



HAL
open science

Estimation Bayésienne non Paramétrique de Systèmes Dynamiques en Présence de Bruits Alpha-Stables

Nouha Jaoua

► **To cite this version:**

Nouha Jaoua. Estimation Bayésienne non Paramétrique de Systèmes Dynamiques en Présence de Bruits Alpha-Stables. Autre. Ecole Centrale de Lille, 2013. Français. NNT : 2013ECLI0003 . tel-00929691

HAL Id: tel-00929691

<https://theses.hal.science/tel-00929691>

Submitted on 13 Jan 2014

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.

N° d'ordre : 213

ECOLE CENTRALE DE LILLE

THÈSE

présentée en vue d'obtenir le grade de

DOCTEUR

Spécialité : Automatique, Génie Informatique, Traitement du Signal et Image

par

Nouha Jaoua

Doctorat délivré par l'Ecole Centrale de Lille

Estimation Bayésienne non Paramétrique de Systèmes Dynamiques en Présence de Bruits Alpha-Stables

Soutenue le 06 Juin 2013 devant le jury d'examen :

Président	M. Laurent Clavier	Professeur, TELECOM Lille 1
Rapporteur	M. Thierry Chonavel	Professeur, TELECOM Bretagne
Rapporteur	M. Jean-François Giovannelli	Professeur, Université de Bordeaux 1
Membre	M. Nouredine Ellouze	Professeur, Ecole Nationale d'Ingénieurs de Tunis
Membre	M. François Caron	Chargé de Recherche, INRIA
Membre	M. François Septier	Maître de Conférences, TELECOM Lille 1
Directeur de thèse	M. Emmanuel Duflos	Professeur, Ecole Centrale de Lille
Directeur de thèse	M. Philippe Vanheeghe	Professeur, Ecole Centrale de Lille

Thèse préparée au Laboratoire d'Automatique, Génie Informatique et Signal
(LAGIS, UMR CNRS 8219)

Ecole Doctorale SPI 072 (Lille I, Lille III, Artois, ULCO, UVHC, EC Lille)

PRES Université Lille Nord-de-France

Estimation Bayésienne non Paramétrique de Systèmes Dynamiques en Présence de Bruits Alpha-Stables

Résumé : Dans un nombre croissant d'applications, les perturbations rencontrées s'éloignent fortement des modèles classiques qui les modélisent par une gaussienne ou un mélange de gaussiennes. C'est en particulier le cas des bruits impulsifs que nous rencontrons dans plusieurs domaines, notamment celui des télécommunications. Dans ce cas, une modélisation mieux adaptée peut reposer sur les distributions α -stables. C'est dans ce cadre que s'inscrit le travail de cette thèse dont l'objectif est de concevoir de nouvelles méthodes robustes pour l'estimation conjointe état-bruit dans des environnements impulsifs. L'inférence est réalisée dans un cadre bayésien en utilisant les méthodes de Monte Carlo séquentielles. Dans un premier temps, cette problématique a été abordée dans le contexte des systèmes de transmission OFDM en supposant que les distorsions du canal sont modélisées par des distributions α -stables symétriques. Un algorithme de Monte Carlo séquentiel a été proposé pour l'estimation conjointe des symboles OFDM émis et des paramètres du bruit α -stable. Ensuite, cette problématique a été abordée dans un cadre applicatif plus large, celui des systèmes non linéaires. Une approche bayésienne non paramétrique fondée sur la modélisation du bruit α -stable par des mélanges de processus de Dirichlet a été proposée. Des filtres particuliers basés sur des densités d'importance efficaces sont développés pour l'estimation conjointe du signal et des densités de probabilité des bruits.

Mots-clés : bruit impulsif, distributions α -stables, inférence bayésienne, méthodes de Monte Carlo, filtrage particulière, estimation conjointe, systèmes OFDM, estimation non paramétrique de densité, mélange de processus de Dirichlet.

Nonparametric Bayesian Estimation of Dynamical Systems in the Presence of Alpha-Stable Noise

Abstract : In signal processing literature, noise's sources are often assumed to be Gaussian. However, in many fields the conventional Gaussian noise assumption is inadequate and can lead to the loss of resolution and/or accuracy. This is particularly the case of noise that exhibits impulsive nature. The latter is found in several areas, especially telecommunications. α -stable distributions are suitable for modeling this type of noise. In this context, the main focus of this thesis is to propose novel methods for the joint estimation of the state and the noise in impulsive environments. Inference is performed within a Bayesian framework using sequential Monte Carlo methods. First, this issue has been addressed within an OFDM transmission link assuming a symmetric α -stable model for channel distortions. For this purpose, a particle filter is proposed to include the joint estimation of the transmitted OFDM symbols and the noise parameters. Then, this problem has been tackled in the more general context of nonlinear dynamic systems. A flexible Bayesian nonparametric model based on Dirichlet Process Mixtures is introduced to model the α -stable noise. Moreover, sequential Monte Carlo filters based on efficient importance densities are implemented to perform the joint estimation of the state and the unknown measurement noise density.

Keywords : impulsive noise, α -stable distributions, Bayesian inference, Monte Carlo methods, particle filtering, joint estimation, OFDM systems, nonparametric density estimation, Dirichlet process mixture.

A ma mère Nabiha

A mon père Noureddine

A ma sœur Nadia et mon beau-frère Thameur

A mon frère Nabil

A mon frère Nizar et ma belle-sœur Sabeh

Aux petits anges de ma famille : Malak, Omar, Ahmed et Eya

A mon oncle Ali et son épouse Radhia

A l'homme de ma vie Amine

A toute ma famille et ma belle famille

A tous mes amis et amies

Je vous dédie ce travail en témoignage de mon grand amour.

Remerciements

Je tiens tout d'abord à exprimer ma gratitude à Messieurs Emmanuel Duflos et Philippe Vanheeghe, Professeurs à l'Ecole Centrale de Lille pour le co-encadrement de cette thèse. Je leur adresse mes plus sincères remerciements pour la confiance qu'ils m'ont témoignée tout au long de cette thèse, leur soutien moral et intellectuel ainsi que leurs grandes qualités humaines.

Je remercie Messieurs Thierry Chonavel, Professeur à TELECOM Bretagne, et Jean-François Giovannelli, Professeur à l'Université de Bordeaux 1, pour l'intérêt qu'ils ont bien voulu porter à ce travail en acceptant d'en être les rapporteurs. Je remercie également Messieurs Laurent Clavier, Professeur à TELECOM Lille 1, et François Caron, Chargé de Recherche à l'INRIA, de participer à ce jury de thèse.

Je souhaite également exprimer toute ma reconnaissance à Monsieur Nouredine Ellouze, Professeur à l'Ecole Nationale d'Ingénieurs de Tunis, de m'avoir donné l'occasion d'entrer dans le monde de recherche et à qui je dois tout. Je suis particulièrement honoré de son présence dans ce jury de thèse.

Je souhaite également remercier Monsieur François Septier, Maître de Conférences à TELECOM Lille 1, pour la part très importante qu'il a prise dans ce travail. Merci pour son soutien continu et pour la grande disponibilité dont il a fait preuve envers moi durant ces trois années de thèse.

Mes remerciements les plus chaleureux vont à l'ensemble du personnel technique et administratif du LAGIS pour avoir contribué à instaurer une ambiance de travail conviviale et sympathique.

Je tiens également à remercier les nombreuses personnes que j'ai eu la chance de rencontrer pendant cette thèse. Je pense surtout à mes collègues du LAGIS avec qui

j'ai passé de très bons moments. Merci spécialement à Hachem, Sonia et Karama pour leur amitié et leur soutien moral considérable.

Enfin, aucun mot ne saurait exprimer ma reconnaissance et ma gratitude à l'égard de ma famille. Nul mot ne suffirait pour remercier mes parents, pour leur soutien, leurs sacrifices démesurés, leur affection et leurs encouragements permanents. Je suis vraiment fière d'être leur fille. Je remercie infiniment ma sœur, mes frères et mon beau-frère pour tout l'amour, la présence et la confiance dont ils ont fait preuve. Je remercie aussi mes nièces et mes neveux pour le bonheur et la joie qu'ils apportent à ma vie. Je suis aussi redevable à mon oncle, Ali, et son épouse, Radhia, qui m'ont accompagné tout au long de mon parcours universitaire. Je les remercie pour leur amour, leur présence et leurs encouragements permanents. Je conclus en remerciant mon époux, Amine, pour avoir cru en moi et pour m'avoir soutenu par son amour.

Table des matières

Dédicaces	5
Remerciements	7
Table des figures	13
Liste des tableaux	17
Liste des algorithmes	20
Liste des notations	21
Liste des acronymes	23
Introduction	25
1 Bruit impulsif et distributions α-stables	31
1.1 Les distributions α -stables	33
1.1.1 Définitions	33
1.1.2 Fonction caractéristique	35
1.1.3 Différentes paramétrisations	36
1.1.4 Densité de probabilité	37
1.1.5 Propriétés des distributions α -stables	41
1.1.6 Représentation analytique de la densité de probabilité des distributions α -stables	45

1.1.7	Algorithme de simulation	50
1.1.8	Estimation des paramètres des distributions α -stables	50
1.2	Autres modèles du bruit impulsif	52
1.2.1	Modèle de mélange de gaussiennes	52
1.2.2	Loi de Laplace	53
1.2.3	Loi gaussienne généralisée	54
1.2.4	Modèle de Middleton classe A	55
1.3	Conclusion	56
2	Inférence bayésienne	59
2.1	Formulation du problème	59
2.1.1	Système dynamique	59
2.1.2	Inférence bayésienne	60
2.2	Méthodes déterministes	63
2.2.1	Filtre de Kalman	64
2.2.2	Filtre de Kalman étendu	65
2.2.3	Filtre de Kalman sans parfum	67
2.3	Méthodes de Monte Carlo	70
2.3.1	Intégration de Monte Carlo	70
2.3.2	Les méthodes de Monte Carlo séquentielles	74
2.3.3	Filtre particulaire Rao-Blackwellisé	81
2.3.4	Méthodes de Monte Carlo par chaînes de Markov	82
2.4	Inférence bayésienne dans le cas des distributions α -stables	86
2.4.1	Estimation bayésienne des paramètres de la distribution α -stable	87
2.4.2	Estimation bayésienne conjointe du signal et des paramètres du bruit α -stable	89
2.4.3	Estimation bayésienne des mélanges finis de distributions α - stables	89
2.5	Conclusion	90
3	Estimation conjointe état-bruit dans les systèmes OFDM	93
3.1	Les systèmes OFDM	95
3.1.1	Principe de l'OFDM	95
3.1.2	Implémentation numérique d'un système OFDM	96

3.1.3	Etat de l'art sur les méthodes de réduction du bruit impulsif dans les systèmes OFDM	100
3.2	Modèle statistique	103
3.2.1	Modèle d'observation	104
3.2.2	Modèle d'évolution de l'état	105
3.2.3	Le modèle statistique	107
3.3	Filtre particulaire Rao-Blackwellisé pour l'estimation conjointe du signal OFDM et des paramètres du bruit α -stable	108
3.3.1	Filtre de Kalman pour l'estimation de l'état	108
3.3.2	Lois d'importance	109
3.3.3	Evaluation des poids	111
3.3.4	Estimées MMSE du signal OFDM et des paramètres du bruit	111
3.3.5	Algorithme	112
3.4	Simulations	112
3.5	Conclusion	121
4	DP pour l'estimation non paramétrique de densités	125
4.1	Problème d'estimation de densité	126
4.2	Processus de Dirichlet	127
4.2.1	Représentation stick-breaking	128
4.2.2	Distribution <i>a posteriori</i>	130
4.2.3	Représentation en urne de Polya	131
4.2.4	Clustering	132
4.2.5	Processus du restaurant chinois	132
4.2.6	Nombre de valeurs distinctes	133
4.3	Mélange de processus de Dirichlet	134
4.4	Estimation de la densité dans le cas d'un DPM	136
4.4.1	Algorithme d'Escobar	137
4.4.2	Algorithme de Neal	139
4.5	Conclusion	140
5	DPMs pour l'estimation conjointe état-bruit impulsif	143
5.1	Modélisation du bruit d'observation avec les DPMs	145
5.1.1	Modèle de bruit avec DPM de distributions gaussiennes	146

5.1.2	Modèle de bruit avec DPM de distributions de Cauchy	147
5.1.3	Modèle de bruit avec DPM de distributions α -stables	147
5.2	Filtre particulaire pour l'estimation conjointe de l'état et de la densité de probabilité du bruit d'observation	149
5.2.1	Cas d'un bruit d'observation modélisé par un DPM de distri- butions gaussiennes	151
5.2.2	Cas d'un bruit d'observation modélisé par un DPM de distri- butions de Cauchy	155
5.2.3	Cas d'un bruit d'observation modélisé par un DPM de distri- butions α -stables	159
5.2.4	Une troisième densité d'importance	160
5.3	Exemple de mise en œuvre	164
5.3.1	Bruit d'observation modélisé avec un DPM de distributions gaussiennes	164
5.3.2	Bruit d'observation modélisé avec un DPM de distributions de Cauchy	167
5.3.3	Bruit d'observation modélisé avec un DPM de distributions α -stable	169
5.3.4	Comparaison	170
5.3.5	Discussion	174
5.4	Conclusion	177
	Conclusion	179
	Bibliographie	183
	Annexe A : Densité de probabilité de la somme d'une variable aléatoire gaussienne et d'une variable aléatoire de Cauchy	199

Table des figures

1.1	pdf de la distribution α -stable pour différentes valeurs de β	39
1.2	pdf de la distribution α -stable pour différentes valeurs de α	40
1.3	Queue de la pdf de la distribution α -stable pour différentes valeurs de α	41
1.4	pdf de la distribution α -stable pour différentes valeurs de γ	42
1.5	pdf de la distribution α -stable pour différentes valeurs de μ	43
1.6	pdf d'un mélange de deux distributions gaussiennes	52
1.7	pdf de la distribution de Laplace pour différentes valeurs de μ et b . .	53
1.8	pdf de la distribution gaussienne généralisée pour différentes valeurs de μ , σ et p	54
2.1	Représentation graphique du modèle dynamique	61
2.2	Echantillonnage d'importance	74
3.1	Schéma bloc d'un système OFDM	97
3.2	Insertion du préfixe cyclique.	99
3.3	Construction du vecteur $\mathbf{s}_{n,t}$	105
3.4	Les boîtes à moustaches du MSE de l'estimation de α en fonction du nombre de particules pour différentes valeurs de γ	116
3.5	Les boîtes à moustaches du MSE de l'estimation de γ en fonction du nombre de particules pour différentes valeurs de γ	117
3.6	MSE de l'estimation du signal OFDM en fonction de γ pour différentes valeurs de α	118
3.7	MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α	118

3.8	Les boîtes à moustaches du MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $\gamma=0.1$	119
3.9	Les boîtes à moustaches du MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $\gamma=0.01$	119
3.10	Les boîtes à moustaches du MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $\gamma=0.001$	120
3.11	Les boîtes à moustaches du MSE de l'estimation de α pour $\gamma = 10$	121
3.12	Les boîtes à moustaches du MSE de l'estimation de α pour $\gamma = 1$	122
3.13	Les boîtes à moustaches du MSE de l'estimation de α pour $\gamma = 0.1$	123
3.14	Performance en terme de BER de la méthode proposée FPRB-ECSB pour $\alpha = 0.8$	123
3.15	Performance en terme de BER de la méthode proposée FPRB-ECSB pour $\alpha = 1.1$	124
3.16	Performance en terme de BER de la méthode proposée FPRB-ECSB pour $\alpha = 1.4$	124
4.1	Représentation stick-breaking	129
4.2	Réalisations de processus de Dirichlet pour différentes valeurs de α	130
4.3	Modèle statistique hiérarchique d'un DPM	134
4.4	Représentation de la densité F associée à une réalisation d'un DP	135
5.1	Estimation de l'état dans le cas où la densité réelle du bruit d'observation est une distribution α -stable modélisée par un DPM de distributions gaussiennes	166
5.2	Estimation de la densité du bruit dans le cas où le bruit d'observation suit une distribution α -stable modélisé par un DPM de distributions gaussiennes.	167
5.3	Estimation de l'état dans le cas où la densité réelle du bruit d'observation est un mélange de deux distributions α -stables modélisée par un DPM de distributions gaussiennes	168
5.4	Estimation de la densité du bruit dans le cas où le bruit d'observation suit un mélange de deux distributions α -stables modélisé par un DPM de distributions gaussiennes	169

5.5	Estimation de l'état dans le cas où la densité réelle du bruit d'observation est un mélange de deux distributions de Cauchy modélisée par un DPM de distributions de Cauchy	170
5.6	Estimation de la densité du bruit dans le cas où le bruit d'observation suit un mélange de deux distributions de Cauchy modélisé par un DPM de distributions de Cauchy	171
5.7	Estimation de l'état dans le cas où la densité réelle du bruit d'observation est un mélange de deux distributions α -stables modélisée par un DPM de Cauchy	172
5.8	Estimation de la densité du bruit dans le cas où le bruit d'observation suit un mélange de deux distributions α -stables modélisé par un DPM de distributions de Cauchy	173
5.9	Estimation de l'état dans le cas où la densité réelle du bruit d'observation est un mélange de deux distributions α -stables modélisée par un DPM de distributions α -stables	174
5.10	Estimation de la densité du bruit dans le cas où le bruit d'observation suit un mélange de deux distributions α -stables modélisé par un DPM de distributions α -stables	175
5.11	Estimation de la densité dans le cas où la densité réelle est un mélange de trois distributions α -stables	175
5.12	Estimation des queues de la densité avec les différentes modélisations proposées dans le cas où la densité réelle est un mélange de trois distributions α -stables	176
5.13	Evolution temporelle de la distance de Kullback-Leibler avec les différentes modélisations proposées dans le cas où la densité réelle est un mélange de trois distributions α -stables	176
5.14	Evolution temporelle de la distance de Hellinger avec les différentes modélisations proposées dans le cas où la densité réelle est un mélange de trois distributions α -stables	177

Liste des tableaux

3.1	Types de modulation	97
3.2	Biais et variance de l'estimation du signal OFDM en fonction du nombre de particules pour différentes valeurs de γ avec $\alpha = 1.4$	115
5.1	Comparaison entre les trois modélisations proposées en termes des distances de Kullback-Leibler et de Hellinger dans le cas où la densité réelle est un mélange de trois distributions α -stables	174

Liste des algorithmes

1	Approximation de la pdf d'une distribution α -stable symétrique par un mélange fini de gaussiennes	49
2	Simulation de variables aléatoires α -stables $S_1(\alpha, \beta)$	50
3	Filtre de Kalman	65
4	Filtre de Kalman étendu	66
5	Filtre de Kalman sans parfum	69
6	Echantillonnage d'importance séquentiel	76
7	Rééchantillonnage systématique	77
8	Filtre Particulaire	79
9	Filtre Particulaire Rao-Blackwellisé	83
10	Algorithme de Metropolis-Hastings	85
11	Echantillonneur de Gibbs	85
12	Filtre particulaire Rao-Blackwellisé pour l'estimation conjointe du signal OFDM et des paramètres du bruit α -stable	113
13	Algorithme de Gibbs pour échantillonner selon $p(\theta_{1:n} \mathbf{y}_{1:n})$	138
14	Algorithme de Gibbs pour échantillonner selon $p(z_{1:n}, U_{X(z_{1:n})} \mathbf{y}_{1:n})$	141
15	Filtre particulaire pour l'estimation séquentielle de la densité <i>a posteriori</i> $p(\mathbf{x}_{0:t}, \theta_{1:t} \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de gaussiennes	152
16	Filtre particulaire pour l'estimation séquentielle de la densité <i>a posteriori</i> $p(\mathbf{x}_{0:t}, \theta_{1:t} \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de Cauchy	156

-
- 17 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de distributions α -stables 161
- 18 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ 165

Liste des notations

x	scalaire x
\mathbf{x}	vecteur \mathbf{x}
\mathbf{X}	matrice \mathbf{X}
$\mathbf{0}_{n \times m}$	matrice nulle de dimension $n \times m$
\mathbf{I}_n	matrice identité de dimension $n \times n$
$(\cdot)^T$	transposée d'une matrice
$(\cdot)^*$	conjuguée d'une matrice
$(\cdot)^H$	conjuguée Hermitienne d'une matrice
$\Re(\cdot)$	partie réelle
$\Im(\cdot)$	partie imaginaire
$\mathbf{x}_{d:f}$	l'ensemble $\{\mathbf{x}_d, \mathbf{x}_{d+1}, \dots, \mathbf{x}_{f-1}, \mathbf{x}_f\}$
\mathbf{x}_{-k}	l'ensemble $\{\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n\}$ où n est l'indice final
$\tilde{\mathbf{x}}_t^{(i)}$	$i^{\text{ième}}$ particule avant rééchantillonnage
$\mathbf{x}_t^{(i)}$	$i^{\text{ième}}$ particule après rééchantillonnage
$\tilde{x}_{t,i}^{(j)}$	$j^{\text{ième}}$ particule auxiliaire associé à la $i^{\text{ième}}$ particule
$\tilde{\omega}_t^{(i)}$	poids associé à la $i^{\text{ième}}$ particule avant rééchantillonnage
$\omega_t^{(i)}$	poids associé à la $i^{\text{ième}}$ particule après rééchantillonnage
$\tilde{\omega}_{t,i}^{(j)}$	poids associé à la $j^{\text{ième}}$ particule auxiliaire
$\mathbb{E}[\cdot]$	espérance
$\text{Var}[\cdot]$	variance
$\mathcal{S}(\alpha, \beta, \gamma, \mu)$	distribution α -stable d'exposant caractéristique α , de paramètre de symétrie β , de paramètre de dispersion γ et de paramètre de position μ
$\mathcal{N}(\mu, \Sigma)$	distribution gaussienne de moyenne μ et de matrice de covariance Σ

$\mathcal{N}(\mathbf{x}; \mu, \Sigma)$	densité de probabilité gaussienne de moyenne μ et de matrice de covariance Σ évaluée au point \mathbf{x}
$\mathcal{C}(a, b)$	loi de Cauchy de paramètre d'échelle a et de paramètre de position b
$\mathcal{IG}(a, b)$	loi inverse-gamma de paramètre de forme a et de paramètre d'échelle b
$\mathcal{U}[a, b]$	loi uniforme sur l'intervalle $[a, b]$
$\mathcal{B}(a, b)$	loi beta de paramètres a et b
$\text{Multinomial}(\pi)$	loi multinomiale de paramètre π
$\mathcal{D}(a_0, \dots, a_n)$	distribution de Dirichlet de paramètres a_0, \dots, a_n
$DP(\mathbb{G}_0, \alpha)$	processus de dirichlet de distribution de base \mathbb{G}_0 et de paramètre d'échelle α
$\mathcal{NIW}(\mu, \kappa, \nu, \Lambda)$	distribution normale-inverse Wishart de paramètres μ, κ, ν et Λ
$\mathcal{W}(\nu, \Lambda)$	distribution Wishart de degré de libertés ν et de paramètre d'échelle Λ
$\stackrel{\text{loi}}{=}$	égalité en loi
$\xrightarrow{\text{loi}}$	limite en loi
$\xrightarrow{\text{p.s.}}$	Convergence presque sûre

Liste des acronymes

AR	Auto-Regressive
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CS	Compressed Sensing
DAB	Digital Audio Broadcasting
DFT	Discrete Fourier Transform
DPM	Dirichlet Process Mixture
DSL	Digital Subscriber Line
DVB-T	Digital Video Broadcasting-Terrestrial
EKF	Extended Kalman Filter
EM	Expectation-Maximization
FFT	Fast Fourier Transform
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
IS	Importance Sampling
ISI	Inter-Symbol Interference
MAI	Multiple Access Interference
MAP	Maximum A Posteriori
MCMC	Markov Chain Monte Carlo
MMSE	Minimum Mean Square Error
MSE	Mean Squared Error
OFDM	Orthogonal Frequency Division Multiplexing

pdf	probability density function
PLC	Power Line Communication
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RPM	Random Probability Measure
RS	Reed-Solomon
SMC	Sequential Monte Carlo
SMiN	Scale Mixture of Normals
SNR	Signal to Noise Ratio
TH-UWB	Time Hopping-Ultra WideBand
TVAR	Time-Varying Auto-Regressive
UKF	Unscented Kalman Filter
UMTS	Universal Mobile Transmission System
UT	Unscented Transform
UWB	Ultra WideBand
UWB-IR	Ultra WideBand-Impulse Radio

Introduction

Contexte

Le domaine du traitement statistique du signal a été toujours dominé par l'hypothèse de gaussianité pour la modélisation du bruit. Dans plusieurs situations, cette hypothèse est pertinente et peut être justifiée par le biais du théorème de la Limite Centrale. De plus, elle permet généralement d'aboutir à des solutions analytiques de manière simple et rapide. Par exemple, en théorie des communications, l'hypothèse du bruit additif blanc gaussien permet de simplifier considérablement la conception des récepteurs dans les systèmes de transmission. Cependant, dans un nombre croissant d'applications, l'hypothèse classique sur la nature gaussienne du bruit n'est plus vérifiée. C'est en particulier le cas de nombreux systèmes de communication, tels que la transmission sur le réseau électrique, les communications hautes fréquences ou bien les communications sous-marines. Dans de tels systèmes, des bruits à faible probabilité d'apparition mais de très fortes amplitudes dits de nature impulsive interviennent. Les modèles classiques basés sur les distributions gaussiennes sont insuffisants pour bien représenter de tels phénomènes et la mise au point de nouveaux modèles plus réalistes s'avère nécessaire. En effet, la modélisation d'un bruit de nature impulsive par une distribution gaussienne peut considérablement dégrader les performances des systèmes de transmission. Ceci est dû principalement à la grande variabilité de ce type de bruit. Cette variabilité n'est pas compatible avec les distributions gaussiennes mais peut être très bien représentée avec des distributions non gaussiennes à variance infinie. Dans le cadre des distributions non gaussiennes à variance infinie sont apparues les distributions α -stables. Elles font

partie d'une classe très riche de distributions de probabilité qui englobent les lois de Gauss, Cauchy et Lévy et qui autorisent l'asymétrie et les queues lourdes. Ces distributions présentent des propriétés très intéressantes qui les rendent très adaptées pour la modélisation des processus impulsifs. Des applications de ces distributions ont été proposées dans plusieurs domaines tels que finance, traitement du signal, astronomie et télécommunications. Cette modélisation est en parfait accord avec les limitations observées.

Position du problème

La compensation de l'effet du bruit constitue un problème central dans la plupart des applications en traitement du signal. Il est aussi un problème fondamental en statistique que ce soit pour des objectifs de détection ou d'estimation. Ce problème a été largement étudié dans la littérature et plusieurs solutions ont été proposées lorsque le bruit est modélisé par une loi gaussienne. Cependant, la littérature reste relativement pauvre lorsqu'il s'agit d'un bruit non gaussien de nature impulsive.

Ce travail de thèse s'inscrit dans ce contexte et vise à développer de nouvelles techniques de réduction du bruit impulsif. Le cas où le bruit est modélisé par les lois α -stables est particulièrement très intéressant. Tout d'abord, d'un point de vue théorique, ceci permet d'apporter des solutions au problème de l'inférence dans le cas des distributions α -stables. Ensuite, d'un point de vue pratique, ceci permet la circulation de l'information avec une grande fiabilité même dans les environnements les plus complexes.

Objectifs et contributions

L'objectif principal de ce travail de thèse est de concevoir de nouvelles méthodes robustes, basées sur l'inférence bayésienne, pour l'estimation conjointe des données et du bruit dans des environnements impulsifs. Nous attacherons une importance particulière :

- Aux environnements présentant des phénomènes impulsifs admettant des distributions pour lesquelles les moments d'ordre supérieur ou égal à 2 n'existent pas et en particulier les distributions α -stables.

- Aux applications pour lesquelles les données arrivent séquentiellement.

Ce problème a été peu étudié dans la littérature. En effet, lorsque les bruits sont modélisés par des distributions α -stables, les méthodes classiques ne s'appliquent plus car la densité de probabilité n'as pas d'expression analytique explicite et les moments d'ordre 2 ou supérieur d'existent pas. Pour contribuer à la résolution de ce problème, nous avons proposés une approche paramétrique et une autre non paramétrique basées sur les méthodes de Monte Carlo séquentielles :

- Approche paramétrique : dans un premier temps, nous avons proposé une méthode de réduction du bruit impulsif dans les systèmes de transmission OFDM. Le bruit impulsif est modélisé par une distribution α -stable symétrique. L'approche proposée repose sur la décomposition en mélange d'échelles de gaussiennes d'une variable aléatoire α -stable symétrique. Une telle décomposition permet de représenter une variable aléatoire α -stable symétrique comme le produit d'une variable aléatoire gaussienne et d'une variable aléatoire α -stable positive. Par conséquent, conditionnellement à la variable aléatoire α -stable positive, nous obtenons un modèle linéaire gaussien. Nous avons tenu profit de cette structure particulière du modèle pour proposer un filtre particulière Rao-Blackwellisé qui permet l'estimation conjointe des symboles OFDM émis et des paramètres du bruit impulsif.
- Approche non paramétrique : dans un deuxième temps, nous avons proposé une approche non paramétrique bayésienne qui permet l'estimation conjointe de l'état et de la densité de probabilité du bruit impulsif dans les systèmes dynamiques non linéaires. Dans le cadre de cette approche, nous ne restreignons pas notre étude au cas des bruits symétriques et unimodales ; elle peut aussi s'appliquer dans un contexte plus général où le bruit est asymétrique et multimodal. La méthodologie proposée repose sur l'introduction des mélanges de processus de Dirichlet pour modéliser le bruit impulsif. Trois modélisations différentes ont été proposées : mélange de processus de Dirichlet de distributions gaussiennes, mélange de processus de Dirichlet de distributions de Cauchy et mélange de processus de Dirichlet de distributions α -stables. Des algorithmes de Monte Carlo séquentiels basés sur des lois d'importance efficaces ont été développés.

Organisation du manuscrit

Ce manuscrit de thèse est organisé comme suit :

Le **chapitre 1** est consacré à l'introduction des modèles non gaussiens adaptés à la modélisation des phénomènes impulsifs. Une attention particulière est apportée à la famille des distributions α -stables qui fera l'objet de la première section de ce chapitre. Dans la deuxième section, nous présentons d'autres lois couramment utilisées dans la littérature pour la modélisation des bruits impulsifs dans le domaine des télécommunications.

Le **chapitre 2** est dédié à l'inférence bayésienne. Les équations d'état et d'observation du système dynamique ainsi que l'ensemble des hypothèses posées sont premièrement données. Après avoir introduit le formalisme bayésien, nous présentons quelques méthodes déterministes pour l'inférence de l'état. Ensuite, les bases des méthodes de Monte Carlo sont données. Finalement, nous dressons un état de l'art des méthodes bayésiennes permettant l'inférence dans le cas des distributions α -stables.

Le **chapitre 3** traite le problème de réduction du bruit impulsif dans les systèmes OFDM. Dans un premier temps, les systèmes de transmission OFDM sont brièvement présentés. Ensuite, nous dressons un état de l'art des méthodes existantes permettant la compensation de l'effet du bruit impulsif dans les systèmes OFDM. Après avoir introduit le modèle dynamique de l'état, nous décrivons l'approche proposée. De nombreuses simulations sont finalement présentées afin d'illustrer la validité et la robustesse de notre approche.

Le **chapitre 4** donne une revue de l'estimation non paramétrique de densité de probabilité. Des modèles non paramétriques basés sur les mélanges de processus de Dirichlet sont introduits et plusieurs algorithmes d'estimation sont fournis.

Le **chapitre 5** s'intéresse à l'estimation conjointe de l'état et de la densité du bruit d'observation dans les systèmes dynamiques non linéaires. Tout d'abord, nous introduisons trois modélisations du bruit d'observation α -stable basées sur les mélanges de processus de Dirichlet. Ensuite, des algorithmes de Monte Carlo séquentiels permettant l'estimation conjointe de l'état et de la densité du bruit d'observation ainsi que des densités d'importance efficaces sont proposés. Finalement, les différents algorithmes proposés sont testés sur un exemple illustratif et une étude comparative des trois modèles est menée.

Nous terminons ce manuscrit par une synthèse des problèmes abordés, des principales contributions apportées et des perspectives envisagées.

CHAPITRE 1

Bruit impulsif et distributions α -stables

EN modélisation statistique, l'une des lois les plus utilisées est la loi gaussienne. Ses différentes propriétés, comme la stabilité, la caractérisation parfaite par les moments d'ordre 1 et d'ordre 2 ou bien encore le théorème de la limite centrale rendent cette loi bien adaptée dans de nombreuses situations. Cependant, bien que les calculs d'inférence statistique soient simples, l'hypothèse de gaussianité s'avère trop restrictive dans plusieurs applications pour lesquelles il faut prendre en compte une plus grande variabilité des perturbations rencontrées. C'est en particulier le cas des bruits impulsifs que nous rencontrons dans de nombreux domaines, notamment les télécommunications. Ce type de bruit se retrouve par exemple dans certains phénomènes de propagation, comme le bruit subi à la réception d'un signal radar dû aux échos multiples, ou le bruit reçu par un sonar dû à la propagation sous-marine. On peut également considérer comme un bruit impulsif les perturbations électromagnétiques créés par un orage. De plus, dans certains systèmes de communication, de nombreux phénomènes rencontrés présentent un caractère impulsif. Citons par exemple les bruits rencontrés lors de transmissions sur le réseau électrique, transmission large bande par courant porteur en ligne (PLC, Power Line Communication)[ZD02, MSG05, CD89], les lignes d'abonnés numériques, les communications hautes fréquences, les systèmes de transmission mobile (UMTS, Universal Mobile Transmission System) [SAC04], les systèmes à bande ultra large (UWB, Ultra WideBand), les interférences dans les réseaux *ad hoc* [Sou92, IH96], les interférences dans des systèmes multi utilisateurs sans fils, etc.

Un bruit de nature impulsive est caractérisé par une faible probabilité d'appari-

tion mais de très fortes amplitudes. La modélisation de ce type de bruit par une loi gaussienne n'est pas pertinente à cause de sa grande variabilité. Il convient alors de modéliser ce type de bruit non pas par une gaussienne mais par une distribution qui n'admet pas forcément une variance finie.

Dans le cadre des distributions non gaussiennes à variance infinie sont apparues les distributions α -stables. Elles font partie d'une classe riche de distributions de probabilité qui englobent les lois de Gauss, de Cauchy et de Lévy et qui autorise l'asymétrie et les queues lourdes. Ces distributions ont attiré l'attention de beaucoup de chercheurs en statistique et en traitement du signal principalement pour deux raisons. La première provient du théorème de la limite centrale généralisé qui stipule que la somme de variables aléatoires indépendantes et identiquement distribuées converge vers une variable α -stable. Quand la variance des variables est finie nous obtenons le cas particulier des variables gaussiennes. L'utilisation de variables à variance infinie permet de modéliser des phénomènes présentant d'importantes variabilités. La deuxième raison est due à la propriété de stabilité qui affirme que toute combinaison linéaire de variables aléatoires α -stables est également une variable aléatoire α -stable.

Le but de ce chapitre n'est pas de faire une description exhaustive des modèles non gaussiens, il s'agit seulement d'introduire ceux qui sont particulièrement adaptés à la modélisation des phénomènes impulsifs. Une attention particulière est portée sur la famille des distributions α -stables. La première section est dédiée à la présentation de ces lois. Tout d'abord, les différentes définitions et diverses propriétés des distributions α -stables sont données. Ensuite, nous exposons quelques approches permettant l'approximation ou l'évaluation de leur densité de probabilité. L'algorithme de simulation des variables aléatoires α -stables est ensuite donné. Nous terminons l'étude des distributions α -stable par la présentation d'un état de l'art des méthodes d'estimation de ses paramètres. Dans la deuxième section, nous nous intéressons à d'autres modèles couramment utilisés dans la littérature pour la modélisation des bruits impulsifs, notamment dans le domaine des télécommunications.

1.1 Les distributions α -stables

Les distributions α -stables ont été introduites en 1925 par le mathématicien français Paul Lévy. D'autres mathématiciens ont contribué plus tard à l'étude approfondie de ces lois.

Les distributions α -stables fournissent des modèles appropriés pour de nombreux phénomènes observés dans divers domaines. En effet, des applications de ces distributions ont déjà été proposées en astronomie, en finance, en télécommunications, etc. En astronomie, Holtsmark [Hol19] a montré que la force gravitationnelle exercée par le système stellaire suit une distribution α -stable. En finance, Mandelbort [Man63] a utilisé les distributions α -stables pour modéliser les variations de prix pratiquées sur certaines bourses des valeurs. Par la suite, ce modèle a été validé par Fama [Fam65] sur le prix du marché des actions. En télécommunications, il a été par exemple prouvé que les distributions α -stables sont adéquats pour modéliser l'interférence d'accès multiple (MAI, Multiple Access Interference) dans les réseaux *ad hoc* [Sou92, IH96, PCG⁺06].

Plusieurs ouvrages ont été consacrés à l'étude des distributions α -stables. Une revue détaillée de ces distributions dans le cas univarié est disponible dans le livre de Zolotarev [Zol86]. De plus, Samorodnitsky et Taqqu [ST94] ont étudié de manière approfondie plusieurs propriétés de ces distributions dans le cas univarié comme dans le cas multivarié. Dans leur ouvrage [JW93], Janicki et Weron s'intéressent à l'étude des processus stochastiques α -stables. Une étude détaillée des applications des distributions α -stables dans le domaine du traitement du signal peut être trouvée dans le livre de Nikias et Shao [NS95]. La synthèse sur les distributions α -stables, présentée dans cette section, est principalement issue de ces ouvrages.

1.1.1 Définitions

Dans ce paragraphe, nous présentons trois définitions équivalentes d'une variable aléatoire stable :

Définition 1.1. *Une variable aléatoire X est dite stable si et seulement si pour tous*

réels positifs a et b , il existe un réel positif c et un réel d , tels que¹

$$aX_1 + bX_2 \stackrel{\text{loi}}{=} cX + d \quad (1.1)$$

où X_1 et X_2 sont deux variables aléatoires indépendantes de même loi que X . Si $d = 0$, on dit que X est strictement stable.

La loi de la somme de deux variables aléatoires indépendantes est le produit de convolution des lois de ces deux variables. La définition 1.1 fait intervenir la stabilité par produit de convolution, et justifie par conséquent la nomination "stable". Une généralisation de la définition 1.1 pour un nombre fini de copies indépendantes d'une variable stable aboutit à la définition suivante :

Définition 1.2. Une variable aléatoire X est dite stable si et seulement si pour tout entier $n \geq 2$, il existe un réel positif c_n et un réel d_n , tels que

$$X_1 + X_2 + \dots + X_n \stackrel{\text{loi}}{=} c_n X + d_n \quad (1.2)$$

où X_1, \dots, X_n sont des variables aléatoires indépendantes de même loi que X . Si $d_n = 0$, X est dite strictement stable.

La constante c_n est nécessairement de la forme $n^{\frac{1}{\alpha}}$ où α est un réel tel que $\alpha \in]0, 2]$. Le réel α est unique et ne dépend que de la loi de la variable aléatoire X , d'où la nomination " α -stable". La définition suivante repose sur le théorème de la limite centrale généralisé :

Définition 1.3. Une variable aléatoire X est dite stable si et seulement elle a un domaine d'attraction. C'est-à-dire qu'il existe une suite de variables aléatoires indépendantes et identiquement distribuées $\{X_i\}_{i>0}$, un réel strictement positif a_n et un réel b_n tels que² :

$$a_n(X_1 + \dots + X_n) - b_n \xrightarrow{\text{loi}} X \quad (1.3)$$

Cette définition montre que les lois stables sont les seules lois qui sont limites de sommes normalisées de variables aléatoires indépendantes et identiquement distribuées. Ce résultat est très utile en pratique car il permet d'approcher une suite

1. L'écriture $\stackrel{\text{loi}}{=}$ indique une égalité en loi

2. L'écriture $\xrightarrow{\text{loi}}$ indique une limite en loi

de variables aléatoires indépendantes et identiquement distribuées par une variable aléatoire stable.

1.1.2 Fonction caractéristique

Une autre alternative couramment utilisée pour définir les distributions α -stables repose sur l'expression de leur fonction caractéristique :

Définition 1.4 (Lévy-Khintchine). *Une variable aléatoire X est dite distribuée selon une loi stable si et seulement il existe quatre paramètres uniques : $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma > 0$ et $\mu \in \mathbb{R}$ tels que la fonction caractéristique de X notée par $\phi(t)$, s'écrit pour tout $t \in \mathbb{R}$:*

$$\phi(t) = \begin{cases} \exp\left(i\mu t - \gamma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{sgn}(t) \tan\left(\frac{\alpha\pi}{2}\right)\right]\right) & \text{si } \alpha \neq 1 \\ \exp\left(i\mu t - \gamma |t|^\alpha \left[1 + i\beta \operatorname{sgn}(t) \frac{2}{\pi} \log |t|\right]\right) & \text{si } \alpha = 1 \end{cases} \quad (1.4)$$

avec

$$\operatorname{sgn}(t) = \begin{cases} 1, & \text{si } t > 0 \\ 0, & \text{si } t = 0 \\ -1, & \text{si } t < 0 \end{cases} \quad (1.5)$$

Une distribution α -stable est entièrement définie par quatre paramètres :

- α est appelé exposant caractéristique de la loi α -stable. Plus α est petit, plus les queues de la distribution sont lourdes et les valeurs extrêmes sont probables ;
- μ est appelé paramètre de position. Il caractérise la moyenne de la loi lorsque α est strictement supérieur à 1 ;
- γ est appelé paramètre de dispersion. Il mesure la dispersion de la distribution autour du paramètre de position μ ;
- β est appelé paramètre de symétrie. Lorsque ce paramètre vaut 0, la distribution est symétrique par rapport au paramètre de position μ . Si de plus μ vaut 0, la distribution est dite α -stable symétrique de fonction caractéristique $\phi(t) = \exp(-\gamma|t|^\alpha)$.

Notation 1.1. *La notation $\mathcal{S}_\alpha(\mu, \beta, \gamma)$ est couramment utilisée dans la littérature pour désigner une distribution α -stable de paramètres α , μ , β et γ . Cependant, les distributions α -stables admettent plusieurs représentations possibles de leur fonction caractéristique. Ces différentes paramétrisations ont toujours constitué une source*

de confusion. Il est donc indispensable de différencier entre elles en adoptant des notations adéquates évitant la moindre ambiguïté. Dans ce manuscrit, une distribution α -stable de paramètres α , β , γ et μ sera notée $\mathcal{S}_k(\alpha, \beta, \gamma, \mu)$ où l'indice k fait référence au numéro de la paramétrisation [Nol12]. Le paragraphe suivant sera dédié à la présentation de ces différentes paramétrisations.

1.1.3 Différentes paramétrisations

La représentation précédente de la fonction caractéristique est appelée paramétrisation standard. C'est une variante de la paramétrisation forme (A) de Zolotarev [Zol86] et elle est notée $\mathcal{S}_1(\alpha, \beta, \gamma, \mu)$. Elle fournit une expression simple de la fonction caractéristique et elle est adaptée pour l'étude des propriétés algébriques des distributions α -stables. Cependant, elle a le désavantage d'être discontinue lorsque $\alpha = 1$ et $\beta \neq 0$.

Pour contourner ce problème, Zolotarev [Zol86] a proposé une autre paramétrisation, appelée forme (M). Avec cette paramétrisation, la fonction caractéristique s'exprime comme suit :

$$\phi(t) = \begin{cases} \exp\left(i\mu_0 t - \gamma^\alpha |t|^\alpha \left[1 + i\beta \operatorname{sgn}(t) \tan\left(\frac{\alpha\pi}{2}\right) (|\gamma t|^{1-\alpha} - 1)\right]\right) & \text{si } \alpha \neq 1 \\ \exp\left(i\mu_0 t - \gamma |t|^\alpha \left[1 + i\beta \operatorname{sgn}(t) \frac{2}{\pi} \log(\gamma|t|)\right]\right) & \text{si } \alpha = 1 \end{cases} \quad (1.6)$$

Dans ce cas, la distribution est notée par $\mathcal{S}_0(\alpha, \beta, \gamma, \mu_0)$. Cette paramétrisation est plus commode pour l'estimation et l'inférence statistique dans le cas des distributions α -stables car l'expression de la fonction caractéristique obtenue est continue par rapport à tous les paramètres. Les paramètres α , β et γ sont identiques à ceux de la paramétrisation standard (cf. définition 1.4), μ et μ_0 sont en revanche reliés par les relations suivantes :

$$\mu_0 = \begin{cases} \mu + \beta\gamma \tan \frac{\pi\alpha}{2} & \text{si } \alpha \neq 1 \\ \mu + \beta \frac{2}{\pi} \gamma \log \gamma & \text{si } \alpha = 1 \end{cases} \quad (1.7)$$

Une autre paramétrisation, également proposée par Zolotarev [Zol86], fournit une

expression de la fonction caractéristique sous la forme suivante :

$$\phi(t) = \begin{cases} \exp\left(i\mu t - \gamma_2^\alpha |t|^\alpha \exp\left[-i\beta_2 \text{sign}(t) \frac{\pi}{2} \min(\alpha, 2 - \alpha)\right]\right) & \text{si } \alpha \neq 1 \\ \exp\left(i\mu t - \gamma_2 |t|^\alpha \left[1 + i\beta_2 \text{sign}(t) \frac{2}{\pi} \log \gamma_2 |t|\right]\right) & \text{si } \alpha = 1 \end{cases} \quad (1.8)$$

Les paramètres α et μ sont les mêmes que pour la paramétrisation standard. Les autres paramètres satisfont les relations suivantes :

$$\beta = \cot\left(\frac{\pi\alpha}{2}\right) \tan\left(\frac{\pi\beta_2}{2} \min(\alpha, 2 - \alpha)\right) \quad (1.9)$$

$$\gamma = \gamma_2 \left[\cos\left(\frac{\pi\beta_2}{2} \min(\alpha, 2 - \alpha)\right) \right]^{\frac{1}{\alpha}} \quad (1.10)$$

1.1.4 Densité de probabilité

Dans le cas général, la densité de probabilité des distributions α -stables n'admet pas de forme analytique. Cependant, une expression sous forme d'une intégrale de cette densité peut être obtenue en appliquant la transformée de Fourier inverse à la fonction caractéristique $\phi(t)$:

$$f(x; \alpha, \beta, \gamma, \mu) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-itx) \phi(t) dt \quad (1.11)$$

Notation 1.2. La densité de probabilité et la fonction de répartition d'une distribution α -stable seront respectivement notées $f_k(x; \alpha, \beta, \gamma, \mu)$ et $F_k(x; \alpha, \beta, \gamma, \mu)$ où l'indice k fait référence au numéro de la paramétrisation.

1.1.4.1 Cas particuliers

La forme explicite de la densité de probabilité des distributions α -stables n'existe que dans les trois cas suivants :

1. Lorsque $\alpha = 2$, on retrouve une **loi normale** avec une moyenne μ et une variance $2\gamma^2$.
2. Lorsque $\alpha = 1$ et $\beta = 0$, on retrouve une **loi de Cauchy** de paramètre de position μ et de paramètre d'échelle γ .

3. Lorsque $\alpha = \frac{1}{2}$ et $\beta = 1$, on retrouve une **loi de Lévy** de paramètre de position μ et de paramètre d'échelle γ .

1.1.4.2 Propriétés de la densité de probabilité

Bien que la densité de probabilité des distributions α -stables n'admet d'expression analytique que dans les cas particuliers cités au paragraphe 1.1.4.1, on dispose de plusieurs propriétés intéressantes de cette densité.

Propriété 1.1 (Continuité). *La densité de probabilité d'une distribution α -stable est de classe C^∞ .*

Propriété 1.2 (Support). *Le support des distributions α -stables pour les différentes paramétrisations est donné par :*

$$\text{le support de } f_0(x; \alpha, \beta, \gamma, \mu) = \begin{cases} [\mu - \gamma \tan \frac{\pi\alpha}{2}, +\infty[& \text{si } \alpha < 1 \text{ et } \beta = 1 \\] - \infty, \mu + \gamma \tan \frac{\pi\alpha}{2}] & \text{si } \alpha < 1 \text{ et } \beta = -1 \\] - \infty, +\infty[& \text{sinon} \end{cases} \quad (1.12)$$

$$\text{le support de } f_1(x; \alpha, \beta, \gamma, \mu) = \begin{cases} [\mu, +\infty[& \text{si } \alpha < 1 \text{ et } \beta = 1 \\] - \infty, \mu] & \text{si } \alpha < 1 \text{ et } \beta = -1 \\] - \infty, +\infty[& \text{sinon} \end{cases} \quad (1.13)$$

Propriété 1.3 (Mode). *Les distributions α -stables sont unimodales. Dans le cas des distributions α -stables symétriques avec une valeur de α strictement supérieur à 1, le mode correspond à la moyenne.*

1.1.4.3 Interprétation des paramètres

Propriété 1.4 (réflexion). *Soient $X_1 \sim \mathcal{S}_k(\alpha, \beta, 1, 0)$ et $X_2 \sim \mathcal{S}_k(\alpha, -\beta, 1, 0)$, on a alors le résultat suivant :*

$$X_1 \stackrel{\text{loi}}{=} -X_2$$

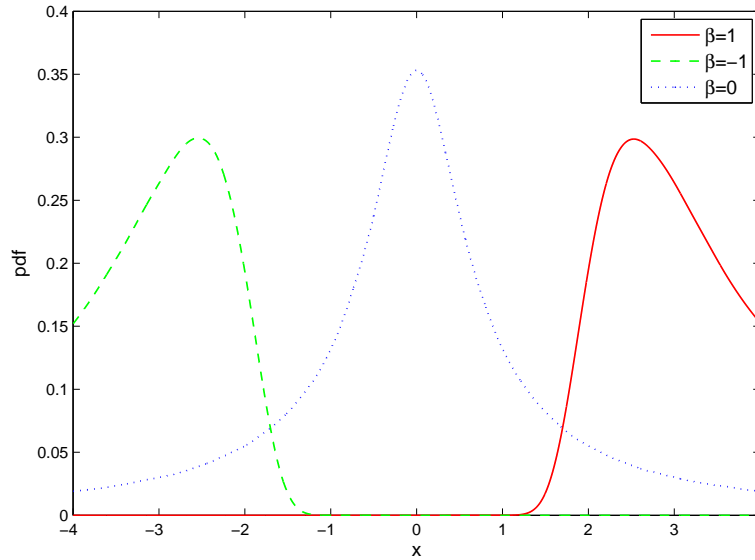
Par conséquent, la densité de probabilité et la fonction de répartition satisfont :

$$f_k(x; \alpha, \beta, 1, 0) = f_k(-x; \alpha, -\beta, 1, 0) \text{ et } F_k(x; \alpha, \beta, 1, 0) = 1 - F_k(-x; \alpha, -\beta, 1, 0)$$

A partir de cette propriété, on peut déduire que lorsque $\beta = 0$, la densité est symétrique. Lorsque $\beta > 0$ ($\beta = 1$), la distribution est asymétrique (totalement

asymétrique) à droite alors que lorsque $\beta < 0$ ($\beta = -1$) la distribution est asymétrique (totalement asymétrique) à gauche. La figure 1.1 montre le comportement de la densité de probabilité de la distribution α -stable pour différentes valeurs de β .

FIGURE 1.1 Densité de probabilité de la distribution α -stable de paramètres $\alpha = 0.8$, $\gamma = 1$, $\mu = 0$ et pour différentes valeurs de β .



Propriété 1.5 (Comportement des queues de distribution). Soit $X \sim \mathcal{S}_0(\alpha, \beta, \gamma, \mu)$ avec $\alpha < 2$ et $-1 < \beta \leq 1$, alors nous avons les deux résultats suivants :

$$\lim_{x \rightarrow \infty} P(X > x) = \gamma^\alpha C_\alpha (1 + \beta) x^{-\alpha} \quad (1.14)$$

$$\lim_{x \rightarrow \infty} f_0(x; \alpha, \beta, \gamma, \mu) = -\alpha \gamma^\alpha C_\alpha (1 + \beta) x^{-(\alpha+1)} \quad (1.15)$$

où $C_\alpha = \frac{\Gamma(\alpha)}{\pi} \sin\left(\frac{\pi\alpha}{2}\right)$. En utilisant la propriété 1.4, on obtient des résultats similaires lorsque $-1 \leq \beta < 1$:

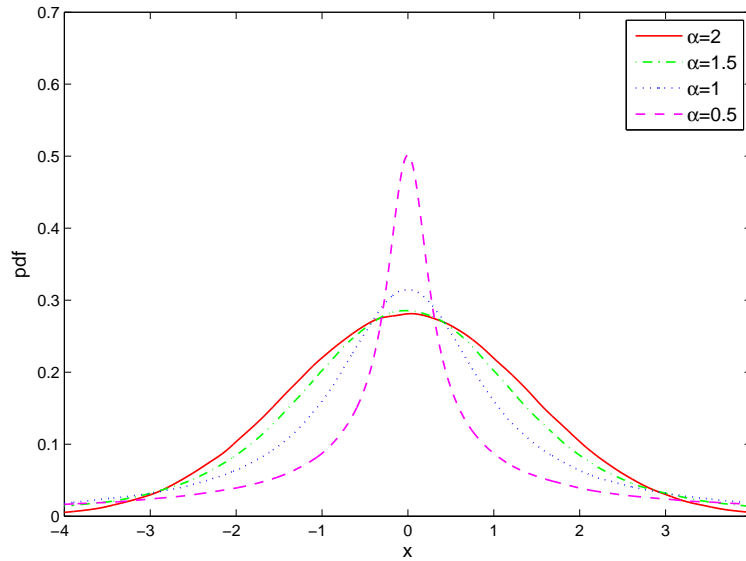
$$\lim_{x \rightarrow \infty} P(X < -x) = \gamma^\alpha C_\alpha (1 - \beta) x^{-\alpha} \quad (1.16)$$

$$\lim_{x \rightarrow \infty} f_0(-x; \alpha, \beta, \gamma, \mu) = -\alpha \gamma^\alpha C_\alpha (1 - \beta) x^{-(\alpha+1)} \quad (1.17)$$

D'après cette propriété, on peut remarquer que plus α est petit plus les queues de distribution sont lourdes, ceci s'interprète par le fait que x peut prendre de très

grandes valeurs avec une probabilité non négligeable. Pour une meilleure illustration de ces résultats, nous avons présenté, dans la figure 1.2, la densité de probabilité de la distribution α -stable pour différentes valeurs de α . La figure 1.3 permet de bien voir l'allongement des queues de distributions.

FIGURE 1.2 Densité de probabilité de la distribution α -stable de paramètres $\beta = 0$, $\gamma = 1$, $\mu = 0$ et pour différentes valeurs de α .



Propriété 1.6 (Changement de variables). Soit $Y \sim \mathcal{S}_0(\alpha, \beta, 1, 0)$, alors

$$X = \begin{cases} \gamma(Y - \beta\frac{\pi\alpha}{2}) + \mu & \text{si } \alpha \neq 1 \\ \gamma Y + \mu & \text{si } \alpha = 1 \end{cases} \quad (1.18)$$

suit une loi $\mathcal{S}_0(\alpha, \beta, \gamma, \mu)$.

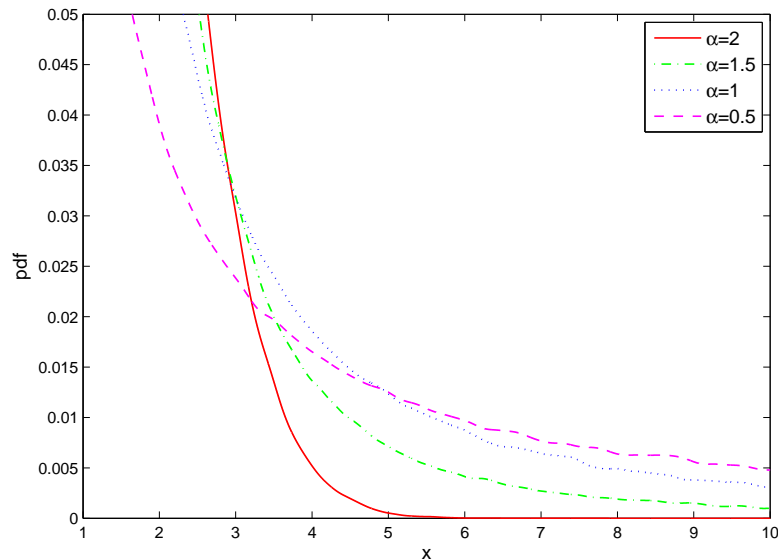
Si $Y \sim \mathcal{S}_1(\alpha, \beta, 1, 0)$, alors

$$X = \begin{cases} \gamma Y + \mu & \text{si } \alpha \neq 1 \\ \gamma Y + \mu + \frac{2}{\pi}\beta\gamma \log(\gamma) & \text{si } \alpha = 1 \end{cases} \quad (1.19)$$

suit une loi $\mathcal{S}_1(\alpha, \beta, \gamma, \mu)$.

Cette propriété confirme que γ et μ représentent respectivement le paramètre d'échelle et le paramètre de position de la distribution α -stable. Les figures 1.4 et

FIGURE 1.3 Queue de la densité de probabilité de la distribution α -stable de paramètres $\beta = 0$, $\gamma = 1$, $\mu = 0$ et pour différentes valeurs de α .



