



HAL
open science

Traitement de l'information en mode comptage appliqué aux détecteurs spectrométriques

Rémi Perenon

► **To cite this version:**

Rémi Perenon. Traitement de l'information en mode comptage appliqué aux détecteurs spectrométriques. Autre. Université de Grenoble, 2013. Français. NNT : 2013GRENT079 . tel-00957594

HAL Id: tel-00957594

<https://theses.hal.science/tel-00957594v1>

Submitted on 10 Mar 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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **SIPT**

Arrêté ministériel : 7 août 2006

Présentée par

PERENON Rémi

Thèse dirigée par **GRANGEAT Pierre**

et codirigée par **MOHAMMAD-DJAFARI Ali**

préparée au sein du **CEA-Leti - Département des micro-Technologies pour la Biologie et la Santé**

et de **EEATS (Electronique Electrotechnique Automatique Traitement du Signal) - Spécialité Signal Image Parole Télécoms**

Traitement de l'information en mode comptage appliqué aux détecteurs spectrométriques

Thèse soutenue publiquement le **8 Octobre 2013**,

devant le jury composée de :

M. JUTTEN Christian

Professeur à l'Université Joseph Fourier de Grenoble, Président

M. IDIER Jérôme

Directeur de Recherche au CNRS, Rapporteur

M. JOUBERT Pierre-Yves

Professeur à l'Université de Paris Sud, Rapporteur

M. LEMOINE Jérôme

Professeur à l'Université Claude Bernard de Lyon, Examineur

M. GRANGEAT Pierre

Directeur de Recherche au CEA, Directeur de thèse

M. MOHAMMAD-DJAFARI Ali

Directeur de Recherche au CNRS, Co-Directeur de thèse



Table des matières

I	Introduction	1
	Introduction générale	3
	Structure du document	3
II	Chapitres introductifs	5
1	Protéomique et spectrométrie de masse	7
1.1	La protéomique	8
1.1.1	Les protéines	8
1.1.2	Protéomique	9
1.2	La spectrométrie de masse	10
1.2.1	Analyser un mélange	10
1.2.2	Enjeux liés à la spectrométrie de masse	10
1.2.3	Analyseurs existants	11
1.2.4	Fragmentation	14
1.2.5	Quantification absolue	14
1.2.6	Traitement des données	14
2	Les systèmes électromécaniques nanométriques	17
2.1	Description des capteurs	19
2.1.1	Les MEMS et les NEMS	19
2.1.2	Les NEMS capteurs de masse	19
2.1.3	Modèle des chutes de fréquence	23
2.2	Suivi de la fréquence de résonance du capteur	27
2.2.1	Mesure de la réponse du capteur à une excitation harmonique	27
2.2.2	Boucle de contrôle pour le suivi de la fréquence de résonance	29
2.3	La spectrométrie de masse à base de NEMS	32
2.3.1	Principe général	32
2.3.2	Hypothèses prises sur le modèle physique	32
2.3.3	Le problème de la source	33
2.4	Sessions expérimentales	35
2.4.1	Sessions du mois de Décembre 2011	35
2.4.2	Sessions des mois d’Août et de Septembre 2012	35
2.4.3	Sessions du mois de Décembre 2012	37
2.4.4	Sessions du mois de Février 2013	38

