARAUT : *Analysis on Lie Group : an Introduction*. Cambridge University Press, 2008.
- [54] H. KARCHER : Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30:509–541, 1977.
- [55] X. PENNEC : Intrinsic statistics on Riemannian manifolds : basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25:127–155, 2006.
- [56] S. SAID, L. BOMBRUN, Y. BERTHOUMIEU et J. H. MANTON : Riemannian Gaussian distributions on the space of symmetric positive definite matrices. *IEEE Transactions on Information Theory*, 63:2153–2170, 2017.
- [57] C. BORDIN et M. G.S. BRUNO : Particle filtering on the complex Stiefel manifold with application to subspace tracking. *In International Conference on Acoustic Speech and Signal Processing*, 2020.
- [58] R. B. ABDALLAH, A. BRELOY, M. N. EL CORSO et David LAUTRU. : Bayesian signal subspace estimation with compound Gaussian sources. *Signal Processing, Elsevier*, 167, 2019.
- [59] T. SEI, H. SHIBATA, A. TAKEMURA, K. OHARA et N. TAKAYAMA. : Properties and applications of Fisher distribution on the rotation group. *Journal of Multivariate Analysis*, 116, 2011.
- [60] P. T. FLETCHER, S. C. JOSHI, C. LU et S. M. PIZER : Gaussian distributions on Lie groups and their application to statistical shape analysis. *In International Conference on Information Processing in Medical Analysis*, 2003.
- [61] P. SMITH, T. DRUMMOND et K. ROUSSOPOULOS : MAP trajectories by representing, propagating and combining pdfs over groups. *In International Conference on Computer Vision*, 2003.
- [62] T. BARFOOT et P. T. FURGALE : Associating uncertainty with three-dimensional poses for use in estimation problem. *IEEE Transactions on Robotics*, 30(3):679–693, Juin

- 2014.
- [63] G. BOURMAUD, R. MÉGRET, A. GIREMUS et Y. BERTHOUMIEU : From intrinsic optimization to iterated extended Kalman filtering on Lie groups. *Journal of Mathematical Imaging and Vision*, 55(3):284–303, Janvier 2016.
 - [64] A. W. LONG, K. C. WOLFE, M. MASHNER et G. S. CHIRIKJIAN : The banana distribution is Gaussian : a localization study with exponential coordinates. *Robotics : Science and Systems*, 2012.
 - [65] C. UDRISTE : *Convex Functions and Optimization Methods on Riemannian Manifolds*.
 - [66] N. BOUMAL : *Optimization and estimation on manifolds*. Thèse de doctorat, Catholic University of Louvain, Louvain-la-Neuve, 2014.
 - [67] J. ZHOU, S. SAID et Y. BERTHOUMIEU. : Online estimation of MGGD : the Riemannian averaged natural gradient method. In *International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, 2019.
 - [68] J. P. DEDIEU, P. PRIOURET et G. MALAJOVICH : Newton method on Riemannian manifolds : Covariant alpha-theory. *Journal of Numerical Analysis* , 23, 2003.
 - [69] A. BARACHANT, S. BONNET, M. CONGEDO et C. JUTTEN : Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing, Elsevier* , 112:172–178, 2013.
 - [70] R. FLETCHER : *Practical Methods of Optimization*. Wiley, Oxford, 1996.
 - [71] B. OWREN et B. WELFERT : The Newton iteration on Lie groups. *BIT Numerical Mathematics* , 40, 2000.
 - [72] P. Y. LEE et J. B. MOORE : Gauss-Newton-on-manifold for pose estimation. *Journal of Industrial and Management Optimization* , 4, 2005.
 - [73] C. J. TAYLOR et D. J. KRIEGMAN : Minimization on the Lie group $SO(3)$ and related manifolds. 1994.
 - [74] D. MIKA et J. JOZWIK : Lie groups methods in blind signal processing. *MDPI Sensors*, 2020.
 - [75] J. CESIC, I. MARKOVIC et I. PETROVIC : Moving object tracking employing rigid body motion on matrix Lie groups. In *International Conference on Information Fusion*, 2016.
 - [76] H. C. BERG : *E. Coli in motion*. Springer, 2004.
 - [77] H. SNOUSSI et C. RICHARD : Tracking on the Riemannian manifold of multivariate normal distributions. In *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, 2009.
 - [78] A. BARRAU et S. BONNABEL : Intrinsic filtering on Lie groups with applications to attitude estimation. *IEEE Transactions on Automatic Control*, 60(2):439–449, Février 2015.
 - [79] M. PILTÉ, S. BONNABEL et F. BARBARESCO : Tracking the Frenet-Serret frame associated to a highly maneuvering target in 3D. In *Annual Conference on Decision and Control*, 2017.
 - [80] J. ČESIĆ, I. MARKOVIC, M. BUKAL et I. PETROVIĆ : Extended information filter on matrix Lie groups. *Elsevier, Automatica*, 82, Janvier 2017.
 - [81] M. BROSSARD, S. BONNABEL et A. BARRAU : Unscented Kalman filter on Lie groups. In *International Conference on Intelligent Robots and Systems*, 2018.
 - [82] C. CHOI et H. I. CHRISTENSEN : Robust 3D visual tracking using particle filtering on the $SE(3)$ group. In *IEEE International Conference on Robotics and Automation*, 2011.
 - [83] I. LISECKIA et R. TRIGGIANI : *Control Theory for Partial Differential Equations* :. Cambridge University Press, 2000.