1.5 montrent l'effet de ces deux paramètres sur la densité de probabilité.

Notation 1.3. La distribution $\mathcal{S}_k(\alpha, \beta, 1, 0)$ est appelée *distribution standard* et elle sera notée $\mathcal{S}_k(\alpha, \beta)$. De même, la densité de probabilité et la fonction de répartition associées à cette distribution seront respectivement notées $f_k(x; \alpha, \beta)$ et $F_k(x; \alpha, \beta)$.

1.1.5 Propriétés des distributions α -stables

Dans cette section, sont présentées les propriétés les plus importantes des distributions α -stables.

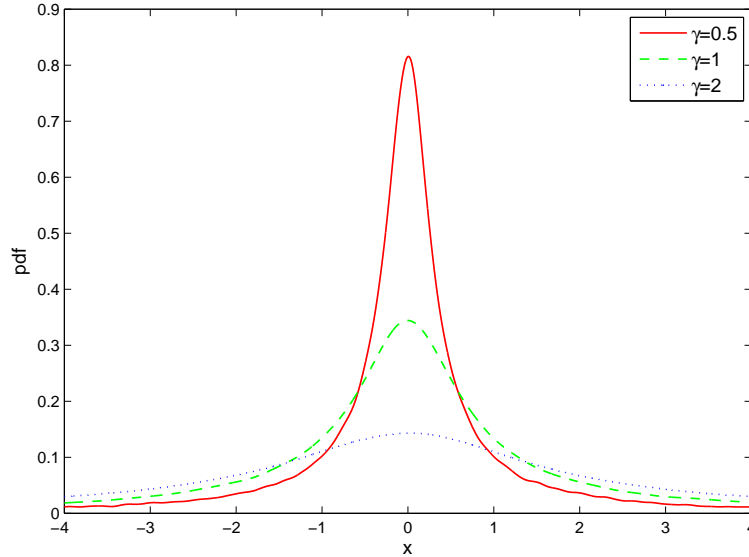
1.1.5.1 Stabilité

En utilisant la fonction caractéristique d'une distribution α -stable, on peut généraliser la propriété de stabilité donnée par la définition 1.1 comme suit :

Propriété 1.7. Si X_1, \dots, X_n sont des variables aléatoires indépendantes de loi stable avec les mêmes paramètres α et β , alors toute combinaison linéaire de la forme

$$\sum_{i=1}^n a_i X_i \tag{1.20}$$

FIGURE 1.4 Densité de probabilité de la distribution α -stable de paramètres $\alpha = 0.8$, $\beta = 0$, $\mu = 0$ et pour différentes valeurs de γ .



soit une loi stable avec les mêmes paramètres α et β .

1.1.5.2 Calcul des moments

En exploitant la propriété 1.5, on peut montrer que dans le cas des distributions α -stables, les moments d'ordre $p > \alpha$ n'existent pas pour $\alpha < 2$.

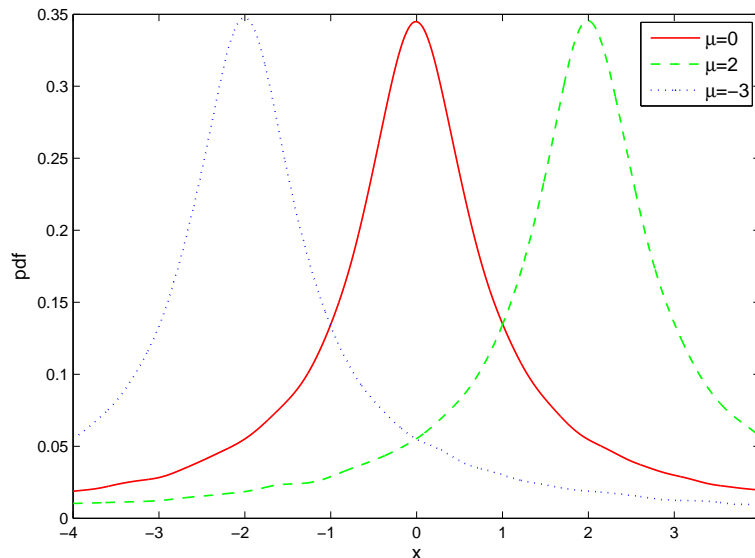
Propriété 1.8 (Moments). *Soit X une variable aléatoire de loi $\mathcal{S}_k(\alpha, \beta, \gamma, \mu)$ alors $\mathbb{E}[|X|^p] < \infty$ si et seulement si $0 < p < \alpha$.*

Les moments d'ordre supérieur ou égal à 2 des lois α -stables sont infinis, sauf dans le cas gaussien ($\alpha = 2$). Et lorsque α est strictement inférieur à 1, c'est la moyenne qui devient infinie. La démonstration de cette propriété est détaillée dans [NS95]. Lorsque α est supérieur ou égal à 1, la valeur de la moyenne est donnée par la propriété suivante :

Propriété 1.9 (Moyenne). *Soit X une variable aléatoire de loi $\mathcal{S}_0(\alpha, \beta, \gamma, \mu)$ avec $\alpha \geq 1$, alors*

$$\mathbb{E}[X] = \mu - \beta\gamma \tan \frac{\pi\alpha}{2} \quad (1.21)$$

FIGURE 1.5 Densité de probabilité de la distribution α -stable de paramètres $\alpha = 0.8$, $\beta = 0$, $\gamma = 1$ et pour différentes valeurs de μ .



Si X est de loi $\mathcal{S}_1(\alpha, \beta, \gamma, \mu)$ avec $\alpha \geq 1$, alors

$$\mathbb{E}[X] = \mu \quad (1.22)$$

1.1.5.3 Propriétés algébriques

Propriété 1.10. Soient X une variable aléatoire de loi $\mathcal{S}_0(\alpha, \beta, \gamma, \mu)$ et deux réels $a \neq 0$ et b , alors

$$aX + b \sim \mathcal{S}_0(\alpha, \text{sgn}(a)\beta, |a|\gamma, a\mu + b) \quad (1.23)$$

Si X est de loi $\mathcal{S}_1(\alpha, \beta, \gamma, \mu)$, alors

$$aX + b \sim \begin{cases} \mathcal{S}_1(\alpha, \text{sgn}(a)\beta, |a|\gamma, a\mu + b) & \text{si } \alpha \neq 1 \\ \mathcal{S}_1(\alpha, \text{sgn}(a)\beta, |a|\gamma, a(\mu - \beta\gamma\frac{2}{\pi}\log|a|) + b) & \text{si } \alpha = 1 \end{cases} \quad (1.24)$$

Propriété 1.11. Soient $X_1 \sim \mathcal{S}_0(\alpha, \beta_1, \gamma_1, \mu_1)$ et $X_2 \sim \mathcal{S}_0(\alpha, \beta_2, \gamma_2, \mu_2)$ deux va-

riables aléatoires indépendantes, alors : $X_1 + X_2 \sim S_0(\alpha, \beta, \gamma, \mu)$ avec :

$$\begin{aligned}\beta &= \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha} \\ \gamma &= \sqrt[\alpha]{\gamma_1^\alpha + \gamma_2^\alpha} \\ \mu &= \begin{cases} \mu_1 + \mu_2 + \tan \frac{\pi\alpha}{2} (\beta\gamma - \beta_1\gamma_1 - \beta_2\gamma_2) & \text{si } \alpha \neq 1 \\ \mu_1 + \mu_2 + \frac{2}{\pi} (\beta\gamma \log \gamma - \beta_1\gamma_1 \log \gamma_1 - \beta_2\gamma_2 \log \gamma_2) & \text{si } \alpha = 1 \end{cases}\end{aligned}\quad (1.25)$$

Si $X_1 \sim S_1(\alpha, \beta_1, \gamma_1, \mu_1)$ et $X_2 \sim S_1(\alpha, \beta_2, \gamma_2, \mu_2)$, alors $X_1 + X_2 \sim S_1(\alpha, \beta, \gamma, \mu)$ avec :

$$\begin{aligned}\beta &= \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha} \\ \gamma &= \sqrt[\alpha]{\gamma_1^\alpha + \gamma_2^\alpha} \\ \mu &= \mu_1 + \mu_2\end{aligned}\quad (1.26)$$

Le théorème suivant montre qu'une variable aléatoire α -stable symétrique peut s'écrire comme le produit de deux variables aléatoires indépendantes, l'une suit une loi α -stable symétrique et l'autre suit une loi α -stable positive [ST94, Fel68].

Théorème 1.1 (Théorème de mélange d'échelles). *Soient X une variable aléatoire de loi $\mathcal{S}_1(\alpha', 0, \gamma, 0)$ et α un réel tel que $0 < \alpha < \alpha'$. Soit Y une variable aléatoire de loi $\mathcal{S}_1(\frac{\alpha}{\alpha'}, 1, \cos(\frac{\pi\alpha}{2\alpha'})^{\frac{\alpha'}{\alpha}}, 0)$. Si X et Y sont indépendantes alors :*

$$Z = XY^{\frac{1}{\alpha'}} \sim \mathcal{S}_1(\alpha, 0, \gamma, 0) \quad (1.27)$$

La démonstration de ce théorème est donnée dans [ST94, Fel68].

Un cas particulier du théorème 1.1 est obtenu en posant $\alpha' = 2$. Ceci conduit à l'énoncé suivant :

Corollaire 1.1. *Soit X une variable aléatoire gaussienne tel que $X \sim \mathcal{N}(0, 2\gamma^2)$ et soit Y une variable aléatoire positive de loi $\mathcal{S}_1(\frac{\alpha}{2}, 1, \cos(\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}, 0)$ et indépendante de X . Alors*

$$Z = XY^{\frac{1}{2}} \sim \mathcal{S}_1(\alpha, 0, \gamma, 0) \quad (1.28)$$

Ceci montre qu'une variable aléatoire α -stable symétrique peut être représentée en fonction d'une variable aléatoire gaussienne et d'une variable aléatoire α -stable

positive. Ce résultat est très intéressant vu qu'il présente les distributions α -stables comme des distributions gaussiennes conditionnelles.

1.1.6 Représentation analytique de la densité de probabilité des distributions α -stables

A l'exception de la loi de Gauss, la loi de Cauchy et la loi de Lévy, les distributions α -stables n'ont pas d'expression analytique pour leur densité de probabilité. Cela a constitué la principale cause du peu d'intérêt qu'ont suscité ces lois en traitement statistique du signal. L'accroissement incessant des performances des calculateurs a permis d'envisager l'utilisation de méthodes d'approximation numériques pour l'évaluation directe des densités α -stables. Dans ce cadre, Nolan [Nol97, Nol99] a prouvé l'existence d'une intégrale exacte avec des bornes d'intégrations finies équivalente à celle de l'équation (1.11) et il a mis à profit les techniques d'intégration numérique pour l'évaluer. Cette approche, bien que stable numériquement, reste très coûteuse en temps de calcul. Dans ce paragraphe, nous présentons d'abord la méthode proposée par Nolan [Nol97, Nol99]. Ensuite, nous exposons d'autres solutions permettant d'approcher la densité de probabilité des distributions α -stables.

1.1.6.1 Evaluation numérique de la densité de probabilité des distributions α -stables

Nolan [Nol97, Nol99] a proposé un algorithme qui permet l'évaluation de la densité de probabilité d'une distribution α -stable standard $\mathcal{S}_0(\alpha, \beta)$ par intégration numérique directe. Cet algorithme repose sur une expression plus simplifiée de l'intégrale (1.11) avec des bornes d'intégration finies. Pour se ramener à une telle expression, Nolan définit les quantités suivantes [Nol97] :

$$\zeta = \begin{cases} -\beta \tan \frac{\pi\alpha}{2} & \text{si } \alpha \neq 1 \\ 0 & \text{si } \alpha = 1 \end{cases}$$

$$\theta_0 = \begin{cases} \frac{1}{\alpha} \arctan \left(\beta \tan \frac{\pi\alpha}{2} \right) & \text{si } \alpha \neq 1 \\ \frac{\pi}{2} & \text{si } \alpha = 1 \end{cases}$$

$$c_1 = \begin{cases} \frac{1}{\pi}(\frac{\pi}{2} - \theta_0) & \text{si } \alpha < 1 \\ 0 & \text{si } \alpha = 1 \\ 1 & \text{si } \alpha > 1 \end{cases}$$

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha \theta_0)^{\frac{1}{\alpha-1}} \left(\frac{\cos \theta}{\sin \alpha(\theta_0 + \theta)} \right)^{\frac{\alpha}{\alpha-1}} \frac{\cos(\alpha \theta_0 + (\alpha-1)\theta)}{\cos \theta} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} \frac{\frac{\pi}{2} + \beta \theta}{\cos \theta} \exp\left(\frac{1}{\beta} \left(\frac{\pi}{2} + \beta \theta\right) \tan \theta\right) & \text{si } \alpha = 1, \beta \neq 0 \end{cases}$$

La densité de probabilité de la distribution α -stable $\mathcal{S}_0(\alpha, \beta)$ est alors donnée par le théorème suivant :

Théorème 1.2. *Soit $X \sim \mathcal{S}_0(\alpha, \beta)$ alors la densité de probabilité de X est donnée par :*

$$f_0(x; \alpha, \beta) = \begin{cases} \frac{\alpha(x - \zeta)^{\frac{1}{\alpha-1}}}{\pi |\alpha - 1|} \int_{-\theta_0}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp\left(-\frac{\alpha}{\alpha-1}(x - \zeta)V(\theta; \alpha, \beta)\right) d\theta & \text{si } \alpha \neq 1, x > \zeta \\ \frac{\Gamma(1 + \frac{1}{\alpha}) \cos \theta_0}{\pi^{2\alpha} \sqrt{1 + x^2}} & \text{si } \alpha \neq 1, x = \zeta \\ f_0(-x; \alpha, -\beta) & \text{si } \alpha \neq 1, x < \zeta \\ \frac{1}{2|\beta|} \exp\left(\frac{-\pi x}{2\beta}\right) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp\left(-\exp\left(\frac{-\pi x}{2\beta}\right)V(\theta; \alpha, \beta)\right) d\theta & \text{si } \alpha = 1, \beta \neq 0 \\ \frac{1}{\pi(1 + x^2)} & \text{si } \alpha = 1, \beta = 0 \end{cases} \quad (1.29)$$

La densité de probabilité d'une distribution α -stable $\mathcal{S}_1(\alpha, \beta, \gamma, \mu)$ est alors calculée en utilisant les formules suivantes [Nol99] :

$$f_1(x; \alpha, \beta, \gamma, \mu) = \begin{cases} \frac{1}{\gamma} f_0\left(\frac{x - (\gamma\beta \tan \frac{\alpha\pi}{2} + \mu)}{\gamma}; \alpha, \beta\right) & \text{si } \alpha \neq 1 \\ \frac{1}{\gamma} f_0\left(\frac{x - (\frac{2}{\pi}\gamma\beta \log \gamma + \mu)}{\gamma}; \alpha, \beta\right) & \text{si } \alpha = 1 \end{cases}$$

La principale difficulté de l'évaluation de la densité (1.29) réside dans l'approximation numérique de l'intégrale suivant :

$$\int_{-\theta_0}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp\left(-\frac{\alpha}{\alpha-1}(x - \zeta)V(\theta; \alpha, \beta)\right) d\theta \quad (1.30)$$

Pour contourner cette difficulté, Nolan propose une autre expression plus facilement

calculable pour le cas où $x > \zeta$ ³ :

$$f_0(x; \alpha, \beta) = c_2(x; \alpha, \beta) \int_{-\theta_0}^{\frac{\pi}{2}} g(\theta; x, \alpha, \beta) \exp(-g(\theta; x, \alpha, \beta)) d\theta \quad (1.31)$$

où

$$c_2(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\pi|\alpha-1|(x-\zeta)} & \text{si } \alpha \neq 1 \\ \frac{1}{2|\beta|} & \text{si } \alpha = 1 \end{cases} \quad (1.32)$$

et

$$g(\theta; x, \alpha, \beta) = \begin{cases} (x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta) & \text{si } \alpha \neq 1 \\ \exp(-\frac{\pi x}{2\beta}) V(\theta; \alpha, \beta) & \text{si } \alpha = 1 \end{cases} \quad (1.33)$$

Des problèmes d'instabilité numérique apparaissent lorsque α est proche de 0 ou de 1. En effet, dans le cas où α est proche de 0, le problème est lié à la présence d'un pic dans la densité de probabilités. Dans le cas où α est proche de 1, la fonction $V(\theta; \alpha, \beta)$ varie rapidement, rendant très difficile l'approximation de l'intégrale.

1.1.6.2 Développement en séries entières

Les densités de probabilité des distributions α -stables admettent des développements en séries entières. Dans le cas d'une distribution α -stable standard, Bergström [Ber52] a démontré que la densité de probabilité est développable en série entière convergente $\forall x > 0$:

$$f_2(x; \alpha, \beta) = \begin{cases} \frac{1}{\pi} \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) \sin\left(\frac{k\alpha\pi}{2} K(\alpha)\right) x^{-\alpha k - 1} & \text{si } 0 < \alpha < 1 \\ \frac{1}{\pi} \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{k!} \Gamma(k/\alpha + 1) \sin\left(\frac{k\pi}{2} K(\alpha)\right) x^{k-1} & \text{si } 1 < \alpha < 2 \end{cases} \quad (1.34)$$

où $K(\alpha) = \beta + \min(\alpha, 2 - \alpha)$.

Bergström [Ber52] a également proposé des développements asymptotiques de la densité de probabilité d'une distribution α -stable standard. Le développement asymp-

3. Lorsque $x = \zeta$ la densité admet une expression analytique, et lorsque $x < \zeta$ son expression est directement déduite du cas $x > \zeta$.

totique quand $x \rightarrow +\infty$ est défini pour $1 < \alpha < 2$ par la relation suivante :

$$f_2(x; \alpha, \beta) = \frac{1}{\pi} \sum_{k=1}^n \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) \sin\left(\frac{k\alpha\pi}{2} K(\alpha)\right) x^{-\alpha k - 1} + O(x^{-\alpha(n+1)-1}) \quad (1.35)$$

et celui quand $x \rightarrow 0^+$ est donné pour $0 < \alpha < 1$ comme suit :

$$f_2(x; \alpha, \beta) = \frac{1}{\pi} \sum_{k=1}^n \frac{(-1)^{k-1}}{k!} \Gamma(k/\alpha + 1) \sin\left(\frac{k\pi}{2} K(\alpha)\right) x^{k-1} + O(x^{n+1}) \quad (1.36)$$

1.1.6.3 Approximation par un mélange fini de gaussiennes

La méthode d'approximation des densités de probabilité des distributions α -stables symétriques a été introduite par Kuruoglu *et al* [KMF98]. La méthode proposée est essentiellement basée sur le résultat donné par le corollaire 1.1 qui représente une variable aléatoire α -stable symétrique comme le produit d'une variable aléatoire gaussienne et d'une variable aléatoire α -stable positive.

Soit Z une variable aléatoire α -stable symétrique de loi $\mathcal{S}_1(\alpha, 0, \gamma, 0)$. D'après le corollaire 1.1, Z peut s'écrire :

$$Z = XY^{\frac{1}{2}} \quad (1.37)$$

où $X \sim \mathcal{N}(0, 2\gamma^2)$ et $Y \sim \mathcal{S}_1(\frac{\alpha}{2}, 1, \cos(\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}, 0)$. En définissant la variable aléatoire $V = Y^{\frac{1}{2}}$, la densité de probabilité de Z peut alors se décomposer de la façon suivante :

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{Z|V}(z|v) f_V(v) J(z, v) dv \quad (1.38)$$

où $f_V(\cdot)$ et $f_{Z|V}(\cdot)$ représentent respectivement la densité de probabilité de V et la densité de probabilité conditionnelle de Z sachant V , et $J(z, v)$ est le Jacobien de Z par rapport à V . Or X suit une loi normale, par conséquent pour une réalisation $V = v$ donnée, $f_{Z|V}(z|v)$ est conditionnellement distribuée selon une loi normale. L'équation (1.38) peut alors se réécrire comme suit [AM74] :

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{z^2}{2v^2}\right) f_V(v) v^{-1} dv \quad (1.39)$$

Les densités de probabilité de Z pouvant s'écrire sous la forme de l'équation (1.39) sont appelées mélanges d'échelles de gaussiennes (SMiN, Scale Mixture of Normals)

et la fonction $f_V(v)$ est dite densité de mélange [AM74]. La génération de N échantillons à partir de l'équation (1.39) permet d'obtenir une approximation de la densité de probabilité d'une distribution α -stable symétrique par un mélange fini de gaussiennes :

$$f_1(z; \alpha, 0, \gamma, \mu) \approx \frac{\sum_{i=1}^N \frac{1}{\sqrt{2\pi v_i}} \exp\left(-\frac{(z-\mu)^2}{2v_i}\right) f_V(v_i)}{\sum_{i=1}^N f_V(v_i)} \quad (1.40)$$

Pour une bonne approximation, il faut absolument prendre N assez grand ce qui va aboutir à un algorithme très coûteux en temps de calcul. Afin de réduire cette complexité, Kuruoglu *et al* [KMF98] proposent d'utiliser un nombre réduit d'échantillons et puis d'améliorer l'approximation obtenue à l'aide de l'algorithme espérance-maximisation (EM, Expectation-Maximization). La méthode proposée pour l'approximation de la densité de probabilité des distributions α -stables symétriques par un mélange fini de gaussiennes est finalement donnée par l'algorithme 1.

Algorithme 1 Approximation de la pdf d'une distribution α -stable symétrique par un mélange fini de gaussiennes

- Calculer la fonction caractéristique $\phi_Y(t)$ de la variable α -stable positive Y de loi $\mathcal{S}_1(\frac{\alpha}{2}, 1, \cos((\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}), 0)$ en utilisant l'équation (1.4).
- Générer un ensemble de N points équidistants $\{v_i\}_{i=0}^N$.

pour $i = 0$ à N **faire**

- Evaluer la densité de probabilité de la distribution α -stable positive $f_Y(v_i)$ en appliquant la transformée de Fourier rapide inverse à la fonction caractéristique $\phi_Y(t)$:

$$f_Y(v_i) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-itv_i) \phi_Y(t) dt$$

- Calculer la densité de mélange $f_V(v_i)$:

$$f_V(v_i) = 2v_i f_Y(v_i^2)$$

fin pour

- Approcher la densité de probabilité de la densité de probabilité α -stable symétrique en utilisant l'équation (1.40).
 - Faire appel à l'algorithme EM en prenant les échantillons de la densité de mélange comme valeurs initiales pour les coefficients des composantes gaussiennes.
-

1.1.7 Algorithme de simulation

Chambers *et al* [CMS76] ont développé un algorithme qui permet de tirer des échantillons selon les distributions α -stables. L'algorithme proposé permet de générer des lois $S_1(\alpha, \beta)$. Pour obtenir des loi $S_1(\alpha, \beta, \gamma, \mu)$, il suffit d'utiliser la propriété 1.6 du changement de variable. D'une manière similaire, les lois $S_0(\alpha, \beta, \gamma, \mu)$ peuvent être obtenues en exploitant l'équation (1.7).

L'algorithme 2 décrit les différentes étapes de la méthode de Chambers *et al* [CMS76] pour générer des variables aléatoires α -stables $S_1(\alpha, \beta)$.

Algorithme 2 Simulation de variables aléatoires α -stables $S_1(\alpha, \beta)$

- Générer deux variables aléatoires V et W respectivement distribuées selon une loi uniforme sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$ et une loi exponentielle de paramètre 1 :

$$\begin{aligned} V &\sim \mathcal{U}\left]-\frac{\pi}{2}, \frac{\pi}{2}[\right. \\ W &\sim \mathcal{E}(1) \end{aligned}$$

- Calculer les quantités :

$$S_{\alpha, \beta} = \left(1 + \beta^2 \left(\tan \frac{\alpha\pi}{2}\right)^2\right)^{\frac{1}{2\alpha}}$$

$$B_{\alpha, \beta} = \frac{\arctan(\beta \tan \frac{\alpha\pi}{2})}{\alpha}$$

- Générer une variable X qui suit une distribution α -stable $S_1(\alpha, \beta)$:

$$X = \begin{cases} S_{\alpha, \beta} \times \frac{\sin(\alpha(V + B_{\alpha, \beta}))}{\sqrt[\alpha]{\cos V}} \times \left[\frac{\cos(V - \alpha(V + B_{\alpha, \beta}))}{W} \right]^{\frac{1-\alpha}{\alpha}} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta V \right) \tan V - \beta \log \left(\frac{\frac{\pi}{2} W \cos V}{\frac{\pi}{2} + \beta V} \right) \right] & \text{si } \alpha = 1 \end{cases}$$

1.1.8 Estimation des paramètres des distributions α -stables

Dans la littérature, de nombreux travaux ont été menés afin d'estimer les paramètres des distributions α -stables. La plupart des méthodes proposées traite le

problème dans le cas particulier des distributions α -stables symétriques, citons par exemple [FR68, FR71, BY90, MN95, BN99]. Pour les méthodes qui abordent le problème dans le cas général, nous pouvons citer [Pre72, DuM73, Kou80, Kou81, McC86, KW95, MFM00, Kur01]. Dans ce dernier cas, les méthodes proposées aboutissent à des estimateurs avec de faibles propriétés de convergence ou nécessitant des temps de calcul élevés. Parmi les travaux pionniers dans ce contexte, nous pouvons citer celui de Fama et Roll [FR68, FR71]. Dans cette contribution, les auteurs proposent une approche basée sur l'utilisation des quantiles pour l'estimation des paramètres α et γ dans le cas des distributions α -stables symétriques et lorsque $1 \leq \alpha \leq 2$. Cette méthode est très simple à implémenter. Cependant les estimateurs obtenus sont asymptotiquement biaisés et les conditions imposées aux paramètres sont trop restrictives. Cette méthode a été ensuite généralisée par McCulloch [McC86] pour l'estimation de tous les paramètres de la distribution α -stable. D'autres approches basées sur l'utilisation de la fonction caractéristique empirique ont été également proposées. L'idée d'utiliser la fonction caractéristique empirique pour l'estimation des paramètres des distributions α -stables a été initialement introduite par Press avec sa méthode des moments [Pre72], mais ensuite, plusieurs autres modifications ont été proposées telles que : la méthode de Paulson *et al* [PHL75], la méthode de régression [Kou80, Kou81] et récemment la méthode de Koyon et Williams [KW95]. La méthode du maximum de vraisemblance est généralement très efficace pour résoudre ce type de problème et permet d'obtenir les meilleurs estimateurs possibles. Néanmoins, sa mise en place dans le cas des distributions α -stables est entravée par l'absence d'une expression analytique de leur densité de probabilité et nécessite le recours à des méthodes additionnelles pour l'approximation de la densité de probabilité. Plusieurs contributions ont été proposées dans cette direction [DuM73, BY90]. DuMouchel [DuM73] propose de maximiser une approximation multinomiale de la fonction de vraisemblance dans le cas où $\beta = \mu = 0$ mais cette méthode est très coûteuse en temps de calcul. L'application directe de la méthode du maximum de vraisemblance a été implémentée par Brorsen et Yang [BY90] en considérant une représentation intégrale de la densité de probabilité des distributions α -stables symétriques. Les résultats obtenus sont très satisfaisants. Cependant, la complexité de calcul reste énorme vu la nécessité d'évaluer numériquement l'intégrale à chaque itération. Des approches bayésiennes ont été également proposées [Buc95, God99, God00, Tsi99, Lom07]. Ces approches seront détaillées dans le chapitre suivant.

1.2 Autres modèles du bruit impulsif

Dans cette section sont présentés d'autres modèles fréquemment utilisés dans la littérature pour modéliser les bruits impulsifs.

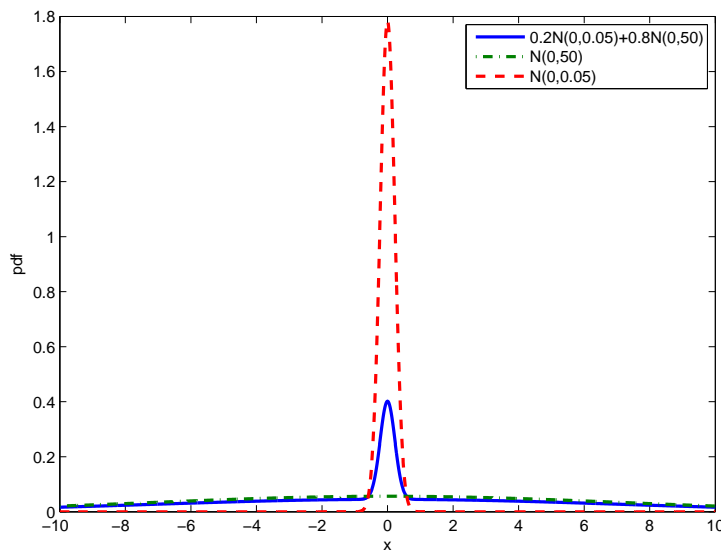
1.2.1 Modèle de mélange de gaussiennes

Le modèle de mélange de gaussiennes correspond à une densité de probabilité sous la forme d'une somme pondérée de gaussiennes :

$$f(x) = \sum_{k=1}^N \pi_k \mathcal{N}(x|\mu_k, \Sigma_k) \quad (1.41)$$

où N est le nombre de gaussiennes du mélange, $\mathcal{N}(x|\mu_k, \Sigma_k)$ est une gaussienne de moyenne μ_k et de matrice de covariance Σ_k et π_k est le poids associé à la $k^{\text{ième}}$ gaussienne représentant sa probabilité *a priori*.

FIGURE 1.6 Densité de probabilité d'un mélange de deux gaussiennes $0.2\mathcal{N}(0, 0.05) + 0.8\mathcal{N}(0, 50)$.



Ce modèle a été utilisé dans le domaine des télécommunications pour modéliser les MAI qui apparaissent dans les systèmes à bande ultra large à saut temporel (TH-UWB, Time Hopping-Ultra WideBand) [BY09, HB08] et les systèmes

radio impulsions à bande ultra large (UWB-IR, Ultra WideBand-Impulse Radio) [ECD08, ML07, CD05].

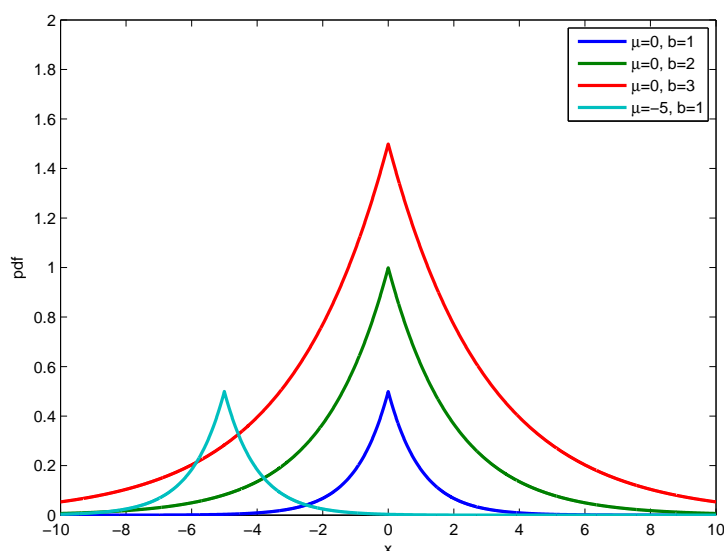
1.2.2 Loi de Laplace

La loi de Laplace, appelée aussi loi double exponentielle, est la distribution de la différence de deux variables aléatoires exponentielles indépendantes. Sa densité de probabilité est donnée par :

$$f(x|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right) \quad (1.42)$$

où $\mu \in \mathbb{R}$ est le paramètre de position et $b > 0$ est le paramètre d'échelle.

FIGURE 1.7 Densité de probabilité de la distribution de Laplace pour différentes valeurs de μ et b .



Cette loi a été souvent utilisée dans la littérature pour modéliser les bruits impulsifs [SB10, TC94]. Dans le domaine des télécommunications, elle a été utilisée pour modéliser les MAI dans les systèmes TH-UWB [BY09, HB08, BH08, BH06].

1.2.3 Loi gaussienne généralisée

La loi gaussienne généralisée est une généralisation des lois de Gauss et de Laplace. Elle est décrite par la densité de probabilité suivante :

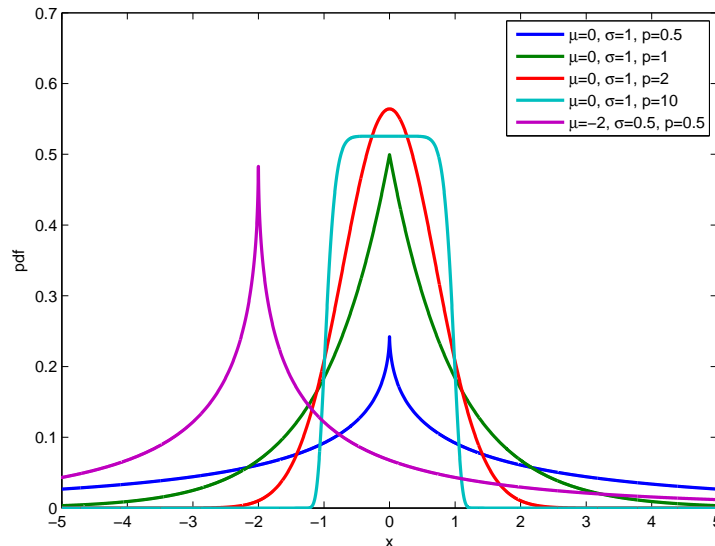
$$f_p(x) = \frac{p}{2\sigma\Gamma(\frac{1}{p})} \exp\left(-\left|\frac{x}{\sigma}\right|^p\right) \quad (1.43)$$

où $\Gamma(\cdot)$ est la fonction gamma donnée par :

$$\Gamma(a) = \int_0^{\infty} t^{a-1} \exp(-t) dt, a > 0 \quad (1.44)$$

Cette loi est définie par trois paramètres : $\mu \in \mathbb{R}$ est le paramètre de position, $\sigma > 0$ est le paramètre d'échelle et $p > 0$ est le paramètre de forme qui caractérise l'impulsivité. Notons que pour $p = 2$, on retrouve la loi gaussienne, pour $p = 1$, on retrouve la loi de Laplace et lorsque $p \rightarrow \infty$, on retrouve la loi uniforme sur $[\mu - \sigma, \mu + \sigma]$.

FIGURE 1.8 Densité de probabilité de la distribution gaussienne généralisée pour différentes valeurs de μ , σ et p .



Cette loi a été largement utilisée pour modéliser les bruits de nature impulsive, citons par exemple la modélisation de bruit atmosphérique [MT72] ainsi que la mo-

délisation des MAI dans les systèmes TH-UWB [BY09, BSF08], les systèmes radio impulsions [Fio06], les réseaux *ad hoc* [EGCR07] et dans les systèmes multi utilisateurs [ZB02]. Malgré le succès relatif de la famille des lois gaussiennes généralisées, elle présente quelques limitations. En effet, lorsque $p < 1$, la forme pointue de la distribution est non appropriée à certaines situations pratiques du bruit. On peut également noter que les queues de distributions gaussiennes généralisées se distinguent par une décroissance exponentielle contrairement au comportement algébrique⁴ des queues de distributions des processus impulsifs rencontrés dans plusieurs applications.

1.2.4 Modèle de Middleton classe A

Le modèle de Middleton est le fruit des travaux menés par Middleton pour développer un modèle statistico-physique des interférences électromagnétiques. Ce modèle se décline en trois classes :

- **la classe A** représente les bruits caractérisés par des bandes de largeur inférieure ou comparable à celle du récepteur ;
- **la classe B** correspond aux bruits ayant des bandes de fréquences plus larges que celle du récepteur ;
- **la classe C** est définie comme la somme des deux classes A et B.

Dans la littérature, un intérêt particulier est accordé à la classe A vu ses qualités pour la modélisation des bruits impulsifs. Sa densité de probabilité est donnée par :

$$f_A(x) = \sum_{m=0}^{\infty} K_m \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left(-\frac{x^2}{2\sigma_m^2}\right) \quad (1.45)$$

avec

$$K_m = \exp(-A) \frac{A^m}{m!},$$

$$\sigma_m = \frac{\frac{m}{A} + \Gamma}{1 + \Gamma},$$

et

$$\Gamma = \frac{\sigma_G^2}{\Omega_{2A}}.$$

4. On dit qu'une variable aléatoire X a une queue algébrique s'il existe deux réels c et $\alpha > 0$ tels que $p(|X| > x) \sim cx^{-\alpha}$, quand $x \rightarrow \infty$

Cette densité peut être vue comme étant une somme pondérée infinie de gaussiennes centrées, de variance σ_m^2 et dont le coefficient de pondération K_m suit une loi de poisson. Le bruit de Middleton classe A est entièrement définie à partir de l'ensemble des trois paramètres $P_3 = (A, \Gamma, \Omega_{2A})$:

- Le paramètre A est appelé indice de recouvrement (overlap index). Il représente le nombre moyen d'apparition des sources de perturbation au niveau du récepteur, multiplié par la durée moyenne d'une émission typique du bruit. Une faible valeur de A implique un bruit fortement impulsif (plus A est faible, plus le nombre des sources de perturbation est petit). Cependant, plus A est grand, plus on s'approche d'un bruit gaussien ;
- Le paramètre Γ est appelé facteur gaussien. Il est défini comme étant le rapport entre la puissance de la composante gaussienne du bruit σ_G^2 et la puissance de la composante impulsive Ω_{2A} ;
- Le paramètre Ω_{2A} représente la puissance de la composante impulsive du bruit.

En outre de son utilisation comme modèle pour les interférences électromagnétiques, le modèle de bruit de Middleton Classe A est couramment adopté dans la littérature pour modéliser d'autres types de bruits impulsifs dans plusieurs domaines, en particulier les télécommunications [KOI12, RR10, WST09]. Ce modèle est utilisé dans la majorité dans travaux de modélisation du bruit impulsif dans les communications PLC [NUDM05, MUKM02, HV01, YTM89]. Il est également utilisé pour modéliser les interférences multi-utilisateurs dans les systèmes TH-UWB [HB08, DK06] ainsi que le interférences de radiofréquences [NGS⁺08].

1.3 Conclusion

Dans ce chapitre, nous avons présenté des modèles fréquemment utilisés dans la littérature pour la modélisation des phénomènes impulsifs. La première section de ce chapitre a été dédiée à l'introduction des distributions α -stables. D'autres modèles de bruit impulsifs ont été également présentés dans la deuxième section. Les distributions α -stables présentent plusieurs propriétés intéressantes qui les rendent adaptées pour la modélisation des phénomènes impulsifs. De plus, leur choix a été validé dans plusieurs domaines en particulier dans le domaine des télécommunications. Dans cette thèse, notre travail de recherche s'est alors axé sur la réduction

du bruit dans les environnements présentant des phénomènes impulsifs modélisés par les distributions α -stables. Les méthodes proposées sont basées sur l'inférence bayésienne qui fera l'objet du chapitre suivant.

CHAPITRE 2

Inférence bayésienne

Ce chapitre fixe le cadre d'étude choisi dans cette thèse et présente les différents algorithmes utilisés. Dans le paragraphe 2.1, nous introduisons le système dynamique, ensuite le formalisme bayésien est présenté. Le paragraphe 2.2 introduit quelques algorithmes déterministes permettant de réaliser l'inférence, notamment le filtre de Kalman ainsi que ses différentes extensions : filtre de Kalman étendu et filtre de kalman sans parfum. Dans le paragraphe 2.3, les méthodes de Monte Carlo sont présentées. Un état de l'art sur les méthodes existantes traitant le problème de l'inférence bayésienne dans le cas des distributions α -stable est finalement exposé dans le paragraphe 2.4.

2.1 Formulation du problème

2.1.1 Système dynamique

Nous restreignons notre étude aux systèmes dynamiques à temps discret, c'est-à-dire que le modèle du système est décrit par des équations aux différences.

La séquence des états cachés $\{\mathbf{x}_t, t \in \mathbb{N}\}$ est considérée comme étant un processus de Markov sur un espace d'état \mathcal{X} discret, de loi initiale $p(\mathbf{x}_0)$ et de noyau de

transition $p(\mathbf{x}_t|\mathbf{x}_{t-1})$. Le modèle d'évolution de l'état \mathbf{x}_t est entièrement défini par :

$$\begin{cases} p(\mathbf{x}_0) \\ p(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad t \geq 1 \end{cases} \quad (2.1)$$

Il peut être également réécrit sous la forme

$$\begin{cases} p(\mathbf{x}_0) \\ \mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \mathbf{w}_t) \quad t \geq 1 \end{cases} \quad (2.2)$$

où f_t est une fonction linéaire ou non linéaire modélisant l'évolution de l'état \mathbf{x}_t et \mathbf{w}_t est le bruit d'état distribué selon une distribution quelconque.

Les observations $\{\mathbf{y}_t, t \in \mathbb{N}^*\}$, $\mathbf{y}_t \in \mathcal{Y}$ sont supposées indépendantes conditionnellement aux états cachés $\{\mathbf{x}_t, t \in \mathbb{N}\}$. Chaque observation \mathbf{y}_t est reliée à l'état caché \mathbf{x}_t par une densité de probabilité $p(\mathbf{y}_t|\mathbf{x}_t)$, appelée par la suite loi d'observation, définie par un modèle d'observation de la forme

$$\mathbf{y}_t = g_t(\mathbf{x}_t, \mathbf{v}_t) \quad t \geq 1 \quad (2.3)$$

où g_t est une fonction linéaire ou non linéaire modélisant la relation entre l'état et l'observation, et \mathbf{v}_t est le bruit d'observation distribué selon une distribution quelconque.

En résumé, la représentation d'état du système est composée du modèle d'évolution de l'état (2.2) et du modèle d'observation (2.3), elle s'écrit sous la forme :

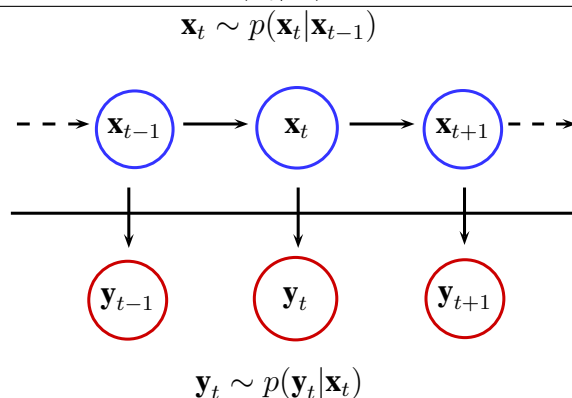
$$\begin{cases} \mathbf{x}_0 \sim p(\mathbf{x}_0) \\ \mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \mathbf{w}_t) \quad t \geq 1 \\ \mathbf{y}_t = g_t(\mathbf{x}_t, \mathbf{v}_t) \quad t \geq 1 \end{cases} \quad (2.4)$$

Une représentation graphique de ce modèle est illustrée par la figure 2.1.

2.1.2 Inférence bayésienne

Etant donnée le modèle dynamique décrit par l'équation 2.4, un problème majeur consiste à produire à un instant t une inférence sur la séquence d'états $\mathbf{x}_{0:t}$ en se basant sur les observations $\mathbf{y}_{1:t}$. L'approche bayésienne offre un cadre cohérent,

FIGURE 2.1 Représentation graphique du modèle dynamique. Un processus de Markov caché défini par la loi d'évolution de l'état $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ est observé par une observation \mathbf{y}_t selon la loi d'observation $p(\mathbf{y}_t|\mathbf{x}_t)$.



complet et exhaustif pour mener une inférence lorsque les informations disponibles sont entachées d'incertitude. Elle repose sur la détermination de la densité *a posteriori* $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:\tau})$. En effet, cette densité rassemble toute l'information disponible sur l'état caché lorsque l'on connaît les observations et les informations *a priori*. L'information fournie par les observations est contenue dans la densité de probabilité $p(\mathbf{y}_{1:\tau}|\mathbf{x}_{0:t})$, appelée fonction de vraisemblance. L'incertitude sur le vecteur de l'état caché est décrite par la densité de probabilité $p(\mathbf{x}_{0:t})$ appelée distribution *a priori*. La règle de Bayes permet d'exprimer la densité *a posteriori* $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:\tau})$ en fonction des différentes densités $p(\mathbf{y}_{1:\tau}|\mathbf{x}_{0:t})$, $p(\mathbf{x}_{0:t})$ et $p(\mathbf{y}_{1:\tau})$:

$$p(\mathbf{x}_{0:t}|\mathbf{y}_{1:\tau}) = \frac{p(\mathbf{y}_{1:\tau}|\mathbf{x}_{0:t})p(\mathbf{x}_{0:t})}{p(\mathbf{y}_{1:\tau})} \quad (2.5)$$

Cette densité est généralement utilisée pour résoudre trois types de problèmes :

- le filtrage, qui consiste à extraire de l'information à l'instant t à partir de données disponibles jusqu'à l'instant t inclus, c'est-à-dire dans l'équation (2.5) $\tau = t$;
- le lissage, qui consiste à extraire de l'information à l'instant t en utilisant quelques données futures, c'est-à-dire dans l'équation (2.5) $\tau > t$;
- la prédiction, qui consiste à prédire la valeur d'une quantité d'intérêt à un instant t futur par rapport aux données prises en compte, c'est-à-dire dans l'équation (2.5) $\tau < t$.

Dans cette thèse, nous considérons le problème de filtrage qui consiste à estimer à un instant t la séquence des états $\mathbf{x}_{0:t}$ ou l'état caché \mathbf{x}_t à partir des observations $y_{0:t}$. Dans le cadre d'une estimation hors ligne, le problème revient à déterminer la densité *a posteriori* complète $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T})$ ou sa densité marginale $p(\mathbf{x}_t|\mathbf{y}_{1:T})$ appelée aussi densité *a posteriori* de lissage, une fois que toutes les observations sont collectées. Cependant, ce traitement n'est pas adapté dans le contexte des applications temps réel qui nous intéressent où le problème d'estimation doit être résolu en traitant les observations au fur et à mesure de leur disponibilité. Ceci revient à faire de l'estimation en ligne où on cherche à estimer la densité *a posteriori* $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ ou sa densité marginale $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ appelée aussi densité *a posteriori* de filtrage.

Notre objectif est alors d'estimer la densité *a posteriori* $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ et, en particulier, la densité de filtrage qui lui est associée :

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \int_{\mathcal{X}^t} p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t-1} \quad (2.6)$$

Ces deux densités peuvent être évaluées de façon récursive. Pour la densité *a posteriori*, une expression récursive peut être obtenue en utilisant la règle de Bayes :

$$p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = p(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}) \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \quad (2.7)$$

Le calcul récursif de la densité de filtrage $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ se fait en deux étapes à savoir : une étape de prédiction et une étape de correction. La densité marginale $p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$ étant valable à l'instant t , l'étape de prédiction consiste alors à déterminer la densité *a priori* de l'état $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ en appliquant l'équation de Chapman-Kolmogorov :

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int_{\mathcal{X}} p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1} \quad (2.8)$$

Notons que l'équation (2.8) est déduite du fait que $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_{1:t-1}) = p(\mathbf{x}_t|\mathbf{x}_{t-1})$ puisque l'état \mathbf{x}_t est modélisé par un processus de Markov (cf. figure 2.1). L'étape de correction consiste à mettre à jour la densité *a posteriori*, quand une nouvelle mesure \mathbf{y}_t sera disponible à l'instant t , en utilisant la règle de Bayes :

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) \frac{p(\mathbf{y}_t|\mathbf{x}_t)}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \quad (2.9)$$

où la constante de normalisation $p(\mathbf{y}_t|\mathbf{y}_{1:t-1})$ est calculée par :

$$p(\mathbf{y}_t|\mathbf{y}_{1:t-1}) = \int_{\mathcal{X}} p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})d\mathbf{x}_t \quad (2.10)$$

A partir de la densité *a posteriori* $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ ou marginale $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, une estimation de l'état caché peut être obtenue en utilisant un estimateur bayésien. Les estimateurs bayésiens les plus utilisés dans les applications du traitement du signal sont l'estimateur MMSE (Minimum Mean Square Error) et l'estimateur MAP (Maximum A Posteriori). Ces deux estimateurs sont respectivement définis par :

$$\begin{aligned} \hat{\mathbf{x}}_t^{MMSE} &= \mathbb{E}[\mathbf{x}_t|\mathbf{y}_{1:t}] \\ &= \int_{\mathcal{X}^{t+1}} \mathbf{x}_t p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \\ &= \int_{\mathcal{X}} \mathbf{x}_t p(\mathbf{x}_t|\mathbf{y}_{1:t}) d\mathbf{x}_t \end{aligned} \quad (2.11)$$

et

$$\begin{aligned} \hat{\mathbf{x}}_t^{MAP} &= \arg \max_{\mathbf{x}_{0:t} \in \mathcal{X}^{t+1}} [p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})] \\ &= \arg \max_{\mathbf{x}_t \in \mathcal{X}} [p(\mathbf{x}_t|\mathbf{y}_{1:t})] \end{aligned} \quad (2.12)$$

2.2 Méthodes déterministes

Les équations (2.8) et (2.9) donnent la solution optimale du filtrage bayésien. Cette propagation récursive de la densité *a posteriori* est une solution théorique. En effet, elle ne peut être déterminée explicitement que dans certains cas. Dans le cas où la fonction f_t est linéaire en \mathbf{x}_t et \mathbf{w}_t , la fonction g_t est linéaire en \mathbf{x}_t et les bruits \mathbf{w}_t et \mathbf{v}_t sont gaussiens, on parle de modèle linéaire à bruits additifs gaussiens, ou plus simplement de modèle linéaire gaussien. Pour ce type de modèle, la solution explicite des équations (2.8) et (2.9) est donnée par le filtre de Kalman [Kal60]. Dans le cas où les fonctions f_t ou g_t sont non linéaires, des filtres non optimaux déterministes basés sur l'algorithme du filtre de Kalman peuvent être utilisés. Le filtre de Kalman étendu (EKF, Extended Kalman Filter) nécessite la linéarisation locale des fonctions non linéaires d'évolution et d'observation. Quant au filtre de Kalman sans parfum (UKF, Unscented Kalman Filter), il procède à une approximation de la densité de

probabilité par un nuage de points générés de façon déterministe en utilisant la transformée sans parfum (UT, Unscented Transform).

2.2.1 Filtre de Kalman

Dans le cas linéaire gaussien, le modèle initial défini par les équations (2.2) et (2.3) peut se réécrire sous la forme suivante :

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{C}_t \mathbf{w}_t \quad t \geq 1 \quad (2.13)$$

$$\mathbf{y}_t = \mathbf{G}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{v}_t \quad t \geq 1 \quad (2.14)$$

où \mathbf{F}_t et \mathbf{G}_t représentent respectivement la matrice d'évolution de l'état et la matrice d'observation, \mathbf{w}_t et \mathbf{v}_t sont des bruits additifs, gaussiens, mutuellement indépendants et indépendants de l'état initiale \mathbf{x}_0 , centrés, de matrices de covariance respectives \mathbf{Q}_t et \mathbf{R}_t et de matrices de transition respectives \mathbf{C}_t et \mathbf{D}_t . L'état initial \mathbf{x}_0 est distribué selon une densité gaussienne de moyenne $\hat{\mathbf{x}}_0$ et de matrice de covariance Σ_0 . Il est supposé être indépendant des bruits \mathbf{w}_t et \mathbf{v}_t .

Grâce aux hypothèses posées ci-dessus, les densités de probabilité $p(\mathbf{x}_t | \mathbf{y}_{1:t-1})$ et $p(\mathbf{x}_t | \mathbf{y}_{1:t})$ sont des distributions gaussiennes définies respectivement par les paramètres $\{\hat{\mathbf{x}}_{t|t-1}, \Sigma_{t|t-1}\}$ et $\{\hat{\mathbf{x}}_{t|t}, \Sigma_{t|t}\}$:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) = \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \Sigma_{t|t-1}) \quad (2.15)$$

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}) = \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}, \Sigma_{t|t}) \quad (2.16)$$

Le filtre de Kalman [Kal60] fournit la solution optimale aux équations (2.8) et (2.9) en évaluant les deux distributions gaussiennes définies ci-dessus (cf. équations (2.15) et (2.16)). Ces deux distributions sont calculées de manière récursive suivant les étapes de prédiction et de correction. L'étape de prédiction du filtre de Kalman consiste à calculer la moyenne $\hat{\mathbf{x}}_{t|t-1}$, appelée la prédiction de l'état, et la matrice de covariance $\Sigma_{t|t-1}$, appelée la matrice de covariance de l'erreur de prédiction, de la densité $p(\mathbf{x}_t | \mathbf{y}_{1:t-1})$. L'étape de correction du filtre de Kalman consiste à calculer la moyenne $\hat{\mathbf{x}}_{t|t}$ et la matrice de covariance $\Sigma_{t|t}$ de la densité $p(\mathbf{x}_t | \mathbf{y}_{1:t})$ respectivement appelés l'estimée de l'état et la matrice de covariance de l'erreur d'estimation. Le principe du filtre de Kalman est donnée par l'algorithme 3.

Algorithme 3 Filtre de Kalman

- A l'instant $t = 0$

- $\hat{\mathbf{x}}_{0|0} = \hat{\mathbf{x}}_0$
- $\Sigma_{0|0} = \Sigma_0$

- A l'instant $t > 0$

Prédiction

- Prédiction de l'état : $\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1}$
- Covariance de l'erreur de prédiction : $\Sigma_{t|t-1} = \mathbf{F}_t \Sigma_{t-1|t-1} \mathbf{F}_t^T + \mathbf{C}_t \mathbf{Q}_t \mathbf{C}_t^T$
- Prédiction de l'observation : $\hat{\mathbf{y}}_{t|t-1} = \mathbf{G}_t \hat{\mathbf{x}}_{t|t-1}$

Correction

- Innovation : $\tilde{\mathbf{y}}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$
- Covariance de l'innovation : $\mathbf{S}_t = \mathbf{G}_t \Sigma_{t|t-1} \mathbf{G}_t^T + \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t^T$
- Gain de Kalman : $\mathbf{K}_t = \Sigma_{t|t-1} \mathbf{G}_t^T \mathbf{S}_t^{-1}$
- Estimée de l'état : $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \tilde{\mathbf{y}}_t$
- Covariance de l'erreur d'estimation : $\Sigma_{t|t} = \Sigma_{t|t-1} - \mathbf{K}_t \mathbf{S}_t \mathbf{K}_t^T$

2.2.2 Filtre de Kalman étendu

Le filtre de Kalman est l'un des outils les plus puissants du traitement statistique du signal pour le filtrage. Cependant, sa mise en œuvre n'est possible que sur les modèles linéaires gaussiens. Dans le cas d'un modèle non linéaire avec des bruits additifs, il n'est plus possible d'écrire les équations (2.2) et (2.3) sous les formes (2.13) et (2.14) et par conséquent l'application du filtre de Kalman n'est plus possible. En revanche, lorsque le modèle est faiblement non linéaire, l'utilisation du filtre de Kalman étendu se présente comme étant une solution adaptée. L'idée de l'EKF est de linéariser les fonctions du modèle autour de l'estimée courante et d'appliquer le filtre de Kalman. Il fait l'hypothèse que la densité de probabilité *a posteriori* $p(\mathbf{x}_t | \mathbf{y}_{1:t})$, qui n'est plus gaussienne dans ce cas, peut être approchée par une densité gaussienne.

Considérons le modèle non linéaire gaussien suivant :

$$\mathbf{x}_t = f_t(\mathbf{x}_{t-1}) + \mathbf{C}_t \mathbf{w}_t \quad (2.17)$$

$$\mathbf{y}_t = g_t(\mathbf{x}_t) + \mathbf{D}_t \mathbf{v}_t \quad (2.18)$$

où f_t et g_t sont deux fonctions non linéaires et supposées dérivables, \mathbf{w}_t et \mathbf{v}_t sont des bruits additifs, gaussiens, centrés, de matrices de covariance respectives \mathbf{Q}_t et \mathbf{R}_t et de matrices de transition respectives \mathbf{C}_t et \mathbf{D}_t . Les bruits \mathbf{w}_t et \mathbf{v}_t sont supposés

indépendants entre eux et indépendants de l'état initiale \mathbf{x}_0 .