III	Mise en place du cadre de travail	39
3	Modélisation du système de mesure et formalisation du problème	41
3.1	Généralités sur la modélisation	43
3.1.1	Se donner un modèle de mesure	43
3.1.2	Identifier et modéliser les grandeurs d'intérêt	44
3.1.3	Prendre en compte les incertitudes	45
3.2	Modélisation des comportements physiques	50
3.2.1	Schéma-bloc du système de mesure	50
3.2.2	Modélisation des blocs	53
3.2.3	Résumé des équations et modèle de signal	57
3.3	Expression mathématique des grandeurs d'intérêt	59
3.3.1	Modélisation de l'arrivée des molécules	59
3.3.2	Événements élémentaires et paramètres collectifs	62
3.4	Modélisation des incertitudes	64
3.4.1	Description des incertitudes	64
3.4.2	Prise en compte des perturbations par une loi normale	67
3.4.3	Synthèse	70
3.5	Détecter - quantifier - compter	74
3.5.1	Détecter	74
3.5.2	Quantifier	74
3.5.3	Compter	74
4	Approches suivies	77
4.1	L'approche problème inverse	79
4.1.1	Problèmes bien-posés et problèmes mal-posés	80
4.1.2	Méthodes d'inversion analytiques	81
4.1.3	Méthodes d'inversion par décomposition	82
4.1.4	Méthodes d'inversion algébriques	83
4.2	Le cadre bayésien	88
4.2.1	Principes généraux	88
4.2.2	Estimateurs bayésiens	90
4.2.3	Choix de modèle	91
4.2.4	Modélisation bayésienne hiérarchique	92
4.2.5	Liens entre loi <i>a priori</i> et pénalisation	93
4.2.6	Méthodes de calcul bayésien	94
4.3	Les méthodes MCMC	96
4.3.1	Principe du MCMC	96
4.3.2	L'algorithme de Metropolis-Hastings	97
4.3.3	Méthodes dérivées de Metropolis-Hastings	98
IV	Méthodologies développées	101
5	Traitement mono-mode – Déconvolution impulsionnelle	103
5.1	Cadre de travail	105
5.2	Méthodes de référence	107
5.2.1	État de l'art – la méthode publiée par Caltech	107
5.2.2	Méthodes directes adaptées	107

5.2.3	Approches problème inverse adaptées	111
5.2.4	Approche pénalisation et approche bayésienne	114
5.3	Développement d'une méthode d'inversion probabiliste reposant sur un choix de modèle	122
5.3.1	Justification de l'approche employée	122
5.3.2	Modélisation statistique des grandeurs	124
5.3.3	Estimation	127
5.4	Résultats	132
5.4.1	Définitions des critères de performance	132
5.4.2	Résultats sur données simulées	133
5.4.3	Validation de la méthode sur données réelles	138
5.4.4	Discussion	138
6	Traitement multi-mode – Détection et quantification	141
6.1	Cadre de travail	143
6.2	Méthodes de référence	145
6.2.1	État de l'art – méthodes publiées	145
6.2.2	Adaptation des méthodes de déconvolution mono-mode	146
6.2.3	Adaptation d'un modèle Bernoulli-Gaussien	146
6.3	Développement d'une méthode d'inversion probabiliste multi-mode	148
6.3.1	Justification de l'approche employée	148
6.3.2	Modélisation statistique des grandeurs	149
6.3.3	Estimation	153
6.4	Résultats	155
6.4.1	Définitions des critères de performance	155
6.4.2	Résultats sur données simulées	155
6.4.3	Validation de la méthode sur données réelles	160
6.4.4	Discussion	161
7	Traitement multi-mode – Reconstruction de spectre de masse	163
7.1	Cadre de travail	165
7.2	Méthodes de référence	167
7.2.1	État de l'art – la méthode publiée par Caltech	167
7.2.2	Méthodes connexes	167
7.3	Adaptation de la méthode de Caltech à notre algorithme	170
7.3.1	Principe	170
7.3.2	Exemple de résultat sur données expérimentales	172
7.3.3	Commentaires	173
7.4	Développement d'une méthode de reconstruction d'un spectre de raies	176
7.4.1	Cadre de travail utilisé	176
7.4.2	Modélisation statistique des grandeurs et estimation	176
7.4.3	Quelques résultats préliminaires sur signaux synthétiques	180