- [84] G. BOURMAUD, R. MÉGRET, M. ARNAUDON et A. GIREMUS : Continuous-discrete extended Kalman filter on matrix Lie groups. *Journal of Mathematical Imaging and Vision*, 51(1):284–303, Janvier 2015.
- [85] S. SARKKA : *Bayesian filtering and smoothing*. IMS, 2004.
- [86] B.-N. VO, M. MALLICK, Y. BAR-SHALOM, S. CORALUPPI, R. OSBORNE III, R. MAHLER et B.-T. VO : *Multitarget Tracking*. Wiley Encyclopedia of Electrical and Electronics Engineering, 2015.
- [87] J. M. SELIG : *Geometrical Methods in Robotics*, chapitre 4, page 52. Springer, 1996.
- [88] W. RUCKLIDGE : *Efficient Visual Recognition Using the Hausdorff Distance*. Springer-Verlag, 1996.
- [89] M. PEREYRA, N. DOBIGEON, H. BATATIA et J.-Y. TOURNERET : Computing the Cramér-Rao bound of Markov random field parameters : Application to the Ising and the Potts models. *IEEE Signal Processing Letters*, 21:47–50, 2014.
- [90] H. CRAMÉR : *Mathematical Methods of Statistics*. Princeton University Press, 1946.
- [91] D. A. S. FRASER et I. GUTTMAN : *Bhattacharyya Bounds without Regularity Assumptions*. Annals of Mathematical Statistic, 1952.
- [92] K. L. BELL, Y. STEINBERG, Y. EPHRAIM et H. L. VAN TREES : Extended Ziv-Zakai lower bound for vector parameter estimation. *IEEE Transactions on Information Theory*, 63:624–637, Mars 1997.
- [93] A. RENAUX, P. FORSTER, P. LARZABAL, C. D. RICHMOND et A. NEHORAI : A fresh look at the Bayesian bounds of the Weiss-Weinstein family. *IEEE Transactions on Signal Processing*, 56:5334 – 5352, Mars 2008.
- [94] P. TICHAVSKY, C. H. MURAVCHIK et A. NEHORAI : Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46:1386 – 1396, Mars 1998.
- [95] C. REN, J. GALY, E. CHAUMETTE, P. LARZABAL et A. RENAUX : A Ziv-Zakai type bound for hybrid parameter estimation. Florence, Italie, Janvier 2014. *International Conference on Acoustics, Speech and Signal Processing*.
- [96] W. RING et B. WIRTH : Optimization methods on Riemannian manifolds and their application to shape space. *SIAM J. Optim.*, 22(2), 596–627, 2012.
- [97] S. T. SMITH : Covariance, subspace, and intrinsic Cramér-Rao bounds. *IEEE Transactions On Signal Processing*, 53(5):1610–1630, 2005.
- [98] N. BOUMAL : On intrinsic Cramér-Rao bounds for Riemannian submanifolds and quotient manifolds. *IEEE Transactions On Signal Processing*, 61(7):1809–1821, 2013.
- [99] A. BRELOY, G. GINOLHAC, A. RENAUX et F. BOUCHARD : Intrinsic Cramér-Rao bounds for scatter and shape matrices estimation in CES distribution. *IEEE Signal Processing Letters*, 26(2):262–266, Feb.2019.
- [100] A. TERRAS : *Harmonic analysis on symmetric spaces and applications*. Springer-Verlag, 1988.
- [101] S. BONNABEL et A. BARRAU : An intrinsic Cramér-Rao bound on Lie groups. Paris, France, Juin 2015. *International Conference on Networked Geometric Science of Information*.
- [102] G. S. CHIRIKJIAN : Information-theoretic inequalities on unimodular Lie groups. *Journal of Geometric Mechanics*, 2:119–158, 2010.
- [103] U. GRENANDER, M. I. MILLER et A. STRIVASTAVA : Hilbert-Schmidt lower bounds for estimators on matrix Lie groups for atr. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 20:790–802, 1998.