Etant donné que les fonctions f_t et g_t sont supposées dérivables, elles peuvent être respectivement linéarisées autour de l'état courant $\widehat{\mathbf{x}}_{t-1|t-1}$ et l'état prédit $\widehat{\mathbf{x}}_{t|t-1}$:

$$f_t(\mathbf{x}_{t-1}) \simeq f_t(\widehat{\mathbf{x}}_{t-1|t-1}) + \underbrace{\frac{\partial f_t}{\partial \mathbf{x}_{t-1}} \Big|_{\mathbf{x}_{t-1}=\widehat{\mathbf{x}}_{t-1|t-1}}}_{\widehat{\mathbf{F}}_t} (\mathbf{x}_{t-1} - \widehat{\mathbf{x}}_{t-1|t-1}) \quad (2.19)$$

$$g_t(\mathbf{x}_t) \simeq g_t(\widehat{\mathbf{x}}_{t|t-1}) + \underbrace{\frac{\partial g_t}{\partial \mathbf{x}_t} \Big|_{\mathbf{x}_t=\widehat{\mathbf{x}}_{t|t-1}}}_{\widehat{\mathbf{G}}_t} (\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) \quad (2.20)$$

Le modèle défini par les équations (2.17) et (2.18) peut alors s'écrire sous la forme du modèle linéarisé suivant :

$$\mathbf{x}_t \approx \widehat{\mathbf{F}}_t \mathbf{x}_{t-1} + f_t(\widehat{\mathbf{x}}_{t-1|t-1}) - \widehat{\mathbf{F}}_t \mathbf{x}_{t-1|t-1} + \mathbf{C}_t \mathbf{w}_t \quad (2.21)$$

$$\mathbf{y}_t \approx \widehat{\mathbf{G}}_t \mathbf{x}_t + g_t(\widehat{\mathbf{x}}_{t|t-1}) - \widehat{\mathbf{G}}_t \widehat{\mathbf{x}}_{t|t-1} + \mathbf{D}_t \mathbf{v}_t \quad (2.22)$$

En appliquant les équations du filtre de Kalman sur le modèle précédent, on obtient les récursions de l'EKF qui conduisent à l'algorithme 4.

Algorithme 4 Filtre de Kalman étendu

- A l'instant $t = 0$

- $\widehat{\mathbf{x}}_{0|0} = \widehat{\mathbf{x}}_0$
- $\widehat{\Sigma}_{0|0} = \widehat{\Sigma}_0$

- A l'instant $t > 0$

Prédiction

- Prédiction de l'état : $\widehat{\mathbf{x}}_{t|t-1} = f_t(\widehat{\mathbf{x}}_{t-1|t-1})$
- Covariance de l'erreur de prédiction : $\widehat{\Sigma}_{t|t-1} = \widehat{\mathbf{F}}_t \widehat{\Sigma}_{t-1|t-1} \widehat{\mathbf{F}}_t^T + \mathbf{C}_t \mathbf{Q}_t \mathbf{C}_t^T$
- Prédiction de l'observation : $\widehat{\mathbf{y}}_{t|t-1} = g_t(\widehat{\mathbf{x}}_{t|t-1})$

Correction

- Innovation : $\widetilde{\mathbf{y}}_t = \mathbf{y}_t - \widehat{\mathbf{y}}_{t|t-1}$
 - Covariance de l'innovation : $\mathbf{S}_t = \widehat{\mathbf{G}}_t \widehat{\Sigma}_{t|t-1} \widehat{\mathbf{G}}_t^T + \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t^T$
 - Gain de Kalman : $\mathbf{K}_t = \widehat{\Sigma}_{t|t-1} \widehat{\mathbf{G}}_t^T \mathbf{S}_t^{-1}$
 - Estimée de l'état : $\widehat{\mathbf{x}}_{t|t} = \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \widetilde{\mathbf{y}}_t$
 - Covariance de l'erreur d'estimation : $\widehat{\Sigma}_{t|t} = \widehat{\Sigma}_{t|t-1} - \mathbf{K}_t \mathbf{S}_t \mathbf{K}_t^T$
-

2.2.3 Filtre de Kalman sans parfum

Le filtre de Kalman sans parfum, proposé par Julier *et al* [JU97], est une autre extension du filtre de Kalman aux systèmes non linéaires. Comme pour l'EKF, son principe repose sur l'approximation de la densité *a posteriori* par une densité gaussienne. Mais au lieu de procéder à la linéarisation des fonctions non linéaires, il approche la densité de probabilité par un ensemble de points pondérés convenablement choisis de façon déterministe, ces points étant transformés par les fonctions non linéaires f_t et g_t afin d'obtenir une nouvelle densité de probabilité. Cette approximation est appelée la transformée sans parfum.

2.2.3.1 Transformée sans parfum

L'UT [WVdM00, JU97] est une méthode permettant de calculer les statistiques d'une variable aléatoire qui subit une transformation non linéaire. Son principe est décrit dans les paragraphes suivants.

Considérons le modèle d'état suivant :

$$\mathbf{y} = f(\mathbf{x})$$

avec \mathbf{x} une variable aléatoire de dimension n , de moyenne $\bar{\mathbf{x}}$ et de covariance \mathbf{P}_{xx} , et \mathbf{y} une variable aléatoire de statistiques à déterminer. Afin de trouver la moyenne $\bar{\mathbf{y}}$ et la matrice de covariance \mathbf{P}_{yy} de \mathbf{y} , la transformée sans parfum opère en trois étapes :

Etape 1

Construction de manière déterministe d'un ensemble de $2n+1$ points pondérés, tel que leur moyenne et covariance soient respectivement égaux à $\bar{\mathbf{x}}$ et \mathbf{P}_{xx} . Ces points sont représentatifs de la forme de la densité de probabilité de \mathbf{x} :

$$\begin{aligned} \chi_0 &= \bar{\mathbf{x}} & W_0 &= \frac{k}{n+k} \\ \chi_i &= \bar{\mathbf{x}} + \left(\sqrt{(n+k)\mathbf{P}_{xx}} \right)_i & W_i &= \frac{1}{2(n+k)} \\ \chi_{i+n} &= \bar{\mathbf{x}} - \left(\sqrt{(n+k)\mathbf{P}_{xx}} \right)_i & W_{i+n} &= \frac{1}{2(n+k)} \end{aligned} \quad (2.23)$$

où $k \in \mathbb{R}$ est un paramètre d'échelle qui contrôle la répartition des points χ_i

autour de la moyenne \bar{x} , $\left(\sqrt{(n+k)P_{xx}}\right)_i$ est la $i^{\text{ième}}$ colonne de la décomposition de Cholesky de la matrice racine carrée de $(n+k)\mathbf{P}_{xx}$ et W_i est le poids associé au $i^{\text{ième}}$ point.

Etape 2

Propager chaque point χ_i au travers de la fonction non linéaire f afin d'obtenir un nuage de points transformés de moyenne \bar{y} et de matrice de covariance \mathbf{P}_{yy} :

$$\mathcal{Y}_i = f(\chi_i)$$

Etape 3

L'estimation de la moyenne \bar{y} et l'estimation de la matrice de covariance \mathbf{P}_{yy} sont respectivement données par la moyenne et la covariance pondérées des points transformés :

$$\bar{y} = \sum_{i=0}^{2n} \mathcal{Y}_i W_i$$

$$\mathbf{P}_{yy} = \sum_{i=0}^{2n} W_i (\mathcal{Y}_i - \bar{y})(\mathcal{Y}_i - \bar{y})^T$$

2.2.3.2 Algorithme

Nous considérons le modèle non linéaire gaussien défini par le modèle d'évolution de l'état (2.17) et le modèle d'observation (2.18). Nous notons $\mathbf{x}_t^a = [\mathbf{x}_t^T \quad \mathbf{w}_t^T \quad \mathbf{v}_t^T]^T$ le vecteur d'état augmenté par les bruits d'évolution et d'observation de dimension $n^a = n + q + r$, où n , q et r sont respectivement les dimensions du vecteur d'état et des vecteurs des bruits d'évolution et d'observation.

L'UKF est donné par l'algorithme 5.

D'une manière générale, ce filtre fournit de meilleurs résultats que l'EKF car l'effet des non linéarités est mieux pris en compte.

Algorithme 5 Filtre de Kalman sans parfum

- A l'instant $t = 0$

$$- \hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0]$$

$$- \Sigma_{0|0} = \mathbb{E}[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T]$$

$$- \hat{\mathbf{x}}_0^a = \mathbb{E}[\mathbf{x}_0^a] = \begin{bmatrix} \hat{\mathbf{x}}_{0|0}^T & \mathbf{0}_{1 \times q} & \mathbf{0}_{1 \times r} \end{bmatrix}^T$$

$$- \Sigma_0^a = \mathbb{E}[(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)^T] = \begin{pmatrix} \Sigma_{0|0} & 0 & 0 \\ 0 & \mathbf{Q}_0 & 0 \\ 0 & 0 & \mathbf{R}_0 \end{pmatrix}$$

- A l'instant $t > 0$

Calcul des points d'approximation et des poids associés

$$- \hat{\mathbf{x}}_{t-1}^a = \begin{bmatrix} \hat{\mathbf{x}}_{t-1|t-1}^T & \mathbf{0}_{1 \times q} & \mathbf{0}_{1 \times r} \end{bmatrix}^T$$

$$- \Sigma_{t-1}^a = \begin{pmatrix} \Sigma_{t-1|t-1} & 0 & 0 \\ 0 & \mathbf{Q}_{t-1} & 0 \\ 0 & 0 & \mathbf{R}_{t-1} \end{pmatrix}$$

$$- \chi_{t-1}^a = [\hat{\mathbf{x}}_{t-1}^a \quad \hat{\mathbf{x}}_{t-1}^a + \sqrt{(n^a + k)\Sigma_{t-1}^a} \quad \hat{\mathbf{x}}_{t-1}^a - \sqrt{(n^a + k)\Sigma_{t-1}^a}]$$

$$- \text{Calculer les poids } \{W_i\}_{i=1}^{2n} \text{ en utilisant l'équation (2.23)}$$

Prédiction

$$- \chi_{t|t-1}^x = f_t(\chi_{t-1}^x, \chi_{t-1}^w)$$

$$- \hat{\mathbf{x}}_{t|t-1} = \sum_{i=0}^{2n^a} W_i \chi_{i,t|t-1}^x$$

$$- \Sigma_{t|t-1} = \sum_{i=0}^{2n^a} W_i (\chi_{i,t|t-1}^x - \hat{\mathbf{x}}_{t|t-1})(\chi_{i,t|t-1}^x - \hat{\mathbf{x}}_{t|t-1})^T$$

$$- \mathcal{Y}_{t|t-1} = g_t(\chi_{t|t-1}^x, \chi_{t-1}^v)$$

$$- \hat{\mathbf{y}}_{t|t-1} = \sum_{i=0}^{2n^a} W_i \mathcal{Y}_{i,t|t-1}$$

Correction

$$- \tilde{\mathbf{y}}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$$

$$- \mathbf{S}_t = \sum_{i=0}^{2n^a} W_i (\mathcal{Y}_{i,t|t-1} - \hat{\mathbf{y}}_{t|t-1})(\mathcal{Y}_{i,t|t-1} - \hat{\mathbf{y}}_{t|t-1})^T + \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t^T$$

$$- \mathbf{P}_{\mathbf{x}_t \mathbf{y}_t} = \sum_{i=0}^{2n^a} W_i (\chi_{i,t|t-1}^x - \hat{\mathbf{x}}_{t|t-1})(\mathcal{Y}_{i,t|t-1} - \hat{\mathbf{y}}_{t|t-1})^T$$

$$- \mathbf{K}_t = \mathbf{P}_{\mathbf{x}_t \mathbf{y}_t} \mathbf{S}_t^{-1}$$

$$- \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \tilde{\mathbf{y}}_t$$

$$- \Sigma_{t|t} = \Sigma_{t|t-1} - \mathbf{K}_t \mathbf{S}_t \mathbf{K}_t^T$$

2.3 Méthodes de Monte Carlo

Le filtre de Kalman apporte une solution optimale au problème de filtrage bayésien dans le cas des modèles linéaires gaussiens. Dans le cas des modèles non linéaires, les algorithmes déterministes de l'EKF et de l'UKF fournissent des solutions sous optimales en approchant la densité *a posteriori* par une distribution gaussienne. Cependant, ces techniques sont inadaptées aux modèles fortement non linéaires avec des bruits très importants. En effet, dans ce cas de figure, la densité *a posteriori* peut présenter un caractère très piqué et/ou multimodal, et son approximation par une distribution gaussienne n'est plus réaliste et peut conduire à des erreurs d'estimation importantes. Dans ce cas, il faut donc avoir recours à des méthodes d'approximation numériques comme les méthodes de Monte Carlo afin d'appréhender n'importe quel type de densité *a posteriori*.

Les méthodes de Monte Carlo tirent leurs origines des recherches menées par Metropolis, Ulam et Von Neumann [MU49], notamment au cours de la seconde guerre mondiale dans le cadre du projet américain Manhattan concernant le développement de l'arme nucléaire. Le nom de la méthode fait référence aux jeux de hasard pratiqués dans les casinos et plus particulièrement dans celui situé à Monte Carlo, une ville de la principauté de Monaco. Depuis une vingtaine d'années, ces méthodes ont véritablement pris leur essor et de nouvelles méthodes spécifiques telles que les méthodes de Monte Carlo séquentielles (SMC, Sequential Monte Carlo) ont été développées [GSS93].

Dans cette section, nous passons en revue les méthodes de Monte Carlo en se basant principalement sur les ouvrages [DDFG01, Che03] ainsi que les articles de synthèse [DGA00, AMGC02, DW05, CGM07].

2.3.1 Intégration de Monte Carlo

L'estimation dans le cadre de l'inférence bayésienne nécessite le calcul d'intégrales du type

$$I(h) = \mathbb{E}_{p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})} [h(\mathbf{x}_{0:t})] = \int_{\mathcal{X}^{t+1}} h(\mathbf{x}_{0:t}) p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \quad (2.24)$$

Dans le cas de densités de probabilité complexes, le calcul analytique de cette intégrale n'est pas envisageable, et les techniques classiques d'intégration ne sont pas suffisamment précises et/ou sont trop coûteuses en temps de calcul.

Considérons N réalisations de la variable aléatoire $\mathbf{x}_{0:t}$ indépendantes et identiquement distribuées selon $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$, appelées particules et notées $\tilde{\mathbf{x}}_{0:t}^{(i)}$:

$$\tilde{\mathbf{x}}_{0:t}^{(i)} \sim p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) \quad \text{pour } i = 1, \dots, N \quad (2.25)$$

L'approximation de Monte Carlo de la densité *a posteriori* est alors donnée par la distribution empirique suivante :

$$P_N(d\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{\mathbf{x}}_{0:t}^{(i)}}(d\mathbf{x}_{0:t}) \quad (2.26)$$

où $\delta_x(\cdot)$ est la fonction de Dirac au point x . En se basant sur cette approximation, l'intégrale $I(h)$, définie dans l'équation (2.24), peut être approchée par :

$$\hat{I}_N(h) = \int_{\mathcal{X}^{t+1}} h(\mathbf{x}_{0:t}) P_N(d\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = \frac{1}{N} \sum_{i=1}^N h(\tilde{\mathbf{x}}_{0:t}^{(i)}) \quad (2.27)$$

Il est possible de montrer que cette estimation de l'intégrale $I(h)$ est non biaisée. De plus, si la variance *a posteriori* de $h(\mathbf{x}_{0:t})$, notée σ_h^2 , est finie

$$\sigma_h^2 \triangleq E_{p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})} [h(\mathbf{x}_{0:t})^2] - I(h)^2 < +\infty \quad (2.28)$$

alors la variance de $\hat{I}_N(h)$ est donnée par

$$\widehat{I}_N(h) = \frac{\sigma_h^2}{N} \quad (2.29)$$

Ainsi, d'après la loi forte des grands nombres, $\hat{I}_N(h)$ converge presque sûrement vers $I(h)$ lorsque N tend vers l'infini

$$\hat{I}_N(h) \xrightarrow[N \rightarrow \infty]{p.s.} I(h) \quad (2.30)$$

et la vitesse de convergence est donnée par le théorème de la limite centrale :

$$\sqrt{N} \left[\hat{I}_N(h) - I(h) \right] \xrightarrow[N \rightarrow \infty]{loi} \mathcal{N}(0, \sigma_h^2) \quad (2.31)$$

Ce résultat met en avant un point important des méthodes de Monte Carlo. En

effet, à partir d'un ensemble de N échantillons $\left\{ \tilde{\mathbf{x}}_{0:t}^{(i)} \right\}_{i=1}^N$, on peut facilement trouver une approximation de l'intégrale $I(h)$ de n'importe quelle fonction h et la vitesse de convergence du calcul de cette approximation est indépendante de la dimension de l'espace d'état. Avec les méthodes classiques d'intégration, la vitesse de convergence augmente de façon exponentielle avec la dimension de l'espace d'état. Malheureusement, un problème se pose en pratique. Il est souvent impossible de générer des échantillons selon la densité *a posteriori*. Autrement dit, l'obtention des particules n'est pas un problème trivial et le recours à des algorithmes spécifiques est nécessaire. Il existe de nombreuses méthodes d'échantillonnage selon une loi $p(x)$, notamment la méthode de l'inverse de la fonction de répartition [Dev86], la méthode d'acceptation/rejet, l'échantillonnage d'importance (IS, Importance Sampling), les méthodes de Monte Carlo par chaînes de Markov (MCMC, Markov Chain Monte Carlo) [Has70], etc. Nous présentons ci-dessous l'échantillonnage d'importance.

2.3.1.1 Echantillonnage d'importance

L'objectif de l'échantillonnage d'importance est de générer des particules dans les régions d'importance de la densité *a posteriori*, c'est-à-dire dans les régions de l'espace d'état dans lesquelles la densité *a posteriori* prend des valeurs élevées. Il repose sur une densité de probabilité $\pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$, appelée densité d'importance, dont le support contient celui de $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ et à partir de laquelle il est possible de générer des échantillons. Sous ces hypothèses, l'intégrale définie dans l'équation (2.24) peut alors être reformulée de la manière suivante :

$$\begin{aligned}
 I(h) &= \int_{\mathcal{X}^{t+1}} h(\mathbf{x}_{0:t}) p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \\
 &= \int_{\mathcal{X}^{t+1}} h(\mathbf{x}_{0:t}) \frac{p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})}{\pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})} \pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \\
 &= \frac{\int_{\mathcal{X}^{t+1}} h(\mathbf{x}_{0:t}) \omega(\mathbf{x}_{0:t}) \pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t}}{\int_{\mathcal{X}^{t+1}} \omega(\mathbf{x}_{0:t}) \pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) d\mathbf{x}_{0:t}}
 \end{aligned} \tag{2.32}$$

où $\omega(\mathbf{x}_{0:t})$ est le poids d'importance non normalisé, défini par :

$$\omega(\mathbf{x}_{0:t}) = \frac{p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})}{\pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})} \tag{2.33}$$

Considérons un ensemble de N échantillons $\{\tilde{\mathbf{x}}_{0:t}^{(i)}\}_{i=1}^N$ générés selon la densité d'importance $\pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$:

$$\tilde{\mathbf{x}}_{0:t}^{(i)} \sim \pi(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) \quad (2.34)$$

En appliquant la méthode de Monte Carlo au numérateur et au dénominateur de l'équation (2.32), l'intégrale $I(h)$ peut être approchée comme suit :

$$\begin{aligned} \hat{I}_N(h) &= \frac{\frac{1}{N} \sum_{i=1}^N \omega(\tilde{\mathbf{x}}_{0:t}^{(i)}) h(\tilde{\mathbf{x}}_{0:t}^{(i)})}{\frac{1}{N} \sum_{i=1}^N \omega(\tilde{\mathbf{x}}_{0:t}^{(i)})} \\ &= \sum_{i=1}^N \tilde{\omega}_t^{(i)} h(\tilde{\mathbf{x}}_{0:t}^{(i)}) \end{aligned} \quad (2.35)$$

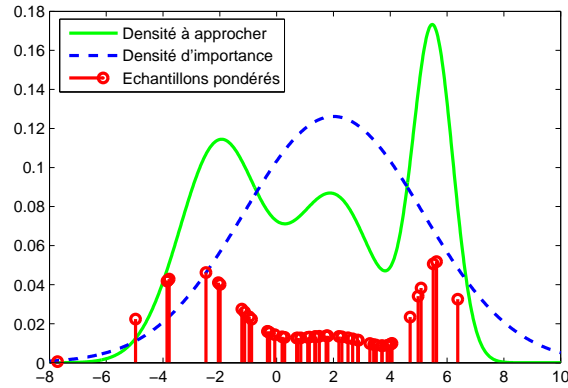
où $\tilde{\omega}_t^{(i)}$ est le poids d'importance normalisé, donnée par :

$$\tilde{\omega}_t^{(i)} = \frac{\omega(\tilde{\mathbf{x}}_{0:t}^{(i)})}{\sum_{j=1}^N \omega(\tilde{\mathbf{x}}_{0:t}^{(j)})} \quad (2.36)$$

Exemple 2.1. *Nous cherchons à approcher la densité $p(x) = 0.4\mathcal{N}(x; -2, 2) + 0.3\mathcal{N}(x; 2, 2) + 0.3\mathcal{N}(x; 5.5, 0.5)$. La densité d'importance choisie est une distribution gaussienne de moyenne 2 et de variance 10 : $\pi(x) = \mathcal{N}(x; 2, 10)$. Nous considérons un ensemble de $N = 30$ échantillons générés selon π . La densité p est approchée par la distribution empirique $P_N(dx) = \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta_{\tilde{x}^{(i)}}(dx)$, où $\tilde{\omega}_t^{(i)} = \frac{p(\tilde{x}^{(i)})}{\pi(\tilde{x}^{(i)})}$. Sur la figure 2.2 sont représentées la densité à approcher p , la densité d'importance π ainsi que l'approximation par la distribution empirique $P_N(dx)$ de cette densité.*

L'échantillonnage d'importance, tel qu'il est présenté ici, est inadapté pour la mise en œuvre d'une procédure séquentielle pour l'estimation de la densité *a posteriori*. En effet, lorsqu'une nouvelle observation \mathbf{y}_t est disponible, il faut redéterminer l'ensemble des particules pour tous les instants. On obtient donc une complexité de calculs qui augmente avec le temps. Cependant, l'échantillonnage d'importance peut être modifié afin que le calcul d'une approximation de la densité *a posteriori* puisse être réalisé sans avoir besoin de modifier l'ensemble des trajectoires des particules $\{\tilde{\mathbf{x}}_{0:t-1}^{(i)}\}_{i=1}^N$. La méthode correspondante porte le nom d'échantillonnage d'impor-

FIGURE 2.2 Echantillonnage d'importance. La densité $p(x) = 0.4\mathcal{N}(x; -2, 2) + 0.3\mathcal{N}(x; 2, 2) + 0.3\mathcal{N}(x; 5.5, 0.5)$ est approchée avec un système de $N = 30$ particules pondérées distribuées selon la densité d'importance $\pi(x) = \mathcal{N}(x; 2, 10)$.



tance séquentiel et elle est la base des méthodes de Monte Carlo séquentielles qui sont abordées dans la section suivante.

2.3.2 Les méthodes de Monte Carlo séquentielles

Les méthodes de Monte Carlo séquentielles, communément appelées méthodes de filtrage particulaire, sont des méthodes numériques permettant d'approcher la densité de probabilité *a posteriori* au moyen de la distribution de probabilité empirique d'un système de particules. Entre deux instants d'observation, les particules se déplacent de façon indépendante selon la dynamique de l'état caché, et sitôt qu'une nouvelle observation est disponible, un rééchantillonnage a lieu, où les particules sont choisies en fonction de leur adéquation à la nouvelle observation (quantifiée par la fonction de vraisemblance). Sous l'effet du rééchantillonnage, qui constitue l'étape essentielle de la méthode, les particules se concentrent automatiquement dans les régions d'intérêt de l'espace d'état.

2.3.2.1 Echantillonnage d'importance séquentiel

L'aspect séquentiel du problème d'estimation de la densité *a posteriori* se résume dans l'utilisation du résultat à l'instant $t - 1$ pour estimer la densité à l'instant t . On rappelle qu'entre deux instants $t - 1$ et t , la densité *a posteriori* admet la formulation

réursive donnée par l'équation (2.7) :

$$p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = p(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}) \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}$$

Afin d'appliquer séquentiellement l'algorithme d'échantillonnage d'importance, la densité d'importance doit vérifier la forme réursive suivante :

$$\begin{aligned} \pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) &= \pi_{t-1}(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1})q_t(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t}) \\ &= \pi_0(\mathbf{x}_0) \prod_{k=1}^t q_k(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) \end{aligned} \quad (2.37)$$

A l'instant 0, les particules sont échantillonnées selon $\pi_0(\mathbf{x}_0)$. La densité d'importance $q_t(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t})$ permet de propager les particules de l'instant $t-1$ à l'instant t . On peut noter que la propagation augmente la dimension de la trajectoire du vecteur d'état $\mathbf{x}_{0:t}$. Construites de cette manière, les particules à l'instant t sont distribuées selon $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$. Pour corriger cette différence avec la densité *a posteriori*, il est nécessaire de pondérer les particules en calculant les poids d'importance. En utilisant les équations (2.7) et (2.37) dans l'équation (2.33), nous obtenons l'expression réursive suivante pour la mise à jour des poids :

$$\begin{aligned} \tilde{\omega}_t^{(i)} &= \frac{p(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{1:t})}{\pi(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{1:t})} \\ &= p(\tilde{\mathbf{x}}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1}) \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})p(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})}{\pi_{t-1}(\tilde{\mathbf{x}}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1})q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \\ &= \frac{p(\tilde{\mathbf{x}}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1})}{\pi_{t-1}(\tilde{\mathbf{x}}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1})} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})p(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \\ &= \tilde{\omega}_{t-1}^{(i)} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})p(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \end{aligned} \quad (2.38)$$

soit

$$\tilde{\omega}_t^{(i)} \propto \tilde{\omega}_{t-1}^{(i)} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})p(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})} \quad (2.39)$$

L'utilisation d'une formule définie à une constante près ne pose pas de problème, les poids étant ensuite normalisés par :

$$\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{i=1}^N \tilde{\omega}_t^{(i)}} \quad (2.40)$$

Finalement, la densité *a posteriori* est approchée par :

$$P_N(d\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta_{\tilde{\mathbf{x}}_{0:t}^{(i)}}(d\mathbf{x}_{0:t}) \quad (2.41)$$

L'échantillonnage d'importance séquentiel est détaillé par l'algorithme 6.

Algorithme 6 Echantillonnage d'importance séquentiel

- A l'instant $t = 0$
 - pour** $i = 1$ à N **faire**
 - Générer $\tilde{\mathbf{x}}_0^{(i)} \sim \pi_0(\mathbf{x}_0)$
 - Initialiser les poids d'importance $\tilde{\omega}_0^{(i)} = \frac{1}{N}$
 - fin pour**
 - A l'instant $t > 0$
 - pour** $i = 1$ à N **faire**
 - Générer $\tilde{\mathbf{x}}_t^{(i)} \sim q_t(\mathbf{x}_t | \tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})$
 - Définir $\tilde{\mathbf{x}}_{0:t}^{(i)} = \{\tilde{\mathbf{x}}_{0:t-1}^{(i)}, \tilde{\mathbf{x}}_t^{(i)}\}$
 - Calculer les poids d'importance $\tilde{\omega}_t^{(i)} \propto \tilde{\omega}_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{(i)})p(\tilde{\mathbf{x}}_t^{(i)} | \tilde{\mathbf{x}}_{t-1}^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)} | \tilde{\mathbf{x}}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})}$
 - fin pour**
 - pour** $i = 1$ à N **faire**
 - Normaliser les poids d'importance $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^N \tilde{\omega}_t^{(j)}}$
 - fin pour**
-

2.3.2.2 Dégénérescence des poids

Une limitation connue de l'algorithme d'échantillonnage d'importance séquentiel est que la variance des poids augmente avec le temps et ceci à cause du caractère multiplicatif de la formule récursive des poids [KLW94]. Ce phénomène est appelé la dégénérescence des poids et se traduit par une diminution significative du nombre

des particules utiles au cours du temps. En effet, après plusieurs itérations, la plupart des particules ont des poids nuls à l'exception de quelques-unes ayant des poids très importants. Ceci fausse l'approximation de la densité *a posteriori* et provoque la divergence du filtre. Afin de remédier à ce problème, une étape de rééchantillonnage est ajoutée à l'algorithme 6. Le principe du rééchantillonnage est de dupliquer les particules ayant un poids fort au détriment de celles ayant un poids faible tout en restant à un nombre de particules constant.

Plusieurs méthodes de rééchantillonnage existent dans la littérature. Les plus fréquemment rencontrées sont : le rééchantillonnage multinomial [GSS93], le rééchantillonnage résiduel [LC98] et le rééchantillonnage systématique [Kit96]. Les algorithmes de rééchantillonnage systématique et de rééchantillonnage résiduel sont les plus intéressants dans la mesure où ils améliorent la qualité d'estimation du filtre tout en réduisant la complexité du calcul [HSG06]. Compte tenu de sa facilité d'implémentation, c'est la méthode de rééchantillonnage systématique qui est la plus utilisée [DC05]. L'algorithme 7 décrit le rééchantillonnage systématique.

Algorithme 7 Rééchantillonnage systématique

Calcul de la somme cumulative des poids

$$c_1 = \tilde{\omega}_t^{(1)}$$

pour $i = 2$ à N **faire**

$$c_i = c_{i-1} + \tilde{\omega}_t^{(i)}$$

fin pour

$$u \sim \mathcal{U}\left[0, \frac{1}{N}\right]$$

$$i = 1$$

pour $j = 1$ à N **faire**

$$\text{Générer } u_j = u + \frac{j-1}{N}$$

tant que $u_j > c_i$ **faire**

$$i = i + 1$$

fin tant que

$$\mathbf{x}_t^{(j)} = \tilde{\mathbf{x}}_t^{(i)}$$

$$\omega_t^{(j)} = \frac{1}{N}$$

fin pour

Il est important de noter que, quel que soit la méthode utilisée, l'étape de rééchantillonnage entraîne des erreurs de Monte Carlo. Il est donc préférable de ne pas passer par l'étape de rééchantillonnage à chaque cycle de calcul et d'introduire un indicateur de rééchantillonnage. Plusieurs critères ont été proposés dans la litté-

ture dont le plus utilisé est celui introduit dans [KLW94, Liu96]. Il est basé sur la variance empirique des poids et il est défini par :

$$\begin{aligned} N_{eff} &= \frac{N}{1 + \text{Var}[\tilde{\omega}_t]} \\ &= \frac{N}{1 + \mathbb{E}[\tilde{\omega}_t^2]} \end{aligned} \quad (2.42)$$

Ce critère peut être estimé par :

$$N_{eff} \approx \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}. \quad (2.43)$$

Le principe est alors de rééchantillonner lorsque N_{eff} devient inférieur à un certain seuil N_{seuil} .

2.3.2.3 Algorithme

Les différentes étapes du filtrage particulaire sont données par l'algorithme 8.

2.3.2.4 Choix de la densité d'importance

La densité d'importance joue un rôle primordial dans les algorithmes de filtrage particulaire et son choix conditionne le bon fonctionnement du filtre. Elle doit être choisie de préférence pour minimiser la variance des poids. Les principaux choix possibles sont :

- **Densité d'importance optimale :**

La densité d'importance optimale a été introduite par Zaritskii *et al* [ZS75] et elle est donnée par :

$$q_t(\mathbf{x}_t | \mathbf{x}_{t-1}, y_t) = p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y}_t) \quad (2.44)$$

Dans ce cas, les poids sont calculés comme suit :

$$\tilde{\omega}_t^{(i)} \propto \tilde{\omega}_{t-1}^{(i)} p(\mathbf{y}_t | \tilde{\mathbf{x}}_{t-1}^{(i)}) \quad (2.45)$$

avec

$$p(\mathbf{y}_t | \tilde{\mathbf{x}}_{t-1}^{(i)}) = \int_{\mathcal{X}} p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \tilde{\mathbf{x}}_{t-1}^{(i)}) d\mathbf{x}_t \quad (2.46)$$

Algorithme 8 Filtre Particulaire

- A l'instant $t = 0$

pour $i = 1$ à N **faire**

- Echantillonner $\mathbf{x}_0^{(i)} \sim \pi_0(\mathbf{x}_0)$

- Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$

fin pour

- A l'instant $t > 0$

pour $i = 1$ à N **faire**

- Echantillonner $\tilde{\mathbf{x}}_t^{(i)} \sim q_t(\mathbf{x}_t | \tilde{\mathbf{x}}_{t-1}^{(i)})$

- Calculer les poids $\tilde{\omega}_t^{(i)} = \omega_{t-1}^{(i)} \frac{p(\tilde{\mathbf{x}}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}) p(y_t | \tilde{\mathbf{x}}_t^{(i)})}{q_t(\tilde{\mathbf{x}}_t | \mathbf{x}_{t-1}^{(i)})}$

fin pour

pour $i = 1$ à N **faire**

- Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{i=1}^N \tilde{\omega}_t^{(i)}}$

fin pour

- Calculer $N_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}$

si $N_{eff} \leq N_{seuil}$ **alors**

- Rééchantillonner les particules en utilisant l'algorithme 7

sinon

pour $i = 1$ à N **faire**

- Mettre à jour les particules $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)}$

- Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

fin pour

fin si

Ce choix est optimal au sens où cette densité minimise la variance des poids conditionnellement aux états passés $\mathbf{x}_{0:t-1}$ et aux observations $\mathbf{y}_{0:t}$. Cependant, il est rarement applicable en pratique, car en général il n'est pas toujours possible d'échantillonner selon la densité $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$ et/ou de calculer l'intégrale (2.46) qui est indispensable pour l'évaluation des poids.

- **Loi d'évolution de l'état :**

On choisit la loi d'évolution de l'état comme densité d'importance :

$$q_t(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) = p(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad (2.47)$$

Dans ce cas, les poids sont calculés comme suit :

$$\tilde{\omega}_t^{(i)} \propto \tilde{\omega}_{t-1}^{(i)} p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}) \quad (2.48)$$

C'est le choix adopté par Gordon *et al* [GSS93] et l'algorithme ainsi construit est appelé filtre bootstrap. Il a l'avantage d'être facile à implémenter. Cependant, il est souvent peu efficace vu que l'information apportée à l'instant t par la nouvelle observation \mathbf{y}_t n'est pas prise en compte pour la mise à jour des particules, ce qui se traduit par une grande variance des poids.

- **Pas de filtre de Kalman étendu (ou sans parfum) :**

La loi choisie pour la densité d'importance est une loi gaussienne

$$q_t(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}, y_t) = \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}^{(i)}, \Sigma_{t|t}^{(i)})$$

dont la moyenne $\hat{\mathbf{x}}_{t|t}^{(i)}$ et la matrice de covariance $\Sigma_{t|t}^{(i)}$ sont calculées par un pas d'EKF ou d'UKF en prenant comme statistiques au temps précédent $\hat{\mathbf{x}}_{t-1|t-1}^{(i)} = \mathbf{x}_{t-1}^{(i)}$ et $\Sigma_{t-1|t-1}^{(i)} = 0$. En procédant ainsi, on approche, en réalité, la densité d'importance optimale par une densité gaussienne, par linéarisation locale pour l'EKF et par l'utilisation d'une transformée sans parfum pour l'UKF. Ces deux choix sont efficaces, mais assez coûteux à mettre en œuvre en termes de nombre d'opérations, car à chaque instant t un pas de filtre de Kalman est réalisé pour chaque particule.

2.3.3 Filtre particulière Rao-Blackwellisé

Le principe de base de la technique de Rao-Blackwellisation [CR96] est d'exploiter la linéarité de certaines composantes du vecteur d'état du modèle, de façon à réduire la dimension de l'espace d'état. Ainsi, on aura besoin d'utiliser le filtrage particulière que sur la partie non linéaire. L'autre partie est traitée par filtrage de Kalman. L'estimation du vecteur d'état est alors effectuée en combinant un filtre particulière avec une banc de filtres de Kalman. Cette combinaison est appelée filtre particulière Rao-Blackwellisé [LC98, DGA00, CL00, AD00, SGN05]. Un tel filtre permet de réduire la variance de l'erreur et, par conséquent, d'améliorer l'efficacité de l'inférence, car une partie du problème est résolu de façon optimale [DGA00].

Rappelons le modèle considéré, décrit par l'équation (2.4) :

$$\begin{cases} \mathbf{x}_0 \sim p(\mathbf{x}_0) \\ \mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \mathbf{w}_t) & t \geq 1 \\ \mathbf{y}_t = g_t(\mathbf{x}_t, \mathbf{v}_t) & t \geq 1 \end{cases}$$

Dans certains cas, le vecteur d'état \mathbf{x}_t peut être décomposé en deux parties $\mathbf{x}_t = [\mathbf{x}_t^l \quad \mathbf{x}_t^n]^T$ de sorte que, conditionnellement à \mathbf{x}_t^n , le modèle est linéaire gaussien par rapport à \mathbf{x}_t^l . Le modèle dynamique initial (2.4) peut alors se réécrire sous la forme du système dynamique conditionnellement linéaire gaussien suivant :

$$\begin{cases} \mathbf{x}_t^n \sim p(\mathbf{x}_t^n | \mathbf{x}_{t-1}^n) \\ \mathbf{x}_t^l = \mathbf{F}_t(\mathbf{x}_{t-1}^n) \mathbf{x}_{t-1}^l + \mathbf{C}_t(\mathbf{x}_{t-1}^n) \mathbf{w}_t \\ \mathbf{y}_t = \mathbf{G}_t(\mathbf{x}_t^n) \mathbf{x}_t^l + \mathbf{D}_t(\mathbf{x}_t^n) \mathbf{v}_t \end{cases} \quad (2.49)$$

Etant donné le modèle statistique défini par l'équation (2.49), la densité *a posteriori* peut être décomposée comme suit :

$$p(\mathbf{x}_t^l, \mathbf{x}_{0:t}^n | \mathbf{y}_{1:t}) = p(\mathbf{x}_t^l | \mathbf{x}_{0:t}^n, \mathbf{y}_{1:t}) p(\mathbf{x}_{0:t}^n | \mathbf{y}_{1:t}) \quad (2.50)$$

La densité $p(\mathbf{x}_t^l | \mathbf{x}_{0:t}^n, \mathbf{y}_{1:t}) = \mathcal{N}(\mathbf{x}_t^l; \hat{\mathbf{x}}_{t|t}^l(\mathbf{x}_{0:t}^n), \Sigma_{t|t}(\mathbf{x}_{0:t}^n))$ peut se calculer à l'aide d'un filtre de Kalman. La densité $p(\mathbf{x}_{0:t}^n | \mathbf{y}_{1:t})$ va être approchée par filtrage particulière :

$$P_N(\mathbf{x}_{0:t}^n | \mathbf{y}_{1:t}) = \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta_{\tilde{\mathbf{x}}_{0:t}^{n,(i)}}(\mathbf{x}_{0:t}^n) \quad (2.51)$$

La densité marginale $p(\mathbf{x}_t^l | \mathbf{y}_{1:t})$ peut alors être approchée par un mélange de N distributions gaussiennes comme suit :

$$\begin{aligned}
p(\mathbf{x}_t^l | \mathbf{y}_{1:t}) &= \int p(\mathbf{x}_t^l | \mathbf{x}_{0:t}^n, \mathbf{y}_{1:t}) p(\mathbf{x}_{0:t}^n | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t}^n \\
&\simeq \int p(\mathbf{x}_t^l | \mathbf{x}_{0:t}^n, \mathbf{y}_{1:t}) \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta_{\tilde{\mathbf{x}}_{0:t}^{n,(i)}}(\mathbf{x}_{0:t}^n) d\mathbf{x}_{0:t}^n \\
&\simeq \sum_{i=1}^N \tilde{\omega}_t^{(i)} p(\mathbf{x}_t^l | \tilde{\mathbf{x}}_{0:t}^{n,(i)}, \mathbf{y}_{1:t}) \\
&\simeq \sum_{i=1}^N \tilde{\omega}_t^{(i)} \mathcal{N}(\mathbf{x}_t^l; \hat{\mathbf{x}}_{t|t}^l(\tilde{\mathbf{x}}_{0:t}^{n,(i)}), \Sigma_{t|t}(\tilde{\mathbf{x}}_{0:t}^{n,(i)}))
\end{aligned} \tag{2.52}$$

Dans ce cas, les poids sont mis à jour en utilisant la relation :

$$\tilde{\omega}_t^{(i)} \propto \tilde{\omega}_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t-1}) p(\tilde{\mathbf{x}}_t^{n,(i)} | \mathbf{x}_{0:t-1}^{n,(i)})}{q_t(\tilde{\mathbf{x}}_t^{n,(i)} | \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t})} \tag{2.53}$$

où la densité $p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t-1})$ est une distribution gaussienne

$$p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t-1}) = \mathcal{N}(\mathbf{y}_t; \hat{\mathbf{y}}_{t|t-1}(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)}), \mathbf{S}_t(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)})) \tag{2.54}$$

dont les paramètres $\hat{\mathbf{y}}_{t|t-1}(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)})$ et $\mathbf{S}_t(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)})$ sont respectivement la prédiction de l'observation et la covariance de l'innovation du filtre de Kalman (cf. algorithme 3).

L'algorithme 9 résume les différentes étapes du filtre particulaire Rao-Blackwellisé.

2.3.4 Méthodes de Monte Carlo par chaînes de Markov

Les méthodes MCMC sont un ensemble de méthodes consistant à simuler selon une densité de probabilité donnée π à partir de la construction d'une chaîne de Markov ergodique $\{\xi^{(i)}\}_{i=1}^N$ ayant pour loi stationnaire la densité d'intérêt π . La construction de cette chaîne se fait de manière itérative en se basant sur :

- un échantillon initial $\xi^{(0)} \sim \pi_0$;
- un noyau de transition $K(\xi | \xi')$.

Le noyau de transition doit vérifier certaines conditions [RC99]. En particulier, pour garantir que π est bien une loi stationnaire de la chaîne $\{\xi^{(i)}\}_{i=1}^N$, les algorithmes

Algorithme 9 Filtre Particulaire Rao-Blackwellisé

• A l'instant $t = 0$
pour $i = 1$ à N **faire**
 – Générer $\mathbf{x}_0^{n,(i)} \sim \pi_0(\mathbf{x}_0^n)$
 – Faire $(\tilde{\mathbf{x}}_{0|0}^l, \Sigma_{0|0}) = (\mathbf{x}_0^l, \Sigma_0)$
 – Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$
fin pour
 • A l'instant $t > 0$
pour $i = 1$ à N **faire**
 – Générer $\tilde{\mathbf{x}}_t^{n,(i)} \sim q_t(\mathbf{x}_t^n | \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t})$
 – Calculer $\tilde{\mathbf{x}}_{t|t}^{l,(i)}(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)})$ et $\Sigma_{t|t}^{(i)}(\tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)})$ à l'aide d'un pas de filtre de Kalman
 – Calculer les poids $\tilde{\omega}_t^{(i)} \propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{n,(i)}, \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t-1}) p(\tilde{\mathbf{x}}_t^{n,(i)} | \mathbf{x}_{0:t-1}^{n,(i)})}{q_t(\tilde{\mathbf{x}}_t^{n,(i)} | \mathbf{x}_{0:t-1}^{n,(i)}, \mathbf{y}_{1:t})}$
fin pour
pour $i = 1$ à N **faire**
 – Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^N \tilde{\omega}_t^{(j)}}$
fin pour
 – Calculer $N_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}$
si $N_{eff} \leq N_{seuil}$ **alors**
 – Rééchantillonner les particules en utilisant l'algorithme 7
sinon
 pour $i = 1$ à N **faire**
 – Mettre à jour les particules $\mathbf{x}_t^{n,(i)} = \tilde{\mathbf{x}}_t^{n,(i)}$
 – Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$
 fin pour
fin si

MCMC se fondent généralement sur une condition dite de balance détaillée :

$$\pi(\xi)K(\xi|\xi') = \pi(\xi')K(\xi'|\xi) \quad \forall \xi, \xi' \in \Psi \quad (2.55)$$

Principalement, on distingue deux grandes familles d'algorithmes MCMC qui permettent de construire un tel noyau de transition : l'algorithme de Metropolis-Hastings et l'échantillonneur de Gibbs.

2.3.4.1 Algorithme de Metropolis-Hastings

L'algorithme de Metropolis-Hastings a été initialement développé par Metropolis *et al* [MRR⁺53] en 1953 pour la physique particulaire. Il a été, ensuite, généralisé par Hastings [Has70] en 1970 dans un cadre plus statistique. Pour simuler selon une densité de probabilité donnée π (aussi appelée loi cible), l'algorithme de Metropolis-Hastings introduit une loi candidate r (aussi appelée loi instrumentale) vérifiant les conditions suivantes :

- la loi candidate est simulable directement ;
- le support de la loi candidate doit inclure le support de la loi cible : $r(\xi|\xi') \neq 0$ si $\pi(\xi) \neq 0$.

Les différentes étapes de l'échantillonnage par Metropolis-Hastings sont décrites par l'algorithme 10.

L'algorithme de Metropolis-Hasting peut être appliqué à une grande variété de problèmes car sa seule restriction est que la loi cible soit connue à une constante multiplicative près.

2.3.4.2 Echantillonneur de Gibbs

L'échantillonneur de Gibbs tire son nom des champs aléatoires de Gibbs où il a été utilisé pour la première fois par Geman et Geman [GG84]. Cette méthode est très pratique lorsque la loi cible π est définie sur un espace multidimensionnel. Supposons que $\Psi = \Psi_1 \times \dots \times \Psi_m$ est un espace produit et supposons que pour tout $\xi = (\xi_1, \dots, \xi_m) \in \Psi$, les densités conditionnelles $\pi_k(\xi_k|\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_m)$, $k = 1..m$ sont facilement simulables. L'échantillonneur de Gibbs consiste alors à échantillonner les composants ξ_k selon les densités conditionnelles. L'algorithme 11 décrit ce principe.

L'échantillonneur de Gibbs ne peut s'appliquer que lorsque la dimension de l'es-

Algorithme 10 Algorithme de Metropolis-Hastings

• Initialisation

– Initialiser $\xi^{(0)} \sim \pi_0(\xi)$

• Itérations

pour $i = 1$ à N **faire**– Echantillonner $\xi^* \sim r(\xi|\xi^{(i-1)})$

– Calculer la probabilité d'acceptation

$$\rho(\xi^{(i-1)}, \xi^*) = \min \left\{ 1, \frac{\pi(\xi^*)r(\xi^{(i-1)}|\xi)}{\pi(\xi^{(i-1)})r(\xi|\xi^{(i-1)})} \right\}$$

– Acceptation-Rejet

$$\xi^{(i)} = \begin{cases} \xi^* & \text{avec la probabilité } \rho(\xi^{(i-1)}, \xi^*) \\ \xi^{(i-1)} & \text{avec la probabilité } 1 - \rho(\xi^{(i-1)}, \xi^*) \end{cases}$$

fin pour

Algorithme 11 Echantillonneur de Gibbs

• Initialisation

– Initialiser $\xi^{(0)} \sim (\xi_1^{(0)}, \dots, \xi_m^{(0)})$

• Itérations

pour $i = 1$ à N **faire****pour** $k = 1$ à m **faire**– Echantillonner $\xi_k^{(i)} \sim \pi_k(\xi_k|\xi_1^{(i)}, \dots, \xi_{k-1}^{(i)}, \xi_{k+1}^{(i-1)}, \xi_m^{(i-1)})$ **fin pour****fin pour**

pace est multidimensionnel. De plus, il nécessite la détermination explicite des lois conditionnelles.

2.4 Inférence bayésienne dans le cas des distributions α -stables

Dans le paragraphe 2.1 a été défini le modèle dynamique composé du modèle d'évolution de l'état et du modèle d'observation. Des méthodes déterministes et de Monte Carlo permettant de réaliser l'inférence bayésienne dans le cadre d'un tel modèle dynamique ont été ensuite présentées dans les paragraphes 2.2 et 2.3. Notre travail de thèse est axé particulièrement sur l'estimation des signaux affectés par des bruits additifs impulsifs modélisés par des lois α -stables. Le modèle dynamique (2.4) sera alors modifié en considérant que le bruit d'observation est distribué selon une distribution α -stable. La modification apportée à notre modèle bien que mineure fait toute la différence et réaliser une inférence bayésienne pour ce type de modèle est un problème complexe. En effet, lorsque les bruits sont modélisés par des distributions α -stables, l'application directe des méthodes classiques présentées dans les paragraphes 2.2 et 2.3 n'est plus possible, car la densité de probabilité n'a pas d'expression analytique explicite et ses moments d'ordre supérieur ou égal à 2 sont infinis. Une solution évidente à ce problème consiste à appliquer directement les méthodes de Monte Carlo en calculant la densité de probabilité numériquement à l'aide de la technique proposée par Nolan [Nol97, Nol99]. Cependant, une telle solution risque d'être très complexe en temps de calcul. La bibliographie fait état de plusieurs travaux dont l'objectif est d'apporter des solutions efficaces à ce problème. Ces travaux peuvent être répartis en trois catégories : des travaux s'intéressant seulement à l'estimation des paramètres de la distribution α -stable [Buc95, Lom07, Tsi99, God99, GK99, God00], des travaux qui s'attachent à l'estimation conjointe du signal et des paramètres du bruit α -stable [LG04, LG06, VVFPCA FR11, MBA⁺05] et finalement des travaux qui traitent l'estimation d'un mélange fini de distributions α -stables [Cas04, SGKR09, SGKR06, SGKR10].

2.4.1 Estimation bayésienne des paramètres de la distribution α -stable

Parmi les travaux pionniers dans ce contexte, nous citons celui de Buckle [Buc95] basé sur les méthodes MCMC. Dans cette contribution, l'auteur considère un vecteur d'observations $\mathbf{y}_{1:n} = \{\mathbf{y}_i\}_{i=1}^n$ distribuées selon une distribution α -stable

$$\mathbf{y}_i \sim S_2(\alpha, \beta, \gamma, \mu), \quad i = 1, \dots, n \quad (2.56)$$

et cherche à estimer dans un cadre bayésien les paramètres de la distribution α -stable. Pour ce faire, il tire profit d'une propriété mathématique qui permet d'exprimer analytiquement la densité de probabilité de $\mathbf{y}_{1:n}$ conditionnellement à une variable auxiliaire et met en place un échantillonneur de Gibbs pour l'estimation des différents paramètres. Cette approche présente toutefois quelques difficultés. L'échantillonnage selon la densité de probabilité de la variable auxiliaire n'est pas simple et nécessite le recours à la méthode d'acceptation/rejet. De plus, à chaque itération de l'algorithme, il faut simuler un échantillon de même taille que celle du vecteur d'observations, aboutissant ainsi à un algorithme très coûteux en temps de calcul, en particulier lorsqu'il s'agit d'un vecteur d'observations de grande taille.

Dans ce même contexte, Lombardi [Lom07] propose une méthode basée sur l'algorithme de Metropolis-Hastings à marche aléatoire pour l'estimation des paramètres d'une distribution α -stable. Cette approche repose sur l'approximation de la fonction de vraisemblance à l'aide d'une méthode introduite par Mittinink *et al* [MDC99]. Cette méthode consiste à évaluer la fonction de vraisemblance pour un ensemble donné de points en appliquant la transformée de Fourier rapide inverse à la fonction caractéristique. On obtient ainsi la valeur exacte de la fonction de vraisemblance pour chaque point de l'ensemble considéré. Les valeurs de la fonction de vraisemblance pour les observations se situant entre les points de l'ensemble de départ sont ensuite obtenues par interpolation linéaire. Cependant, cette méthode présente un inconvénient majeur. En effet, les points pour lesquels la transformée de Fourier rapide inverse est calculé, doivent être équidistants. Par conséquent, pour pouvoir atteindre les observations se trouvant sur les extrémités des queues de distribution, on est obligé de considérer un ensemble de départ avec un grand nombre de points, ce qui aboutit à un algorithme très coûteux en temps de calcul. Pour pallier à ce

problème, Lombardi n'applique la transformée de Fourier rapide inverse que sur un intervalle susceptible de contenir une grande majorité des observations et pour les points se situant en dehors de cet intervalle, il utilise le développement en séries entières de Bergström [Ber52] (cf. paragraphe 1.1.6.2).

Des travaux similaires ont été menés par Tsionas [Tsi99] et Godsill *et al* [God99, GK99, God00] pour considérer cette problématique dans les processus autorégressifs (AR, Auto-Regressive) avec des bruit additifs α -stable symétriques. Tsionas propose une approche MCMC basée sur l'échantillonnage de Gibbs afin d'estimer les paramètres des perturbations modélisées par les distributions α -stable symétriques dans des modèles économiques. Godsill *et al* propose une méthode MCEM (Monte Carlo Expectation Maximization) pour l'estimation conjointe des paramètres du processus AR et ceux du bruit α -stable. Ces méthodes reposent sur la décomposition de la distribution α -stable symétrique en SMiN. Une telle décomposition permet d'exprimer une variable α -stable symétrique v_t comme le produit d'une variable gaussienne et d'une variable α -stable positive :

$$v_t \sim S_1(\alpha, 0, \gamma, 0) \rightarrow v_t \sim \mathcal{N}(0, \lambda_t \gamma^2), \quad \lambda_t \sim S_1\left(\frac{\alpha}{2}, 1, 2\left(\cos \frac{\pi\alpha}{4}\right)^{\frac{2}{\alpha}}, 0\right) \quad (2.57)$$

Cette écriture est très utile en pratique vu que conditionnellement à λ_t et γ , la variable α -stable v_t est gaussienne et par conséquent les méthodes classiques peuvent être appliquées.

Plus récemment, Gencaga *et al* [GEK08] généralisent les approches introduites par Tsionas et Godsill *et al* pour les processus autorégressifs variants dans le temps (TVAR, Time-Varying Auto-Regressive). La méthode proposée utilise un filtre particulaire pour l'estimation des paramètres du processus TVAR et un échantillonneur de Gibbs pour l'estimation des paramètres de la distribution α -stable.

Des travaux ont été également réalisés pour traiter cette problématique dans le cas où les paramètres de la distribution α -stable varient dans le temps [GEK08, HZK11]. Dans [GEK08], les auteurs considèrent un processus TVAR avec un bruit additif α -stable et proposent un filtre Bootstrap pour l'estimation conjointe des paramètres du processus TVAR et ceux du bruit α -stable. Une marche aléatoire est utilisée pour modéliser l'évolution au cours du temps de ces différents paramètres. Une approche similaire a été proposé par Huang *et al* [HZK11] dans le cas des processus autorégressifs variants dans le temps à moyenne mobile (TV ARMA, Time-Varying

Auto-Regressive Moving Average). Ces approches sont toutefois très coûteuses en temps de calcul vu qu'à chaque itération du filtre particulaire et pour chacune des particules, la vraisemblance est calculée en utilisant la technique proposée par Nolan [Nol97, Nol99].

2.4.2 Estimation bayésienne conjointe du signal et des paramètres du bruit α -stable

Il existe à notre connaissance peu de travaux dans la littérature traitant le problème de l'estimation conjointe de l'état et des paramètres du bruit dans les modèles dynamiques linéaires ou non linéaires en présence d'un bruit α -stable. De plus, la plupart des études menées considèrent le cas particulier du bruit α -stable symétrique. Parmi ces travaux, nous citons ceux de Lombardi et Godsill [LG04, LG06] qui proposent une approche de filtrage bayésien dans les modèles dynamiques linéaires avec un bruit d'observation α -stable symétrique. La méthode proposée repose aussi sur la décomposition de la distribution α -stable symétrique en SMiN. Grâce à cette décomposition, un modèle linéaire conditionnellement Gaussien est obtenu. Les auteurs mettent à profit cette structure avantageuse du modèle et utilisent un filtre particulaire Rao-Blackwellisé.

Cette approche a été récemment étendue dans [VVFPCA^{FR}11] au cas des modèles non linéaires en utilisant une méthode hybride UKF/SMC. En effet, un UKF est mis en oeuvre pour l'estimation de l'état. Ensuite, l'estimation des paramètres du bruit α -stable est réalisée avec un filtre particulaire.

Un autre travail a été également proposé pour traiter cette même problématique dans le cas où c'est le bruit d'état qui est α -stable symétrique [MBA⁺05]. Pour ce faire, les auteurs utilisent un filtre Bootstrap en choisissant comme densité d'importance pour les particules d'état une distribution α -stable symétrique.

2.4.3 Estimation bayésienne des mélanges finis de distributions α -stables

Le problème d'estimation d'un mélange fini de distributions α -stable a été initialement introduit par Casarin [Cas04]. Dans cette contribution, l'auteur utilise la méthode proposée par Buckle [Buc95] pour l'estimation des paramètres de chaque

distribution en supposant que le nombre de composantes du mélange est connu. Une autre contribution a été proposée par Gonzalez *et al* [SGKR09]. Dans ce travail, l'estimation des paramètres de chaque composante du mélange est réalisée en utilisant la méthode introduite dans [Lom07]. Le nombre de composantes du mélange est supposé inconnu et il est estimé en utilisant les méthodes MCMC à sauts réversibles [GRE95]. Des travaux similaires ont été également proposés dans le cas d'un mélange fini de distributions α -stable symétriques [SGKR06, SGKR10]. Les auteurs exploitent la décomposition de la distribution α -stable en SMiN et obtiennent ainsi un mélange fini de gaussiennes. La méthode proposée est similaire à celle présentée dans [RG97] pour l'estimation d'un mélange fini de gaussiennes. La variable α -stable positive est estimée en utilisant l'approche de Godsill *et al* [GK99].