V	Conclusion	185
VI	Annexes	193
A	Quelques lois de probabilité	197
A.1	Loi normale	197
A.2	Loi uniforme	197
A.3	Loi Gamma	197
A.4	Loi Binomiale	198
A.5	Loi Bêta	198
B	Pré-traitements	199
B.1	Recalage des horloges	199
B.2	Suppression du bruit de secteur	200
C	Étude sur le bruit	201
C.1	Présentation des signaux de bruit et analyse	201
C.2	Matrice de covariance	209
D	Méthodes d'inversion du système d'équations non-linéaires	215
D.1	Contexte	215
D.2	Méthode de recherche d'une solution aux équations non-linéaires	215
D.3	Méthode de recherche d'une solution aux équations non-linéaires	216
E	Détails des calculs	219
E.1	Démonstration de la relation (4.42)	219
E.2	Calcul de la probabilité d'acceptation dans le cadre d'un échantillonneur de Gibbs	220
E.3	Marginalisation de l'inverse variance d'une loi normale	220
E.4	Calculs d'acceptation des algorithmes RJMCMC	221
F	Bibliographie personnelle	225
F.1	Travail portant sur la déconvolution mono-mode	225
F.2	Travail portant sur la déconvolution multi-mode	226
F.3	Travail portant sur la reconstruction de spectre	226
VII	Bibliographies	227

Table des figures

1.1	Un exemple de signal de sortie d'un dispositif de spectrométrie de masse	15
2.1	Illustration d'un NEMS capteur de masse en configuration <i>double-clamped</i>	20
2.2	Illustration d'un NEMS capteur de masse en configuration <i>cantilever</i>	21
2.3	Illustration du décalage en fréquence des résonateurs NEMS par ajout de masse .	22
2.4	Fonctions de gain variable pour une configuration <i>double-clamped</i> sur les trois premières harmoniques	25
2.5	Fonctions de gain variable pour une configuration <i>cantilever</i> sur les trois premières harmoniques	26
2.6	Illustration du système de mesure à base de NEMS comme capteur de masse en configuration <i>double-clamped</i> (à gauche) et <i>cantilever</i> (à droite)	28
2.7	Diagramme de Bode du gain du système	30
2.8	Illustration de la boucle auto-oscillante	31
2.9	Illustration de la boucle à verrouillage de phase	31
2.10	Schéma de principe de génération des nano-agrégats	34
2.11	Photographie du banc expérimental du CEA/INAC	36
2.12	Spectre de masse typique lors des sessions expérimentales de Décembre 2011 . . .	36
2.13	Spectre de masse typique lors des sessions expérimentales de Février 2013	38
3.1	Entrée et sortie du système	50
3.2	Exemple de signal sur une harmonique donnée à la sortie du dispositif	51
3.3	Modélisation en blocs du système de mesure	52
3.4	Exemples de la réponse du système à une adsorption sur le mode 1 (en haut) et le mode 2 (en bas)	55
3.5	Exemple de réponse du système et ajustement sur le mode 1 (en haut) et le mode 2 (en bas)	56
3.6	Illustration du caractère "non-entier" des instants de chutes de fréquences	57
3.7	Modélisation hiérarchique du problème	63
3.8	Illustration du caractère additif du bruit pour un système à deux modes	64
3.9	Modification des relations hiérarchiques due au caractère additif et indépendant du bruit	65
3.10	Histogramme des échantillons du signal A après pré-traitement et loi normale ajustée	68
3.11	Influence du pré-traitement sur la réponse impulsionnelle	71