- [104] N. BERGMAN : *Sequential Monte Carlo Methods In Practice*, chapitre 15. Statistics for Engineering and Information Science, Springer, 2001.
- [105] S. BUZZI, M. LOPS et L. VENTURINO : Track before detect procedures for early detection of moving target from airborne radars. *IEEE Transactions on aerospace and electronic systems*, 41, 2005.
- [106] D. CLARK : *Multi-target tracking with the probability density hypothesis filter*. Thèse de doctorat, Heriot-Watt University, 2006.
- [107] N. MIOLANE et X. PENNEC : Computing bi-invariant pseudo-metrics on Lie groups for consistent statistics. *Entropy*, 17, 2015.

Titre : Méthodes statistiques fondées sur les groupes de Lie pour le suivi d'un amas de débris spatiaux

Résumé : Dans le contexte de la surveillance spatiale, nous nous intéressons à un amas de débris évoluant en orbite autour de la Terre et observé par un capteur radar. Il est alors constaté que l'ensemble de ses réflecteurs prépondérants se disperse selon une forme "banane" due au mouvement contraint par les lois de Kepler. Cette répartition est représentative d'échantillons gaussiens concentrés sur le groupe de Lie $SE(3)$ et peut être complètement caractérisée par une matrice de covariance inconnue. Nous proposons dans cette thèse une reformulation originale sur groupe de Lie du modèle d'observation de l'amas. Ce dernier est alors modélisé comme une cible étendue caractérisée par sa forme et son centroïde. Cette manière, nous reconsidérons l'estimation de sa forme et de son centroïde comme un problème d'inférence sur variété. Sa géométrie est ainsi intrinsèquement prise en compte. Deux algorithmes sur groupes de Lie sont alors proposés afin d'estimer respectivement de manière statique et dynamique les paramètres de l'amas. Dans une première partie, l'enjeu de la surveillance spatiale et les principales méthodes de pistage de débris sont rappelés. Dans une seconde partie, les fondements des groupes de Lie sont présentés. La troisième partie est axée sur les contributions de la thèse et propose un modèle et deux algorithmes d'estimation de la forme et du centre de gravité d'un amas qui sont ensuite testés sur différents scénarios de simulation. La dernière partie est consacrée à une contribution théorique dans laquelle est mise en place une borne d'erreur d'estimation bayésienne sur groupe de Lie.

Mots-clés : débris spatiaux, pistage de cible étendue, groupe de Lie, filtrage bayésien, optimisation sur groupe de Lie.

Title : Lie group-based statistical methods for tracking a cluster of space debris

Abstract : In the context of space surveillance, we are interested in a cluster of debris evolving in orbit around the Earth and observed by a radar sensor. It is then observed that the debris spreads out taking a bananoid shape due to their movement constrained by Kepler's laws. This distribution is representative of concentrated Gaussian samples on the Lie group $SE(3)$ and can be completely characterized by an unknown covariance matrix. We propose in this thesis an original reformulation of the cluster observation model on Lie groups. The latter is then modeled as an extended target characterized by its shape and its centroid. In this way, we reconsider its estimation as a manifold inference problem. The geometry of the cluster is thus intrinsically taken into account. Two algorithms on Lie groups are then proposed in order to estimate respectively statically and dynamically the parameters of the cluster. In the first part of the manuscript, the issue of space surveillance is underlined and the main methods for tracking debris are recalled. In a second part, the foundations of Lie groups are presented. The third part focuses on the contributions of the thesis and proposes a model and two algorithms for estimating the shape and centroid of a cluster which are then tested on different simulation scenarios. The last part is devoted to a theoretical contribution in which is proposed a bound for Bayesian estimation error on Lie groups.

Keywords : space debris, extended target tracking, Lie group, Bayesian filtering, optimization on Lie group.