2.5 Conclusion

Dans ce chapitre, nous avons présenté les bases de l'inférence bayésienne dans les modèles de Markov cachés. En effet, l'approche bayésienne offre un cadre cohérent, complet et exhaustif pour mener une inférence lorsque les informations disponibles sont entachées d'incertitude. Elle repose essentiellement sur la détermination de la densité *a posteriori*. Cette densité rassemble toute l'information véhiculée par les observations sur l'état caché. Lorsque le modèle est linéaire gaussien, une solution optimale est fournie par le filtre de Kalman. Sinon, lorsque que le modèle est faiblement non linéaire, des filtres non optimaux déterministes basés sur le filtre de Kalman peuvent être utilisés. Dans le cas général, une approximation numérique de cette densité peut être obtenue en utilisant les méthodes de Monte Carlo. Dans cette thèse, nous nous intéressons à l'estimation des signaux affectés par des bruits additifs impulsifs modélisés par des lois α -stables. Dans ce cas, les mécanismes classiques de l'inférence bayésienne ne peuvent pas être directement appliqués car la densité de probabilité des distributions α -stables n'a pas d'expression analytique explicite et ses moments d'ordre supérieur ou égal à 2 sont infinis. Dans la littérature, de nombreuses études ont été menées afin d'apporter des solutions à ce problème. Cependant, la plupart des travaux proposés considèrent le cas particulier des distributions α -stables symétriques. Les travaux traitant cette problématique dans le cas général font appel à des techniques supplémentaires pour évaluer numériquement la

pdf des distributions α -stables, aboutissant ainsi à des algorithmes très complexes en temps de calcul. Afin de contribuer à la résolution de ce problème, des approches paramétriques et non paramétriques basées sur les méthodes de Monte Carlo séquentielles sont proposées dans les chapitres qui suivent.

Estimation conjointe état-bruit dans les systèmes OFDM

LE développement des systèmes de communication numérique a conduit à l'augmentation des besoins en débits. Dans ce contexte, les systèmes de transmission multi-porteuses ont suscité un vif intérêt car ils répondent à ces besoins tout en garantissant une grande efficacité spectrale et une excellente robustesse vis-à-vis des canaux sélectifs en fréquence.

La technique de multiplexage par division en fréquences orthogonales (OFDM, Orthogonal Frequency Division Multiplexing) fait partie des systèmes de transmission multi-porteuses. Elle a été introduite pour la première fois en 1966 par Chang [Cha66]. Délaissée ensuite à cause de sa grande complexité d'implémentation, l'OFDM dû son retour en grâce, vers le milieu des années 1980, aux travaux présentés par Weinstein et Ebert [WE71] qui proposent une modélisation discrète des systèmes multi-porteuses basée sur la transformée de Fourier discrète (DFT, Discrete Fourier Transform) et la transformée de Fourier discrète inverse (IDFT, Inverse Discrete Fourier Transform). Depuis lors, l'OFDM a connu un succès considérable et, de nos jours cette technique est présente dans de nombreux systèmes de communication filaire ou sans fil. En particulier, on retrouve cette technique de modulation dans les normes de la radiodiffusion numérique (DAB, Digital Audio Broadcasting), de la télévision numérique terrestre (DVB-T, Digital Video Broadcasting-Terrestrial), de communications numériques hauts débits filaires (DSL, Digital Subscriber Line), ainsi que dans les réseaux sans fil tels que la norme américaine IEEE-802.11 (WiFi)

et son équivalent européen HiperLAN/2. Elle a été également adoptée dans les standards d'accès sans fil à bande large tel que la norme 802.16 (WiMAX) d'IEEE.

Malheureusement, la plupart des systèmes de communication qu'ils soient filaires ou sans fil, sont soumis à de nombreuses dégradations, entre autres les bruits impulsifs. Les techniques de transmission multi-porteuses de type OFDM sont généralement robustes vis-à-vis de ce type de perturbations grâce à la DFT qui permet de répartir l'énergie du bruit impulsif entre les différentes sous-porteuses. Cependant, lorsque cette énergie dépasse un certain seuil, la présence du bruit impulsif peut considérablement affecter la qualité de transmission [Gho96]. De nombreuses études ont été menées pour analyser l'effet du bruit impulsif sur une transmission OFDM dans le cadre de certains systèmes de communications à savoir : PLC [ANK06], DVB-T [Wu99, WPC⁺00], DSL [CANN08], etc. Ces travaux viennent tous confirmer que les bruits impulsifs réduisent considérablement les performances des systèmes de communication à moins qu'une technique efficace de compensation soit intégrée dans la chaîne de transmission. Dans la littérature, plusieurs approches ont été proposées pour faire face aux effets néfastes du bruit impulsif dans les systèmes OFDM [Zhi08, MUKM02, Zhi03, HFII06, Wol83, ADA05, ADAA08, CANN08, HV03].

Dans ce chapitre, nous proposons une méthode de compensation de l'effet des perturbations de nature impulsive dans une transmission OFDM. L'approche proposée repose sur la modélisation du bruit impulsif par une distribution α -stable symétrique. Différents facteurs nous ont amené à adopter cette modélisation. Tout d'abord la grande variabilité du bruit en question qui n'est pas compatible avec les distributions gaussiennes mais, au contraire, est très bien représentée par des distributions α -stables. Ensuite, les propriétés mêmes des distributions α -stables et le fait que le cas gaussien n'en soit qu'un cas particulier, sont des avantages de cette modélisation. De plus, ce choix a été justifié théoriquement et/ou expérimentalement dans le cadre de plusieurs systèmes de communication. En effet, de nombreuses études ont montré que le MAI dans les réseaux *ad hoc* suit une distribution α -stable [Sou92, IH96, PCG⁺06]. En outre, Nassar *et al* [NGS⁺08] ont prouvé expérimentalement que cette distribution fournit un modèle approprié pour l'interférence radiofréquence (RFI, Radio Frequency Interference) dans la communication sans fil entre les émetteur-récepteur intégrés dans les ordinateurs portables. De plus, Win *et al* [WPS09] présente un cadre général pour la modélisation de l'interférence dans les réseaux sans fil aboutissant à une distribution α -stable. Enfin, dans plu-

sieurs travaux [JDV⁺11, TN95b, AIH94, EGCA⁺10], les distributions α -stables ont été adopté comme modèle des bruits à caractère impulsif rencontrés dans le contexte d'une transmission OFDM.

L'approche que nous proposons pour la réduction du bruit impulsif dans les systèmes OFDM repose sur l'inférence bayésienne et plus particulièrement les méthodes de Monte Carlo séquentielles et elle permet l'estimation conjointe du signal OFDM et des paramètres du bruit impulsif. Cette approche s'appuie sur les travaux de Septier *et al* [SDMRG10, SDMRG08, Sep08] dans lesquels une modélisation originale du signal OFDM dans le domaine temporel par processus autorégressif a été introduite, ainsi que ceux de Lombardi et Godsill [LG04, LG06] qui traitent le problème de l'estimation conjointe de l'état et des paramètres du bruit dans les modèles dynamiques linéaire en présence d'un bruit d'observation α -stable symétrique.

Ce chapitre est organisé comme suit. Tout d'abord, nous présentons les systèmes de transmission OFDM. Ensuite, nous dressons un état de l'art sur les approches existantes traitant le problème de réduction du bruit impulsif dans le cadre de ces systèmes. Après avoir introduit le modèle dynamique, nous décrivons le filtre particulier proposé pour l'estimation conjointe des symboles OFDM émis et des paramètres du bruit α -stable symétrique. Enfin, des résultats sont donnés afin d'illustrer la validité de notre approche. De nombreuses simulations sont effectuées et les performances de notre algorithme sont comparées à celles des méthodes existantes.

3.1 Les systèmes OFDM

3.1.1 Principe de l'OFDM

L'OFDM fait partie des techniques de modulation multi-porteuses. Ces techniques consistent à transmettre des données numériques en parallèle, à l'opposé des systèmes de transmission conventionnels qui transmettent les données séquentiellement, le spectre de chaque donnée occupe alors toute la bande passante allouée à la transmission. Le principe des techniques de modulation multi-porteuses est de diviser la bande de fréquences du canal de transmission en plusieurs sous-bandes, appelées aussi sous-porteuses. La bande passante allouée à la transmission est alors partagée entre les différentes sous-porteuses. Ainsi, chaque sous-porteuse occupe une bande de

fréquences de fréquence centrale différente. Pour assurer une grande efficacité spectrale¹, l'écart entre les fréquences centrales des sous-porteuses doit être le plus petit possible. Cependant, avec les techniques de modulation multi-porteuses classiques, cet écart est choisi suffisamment grand pour éviter les interférences entre les spectres des signaux émis par les différentes sous-porteuses et permettre ainsi au récepteur de les séparer et de récupérer le signal émis par chaque sous-porteuse. La spécificité de l'OFDM par rapport aux autres techniques de modulation multi-porteuses réside dans le fait qu'elle autorise un recouvrement entre les sous-porteuses, ce qui permet d'augmenter l'efficacité spectrale. Cependant, les sous-porteuses doivent respecter une contrainte d'orthogonalité. Cette contrainte est vérifiée lorsque le spectre d'une sous-porteuse est nul aux fréquences centrales des autres sous-porteuses. Par conséquent, les signaux des différentes sous-porteuses se chevauchent mais, grâce à l'orthogonalité, n'interfèrent pas.

La mise en place des systèmes OFDM présente une difficulté majeure qui réside dans la nécessité de disposer d'un banc d'oscillateurs pour générer les sous-porteuses. En raison de leur complexité de mise en œuvre, les systèmes OFDM ont été abandonnés pendant plusieurs années. En 1971, Weinstein et Albert [WE71] proposent une méthode très simple pour la mise en place d'un système OFDM basée sur l'utilisation de la DFT et de l>IDFT pour mettre en application les processus de modulation et de démodulation. Pour réduire la complexité, les algorithmes de la transformée de Fourier rapide (FFT, Fast Fourier Transform) et de la transformée de Fourier rapide inverse (IFFT, Inverse Fast Fourier Transform) sont souvent utilisés en pratique.

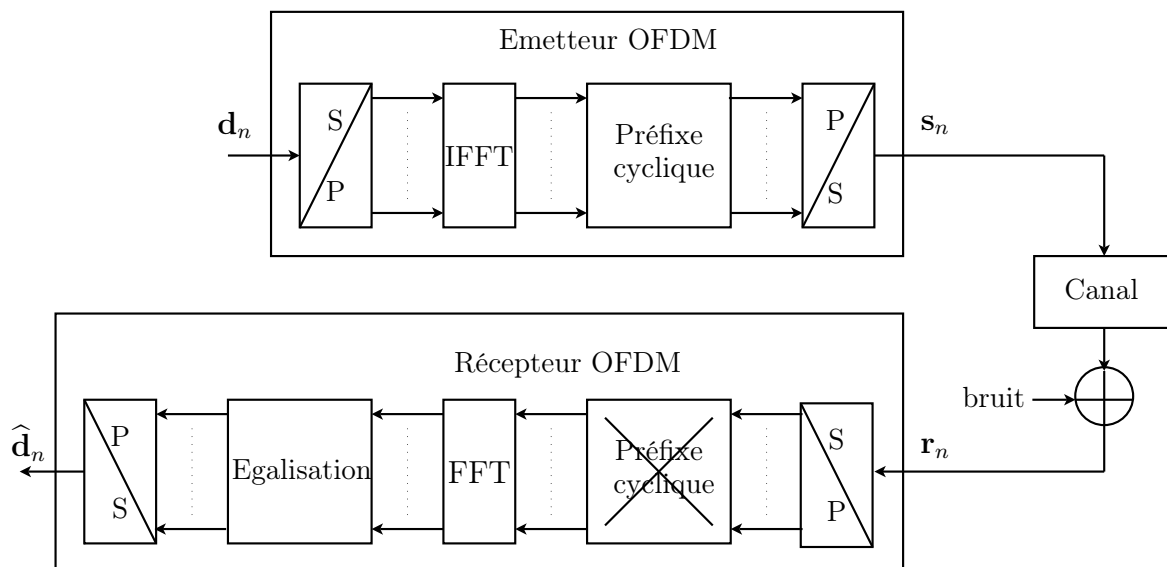
3.1.2 Implémentation numérique d'un système OFDM

La technique OFDM consiste à transmettre un flux de données en parallèle modulées sur plusieurs sous-porteuses. Nous notons N le nombre de sous-porteuses utilisées pour la transmission. Le schéma de principe d'une chaîne de transmission OFDM est donné figure 3.1.

L'émetteur OFDM commence par effectuer le codage d'un flux binaire délivré par une source d'informations. Un flux de symboles de données ayant des valeurs complexes est alors obtenu. Plus précisément, ces symboles de données sont des

1. L'efficacité spectrale est définie comme étant le débit binaire par unité de fréquence. Elle s'exprime en bit/s/Hz.

FIGURE 3.1 Schéma bloc d'un système OFDM



symboles q -aires formés par groupement de q bits et dont les valeurs appartiennent à un alphabet de taille 2^q . On note T_s la durée de chacun de ces symboles. Suivant le nombre de bits regroupés ensemble, on distingue plusieurs types de modulation : modulation par déplacement de phase à 2 états (BPSK, Binary Phase Shift Keying), modulation par déplacement de phase à 4 états (QPSK, Quadrature Phase Shift Keying), modulation d'amplitude en quadrature à 16 états (16-QAM, 16-Quadrature Amplitude Modulation), modulation d'amplitude en quadrature à 64 états (64-QAM, 64-Quadrature Amplitude Modulation), etc. La table 3.1 donne les différents types de modulation.

TABLE 3.1 – Types de modulation

Modulation	Nombre de bits
BPSK	1
QPSK	2
16-QAM	4
64-QAM	6

Les symboles de données sont regroupés par blocs de N . L'émetteur OFDM

effectue alors la conversion série-parallèle de chaque bloc de N symboles de données. Le $n^{\text{ième}}$ bloc est défini comme suit :

$$\mathbf{d}_n = \left[d_{n,0} \quad \cdots \quad d_{n,N-1} \right]^T \quad (3.1)$$

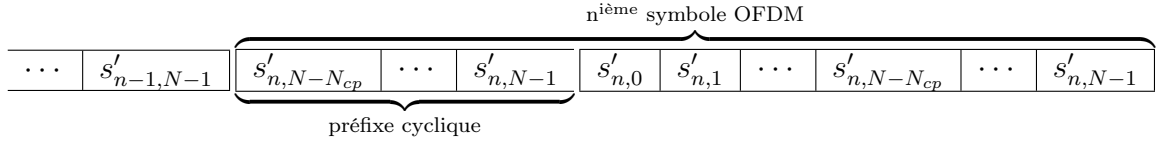
où $d_{n,i}$ désigne le symbole de données du $n^{\text{ième}}$ bloc qui sera modulée sur la $i^{\text{ième}}$ sous-porteuse. La durée d'un bloc est N fois plus grande que celle d'un symbole de donnée en série : $T = NT_s$. Ensuite, une IFFT fournit le symbole OFDM dans le domaine temporel. Celui-ci est constitué de N échantillons $\{s'_{n,t}\}_{t=1}^{N-1}$. $s'_{n,t}$ est le $t^{\text{ième}}$ échantillon du $n^{\text{ième}}$ symbole OFDM et il s'exprime comme suit :

$$s'_{n,t} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} d_{n,i} \exp\left(\frac{j2\pi it}{N}\right) \quad \text{pour } t = 0, \dots, N-1 \quad (3.2)$$

Nous considérons un canal de transmission $h(t, \tau)$ sélectif en fréquence variant dans le temps à trajets multiples. Nous notons L le nombre de ses trajets de propagation. Nous supposons que le canal est invariant durant la transmission d'un symbole OFDM et nous notons $h_{n,l}, l = 0, \dots, L-1$ la réponse impulsionnelle de son $l^{\text{ième}}$ trajet de propagation pour un symbole OFDM donné n . A cause du comportement multi-trajets du canal de transmission, le symbole OFDM transmis prend différents retards pour arriver au récepteur par différents chemins de propagation. Du point de vue récepteur, le canal présente une dispersion temporelle dans laquelle la durée du symbole reçu est étalée. Cette prolongation de la durée du symbole est à l'origine du phénomène de l'interférence entre symboles (ISI, Inter-Symbol Interference) qui se manifeste par le chevauchement du symbole OFDM courant avec le symbole OFDM précédent. Afin de remédier au problème de l'ISI, un préfixe cyclique de longueur N_{cp} est inséré entre deux symboles OFDM consécutifs. En fait, il consiste à recopier les N_{cp} derniers échantillons du symbole OFDM en début de symbole, comme illustré dans la figure 3.2. Le signal OFDM obtenu est constitué de $N + N_{cp}$ échantillons $\{s_{n,t}\}_{t=0}^{N+N_{cp}-1}$ qui sont définis comme suit :

$$s_{n,t} = \begin{cases} s'_{n,N-N_{cp}+t} & \text{si } 0 \leq t \leq N_{cp} - 1 \\ s'_{n,t-N_{cp}} & \text{si } N_{cp} \leq t \leq N + N_{cp} - 1 \end{cases} \quad (3.3)$$

Sa durée totale est égale à $T' = T + T_{cp}$, où T_{cp} est la durée du préfixe cyclique. Pour

FIGURE 3.2 Insertion du préfixe cyclique.

que l'ISI soit entièrement supprimé, la durée du préfixe cyclique doit être supérieure au retard maximum τ_{max} apporté par le canal :

$$T_{cp} > \tau_{max} \quad (3.4)$$

Ainsi, les derniers échos du symbole OFDM précédent auront lieu durant cet intervalle et le symbole OFDM courant ne sera plus perturbé.

Du côté du récepteur OFDM, les opérations inverses sont réalisées. Le récepteur commence par effectuer la conversion série parallèle du signal reçu. Le vecteur obtenu, noté \mathbf{r}_n , s'exprime comme suit :

$$\underbrace{\begin{bmatrix} r_{n, N+N_{cp}-1} \\ r_{n, N+N_{cp}-2} \\ \vdots \\ r_{n, N_{cp}} \\ \vdots \\ r_{n, 0} \end{bmatrix}}_{\mathbf{r}_n} = \underbrace{\begin{bmatrix} h_{n,0} & h_{n,1} & \cdots & h_{n,L-1} & 0 & \cdots & 0 \\ 0 & h_{n,0} & h_{n,1} & \cdots & h_{n,L-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & h_{n,0} & h_{n,1} & \cdots & h_{n,L-1} \end{bmatrix}}_{\mathbf{H}_n} \underbrace{\begin{bmatrix} s_{n, N+N_{cp}-1} \\ s_{n, N+N_{cp}-2} \\ \vdots \\ s_{n, N_{cp}} \\ \vdots \\ s_{n, 0} \\ s_{n-1, N+N_{cp}-1} \\ \vdots \\ s_{n-1, N+N_{cp}-L+1} \end{bmatrix}}_{\mathbf{s}_n} + \mathbf{w}_n \quad (3.5)$$

où \mathbf{r}_n , \mathbf{s}_n et \mathbf{w}_n sont des vecteurs de tailles respectives $(N+N_{cp}, 1)$, $(N+N_{cp}+L-1, 1)$, $(N+N_{cp}, 1)$ et \mathbf{H}_n est une matrice de taille $(N+N_{cp}, N+N_{cp}+L-1)$. Le vecteur \mathbf{w}_n correspond au bruit additif qui permet de prendre en compte les différentes perturbations liées au canal de transmission. Après la conversion série parallèle, le récepteur procède à la suppression du préfixe cyclique. Ensuite, une FFT est réalisée afin de récupérer les N échantillons émis. Afin de corriger les distorsions apportées sur le canal de transmission, une opération d'égalisation est effectuée. Enfin, après la conversion parallèle série, une opération de décodage permet de restituer les symboles

d'origine.

Généralement, le bruit \mathbf{w}_n est supposé être un bruit blanc gaussien circulaire. Cette hypothèse permet de simplifier les différents traitements et d'obtenir des solutions analytiques. Cependant, elle est mal adaptée pour la modélisation des phénomènes qui présentent une grande variabilité comme les bruits impulsifs. Or, dans les systèmes de communication, de nombreuses perturbations rencontrées sont de nature impulsive. Citons par exemple les bruits rencontrés sur le réseau électrique [YES07, ZD02, MSG05, CD89] ou dans les systèmes UMTS [SAC04], dans les systèmes UWB [PCG⁺06], ainsi que les interférences présentes dans les réseaux *ad hoc* [Sou92, IH96]. Dans ces systèmes, le signal émis peut être perturbé par un bruit de type impulsif de forte amplitude en présence d'un bruit gaussien beaucoup plus faible. Dans cette thèse nous supposons que le bruit \mathbf{w}_n admet deux composantes, une composante gaussienne \mathbf{v}_n et une autre impulsive \mathbf{u}_n :

$$\mathbf{w}_n = \mathbf{v}_n + \mathbf{u}_n \quad (3.6)$$

3.1.3 Etat de l'art sur les méthodes de réduction du bruit impulsif dans les systèmes OFDM

3.1.3.1 Méthodes de traitement du signal

Plusieurs méthodes ont été proposées dans la littérature afin de limiter les erreurs engendrées par les bruits impulsifs dans le contexte d'une transmission OFDM. Ces méthodes peuvent être divisées en trois catégories, les méthodes qui agissent dans le domaine temporel, les approches fréquentielles qui traitent le bruit impulsif dans le domaine fréquentiel et les méthodes temps-fréquence.

Des approches classiques, souvent utilisées dans la littérature, sont le Blanking et le Clipping. Leur principe est simple et consiste à annuler ou limiter les échantillons du signal dont l'amplitude dépasse un certain seuil.

- Blanking

$$r_t = \begin{cases} r_t, & \text{if } |r_t| < T_b \\ 0, & \text{sinon} \end{cases}$$

- Clipping

$$r_t = \begin{cases} r_t, & \text{if } |r_t| < T_c \\ T_c \exp(j \arg(r_t)), & \text{sinon} \end{cases}$$

Dans [Zhi08], Zhidkov propose une contribution qui consiste à déterminer les seuils optimaux de ces deux méthodes en vue d'améliorer leurs performances.

Les méthodes Blanking et Clipping réduisent l'effet du bruit impulsif mais reste sous-optimales. En plus, la procédure de détection de l'impulsion est souvent entachée d'erreurs.

Dans [MUKM02], Matsuo *et al* proposent une méthode utilisant un module de détection pour l'atténuation des bruits impulsifs dans les systèmes OFDM. L'approche consiste à déterminer dans le domaine temporel les amplitudes les plus importantes du bruit grâce à un procédé itératif, puis à les soustraire du signal reçu. D'autres algorithmes similaires ont été proposés dans la littérature. Dans [HFII06], le seuil utilisé pour la sélection des amplitudes importantes est mis à jour dans chaque itération, et dans [Zhi03] l'algorithme de Matsuo *et al* est appliqué dans le domaine fréquentiel.

L'algorithme de Matsuo et ses optimisations offrent de bonnes performances. Cependant, il faut noter que ces performances se font au détriment d'une complexité en temps de calcul accrue. En plus, cet algorithme n'est pas adapté aux bruits très impulsifs vu que la puissance totale d'impulsion est répartie en fréquence sur l'ensemble des sous porteuses, et la décision du maximum de vraisemblance devient par conséquent moins fiable.

3.1.3.2 Codage de canal

Une autre solution qui permet de corriger les erreurs engendrées par le bruit impulsif consiste à effectuer un codage de canal. Plusieurs approches, faisant intervenir un codage du canal, ont été proposées dans la littérature pour la réduction du bruit impulsif dans une transmission OFDM [Wol83, ADA05, ADAA08, CANN08]. La plupart de ces travaux est basée sur l'utilisation de la similitude entre la transformation de Fourier dans un système OFDM et le codage Reed-Solomon (RS) dans le corps des nombres complexes. Plus précisément, un système de modulation élargi par des zéros et faisant appel à l>IDFT, tel que l'OFDM, peut être considéré comme

un codeur RS et le récepteur correspondant peut contenir un module supplémentaire de décodage RS avec localisation et correction d'erreurs dans le domaine fréquentiel. En effet, le code correcteur d'erreurs, introduit à l'émission avant le modulateur OFDM, permet de détecter et de corriger les erreurs qui se produisent lors d'une transmission, grâce à une redondance de l'information.

Parmi les travaux pionniers dans ce contexte, nous citons celui de Wolf [Wol83]. Dans cette contribution, l'opération de décodage est basée sur l'utilisation des sous-porteuses d'amplitude nulle considérées comme syndromes. Ces sous-porteuses ne contiennent à la réception que des composantes de bruit transformées dans le domaine fréquentiel et peuvent servir à déterminer le reste des composantes inconnues au niveau des sous-porteuses d'information. Ces composantes seront par la suite soustraites du signal reçu afin d'extraire le signal non bruité.

Récemment, des études similaires ont été menées par Abdelkefi *et al* [ADA05, ADAA08] pour la compensation du bruit impulsif dans les systèmes OFDM. L'approche proposée dans [ADA05] repose sur des techniques semblables à celles utilisées par Wolf [Wol83]. Cependant, ces techniques sont appliquées dans un contexte plus général où les symboles pilotes ne sont pas répartis de manière consécutive. Dans [ADAA08], les auteurs proposent une structure en cascade en vue d'améliorer leur ancien algorithme présenté dans [ADA05]. Cette nouvelle structure permet une implantation plus facile et plus fiable.

Dans [CANN08], Caire *et al* utilisent des techniques de programmation convexe, plus précisément le "Compressed Sensing" (CS) qui permettent de reconstruire un sparse signal en utilisant des projections linéaires sur un espace de petite dimension. Le CS peut être mis en œuvre pour l'estimation du bruit impulsif au niveau du récepteur grâce à la rareté du bruit en question dans le domaine temporel. Dans la méthode proposée, les porteuses pilotes sont utilisées pour estimer le bruit impulsif.

Bien que les techniques de décodage puissent aboutir à une réduction significative des effets du bruit impulsif, elles présentent quelques inconvénients. En effet, ces algorithmes sont généralement d'une complexité en temps de calcul élevée. En plus, ces méthodes sont très sensibles au bruit additif gaussien et leurs performances se dégradent lorsqu'il s'agit d'un bruit impulsif superposé à un bruit gaussien. Aussi, ces approches sont basées sur l'utilisation des porteuses pilotes qui ne peut pas être justifiée pour un objectif autre que la signalisation.

L'idée d'utiliser une technique de décodage itérative pour l'atténuation des bruits

impulsifs a été proposée par Haring et Vink [HV03]. Cette approche réduit considérablement les effets néfastes des bruits impulsifs. Cependant, sa complexité en temps de calcul est très élevée et elle introduit des retards de communications.

3.2 Modèle statistique

Dans ce chapitre, nous introduisons une nouvelle approche pour la compensation du bruit impulsif dans le contexte d'une transmission OFDM. La méthode proposée repose sur l'inférence bayésienne, et plus particulièrement sur les méthodes de Monte Carlo séquentielles et permet l'estimation conjointe du signal OFDM et des paramètres du bruit impulsif. Cette approche opère dans le domaine temporel. Au niveau de la chaîne de réception d'un système OFDM, notre algorithme se place donc juste avant le bloc de la FFT (cf. figure 3.1).

Comme il a été vu au paragraphe 3.1.3, de nombreuses approches ont été proposées dans la littérature afin de traiter le problème de réduction du bruit impulsif dans les systèmes multiporteuses. Cependant, toutes ces approches souffrent de plusieurs inconvénients. La plupart de ces méthodes sont basées sur l'utilisation des porteuses pilotes. Un problème posé par l'utilisation des porteuses pilotes réside dans le fait que l'on doit réserver un certain nombre des fréquences de transmission à ces porteuses. Ainsi, toutes les fréquences disponibles pour la transmission des données ne peuvent pas être utilisées, ce qui restreint le débit de transmission. En outre, certains standards de communication ne spécifient pas l'utilisation de porteuses pilotes, et ne permettent donc pas l'application des méthodes basées sur ce principe. En plus, plusieurs de ces approches supposent que les paramètres du bruit impulsif sont parfaitement connus du récepteur, ce qui ne s'avère pas être une hypothèse très réaliste. Contrairement aux travaux existants, l'approche proposée dans cette thèse ne repose pas sur l'utilisation des porteuses pilotes et les paramètres du bruit impulsif sont inconnus du récepteur.

L'approche proposée pour la compensation du bruit impulsif est fondée sur l'inférence bayésienne. Dans un tel cadre, le processus d'estimation repose essentiellement sur la connaissance d'un modèle statistique qui permet de décrire au mieux le système dynamique étudié. Afin de mettre en œuvre notre algorithme, nous avons utilisé une modélisation du système OFDM qui a été introduite par Septier *et al*

[SDMRG10, SDMRG08, Sep08]. Les modèles d'évolution et d'observation de cette modélisation sont définis dans les paragraphes suivants.

3.2.1 Modèle d'observation

Le signal OFDM reçu $r_{n,t}$ peut s'écrire avant la suppression du préfixe cyclique :

$$r_{n,t} = \sum_{l=0}^{L-1} h_{n,l} s_{n,t-l} + v_{n,t} + u_{n,t} \quad \text{pour } t = 0, \dots, N + N_{cp} - 1 \quad (3.7)$$

où $v_{n,t}$ et $u_{n,t}$ représentent respectivement les composantes gaussienne et impulsive du bruit. Cette équation d'observation prend en considération l'ISI due aux multi-trajets en utilisant la définition suivante de $s_{n,t}$ pour $t < 0$:

$$s_{n,t} = s_{n-1, N+N_{cp}+t} \quad \text{si } t < 0 \quad (3.8)$$

L'équation (3.7) peut être réécrite sous forme matricielle de la manière suivante :

$$r_{n,t} = \mathbf{h}_n^T \mathbf{s}_{n,t} + v_{n,t} + u_{n,t} \quad \text{pour } t = 0, \dots, N + N_{cp} - 1 \quad (3.9)$$

où

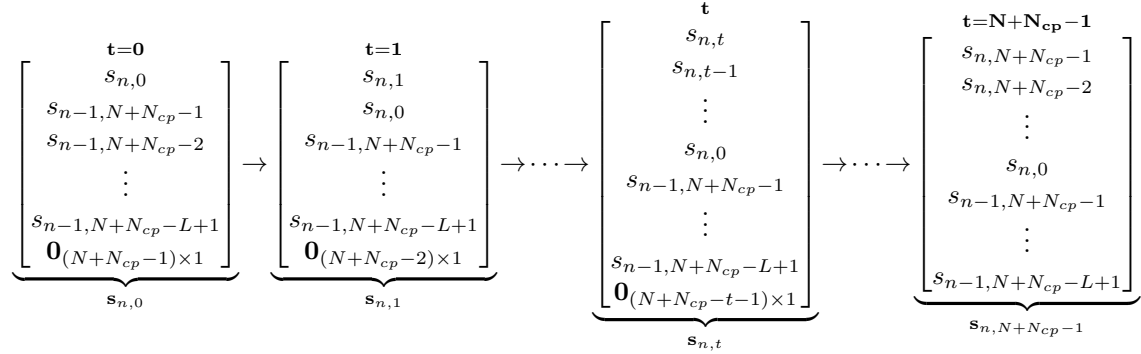
$$\mathbf{h}_n = \left[h_{n,0} \quad \cdots \quad h_{n,L-1} \quad \mathbf{0}_{1 \times (N+N_{cp}-1)} \right]^T \quad (3.10)$$

$$\mathbf{s}_{n,t} = \left[s_{n,t} \quad \cdots \quad s_{n,-L+1} \quad \mathbf{0}_{1 \times (N+N_{cp}-t-1)} \right]^T \quad (3.11)$$

Le vecteur $\mathbf{s}_{n,t}$ est défini comme étant la concaténation des différents échantillons $\{s_{n,t}\}_{t=0}^{N+N_{cp}-1}$ du signal OFDM transmis. Il est de dimension fixe $N + N_{cp} + L - 1$. Pour tenir compte de l'ISI, nous initialisons le vecteur $\mathbf{s}_{n,t}$ comme suit :

$$\mathbf{s}_{n,0} = \left[s_{n,0} \quad s_{n-1, N+N_{cp}-1} \quad \cdots \quad s_{n-1, N+N_{cp}-L+1} \quad \mathbf{0}_{1 \times (N+N_{cp}-t-1)} \right]^T$$

Par la suite, à chaque instant t , $\mathbf{s}_{n,t}$ est obtenu en décalant les différents échantillons de $\mathbf{s}_{n,t-1}$ et ajoutant le nouveau échantillon $s_{n,t}$ qui arrive (figure 3.3).

FIGURE 3.3 Construction du vecteur $\mathbf{s}_{n,t}$ 

La composante impulsive du bruit $u_{n,t}$ est modélisée par une distribution α -stable symétrique :

$$u_{n,t} \sim \mathcal{S}_1(\alpha, 0, \gamma, 0)$$

En utilisant le corollaire du théorème de mélange d'échelles 1.1, le bruit $u_{n,t}$ peut s'exprimer comme le produit d'une variable aléatoire gaussienne et d'une variable aléatoire α -stable positive [God99, GK99, God00] :

$$u_{n,t} = \sqrt{\lambda_{n,t}} w_{n,t} \quad \text{avec} \quad \lambda_{n,t} \sim \mathcal{S}_1\left(\frac{\alpha}{2}, 1, \cos\left(\frac{\pi\alpha}{4}\right)^{\frac{2}{\alpha}}, 0\right) \quad \text{et} \quad w_{n,t} \sim \mathcal{N}(0, 2\gamma^2)$$

De plus, en supposant que la composante gaussienne du bruit $v_{n,t}$ est de moyenne nulle et de variance σ_v^2 , on aboutit à l'équation d'observation suivante :

$$r_{n,t} = \mathbf{h}_n^T \mathbf{s}_{n,t} + d_{n,t} w_{n,t} \quad (3.12)$$

où $d_{n,t} = \sqrt{2\gamma^2 \lambda_{n,t} + \sigma_v^2}$.

3.2.2 Modèle d'évolution de l'état

Le préfixe cyclique correspond à la recopie des derniers échantillons du symbole OFDM en début de symbole

$$s_{n,t} = s_{n,t+N} \quad \forall \quad t \in [0, N_{cp} - 1] \quad (3.13)$$

Par conséquent, le signal OFDM peut être considéré comme un processus cyclostationnaire de période N . En se basant sur cette propriété et l'hypothèse que les échantillons d'un symbole OFDM sont indépendants et identiquement distribués de puissance 1, $\mathbb{E}[|s_{n,k}|^2] = 1$, nous pouvons énoncer la relation suivante [SDMRG10, SDMRG08, Sep08] :

$$\mathbb{E}[s_{n,k}s_{n,l}^*] = \begin{cases} 1 & \text{si } k = l + zN \text{ avec } z \in \mathbb{Z}, \quad k \in [0, N + N_{cp} - 1] \\ 0 & \text{sinon} \end{cases} \quad (3.14)$$

D'après le théorème de la limite centrale, l'enveloppe du signal OFDM peut être approchée par une variable aléatoire gaussienne circulaire. Dans la littérature, ce théorème a été déjà utilisé pour l'approximation d'un signal OFDM par un vecteur aléatoire gaussien circulaire [DW01, CZQ07]. Dans [WGK10], les auteurs ont démontré que l'enveloppe complexe d'un signal OFDM en bande de base converge lentement vers un processus gaussien. En se basant sur ces résultats, Septier *et al* [SDMRG10, SDMRG08, Sep08] ont proposé une modélisation originale du signal OFDM dans le domaine temporel par un processus AR. Selon ce modèle, le signal OFDM $\mathbf{s}_{n,t}$ peut s'écrire à chaque instant t

$$\mathbf{s}_{n,t} = \mathbf{A}_{n,t}\mathbf{s}_{n,t-1} + \mathbf{b}_{n,t}$$

où $\mathbf{A}_{n,t}$ est la matrice de transition définie par :

$$\begin{bmatrix} \xi_{n,t}^T \\ \mathbf{I}_{(N+N_{cp}+L-2)} & \mathbf{0}_{(N+N_{cp}+L-2) \times 1} \end{bmatrix}$$

En utilisant la relation (3.13), $\xi_{n,t}$ est donné par

$$\xi_{n,t}^T = \begin{cases} \left[\mathbf{0}_{1 \times (N+N_{cp}+L-1)} \right]^T & \text{si } 0 \leq t \leq N-1 \\ \left[\mathbf{0}_{1 \times (N-1)} \quad 1 \quad \mathbf{0}_{1 \times (N_{cp}+L-1)} \right]^T & \text{si } N \leq t \leq N + N_{cp} - 1 \end{cases}$$

Finalement, $\mathbf{b}_{n,t}$ est un bruit blanc gaussien circulaire de moyenne nulle et de matrice de covariance définie comme suit :

$$\mathbb{E}[\mathbf{b}_{n,t}\mathbf{b}_{n,t}^H] = \begin{bmatrix} \sigma_{b_{n,t}}^2 & 0 & \cdots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 0 \end{bmatrix}$$

où

$$\sigma_{b_{n,t}}^2 = \begin{cases} 1, & \text{si } 0 \leq t \leq N - 1 \\ 0, & \text{si } N \leq t \leq N + N_{cp} - 1 \end{cases}$$

3.2.3 Le modèle statistique

En conclusion, nous aboutissons au modèle dynamique suivant :

$$\begin{cases} \mathbf{s}_{n,t} = \mathbf{A}_{n,t}\mathbf{s}_{n,t-1} + \mathbf{b}_{n,t} \\ r_{n,t} = \mathbf{h}_{n,t}^T\mathbf{s}_{n,t} + d_{n,t}w_{n,t} \end{cases} \quad (3.15)$$

Etant donné ce modèle dynamique, l'objectif visé est d'estimer à chaque instant t le vecteur d'état $\mathbf{s}_{n,t}$ ainsi que les paramètres du bruit α -stable α , γ et $\lambda_{n,t}$ conditionnellement aux observations $r_{n,t}$. Par souci de clarté des notations, l'indice du symbole OFDM n est omis dans ce qui suit. L'exposant caractéristique α et le paramètre de dispersion γ sont des paramètres statiques, donc

$$\alpha_t = \alpha_{t-1} \quad \text{et} \quad \gamma_t = \gamma_{t-1} \quad (3.16)$$

Nous notons $\boldsymbol{\theta}_t = [\alpha_t \quad \gamma_t \quad \lambda_t]^T$ le vecteur des paramètres à estimer. L'approche que nous proposons pour l'estimation conjointe de \mathbf{s}_t et $\boldsymbol{\theta}_t$ repose sur les techniques de filtrage particulière Rao-Blackwellisé. Cette approche fera l'objet de la section suivante.

3.3 Filtre particulaire Rao-Blackwellisé pour l'estimation conjointe du signal OFDM et des paramètres du bruit α -stable

Dans le cadre de l'estimation Bayésienne, l'estimation conjointe du vecteur d'état \mathbf{s}_t et du vecteur des paramètres du bruit $\boldsymbol{\theta}_t$ revient à calculer la densité *a posteriori* $p(\mathbf{s}_t, \boldsymbol{\theta}_t | r_{0:t})$. Cependant, le calcul analytique de cette densité n'est pas possible. Par conséquent, nous proposons de l'approcher numériquement en utilisant les méthodes de Monte Carlo séquentielles. A partir du modèle statistique défini par le système d'équations (3.15), nous pouvons remarquer que conditionnellement à $\boldsymbol{\theta}_t$ notre modèle est linéaire gaussien. Cette structure particulière du modèle peut être mise à profit en utilisant un filtre particulaire Rao-Blackwellisé.

En marginalisant selon le vecteur d'état \mathbf{s}_t , la densité *a posteriori* peut s'écrire comme suit :

$$p(\mathbf{s}_t, \boldsymbol{\theta}_t | r_{0:t}) = p(\mathbf{s}_t | \boldsymbol{\theta}_t, r_{0:t}) p(\boldsymbol{\theta}_t | r_{0:t}) \quad (3.17)$$

Dans (3.17), la première densité peut être déterminée par un filtre de Kalman. Cependant, le calcul analytique de la deuxième densité n'est pas envisageable, c'est pour cette raison que nous proposons de l'approcher par un filtre particulaire :

$$p(\boldsymbol{\theta}_t | r_{0:t}) \simeq \sum_{i=1}^M \tilde{\omega}_t^{(i)} \delta_{\tilde{\boldsymbol{\theta}}_t^{(i)}}(\boldsymbol{\theta}_t) \quad (3.18)$$

En substituant (3.18) dans (3.17), nous obtenons une approximation de la densité *a posteriori* par :

$$p(\mathbf{s}_t, \boldsymbol{\theta}_t | r_{0:t}) \simeq \sum_{i=1}^M \tilde{\omega}_t^{(i)} p(\mathbf{s}_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t}) \quad (3.19)$$

où $p(\mathbf{s}_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t})$ est une densité gaussienne dont la moyenne $\hat{\mathbf{s}}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ et la matrice de covariance $\boldsymbol{\Sigma}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ sont évaluées en utilisant un filtre de Kalman.

3.3.1 Filtre de Kalman pour l'estimation de l'état

Les paramètres $\hat{\mathbf{s}}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ et $\boldsymbol{\Sigma}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ de la densité gaussienne $p(\mathbf{s}_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t})$ sont calculés de façon récurrente pour chaque particule i à l'aide d'un filtre de Kalman.

Les équations de prédiction du filtre de Kalman sont les suivantes :

$$\begin{cases} \widehat{\mathbf{s}}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \mathbf{A}_t \widehat{\mathbf{s}}_{t-1|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) \\ \boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \mathbf{A}_t \boldsymbol{\Sigma}_{t-1|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) \mathbf{A}_t^H + \mathbb{E}[\mathbf{b}_t \mathbf{b}_t^H] \\ \widehat{r}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \mathbf{h}^T \widehat{\mathbf{s}}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) \end{cases} \quad (3.20)$$

et les équations de correction sont :

$$\begin{cases} S_t(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \mathbf{h}^T \boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) \mathbf{h}^* + d_t^2(\tilde{\boldsymbol{\theta}}_t^{(i)}) \\ \mathbf{K}_t(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) (\mathbf{h}^T)^H (S_t(\tilde{\boldsymbol{\theta}}_t^{(i)}))^{-1} \\ \widehat{\mathbf{s}}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \widehat{\mathbf{s}}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) + \mathbf{K}_t(\tilde{\boldsymbol{\theta}}_t^{(i)}) (r_t - \widehat{r}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})) \\ \boldsymbol{\Sigma}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)}) = \boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) - \mathbf{K}_t(\tilde{\boldsymbol{\theta}}_t^{(i)}) \mathbf{h}^T \boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}) \end{cases} \quad (3.21)$$

Une fois la densité *a posteriori* du symbole OFDM $p(\mathbf{s}_t | \boldsymbol{\theta}_t, r_{0:t})$ obtenue, nous approchons la densité $p(\boldsymbol{\theta}_t | r_{0:t})$ par un filtre particulaire. Ceci nécessite de choisir les densités d'importance selon lesquelles les particules seront générées.

3.3.2 Lois d'importance

La densité d'importance optimale du vecteur des paramètres inconnus $\boldsymbol{\theta}_t$ est donnée par :

$$q(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{0:t-1}^{(i)}, r_{0:t}) = p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{0:t-1}^{(i)}, r_{0:t}) \quad (3.22)$$

Cependant, cette densité ne peut pas être utilisée car les poids associés ne peuvent pas se calculer analytiquement. La densité d'importance choisie pour faire évoluer les particules $\boldsymbol{\theta}_t^{(i)}$ est la loi d'évolution $p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}^{(i)})$. Cette densité peut s'écrire comme suit [LG06, LG04] :

$$p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}^{(i)}) = p(\lambda_t | \alpha_t, \gamma_t) p(\alpha_t | \alpha_{t-1}^{(i)}) p(\gamma_t | \gamma_{t-1}^{(i)}) \quad (3.23)$$

La variable auxiliaire λ_t suit une distribution α -stable positive. Sa densité d'importance est donnée par :

$$p(\lambda_t) \sim \mathcal{S}_1 \left(\frac{\alpha_t}{2}, 1, \cos \left(\left(\frac{\pi \alpha_t}{4} \right)^{\frac{2}{\alpha_t}} \right), 0 \right) \quad (3.24)$$

L'exposant caractéristique α et le paramètre de dispersion γ sont des paramètres statiques. Or, l'estimation de tels paramètres par le filtrage particulaire favorise davantage le problème de dégénérescence des poids. Pour contourner ce problème, nous utilisons la solution proposée par Doucet *et al* [DDFG01]. Le principe de cette méthode ainsi que les densités d'importance des paramètres α et γ sont détaillés dans les paragraphes suivants.

3.3.2.1 Echantillonnage de l'exposant caractéristique α

La méthode d'évolution du paramètre statique, proposée dans [DDFG01], consiste à supposer que le paramètre α varie dans le temps en ajoutant une petite perturbation à chaque instant t :

$$\alpha_t = \alpha_{t-1} + \zeta_t \quad \text{avec} \quad \zeta_t \sim \mathcal{N}(0, \Sigma_{\zeta_t})$$

Dans le cas où α est statique, la densité de probabilité *a posteriori* $p(\alpha|r_{0:t})$ est caractérisé par sa moyenne $\bar{\alpha}_t$ et sa variance Σ_t^α . Par contre, dans le cas du paramètre artificiel, la variance sera $\Sigma_t^\alpha + \Sigma_{\zeta_t}$. On peut exprimer l'évolution de cette variance comme suit :

$$\Sigma_t^\alpha = \Sigma_{t-1}^\alpha + \Sigma_{\zeta_t} + 2\text{cov}(\alpha_{t-1}, \zeta_t)$$

Pour éviter la perte d'information provoquée par le fait de considérer que le paramètre statique est variant dans le temps, il faut choisir $\text{cov}(\alpha_{t-1}, \zeta_t) = -\frac{\Sigma_{\zeta_t}}{2}$. Cette condition est vérifiée lorsqu'on suppose que :

$$\Sigma_{\zeta_t} = \Sigma_t^\alpha \left(\frac{1}{\delta} - 1 \right)$$

où δ est un facteur qui doit être compris entre 0 et 1.

La loi d'évolution de α_t peut alors s'écrire :

$$p(\alpha_t|\alpha_{t-1}) = \mathcal{N}(\alpha_t; d\alpha_{t-1} + (1-d)\bar{\alpha}_{t-1}, h^2\Sigma_{t-1}^\alpha)$$

avec

$$d = \frac{3\delta - 1}{2\delta} \quad \text{et} \quad h^2 = 1 - d^2$$

3.3.2.2 Echantillonnage du paramètre de dispersion γ

La densité d'importance du paramètre de dispersion γ est obtenue de la même manière que celle de l'exposant caractéristique α :

$$p(\gamma_t | \gamma_{t-1}) = \mathcal{N}(\gamma_t; d\gamma_{t-1} + (1-d)\bar{\gamma}_{t-1}, h^2 \Sigma_{t-1}^\gamma)$$

où $\bar{\gamma}_{t-1}$ et Σ_{t-1}^γ sont respectivement la moyenne et la variance de Monte Carlo de la densité $p(\gamma | r_{0:t-1})$.

3.3.3 Evaluation des poids

Avec la densité d'importance définie dans (3.23), nous obtenons l'expression suivante pour la mise à jour des poids :

$$\begin{aligned} \tilde{\omega}_t^{(i)} &\propto \frac{p(r_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t-1}) p(\tilde{\boldsymbol{\theta}}_t^{(i)} | \boldsymbol{\theta}_{t-1}^{(i)})}{q(\tilde{\boldsymbol{\theta}}_t^{(i)} | \boldsymbol{\theta}_{0:t-1}^{(i)}, r_{0:t})} \\ &\propto \tilde{\omega}_{t-1}^{(i)} p(r_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t-1}) \end{aligned} \quad (3.25)$$

Dans notre cas, la densité $p(r_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t-1})$ est une distribution gaussienne

$$p(r_t | \tilde{\boldsymbol{\theta}}_t^{(i)}, r_{0:t-1}) = \mathcal{N}(r_t; \hat{r}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}), S_t(\tilde{\boldsymbol{\theta}}_t^{(i)})) \quad (3.26)$$

où $\hat{r}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ et $S_t(\tilde{\boldsymbol{\theta}}_t^{(i)})$ sont respectivement la prédiction de l'observation et la covariance de l'innovation du filtre de kalman (cf. équations (3.20) et (3.21)).

3.3.4 Estimées MMSE du signal OFDM et des paramètres du bruit

A chaque instant t , les estimées MMSE du vecteur $\hat{\mathbf{s}}_t$ et des paramètres du bruit $\hat{\gamma}_t$ et $\hat{\alpha}_t$ sont respectivement obtenues par les expressions suivantes :

$$\hat{\mathbf{s}}_t = \sum_{i=1}^M \omega_t^{(i)} \hat{\mathbf{s}}_{t|t}(\boldsymbol{\theta}_t^{(i)}) \quad (3.27)$$

$$\hat{\gamma}_t = \sum_{i=1}^M \omega_t^{(i)} \gamma_t^{(i)} \quad (3.28)$$

$$\hat{\alpha}_t = \sum_{i=1}^M \omega_t^{(i)} \alpha_t^{(i)} \quad (3.29)$$

L'estimée MMSE du symbole OFDM transmis est alors donnée par $\hat{\mathbf{S}}_{N+N_{cp}-1}$.

3.3.5 Algorithme

Le filtre particulière Rao-Blackwellisé proposé pour l'estimation conjointe du signal OFDM transmis et des paramètres du bruit α -stable est résumé dans l'algorithme 12.

3.4 Simulations

Les différentes simulations sont réalisées en appliquant l'algorithme 12 sur un système OFDM ayant les caractéristiques suivantes :

- le nombre de porteuses est fixé à $N = 64$;
- La longueur du préfixe cyclique est $N_{cp} = 8$;
- La modulation utilisée sur chaque porteuse est de type 16-QAM ;
- Le canal de transmission est de type Rayleigh sélectif en fréquence à $L = 4$ trajets.

Pour les paramètres du filtre particulière, nous avons choisi une loi uniforme $\mathcal{U}[0.2, 2]$ comme *a priori* pour l'exposant caractéristique α et une loi inverse gamma $\mathcal{IG}(a^\gamma, b^\gamma)$ comme *a priori* pour le paramètre de dispersion γ . Le facteur δ est fixé à 0.95 [DDFG01].

Dans l'étude d'une chaîne de transmission numérique, les performances du système sont généralement évaluées en fonction du rapport signal sur bruit (SNR, Signal to Noise Ratio) défini comme étant le rapport des puissances du signal P_S et du bruit P_B :

$$SNR = \frac{P_S}{P_B} \quad (3.30)$$

Cependant, cette expression standard du SNR ne peut pas être utilisée dans le cas de bruits α -stables à cause de leur variance infinie. Dans la littérature, d'autres

Algorithme 12 Filtre particulaire Rao-Blackwellisé pour l'estimation conjointe du signal OFDM et des paramètres du bruit α -stable

• **Initialisation**

pour $i = 1$ à M **faire**

– Echantillonner $\tilde{\boldsymbol{\theta}}_0^{(i)} \sim p_0(\boldsymbol{\theta}_0)$

– Initialiser les poids du filtre particulaire $\omega_0^{(i)} = \frac{1}{M}$

fin pour

• **Itérations**

pour $t = 1$ à $N + N_{cp} - 1$ **faire**

pour $i = 1$ à M **faire**

– Echantillonner $\tilde{\boldsymbol{\theta}}_t^{(i)} \sim p(\boldsymbol{\theta}_t^{(i)} | \boldsymbol{\theta}_{t-1}^{(i)})$

– Calculer $\hat{\mathbf{s}}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})$, $\boldsymbol{\Sigma}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})$, $\hat{\mathbf{s}}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$, $\boldsymbol{\Sigma}_{t|t}(\tilde{\boldsymbol{\theta}}_t^{(i)})$ et $S_t(\tilde{\boldsymbol{\theta}}_t^{(i)})$ en faisant un pas de filtre de Kalman avec le système dynamique (3.15) à partir de $\hat{\mathbf{s}}_{t-1|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})$, $\boldsymbol{\Sigma}_{t-1|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)})$, $\tilde{\boldsymbol{\theta}}_t^{(i)}$ et r_t

– Calculer les poids $\tilde{\omega}_t^{(i)} = \omega_{t-1}^{(i)} \mathcal{N}(r_t; \mathbf{h}^T \hat{\mathbf{s}}_{t|t-1}(\tilde{\boldsymbol{\theta}}_t^{(i)}), S_t(\tilde{\boldsymbol{\theta}}_t^{(i)}))$

fin pour

pour $i = 1$ à M **faire**

– Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^M \tilde{\omega}_t^{(j)}}$

fin pour

– Calculer $N_{eff} = \frac{1}{\sum_{i=1}^M \tilde{\omega}_t^{(i)}}$

si $N_{eff} \leq N_{seuil}$ **alors**

– Rééchantillonner les particules $\tilde{\boldsymbol{\theta}}_t^{(i)}$ en utilisant le rééchantillonnage systématique

– Réinitialiser les poids $\tilde{\omega}_t^{(i)} = \frac{1}{M}$

fin si

– Mettre à jour des particules $\boldsymbol{\theta}_t^{(i)} = \tilde{\boldsymbol{\theta}}_t^{(i)}$

– Mettre à jour des poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

fin pour

expressions du SNR adaptées aux bruits α -stables ont été proposées :

- Le rapport signal sur bruit généralisé [TN95a]

$$generalized - SNR = 10 \log \left(\frac{1}{\gamma N} \sum_{i=1}^N |s(t)|^2 \right) \quad (3.31)$$

- Le pseudo rapport signal sur bruit [TN95a]

$$pseudo - SNR = 10 \log \left(\frac{\sum_{t=1}^N |s(t)|^2}{\sum_{t=1}^N |n(t)|^2} \right) \quad (3.32)$$

- Le rapport signal sur bruit géométrique basé sur la puissance géométrique [GPA06]

$$geometric - SNR = \frac{1}{2C_g} \cdot \frac{S^2}{S_0} \quad (3.33)$$

où S est l'amplitude du signal, $C_g = \exp(\gamma)^2$ et S_0 est la puissance géométrique du bruit définie par : $\exp(E \log |n|)$.

Les systèmes de transmission sont généralement comparés en analysant l'évolution du taux d'erreur binaire (BER, Bit Error Rate) en fonction du SNR. L'utilisation de l'un des critères définis ci-dessus pour évaluer les performances de notre système représente un inconvénient pour comparer nos résultats avec les résultats qui se trouvent dans la littérature. Pour cette raison, nous proposons d'évaluer les performances de notre système en utilisant la variation de la dispersion γ qui est proportionnel à la variance dans le cas d'un bruit blanc gaussien. Dans notre cadre d'étude, le signal OFDM est aussi entaché d'une distorsion gaussienne. Pour tenir compte de l'effet de cette composante dans l'évaluation des performances de notre système, nous proposons d'utiliser la variation de la dispersion en fixant le rapport $10 \log_{10}(\sigma_v/\gamma)$.

Tout d'abord, nous nous attachons à étudier la performance du filtre particulaire proposé en fonction du nombre de particules utilisées. Cette étude vise principalement à déterminer le nombre optimal de particules qui permet de réaliser un bon compromis entre efficacité et faible complexité en temps de calcul. La performance de notre algorithme est étudiée en termes du biais et de la variance d'estimation du signal OFDM. Ces quantités sont respectivement calculées pour un symbole OFDM

2. $\gamma = 0,577215665$ est la constante d'Euler

donné par :

$$\mathbf{B} = \frac{1}{N + N_{cp} - 1} \left[\sum_{i=0}^{N+N_{cp}-1} (\Re(\mathbf{s}_{N+N_{cp}-1}[i] - \widehat{\mathbf{s}}_{N+N_{cp}-1}[i]) + \Im(\mathbf{s}_{N+N_{cp}-1}[i] - \widehat{\mathbf{s}}_{N+N_{cp}-1}[i])) \right] \quad (3.34)$$

$$\mathbf{V} = \frac{1}{N + N_{cp} - 1} (\mathbf{s}_{N+N_{cp}-1} - \widehat{\mathbf{s}}_{N+N_{cp}-1})(\mathbf{s}_{N+N_{cp}-1} - \widehat{\mathbf{s}}_{N+N_{cp}-1})^H - \mathbf{B}^2 \quad (3.35)$$

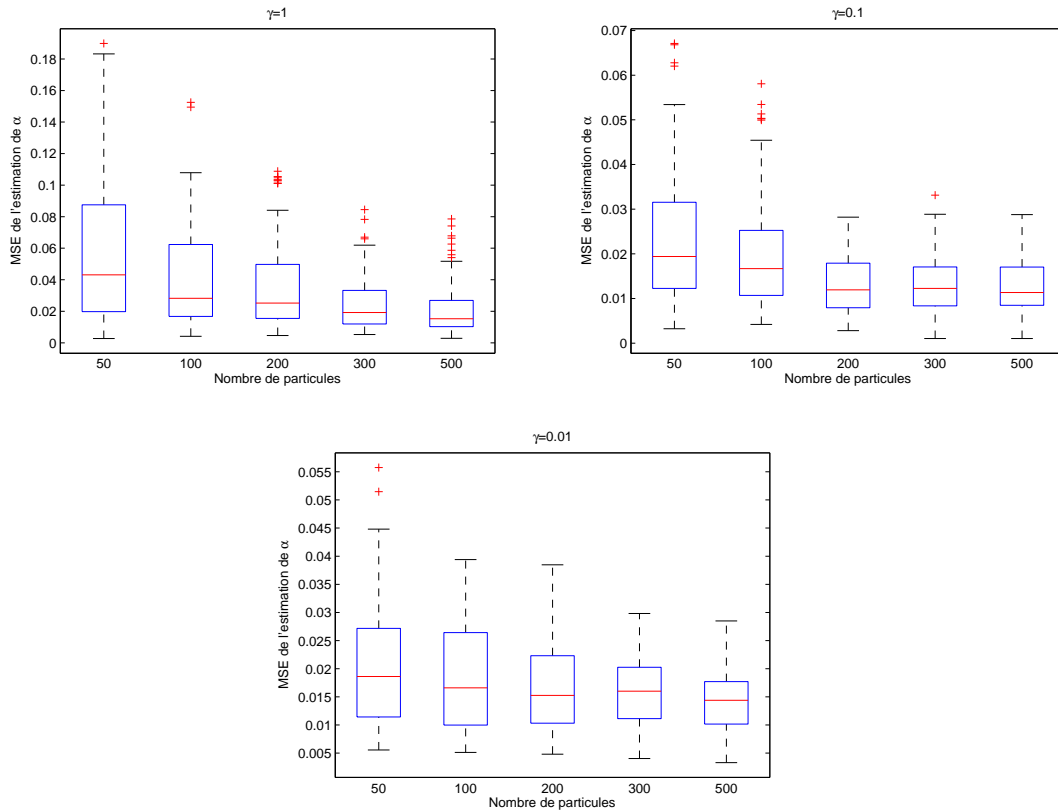
Nous avons réalisé 100 simulations avec des signaux OFDM et des réalisations de bruits différents. Les paramètres du bruit α -stable sont fixés à $\alpha = 1.4$ et $\mu = 0$. Plusieurs valeurs de γ ont été considérées $\gamma = 1$, $\gamma = 0.1$ et $\gamma = 0.01$. Dans la table 3.2, la moyenne de \mathbf{B} et la moyenne de \mathbf{V} pour les 100 simulations sont reportées en fonction du nombre de particules. Les figures 3.4 et 3.5 représentent respectivement les boîtes à moustaches de l'erreur quadratique moyenne (MSE, Mean Squared Error) d'estimation de α et γ en fonction du nombre de particules pour différentes valeurs de γ . D'après les résultats illustrés dans le tableau et dans les figures, nous pouvons tout d'abord constater que la performance du filtre proposé augmente avec le nombre de particules. Cependant, même avec un petit nombre de particules, les résultats d'estimation obtenus restent satisfaisants. De plus, nous remarquons qu'un nombre de particules supérieur à 300 n'améliore pas significativement les performances. En revanche, ceci augmente considérablement les temps de calcul du filtre particulaire. Par conséquent, nous pouvons conclure qu'il sera judicieux d'utiliser un nombre de particules entre 200 et 300. Désormais, le nombre de particules M est fixé à 200.