3.12	Illustration du gain en temps de calcul par la méthode de calcul de vraisemblance proposée et un filtre d'ordre 10 et 2 harmoniques	73
4.1	Illustration du problème direct	79
4.2	Illustration du principe de l'inversion	79
4.3	Exemple de modèle hiérarchique	92
5.1	Modélisation hiérarchique du problème de détection	106
5.2	Illustration de la démarche « détection de rupture »	108
5.3	Illustration de la démarche « détection de pics »	110
5.4	Illustration de la pénalisation L_0	116
5.5	Illustration de la loi double-exponentielle avec $a = 1$	117
5.6	Illustration de la loi de Student-t avec $s = 1$	118
5.7	Vision hiérarchique du modèle Bernoulli-Gaussien	120
5.8	Illustration du modèle Bernoulli-Gaussien avec $r = 10$	120
5.9	Illustration du caractère mal-posé du problème	123
5.10	Structure de l'algorithme hybride RJMCMC / MCMC mono-mode pendant le temps de chauffe	128
5.11	Structure de l'algorithme hybride RJMCMC / MCMC mono-mode après le temps de chauffe	129
5.12	Courbe ROC mono-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption entiers	136
5.13	Courbe ROC mono-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption non entiers	137
5.14	Illustration de la détection mono-mode sur un signal expérimental	139
6.1	Modélisation hiérarchique du problème de détection – quantification	144
6.2	Vision hiérarchique du modèle Bernoulli-Gaussien multi-mode	147
6.3	Illustration de la structure des chutes de fréquence	150
6.4	Structure de l'algorithme hybride RJMCMC / MCMC multi-mode pendant le temps de chauffe	153
6.5	Structure de l'algorithme hybride RJMCMC / MCMC multi-mode après le temps de chauffe	154
6.6	Courbe ROC multi-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption entiers	157
6.7	Courbe ROC multi-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption non entiers	159
6.8	Illustration de la détection multi-mode sur un signal expérimental	161
7.1	Modélisation hiérarchique du problème de détection – quantification – comptage	166
7.2	Structure de l'algorithme hybride RJMCMC / MCMC en reconstruction de spectre après le temps de chauffe	170
7.3	Reconstruction d'un spectre sur données réelles	174
7.4	Structure de l'algorithme hybride RJMCMC / MCMC en reconstruction de spectre de masse parcimonieux	179
7.5	Reconstruction d'un spectre parcimonieux « sans classification »	183
7.6	Reconstruction d'un spectre parcimonieux « avec classification »	184

C.1	Illustration du signal A correspondant un bruit seul pour les modes 1 (en haut) et 2 (en bas)	202
C.2	Illustration du signal B correspondant un bruit seul pour les modes 1 (en haut) et 2 (en bas)	203
C.3	Densité Spectrale de Puissance (DSP) du bruit du mode 1 correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)	204
C.4	Densité Spectrale de Puissance (DSP) du bruit du mode 2 correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)	205
C.5	Densité Spectrale de Puissance (DSP) du bruit du mode 1 après pré-traitement et dérivation correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)	206
C.6	Densité Spectrale de Puissance (DSP) du bruit du mode 2 après pré-traitement et dérivation correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)	207
C.7	Auto-corrélation des signaux après-pré-traitement et dérivation pour les modes 1 (en haut) et 2 (en bas)	208
C.8	Estimation de la matrice de covariance du bruit pour $V = 10$	210
C.9	Estimation de la matrice de covariance du bruit pour $V = 10$ après égalisation	211
C.10	Illustration du blanchiment du bruit par $\mathbf{\Gamma}^{-\frac{1}{2}}$	214
C.11	Illustration du blanchiment du bruit par notre méthode	214
D.1	Illustration du rapport des fonctions de gain variables en fonction de la position	216