TABLE 3.2 – Biais et variance de l'estimation (entre parenthèses) du signal OFDM en fonction du nombre de particules pour différentes valeurs de γ avec $\alpha = 1.4$ et $10 \log_{10}(\sigma/\gamma) = 0$.

	M=50	M=100	M=200	M=300	M=500
$\gamma = 1$	0.9609 (0.3108)	0.9390 (0.2940)	0.9330 (0.2206)	0.9306 (0.2831)	0.9177 (0.2539)
$\gamma = 0.1$	0.3634 (0.2883)	0.3574 (0.2834)	0.3478 (0.2716)	0.3422 (0.2571)	0.3330 (0.2480)
$\gamma = 0.01$	0.0528 (0.0500)	0.0523 (0.0495)	0.0519 (0.0492)	0.0519 (0.0492)	0.0509 (0.0483)

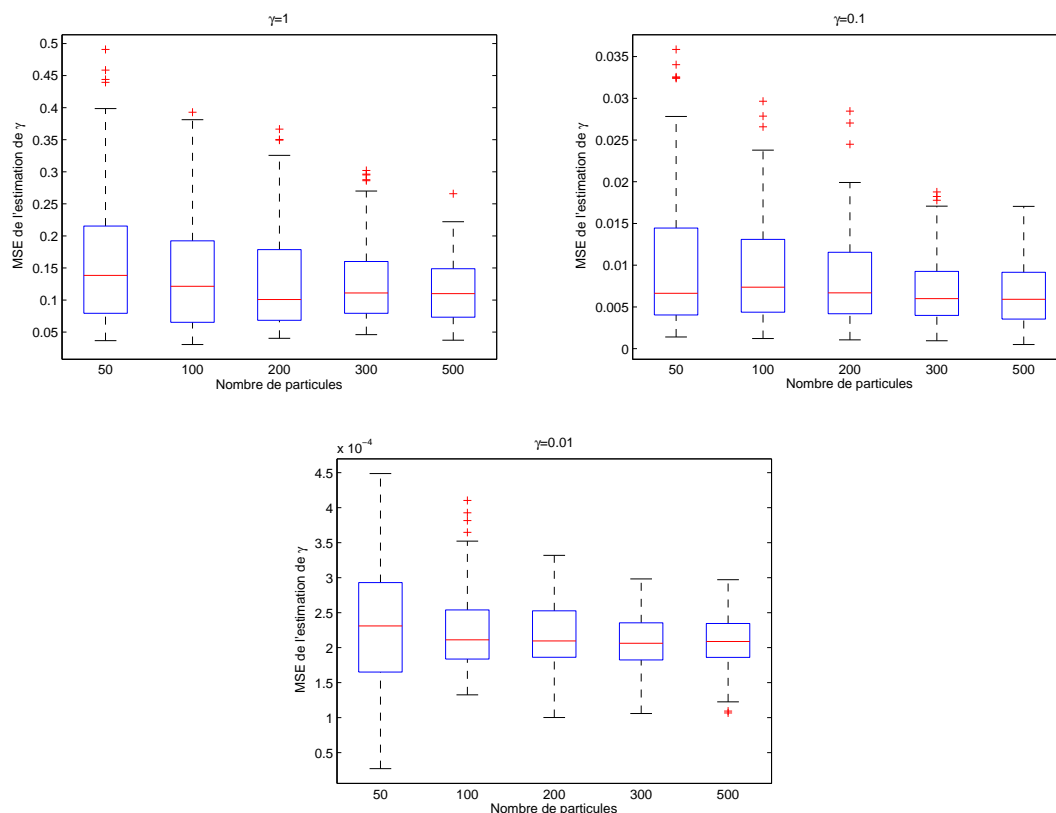
Maintenant, nous nous proposons d'étudier l'influence de l'exposant caractéristique α sur la performance de notre filtre particulaire. Pour ce faire, plusieurs valeurs

FIGURE 3.4 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation de l'exposant caractéristique α en fonction du nombre de particules pour différentes valeurs de γ avec $\alpha = 1.4$ et $10 \log_{10}(\sigma/\gamma) = 0$.



de α ont été considérées à savoir : $\alpha = 0.8$, $\alpha = 1.1$, $\alpha = 1.4$ et $\alpha = 1.7$. Dans cette étude, la performance de notre algorithme est évaluée en termes de l'erreur quadratique moyenne. La figure 3.6 montre l'évolution du MSE de l'estimation du signal OFDM en fonction de γ pour différentes valeurs de α . A partir de cette figure, nous pouvons remarquer que la performance du filtre particulaire proposé se dégrade quand la valeur de l'exposant caractéristique α diminue. Ce résultat peut être expliqué par le fait qu'en diminuant la valeur de α , l'impulsivité du bruit d'observation augmente, rendant l'inférence plus délicate. Les résultats obtenus pour l'estimation du paramètre de dispersion γ sont illustrés dans les figures 3.7-3.10. Sur la figure 3.7, sont tracées les courbes d'évolution du MSE de l'estimation de γ pour les différentes valeurs de α considérées. Comme pour le signal OFDM, nous remarquons que l'augmentation de la valeur de l'exposant caractéristique α améliore la qualité

FIGURE 3.5 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation du paramètre de dispersion γ en fonction du nombre de particules pour différentes valeurs de γ avec $\alpha = 1.4$ et $10 \log_{10}(\sigma/\gamma) = 0$.



d'estimation de γ . En revanche, l'écart entre les courbes diminue avec γ et à partir d'une valeur de $\gamma = 0.001$ les différentes courbes se rejoignent. L'influence de l'exposant caractéristique sur l'estimation de γ est plus significative lorsque la puissance du bruit impulsif est grande. Ces résultats sont également confirmés par les boîtes à moustaches du MSE d'estimation de γ en fonction de α représentées dans les figures 3.8-3.10. En effet, pour $\gamma = 0.1$ (cf. figure 3.8), nous pouvons bien voir la diminution de la dispersion et du biais quand la valeur de α augmente. Cependant, cette diminution est moins significative pour $\gamma = 0.01$ (cf. figure 3.9) et presque inexistant pour $\gamma = 0.001$ (cf. figure 3.10). Les figures 3.11-3.13 représentent respectivement les boîtes à moustaches du MSE d'estimation de l'exposant caractéristique α pour $\gamma = 10$, $\gamma = 1$ et $\gamma = 0.1$. A partir de ces figures, nous pouvons constater que la qualité d'estimation augmente avec α . Cependant, pour $\alpha = 1.7$, nous remarquons une

FIGURE 3.6 MSE de l'estimation du signal OFDM en fonction de $10\log_{10}(\gamma)$ pour différentes valeurs de α avec $10\log_{10}(\sigma/\gamma) = 0$ (gauche) et $10\log_{10}(\sigma/\gamma) = -10$ (droite)

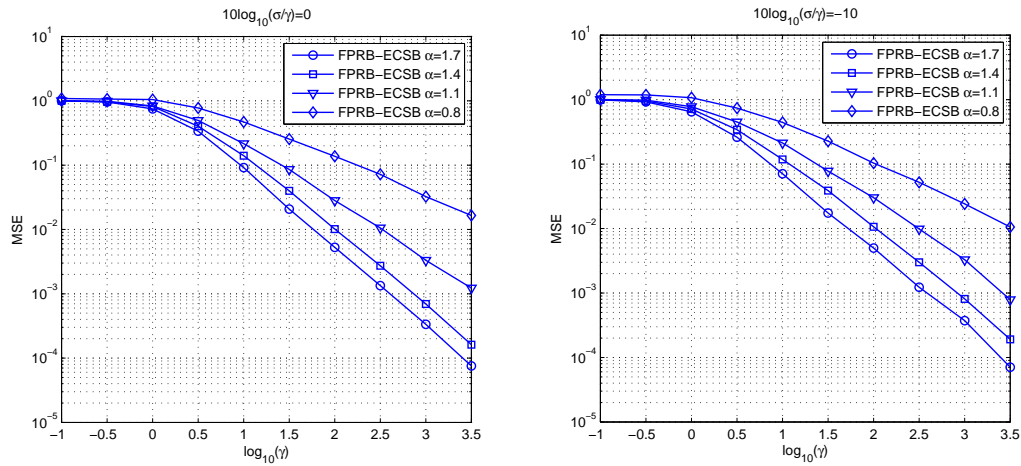


FIGURE 3.7 Erreur quadratique moyenne de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $10\log_{10}(\sigma/\gamma) = 0$ (gauche) et $10\log_{10}(\sigma/\gamma) = -10$ (droite)

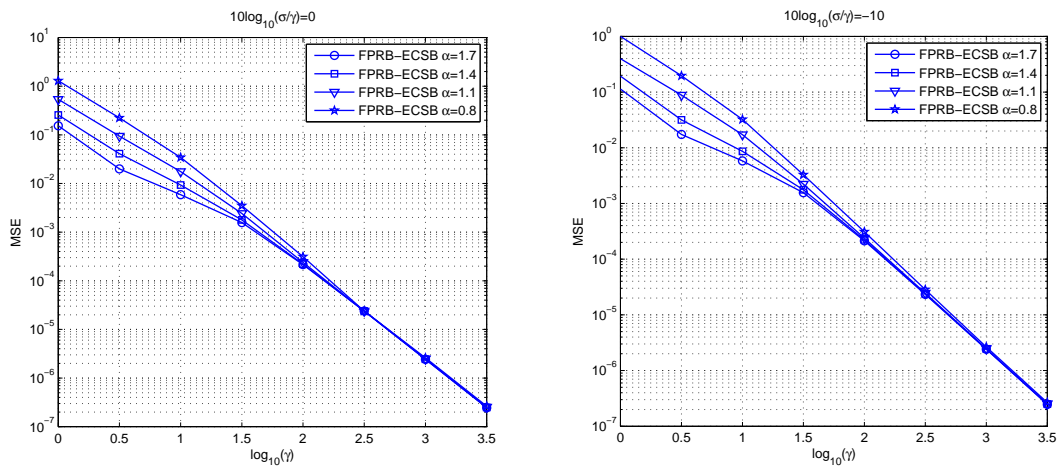


FIGURE 3.8 Les boîtes à moustaches du MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $\gamma=0.1$ et $10 \log_{10}(\sigma/\gamma) = 0$ (gauche) et $10 \log_{10}(\sigma/\gamma) = -10$ (droite)

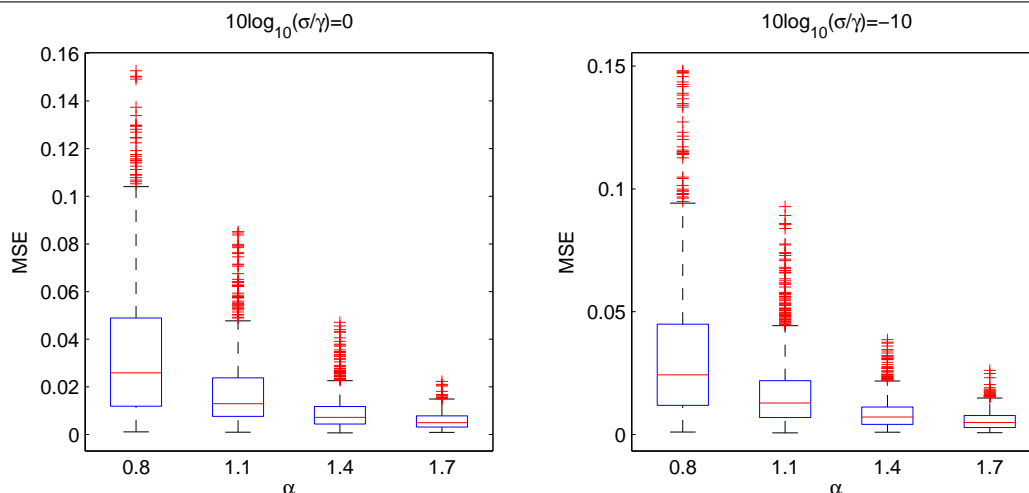


FIGURE 3.9 Les boîtes à moustaches du MSE de l'estimation du paramètre de dispersion en fonction de γ pour différentes valeurs de α avec $\gamma=0.01$ et $10 \log_{10}(\sigma/\gamma) = 0$ (gauche) et $10 \log_{10}(\sigma/\gamma) = -10$ (droite)

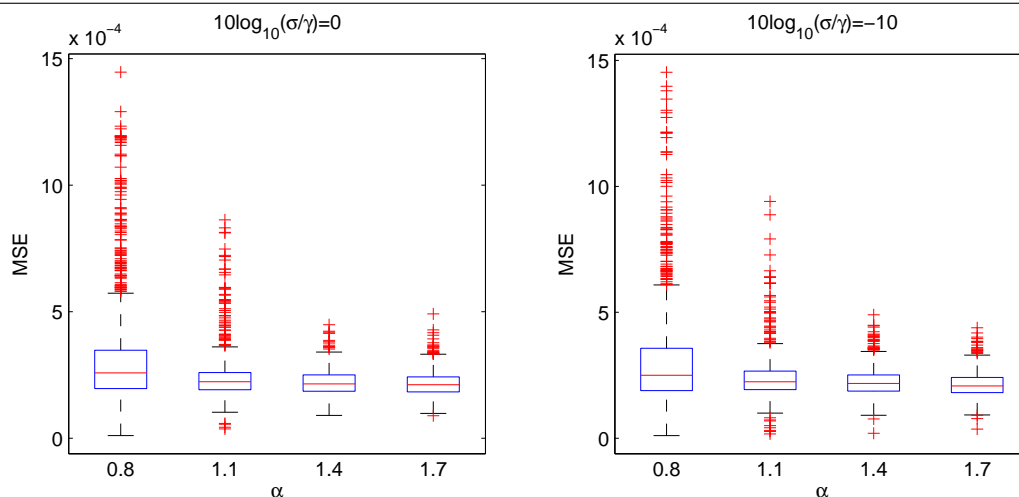
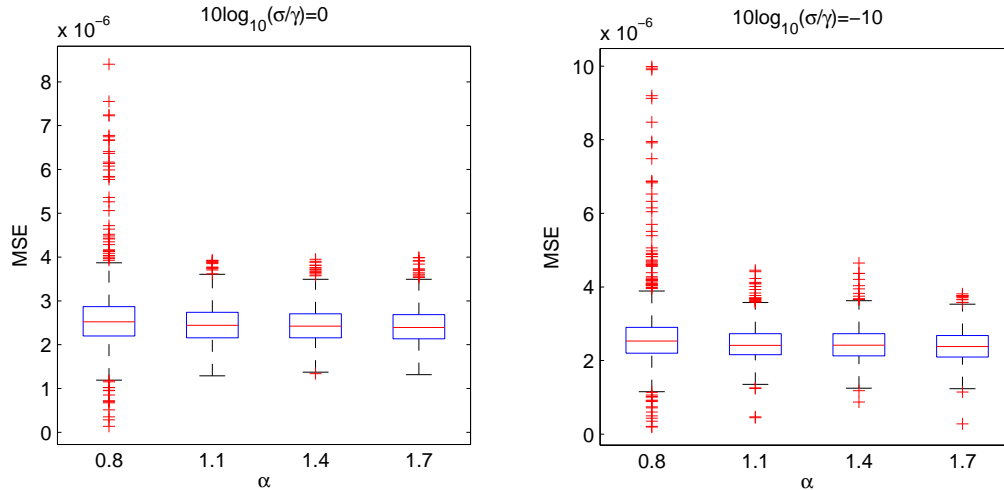


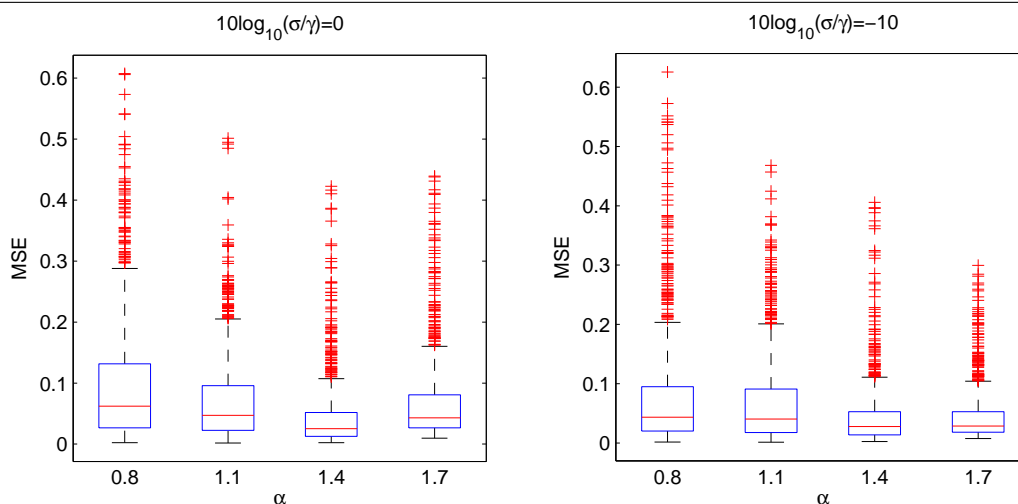
FIGURE 3.10 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation du paramètre de dispersion pour différentes valeurs de α avec $\gamma=0.001$ et $10\log_{10}(\sigma/\gamma) = 0$ (gauche) et $10\log_{10}(\sigma/\gamma) = -10$ (droite)



dégradation des performances qui se traduit dans les boîtes à moustaches par une augmentation de la dispersion et du biais d'estimation. Cette dégradation devient de plus en plus remarquable lorsque la puissance du bruit impulsif diminue (valeur de γ plus petite). En effet, pour $\alpha = 1.7$, on est proche du cas gaussien et notre bruit est dominé par sa composante gaussienne. Par conséquent, on n'a pas assez d'informations pour estimer le paramètre α qui caractérise la composante impulsive de notre bruit.

Dans les systèmes de transmission numériques, les performances sont généralement évaluées en termes de BER. Les figures 3.14, 3.15 et 3.16 représentent les courbes d'évolution du taux d'erreur binaire obtenu avec la méthode proposée en fonction de γ pour différentes valeurs de α . Pour fins de comparaison, nous traçons également les courbes obtenues avec un égaliseur de canal MMSE et celles obtenues avec deux méthodes classiques de compensation du bruit impulsif à savoir le Blanking et le Clipping (cf. paragraphe 3.1.3). A partir de ces figures, nous pouvons constater que les résultats obtenus avec notre filtre particulière sont globalement meilleurs que ceux obtenus avec les autres méthodes. De plus, nous pouvons remarquer que l'amélioration apportée par notre méthode est particulièrement significative lorsqu'il s'agit d'un bruit extrêmement impulsif (une petite valeur de α).

FIGURE 3.11 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation de l'exposant caractéristique α pour $\gamma = 10$ et $10 \log_{10}(\sigma/\gamma) = 0$ (gauche) et $10 \log_{10}(\sigma/\gamma) = -10$ (droite).



3.5 Conclusion

Dans ce chapitre, nous avons traité le problème de compensation du bruit impulsif dans les systèmes de transmission OFDM. A cet effet, une approche bayésienne paramétrique a été proposée. Contrairement aux méthodes existantes, nous supposons que les paramètres du bruit impulsif sont inconnus du récepteur. Nous avons considéré un modèle basé sur les distributions α -stables symétriques pour ce bruit. Un filtre particulière Rao-Blackwellisé permettant l'estimation conjointe des symboles OFDM et des paramètres du bruit α -stable a été présenté. Les performances du filtre proposé ont été étudiées à travers plusieurs simulations numériques. Les résultats obtenus montrent l'efficacité du schéma proposé pour l'estimation du symbole OFDM transmis et des paramètres du bruit α -stable. De plus, notre filtre permet d'améliorer les performances en termes de taux d'erreur binaire comparé aux méthodes classiques de Blanking et de Clipping.

Dans les chapitres qui suivent, nous étudions cette problématique dans un cadre plus large, celui des systèmes non linéaires. De plus, nous ne nous limitons pas au cas des bruits α -stables symétriques. Les solutions proposées font appel à des modèles bayésiens non paramétriques.

FIGURE 3.12 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation de l'exposant caractéristique α pour $\gamma = 1$ et $10 \log_{10}(\sigma/\gamma) = 0$ (gauche) et $10 \log_{10}(\sigma/\gamma) = -10$ (droite).

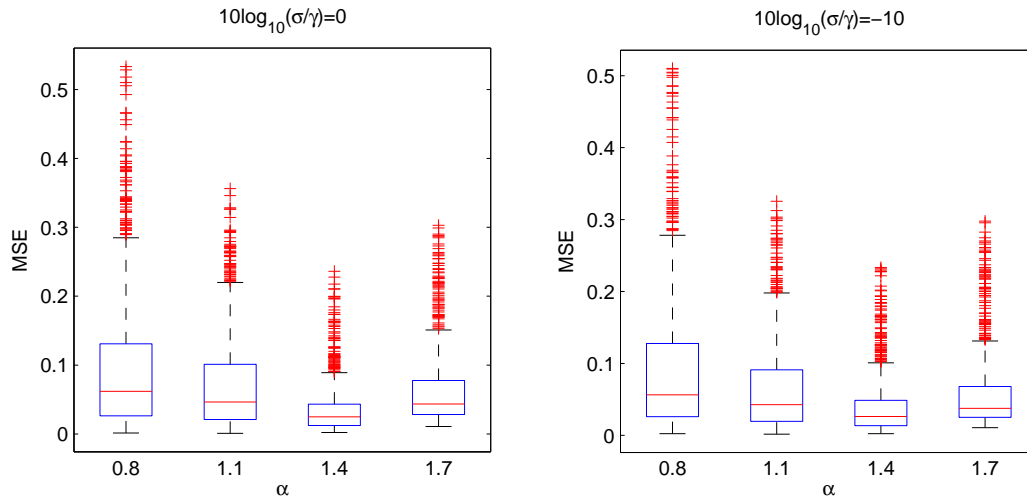


FIGURE 3.13 Les boîtes à moustaches de l'erreur quadratique moyenne de l'estimation de l'exposant caractéristique α pour $\gamma = 0.1$ et $10 \log_{10}(\sigma/\gamma) = 0$ (gauche) et $10 \log_{10}(\sigma/\gamma) = -10$ (droite).

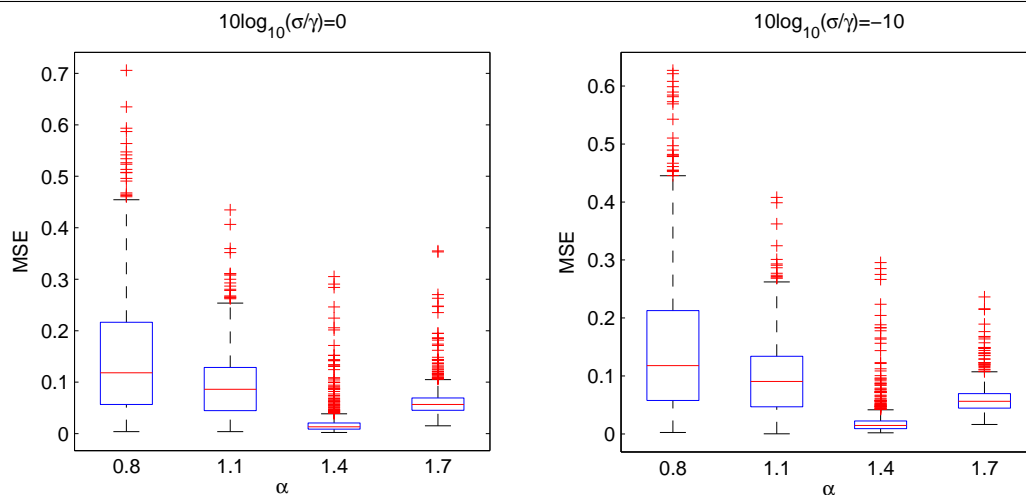


FIGURE 3.14 Performance en terme de taux d'erreur binaire de la méthode proposée FPRB-ECSB pour différentes valeurs du rapport $10 \log_{10}(\sigma/\gamma)$ avec $\alpha = 0.8$.

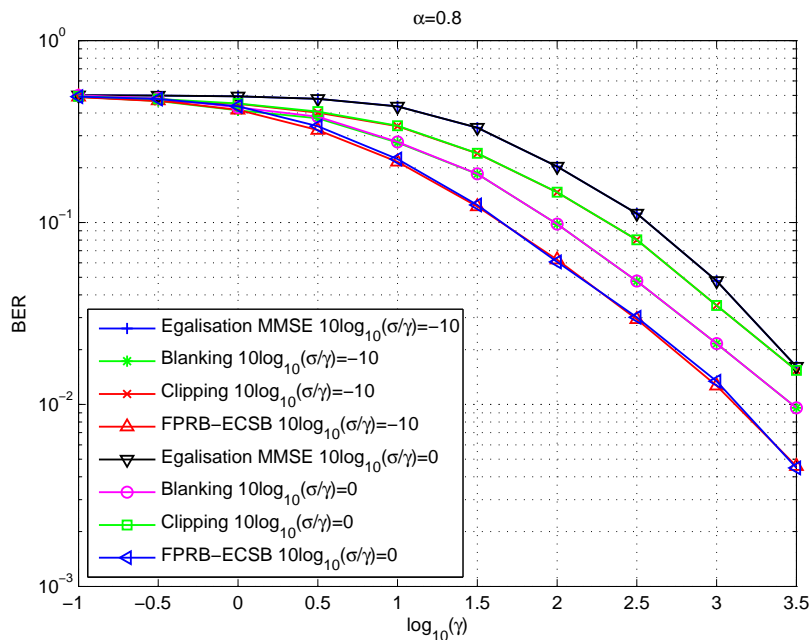


FIGURE 3.15 Performance en terme de taux d'erreur binaire de la méthode proposée FPRB-ECSB pour différentes valeurs du rapport $10 \log_{10}(\sigma/\gamma)$ avec $\alpha = 1.1$.

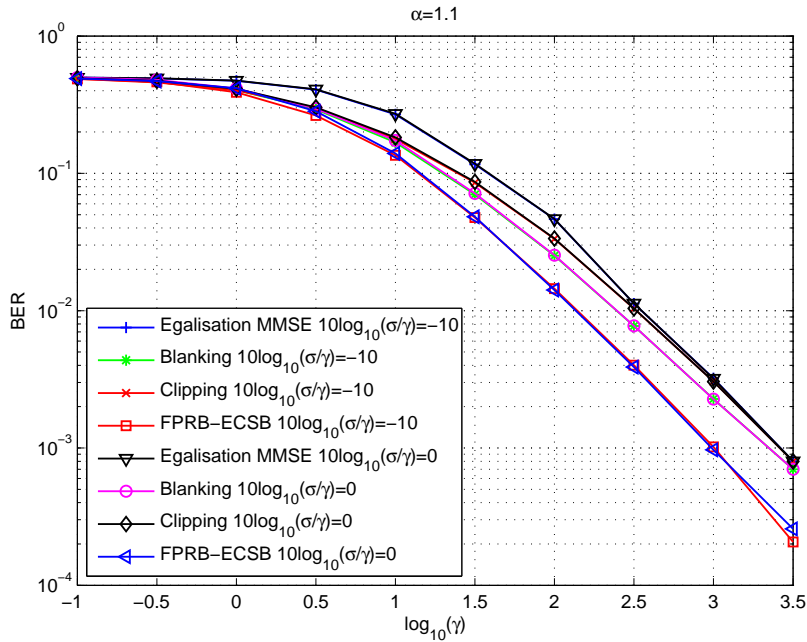
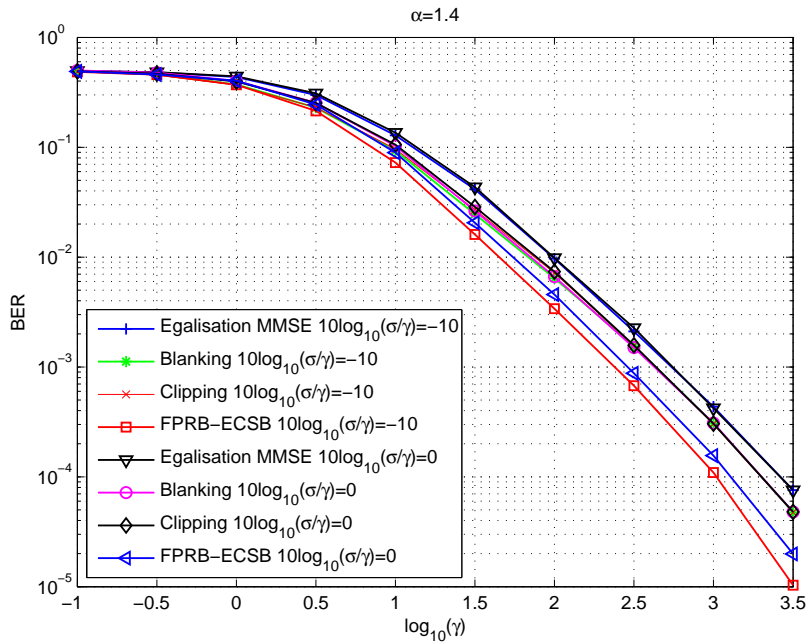


FIGURE 3.16 Performance en terme de taux d'erreur binaire de la méthode proposée FPRB-ECSB pour différentes valeurs du rapport $10 \log_{10}(\sigma/\gamma)$ avec $\alpha = 1.4$.



CHAPITRE 4

DP pour l'estimation non paramétrique de densités

Dans le chapitre 3, nous avons présenté une approche bayésienne paramétrique pour l'estimation conjointe du signal et des paramètres du bruit α -stable dans les systèmes OFDM. Les méthodes bayésiennes paramétriques consistent à modéliser le système à partir d'un ensemble fini de paramètres. L'inconvénient de ces méthodes est que l'on n'explore que quelques directions dans l'espace des paramètres possibles et que l'on ne s'adapte pas à de nouveaux modes qui pourraient survenir. Pour gagner en robustesse, il est alors souhaitable d'envisager des modèles qui ne sont pas limités aux paramétrisations finies, et donc de définir des distributions *a priori* sur des espaces de dimension infinie. Les méthodes bayésiennes non paramétriques permettent d'éviter les hypothèses souvent restrictives des modèles paramétriques en définissant des distributions sur les espaces fonctionnels. Elles permettent à la fois d'apprendre le comportement stochastique de commutation entre différents modes et de déterminer de manière dynamique de nouveaux modes au fur et à mesure que ceux-ci sont observés. Les modèles bayésiens non paramétriques les plus utilisés sont actuellement les processus de Dirichlet. Dans le chapitre 5, nous nous intéressons à l'estimation conjointe de l'état et de la densité de probabilité du bruit dans les modèles dynamiques non linéaires. L'estimation de la densité est réalisée à l'aide des processus de Dirichlet. Ces modèles ne sont pas encore très développés en traitement du signal. L'objectif de ce chapitre est de présenter les bases de l'estimation de la densité avec les processus de Dirichlet, nécessaires à la compréhension de la suite du

manuscrit. Une synthèse sur ce sujet est également disponible dans [Teh10, Car06].

Ce chapitre est organisé comme suit. Dans un premier temps, nous exposons le problème d'estimation de la densité de probabilité. Ensuite, nous introduisons les processus de Dirichlet ainsi que leurs principales caractéristiques. Puis, le principe de l'estimation de la densité de probabilité par les mélanges de processus de Dirichlet est décrit. Finalement, des algorithmes basés sur les méthodes MCMC permettant l'estimation de la densité de probabilité sont présentés.

4.1 Problème d'estimation de densité

Considérons un ensemble d'observations $\mathbf{y}_{1:n} = \{\mathbf{y}_i\}_{i=1}^n$, $\mathbf{y}_{1:n} \in \mathcal{Y}^n$, distribuées selon une densité de probabilité inconnue F :

$$\mathbf{y}_i \sim F(\mathbf{y}_i), \quad i = 1, \dots, n \quad (4.1)$$

L'objectif visé est d'estimer dans un cadre bayésien la densité de probabilité F à partir de la connaissance de l'ensemble d'observations $\mathbf{y}_{1:n}$. Nous nous intéressons tout particulièrement à la classe des densités de probabilité pouvant s'écrire sous la forme du modèle de mélange suivant :

$$F(\mathbf{y}) = \int_{\Theta} f(\mathbf{y}|\theta) d\mathbb{G}(\theta) \quad (4.2)$$

avec $\theta \in \Theta$ une variable latente, $f(\cdot|\theta)$ la densité mélangée connue et \mathbb{G} la distribution de mélange. Cette dernière est supposée inconnue de distribution *a priori* $p(\mathbb{G})$. \mathbb{G} est alors appelée mesure de probabilité aléatoire (RPM, Random Probability Measure). Les équations (4.1) et (4.2) peuvent être reformulées sous la forme hiérarchique suivante :

$$\begin{aligned} \mathbb{G} &\sim P(\mathbb{G}), \\ \theta_i | \mathbb{G} &\sim \mathbb{G}, \quad i = 1, \dots, n \\ \mathbf{y}_i | \theta_i &\sim f(\cdot|\theta_i), \quad i = 1, \dots, n \end{aligned} \quad (4.3)$$

Dans un cadre paramétrique, nous pouvons supposer que la distribution \mathbb{G} (et donc F) est caractérisée par un ensemble de paramètres inconnus de dimension finie

[WDLS99]. La densité F appartient donc à un espace fonctionnel \mathcal{F} de dimension finie. Cependant, contraindre la densité de probabilité à prendre une certaine forme paramétrique donnée peut limiter l'inférence réalisée à partir de tels modèles [MQ04]. Il est alors souhaitable de considérer que la densité F a un support plus large que celui fourni par un modèle paramétrique donné. Les modèles non paramétriques permettent d'éviter les hypothèses souvent restrictives des modèles paramétriques en définissant une distribution *a priori* sur des espaces fonctionnels (de dimension infinie). Les modèles non paramétriques peuvent ainsi être simplement définis comme des modèles paramétriques avec un nombre infini de paramètres.

Dans le cadre des modèles non paramétriques, la distribution aléatoire \mathbb{G} (et donc F) appartient à un espace de fonctions \mathcal{F} de dimension infinie. Dans un cadre bayésien, il est supposé que la RPM \mathbb{G} est distribuée selon une certaine distribution *a priori*, c'est-à-dire une distribution sur l'ensemble des distributions de probabilité. Plusieurs distributions *a priori* pour \mathbb{G} ont été proposées dans la littérature [WDLS99, MQ04, Sud06] dont les plus courantes sont les processus de Dirichlet.

4.2 Processus de Dirichlet

Les processus de Dirichlet définissent une mesure de probabilité dans l'espace des mesures de probabilité [Fer74, Fer73, Set94]. Ils permettent donc de définir, dans le cadre de l'estimation bayésienne, un *a priori* sur une distribution de probabilité inconnue. Un processus de Dirichlet est entièrement défini par deux paramètres : une mesure de probabilité de base \mathbb{G}_0 dans un espace mesurable Θ et un paramètre d'échelle α . Nous le notons par $DP(\mathbb{G}_0, \alpha)$.

Théorème 4.1. *Soit \mathbb{G}_0 une distribution de probabilité définie dans l'espace mesurable Θ , α un réel positif et $\{A_1, \dots, A_K\}$ une partition de Θ :*

$$\bigcup_{k=1}^K A_k = \Theta \quad \text{et} \quad A_k \cap A_j = \emptyset, \quad \forall \quad j \neq k. \quad (4.4)$$

Une distribution de probabilité \mathbb{G} sur Θ est distribuée selon un processus de Dirichlet de distribution de base \mathbb{G}_0 et de facteur d'échelle α , si sa mesure sur n'importe quelle

partition finie suit une distribution de Dirichlet :

$$[\mathbb{G}(A_1), \dots, \mathbb{G}(A_K)] \sim \mathcal{D}(\alpha \mathbb{G}_0(A_1), \dots, \alpha \mathbb{G}_0(A_K)). \quad (4.5)$$

Pour chaque distribution de base \mathbb{G}_0 et facteur d'échelle α , il existe un seul processus stochastique satisfaisant aux conditions ci-dessus, qu'on note $DP(\mathbb{G}_0, \alpha)$.

Les paramètres \mathbb{G}_0 et α jouent un rôle important pour la compréhension des processus de Dirichlet. En effet, en se basant sur l'équation (4.5), il est possible de montrer que l'espérance et la variance d'un processus de Dirichlet pour tout $A \in \Theta$ sont données par :

$$\mathbb{E}_{DP(\mathbb{G}_0, \alpha)}[\mathbb{G}(A)] = \mathbb{G}_0(A) \quad (4.6)$$

$$\text{Var}_{DP(\mathbb{G}_0, \alpha)}[\mathbb{G}(A)] = \frac{\mathbb{G}_0(A)(1 - \mathbb{G}_0(A))}{1 + \alpha} \quad (4.7)$$

Par conséquent, la distribution de base \mathbb{G}_0 représente la valeur moyenne du processus de Dirichlet. Par ailleurs, le paramètre d'échelle α peut être vu comme une variance inverse. Plus α est grand, plus la variance est petite et le processus de Dirichlet est concentré autour de la distribution de base \mathbb{G}_0 .

4.2.1 Représentation stick-breaking

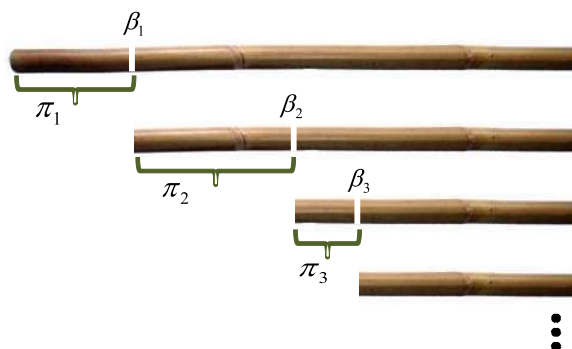
Nous considérons la fonction de masse $\{\beta_k\}_{k=1}^{\infty}$ sur un ensemble infini dénombrable, où les probabilités discrètes sont définies comme suit :

$$\beta_k | \alpha \sim \mathcal{B}(1, \alpha), \quad k = 1, 2, \dots \quad (4.8)$$

$$\pi_k = \beta_k \prod_{l=1}^{k-1} (1 - \beta_l), \quad k = 1, 2, \dots \quad (4.9)$$

où $\mathcal{B}(\cdot, \cdot)$ est la loi bêta. En effet, nous considérons un bâton de longueur initiale 1. A chaque instant k , nous cassons un morceau du bâton dont la longueur correspond à une proportion aléatoire β_k du bâton restant. Les poids π_k correspondent à la longueur des morceaux cassés. Cette façon d'obtenir les poids π_k est appelée construction stick-breaking ; son principe est illustré sur la figure 4.1. La construction stick-breaking est souvent notée $\pi \sim GEM(\alpha)$.

FIGURE 4.1 Représentation stick-breaking



Sethuraman [Set94] a montré qu'avec une probabilité 1, les réalisations d'un processus de Dirichlet $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ peuvent avoir la représentation stick-breaking suivante :

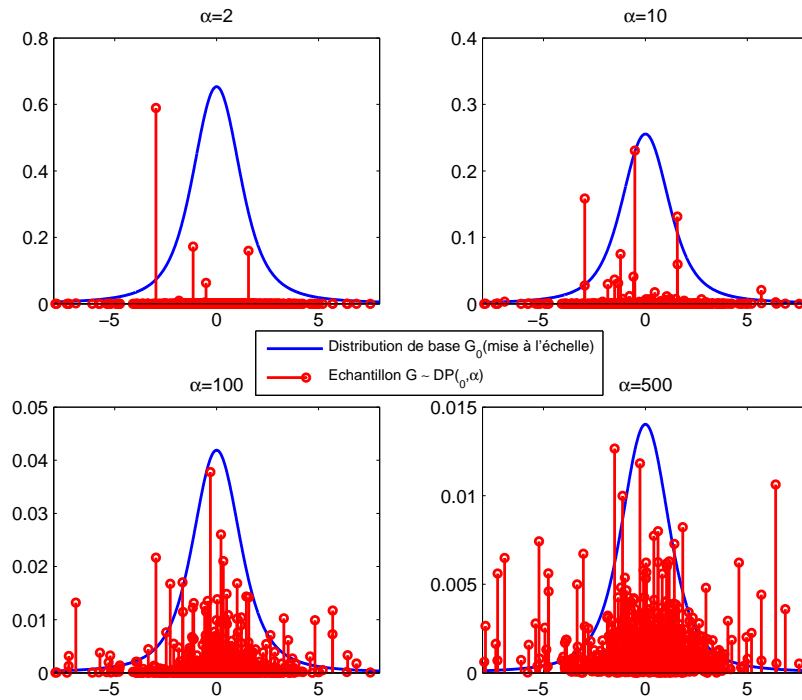
$$\mathbb{G} = \sum_{k=1}^{\infty} \pi_k \delta_{U_k} \quad \text{avec} \quad U_k | \mathbb{G}_0 \sim \mathbb{G}_0, \quad k = 1, 2, \dots \quad (4.10)$$

Les réalisations d'un processus de Dirichlet peuvent donc être vues comme des mélanges infinis dénombrables de mesures de Dirac. La distribution de base \mathbb{G}_0 détermine les positions des composantes discrètes de \mathbb{G} , tandis que le facteur d'échelle α contrôle la variance des poids. Par la suite, les positions $U_k, k = 1, 2, \dots$ de ces fonctions de Dirac seront notées clusters. Sur la figure 4.2 sont représentées les réalisations d'un processus de Dirichlet $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ avec $\mathbb{G}_0 \sim \mathcal{S}_1(1.4, 0, 1, 0)$ et différentes valeurs de α .

Les processus de Dirichlet appartiennent à une classe de processus appelée stick-breaking priors. Cette classe a été introduite par Ishwaran et James [IJ01] et elle inclut l'ensemble des processus pour lesquels la mesure de probabilité aléatoire prend la forme de l'équation (4.10). Les processus appartenant à cette classe diffèrent dans la façon d'obtenir les coefficients β_k . Ces derniers sont tirés, dans le cas général, de la distribution beta $\mathcal{B}(a_k, b_k)$. Parmi ces processus, on peut citer :

- le processus de beta à deux paramètres [IZ00] : $a_k = a$ et $b_k = b$;
- le processus de Poisson-Dirichlet à deux paramètres encore appelé processus de Pitman-Yor [PY97] : $a_k = 1 - a$ et $b_k = b + ka$.

FIGURE 4.2 Représentation des réalisations du processus de Dirichlet $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ de distribution de base $\mathbb{G}_0 \sim \mathcal{S}_1(1.4, 0, 1, 0)$ pour différentes valeurs du facteur d'échelle α . Ces réalisations sont obtenus à partir de la représentation stick-breaking (4.10).



4.2.2 Distribution *a posteriori*

Une motivation fondamentale pour l'utilisation du processus de Dirichlet comme RPM est la simplicité de la mise à jour de la distribution *a posteriori*. Soient n échantillons $\theta_1, \dots, \theta_n$ indépendants et identiquement distribués selon une distribution inconnue \mathbb{G}

$$\theta_1, \dots, \theta_n | \mathbb{G} \sim \mathbb{G} \quad (4.11)$$

et supposons que la RPM est distribuée selon un processus de Dirichlet de distribution de base \mathbb{G}_0 et de facteur d'échelle α , soit

$$\mathbb{G} \sim DP(\mathbb{G}_0, \alpha) \quad (4.12)$$

Nous nous intéressons à la détermination de la distribution *a posteriori* $p(\mathbb{G}|\theta_1, \dots, \theta_n)$. Soit $\{A_1, \dots, A_K\}$ une partition finie de Θ . En se basant sur l'équation (4.5) et le fait que la distribution de Dirichlet est une distribution *a priori* conjuguée pour la loi multinomiale, on peut montrer que :

$$[\mathbb{G}(A_1), \dots, \mathbb{G}(A_K)]|\theta_1, \dots, \theta_n \sim \mathcal{D} \left(\alpha \mathbb{G}_0(A_1) + \sum_{i=1}^n \delta_{\theta_i}(A_1), \dots, \alpha \mathbb{G}_0(A_K) + \sum_{i=1}^n \delta_{\theta_i}(A_K) \right) \quad (4.13)$$

L'expression ci-dessus est vraie pour toute partition finie de Θ , par conséquent en se basant sur le théorème 4.1, on peut déduire que la distribution *a posteriori* $p(\mathbb{G}|\theta_1, \dots, \theta_n)$ est également un processus de Dirichlet :

$$\mathbb{G}|\theta_1, \dots, \theta_n \sim DP(\mathbb{G}'_0, \alpha') \quad (4.14)$$

avec $\mathbb{G}'_0 = \frac{\alpha}{\alpha+n} \mathbb{G}_0 + \frac{1}{\alpha+n} \sum_{i=1}^n \delta_{\theta_i}$ et $\alpha' = \alpha + n$.

4.2.3 Représentation en urne de Polya

Considérons une RPM \mathbb{G} distribuée selon un processus de Dirichlet de distribution de base \mathbb{G}_0 et de paramètre d'échelle α :

$$\mathbb{G} \sim DP(\mathbb{G}_0, \alpha) \quad (4.15)$$

et soit $\{\theta_i\}_{i=1}^n$ un ensemble de n variables indépendantes et identiquement distribuées selon \mathbb{G} :

$$\theta_1, \dots, \theta_n | \mathbb{G} \sim \mathbb{G} \quad (4.16)$$

Blackwell et MacQueen [BM73] ont introduit une représentation en urne de Polya de la distribution prédictive qui peut être obtenue en marginalisant selon la RPM \mathbb{G} :

$$\theta_{n+1} | \theta_1, \dots, \theta_n \sim \frac{\alpha}{\alpha+n} \mathbb{G}_0 + \sum_{i=1}^n \frac{1}{\alpha+n} \delta_{\theta_i} \quad (4.17)$$

Autrement dit, sachant les variables latentes $\theta_{1:n}$ déjà échantillonnées, la probabilité que le nouvel échantillon θ_{n+1} soit identique à un échantillon précédent est $\frac{n}{\alpha+n}$, alors que la probabilité qu'il soit distribué indépendamment selon la distribution de base \mathbb{G}_0 est $\frac{\alpha}{\alpha+n}$.

La représentation en urne de Polya est très utile en pratique car elle permet d'échantillonner selon un processus de Dirichlet sans la construction explicite de la RPM.

4.2.4 Clustering

Nous avons vu précédemment que les réalisations d'un processus de Dirichlet sont des mélanges infinis dénombrables de mesures de Dirac. Par conséquent, la probabilité que plusieurs échantillons aient la même valeur, est strictement positive. Cela implique une propriété de clustering pour les processus de Dirichlet où les échantillons sont répartis en clusters $\{U_k\}_{k=1}^{\infty}$. On dit alors que ces échantillons sont associés au même cluster. De ce fait, la représentation en urne de Polya de la distribution prédictive peut aussi se réécrire en fonction de l'ensemble des clusters comme suit :

$$\theta_{n+1} | \theta_1, \dots, \theta_n \sim \frac{\alpha}{\alpha + n} \mathbb{G}_0 + \sum_{k=1}^K \frac{n_k}{\alpha + n} \delta_{U_k} \quad (4.18)$$

où $K \leq n$ est le nombre de valeurs distinctes de θ_i et n_k représente le nombre d'échantillons θ_i qui ont pris la valeur du cluster U_k . L'ensemble $\{U_k, k \in \mathbb{N} | \exists i \in 1, \dots, n, \theta_i = U_k\}$ est appelé l'ensemble des clusters animés. Sa cardinalité est bien évidemment K .

Dans la littérature, l'effet de clustering est souvent illustré à l'aide du processus du restaurant chinois. Ce dernier sera présenté dans le paragraphe suivant.

4.2.5 Processus du restaurant chinois

Soit $\mathbf{z} = (z_1, \dots, z_n)$ le vecteur des variables d'allocations qui attribue chaque variable latente θ_i à son cluster :

$$\theta_i = U_{z_i}, \quad i = 1, \dots, n \quad (4.19)$$

D'après l'équation (4.18), il est possible de montrer que la distribution prédictive des variables d'allocations peut s'écrire de la manière suivante :

$$p(z_{n+1} = z | z_1, \dots, z_n, \alpha) = \frac{\alpha}{\alpha + n} \delta_z(K + 1) + \frac{1}{\alpha + n} \sum_{k=1}^K n_k \delta_z(k) \quad (4.20)$$

où $n_k = \sum_{i=1}^n \delta_{z_i}(k)$ est le nombre de variables d'allocations prenant la valeur k et $K + 1$ est une valeur non encore observée.

La distribution sur des partitions induites par les distributions conditionnelles dans l'équation (4.20) porte le nom de processus du restaurant chinois. Dans cette métaphore, nous avons un restaurant chinois où se trouve une infinité de tables. Nous supposons que chacune de ces tables peut être occupée par un nombre infini de clients et que seul un nombre fini de tables est occupé par une ou plusieurs personnes. Un nouveau client qui entre a alors deux possibilités. Soit il se joint à une table déjà occupée (indiqué par k) avec une probabilité proportionnelle au nombre n_k de personnes à cette table. Soit il s'assoit à une nouvelle table $K + 1$ avec une probabilité proportionnelle à α .

4.2.6 Nombre de valeurs distinctes

D'après l'équation (4.20), on peut déduire que lorsque

$$z_i | \pi \sim \pi \quad \text{avec} \quad \pi | \alpha \sim GEM(\alpha) \quad (4.21)$$

il est possible de déterminer une expression pour la distribution prédictive de z_i en marginalisant selon π . Il est également possible de déterminer la distribution du nombre de valeurs distinctes de z_i (le nombre des tables occupées dans le processus du restaurant chinois) résultant de n tirages selon π . Soit K le nombre de valeurs distinctes de $\{z_i\}_{i=1}^n$, Antoniak [Ant74] a montré que la probabilité conditionnelle de M est donnée par

$$p(K|n, \alpha) = \frac{\Gamma(\alpha)}{\Gamma(\alpha + n)} s(n, K) \alpha^K \quad (4.22)$$

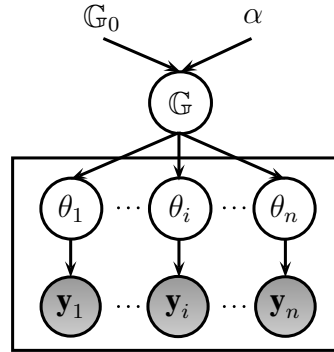
où $s(n, K)$ sont les nombres de Stirling de première espèce non signés [AS64].

En utilisant l'équation (4.22), Antoniak [Ant74] a aussi montré que :

$$\mathbb{E}[K|n, \alpha] \approx \alpha \log \left(\frac{\alpha + n}{\alpha} \right) \quad \text{lorsque} \quad n \rightarrow +\infty \quad (4.23)$$

Ainsi le facteur d'échelle α règle le nombre de valeurs distinctes. Quand α tend vers 0, la plupart des échantillons partagent la même valeur, tandis que lorsque α tend vers l'infini les échantillons sont quasiment indépendants et identiquement distribués selon la distribution de base \mathbb{G}_0 .

FIGURE 4.3 Modèle hiérarchique d'un mélange de processus de Dirichlet.



4.3 Mélange de processus de Dirichlet

En choisissant un processus de Dirichlet comme distribution *a priori* pour la RPM \mathbb{G} , le problème d'estimation de densité (4.3) peut être reformulé selon le modèle hiérarchique suivant :

$$\begin{aligned}
 \mathbb{G} &\sim DP(\mathbb{G}_0, \alpha), \\
 \theta_i | \mathbb{G} &\sim \mathbb{G}, \quad i = 1, \dots, n \\
 \mathbf{y}_i | \theta_i &\sim f(\cdot | \theta_i), \quad i = 1, \dots, n
 \end{aligned}
 \tag{4.24}$$

La distribution de mélange \mathbb{G} est distribuée selon un processus de Dirichlet de distribution de base \mathbb{G}_0 et de paramètre d'échelle α . Chacune des variables latentes θ_i est ensuite distribuée selon \mathbb{G} . Les observations \mathbf{y}_i sont finalement distribuées selon la densité mélangée $f(\cdot | \theta_i)$. Ce modèle hiérarchique, illustré sur la figure 4.3, est connu sous le nom de mélange de processus de Dirichlet (DPM, Dirichlet Process Mixture)[Ant74].

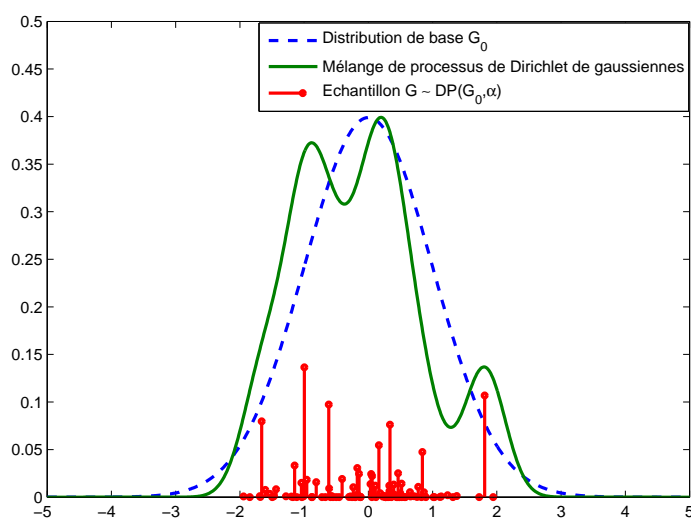
En utilisant les équations (4.2) et (4.10), une estimation de la densité inconnue F peut s'exprimer par :

$$F(\cdot) = \sum_{k=1}^{\infty} \pi_k f(\cdot | U_k)
 \tag{4.25}$$

avec $U_k \sim \mathbb{G}_0$, $\pi_k = \beta_k \prod_{l=1}^{k-1} (1 - \beta_l)$ et $\beta_k \sim \mathcal{B}(1, \alpha)$. La distribution inconnue F peut ainsi être vue comme un mélange infini dénombrable de densités de probabilité $f(\cdot | U_k)$.

La figure 4.4 représente une réalisation $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ d'un processus de Dirichlet de distribution de base $\mathbb{G}_0 = \mathcal{N}(0, 1)$ et de facteur d'échelle $\alpha = 10$ ainsi que la densité $F = \int_{\mathbb{G}} f(\cdot|\theta)d(\mathbb{G})$ associée avec une densité mélangée gaussienne $f(\cdot|\theta) = \mathcal{N}(\cdot; \theta, 0.1)$.

FIGURE 4.4 Représentation de la densité F associée à une réalisation $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ d'un processus de Dirichlet de distribution de base $\mathbb{G}_0 = \mathcal{N}(0, 1)$ et de facteur d'échelle $\alpha = 10$ avec une densité mélangée $f(\cdot|\theta) = \mathcal{N}(\cdot; \theta, 0.1)$.



Le mélange de processus de Dirichlet peut aussi être obtenu comme la limite de modèles de mélanges finis. Considérons, pour $p \in \mathbb{N}^*$, le modèle fini Dirichlet-multinomial suivant :

$$\begin{aligned}
 U_1, \dots, U_p | \mathbb{G}_0 &\sim \mathbb{G}_0 \\
 \pi_1, \dots, \pi_p | \alpha &\sim \mathcal{D}\left(\frac{\alpha}{p}, \dots, \frac{\alpha}{p}\right) \\
 z_i | \pi_{1:p} &\sim \text{Multinomial}(\pi_{1:p}), \quad i = 1, \dots, n \\
 \mathbf{y}_i | z_i, U &\sim f(\cdot | U_{z_i}), \quad i = 1, \dots, n
 \end{aligned} \tag{4.26}$$

Les variables $z_i \in \{1, \dots, p\}$, pour $i = 1, \dots, n$, sont appelées variables d'allocation d'une observation \mathbf{y}_i à un cluster U_{z_i} , et l'on a la relation $\theta_i = U_{z_i}$. Le vecteur de probabilité $\pi_{1:p}$ est distribué selon une distribution de Dirichlet $\mathcal{D}(\frac{\alpha}{p}, \dots, \frac{\alpha}{p})$, et les variables d'allocation z_i sont distribuées selon la loi multinomiale $\pi_{1:p}$. Dans [Nea00], Neal a montré que ce modèle est équivalent, lorsque p tend vers l'infini, au mélange de

processus de Dirichlet défini par l'équation (4.24). En marginalisant selon le vecteur de probabilité $\pi_{1:p}$, lorsque p tend vers l'infini, nous obtenons la représentation en urne de Polya de la distribution prédictive de z_i :

$$\begin{cases} p(z_{n+1} = k | z_1, \dots, z_n) = \frac{n_k}{\alpha+n} & k \in \{z_1, \dots, z_n\} \\ p(z_{n+1} \neq z_i \forall i \leq n | z_1, \dots, z_n) = \frac{\alpha}{\alpha+n} \end{cases} \quad (4.27)$$

Une variable d'allocation prend donc une valeur précédemment échantillonnée k proportionnellement à n_k , et prend une nouvelle valeur (correspondant à un nouveau cluster) proportionnellement à α . On peut facilement remarquer l'équivalence entre cette formulation de la distribution prédictive des variables d'allocation et celle définie par l'équation (4.20).

4.4 Estimation de la densité dans le cas d'un DPM

Dans un contexte bayésien, l'estimation de la densité inconnue F peut être vue comme un problème de prédiction de l'observation \mathbf{y}_{n+1} connaissant l'ensemble des observations passées $\mathbf{y}_{1:n}$. Ceci revient donc à déterminer la densité prédictive $p(\mathbf{y}_{n+1} | \mathbf{y}_{1:n})$ qui s'exprime par :

$$p(\mathbf{y}_{n+1} | \mathbf{y}_{1:n}) = \int_{\Theta} p(\mathbf{y}_{n+1} | \theta_{n+1}) p(\theta_{n+1} | \mathbf{y}_{1:n}) d\theta_{n+1} \quad (4.28)$$

A partir de l'équation (4.24), on sait que la densité $p(\mathbf{y}_{n+1} | \theta_{n+1}) = f(\mathbf{y}_{n+1} | \theta_{n+1})$. La distribution $p(\theta_{n+1} | \mathbf{y}_{1:n})$ peut s'écrire comme suit :

$$p(\theta_{n+1} | \mathbf{y}_{1:n}) = \int_{\Theta^n} p(\theta_{n+1} | \theta_{1:n}) p(\theta_{1:n} | \mathbf{y}_{1:n}) d\theta_{1:n} \quad (4.29)$$

avec $p(\theta_{n+1} | \theta_{1:n}) = \frac{\alpha}{\alpha+n} \mathbb{G}_0(\theta_{n+1}) + \frac{1}{\alpha+n} \sum_{i=1}^n \delta_{\theta_i}(\theta_{n+1})$. L'évaluation directe de la distribution *a posteriori* marginale $p(\theta_{1:n} | \mathbf{y}_{1:n})$ n'est pas possible et nécessite le recours à des techniques d'approximation. Supposons qu'on dispose de N échantillons $\theta_{1:n}^{(1)}, \dots, \theta_{1:n}^{(N)}$ générés selon $p(\theta_{1:n} | \mathbf{y}_{1:n})$, l'approximation de Monte Carlo de la distribution *a posteriori* marginale $p(\theta_{1:n} | \mathbf{y}_{1:n})$ est alors donnée par la distribution empirique

suivante :

$$P_N(\theta_{1:n}|\mathbf{y}_{1:n}) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{1:n}^{(i)}}(\theta_{1:n}) \quad (4.30)$$

A partir des équations (4.30) et (4.29), on peut déduire l'approximation suivante de la distribution $p(\theta_{n+1}|\mathbf{y}_{1:n})$:

$$P_N(\theta_{n+1}|\mathbf{y}_{1:n}) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{n+1}^{(i)}}(\theta_{n+1}) \quad (4.31)$$

où les échantillons $\theta_{n+1}^{(1)}, \dots, \theta_{n+1}^{(N)}$ sont obtenus en propageant chacun des échantillons $\theta_{1:n}^{(i)}$ par le noyau de transition $p(\theta_{n+1}|\theta_{1:n})$. Finalement, en substituant l'équation (4.31) dans l'équation (4.28), on obtient l'approximation suivante de la densité prédictive $p(\mathbf{y}_{n+1}|\mathbf{y}_{1:n})$:

$$p(\mathbf{y}_{n+1}|\mathbf{y}_{1:n}) \simeq \frac{1}{N} \sum_{i=1}^N f(\mathbf{y}_{n+1}|\theta_{n+1}^{(i)}) \quad (4.32)$$

Le problème maintenant est de savoir échantillonner selon la densité *a posteriori* marginale $p(\theta_{1:n}|\mathbf{y}_{1:n})$. Dans la littérature, plusieurs solutions ont été proposées notamment l'algorithme de d'Escobar [Esc94, EW95] et l'algorithme de Neal [Nea00]. Ces deux algorithmes sont présentés dans les paragraphes suivants.

4.4.1 Algorithme d'Escobar

Pour échantillonner selon la densité *a posteriori* marginale $p(\theta_{1:n}|\mathbf{y}_{1:n})$, Escobar *et al* [Esc94, EW95] proposent d'utiliser les méthodes MCMC par échantillonnage de Gibbs. L'idée consiste alors à échantillonner chaque θ_k selon la densité conditionnelle $p(\theta_k|\theta_{-k}, \mathbf{y}_{1:n})$, où $\theta_{-k} = \{\theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_n\}$. A partir de l'équation (4.17), on peut déduire l'expression suivante de la densité *a priori* conditionnelle $p(\theta_k|\theta_{-k})$ en supposant que θ_k est le dernier des n échantillons

$$\theta_k|\theta_{-k} \sim \frac{\alpha}{\alpha + n - 1} \mathbb{G}_0 + \frac{1}{\alpha + n - 1} \sum_{i=1, i \neq k}^n \delta_{\theta_i} \quad (4.33)$$

et en appliquant la règle de Bayes, on aboutit à l'expression suivante de la densité conditionnelle $p(\theta_k|\theta_{-k}, \mathbf{y}_{1:n})$:

$$\theta_k|\theta_{-k}, \mathbf{y}_{1:n} \sim r_k H_k + \sum_{i=1, i \neq k}^n q_{k,i} \delta_{\theta_i} \quad (4.34)$$

avec :

- $H_k \propto f(\mathbf{y}_k|\theta_k)\mathbb{G}_0(\theta_k)$ est la densité *a posteriori* de θ_k basée sur la densité *a priori* \mathbb{G}_0 et l'observation \mathbf{y}_k ;
- $r_k = C\alpha \int_{\Theta} f(\mathbf{y}_k|\theta_k)d\mathbb{G}_0$;
- $q_{k,i} = C f(\mathbf{y}_k|\theta_i)$ où C est une constante de normalisation telle que $\sum_{i=1, i \neq k}^n q_{k,i} + r_k = 1$.

Pour pouvoir échantillonner directement selon la densité conditionnelle $p(\theta_k|\theta_{-k}, \mathbf{y}_k)$, il faut pouvoir calculer l'intégrale $\int_{\Theta} f(\mathbf{y}_k|\theta_k)d\mathbb{G}_0$. Ceci n'est possible que si la densité mélangée $f(\cdot|\theta)$ et la distribution de base \mathbb{G}_0 sont conjuguées. Dans le cas contraire, on peut échantillonner selon $p(\theta_k|\theta_{-k}, \mathbf{y}_k)$ à partir d'un pas de Metropolis-Hastings en choisissant comme densité de probabilité candidate la densité *a priori* conditionnelle $p(\theta_k|\theta_{-k})$.

L'algorithme 13 résume les étapes nécessaires pour échantillonner selon la densité *a posteriori* marginale $p(\theta_{1:n}|\mathbf{y}_{1:n})$.

Algorithme 13 Algorithme de Gibbs pour échantillonner selon $p(\theta_{1:n}|\mathbf{y}_{1:n})$

• **Initialisation**

pour $k = 1$ à n **faire**

– Échantillonner $\theta_k^{(1)}$

fin pour

• **Itérations**

pour $i = 2$ à N **faire**

pour $k = 1$ à n **faire**

– Échantillonner $\theta_k^{(i)} \sim p(\theta_k^{(i)}|\theta_1^{(i)}, \dots, \theta_{k-1}^{(i)}, \theta_{k+1}^{(i-1)}, \theta_n^{(i-1)}, \mathbf{y}_{1:n})$:

– En utilisant l'équation (4.34) si $f(\cdot|\theta)$ et \mathbb{G}_0 sont conjuguées.

– Sinon, en faisant un pas de Metropolis-Hastings :

– Échantillonner $\theta_k^{(i)*} \sim p(\theta_k^{(i)*}|\theta_{-k}^{(i)})$ en utilisant l'équation (4.33)

– Avec la probabilité $\rho(\theta_k^{(i-1)}, \theta_k^{(i)*})$ faire $\theta_k^{(i)} = \theta_k^{(i)*}$, sinon faire $\theta_k^{(i)} = \theta_k^{(i-1)}$

fin pour

fin pour

Cet algorithme produit une chaîne de Markov ergodique, mais la convergence vers la distribution *a posteriori* peut être assez lente, et l'échantillonnage peut s'avérer inefficace. Le problème est qu'il y a souvent des groupes d'observations qui sont associés, avec une forte probabilité, à la même θ_k . Et, puisque l'algorithme ne peut pas changer les θ_k pour plus d'une observation simultanément, un changement des valeurs de θ_k dans un tel groupe ne peut se produire que très rarement, car il nécessite le passage par un état intermédiaire de faible probabilité dans lequel les observations du groupe ne sont pas associées à la même valeur θ_k [Nea00]. Une solution à ce problème a été proposée par Neal [Nea00]. Celle-ci fera l'objet du paragraphe suivant.

4.4.2 Algorithme de Neal

Afin d'éviter le problème de convergence lente de l'algorithme d'Escobar, Neal [Nea00] a proposé un algorithme similaire dans lequel l'échantillonneur de Gibbs est appliqué sur le modèle (4.26). Avec un tel modèle, la densité *a posteriori* d'intérêt devient $p(z_{1:n}, U_{X(z_{1:n})} | \mathbf{y}_{1:n})$, où $X(z_{1:n})$ et $U_{X(z_{1:n})}$ dénotent respectivement l'ensemble des valeurs distinctes prises par les variables d'allocation $z_{1:n}$ et l'ensemble des clusters associés. Cet algorithme procède en deux étapes. On commence par échantillonner les variables d'allocations z_k , afin de définir des groupes d'observations. Ensuite, pour chacun de ces groupes, la valeur de son cluster U_{z_k} est mise à jour. Cette décomposition en deux étapes de l'échantillonnage de Gibbs permet donc de résoudre le problème de l'algorithme précédent, puisque la valeur d'un cluster associé à un groupe d'observations est désormais modifiée pour l'ensemble de celles-ci.

On note $z_{-k} = \{z_i, i = 1, \dots, n \forall i \neq k\}$. Pour chaque variable d'allocation, l'échantillonnage de Gibbs est basé sur la densité conditionnelle $p(z_k = j | z_{-k}, \mathbf{y}_{1:n}, U_{X(z_{-k})})$. En appliquant la règle de Bayes, cette densité s'exprime comme suit :

$$p(z_k = j | z_{-k}, \mathbf{y}_{1:n}, U_{X(z_{-k})}) \propto p(z_k = j | z_{-k}, U_{X(z_{-k})}) f(\mathbf{y}_k | U_j) \quad (4.35)$$

Or, la densité conditionnelle *a priori* est donnée, d'après l'équation (4.27), par :

$$\begin{cases} p(z_k = j | z_{-k}) = \frac{n_{-k,j}}{\alpha + n - 1} & \forall j \in X(z_{-k}) \\ p(z_k \neq z_i \forall i \neq k | z_{-k}) = \frac{\alpha}{\alpha + n - 1} \end{cases} \quad (4.36)$$

où $n_{-k,j}$ est le nombre de $z_i \in z_{-k}$ qui sont égaux à j . En substituant (4.36)

dans (4.35), on obtient l'expression suivante de la densité conditionnelle $p(z_k = j|z_{-k}, \mathbf{y}_{1:n}, U_{X(z_{-k})})$:

$$\begin{cases} p(z_k = j|z_{-k}, \mathbf{y}_k, U_{X(z_{-k})}) = b \frac{n-k,j}{\alpha+n-1} f(\mathbf{y}_k|U_j) & \forall j \in X(z_{-k}) \\ p(z_k \neq z_i \forall i \neq k|z_{-k}, \mathbf{y}_k, U_{X(z_{-k})}) = b \frac{\alpha}{\alpha+n-1} \int_{\Theta} f(\mathbf{y}_k|\theta) d\mathbb{G}_0(\theta) \end{cases} \quad (4.37)$$

où b est une constante de normalisation telle que la somme de ces probabilités vaut 1. De la même façon que l'algorithme précédent, il faut que la densité mélangée $f(\cdot|\theta)$ et la distribution de base \mathbb{G}_0 soient conjuguées pour pouvoir échantillonner directement selon la densité conditionnelle $p(z_k \neq z_i \forall i \neq k|z_{-k}, \mathbf{y}_k, U_{X(z_{-k})})$. Dans le cas contraire, on peut faire un pas de Metropolis-Hastings en utilisant la densité conditionnelle *a priori* $p(z_k = j|z_{-k})$ comme densité de probabilité candidate.

Après l'échantillonnage des variables d'allocation, les clusters animés $U_j, j \in X(z_{1:n})$ sont mis à jour selon la densité suivante :

$$p(U_j|z_{1:n}^{(i)}, \mathbf{y}_{1:n}) \propto \mathbb{G}_0(U_j) \prod_{i|z_i=j} f(\mathbf{y}_i|U_j) \quad (4.38)$$

Le calcul analytique de l'équation (4.38) n'est possible que lorsque $f(\cdot|\theta)$ et \mathbb{G}_0 sont conjuguées. Dans le cas contraire, on peut faire un pas de Metropolis-Hastings en choisissant la distribution de base \mathbb{G}_0 comme densité de probabilité candidate.

Les étapes nécessaires pour échantillonner selon $p(z_{1:n}, U_{X(z_{1:n})}|\mathbf{y}_{1:n})$ sont données par l'algorithme 14.

4.5 Conclusion

Dans ce chapitre, nous avons présenté des méthodes bayésiennes non paramétriques pour l'estimation de densité par les mélanges de processus de Dirichlet. Ces modèles admettent des propriétés très intéressantes qui les rendent particulièrement attractifs dans un cadre d'estimation car elles permettent la mise en œuvre de mécanismes simplifiés des principes d'inférence bayésienne. Tout d'abord, la représentation en urne de Polya permet l'échantillonnage selon la densité *a posteriori* sans la construction explicite de la RPM. De plus, son caractère discret permet d'estimer directement le nombre de composantes du mélange, et de s'affranchir de techniques additionnelles telles que les sauts réversibles. Dans le chapitre suivant, nous mettons

Algorithme 14 Algorithme de Gibbs pour échantillonner selon $p(z_{1:n}, U_{X(z_{1:n})} | \mathbf{y}_{1:n})$

- **Initialisation**

pour $k = 1$ à n **faire**

- Échantillonner $z_k^{(1)}$

fin pour

- **Itérations**

pour $i = 2$ à N **faire**

pour $k = 1$ à n **faire**

- Échantillonner $z_k^{(i)} \sim p(z_k^{(i)} | z_1^{(i)}, \dots, z_{k-1}^{(i)}, z_{k+1}^{(i-1)}, z_n^{(i-1)}, \mathbf{y}_{1:n})$:
- En utilisant l'équation (4.37) si $f(\cdot | \theta)$ et \mathbb{G}_0 sont conjuguées.
- Sinon, en faisant un pas de Metropolis-Hastings.

si $n_{-k, z_k^{(i)}} = 0$ **alors**

- Échantillonner $U_{z_k^{(i)}}^{(i-1)} \sim \mathbb{G}_0(U_{z_k^{(i)}}^{(i-1)}) f(\mathbf{y}_k | U_{z_k^{(i)}}^{(i-1)})$

fin si

fin pour

- Pour $j \in X(z_{1:n}^{(i)})$, échantillonner $U_j^{(i)} \sim p(U_j | z_{1:n}^{(i)}, \mathbf{y}_{1:n})$

- En utilisant l'équation (4.38) si $f(\cdot | \theta)$ et \mathbb{G}_0 sont conjuguées.
- Sinon, en faisant un pas de Metropolis-Hastings.

fin pour

à profit ces propriétés pour proposer une approche non paramétrique bayésienne qui permet l'estimation conjointe de l'état et de la densité du bruit d'observation dans les systèmes dynamiques non linéaires en présence de bruit α -stable.

DPMs pour l'estimation conjointe état-bruit impulsif

Dans cette thèse, nous nous intéressons au problème de l'estimation conjointe du signal et du bruit dans les applications fonctionnant dans des environnements impulsifs. Dans un tel cadre, les perturbations rencontrées s'éloignent fortement du modèle gaussien classique. Comme nous l'avons vu dans le chapitre 1, les distributions α -stables sont adaptées à la modélisation des phénomènes impulsifs. Dans le chapitre 3, une approche bayésienne basée sur les méthodes de Monte Carlo séquentielles a été proposée pour l'estimation conjointe du signal et des paramètres de bruit dans les systèmes de transmission OFDM. Dans ce chapitre, nous nous proposons d'étudier cette problématique dans un cadre applicatif plus large, celui des systèmes non linéaires. À cet effet, nous considérons le modèle dynamique non linéaire suivant :

$$\mathbf{x}_t = g_t(\mathbf{x}_{t-1}) + \mathbf{w}_t \quad (5.1)$$

$$\mathbf{y}_t = h_t(\mathbf{x}_t) + \mathbf{v}_t \quad (5.2)$$

où \mathbf{x}_t est le vecteur d'état caché, \mathbf{y}_t est l'observation, g_t et h_t sont les fonctions connues d'évolution de l'état et d'observation, \mathbf{w}_t est le bruit d'état et \mathbf{v}_t est le bruit d'observation. Nous supposons que le bruit \mathbf{w}_t est gaussien de moyenne μ_t^w et de matrice de covariance Σ_t^w connues. Le bruit \mathbf{v}_t est supposé être α -stable. Aucune restriction n'est imposée sur ce bruit. Il peut être unimodal ou multimodal, symétrique ou non symétrique, etc.

Étant donné ce modèle dynamique, notre objectif est alors d'estimer en ligne l'état \mathbf{x}_t et la densité du bruit \mathbf{v}_t à partir des observations. Pour ce faire, nous utilisons les DPMs introduits dans le chapitre précédent et nous proposons une approche bayésienne non paramétrique qui repose sur la modélisation du bruit \mathbf{v}_t par un mélange de processus de Dirichlet. Trois modélisations différentes du bruit d'observation sont alors formulées : modèle par DPM de distributions gaussiennes, modèle par DPMs de distributions de Cauchy et finalement modèle par DPM de distributions α -stables. Le modèle par DPM de distributions α -stables sera utilisé comme modèle de référence lors du test. Le choix des deux autres modèles a été motivé par plusieurs arguments. Tout d'abord, les lois de Gauss et de Cauchy sont des cas particuliers des distributions α -stables. De ce fait, nos modèles sont déjà valides pour certaines valeurs de α . De plus, le choix du modèle par DPM de distributions gaussiennes est motivé par les travaux de Caron [Car06, CDD⁺08], dans lesquels ce même modèle a été utilisé pour modéliser les densités de probabilité des séquences de bruit inconnues dans les modèles dynamiques linéaires. Ensuite, il est théoriquement possible de justifier l'utilisation d'un DPM de gaussiennes pour approcher une variable issue d'une distribution α -stable symétrique ou d'un mélange de distributions α -stables symétriques. En effet, d'après la représentation SMiN d'une variable α -stable symétrique (cf. corollaire 1.1), une distribution α -stable symétrique peut se représenter comme un mélange infini de distributions gaussiennes. Dans ce mélange, le poids de chaque composante gaussienne est donné par la valeur de la pdf d'une distribution α -stable positive. En se basant sur ce résultat, Kuruoglu *et al* [KMF98] ont proposé une méthode pour approcher la densité de probabilité d'une distribution α -stable symétrique par un mélange fini de distributions gaussiennes. L'obtention d'une bonne approximation nécessite l'utilisation d'un grand nombre de composantes. Ceci aboutit à un algorithme très complexe en temps de calcul vu qu'on est obligé d'évaluer pour chaque composante la pdf d'une distribution α -stable. Pour réduire la complexité de l'algorithme, les auteurs proposent d'utiliser un nombre réduit de composantes et ensuite faire appel à l'algorithme EM pour corriger l'approximation. Or, l'utilisation d'un DPM de distributions gaussiennes permet d'approcher naturellement une distribution α -stable symétrique par un mélange infini de distributions gaussiennes. Ceci découle directement de la définition même d'un DPM. Donc, l'idée d'utiliser un DPM de distributions gaussiennes pour modéliser une distribution α -stable symétrique ou un mélange de distributions α -

stables symétriques est tout à fait pertinente. Et par rapport à la méthode proposée par Kuruoglu *el al*, notre modèle permet de s'affranchir de l'évaluation numérique de la pdf des distributions α -stables ainsi que l'utilisation d'une étape EM. Le seul souci avec ce modèle est que la distribution gaussienne n'est pas une densité à queue lourde donc ce modèle risque de mal estimer les queues de distributions. Le choix du modèle par DPMs de distributions de Cauchy est motivé par le fait que la loi de Cauchy appartient à la classe des distributions à queues lourdes.

Ce chapitre a une double finalité. D'une part, nous essayons de montrer que réaliser une inférence de l'état en présence d'un bruit extrêmement impulsif et de fortes non linéarités est tout à fait possible même lorsque les statistiques de ce bruit sont inconnues et d'autre part, nous voulons montrer qu'un processus α -stable peut être modélisé par une mélange infini de lois de Cauchy ou de Gauss. Ceci est d'un grand intérêt dans le cadre de ces distributions qui n'admettent pas d'expression analytique pour leur fonction de densité de probabilité.

Le reste de ce chapitre est organisé de la façon suivante. Tout d'abord, nous présentons les différentes modélisations proposées pour le bruit d'observation. Ensuite, des algorithmes de Monte Carlo séquentiels permettant l'estimation conjointe de l'état et de la densité du bruit d'observation ainsi que des densités d'importance efficaces sont exposés. Finalement, les algorithmes proposés sont testés sur un exemple illustratif et une étude comparative des trois modèles est menée.

5.1 Modélisation du bruit d'observation avec les DPMs

Notre approche repose principalement sur la modélisation de la densité de probabilité du bruit d'observation F à l'aide d'un DPM comme suit :

$$F(\mathbf{v}_t) = \int_{\Theta} f(\cdot|\theta_t)d\mathbb{G}(\theta_t) \quad (5.3)$$

où θ_t est la variable latente, $f(\cdot|\theta_t)$ est la densité mélangée et \mathbb{G} est la distribution de mélange distribuée selon un processus de Dirichlet de distribution de base \mathbb{G}_0 et de facteur d'échelle $\alpha_{DPM} : \mathbb{G} \sim DP(\mathbb{G}_0, \alpha_{DPM})$. Autrement dit, F peut être vue comme un mélange infini dénombrable de densités de probabilités avec une forme paramétrique dépendant de la densité mélangée adoptée. Dans cette thèse, trois

modèles sont considérés à savoir : modèle avec DPM de distributions gaussiennes, modèle avec DPM de distributions de Cauchy et modèle avec DPM de distributions α -stables. La différence entre ces trois modèles réside dans le choix de la densité mélangée adoptée. Dans cette section, une description détaillée de chacun de ces modèles est donnée.

5.1.1 Modèle de bruit avec DPM de distributions gaussiennes

Ce modèle a été initialement proposé par Caron *et al* pour modéliser les densités de probabilité inconnues des bruits d'état et d'observation [Car06, CDD⁺08].

La densité mélangée $f(.|\theta_t)$ est supposée être une gaussienne de moyenne μ_t et de matrice de covariance Σ_t notée : $\mathcal{N}(\mu_t, \Sigma_t)$. La distribution de base \mathbb{G}_0 est supposée être une distribution normale-inverse Wishart notée $\mathbb{G}_0 = \mathcal{NIW}(\mu_0, \kappa_0, \nu_0, \Lambda_0)$. Cette distribution est souvent utilisée en statistiques pour définir une loi *a priori* sur les paramètres d'une distribution gaussienne multi-variée. Avec une telle distribution de base, les lois *a priori* des paramètres μ_t et Σ_t sont définies comme suit :

$$\mu_t | \Sigma_t \sim \mathcal{N}\left(\mu_0, \frac{\Sigma_t}{\kappa_0}\right) \quad (5.4)$$

$$\Sigma_t^{-1} \sim \mathcal{W}(\nu_0, \Lambda_0^{-1}) \quad (5.5)$$

où $\mathcal{W}(. , .)$ est la loi Wishart.

En résumé, nous avons le modèle hiérarchique suivant :

$$\begin{aligned} \mathbb{G} | \Phi &\sim DP(\mathbb{G}_0, \alpha_{DPM}) \\ \theta_t | \mathbb{G} &\sim \mathbb{G} \\ \mathbf{v}_t | \theta_t &\sim \mathcal{N}(\mu_t, \Sigma_t) \end{aligned} \quad (5.6)$$

où $\Phi = \{\alpha_{DPM}, \mu_0, \kappa_0, \nu_0, \Lambda_0\}$ et $\theta_t = \{\mu_t, \Sigma_t\}$ sont respectivement l'ensemble des hyperparamètres et la variable latente donnant à chaque instant t la moyenne et la matrice de covariance de la distribution gaussienne.

Ce modèle peut s'écrire de façon équivalente $\mathbf{v}_t \sim F(\mathbf{v}_t)$ où $F(\mathbf{v}_t)$ est la densité de probabilité du bruit d'observation définie comme suit :

$$F(\mathbf{v}_t) = \int \mathcal{N}(\mathbf{v}_t; \mu, \Sigma) d\mathbb{G}(\mu, \Sigma) \quad (5.7)$$

La densité F peut être vue comme un mélange infini dénombrable de distributions gaussiennes dont les paramètres sont inconnus et distribués selon la distribution de mélange \mathbb{G} . Cette dernière est distribuée selon un processus de Dirichlet $DP(\mathbb{G}_0, \alpha_{DPM})$.

5.1.2 Modèle de bruit avec DPM de distributions de Cauchy

Avec ce modèle, la densité mélangée $f(\cdot|\theta_t)$ est supposée être une distribution de Cauchy de paramètre d'échelle a_t et de paramètre de position m_t , notée $\mathcal{C}(a_t, m_t)$. La variable latente est donnée dans ce cas par : $\theta_t = \{a_t, m_t\}$. Nous choisissons une loi inverse-gamma de paramètre de forme a^a et de paramètre d'échelle b^a comme loi *a priori* pour a_t et une loi normale de moyenne m^m et de variance Σ^m comme loi *a priori* pour m_t . La distribution de base \mathbb{G}_0 est définie comme étant le produit des lois *a priori* associées aux paramètres de la densité mélangée :

$$\mathbb{G}_0 = \mathcal{IG}(a^a, b^a) \times \mathcal{N}(m^m, \Sigma^m)$$

L'ensemble des hyperparamètres est donné par $\Phi = \{\alpha_{DPM}, a^a, b^a, m^m, \Sigma^m\}$ et nous obtenons le modèle hiérarchique suivant :

$$\begin{aligned} \mathbb{G}|\Phi &\sim DP(\mathbb{G}_0, \alpha_{DPM}) \\ \theta_t|\mathbb{G} &\sim \mathbb{G} \\ \mathbf{v}_t|\theta_t &\sim \mathcal{C}(a_t, m_t) \end{aligned} \tag{5.8}$$

Dans ce cas, la densité de probabilité du bruit d'observation $F(\mathbf{v}_t)$ est définie par

$$F(\mathbf{v}_t) = \int \mathcal{C}(a, m) d\mathbb{G}(a, m)$$

et elle peut être considérée comme un mélange infini dénombrable de distributions de Cauchy.

5.1.3 Modèle de bruit avec DPM de distributions α -stables

La densité mélangée $f(\cdot|\theta_t)$ est distribuée selon une distribution α -stable notée $\mathcal{S}_1(\alpha_t, \beta_t, \gamma_t, \mu_t)$. A chaque instant t , les paramètres de la densité mélangée sont

donnés par la variables latente $\theta_t = \{\alpha_t, \beta_t, \gamma_t, \mu_t\}$. Les lois *a priori* des différents paramètres de la densité mélangée sont choisies comme suit :

- L'exposant caractéristique α_t et le paramètre de symétrie β_t sont uniformément distribués sur leurs supports : $\alpha_t \sim \mathcal{U}[0, 2]$ et $\beta_t \sim \mathcal{U}[-1, 1]$;
- Le paramètre de position μ_t est distribué selon une distribution gaussienne de moyenne m^μ et de variance Σ^μ notée $\mu_t \sim \mathcal{N}(m^\mu, \Sigma^\mu)$;
- Le paramètre de dispersion γ_t suit une distribution inverse-gamma de paramètre de forme a^γ et de paramètre d'échelle b^γ notée $\gamma_t \sim \mathcal{IG}(a^\gamma, b^\gamma)$.

La distribution de base \mathbb{G}_0 est alors définie comme suit :

$$\mathbb{G}_0 = \mathcal{U}[0, 2] \times \mathcal{U}[-1, 1] \times \mathcal{IG}(a^\gamma, b^\gamma) \times \mathcal{N}(m^\mu, \Sigma^\mu)$$

Dans ce cas, l'ensemble des hyperparamètres est donné par : $\Phi = \{\alpha_{DPM}, a^\gamma, b^\gamma, m^\mu, \Sigma^\mu\}$ et nous obtenons le modèle hiérarchique suivant :

$$\begin{aligned} \mathbb{G} | \Phi &\sim DP(\mathbb{G}_0, \alpha_{DPM}) \\ \theta_t | \mathbb{G} &\sim \mathbb{G} \\ \mathbf{v}_t | \theta_t &\sim \mathcal{S}_1(\alpha_t, \beta_t, \gamma_t, \mu_t) \end{aligned} \tag{5.9}$$

La densité de probabilité du bruit d'observation $F(\mathbf{v}_t)$ peut être définie comme étant un mélange infini dénombrable de distributions α -stables comme suit :

$$F(\mathbf{v}_t) = \int \mathcal{S}_1(\alpha, \beta, \gamma, \mu) d\mathbb{G}(\alpha, \beta, \gamma, \mu)$$

Les paramètres des composantes α -stables du mélange sont inconnus et distribués selon la distribution de mélange \mathbb{G} . Cette dernière est distribuée selon un processus de Dirichlet $DP(\mathbb{G}_0, \alpha_{DPM})$.

5.2 Filtre particulière pour l'estimation conjointe de l'état et de la densité de probabilité du bruit d'observation

Etant donné le modèle dynamique défini par les équations (5.1) et (5.2), notre objectif est l'estimation conjointe de l'état \mathbf{x}_t et de la densité inconnue du bruit d'observation F à chaque instant t , conditionnellement aux observations $\mathbf{y}_{1:t}$. Dans ce travail, les hyperparamètres de la distribution de base \mathbb{G}_0 sont supposés fixes et connus. Les variables d'intérêt sont donc le vecteur d'état \mathbf{x}_t , la variable latente θ_t et la distribution de mélange \mathbb{G} . Grâce à la représentation en urne de Polya, il est possible de marginaliser analytiquement selon la distribution de mélange \mathbb{G} et de s'affranchir de son estimation explicite. Il nous reste donc à estimer le vecteur d'état \mathbf{x}_t et la variable latente θ_t . Dans un cadre bayésien, ceci revient à calculer la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$. Cependant, le calcul analytique de cette densité n'est pas possible. Pour cette raison, nous cherchons une approximation de cette densité par les méthodes de Monte Carlo séquentielles. En effet, cette densité est approchée à l'aide de N particules pondérées par la distribution empirique :

$$p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi) \simeq \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta_{\tilde{\mathbf{x}}_{0:t}^{(i)}, \tilde{\theta}_{1:t}^{(i)}}(\mathbf{x}_{0:t}, \theta_{1:t}) \quad (5.10)$$

A partir de cette approximation de la densité *a posteriori*, l'estimation du vecteur d'état \mathbf{x}_t est obtenue grâce à l'estimateur MMSE. Généralement, nous nous intéressons à l'estimée MMSE (cf. équation (2.11)) marginale $\hat{\mathbf{x}}_{t|t}^{MMSE}$ défini par :

$$\begin{aligned} \hat{\mathbf{x}}_{t|t}^{MMSE} &= \int_{\mathcal{X}^{t+1}} \mathbf{x}_t p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}, \Phi) d\mathbf{x}_{0:t} \\ &= \int_{\mathcal{X}^{t+1}} \int_{\Theta^t} \mathbf{x}_t p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi) d\theta_{1:t} d\mathbf{x}_{0:t} \\ &= \int_{\mathcal{X}} \int_{\Theta^t} \mathbf{x}_t p(\mathbf{x}_t, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi) d\theta_{1:t} d\mathbf{x}_t \end{aligned} \quad (5.11)$$

En intégrant l'approximation de la densité *a posteriori* donnée par l'équation (5.10)

dans l'expression précédente, $\widehat{\mathbf{x}}_{t|t}^{MMSE}$ est alors approchée par :

$$\widehat{\mathbf{x}}_{t|t}^{MMSE} \approx \sum_{i=1}^N \tilde{\omega}_t^{(i)} \tilde{\mathbf{x}}_t^{(i)} \quad (5.12)$$

L'estimation de la densité du bruit d'observation est donnée par la densité prédictive *a posteriori*

$$\widehat{F}_t = \mathbb{E} [F | \mathbf{y}_{1:t}] \quad (5.13)$$

Cette densité peut également s'exprimer par :

$$\widehat{F}_t(\mathbf{v}_{t+1}) = p(\mathbf{v}_{t+1} | \mathbf{y}_{1:t}) \quad (5.14)$$

où \mathbf{v}_{t+1} est un échantillon du bruit d'observation qui n'est pas encore observé. En faisant intervenir l'approximation de la densité *a posteriori* donnée par l'équation (5.10), cette densité peut s'écrire

$$\begin{aligned} \widehat{F}_t(\mathbf{v}_{t+1}) &= p(\mathbf{v}_{t+1} | \mathbf{y}_{1:t}) \\ &= \int_{\Theta^t} \int_{\mathcal{X}^{t+1}} p(\mathbf{v}_{t+1} | \theta_{1:t}, \mathbf{x}_{0:t}) p(\theta_{1:t}, \mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\theta_{1:t} d\mathbf{x}_{0:t} \\ &\simeq \sum_{i=1}^N \tilde{\omega}_t^{(i)} p(\mathbf{v}_{t+1} | \tilde{\theta}_{1:t}^{(i)}, \tilde{\mathbf{x}}_{0:t}^{(i)}) \\ &\simeq \sum_{i=1}^N \tilde{\omega}_t^{(i)} p(\mathbf{v}_{t+1} | \tilde{\theta}_{1:t}^{(i)}) \end{aligned} \quad (5.15)$$

En notant θ_{t+1} la variable latente associée à l'échantillon \mathbf{v}_{t+1} , la densité $p(\mathbf{v}_{t+1} | \theta_{1:t})$ peut s'écrire :

$$p(\mathbf{v}_{t+1} | \theta_{1:t}) = \int_{\Theta} f(\mathbf{v}_{t+1} | \theta_{t+1}) p(\theta_{t+1} | \theta_{1:t}) d\theta_{t+1} \quad (5.16)$$

En utilisant la loi de prédiction en urne de Polya, nous obtenons :

$$\begin{aligned} p(\mathbf{v}_{t+1} | \theta_{1:t}) &= \int_{\Theta} f(\mathbf{v}_{t+1} | \theta_{t+1}) \left[\frac{\alpha_{DPM}}{\alpha_{DPM} + t} \mathbb{G}_0 + \frac{1}{\alpha_{DPM} + t} \sum_{k=1}^t \delta_{\theta_k}(\theta_{t+1}) \right] d\theta_{t+1} \\ &= \frac{1}{\alpha_{DPM} + t} \left[\alpha_{DPM} \int_{\Theta} f(\mathbf{v}_{t+1} | \theta_{t+1}) d\mathbb{G}_0(\theta_{t+1}) + \sum_{k=1}^t f(\mathbf{v}_{t+1} | \theta_k) \right] \end{aligned} \quad (5.17)$$

En substituant (5.17) dans (5.15), nous obtenons l'estimation suivante de la densité

du bruit d'observation F :

$$\widehat{F}_t(\mathbf{v}_{t+1}) \simeq \sum_{i=1}^N \frac{\widetilde{\omega}_t^{(i)}}{\alpha_{DPM} + t} \left[\alpha_{DPM} \int_{\Theta} f(\mathbf{v}_{t+1} | \theta_{t+1}) d\mathbb{G}_0(\theta_{t+1}) + \sum_{k=1}^t f(\mathbf{v}_{t+1} | \widetilde{\theta}_k^{(i)}) \right] \quad (5.18)$$

Lorsque $t \gg \alpha_{DPM}$, nous avons

$$\widehat{F}_t \simeq \frac{1}{\alpha_{DPM} + t} \sum_{i=1}^N \widetilde{\omega}_t^{(i)} \sum_{k=1}^t f(\cdot | \widetilde{\theta}_k^{(i)}) \quad (5.19)$$

Le choix de la densité d'importance est un point crucial dans les algorithmes de filtrage particulaire. L'efficacité du filtre particulaire repose en grande partie sur la densité d'importance adoptée. En effet, le rôle de la densité d'importance est de proposer un nouvel état à chaque instant t pour mettre à jour les particules. Par conséquent, pour que le filtre soit efficace, il faut explorer les régions de fortes probabilités de la densité *a posteriori* par un grand nombre de particules. Le choix des densités d'importance est étudié dans les paragraphes suivants. Il faut noter que ce choix dépendra du modèle adopté pour le bruit d'observation.

5.2.1 Cas d'un bruit d'observation modélisé par un DPM de distributions gaussiennes

La densité d'importance $q(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ peut être décomposée de la manière suivante :

$$q(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) = q(\mathbf{x}_t | \theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) q(\theta_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) \quad (5.20)$$

Ainsi, les particules de la variable latente $\theta_t^{(i)}$ et du vecteur d'état $\mathbf{x}_t^{(i)}$ sont respectivement échantillonnées selon les densités d'importance $q(\theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ et $q(\mathbf{x}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$.

Le filtre particulaire proposé pour l'estimation en ligne de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de gaussiennes est donné par l'algorithme 15. Ses différentes étapes sont décrites dans les paragraphes suivants.

Algorithme 15 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de gaussiennes

• **Initialisation**

pour $i = 1$ à N **faire**

- Echantillonner $\mathbf{x}_0^{(i)} \sim p_0(\mathbf{x}_0)$
- Echantillonner $\theta_0^{(i)} \sim \mathbb{G}_0$
- Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$

fin pour

• **Itérations**

pour $t = 1$ à T **faire**

pour $i = 1$ à N **faire**

- Echantillonner $\tilde{\theta}_t^{(i)} \sim p(\theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ avec un pas de Metropolis-Hastings
- Calculer $\left\{ \hat{\mathbf{x}}_{t|t}(\tilde{\theta}_t^{(i)}), \Sigma_{t|t}(\tilde{\theta}_t^{(i)}) \right\}$ en faisant un pas d'EKF ou d'UKF à partir de $\left\{ \hat{\mathbf{x}}_{t-1|t-1}(\theta_{t-1}^{(i)}) = \mathbf{x}_{t-1}^{(i)}, \Sigma_{t-1|t-1}(\theta_{t-1}^{(i)}) = 0, \tilde{\theta}_t^{(i)}, \mathbf{y}_t \right\}$
- Echantillonner $\tilde{\mathbf{x}}_t^{(i)} \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t}(\tilde{\theta}_t^{(i)}), \Sigma_{t|t}(\tilde{\theta}_t^{(i)}))$
- Mettre à jour le poids $\tilde{\omega}_t^{(i)}$ en utilisant l'équation (5.27)

fin pour

pour $i = 1$ à N **faire**

- Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{i=1}^N \tilde{\omega}_t^{(i)}}$

fin pour

- Calculer $N_{eff} = \frac{1}{\sum_{j=1}^N (\tilde{\omega}_t^{(j)})^2}$

si $N_{eff} \leq N_{seuil}$ **alors**

- Rééchantillonner les particules en utilisant l'algorithme 7

sinon

pour $i = 1$ à N **faire**

- Mettre à jour les particules de l'état $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)}$
- Mettre à jour les particules de la variable latente $\theta_t^{(i)} = \tilde{\theta}_t^{(i)}$
- Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

fin pour

fin si

fin pour

5.2.1.1 Loi $q(\theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$

Une solution simple pour échantillonner les particules $\theta_t^{(i)}$ consiste à utiliser la loi d'évolution en urne de Polya. Cependant, cette loi est peu efficace vu qu'elle ne tient pas compte de l'observation à l'instant t , conduisant ainsi à une variance élevée des poids. Afin de garantir une efficacité maximale, nous proposons d'utiliser la densité d'importance optimale qui minimise la variance des poids. Cette densité est définie par :

$$q(\theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) = p(\theta_t | \theta_{1:t-1}^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) \quad (5.21)$$

En appliquant la règle de Bayes, cette densité peut s'exprimer comme suit :

$$\begin{aligned} p(\theta_t | \theta_{1:t-1}, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t}, \Phi) &= \frac{p(\mathbf{y}_t | \theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t-1}, \Phi) p(\theta_t | \theta_{1:t-1}, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}, \Phi)}{p(\mathbf{y}_t | \theta_{1:t-1}, \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}, \Phi)} \\ &= \frac{p(\mathbf{y}_t | \theta_t, \mathbf{x}_{t-1}, \Phi) p(\theta_t | \theta_{1:t-1}, \Phi)}{p(\mathbf{y}_t | \theta_{1:t-1}, \mathbf{x}_{t-1}, \Phi)} \\ &\propto p(\mathbf{y}_t | \theta_t, \mathbf{x}_{t-1}, \Phi) p(\theta_t | \theta_{1:t-1}, \Phi) \end{aligned} \quad (5.22)$$

Pour échantillonner selon cette densité, nous pouvons utiliser un pas de l'algorithme de Metropolis-Hastings en choisissant comme loi candidate la loi d'évolution en urne de Polya $p(\theta_t | \theta_{1:t-1}^{(i)}, \Phi)$. Le choix de cette loi candidate permet de simplifier le calcul de la probabilité d'acceptation, s'exprimant par conséquent comme suit :

$$\rho(\theta_t^{(j-1)}, \theta_t^{(j)*}) = \min \left\{ 1, \frac{p(\mathbf{y}_t | \theta_t^{(j)*}, \mathbf{x}_{t-1}, \Phi)}{p(\mathbf{y}_t | \theta_t^{(j-1)}, \mathbf{x}_{t-1}, \Phi)} \right\} \quad (5.23)$$

où $\theta_t^{(j)*}$ est la variable latente candidate échantillonnée à la $j^{\text{ième}}$ itération de l'algorithme de Metropolis-Hastings selon $p(\theta_t | \theta_{1:t-1}^{(j)}, \Phi)$. Le calcul de la probabilité d'acceptation nécessite l'évaluation de la fonction de vraisemblance $p(\mathbf{y}_t | \theta_t, \mathbf{x}_{t-1}, \Phi)$. Avec le modèle considéré, le bruit d'observation suit à l'instant t une loi normale de moyenne μ_t et de matrice de covariance Σ_t . Par conséquent, la fonction de vraisemblance peut être approchée par une gaussienne en faisant un pas d'EKF ou d'UKF :

$$p(\mathbf{y}_t | \theta_t, \mathbf{x}_{t-1}, \Phi) \simeq \mathcal{N}(\mathbf{y}_t; \hat{\mathbf{y}}_{t|t}(\theta_t), \mathbf{S}_t(\theta_t)) \quad (5.24)$$

où $\widehat{\mathbf{y}}_{t|t}(\theta_t)$ est la prédiction de l'observation définie dans notre cas par :

$$\widehat{\mathbf{y}}_{t|t-1}(\theta_t) = h_t(\widehat{\mathbf{x}}_{t|t-1}) + \mu_t \quad (5.25)$$

et $\mathbf{S}_t(\theta_t)$ est la covariance de l'innovation.

5.2.1.2 Loi $q(\mathbf{x}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$

Pour faire évoluer les particules de l'état $\mathbf{x}_t^{(i)}$, nous pouvons utiliser la loi d'évolution de l'état $p(\mathbf{x}_t | \mathbf{x}_{0:t-1})$. Cependant, cette densité d'importance est peu efficace vu qu'elle ne prend pas en compte l'observation \mathbf{y}_t . Pour garantir une meilleure qualité d'estimation, nous pouvons utiliser un pas de filtre de Kalman étendu ou sans parfum. Avec une telle densité d'importance, la particule $\mathbf{x}_t^{(i)}$ est générée selon une gaussienne

$$\mathbf{x}_t^{(i)} \sim \mathcal{N}(\widehat{\mathbf{x}}_{t|t}(\widetilde{\theta}_t^{(i)}), \Sigma_{t|t}(\widetilde{\theta}_t^{(i)})) \quad (5.26)$$

dont les paramètres sont obtenus à l'aide d'un pas d'EKF ou d'UKF en prenant comme statistiques au temps précédent $\widehat{\mathbf{x}}_{t-1|t-1}(\widetilde{\theta}_t^{(i)}) = \mathbf{x}_{t-1}^{(i)}$ et $\Sigma_{t-1|t-1}(\widetilde{\theta}_t^{(i)}) = 0$.

5.2.1.3 Evaluation des poids

En utilisant les densités d'importance présentées ci-dessus, les poids des particules sont calculés comme suit :

$$\begin{aligned} \widetilde{\omega}_t^{(i)} &\propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \widetilde{\mathbf{x}}_t^{(i)}, \widetilde{\theta}_t^{(i)}, \Phi) p(\widetilde{\mathbf{x}}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}) p(\widetilde{\theta}_t^{(i)} | \theta_{1:t-1}^{(i)}, \Phi)}{q(\widetilde{\mathbf{x}}_t^{(i)} | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) q(\widetilde{\theta}_t^{(i)} | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)} \\ &\propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \widetilde{\mathbf{x}}_t^{(i)}, \widetilde{\theta}_t^{(i)}, \Phi) p(\widetilde{\mathbf{x}}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}) p(\widetilde{\theta}_t^{(i)} | \theta_{1:t-1}^{(i)}, \Phi) p(\mathbf{y}_t | \theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)}{q(\widetilde{\mathbf{x}}_t^{(i)} | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) p(\mathbf{y}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi) p(\widetilde{\theta}_t^{(i)} | \theta_{1:t-1}^{(i)}, \Phi)} \\ &\propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t | \widetilde{\mathbf{x}}_t^{(i)}, \widetilde{\theta}_t^{(i)}, \Phi) p(\widetilde{\mathbf{x}}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}, \Phi) p(\mathbf{y}_t | \theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)}{\mathcal{N}(\mathbf{y}_t; \widehat{\mathbf{y}}_{t|t}(\widetilde{\theta}_t^{(i)}), S_t(\widetilde{\theta}_t^{(i)})) p(\mathbf{y}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)} \end{aligned} \quad (5.27)$$

La densité $p(\mathbf{y}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)$ peut être approchée par une densité gaussienne en utilisant un pas d'EKF ou d'UKF :

$$p(\mathbf{y}_t | \widetilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi) \simeq \mathcal{N}(\mathbf{y}_t; \widehat{\mathbf{y}}_{t|t-1}(\widetilde{\theta}_t^{(i)}), \mathbf{S}_t(\widetilde{\theta}_t^{(i)})) \quad (5.28)$$

où $\widehat{\mathbf{y}}_{t|t-1}(\tilde{\theta}_t^{(i)})$ et $\mathbf{S}_t(\tilde{\theta}_t^{(i)})$ sont respectivement la prédiction de l'observation et la covariance de l'innovation du filtre de Kalman.

En marginalisant selon θ_t , la densité $p(\mathbf{y}_t|\theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)$ peut s'écrire sous la forme de l'intégrale suivant :

$$p(\mathbf{y}_t|\theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi) = \int_{\Theta} p(\mathbf{y}_t|\theta_t, \mathbf{x}_{t-1}^{(i)}, \Phi)p(\theta_t|\theta_{1:t-1}^{(i)}, \Phi)d\theta_t \quad (5.29)$$

La valeur de cette intégrale peut être approchée par intégration de Monte Carlo en échantillonnant selon la loi d'évolution en Urne de Polya $p(\theta_t|\theta_{1:t-1}^{(i)}, \Phi)$:

$$\begin{aligned} p(\mathbf{y}_t|\theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi) &\simeq \sum_{k=1}^p p(\mathbf{y}_t|\theta_t^{(k)}, \mathbf{x}_{t-1}^{(i)}, \Phi) \\ &\simeq \sum_{k=1}^p \mathcal{N}(\mathbf{y}_t; \widehat{\mathbf{y}}_{t|t}(\theta_t^{(k)}), \mathbf{S}_t(\theta_t^{(k)})) \end{aligned} \quad (5.30)$$

où $\theta_t^{(k)} \sim p(\theta_t|\theta_{1:t-1}^{(i)}, \Phi)$, pour $k = 1, \dots, p$.

5.2.2 Cas d'un bruit d'observation modélisé par un DPM de distributions de Cauchy

La densité d'importance $q(\mathbf{x}_t, \theta_t|\mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ est décomposée de la même façon que précédemment. L'algorithme 16 donne les étapes nécessaires pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t}|\mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par DPM de Cauchy. Les lois d'importance utilisées dans cet algorithme pour échantillonner les particules $\mathbf{x}_t^{(i)}$ et $\theta_t^{(i)}$ sont détaillées dans les paragraphes ci-dessous.

5.2.2.1 Loi $q(\theta_t|\mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$

Comme dans le cas d'un bruit d'observation modélisé par un DPM de gaussiennes, les particules de la variable latente $\theta_t^{(i)}$ sont échantillonnées selon la densité d'importance optimale en utilisant un pas de Metropolis-Hastings. Le calcul de la probabilité d'acceptation nécessite l'évaluation de la fonction de vraisemblance $p(\mathbf{y}_t|\theta_t, \mathbf{x}_{t-1}, \Phi)$. Dans ce cas, cette fonction ne peut plus être approchée avec un pas de filtre de Kalman étendu ou sans parfum vu qu'à l'instant t le bruit d'observation

Algorithme 16 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de Cauchy

• **Initialisation**

pour $i = 1$ à N **faire**

- Echantillonner $\mathbf{x}_0^{(i)} \sim p_0(\mathbf{x}_0)$
- Echantillonner $\theta_0^{(i)} \sim \mathbb{G}_0$
- Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$

fin pour

• **Itérations**

pour $t = 1$ à T **faire**

pour $i = 1$ à N **faire**

- Echantillonner $\tilde{\theta}_t^{(i)} \sim p(\theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ avec un pas de Metropolis-Hastings
- Echantillonner $\tilde{\mathbf{x}}_t^{(i)} \sim p(\mathbf{x}_t | \tilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_t, \Phi)$ avec un pas de Metropolis-Hastings
- Mettre à jour le poids $\tilde{\omega}_t^{(i)}$ en utilisant l'équation (5.35)

fin pour

pour $i = 1$ à N **faire**

- Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^N \tilde{\omega}_t^{(j)}}$

fin pour

- Calculer $N_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}$

si $N_{eff} \leq N_{seuil}$ **alors**

- Rééchantillonner les particules en utilisant l'algorithme 7

sinon

pour $i = 1$ à N **faire**

- Mettre à jour les particules de l'état $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)}$
- Mettre à jour les particules de la variable latente $\theta_t^{(i)} = \tilde{\theta}_t^{(i)}$
- Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

fin pour

fin si

fin pour

n'est pas gaussien mais il suit une distribution de Cauchy. En effet, conditionnellement à θ_t et \mathbf{x}_{t-1} , l'équation d'observation (5.2) peut être reformulée par :

$$\mathbf{y}_t = h_t [g_t(\mathbf{x}_{t-1}) + \mathbf{w}_t] + \mathbf{v}_t \quad \text{avec} \quad \mathbf{v}_t \sim \mathcal{C}(a_t, m_t) \quad (5.31)$$

Le premier terme de cette équation peut être approché par une variable aléatoire gaussienne de densité $\mathcal{N}(\mathbf{y}_t; h_t(\hat{\mathbf{x}}_{t|t-1}), \mathbf{S}_t)$ dont les paramètres $\hat{\mathbf{x}}_{t|t-1}$ et \mathbf{S}_t sont respectivement la prédiction de l'état et la covariance de l'innovation obtenues en appliquant un pas d'EKF ou d'UKF sur le modèle dynamique (5.1)-(5.2) sans considération du bruit d'observation. Par conséquent, l'observation \mathbf{y}_t peut être vue comme la somme d'une variable aléatoire gaussienne et d'une variable aléatoire de Cauchy. Ainsi, la fonction de vraisemblance peut être déterminée en utilisant la densité de probabilité de la somme d'une variable aléatoire gaussienne et d'une variable aléatoire suivant une loi de Cauchy notée comme suit :

$$\begin{aligned} p(\mathbf{y}_t | \theta_t, \mathbf{x}_{t-1}, \mathbf{y}_{1:t-1}, \Phi) &\simeq \mathcal{NC}(\mathbf{y}_t; h_t(\hat{\mathbf{x}}_{t|t-1}), \mathbf{S}_t, a_t, m_t) \\ &\simeq \frac{1}{\sqrt{2\pi\mathbf{S}_t}} \exp\left(\frac{-1}{2\mathbf{S}_t} \left(\mathbf{y}_t (\mathbf{y}_t - 2h_t(\hat{\mathbf{x}}_{t|t-1}) - 2m_t) + a_t^2 - (h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t)^2\right)\right) \\ &\times \left[\cos\left(\frac{a_t (h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t)}{\mathbf{S}_t}\right) \cos\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right) \right. \\ &+ \sin\left(\frac{a_t (h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t)}{\mathbf{S}_t}\right) \sin\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right) - \frac{1}{2} \exp\left(\frac{1}{2\mathbf{S}_t} (a_t^2 - (h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t)^2)\right) \\ &\times \left[\exp\left(\frac{1}{2\mathbf{S}_t} (h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t + ia_t)^2\right) \operatorname{erf}\left(\frac{-a_t + i(\mathbf{y}_t - h_t(\hat{\mathbf{x}}_{t|t-1}) - m_t)}{\sqrt{2\mathbf{S}_t}}\right) \right. \\ &\times \left(-\cos\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right) + i \sin\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right)\right) \\ &- \exp\left(\frac{1}{2\mathbf{S}_t} (-h_t(\hat{\mathbf{x}}_{t|t-1}) - m_t + ia_t)^2\right) \operatorname{erf}\left(\frac{-a_t + i(-\mathbf{y}_t + h_t(\hat{\mathbf{x}}_{t|t-1}) + m_t)}{\sqrt{2\mathbf{S}_t}}\right) \\ &\left. \left. \times \left(\cos\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right) + i \sin\left(\frac{\mathbf{y}_t a_t}{\mathbf{S}_t}\right)\right) \right] \right] \end{aligned}$$

Les détails de calcul de cette densité de probabilité sont données en annexe A.