Liste des tableaux

2.1	Valeurs de gain variable pour une configuration <i>double-clamped</i> sur les trois premières harmoniques	25
2.2	Valeurs de gain variable pour une configuration <i>cantilever</i> sur les trois premières harmoniques	25
2.3	Valeurs des masses effectives sur les deux premières harmoniques pour les sessions expérimentales des mois d’Août et de Septembre 2012	37
2.4	Valeurs des masses effectives sur les deux premières harmoniques pour les sessions expérimentales du mois de Décembre 2012	37
3.1	Paramètres des lois normales pour l’ajustement de la figure 3.5	56
3.2	Comparaison des approches processus ponctuels marqués (PPM) et processus à temps discret (PTD)	61
3.3	Paramètres des lois normales pour l’ajustement de la figure 3.10	67
3.4	Coefficients des filtres blanchisseurs	72
4.1	Quelques exemples de lois <i>a priori</i> conjuguées	89
5.1	Valeurs des paramètres de simulation mono-mode	133
5.2	Valeurs des paramètres de l’algorithme pour le traitement des données simulées mono-mode	134
5.3	Valeurs des paramètres des lois <i>a priori</i> pour le traitement des données simulées mono-mode	134
5.4	Temps de calcul des différents algorithmes en mono-mode avec instants d’adsorption entiers	135
5.5	Temps de calcul des différents algorithmes en mono-mode avec instants d’adsorption non-entiers	136
5.6	Valeurs des paramètres du modèle direct pour les données réelles mono-mode	138
5.7	Coefficients des filtres blanchisseurs pour le traitement des données réelles mono-mode	138
5.8	Valeurs des paramètres de l’algorithme pour le traitement des données réelles mono-mode	138
5.9	Valeurs des paramètres des lois <i>a priori</i> pour le traitement des données réelles mono-mode	138
6.1	Valeurs des paramètres de simulation multi-mode	156

6.2	Valeurs des paramètres de l'algorithme pour le traitement des données simulées multi-mode	156
6.3	Valeurs des paramètres des lois <i>a priori</i> pour le traitement des données simulées multi-mode	156
6.4	Temps de calcul des différents algorithmes en multi-mode avec instants d'adsorption entiers	158
6.5	Temps de calcul des différents algorithmes en multi-mode avec instants d'adsorption non-entiers	159
6.6	Valeurs des paramètres du modèle direct pour les données réelles multi-mode . .	160
6.7	Coefficients des filtres blanchisseurs pour le traitement des données réelles multi-mode	160
6.8	Valeurs des paramètres de l'algorithme pour le traitement des données réelles multi-mode	160
6.9	Valeurs des paramètres des lois <i>a priori</i> pour le traitement des données réelles multi-mode	161
6.10	Diamètre estimé des agrégats détectés sur données expérimentales	161
7.1	Valeurs des paramètres du modèle direct pour les données réelles en reconstruction de spectre	172
7.2	Coefficients des filtres blanchisseurs pour le traitement des données réelles en reconstruction de spectre	173
7.3	Valeurs des paramètres de l'algorithme pour le traitement des données réelles en reconstruction de spectre	173
7.4	Valeurs des paramètres des lois <i>a priori</i> pour le traitement des données réelles en reconstruction de spectre	173
7.5	Valeurs des paramètres de simulation en reconstruction de spectre de masse . . .	181
7.6	Valeurs des paramètres de l'algorithme de reconstruction de spectre de masse sur données simulées	181
7.7	Valeurs des paramètres de l'algorithme de reconstruction de spectre de masse parcimonieux sur données simulées	182
7.8	Valeurs des paramètres des lois <i>a priori</i> pour les algorithmes de reconstruction de spectre de masse sur données simulées	182

Remerciements

A l'heure de commencer ce document, c'est-à-dire pour moi de conclure une belle aventure de trois années, je ne peux m'empêcher de regarder par-dessus mon épaule, l'œil humide, et de mesurer non pas le chemin accompli, mais l'aide qui m'a été apportée.

Dans un premier temps, j'adresse mes plus sincères remerciements aux membres de mon jury. Tout d'abord, merci à M. Jutten, qui a accepté de le rôle de Président, et avec qui j'ai pu avoir quelques conversations très intéressantes en conférence.

Je n'oublie pas mes deux rapporteurs, MM. Idier et Joubert, qui ont été d'une disponibilité remarquable pour des discussions aussi bien avant la remise du rapport qu'après celle-ci. Leur écoute, leurs conseils et leurs remarques m'ont permis de rendre (je l'espère) ce document plus clair qu'il ne l'était au départ.

Enfin, merci à M. Lemoine d'avoir accepté d'examiner ce travail.

Comment ne pas associer à ce travail mes deux directeurs de thèse ? Je me souviens de mon premier « pas » dans cette grande aventure lors d'un entretien dans le bureau d'Ali, au cours duquel ce dernier m'avait interrogé au tableau ! S'en suivit quelques jours plus tard une longue entrevue avec Pierre qui finit de me convaincre de me jeter à l'eau.