5.2.2.2 Loi $q(\mathbf{x}_t|\theta_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$

La densité d'importance proposée pour faire évoluer les particules de l'état $\mathbf{x}_t^{(i)}$, dans le cas d'un bruit d'observation modélisé par un DPM de gaussiennes, ne peut plus être utilisée vu que le bruit d'observation n'est plus gaussien mais suit une distribution de Cauchy. Par ailleurs, nous pouvons utiliser la densité d'importance optimale $p(\mathbf{x}_t|\theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi)$ qui peut s'exprimer comme suit en utilisant la règle de Bayes :

$$\begin{aligned} p(\mathbf{x}_t|\theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) &= \frac{p(\mathbf{y}_t|\mathbf{x}_t, \theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \Phi)p(\mathbf{x}_t|\theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \Phi)}{p(\mathbf{y}_t|\theta_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \Phi)} \\ &= \frac{p(\mathbf{y}_t|\mathbf{x}_t, \theta_t, \Phi)p(\mathbf{x}_t|\mathbf{x}_{t-1})}{p(\mathbf{y}_t|\theta_t, \mathbf{x}_{t-1}, \Phi)} \\ &\propto p(\mathbf{y}_t|\mathbf{x}_t, \theta_t, \Phi)p(\mathbf{x}_t|\mathbf{x}_{t-1}) \end{aligned} \quad (5.32)$$

Il est possible de générer des échantillons selon cette densité avec un pas de Metropolis-Hastings en utilisant la loi d'évolution de l'état $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ comme loi candidate. Avec une telle loi candidate, la probabilité d'acceptation est donnée par :

$$\rho(\mathbf{x}_t^{(j-1)}, \mathbf{x}_t^{(j)*}) = \min \left\{ 1, \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(j)*}, \tilde{\theta}_t^{(i)}, \Phi)}{p(\mathbf{y}_t|\mathbf{x}_t^{(j-1)}, \tilde{\theta}_t^{(i)}, \Phi)} \right\} \quad (5.33)$$

La fonction de vraisemblance $p(\mathbf{y}_t|\mathbf{x}_t, \theta_t, \Phi)$ suit dans ce cas une distribution de Cauchy :

$$p(\mathbf{y}_t|\mathbf{x}_t, \theta_t, \Phi) = \mathcal{C}(\mathbf{y}_t; a_t, m_t + g_t(\mathbf{x}_t)) \quad (5.34)$$

5.2.2.3 Evaluation des poids

Avec les densités d'importance présentées ci-dessus, nous obtenons l'expression suivante pour mettre à jour les poids :

$$\begin{aligned} \tilde{\omega}_t^{(i)} &\propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}, \tilde{\theta}_t^{(i)}, \Phi)p(\tilde{\mathbf{x}}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})p(\tilde{\theta}_t^{(i)}|\theta_{1:t-1}^{(i)}, \Phi)}{q(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\theta}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)q(\tilde{\theta}_t^{(i)}|\mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)} \\ &\propto \omega_{t-1}^{(i)} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}, \tilde{\theta}_t^{(i)}, \Phi)p(\tilde{\mathbf{x}}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})p(\tilde{\theta}_t^{(i)}|\theta_{1:t-1}^{(i)}, \Phi)p(\mathbf{y}_t|\theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)p(\mathbf{y}_t|\tilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)}{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}, \tilde{\theta}_t^{(i)}, \Phi)p(\tilde{\mathbf{x}}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})p(\mathbf{y}_t|\tilde{\theta}_t^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)p(\tilde{\theta}_t^{(i)}|\theta_{1:t-1}^{(i)}, \Phi)} \\ &\propto \omega_{t-1}^{(i)} p(\mathbf{y}_t|\theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi) \end{aligned} \quad (5.35)$$

La densité $p(\mathbf{y}_t | \theta_{1:t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}, \Phi)$ peut être approchée par intégration de Monte Carlo en échantillonnant selon la loi d'évolution en Urne de Polya $p(\theta_t | \theta_{1:t-1}^{(i)}, \Phi)$.

5.2.3 Cas d'un bruit d'observation modélisé par un DPM de distributions α -stables

La densité d'importance $q(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ peut également être décomposée comme suit :

$$q(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) = q(\theta_t | \mathbf{x}_t, \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) q(\mathbf{x}_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) \quad (5.36)$$

Ainsi, les particules du vecteur d'état $\mathbf{x}_t^{(i)}$ et de la variable latente $\theta_t^{(i)}$ sont respectivement échantillonnées selon les densités d'importance $q(\mathbf{x}_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ et $q(\theta_t | \tilde{\mathbf{x}}_t^{(i)}, \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$. Nous pouvons utiliser la loi d'évolution de l'état $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ comme densité d'importance pour les particules de l'état :

$$\tilde{\mathbf{x}}_t^{(i)} \sim \mathcal{N}(g_t(\mathbf{x}_{t-1}^{(i)}) + \mu_t^w, \Sigma_t^w) \quad (5.37)$$

Les particules de la variable latente sont échantillonnées en utilisant la loi d'importance optimale :

$$\begin{aligned} p(\theta_t | \mathbf{x}_t, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) &= \frac{p(\mathbf{y}_t | \theta_{1:t}, \mathbf{x}_t, \mathbf{y}_{1:t-1}, \Phi) p(\theta_t | \theta_{1:t-1}, \mathbf{x}_{0:t}, \mathbf{y}_{1:t-1}, \Phi)}{p(\mathbf{y}_t | \theta_{1:t-1}, \mathbf{x}_{0:t}, \mathbf{y}_{1:t-1}, \Phi)} \\ &= \frac{p(\mathbf{y}_t | \theta_t, \mathbf{x}_t, \Phi) p(\theta_t | \theta_{1:t-1}, \Phi)}{p(\mathbf{y}_t | \theta_{1:t-1}, \mathbf{x}_t, \Phi)} \end{aligned} \quad (5.38)$$

Pour ce faire, nous utilisons un pas de Metropolis-Hastings en choisissant la loi d'évolution en urne de Polya comme loi candidate. Avec une telle loi candidate, la probabilité d'acceptation est définie comme suit :

$$\rho(\theta_t^{(j-1)}, \theta_t^{(j)*}) = \min \left\{ 1, \frac{p(\mathbf{y}_t | \theta_t^{(j)*}, \tilde{\mathbf{x}}_t^{(i)}, \Phi)}{p(\mathbf{y}_t | \theta_t^{(j-1)}, \tilde{\mathbf{x}}_t^{(i)}, \Phi)} \right\} \quad (5.39)$$

où $\theta_t^{(j)*}$ est la variable latente candidate échantillonnée à la $j^{\text{ième}}$ itération de l'algorithme de Metropolis-Hastings selon la loi $p(\theta_t | \theta_{1:t-1}^{(i)}, \Phi)$ et la fonction de vraisem-

blance $p(\mathbf{y}_t|\theta_t, \mathbf{x}_t, \Phi)$ est une loi α -stable

$$p(\mathbf{y}_t|\theta_t, \mathbf{x}_t, \Phi) = \mathcal{S}_1(\alpha_t, \beta_t, \gamma_t, \mu_t + h_t(\mathbf{x}_t)) \quad (5.40)$$

calculée numériquement en utilisant l'algorithme proposé par Nolan [Nol97, Nol99] décrit dans le paragraphe 1.1.6.1.

Avec de telles densités d'importance, nous obtenons l'expression suivante pour la mise à jour des poids :

$$\tilde{\omega}_t^{(i)} \propto \omega_{t-1}^{(i)} p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}, \theta_{1:t-1}^{(i)}, \Phi) \quad (5.41)$$

La densité $p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)}, \theta_{1:t-1}^{(i)}, \Phi)$ peut être approchée par intégration de Monte Carlo en échantillonnant selon la loi d'évolution en Urne de Polya $p(\theta_t|\theta_{1:t-1}^{(i)}, \Phi)$.

L'algorithme 17 résume les étapes nécessaires pour estimer séquentiellement la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t}|\mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de distributions α -stables.

Remarque 5.1. *La décomposition de la densité d'importance utilisée avec les deux autres modèles est aussi valable dans le cas d'un bruit d'observation modélisé par un DPM de distributions α -stables. Cependant, il n'est plus possible d'utiliser la densité d'importance optimale pour échantillonner les particules de la variable latente $\tilde{\theta}_t^{(i)}$. En effet, lorsque le bruit d'observation est α -stable, il n'existe pas à notre connaissance de technique efficace pour approcher la vraisemblance $p(\mathbf{y}_t|\theta_t, \mathbf{x}_{t-1}, \Phi)$. Dans ce cas, nous pouvons utiliser la loi d'évolution en Urne de Polya $p(\theta_t|\theta_{1:t-1}, \Phi)$. Or, cette densité d'importance est peu efficace car elle ne tient pas compte de l'observation à l'instant t . Pour cette raison nous avons eu recours à cette nouvelle décomposition de la densité d'importance.*

5.2.4 Une troisième densité d'importance

Dans ce qui précède, la densité d'importance est décomposée en deux densités d'importance. Ceci nous amène à utiliser deux échantillonneurs : un premier pour les particules du vecteur d'état et un autre pour les particules de la variable latente. Maintenant, nous proposons de mettre en place un seul échantillonneur pour les deux variables. Pour ce faire, nous utilisons la densité d'importance optimale conjointe

Algorithme 17 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$ dans le cas d'un bruit d'observation modélisé par un DPM de distributions α -stables

• **Initialisation**

pour $i = 1$ à N **faire**

- Echantillonner $\mathbf{x}_0^{(i)} \sim p_0(\mathbf{x}_0)$
- Echantillonner $\theta_0^{(i)} \sim \mathbb{G}_0$
- Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$

fin pour

• **Itérations**

pour $t = 1$ à T **faire**

pour $i = 1$ à N **faire**

- Echantillonner $\tilde{\mathbf{x}}_t^{(i)} \sim \mathcal{N}(\mathbf{x}_t; g_t(\mathbf{x}_{t-1}^{(i)}) + \mu_t^w, \Sigma_t^w)$
- Echantillonner $\tilde{\theta}_t^{(i)} \sim p(\theta_t | \tilde{\mathbf{x}}_t^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi)$ avec un pas de Metropolis-Hastings
- Mettre à jour le poids $\tilde{\omega}_t^{(i)}$ en utilisant l'équation (5.41)

fin pour

pour $i = 1$ à N **faire**

- Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^N \tilde{\omega}_t^{(j)}}$

fin pour

- Calculer $N_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}$

si $N_{eff} \leq N_{seuil}$ **alors**

- Rééchantillonner les particules en utilisant l'algorithme 7

sinon

pour $i = 1$ à N **faire**

- Mettre à jour les particules de l'état $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)}$
- Mettre à jour les particules de la variable latente $\theta_t^{(i)} = \tilde{\theta}_t^{(i)}$
- Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

fin pour

fin si

fin pour

définie par :

$$q(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) = p(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) \quad (5.42)$$

En utilisant la règle de Bayes, cette densité peut s'écrire :

$$\begin{aligned} p(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t}, \Phi) &= \frac{p(\mathbf{y}_t | \mathbf{x}_{0:t}, \theta_{1:t}, \mathbf{y}_{1:t-1}, \Phi) p(\mathbf{x}_t | \theta_t, \mathbf{x}_{0:t-1}) p(\theta_t | \theta_{1:t-1}, \Phi)}{p(\mathbf{y}_t | \mathbf{x}_{0:t-1}, \theta_{1:t-1}, \mathbf{y}_{1:t-1}, \Phi)} \\ &\propto p(\mathbf{y}_t | \mathbf{x}_t, \theta_t, \Phi) p(\mathbf{x}_t | \theta_t, \mathbf{x}_{0:t-1}) p(\theta_t | \theta_{1:t-1}, \Phi) \end{aligned} \quad (5.43)$$

Afin de trouver une approximation de cette densité, nous proposons d'utiliser la méthode de Monte Carlo et l'échantillonnage d'importance. Pour ce faire, nous considérons, pour chaque couple de particule $(\tilde{\mathbf{x}}_t^{(i)}, \tilde{\theta}_t^{(i)})$, un système de N_{IS} particules auxiliaires $\left\{ \left(\check{x}_{t,i}^{(j)}, \check{\theta}_{t,i}^{(j)} \right) \right\}_{j=1}^{N_{IS}}$. Les particules de la variable latente $\check{\theta}_{t,i}^{(j)}$ sont échantillonnées selon la loi d'évolution en urne de Polya :

$$\check{\theta}_{t,i}^{(j)} \sim p(\theta_t | \theta_{1:t-1}^{(i)}, \Phi) \quad (5.44)$$

Dans le cas où le bruit d'observation est modélisé par un DPM de distributions gaussiennes, les particules du vecteur d'état $\check{x}_{t,i}^{(j)}$ peuvent être échantillonnées selon une densité gaussienne

$$\check{x}_{t,i}^{(j)} \sim \mathcal{N}(\widehat{\mathbf{x}}_{t|t}(\check{\theta}_{t,i}^{(j)}), \Sigma_{t|t}(\check{\theta}_{t,i}^{(j)})) \quad (5.45)$$

dont la moyenne et la matrice de covariance sont obtenues en faisant un pas d'EKF ou d'UKF en prenant comme statistiques à l'instant précédent $\widehat{\mathbf{x}}_{t-1|t-1} = \mathbf{x}_{t-1}^{(i)}$ et $\Sigma_{t-1|t-1} = 0$. Sinon, dans le cas d'un bruit d'observation modélisé par un DPM de distributions de Cauchy ou de distributions α -stables, nous pouvons utiliser la loi d'évolution de l'état $p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)})$:

$$\check{x}_{t,i}^{(j)} \sim \mathcal{N}(g_t(\mathbf{x}_{t-1}^{(i)}) + \mu_t^w, \Sigma_t^w) \quad (5.46)$$

La densité d'importance optimale est alors approchée à l'aide de ce système de particules pondérées par la densité empirique

$$p(\mathbf{x}_t, \theta_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t}, \Phi) \simeq \sum_{j=1}^{N_{IS}} \check{\omega}_{t,i}^{(j)} \delta_{\check{\mathbf{x}}_{t,i}^{(j)}, \check{\theta}_{t,i}^{(j)}}(\mathbf{x}_t, \theta_t) \quad (5.47)$$

où $\check{\omega}_{t,i}^{(j)}$ est le poids non normalisé associé à la $j^{\text{ième}}$ particule auxiliaire défini par :

$$\check{\omega}_{t,i}^{(j)} \propto \frac{p(\mathbf{y}_t | \check{\mathbf{x}}_{t,i}^{(j)}, \check{\theta}_{t,i}^{(j)}, \mathbf{y}_{1:t-1}, \Phi) p(\check{\mathbf{x}}_{t,i}^{(j)} | \mathbf{x}_{t-1}^{(i)}) p(\check{\theta}_{t,i}^{(j)} | \theta_{1:t-1}^{(i)}, \Phi)}{q(\check{\mathbf{x}}_{t,i}^{(j)} | \mathbf{y}_t, \mathbf{x}_{t-1}^{(i)}, \check{\theta}_{t,i}^{(j)}) q(\check{\theta}_{t,i}^{(j)} | \theta_{1:t-1}^{(i)}, \Phi)} \quad (5.48)$$

Pour générer le $i^{\text{ième}}$ couple de particules $(\mathbf{x}_{t,i}^{(i)}, \theta_{t,i}^{(i)})$ selon cette approximation de la densité d'importance optimale conjointe, il suffit de tirer une particule à partir du système $\left\{ (\check{x}_{t,i}^{(j)}, \check{\theta}_{t,i}^{(j)}) \right\}_{j=1}^{N_{IS}}$ en utilisant les poids $\check{\omega}_{t,i}^{(1)}, \check{\omega}_{t,i}^{(2)}, \dots, \check{\omega}_{t,i}^{(N_{IS})}$ comme probabilités de sélection. Une façon de faire est de générer une variable aléatoire J à valeurs dans $\{1, 2, \dots, N_{IS}\}$ en utilisant la loi multinomiale

$$J \sim \text{Multinomial} \left(\check{\omega}_{t,i}^{(1)}, \check{\omega}_{t,i}^{(2)}, \dots, \check{\omega}_{t,i}^{(N_{IS})} \right) \quad (5.49)$$

Le $i^{\text{ième}}$ couple de particules est alors donné par :

$$\left(\tilde{\mathbf{x}}_t^{(i)}, \tilde{\theta}_t^{(i)} \right) = \left(\check{\mathbf{x}}_{t,i}^{(j=J)}, \check{\theta}_{t,i}^{(j=J)} \right) \quad (5.50)$$

Avec une telle densité d'importance, nous obtenons l'expression suivante pour la mise à jour des poids :

$$\tilde{\omega}_t^{(i)} \propto \omega_{t-1}^{(i)} p(\mathbf{y}_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t-1}, \Phi) \quad (5.51)$$

En marginalisant selon \mathbf{x}_t et θ_t , la densité $p(\mathbf{y}_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t-1}, \Phi)$ peut s'écrire :

$$p(\mathbf{y}_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t-1}, \Phi) = \int_{\theta} \int_{\chi} p(\mathbf{y}_t | \mathbf{x}_t, \theta_t, \Phi) p(\mathbf{x}_t | \theta_t, \mathbf{x}_{t-1}^{(i)}, \Phi) p(\theta_t | \theta_{1:t-1}^{(i)}, \Phi) d\mathbf{x}_t d\theta_t \quad (5.52)$$

Cette densité peut être approchée par intégration de Monte Carlo en utilisant le système de particules auxiliaires $\left\{ (\check{x}_{t,i}^{(j)}, \check{\theta}_{t,i}^{(j)}) \right\}_{j=1}^{N_{IS}}$ précédemment échantillonnées. En procédant ainsi, cette densité est alors donnée par la somme des poids non normalisés $\left\{ \check{\omega}_{t,i}^{(j)} \right\}_{j=1}^{N_{IS}}$:

$$p(\mathbf{y}_t | \mathbf{x}_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}, \mathbf{y}_{1:t-1}, \Phi) \simeq \sum_{i=1}^{N_{IS}} \check{\omega}_{t,i}^{(j)} \quad (5.53)$$

L'algorithme 18 résume la procédure complète pour échantillonner selon la den-

sité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$. Cet algorithme peut être utilisé avec les trois modèles proposés pour le bruit d'observation.

5.3 Exemple de mise en œuvre

Dans cette section, nous proposons d'évaluer les performances des algorithmes proposés. Pour ce faire, nous utilisons le système fortement non linéaire suivant, largement utilisé dans la littérature pour analyser les performances des techniques de filtrage particulaire [DDFG01, Kit96, GSS93, Wes93] :

$$\begin{cases} x_t = 0.5x_{t-1} + 25 \frac{x_{t-1}}{1 + x_{t-1}^2} + 8 \cos(1.2t) + w_t \\ y_t = \frac{x_t^2}{20} + v_t \end{cases}$$

où w_t est un bruit gaussien de moyenne nulle et de variance $\sigma_w^2 = 1$.

Dans ce qui suit, nous étudions les performances des filtres particuliers proposés avec les différents modèles de bruit présentés dans la section 5.1. Pour chacun de ces modèles, plusieurs simulations avec des réalisations de bruits d'observation différentes sont effectuées. Nous terminons cette section par une étude comparative des trois modélisations proposées.

5.3.1 Bruit d'observation modélisé avec un DPM de distributions gaussiennes

Dans ce paragraphe, nous supposons que le bruit d'observation v_t est modélisé par un DPM de distributions gaussiennes. Nous commençons par tester notre algorithme en considérant que la densité du bruit v_t est une distribution α -stable : $F \sim \mathcal{S}_1(1.8, 0, 2.2, 0)$. Les hyperparamètres de la distribution de base sont fixés à $\mu_0 = 0$, $\kappa_0 = 0.1$, $\nu_0 = 10$ et $\Lambda_0 = 40$. Le facteur d'échelle α_{DPM} est fixé à 3. L'algorithme 15 est utilisé avec 200 particules. Les particules $\theta_t^{(i)}$ sont échantillonnées selon la densité d'importance optimale définie dans le paragraphe 5.2.1.1 en utilisant un pas de Metropolis-Hastings avec 300 itérations de chauffage. Les particules $x_t^{(i)}$ sont propagées à l'aide d'un pas d'EKF. Les résultats obtenus sont illustrés dans les figures 5.1 et 5.2. La figure 5.1 (haut) représente l'état estimé, l'état réel ainsi que

Algorithme 18 Filtre particulaire pour l'estimation séquentielle de la densité *a posteriori* $p(\mathbf{x}_{0:t}, \theta_{1:t} | \mathbf{y}_{1:t}, \Phi)$

• **Initialisation**

pour $i = 1$ à N **faire**

- Echantillonner $\mathbf{x}_0^{(i)} \sim p_0(\mathbf{x}_0)$
- Echantillonner $\theta_0^{(i)} \sim \mathbb{G}_0$
- Initialiser les poids $\omega_0^{(i)} = \frac{1}{N}$

fin pour

• **Itérations**

pour $t = 1$ à T **faire**

pour $i = 1$ à N **faire**

pour $j = 1$ à N_{IS} **faire**

- Echantillonner $\check{\mathbf{x}}_{t,i}^{(j)}$:
 - En faisant un pas d'EKF ou d'UKF si le bruit d'observation est modélisé par un DPM de gaussiennes (cf. équation (5.45)).
 - Sinon, directement selon la loi d'évolution de l'état (cf. équation (5.46)).
- Echantillonner $\check{\theta}_{t,i}^{(j)} \sim p(\theta_t | \theta_{1:t-1}, \Phi)$
- Calculer le poids $\check{\omega}_{t,i}^{(j)}$ en utilisant l'équation (5.48)

fin pour

– Calculer la somme des poids non normalisés $S_{\check{\omega}} = \sum_{j=1}^{N_{IS}} \check{\omega}_{t,i}^{(j)}$

– Normaliser les poids $\left\{ \check{\omega}_{t,i}^{(j)} \right\}_{j=1}^{N_{IS}}$

– Echantillonner $J \sim \text{Multinomial} \left(\check{\omega}_{t,i}^{(1)}, \check{\omega}_{t,i}^{(2)}, \dots, \check{\omega}_{t,i}^{(N_{IS})} \right)$

– Faire $\tilde{\mathbf{x}}_t^{(i)} = \check{\mathbf{x}}_{t,i}^{(J)}$ et $\tilde{\theta}_t^{(i)} = \check{\theta}_{t,i}^{(J)}$

– Mettre à jour les poids $\tilde{\omega}_t^{(i)} \propto \omega_{t-1}^{(i)} S_{\check{\omega}}$

fin pour

pour $i = 1$ à N **faire**

– Normaliser les poids $\tilde{\omega}_t^{(i)} = \frac{\tilde{\omega}_t^{(i)}}{\sum_{j=1}^N \tilde{\omega}_t^{(j)}}$

fin pour

– Calculer $N_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2}$

si $N_{eff} \leq N_{seuil}$ **alors**

Rééchantillonner les particules en utilisant l'algorithme 7

sinon

pour $i = 1$ à N **faire**

- Mettre à jour les particules de l'état $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)}$
- Mettre à jour les particules de la variable latente $\theta_t^{(i)} = \tilde{\theta}_t^{(i)}$
- Mettre à jour les poids $\omega_t^{(i)} = \tilde{\omega}_t^{(i)}$

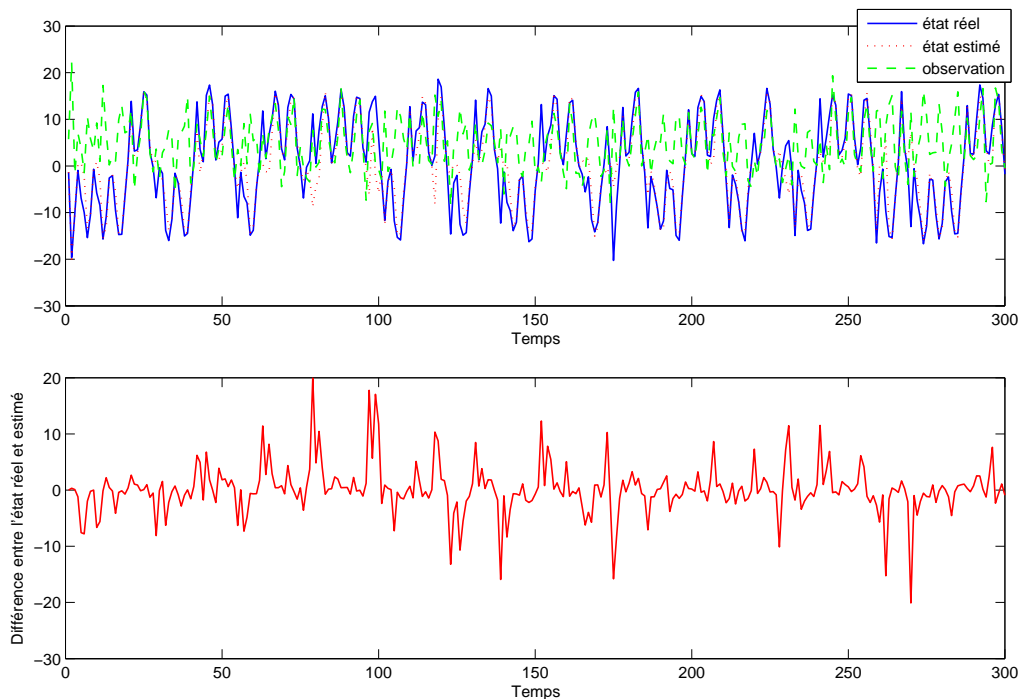
fin pour

fin si

fin pour

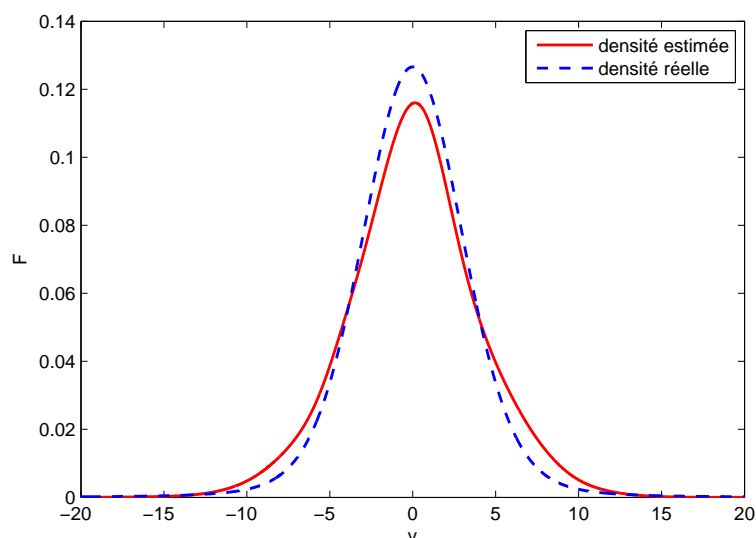
l'observation. L'évolution en fonction du temps de la différence entre l'état réel et l'état estimé est aussi illustrée sur la figure 5.1 (bas). La figure 5.2 montre la densité réelle et la densité estimée du bruit d'observation. D'après ces figures, nous pouvons constater que l'état est correctement estimé. La densité de bruit estimée est très proche de la densité réelle montrant que le modèle utilisé est efficace. Cependant, il faut noter que la densité α -stable considérée s'approche d'une densité gaussienne ($\alpha = 1.8$ proche de 2 et $\beta = 0$).

FIGURE 5.1 Estimation de l'état dans le cas où la densité du bruit d'observation est une distribution α -stable $F \sim \mathcal{S}_1(1.8, 0, 2.2, 0)$ modélisée par un DPM de distributions gaussiennes. (Haut) Etat réel, état estimé et l'observation. (Bas) Evolution temporelle de l'erreur entre l'état réel et l'état estimé.



Maintenant, nous nous proposons de tester la performance de l'algorithme 18 en considérant que le bruit d'observation est distribué selon un mélange de deux distributions α -stables symétriques $F \sim 0.6\mathcal{S}_1(1.8, 0, 2, -5) + 0.4\mathcal{S}_1(1.9, 0, 1.5, 5)$. Les hyperparamètres de la distribution de base sont fixés à $\mu_0 = 0$, $\kappa_0 = 0.1$, $\nu_0 = 10$ et $\Lambda_0 = 10$. Le facteur d'échelle α_{DPM} est fixé à 3. Cet algorithme est appliqué pour $T = 300$ avec $N = 200$ particules et $N_{IS} = 100$ particules auxiliaires. Les figures 5.3

FIGURE 5.2 Densité réelle et densité estimée du bruit d'observation. La densité réelle du bruit d'observation est une distribution α -stable $\mathcal{S}_1(1.8, 0, 2.2, 0)$ modélisée par un DPM de distributions gaussiennes.

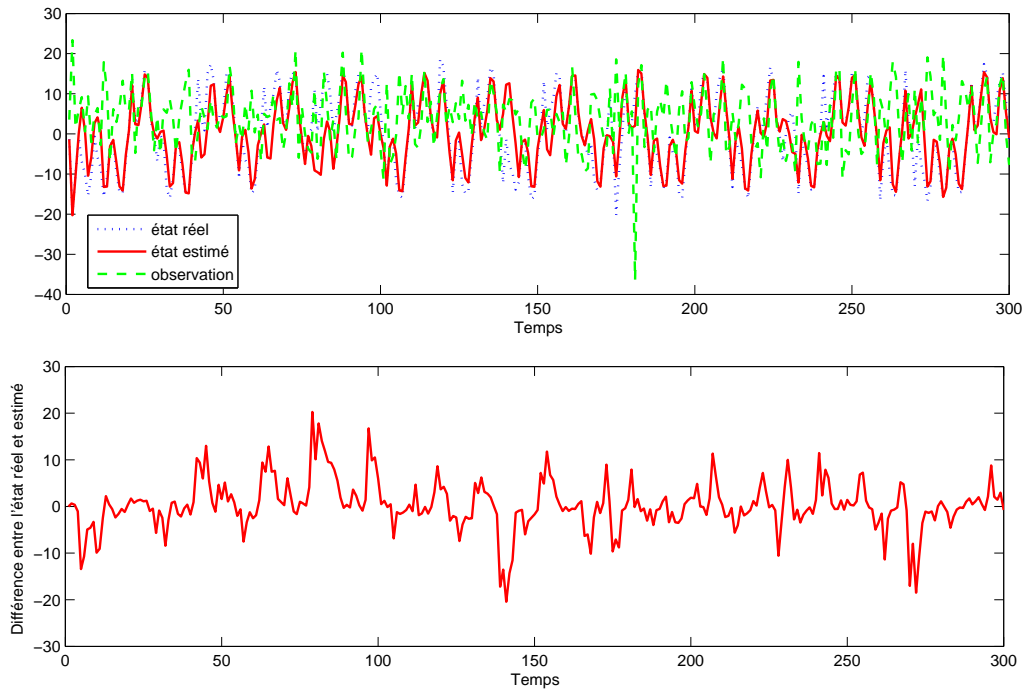


et 5.4 montrent respectivement les résultats d'estimation obtenus pour l'état et pour la densité du bruit d'observation. A partir de la figure 5.3, nous pouvons constater que l'estimation de l'état fournie par le filtre particulaire 18 est satisfaisante. De plus, comme illustré dans la figure 5.4, la densité estimée est très proche de la densité réelle et les deux modes de cette dernière ont été déterminés par notre algorithme.

5.3.2 Bruit d'observation modélisé avec un DPM de distributions de Cauchy

Dans ce paragraphe, nous supposons que le bruit d'observation est modélisé par un DPM de distributions de Cauchy. Nous commençons par tester la performance de la méthode proposée en supposant que la densité du bruit d'observation suit un mélange de deux distributions de Cauchy : $F \sim 0.5\mathcal{C}(1, 0) + 0.5\mathcal{C}(0.5, -10)$. Les hyperparamètres de la distribution de base sont fixés à $a^a = 4$, $b^a = 1$, $m^m = 0$ et $\Sigma^m = 20$. Le facteur d'échelle α_{DPM} est fixé à 3. L'algorithme 18 est utilisé avec $N = 200$ et $N_{IS} = 100$. Les états réel et estimé ainsi que les observations sont tracés dans la figure 5.5 (haut). L'évolution en fonction du temps de la différence entre l'état réel et l'état estimé est aussi illustrée sur la figure 5.5 (bas). La figure

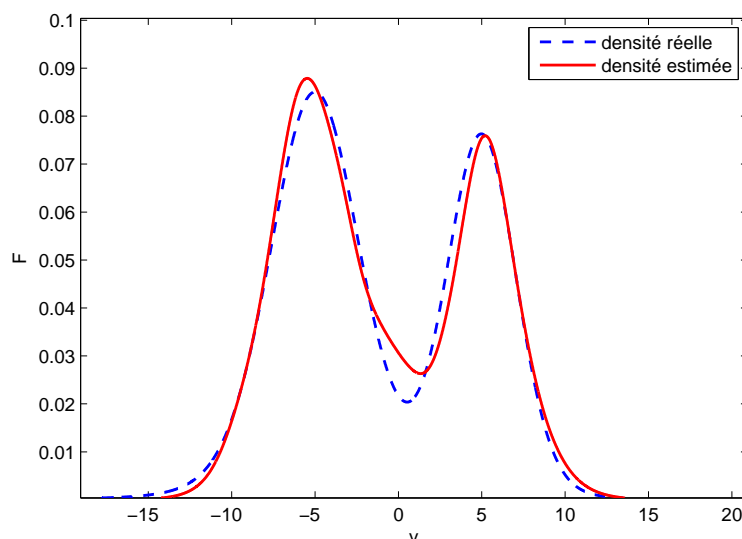
FIGURE 5.3 Estimation de l'état dans le cas où la densité du bruit d'observation est un mélange de deux distributions α -stables $F \sim 0.6\mathcal{S}_1(1.8, 0, 2, -5) + 0.4\mathcal{S}_1(1.9, 0, 1.5, 5)$ modélisée par un DPM de distributions gaussiennes. (Haut) Etat réel, état estimé et l'observation. (Bas) Evolution temporelle de l'erreur entre l'état réel et l'état estimé.



5.6 montre la densité réelle et la densité estimée du bruit d'observation. D'après ces figures, nous pouvons constater que l'état est correctement estimé. La densité de bruit estimée est très proche de la densité réelle montrant que le modèle utilisé est efficace.

L'algorithme 18 est maintenant testé en supposant que la densité du bruit d'observation est un mélange de deux distributions α -stables : $F \sim 0.3\mathcal{S}_1(1.1, 0, 2, 0) + 0.7\mathcal{S}_1(1.2, 0.5, 1.5, -7)$. Les hyperparamètres de la distribution de base \mathbb{G}_0 sont fixés comme suit : $a^a = 5$, $b^a = 4$, $m^m = 0$ et $\Sigma^m = 25$. Le facteur d'échelle α_{DPM} est fixé à 3. Notre filtre particulière est utilisé avec les mêmes paramètres et densités d'importance considérés dans l'exemple de test précédent. Les résultats obtenus sont illustrés dans les figures 5.7 et 5.8. A partir ces figures, nous pouvons remarquer que le résidu entre l'état estimé et l'état réel reste faible en prenant en considération la puissance du bruit ajouté. En ce qui concerne l'estimation de la densité du bruit

FIGURE 5.4 Densité réelle et densité estimée du bruit d'observation. La densité réelle est un mélange de deux distributions α -stables $F \sim 0.6\mathcal{S}_1(1.8, 0, 2, -5) + 0.4\mathcal{S}_1(1.9, 0, 1.5, 5)$ modélisée par un DPM de distributions gaussiennes.

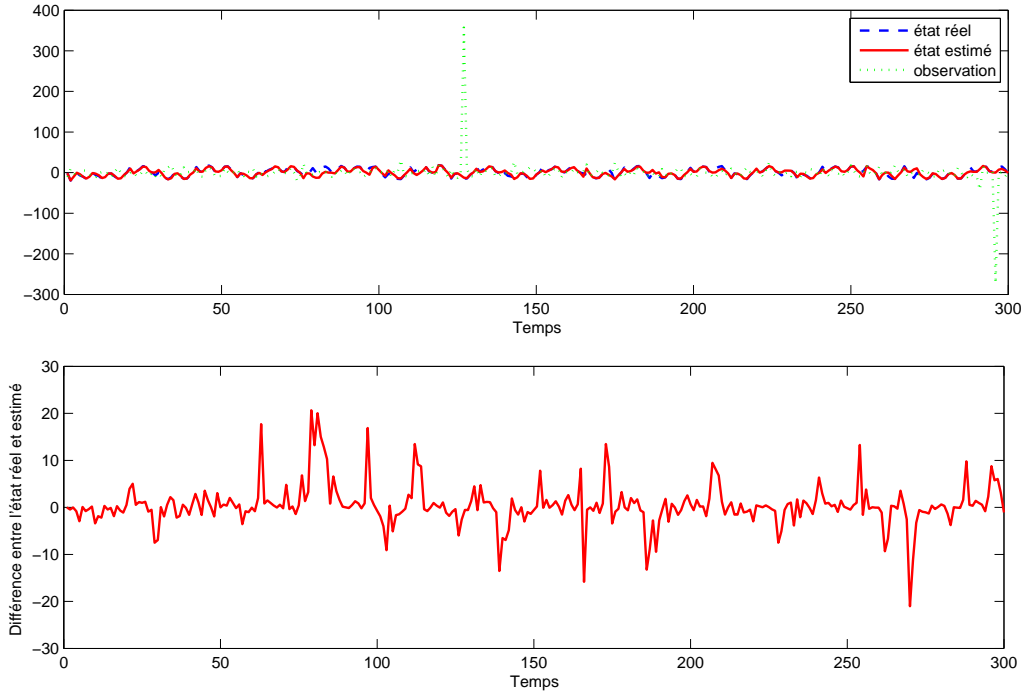


d'observation, notre algorithme arrive à trouver les deux modes de la densité réelle du bruit et les queues de distribution sont bien estimées.

5.3.3 Bruit d'observation modélisé avec un DPM de distributions α -stable

Dans ce paragraphe, la densité du bruit d'observation est modélisée par un DPM de distributions α -stables. Le test des performances de l'algorithme 18 est effectué en considérant que le bruit d'observation est distribué selon un mélange de deux distributions α -stables $F \sim 0.7\mathcal{S}_1(1.4, 0, 2.2, 0) + 0.3\mathcal{S}_1(1.4, 0.5, 1.5, -10)$. Les hyperparamètres de la distribution de base sont fixés à $a^\gamma = 4$, $b^\gamma = 3$, $m^\mu = 0$ et $\Sigma^\mu = 50$. Le facteur d'échelle α_{DPM} est fixé à 3. L'algorithme 18 est appliqué pour $T = 300$ avec $N = 100$ particules et $N_{IS} = 100$ particules auxiliaires. Les figures 5.9 et 5.10 montrent respectivement les résultats d'estimation obtenus pour l'état et la densité du bruit. A partir de la figure 5.9, nous pouvons constater que l'estimation de l'état fournie par le filtre particulaire proposé reste satisfaisante même en présence d'un bruit d'observation fortement impulsif. De plus, comme illustré dans la figure 5.10, la

FIGURE 5.5 Estimation de l'état dans le cas où la densité du bruit d'observation est un mélange de deux distributions de Cauchy $F \sim 0.5\mathcal{C}(1, 0) + 0.5\mathcal{C}(0.5, -10)$ modélisée par un DPM de distributions de Cauchy. (Haut) Etat réel, état estimé et l'observation. (Bas) Evolution temporelle de l'erreur entre l'état réel et l'état estimé.

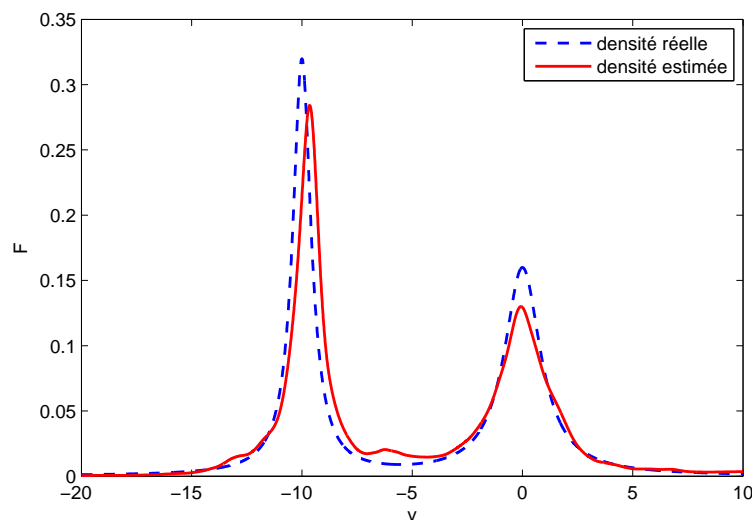


densité estimée est très proche de la densité réelle et les deux modes de cette dernière ont été déterminés par notre algorithme. Nous pouvons ainsi confirmer l'efficacité de l'algorithme proposé.

5.3.4 Comparaison

Dans ce paragraphe, nous nous intéressons à la comparaison des trois modélisations proposées pour modéliser un mélange de distributions α -stables. Pour analyser au mieux la pertinence de ces différents modèles, nous nous plaçons dans le contexte de l'estimation en ligne d'une densité de probabilité distribuée selon un mélange de distributions α -stables à partir des données observées séquentiellement. A cet effet, l'algorithme 18 est appliqué en supposant que les observations $\{\mathbf{y}_t\}_{t=1}^T$ suivent un

FIGURE 5.6 Densité réelle et densité estimée du bruit d'observation. La densité réelle est un mélange de deux distributions de Cauchy $F \sim 0.5\mathcal{C}(1, 0) + 0.5\mathcal{C}(0.5, -10)$ modélisée par un DPM de distributions de Cauchy.



mélange de distributions α -stables :

$$\mathbf{y}_t \sim \sum_{k=1}^K \pi_k \mathcal{S}_1(\alpha_k, \beta_k, \gamma_k, \mu_k), \quad t = 1, \dots, T \quad (5.54)$$

Les trois modélisations sont comparées en termes de distance entre la densité réelle F et la densité estimée \hat{F}_t . Nous utilisons la distance de Kullback-Leibler D_{KL} et le carré de la distance de Hellinger D_H^2 , respectivement définies comme suit :

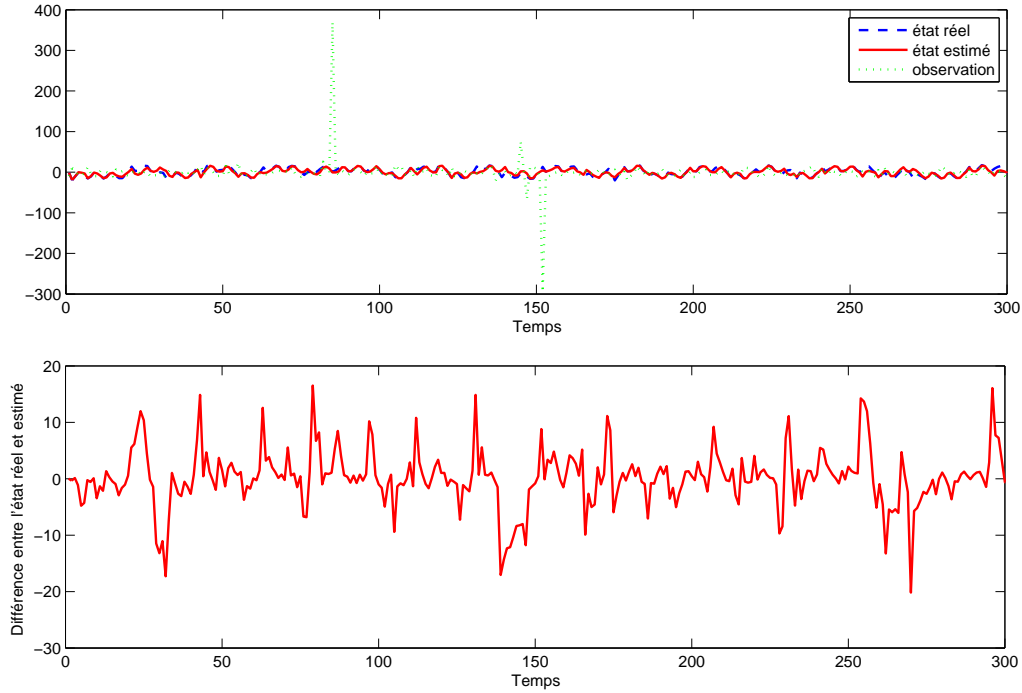
$$D_{KL}(F || \hat{F}_t) = \int_{-\infty}^{+\infty} F(\mathbf{y}) \log \left(\frac{F(\mathbf{y})}{\hat{F}_t(\mathbf{y})} \right) d\mathbf{y} \quad (5.55)$$

$$D_H^2(F, \hat{F}_t) = \int_{-\infty}^{+\infty} \left(\sqrt{F(\mathbf{y})} - \sqrt{\hat{F}_t(\mathbf{y})} \right)^2 d\mathbf{y} \quad (5.56)$$

La valeur de chacune de ces intégrales est approchée numériquement en utilisant la méthode des trapèzes.

la version modifiée de l'algorithme 18 est appliqué avec les trois modèles proposés en supposant que les observations \mathbf{y}_t sont simulées pour $t = 1, \dots, T$ à partir du

FIGURE 5.7 Estimation de l'état dans le cas où la densité du bruit d'observation est un mélange de deux distributions α -stables $F \sim 0.3\mathcal{S}(1.1, 0, 2, 0) + 0.7\mathcal{S}(1.2, 0.5, 1.5, -7)$ modélisée par un DPM de distributions de Cauchy. (Haut) Etat réel, état estimé et l'observation. (Bas) Evolution temporelle de l'erreur entre l'état réel et l'état estimé.



mélange de distributions α -stables suivant :

$$\mathbf{y}_t \sim 0.4\mathcal{S}_1(0.8, 0.5, 1, -7) + 0.2\mathcal{S}_1(1.2, 0, 0.5, 0) + 0.4\mathcal{S}_1(1.5, 0.5, 0.5, 5), \quad t = 1, \dots, 1000 \quad (5.57)$$

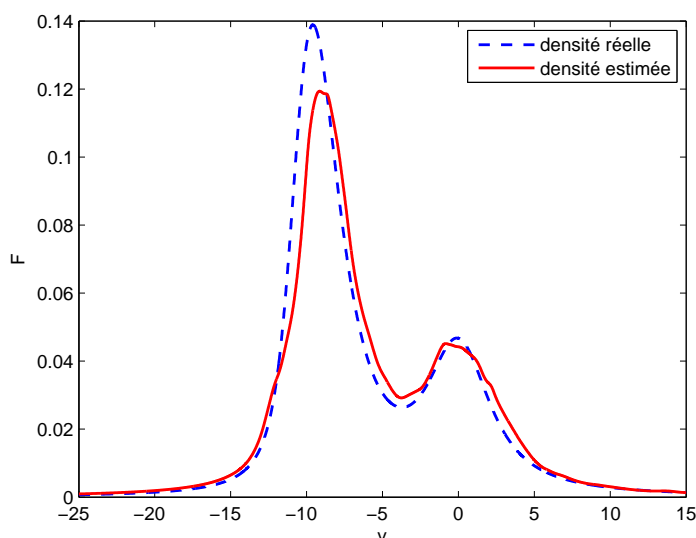
Les hyperparamètres de la distribution de base de chacun des modèles proposés sont fixés comme suit :

- DPM de distributions gaussiennes : $\mu_0 = 0$, $\kappa_0 = 0.01$, $\nu_0 = 8$ et $\Lambda_0 = 2$;
- DPM de distributions de Cauchy : $a^a = 4$, $b^a = 1.5$, $m^m = 0$ et $\Sigma^m = 50$;
- DPM de distributions α -stables : $a^\gamma = 4$, $b^\gamma = 1.5$, $m^\mu = 0$ et $\Sigma^\mu = 50$.

Pour les trois modélisations, le facteur d'échelle α_{DPM} est fixé à 3 et l'algorithme considéré est utilisé avec $N = 200$ particules et $N_{IS} = 100$ particules auxiliaires.

Sur la figure 5.11 sont représentées la densité réelle ainsi que la densité estimée obtenue avec chacun des modèles proposés. La figure 5.12 montre l'allongement des

FIGURE 5.8 Densité réelle et densité estimée du bruit d'observation. La densité réelle est un mélange de deux distributions α -stables $F \sim 0.3\mathcal{S}(1.1, 0, 2, 0) + 0.7\mathcal{S}(1.2, 0.5, 1.5, -7)$ modélisée par un DPM de distributions de Cauchy.



queues des différentes densités illustrées dans la figure 5.11. Les courbes d'évolution en fonction du temps de la distance de Kullback-Leibler pour les différentes modélisations considérées et celles de la distance de Hellinger sont respectivement tracées sur les figures 5.13 et 5.14. Les mesures de la distance de Kullback-Leibler $D_{KL}(F||\hat{F}_T)$ ainsi que celles de la distance de Hellinger $D_H^2(F, \hat{F}_T)$ obtenues à l'instant $t = T$ pour chacun des trois modèles sont reportées dans la table 5.1. A partir de ces figures, nous pouvons constater la convergence progressive de la densité estimée vers la densité réelle dans le cas des trois modèles considérés. La table 5.1 montre que les mesures des distances les plus élevées sont obtenues dans le cas où le mélange est modélisé par un DPM de distributions gaussiennes, tandis que le modèle basé sur un DPM de distributions α -stables permet d'avoir les mesures des distances les plus faibles, suivi par le modèle basé sur un DPM de Cauchy avec des écarts négligeables. Comme illustré dans les figures 5.11 et 5.12, l'écart des performances entre les différents modèles réside principalement au niveau de l'estimation des queues des densités. En particulier, le modèle par DPM de gaussiennes n'arrive pas à bien estimer les queues de distributions par rapport aux deux autres modèles.

FIGURE 5.9 Estimation de l'état dans le cas où la densité du bruit d'observation est un mélange de deux distributions α -stables $F \sim 0.7\mathcal{S}_1(1.4, 0, 2.2, 0) + 0.3\mathcal{S}_1(1.4, 0.5, 1.5, -10)$ modélisée par un DPM de distributions α -stables. (Haut) Etat réel, état estimé et l'observation. (Bas) Evolution temporelle de l'erreur entre l'état réel et l'état estimé.

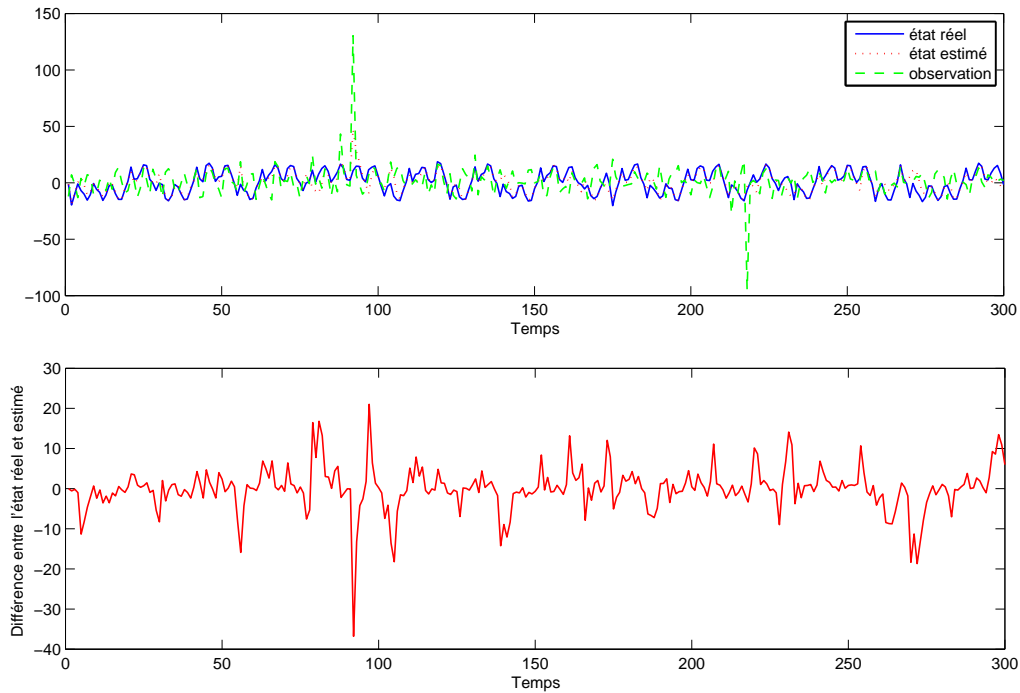


TABLE 5.1 – Comparaison entre les trois modélisations proposées en termes des distances de Kullback-Leibler et de Hellinger dans le cas où la densité réelle est un mélange de trois distributions α -stables $F \sim 0.4\mathcal{S}(0.8, 0.5, 1, -7) + 0.2\mathcal{S}(1.2, 0, 0.5, 0) + 0.4\mathcal{S}(1.5, 0.5, 0.5, 5)$.

Distance	DPM de gaussiennes	DPM de Cauchy	DPM d' α -stable
Kullback-Leibler	0.1304	0.0161	0.0092
Hellinger	0.0138	0.0040	0.0023

5.3.5 Discussion

L'étude comparative menée dans le paragraphe précédent a débouché sur les conclusions suivantes. Le modèle basé sur un DPM de distributions α -stables fournit les meilleures performances en termes de distance entre la densité réelle et la

FIGURE 5.10 Densité réelle et densité estimée du bruit d'observation. La densité réelle est un mélange de deux distributions α -stables $F \sim 0.7\mathcal{S}_1(1.4, 0, 2.2, 0) + 0.3\mathcal{S}_1(1.4, 0.5, 1.5, -10)$ modélisée par un DPM de distributions α -stables.

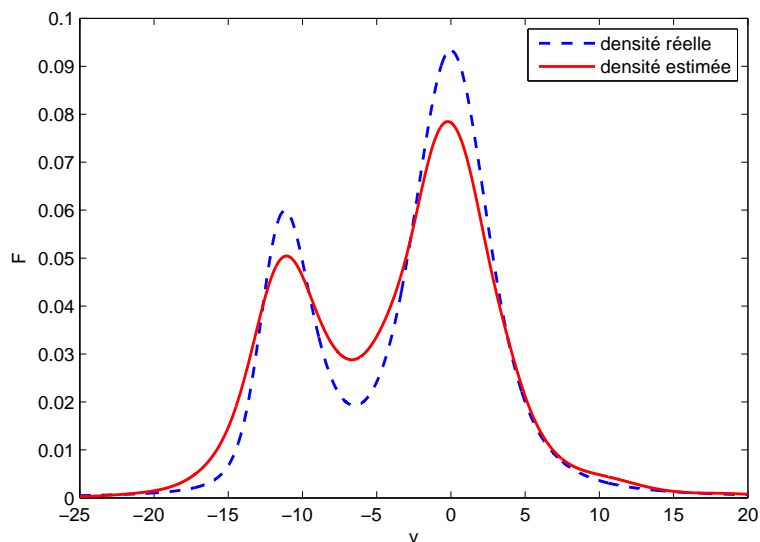
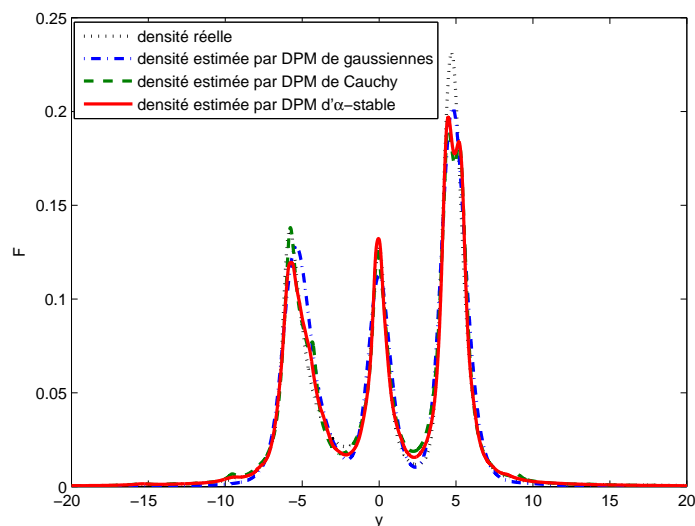


FIGURE 5.11 Densité réelle et densités estimées avec les différentes modélisations proposées. La densité réelle est un mélange de trois distributions α -stables $F \sim 0.4\mathcal{S}(0.8, 0.5, 1, -7) + 0.2\mathcal{S}(1.2, 0, 0.5, 0) + 0.4\mathcal{S}(1.5, 0.5, 0.5, 5)$.



densité estimée. Ces résultats sont tout à fait prévisibles et leurs justifications découlent directement de la définition même d'un DPM. Cependant, ce modèle présente

FIGURE 5.12 Estimation des queues de la distribution avec les différentes modélisations proposées. La densité réelle est un mélange de trois distributions α -stables $F \sim 0.4\mathcal{S}(0.8, 0.5, 1, -7) + 0.2\mathcal{S}(1.2, 0, 0.5, 0) + 0.4\mathcal{S}(1.5, 0.5, 0.5, 5)$.