Je ne remercierai jamais assez Pierre d'avoir su me mettre en confiance au sein de l'équipe et au sein du sujet de thèse. Merci de ta gestion quotidienne, des nombreux points hebdomadaires qui m'ont permis de ne pas couler, de ton soutien constant qui m'a beaucoup aidé dans les nombreux moments de doute (surtout sur la fin !). Merci aussi pour tous ces moments informels, lors des repas, des fins de journée ou des voyages en train.

Je retiendrai du temps passé avec Ali sa disponibilité par téléphone, Webex, ou lorsque je montais au L2S, ainsi que les cours du master ATSI et les discussions sur son fameux tableau noir. Je retiendrai aussi le temps que nous avons pu partager en conférence, que ce soit en amphithéâtre ou en dehors. Ali, merci infiniment pour ces connaissances que vous avez pu me transmettre malgré la distance qui faisait parfois obstacle.

Bien que ce travail soit présenté à titre personnel, de nombreuses personnes m'ont entouré tout au long de ces trois années. Tout d'abord, mes collègues de l'équipe PROTIS, Pierre, Laurent, Pascal, mais aussi les stagiaires qui nous ont aidé (Abrar, Abdoulaye). Merci aussi à Lamine d'avoir fait avec moi un bout de chemin sur les méthodes par décomposition.

Un grand merci à nos collègues physiciens, du Leti-DCOS premièrement, notamment Eric Sage avec lequel j'ai beaucoup travaillé (et beaucoup apprécié travailler), et qui m'a donné ce qui s'apparente à un graal : des données expérimentales ! Merci aussi à Laurent Duraffourg et à

Sébastien Hentz pour nous avoir fréquemment associés à leurs discussions, très riches en informations. Merci également à nos collègues du CEA/INAC Ariel Brenac et Robert Morel, pour leur aide sur les données expérimentales. Je remercie aussi Michael Roukes et Scott Kelber de Caltech. Un grand merci également à Eric Colinet pour son aide tout au long de la thèse.

J'associe aussi des personnes avec lesquelles j'ai pu avoir des discussions m'ayant ouvert l'esprit sur d'autres applications de mon travail : Pierre Puget, Thomas Ernst, Vincent Gouttenoire, Vincent Agache, Frédéric Bottausci, Guillaume Montemont, Fanny Marticke ou Caroline Paulus.

Je voudrais aussi avoir un mot pour les personnes autour de moi qui ont assuré des tâches administratives et/ou hiérarchiques, qui m'ont permis de travailler dans de bonnes conditions matérielles. Je pense à mes chefs de département, de service ou de laboratoire. Je pense aussi à tout le personnel administratif au CEA-Leti comme au L2S envers lequel je suis redevable.

Bien qu'il soit toujours injuste de faire ressortir quelques noms parmi cet ensemble de personnes, je voudrais remercier Michelle, qui jouit désormais de très grandes vacances bien méritées à l'Alpe du Grand Serre, pour son aide précieuse dans toutes les démarches administratives. Merci aussi à Laurence qui a su prendre efficacement le relais, et qui m'a été d'une aide inestimable dans le *rush* final de ce travail.

Je remercie l'ensemble de l'équipe du GPI du L2S, et notamment Thomas pour son explication de texte (salvatrice) sur les métiers de la recherche, Nicolas pour sa grande gentillesse lors de mes passages à Gif-sur-Yvette, et les doctorants du GPI Mircea, Leila, Thomas, Doriano. . .

Je ne peux pas passer à côté de l'ensemble des équipes du LE2S, avec qui j'ai partagé de très bons moments. Merci Régis de m'avoir accueilli dans ton laboratoire et de m'avoir reçu chaque fois que j'en avais le besoin.