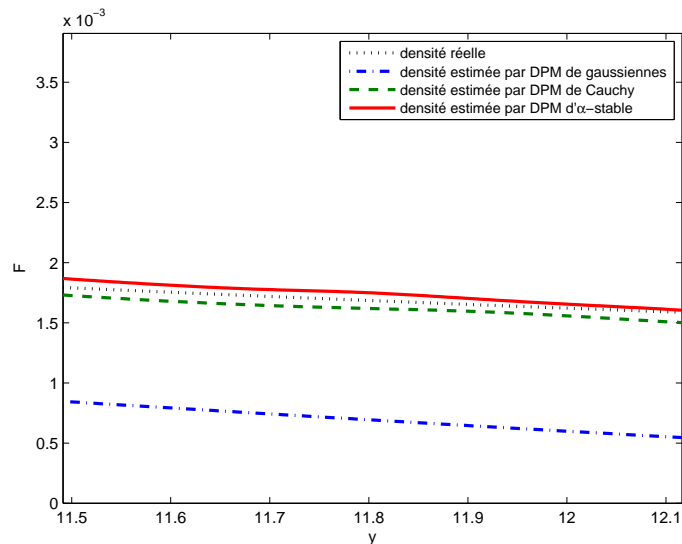
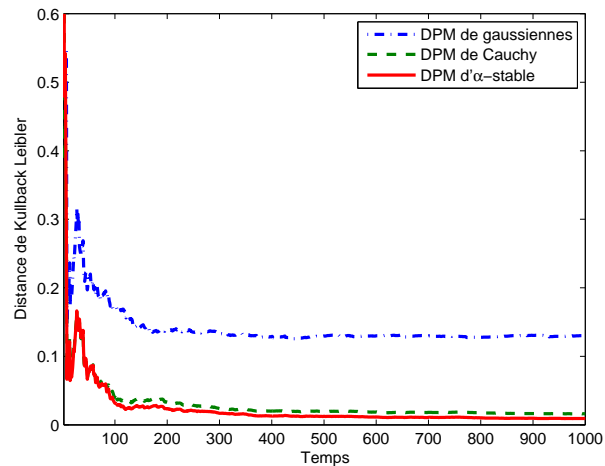
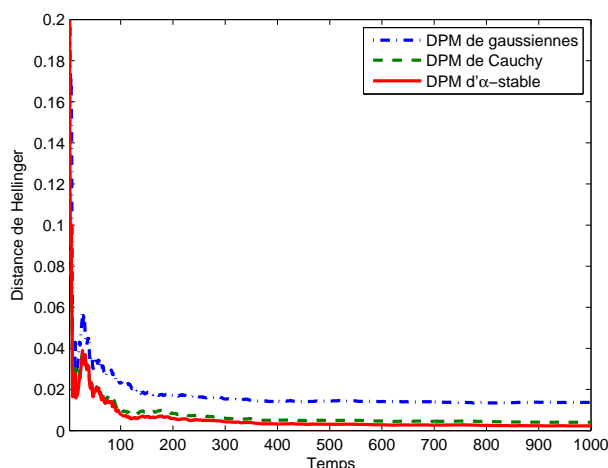


FIGURE 5.13 Evolution temporelle de la distance de Kullback-Leibler entre la densité réelle et la densité estimée avec les différentes modélisations proposées dans le cas où la densité réelle est un mélange de trois distributions α -stables $F \sim 0.4\mathcal{S}(0.8, 0.5, 1, -7) + 0.2\mathcal{S}(1.2, 0, 0.5, 0) + 0.4\mathcal{S}(1.5, 0.5, 0.5, 5)$.



l'inconvénient d'une grande complexité en termes de temps de calcul. Ceci est dû principalement au fait qu'on est obligé d'évaluer numériquement la fonction de la densité de probabilité de la distribution α -stable pour chaque particule. Le modèle

FIGURE 5.14 Evolution temporelle de la distance de Hellinger entre la densité réelle et la densité estimée avec les différentes modélisations proposées dans le cas où la densité réelle est un mélange de trois distributions α -stables $F \sim 0.4\mathcal{S}(0.8, 0.5, 1, -7) + 0.2\mathcal{S}(1.2, 0, 0.5, 0) + 0.4\mathcal{S}(1.5, 0.5, 0.5, 5)$.



basé sur un DPM de distributions de Cauchy permet de réduire considérablement les temps de calculs par rapport au modèle basé sur un DPM de distributions α -stables, tout en offrant des performances presque similaires. De plus, ce modèle permet d'obtenir de meilleurs résultats d'estimation en comparaison avec le modèle basé sur un DPM de distributions gaussiennes, en particulier au niveau des queues des distributions. Cela s'explique par le fait que les distributions de Cauchy se caractérisent par des queues à décroissance algébrique comme celles observées dans le cas des processus α -stables, contrairement à la décroissance exponentielle des queues des distributions gaussiennes. Nous pouvons ainsi conclure que le modèle basé sur un DPM de distributions de Cauchy réalise le meilleur compromis entre performances et complexité en temps de calcul.

5.4 Conclusion

Dans ce chapitre, nous avons abordé la problématique de l'estimation conjointe de l'état et de la densité du bruit dans les modèles dynamiques non linéaires en présence d'un bruit d'observation α -stable. Pour ce faire, des modèles non paramétriques bayésiens basés sur les DPMs ont été introduits pour modéliser le bruit

d'observation. Des algorithmes de Monte Carlo séquentiels basés sur des densités d'importance efficaces ont été ensuite proposés afin de réaliser l'estimation conjointe de l'état et de la densité du bruit. Finalement, ces algorithmes ont été testés sur un modèle dynamique fortement non linéaire et une étude comparative entre les différentes modélisations proposées pour le bruit d'observation a été menée. Les différents résultats obtenus confirment la pertinence des trois modèles introduits pour la modélisation d'un processus α -stable ainsi que la validité et la robustesse des algorithmes proposés pour l'estimation conjointe de l'état et de la densité du bruit d'observation. En particulier, l'étude comparative menée a montré que les DPMs de distributions de Cauchy sont très appropriés pour la modélisation des processus α -stables. Ce résultat est très intéressant dans le cas des distributions α -stables et permet de surmonter les difficultés souvent rencontrées lors de l'inférence en l'absence d'une expression analytique de leur pdf. De plus, dans le cadre des modèles dynamiques non linéaires, nous avons montré que l'estimation de l'état est tout à fait possible même en présence d'un bruit extrêmement impulsif ayant des statistiques inconnues. Finalement, il faut noter que les différents modèles introduits et algorithmes proposés sont valables pour n'importe quel type de bruit. Dans ce chapitre, nous nous sommes focalisés sur le cas α -stable pour rester dans le cadre d'étude de cette thèse.

Conclusion

L'objectif principal de ce travail de thèse est de développer de nouvelles techniques robustes, basées sur l'inférence bayésienne, pour la réduction du bruit dans des environnements impulsifs. Dans de tels environnements, l'hypothèse de gaussianité du bruit n'est plus vérifiée à cause de la grande variabilité des perturbations rencontrées. Une modélisation plus adaptée repose sur les distributions non gaussiennes à queue lourdes en particulier les distributions α -stables. Dans cette thèse, nous nous sommes donc focalisés à l'étude du cas des bruits impulsifs modélisés par les distributions α -stables. L'approche proposée fait appel à des techniques de filtrage particulière pour l'estimation conjointe de l'état et du bruit α -stable. Dans le chapitre 1, nous avons introduit les différents modèles adaptés à la modélisation des phénomènes impulsifs. Nous avons présenté plus en détail les distributions α -stables. Le Chapitre 2 a été consacré à l'inférence bayésienne. Dans le chapitre 3, nous nous sommes intéressés au problème de réduction du bruit α -stable symétrique dans le cadre des systèmes OFDM. Après une brève description de ces systèmes, nous avons décrit les principales approches proposées dans la littérature pour la compensation de l'effet du bruit impulsif dans les systèmes OFDM. Ensuite, nous avons proposé une nouvelle approche pour la réduction du bruit impulsif dans les systèmes OFDM. Contrairement aux méthodes existantes, nous supposons que les paramètres du bruit α stable symétrique sont inconnus du récepteur. Un filtre particulière Rao-Blackwellisé permettant l'estimation conjointe des symboles OFDM émis et des paramètres du bruit α -stable a été proposé. L'efficacité et la robustesse du filtre proposé ont été illustrées à travers plusieurs simulations. Les résultats obtenus montrent une estimation correcte des symboles OFDM émis ainsi que des paramètres

du bruit α -stable symétrique. Notre approche permet d'améliorer les performances en termes de taux d'erreur binaire par rapport aux méthodes classiques de Blanking et de Clipping. Le chapitre 4 donne les éléments de bases de l'estimation non paramétrique bayésienne des densités de probabilités avec les mélanges des processus de Dirichlet. Finalement, dans le chapitre 5, le problème de réduction du bruit impulsif a été étudié dans un contexte plus large celui des systèmes dynamiques non linéaires en présence de bruit impulsif modélisé par des distributions α -stables. Aucune restriction n'est imposée sur le bruit α -stable considéré, il peut être multimodal ou unimodal, symétrique ou asymétrique, etc. A cet effet, une approche non paramétrique bayésienne permettant l'estimation conjointe de l'état et de la densité de probabilité du bruit d'observation a été introduite. L'approche proposée repose sur la modélisation de la densité de probabilité du bruit d'observation à l'aide d'un mélange de processus de Dirichlet. Trois modélisations du bruit d'observation ont été formulées : DPM de distributions gaussiennes, DPM de distributions de Cauchy et DPM de distributions α -stables. Des algorithmes de Monte Carlo séquentiels basés sur des densités d'importance efficaces ont été proposés. Les performances de ces algorithmes ont été étudiées sur un système fortement non linéaire et une étude comparative entre les différentes modélisations proposées pour le bruit d'observation a été menée. Les résultats obtenus confirment la pertinence des trois modèles introduits pour la modélisation d'un processus α -stable ainsi que l'efficacité des algorithmes proposés pour l'estimation conjointe de l'état et de la densité du bruit d'observation. En particulier, nous avons montré que les DPMs de distributions de Cauchy sont très appropriés pour modéliser un processus α -stable. Ce résultat est très intéressant et permet de surmonter les difficultés rencontrées lors de l'inférence dans le cas des distributions α -stables en l'absence d'une expression analytique de leur pdf. Dans le cadre des systèmes dynamiques non linéaires, nous avons montré qu'une estimation de l'état est tout à fait possible même en présence d'un bruit extrêmement impulsif.

Perspectives

Dans le chapitre 3, nous avons proposé un filtre particulaire Rao-Blackwellisé pour l'estimation en ligne des symboles OFDM émis et des paramètres du bruit

α -stable. Nous n'avons traité que le cas d'un bruit α -stable symétrique. Il serait intéressant de traiter le cas général. Une solution envisageable est d'appliquer l'algorithme 18 proposé dans le chapitre 5.

Dans le chapitre 5, nous avons supposé que les hyperparamètres des processus de Dirichlet considérés sont fixes et connus. Il pourrait être envisageable de supposer que ces hyperparamètres sont inconnus. Ceci revient à considérer ces hyperparamètres comme des variables aléatoires tirés selon des distributions *a priori* et de les inclure dans l'inférence. De plus, nous supposons que la distribution du bruit d'observation est fixe. Une extension de ce travail pourrait concerner des bruits d'observation avec des distributions qui varient dans le temps. Une solution à ce problème consiste d'adapter l'approche proposée dans le chapitre 6 de la thèse de Caron [Car06].

Enfin, dans le cadre de cette thèse, notre intérêt s'est focalisé sur les applications temps réel pour lesquels les données arrivent séquentiellement. Par conséquent, les différentes solutions proposées sont dédiés à l'estimation en ligne. Cependant, en utilisant les méthodes MCMC, les différentes approches proposées dans cette thèse pourraient être étendues afin de permettre une estimation hors ligne.

Bibliographie

- [AD00] C. ANDRIEU et A. DOUCET : Particle filtering for partially observed Gaussian state space models. *Journal of the Royal Statistical Society B*, 64:827–836, 2000.
- [ADA05] F. ABDELKEFI, P. DUHAMEL et F. ALBERGE : Impulsive noise cancellation in multicarrier transmission. *IEEE Transactions on Communications*, 53(1):94–106, 2005.
- [ADAA08] F. ABDELKEFI, P. DUHAMEL, F. ALBERGE et J. AYADI : On the use of cascade structure to correct impulsive noise in multicarrier systems. *IEEE Transactions on Communications*, 56(11):1844–1858, 2008.
- [AIH94] S. AMBIKE, J. ILOW et D. HATZINAKOS : Detection for binary transmission in a mixture of Gaussian noise and impulsive noise modeled as an alpha-stable process. *IEEE Signal Processing Letters*, 1(3):55–57, 1994.
- [AM74] D. F. ANDREWS et C. L. MALLOWS : Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(1):99–102, 1974.
- [AMGC02] M. S. ARULAMPALAM, S. MASKELL, N. GORDON et T. CLAPP : A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.
- [ANK06] P. AMIRSHAHI, S. M. NAVIDPOUR et M. KAVEHRAD : Performance analysis of uncoded and coded OFDM broadband transmission

- over low voltage power-line channels with impulsive noise. *IEEE Transactions on Power Delivery*, 21(4):1927–1934, 2006.
- [Ant74] C. E. ANTONIAK : Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics*, 2(6):1152–1174, 1974.
- [AS64] M. ABRAMOWITZ et I. A. STEGUN : *Handbook of Mathematical Functions*. Dover, New York, fifth édition, 1964.
- [Ber52] H. BERGSTRÖM : On some expansions of stable distribution functions. *Arkiv för Matematik*, 2:375–378, 1952.
- [BH06] N. C. BEAULIEU et B. HU : An adaptive threshold soft-limiting UWB receiver with improved performance in multiuser interference. *Dans International Conference on Ultra-Wideband*, pages 405–410, 2006.
- [BH08] N. C. BEAULIEU et B. HU : Soft-limiting receiver structures for time-hopping UWB in multiple-access interference. *IEEE Transactions on Vehicular Technology*, 57(2):810–818, 2008.
- [BM73] D. BLACKWELL et J. MACQUEEN : Ferguson distributions via Polya urn schemes. *The Annals of Statistics*, 1:353–355, 1973.
- [BN99] J. S. BODENSCHATZ et C. L. NIKIAS : Maximum-likelihood symmetric α -stable parameter estimation. *IEEE Transactions on Signal Processing*, 47(5):1382–1384, 1999.
- [BSF08] N. C. BEAULIEU, H. SHAO et J. FIORINA : P-order metric UWB receiver structures with superior performance. *IEEE Transactions on Communications*, 56(10):1666–1676, 2008.
- [Buc95] D. J. BUCKLE : Bayesian inference for stable distributions. *Journal of the American Statistical Association*, 90(430):605–613, 1995.
- [BY90] B. W. BRORSEN et S-R. YANG : Maximum likelihood estimates of symmetric stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 19(4):1459–1464, 1990.
- [BY09] N. C. BEAULIEU et D. J. YOUNG : Designing time-hopping ultra-wide bandwidth receivers for multiuser interference environments. *Proceedings of the IEEE*, 97(2):255–284, 2009.

- [CANN08] G. CAIRE, T. Y. AL NAFFOURI et A. K. NARAYANAN : Impulse noise cancellation in OFDM : an application of compressed sensing. *Dans IEEE International Symposium on Information Theory*, pages 1293–1297, 2008.
- [Car06] F. CARON : *Inférence bayésienne pour la détermination et la sélection de modèles stochastiques*. Thèse de doctorat, Ecole Centrale de Lille / Université des Sciences et Technologie de Lille, 2006.
- [Cas04] R. CASARIN : Bayesian inference for mixtures of stable distributions, 2004.
- [CD89] M. H. L. CHAN et R. W. DONALDSON : Amplitude, width, and interarrival distributions for noise impulses on intrabuilding power line communication networks. *IEEE Transactions on Electromagnetic Compatibility*, 31(3):320–323, 1989.
- [CD05] V. CELLINI et G. DONA : A novel joint channel and multi-user interference statistics estimator for UWB-IR based on Gaussian mixture model. *Dans IEEE International Conference on Ultra-Wideband*, pages 655–660, 2005.
- [CDD⁺08] F. CARON, M. DAVY, A. DOUCET, E. DUFLOS et P. VANHEEGHE : Bayesian inference for linear dynamic models with Dirichlet process mixtures. *IEEE Transactions on Signal Processing*, 56(1):71–84, 2008.
- [CGM07] O. CAPPÉ, S. GODSILL et E. MOULINES : An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE*, 95(5), 2007.
- [Cha66] R. W. CHANG : Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell Systems Technical Journal*, 45:1775–1796, 1966.
- [Che03] Z. CHEN : Bayesian filtering : From Kalman filters to particle filters, and beyond. *Statistics*, pages 1–69, 2003.
- [CL00] R. CHEN et J. S. LIU : Mixture Kalman filters. *Journal of the Royal Statistical Society B*, 62:493–508, 2000.

- [CMS76] J. M. CHAMBERS, C. L. MALLOWS et B. W. STUCK : A method for simulating stable random variables. *Journal of the American Statistical Association*, 71(354):340–344, 1976.
- [CR96] G. CASELLA et C. P. ROBERT : Rao-Blackwellisation of sampling schemes. *Biometrika*, 83(1):81–94, 1996.
- [CZQ07] N. CHEN, G. T. ZHOU et H. QIAN : Power efficiency improvements through peak-to-average power ratio reduction and power amplifier linearization. *EURASIP Journal on Advances in Signal Processing*, 2007(1), 2007.
- [DC05] R. DOUC et O. CAPPÉ : Comparison of resampling schemes for particle filtering. *Dans International Symposium on Image and Signal Processing and Analysis*, pages 64–69, 2005.
- [DDFG01] A. DOUCET, N. DE FREITAS et N. GORDON : *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.
- [Dev86] L. DEVROYE : *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986.
- [DGA00] A. DOUCET, S. GODSILL et C. ANDRIEU : On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.
- [DK06] Y. DHIBI et T. KAISER : On the impulsiveness of multiuser interferences in TH-PPM-UWB systems. *IEEE Transactions on Signal Processing*, 54(7):2853–2857, 2006.
- [DuM73] W. H. DUMOUCHEL : On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics*, 1(5), 1973.
- [DW01] N. DINUR et D. WULICH : Peak-to-average power ratio in high-order OFDM. *IEEE Transactions on Communications*, 49(6):1063–1072, 2001.
- [DW05] A. DOUCET et X. WANG : Monte Carlo methods for signal processing : a review in the statistical signal processing context. *IEEE Signal Processing Magazine*, 22(6):152–170, 2005.

- [ECD08] T. ERSEGHE, V. CELLINI et G. DONA : On UWB impulse radio receivers derived by modeling MAI as a Gaussian mixture process. *IEEE Transactions on Wireless Communications*, 7(6):2388–2396, 2008.
- [EGCA⁺10] H. EL GHANNUDI, L. CLAVIER, N. AZZAOU, F. SEPTIER et P. A. ROLLAND : α -stable interference modeling and Cauchy receiver for an IR-UWB *ad hoc* network. *IEEE Transactions on Communications*, 58(6):1748–1757, 2010.
- [EGCR07] H. EL GHANNUDI, L. CLAVIER et P. A. ROLLAND : Modeling multiple access interference in *Ad Hoc* networks based on IR-UWB signals up-converted to 60 GHz. *Dans European Conference on Wireless Technologies*, pages 106–109, 2007.
- [Esc94] M. D. ESCOBAR : Estimating normal means with a Dirichlet process prior. *Journal of the American Statistical Association*, 89(425):268–277, 1994.
- [EW95] M. D. ESCOBAR et M. WEST : Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90(430):577–588, 1995.
- [Fam65] E. F. FAMA : The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105, 1965.
- [Fel68] W. FELLER : *An Introduction to Probability Theory and its Applications*, volume 1. Wiley, New York, 3rd édition, 1968.
- [Fer73] T. FERGUSON : A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1:209–230, 1973.
- [Fer74] T. FERGUSON : Prior distributions on spaces of probability measures. *The Annals of Statistics*, 2(4):615–629, 1974.
- [Fio06] J. FIORINA : WLC28-2 : A simple IR-UWB receiver adapted to multi-user interferences. *Dans IEEE Global Telecommunications Conference*, pages 1–4, 2006.
- [FR68] E. F. FAMA et R. ROLL : Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, 63(323):817–836, 1968.

- [FR71] E. F. FAMA et R. ROLL : Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66(334):331–338, 1971.
- [GEK08] D. GENCAGA, A. ERTUZUN et E. E. KURUOGLU : Modeling of non-stationary autoregressive alpha-stable processes by particle filters. *Digital Signal Processing*, 18(3), 2008.
- [GG84] S. GEMAN et D. GEMAN : Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, 1984.
- [Gho96] M. GHOSH : Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems. *IEEE Transactions on Communications*, 44(2):145–147, 1996.
- [GK99] S. GODSILL et E. E. KURUOGLU : Bayesian inference for time series with heavy-tailed symmetric α -stable noise processes. *Dans Proceedings of Applications of heavy tailed distributions in economics, engineering and statistics*, Washington, DC, USA, 1999.
- [God99] S. GODSILL : MCMC and EM-based methods for inference in heavy-tailed processes with α -stable innovations. *Dans Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, 1999.
- [God00] S. GODSILL : Inference in symmetric alpha-stable noise using MCMC and the slice sampler. *Dans Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pages 3806–3809, Washington, DC, USA, 2000.
- [GPA06] J. G. GONZALEZ, J. L. PAREDES et G. R. ARCE : Zero-order statistics : A mathematical framework for the processing and characterization of very impulsive signals. *IEEE Transactions on Signal Processing*, 54(10):3839–3851, 2006.
- [GRE95] P. J. GREEN : Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

- [GSS93] N. J. GORDON, D. J. SALMOND et A. F. M. SMITH : Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F Radar and Signal Processing*, 140(2):107–113, 1993.
- [Has70] W. K. HASTINGS : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [HB08] B. HU et N. C. BEAULIEU : On characterizing multiple access interference in TH-UWB systems with impulsive noise models. *Dans IEEE Radio and Wireless Symposium*, pages 879–882, 2008.
- [HFII06] T. HIRAKAWA, M. FUJII, M. ITAMI et K. ITOH : A study on iterative impulse noise reduction in OFDM signal by recovering time domain samples. *Dans Proceeding of the International Symposium on Power-Line Communications and its Applications*, pages 325–330, 2006.
- [Hol19] J. HOLTSMARK : Über die Verbreiterung von Spektrallinien. *Annalen der Physik*, 363:577–630, 1919.
- [HSG06] J. D. HOL, T. B. SCHÖN et F. GUSTAFSSON : On resampling algorithms for particle filters. *Dans Nonlinear Statistical Signal Processing Workshop*, 2006.
- [HV01] J. HARING et A. J. H. VINCK : Coding for impulsive noise channels. *Dans Proceeding of the International Symposium on Power-Line Communications and its Applications*, pages 103–108, 2001.
- [HV03] J. HARING et A. J. H. VINCK : Iterative decoding of codes over complex numbers for impulsive noise channels. *IEEE Transactions on Information Theory*, 49(5):1251–1260, 2003.
- [HZK11] R. HUANG, H. ZHENG et E. E. KURUOGLU : Time-varying ARMA stable process estimation using sequential Monte Carlo. *Signal, Image and Video Processing*, pages 1–8, 2011.
- [IH96] J. ILOW et D. HATZINAKOS : Impulsive noise modeling with stable distributions in fading environments. *Dans IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pages 140–143, 1996.

- [IJ01] H. ISHWARAN et L. F. JAMES : Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96:1316–1332, 2001.
- [IZ00] H. ISHWARAN et M. ZAREPOUR : Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika*, 87(2):371–390, 2000.
- [JDV⁺11] N. JAOUA, E. DUFLOS, P. VANHEEGHE, L. CLAVIER et F. SEPTIER : Impulsive interference mitigation in *ad hoc* networks based on α -stable modeling and particle filtering. *Dans IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3548–3551, 2011.
- [JU97] S. J. JULIER et J. K. UHLMANN : A new extension of the Kalman filter to nonlinear systems. *Dans International Symposium on Aerospace/Defence Sensing, Simulation and Controls*, pages 182–193, Orlando, Florida, 1997.
- [JW93] A. JANICKI et A. WERON : *Simulation and Chaotic Behavior of α -stable Stochastic Processes*. Monographs and Textbooks in Pure and Applied Mathematics. Taylor & Francis, 1993.
- [Kal60] R. E. KALMAN : A new approach to linear filtering and prediction problems. *Journal of the Basic Engineering*, 82(D):35–45, 1960.
- [Kit96] G. KITAGAWA : Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.
- [KLW94] A. KONG, J. S. LIU et W. H. WONG : Sequential imputations and Bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994.
- [KMF98] E. E. KURUOGLU, C. MOLINA et W. J. FITZGERALD : Approximation of α -stable probability densities using finite Gaussian mixtures. *Dans Proceedings of the European Signal Processing Conference*, Rhodes, Greece, 1998.
- [KOI12] T. KITAMURA, K. OHNO et M. ITAMI : Improving of performance of OFDM reception under class A impulsive channel by replica

- signal estimation of impulse. *Dans IEEE International Conference on Consumer Electronics*, pages 620–621, 2012.
- [Kou80] I. A. KOUTROUVELIS : Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, 1980.
- [Kou81] I. A. KOUTROUVELIS : An iterative procedure for the estimation of the parameters of stable laws. *Communications in Statistics - Simulation and Computation*, 10(1):17–28, 1981.
- [Kur01] E. E. KURUOGLU : Density parameter estimation of skewed α -stable distributions. *IEEE Transactions on Signal Processing*, 49(10):2192–2201, 2001.
- [KW95] S. M. KOYON et D. B. WILLIAMS : On the characterization of impulsive noise with α -stable distributions using Fourier techniques. *Dans Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 787–791, 1995.
- [LC98] J. S. LIU et R. CHEN : Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.
- [LG04] M. LOMBARDI et S. GODSILL : Monte Carlo Bayesian filtering and smoothing for TVAR signals in symmetric α -stable noise. *Dans Proceedings of the European Signal Processing Conference*, pages 865–872, Vienna, Austria, 2004.
- [LG06] M. LOMBARDI et S. GODSILL : On-line Bayesian estimation of signals in symmetric α -stable noise. *IEEE Transactions on Signal Processing*, 54:775–779, 2006.
- [Liu96] J. S. LIU : Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, 6:113–119, 1996.
- [Lom07] M. J. LOMBARDI : Bayesian inference for α -stable distributions : A random walk MCMC approach. *Computational Statistics and Data Analysis*, 51:2688–2700, 2007.

- [Man63] B. MANDELBROT : The variation of certain speculative prices. *the Journal of Business*, 36:394–419, 1963.
- [MBA⁺05] L. MIHAYLOVA, P. BRASNETT, A. ACHIM, D. BULL et N. CANAGARAJAH : Particle filtering with alpha-stable distributions. *Dans Workshop on Statistical Signal Processing*, pages 381–386, 2005.
- [McC86] J. H. MCCULLOCH : Simple consistent estimators of stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 15:1109–1136, 1986.
- [MDC99] S. MITTNIK, T. DOGANOGLU et D. CHENYAO : Computing the probability density function of the stable Paretian distribution. *Mathematical and Computer Modelling*, 29(10):235–240, 1999.
- [MFM00] S. MAYMON, J. FRIEDMANN et H. MESSER : A new method for estimating parameters of a skewed alpha-stable distribution. *Dans IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pages 3822–3825, 2000.
- [ML07] J. MITRA et L. LAMPE : Robust detectors for TH IR-UWB systems with multiuser interference. *Dans IEEE International Conference on Ultra-Wideband*, pages 745–750, 2007.
- [MN95] X. MA et C. L. NIKIAS : Parameter estimation and blind channel identification in impulsive signal environments. *IEEE Transactions on Signal Processing*, 43(12):2884–2897, 1995.
- [MQ04] P. MULLER et F. A. QUINTANA : Nonparametric Bayesian data analysis. *Statistical Science*, 19(1):95–110, 2004.
- [MRR⁺53] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER et E. TELLER : Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [MSG05] Y. H. MA, P. L. SO et E. GUNAWAN : Performance analysis of OFDM systems for broadband power line communications under impulsive noise and multipath effects. *IEEE Transactions on Power Delivery*, 20(2):674–682, 2005.

- [MT72] J. MILLER et J. THOMAS : Detectors for discrete-time signals in non-Gaussian noise. *IEEE Transactions on Information Theory*, 18(2):241–250, 1972.
- [MU49] N. METROPOLIS et S. ULAM : The Monte Carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.
- [MUKM02] H. MATSUO, D. UMEHARA, M. KAWAI et Y. MORIHIRO : An iterative detection of OFDM over impulsive noise channel. *Dans Proceeding of the International Symposium on Power-Line Communications and its Applications*, 2002.
- [Nea00] R. M. NEAL : Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9(2):249–265, 2000.
- [NGS+08] M. NASSAR, K. GULATI, A. K. SUJEETH, N. AGHASADEGHI, B. L. EVANS et K. R. TINSLEY : Mitigating near-field interference in laptop embedded wireless transceivers. *Dans IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1405–1408, 2008.
- [Nol97] J. P. NOLAN : Numerical computation of stable densities and distribution functions. *Communications in Statistics-Stochastic Models*, 13(4):759–774, 1997.
- [Nol99] J. P. NOLAN : An algorithm for evaluating stable densities in Zolotarev’s (M) parameterization. *Mathematical and Computer Modelling*, 29(10–12):229–233, 1999.
- [Nol12] J. P. NOLAN : *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston, 2012. In progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- [NS95] C. L. NIKIAS et M. SHAO : *Signal processing with alpha-stable distributions and applications*. John Wiley and Sons, 1995.
- [NUDM05] H. NAKAGAWA, D. UMEHARA, S. DENNO et Y. MORIHIRO : A decoding for low density parity check codes over impulsive noise channels. *Dans International Symposium on Power Line Communications and Its Applications*, pages 85–89, 2005.

- [PCG⁺06] P. C. PINTO, C-C CHONG, A. GIORGETTI, M. CHIARI et M. Z. WIN : Narrowband communication in a Poisson field of ultrawide-band interferers. *Dans IEEE International Conference on Ultra-Wideband*, pages 387–392, 2006.
- [PHL75] A. S. PAULSON, E. W. HOLCOMB et R. A. LEITCH : The estimation of the parameters of the stable laws. *Biometrika*, 62(1):163–170, 1975.
- [Pre72] S. J. PRESS : Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, 67(340):842–846, 1972.
- [PY97] J. PITMAN et M. YOR : The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Annals of Probability*, 25(2):855–900, 1997.
- [RC99] C. P. ROBERT et G. CASELLA : *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1999.
- [RG97] S. RICHARDSON et P. J. GREEN : On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society : Series B (Methodological)*, 59(4):731–792, 1997.
- [RR10] N. ROZIC et J. RADIC : Distribution of the noise squared norm in OFDM systems interfered by class A impulsive noise. *IEEE Communications Letters*, 14(4):318–320, 2010.
- [SAC04] M. G. SANCHEZ, A. V. ALEJOS et I. CUINAS : Urban wide-band measurement of the UMTS electromagnetic environment. *IEEE Transactions on Vehicular Technology*, 53(4):1014–1022, 2004.
- [SB10] H. SHAO et N. C. BEAULIEU : Block coding for impulsive Laplacian noise. *Dans IEEE International Conference on Communications*, pages 1–6, 2010.
- [SDMRG08] F. SEPTIER, Y. DELIGNON, A. MENHAJ-RIVENQ et C. GARNIER : Nonpilot-aided sequential Monte Carlo method to joint signal, phase noise and frequency offset estimation in multicarrier

- systems. *Eurasip Journal on Advances in Signal Processing*, 2008, 2008.
- [SDMRG10] F. SEPTIER, Y. DELIGNON, A. MENHAJ-RIVENQ et C. GARNIER : Pilot-aided sequential Monte Carlo estimation of phase distortions and transmitted symbols in multicarrier systems. *Journal of Electrical and Computer Engineering*, 2010, 2010.
- [Sep08] F. SEPTIER : *Sequential Monte Carlo methods for multicarrier systems in the presence of phase distortions*. Thèse de doctorat, University of Valenciennes / TELECOM Lille1, France, 2008.
- [Set94] J. SETHURAMAN : A constructive definition of Dirichlet priors. *Statistica Sinica*, 4:639–650, 1994.
- [SGKR06] D. SALAS-GONZALEZ, E. E. KURUOGLU et D. P. RUIZ : Estimation of mixtures of symmetric alpha stable distributions with an unknown number of components. *Dans Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 5, page V, 2006.
- [SGKR09] D. SALAS-GONZALEZ, E. E. KURUOGLU et D. P. RUIZ : Finite mixture of α -stable distributions. *Digital Signal Processing*, 19(2), 2009.
- [SGKR10] D. SALAS-GONZALEZ, E. E. KURUOGLU et D. P. RUIZ : Modeling with mixture of symmetric stable distributions using Gibbs sampling. *Signal Processing*, 90(3):774–783, 2010.
- [SGN05] T. SCHON, F. GUSTAFSSON et P.-J. NORDLUND : Marginalized particle filters for mixed linear/nonlinear state-space models. *IEEE Transactions on Signal Processing*, 53(7):2279–2289, 2005.
- [Sou92] E. S. SOUSA : Performance of a spread spectrum packet radio network link in a Poisson field of interferers. *IEEE Transactions on Information Theory*, 38(6):1743–1754, 1992.
- [ST94] G. SAMORODNITSKY et M.S. TAQQU : *Stable non-Gaussian random processes : stochastic models with infinite variance*. Chapman and Hall, London, 1994.

- [Sud06] E. B. SUDDERTH : *Graphical models for visual object recognition and tracking*. Thèse de doctorat, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2006.
- [TC94] M. W. THOMPSON et H-S. CHANG : Coherent detection in Laplace noise. *IEEE Transactions on Aerospace and Electronic Systems*, 30(2):452–461, 1994.
- [Teh10] Y. W. TEH : Dirichlet processes. *Dans Encyclopedia of Machine Learning*. Springer, 2010.
- [TN95a] P. TSAKALIDES et C.L. NIKIAS : Maximum likelihood localization of sources in noise modeled as a stable process. *IEEE Transactions on Signal Processing*, 43(11):2700–2713, 1995.
- [TN95b] G. A. TSIHRINTZIS et C. L. NIKIAS : Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. *IEEE Transactions on Communications*, 43(234):904–914, 1995.
- [Tsi99] E. G. TSIONAS : Monte Carlo inference in econometric models with symmetric stable disturbances. *Journal of Econometrics*, 88(2):365–401, 1999.
- [VVFPCA FR11] J. VILÀ-VALLS, C. FERNANDEZ-PRADES, P. CLOSAS et J. A. FERNÁNDEZ-RUBIO : Bayesian filtering for nonlinear state-space models in symmetric α -stable measurement noise. *Dans Proceedings of the European Signal Processing Conference*, pages 674–678, 2011.
- [WDLS99] S. G. WALKER, P. DAMIEN, P. W. LAUD et A. F. M. SMITH : Bayesian nonparametric inference for random distributions and related functions. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 61(3):485–527, 1999.
- [WE71] S. WEINSTEIN et P. EBERT : Data transmission by frequency-division multiplexing using the discrete Fourier transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, 1971.

- [Wes93] M. WEST : Mixture models, Monte Carlo, Bayesian updating and dynamic models. *Computing Science and Statistics*, 24:325–333, 1993.
- [WGK10] S. WEI, D. L. GOECKEL et P. A. KELLY : Convergence of the complex envelope of bandlimited OFDM signals. *IEEE Transactions on Information Theory*, 56(10):4893–4904, 2010.
- [Wol83] J. WOLF : Redundancy, the discrete Fourier transform, and impulse noise cancellation. *IEEE Transactions on Communications*, 31(3):458–461, 1983.
- [WPC+00] Y. WU, E. PLISZKA, B. CARON, P. BOUCHARD et G. CHOUINARD : Comparison of terrestrial DTV transmission systems : the ATSC 8-VSB, the DVB-T COFDM, and the ISDB-T BST-OFDM. *IEEE Transactions on Broadcasting*, 46(2):101–113, 2000.
- [WPS09] M. Z. WIN, P. C. PINTO et L. A. SHEPP : A mathematical theory of network interference and its applications. *Proceedings of the IEEE*, 97(2):205–230, 2009.
- [WST09] K. C. WIKLUNDH, P. F. STENUMGAARD et H. M. TULLBERG : Channel capacity of Middleton’s class A interference channel. *Electronics Letters*, 45(24):1227–1229, 2009.
- [Wu99] Y. WU : Performance comparison of ATSC 8-VSB and DVB-T COFDM transmission systems for digital television terrestrial broadcasting. *IEEE Transactions on Consumer Electronics*, 45(3):916–924, 1999.
- [WVdM00] E. A. WAN et R. Van der MERVE : The unscented Kalman filter for nonlinear estimation. *Dans The IEEE Symposium on Adaptive Systems for Signal Processing, Communications and Control*, pages 153–158, Lake Louise, Alta., Canada, 2000.
- [YES07] M. S. YOUSUF et M. EL-SHAFAEI : Power line communications : An overview - part i. *Dans International Conference on Innovations in Information Technology*, pages 218–222, 2007.

- [YTM89] K. YAMAUCHI, N. TAKAHASHI et M. MAEDA : Parameter measurement of class A interference on power line. *IEICE Transactions*, E72-E(1):7–9, 1989.
- [ZB02] A. M. ZOUBIR et R. F. BRCICH : Multiuser detection in heavy tailed noise. *Digital Signal Processing*, 12(2–3):262–273, 2002.
- [ZD02] M. ZIMMERMANN et K. DOSTERT : Analysis and modeling of impulsive noise in broad-band powerline communications. *IEEE Transactions on Electromagnetic Compatibility*, 44(1):249–258, 2002.
- [Zhi03] S. V. ZHIDKOV : Impulsive noise suppression in OFDM-based communication systems. *IEEE Transactions on Consumer Electronics*, 49(4):944–948, 2003.
- [Zhi08] S. V. ZHIDKOV : Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers. *IEEE Transactions on Communications*, 56(1):5–9, 2008.
- [Zol86] V. M. ZOLOTAREV : *One-dimensional Stable Distributions*, volume 65. American Mathematical Society, translations of mathematical monographs édition, 1986.
- [ZS75] V. B. ZARITSKII, V. S. Svetnik et L. I. SHIMELEVICH : Monte Carlo technique in problems of optimal data processing. *Automation and Remote Control*, 12:95–103, 1975.

Annexe A : Densité de probabilité de la somme d'une variable aléatoire gaussienne et d'une variable aléatoire de Cauchy

Soit X une variable aléatoire qui suit une loi normale de moyenne μ et de variance σ^2 et Y une variable aléatoire qui suit une loi de Cauchy de paramètre d'échelle a et de paramètre de position b :

$$X \sim \mathcal{N}(\mu, \sigma^2) \quad \text{et} \quad Y \sim \mathcal{C}(a, b)$$

On définit la variable Z comme étant la somme des variables X et Y :

$$Z = X + Y$$

Notre objectif est de déterminer la densité de probabilité de la variable Z . A notre connaissance, il n'existe pas une loi de probabilité usuelle qui caractérise la somme d'une gaussienne et d'une Cauchy. On propose donc d'utiliser le théorème suivant :

Théorème 1. *Soient U et V deux variables aléatoires continues de densités de probabilité respectives $f(u)$ et $g(v)$. Si U et V sont indépendantes, alors la densité de probabilité $h(w)$ de la variable aléatoire W définie par $W = U + V$ est donnée par le produit de convolution des densités f et g :*

$$h(w) = f * g(w) = \int_{-\infty}^{+\infty} f(u)g(w - u)du = \int_{-\infty}^{+\infty} f(w - v)g(v)dv$$

On note Φ le densité de probabilité de Z . D'après le théorème précédent, Φ est

définie comme suit :

$$\begin{aligned}\Phi(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x-t-\mu)^2\right) \times \frac{a}{\pi(a^2+(t-b)^2)} dt \\ &= \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp\left(-\frac{1}{2\sigma^2}(x-t-\mu)^2\right)}{a^2+(t-b)^2} dt\end{aligned}$$

Les techniques classiques de calcul d'intégrales n'ont pas abouti pour déterminer la densité $\Phi(x)$. On a donc eu l'idée d'avoir recours aux équations différentielles. En effet, on propose d'exprimer à l'aide d'une équation différentielle, la densité $\Phi(x)$ en fonction de ses dérivées successives. La résolution de cette équation différentielle va nous permettre de trouver la densité $\Phi(x)$ à une constante près. Pour ce faire, on propose d'appliquer le théorème de Leibniz pour la dérivation d'une intégrale à paramètre :

Théorème 2. Soient I et J deux intervalles non vides de \mathbb{R} et $f : I \times J \mapsto \mathbb{R}$ ou \mathbb{C} . Si pour tout x de I , la fonction $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur J , et si f admet une dérivée partielle $\frac{\partial f}{\partial x}$ vérifiant les hypothèses suivantes :

- pour tout x de I , la fonction $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur J ;
- pour tout t de J , la fonction $x \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue sur I ;
- il existe une fonction g continue par morceaux, positive et intégrable sur J telle que, pour tout $(x, t) \in I \times J$, $\left| \frac{\partial f}{\partial x}(x, t) \right| \leq g(t)$.

alors la fonction $F(x) = \int_J f(x, t) dt$ est dérivable sur I , et pour tout x de I on a :

$$F'(x) = \int_J \frac{\partial f}{\partial x}(x, t) dt$$

On note ϕ la fonction définie par :

$$\begin{aligned}\phi : \mathbb{R}^2 &\mapsto \mathbb{R}^2 \\ (x, t) &\mapsto \frac{\exp\left(-\frac{1}{2\sigma^2}(x-t-\mu)^2\right)}{a^2+(t-b)^2}\end{aligned}$$

La fonction ϕ vérifie les hypothèses du théorème précédent, ainsi que ses dérivées partielles par rapport à x et t . On peut alors définir la dérivée de la densité Φ' en

appliquant le théorème de Leibniz :

$$\begin{aligned}
\Phi'(x) &= \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{-(x-t-\mu) \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^2 a^2 + (t-b)^2} dt \\
&= \frac{\mu-x}{\sigma^2} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\
&+ \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{t \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^2 a^2 + (t-b)^2} dt \\
&= \frac{\mu-x}{\sigma^2} \Phi(x) + I_1
\end{aligned}$$

où I_1 est l'intégrale définie par :

$$I_1 = \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{t \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^2 a^2 + (t-b)^2} dt$$

Les différentes hypothèses du théorème de Leibniz étant toujours vérifiées, on peut l'appliquer une deuxième fois pour exprimer la dérivée seconde de la densité Φ comme suit :

$$\begin{aligned}
\Phi''(x) &= -\frac{1}{\sigma^2} \Phi(x) + \frac{\mu-x}{\sigma^2} \Phi'(x) \\
&+ \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{-(x-t-\mu) t \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^2 a^2 + (t-b)^2} dt \\
&= -\frac{1}{\sigma^2} \Phi(x) + \frac{\mu-x}{\sigma^2} \Phi'(x) + \frac{\mu-x}{\sigma^2} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{t \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\
&+ \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{t^2 \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^4 a^2 + (t-b)^2} dt \\
&= \left(-\frac{1}{\sigma^2}\right) \Phi(x) + \left(\frac{-x+\mu}{\sigma^2}\right) \Phi'(x) + \left(\frac{-x+\mu}{\sigma^2}\right) I_1 + I_2
\end{aligned}$$

où I_2 est l'intégrale définie par :

$$I_2 = \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{t^2 \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{\sigma^4 a^2 + (t-b)^2} dt$$

D'après l'expression de $\Phi'(x)$, l'intégrale I_1 peut s'exprimer comme suit :

$$I_1 = \Phi'(x) + \left(\frac{x - \mu}{\sigma^2} \right) \Phi(x)$$

On obtient donc l'expression suivante de $\Phi''(x)$:

$$\Phi''(x) = \left(-\frac{1}{\sigma^2} \right) \Phi(x) + \left(\frac{-x + \mu}{\sigma^2} \right) \Phi'(x) + \left(\frac{-x + \mu}{\sigma^2} \right) \left[\Phi'(x) + \left(\frac{x - \mu}{\sigma^2} \right) \Phi(x) \right] + I_2$$

En ajoutant et soustrayant $(a^2 - 2bt + b^2)$ à l'intérieur de I_2 , on obtient :

$$\begin{aligned} I_2 &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (t^2 + (a^2 - 2bt + b^2) - (a^2 - 2bt + b^2)) \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (t^2 + a^2 - 2bt + b^2) \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &\quad + \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (2bt) \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &\quad + (-a^2 - b^2) \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (a^2 + (t-b)^2) \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &\quad + \frac{2b}{\sigma^2} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(\frac{t}{\sigma^2} \right) \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &\quad + \frac{-a^2 - b^2}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2)}{a^2 + (t-b)^2} dt \\ &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-\frac{1}{2\sigma^2}(x-t-\mu)^2) dt + \frac{2b}{\sigma^2} I_1 + \frac{-a^2 - b^2}{\sigma^4} \Phi(x) \\ &= I_3 + \frac{2b}{\sigma^2} \left[\Phi'(x) + \left(\frac{x - \mu}{\sigma^2} \right) \Phi(x) \right] + \left(\frac{-a^2 - b^2}{\sigma^4} \right) \Phi(x) \end{aligned}$$

où I_3 est l'intégrale définie par :

$$I_3 = \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2\sigma^2}(x-t-\mu)^2\right) dt$$

En appliquant l'intégrale définie :

$$\int_{-\infty}^{+\infty} \exp(-(ax^2 + bx + c)) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2 - 4ac}{4a}\right)$$

on obtient :

$$\begin{aligned} I_3 &= \frac{1}{\sigma^4} \frac{a}{\pi\sigma\sqrt{2\pi}} \sqrt{2\pi\sigma^2} \exp\left(\frac{4(\mu-x)^2 - 4(\mu-x)^2 2\sigma^2}{4\sigma^4}\right) \\ &= \frac{a}{\sigma^4\pi} \end{aligned}$$

d'où l'expression de I_2

$$I_2 = \frac{a}{\sigma^4\pi} + \frac{2b}{\sigma^2} \left[\Phi'(x) + \left(\frac{x-\mu}{\sigma^2}\right) \Phi(x) \right] + \left(\frac{-a^2 - b^2}{\sigma^4}\right) \Phi(x)$$

et celle de $\Phi''(x)$

$$\begin{aligned} \Phi''(x) &= \left(-\frac{1}{\sigma^2}\right) \Phi(x) + \left(\frac{-x+\mu}{\sigma^2}\right) \Phi'(x) + \left(\frac{-x+\mu}{\sigma^2}\right) \left[\Phi'(x) + \left(\frac{x-\mu}{\sigma^2}\right) \Phi(x) \right] \\ &+ \frac{a}{\sigma^4\pi} + \frac{2b}{\sigma^2} \left[\Phi'(x) + \left(\frac{x-\mu}{\sigma^2}\right) \Phi(x) \right] + \left(\frac{-a^2 - b^2}{\sigma^4}\right) \Phi(x) \\ &= \left[\frac{2b}{\sigma^2} - 2\left(\frac{x-\mu}{\sigma^2}\right) \right] \Phi'(x) + \left[\frac{2b}{\sigma^2} \frac{x-\mu}{\sigma^2} - \frac{a^2 + b^2}{\sigma^4} - \frac{1}{\sigma^2} \right] \Phi(x) + \frac{a}{\sigma^4\pi} \end{aligned}$$

La densité $\Phi(x)$ satisfait donc à l'équation différentielle de second degré suivante :

$$\Phi''(x) + \left[2\left(\frac{x-\mu}{\sigma^2}\right) - \frac{2b}{\sigma^2} \right] \Phi'(x) + \left[\left(\frac{x-\mu}{\sigma^2}\right)^2 - \frac{2b}{\sigma^2} \left(\frac{x-\mu}{\sigma^2}\right) + \left(\frac{a^2 + b^2}{\sigma^4}\right) + \frac{1}{\sigma^2} \right] \Phi(x) = \frac{a}{\sigma^4\pi}$$

Après la résolution de cette équation différentielle à l'aide du logiciel Maple, on

obtient la solution suivante :

$$\begin{aligned} \Phi(x) = & \exp\left(\frac{-1}{2\sigma^2}(x(x-2\mu-2b))\right) \left[C1 \cos\left(\frac{xa}{\sigma^2}\right) + C2 \sin\left(\frac{xa}{\sigma^2}\right) \right. \\ & - \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(2a^2 - 2(\mu+b)^2)\right) \\ & \times \left[\exp\left(\frac{1}{2\sigma^2}(\mu+b+ia)^2\right) \operatorname{erf}\left(\frac{-a+i(x-\mu-b)}{\sqrt{2}\sigma}\right) \left(-\cos\left(\frac{xa}{\sigma^2}\right) + i \sin\left(\frac{xa}{\sigma^2}\right)\right) \right. \\ & \left. \left. - \exp\left(\frac{1}{2\sigma^2}(-\mu-b+ia)^2\right) \operatorname{erf}\left(\frac{-a+i(-x+\mu+b)}{\sqrt{2}\sigma}\right) \left(\cos\left(\frac{xa}{\sigma^2}\right) + i \sin\left(\frac{xa}{\sigma^2}\right)\right) \right] \right] \end{aligned}$$

où $C1$ et $C2$ sont deux constantes à déterminer et erf est la fonction d'erreur définie par :

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-x^2) dx$$

Pour déterminer les deux constantes $C1$ et $C2$, il faut calculer l'expression de départ de $\Phi(x)$ (sous forme d'une intégrale) pour une valeur particulière de x . En effet, pour $x = \mu + b$ on a :

$$\begin{aligned} \Phi(\mu+b) &= \frac{a}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{1}{2\sigma^2}(t-b)^2)}{a^2 + (t-b)^2} dt \\ &= \frac{1}{a\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{a^2}{2\sigma^2}(\frac{t-b}{a})^2)}{1 + (\frac{t-b}{a})^2} dt \end{aligned}$$

En faisant le changement de variable $v = \frac{t-b}{a}$, on obtient :

$$\Phi(\mu+b) = \frac{1}{\pi\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-\frac{a^2}{2\sigma^2}v^2)}{1+v^2} dv$$

En appliquant l'intégrale définie usuelle :

$$\int_{-\infty}^{+\infty} \frac{\exp(-(ct)^2)}{1+t^2} dt = \pi \exp(c^2)(1 - \operatorname{erf}(a))$$

On obtient :

$$\Phi(\mu+b) = \frac{\exp(\frac{a^2}{2\sigma^2})}{\sigma\sqrt{2\pi}} \left[1 - \operatorname{erf}\left(\frac{a}{\sqrt{2}\sigma}\right) \right]$$

D'un autre coté, on a :

$$\begin{aligned}
\Phi(\mu + b) &= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&- \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(-2b^2 - 4\mu b + 2a^2 - 2\mu^2)\right) \\
&\times \left[\exp\left(\frac{1}{2\sigma^2}(\mu + b + ia)^2\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \left(-\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right. \\
&- \left. \exp\left(\frac{1}{2\sigma^2}(-\mu - b + ia)^2\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \left(\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right] \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&- \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(2a^2 - 2(b + \mu)^2)\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \\
&\times \left[\exp\left(\frac{1}{2\sigma^2}((\mu + b)^2 + i2a(\mu + b) - a^2)\right) \left(-\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right. \\
&- \left. \exp\left(\frac{1}{2\sigma^2}((\mu + b)^2 - i2a(\mu + b) - a^2)\right) \left(\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right] \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&- \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (b + \mu)^2)\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \\
&\times \left[\exp\left(\frac{1}{\sigma^2}(ia(\mu + b))\right) \left(-\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right. \\
&- \left. \exp\left(\frac{1}{\sigma^2}(-ia(\mu + b))\right) \left(\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)\right) \right] \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&- \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (b + \mu)^2)\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \\
&\times \left[-\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) \left(\exp\left(\frac{1}{\sigma^2}(ia(\mu + b))\right) + \exp\left(\frac{1}{\sigma^2}(-ia(\mu + b))\right)\right) \right. \\
&+ \left. i \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \left(\exp\left(\frac{1}{\sigma^2}(ia(\mu + b))\right) - \exp\left(\frac{1}{\sigma^2}(-ia(\mu + b))\right)\right) \right] \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&- \frac{\sqrt{2\pi}}{4\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (b + \mu)^2)\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \\
&\times \left[-2 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right)^2 - 2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right] \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right. \\
&+ \frac{\sqrt{2\pi}}{2\pi\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (b + \mu)^2)\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right) \left. \right] \\
&= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right] \\
&+ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{a^2}{2\sigma^2}\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right)
\end{aligned}$$

On obtient donc l'équation suivante :

$$\begin{aligned}\Phi(\mu + b) &= \frac{\exp\left(\frac{a^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \left[1 - \operatorname{erf}\left(\frac{a}{\sqrt{2}\sigma}\right) \right] \\ &= \exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right] \\ &+ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{a^2}{2\sigma^2}\right) \operatorname{erf}\left(\frac{-a}{\sqrt{2}\sigma}\right)\end{aligned}$$

Et puisqu'on a la propriété suivante $\operatorname{erf}(-c) = -\operatorname{erf}(c)$, on obtient :

$$\exp\left(\frac{1}{2\sigma^2}(\mu + b)^2\right) \left[C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \right] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{a^2}{2\sigma^2}\right)$$

donc :

$$C1 \cos\left(\frac{a(\mu + b)}{\sigma^2}\right) + C2 \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (\mu + b)^2)\right)$$

Une solution évidente de cette équation est :

$$C1 = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (\mu + b)^2)\right) \cos\left(\frac{a(\mu + b)}{\sigma^2}\right)$$

et

$$C2 = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{2\sigma^2}(a^2 - (\mu + b)^2)\right) \sin\left(\frac{a(\mu + b)}{\sigma^2}\right)$$

Finalement, la densité de probabilité d'une variable aléatoire X définie comme étant la somme d'une variable aléatoire gaussienne de moyenne μ et de variance σ^2 et d'une variable aléatoire de Cauchy de paramètre d'échelle a et de paramètre de location b , qu'on note $X \sim \mathcal{NC}(\mu, \sigma^2, a, b)$, est définie comme suit :

$$\begin{aligned}\Phi(x) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x(x - 2\mu - 2b) + a^2 - (\mu + b)^2)\right) \\ &\times \left[\cos\left(\frac{a(\mu + b)}{\sigma^2}\right) \cos\left(\frac{xa}{\sigma^2}\right) + \sin\left(\frac{a(\mu + b)}{\sigma^2}\right) \sin\left(\frac{xa}{\sigma^2}\right) - \frac{1}{2} \exp\left(\frac{1}{2\sigma^2}(a^2 - (\mu + b)^2)\right) \right. \\ &\times \left[\exp\left(\frac{1}{2\sigma^2}(\mu + b + ia)^2\right) \operatorname{erf}\left(\frac{-a + i(x - \mu - b)}{\sqrt{2}\sigma}\right) \left(-\cos\left(\frac{xa}{\sigma^2}\right) + i \sin\left(\frac{xa}{\sigma^2}\right)\right) \right. \\ &\left. \left. - \exp\left(\frac{1}{2\sigma^2}(-\mu - b + ia)^2\right) \operatorname{erf}\left(\frac{-a + i(-x + \mu + b)}{\sqrt{2}\sigma}\right) \left(\cos\left(\frac{xa}{\sigma^2}\right) + i \sin\left(\frac{xa}{\sigma^2}\right)\right) \right] \right]\end{aligned}$$

Estimation Bayésienne non Paramétrique de Systèmes Dynamiques en Présence de Bruits Alpha-Stables

Résumé : Dans un nombre croissant d'applications, les perturbations rencontrées s'éloignent fortement des modèles classiques qui les modélisent par une gaussienne ou un mélange de gaussiennes. C'est en particulier le cas des bruits impulsifs que nous rencontrons dans plusieurs domaines, notamment celui des télécommunications. Dans ce cas, une modélisation mieux adaptée peut reposer sur les distributions α -stables. C'est dans ce cadre que s'inscrit le travail de cette thèse dont l'objectif est de concevoir de nouvelles méthodes robustes pour l'estimation conjointe état-bruit dans des environnements impulsifs. L'inférence est réalisée dans un cadre bayésien en utilisant les méthodes de Monte Carlo séquentielles. Dans un premier temps, cette problématique a été abordée dans le contexte des systèmes de transmission OFDM en supposant que les distorsions du canal sont modélisées par des distributions α -stables symétriques. Un algorithme de Monte Carlo séquentiel a été proposé pour l'estimation conjointe des symboles OFDM émis et des paramètres du bruit α -stable. Ensuite, cette problématique a été abordée dans un cadre applicatif plus large, celui des systèmes non linéaires. Une approche bayésienne non paramétrique fondée sur la modélisation du bruit α -stable par des mélanges de processus de Dirichlet a été proposée. Des filtres particuliers basés sur des densités d'importance efficaces sont développés pour l'estimation conjointe du signal et des densités de probabilité des bruits.

Mots-clés : bruit impulsif, distributions α -stables, inférence bayésienne, méthodes de Monte Carlo, filtrage particulière, estimation conjointe, systèmes OFDM, estimation non paramétrique de densité, mélange de processus de Dirichlet.

Nonparametric Bayesian Estimation of Dynamical Systems in the Presence of Alpha-Stable Noise

Abstract : In signal processing literature, noise's sources are often assumed to be Gaussian. However, in many fields the conventional Gaussian noise assumption is inadequate and can lead to the loss of resolution and/or accuracy. This is particularly the case of noise that exhibits impulsive nature. The latter is found in several areas, especially telecommunications. α -stable distributions are suitable for modeling this type of noise. In this context, the main focus of this thesis is to propose novel methods for the joint estimation of the state and the noise in impulsive environments. Inference is performed within a Bayesian framework using sequential Monte Carlo methods. First, this issue has been addressed within an OFDM transmission link assuming a symmetric α -stable model for channel distortions. For this purpose, a particle filter is proposed to include the joint estimation of the transmitted OFDM symbols and the noise parameters. Then, this problem has been tackled in the more general context of nonlinear dynamic systems. A flexible Bayesian nonparametric model based on Dirichlet Process Mixtures is introduced to model the α -stable noise. Moreover, sequential Monte Carlo filters based on efficient importance densities are implemented to perform the joint estimation of the state and the unknown measurement noise density.

Keywords : impulsive noise, α -stable distributions, Bayesian inference, Monte Carlo methods, particle filtering, joint estimation, OFDM systems, nonparametric density estimation, Dirichlet process mixture.