Merci à Vence, pour ces interminables discussions dans NOTRE bureau. On s'en est dit des choses. A tel point qu'aujourd'hui tu découvres le Pays Voironnais (j'ai dû te convaincre) ! Merci à Laurent pour ces grandes discussions si enrichissantes et pour m'avoir épaulé lors de cette difficile étape qu'était la rédaction. Merci à Pascal pour tous ces moments debout, sur une chaise ou sur un vélo à parler bayésien, football, paternité ou saucisses. Merci à Abbas « c'est la vie » pour sa malchance. Merci à Raphaëlle pour organiser les sorties jardin du thé et autres. J'ai une pensée nostalgique envers Alain et Yvon avec qui j'ai pu partager provisoirement un bureau, Alain m'initiant aux joies du dessin technique en toutes circonstances, et Yvon m'informant de l'état du réacteur 3 et de la basket droite de Booba. Merci à Florent pour ses délires, ses BD et ses hamburgers maisons. Merci à Maeva avec laquelle il ne nous a manqué que trois bouts de corde et deux planches de bois pour refaire le monde. Merci à Jean-François qui m'aura beaucoup appris sur l'électronique et conforté dans mon approche du travail. Merci à Stéphane, ses voyages en train et ses goûts musicaux. Merci à Alain et ses « c'est paru cette semaine dans le Canard ». Merci à Pascale pour son caractère trempé et ses chocolats (un mélange explosif !). Merci à Jean avec qui j'ai partagé un « bout d'*open-space* » et quelques outils. Merci à Michel, son ouzo et sa voix portante. Merci à Alexandre pour les grandes discussions sur la tomographie d'impédance. Enfin, merci à Anne-Laure, Florence, Amélie, Sabine, Ilham, Virginie, Gaëlle, Mickaël, Guillaume, William, Florent, Aloïs, Bilal, Joao, Selim, Aurélien, Alexandre, Christophe, Sadok et j'en oublie sûrement.

Enfin, je garde une petite place en bas de page mais une grande place dans mon esprit pour ceux qui me sont chers. Ils sont trop nombreux et « dépasseraient le cadre de ce document ». Merci à Roland et Brigitte pour tout ce qu'ils m'ont apporté depuis le début. Merci à ma femme et à ma fille pour tout l'amour que nous partageons ensemble à la maison.

Notations

Conventions mathématiques

Dans ce document, nous utiliserons les conventions mathématiques suivantes.

- \mathbb{R} : l'ensemble des nombres réels ;
- \mathbb{R}^+ : l'ensemble des nombres réels positifs ou nuls ;
- \mathbb{N} : l'ensemble des entiers naturels ;
- \mathbb{N}^* : l'ensemble des entiers naturels ;
- $\{m, M\}$: l'ensemble des entiers naturels compris entre m et M ;
- $[m, M]$: l'ensemble des réels compris entre m et M ;
- \mathbf{x} : un vecteur temporel discrétisé ;
- \mathbf{x}_i : la i -ème composante du vecteur \mathbf{x} ;
- \mathbf{X} : une matrice ;
- T : l'opérateur de transposition d'une matrice ;
- $tr(\cdot)$: la trace d'une matrice ;
- \mathbf{I} : la matrice identité ;
- $\frac{\mathbf{x}_1}{\mathbf{x}_2}$ le vecteur issu de la division terme à terme du vecteur \mathbf{x}_1 par le vecteur \mathbf{x}_2 ;
- $i_{\mathbb{C}}$: l'unité imaginaire telle que $i_{\mathbb{C}}^2 = -1$;
- \star : le produit de convolution.

Fonctions mathématiques

Nous aurons recours aux entités mathématiques suivantes :

- $\exp(x)$: la fonction exponentielle ;
- $\log(x)$: la fonction logarithme ;
- $\delta(t)$: la distribution de Dirac ;

$$\begin{aligned}\delta(t) &= 1 \text{ si } t = 0 \\ &= 0 \text{ sinon}\end{aligned}$$

- \diamond : opérateur linéaire et distributif sur l'addition, qui permet la multiplication de distributions

de Dirac comme suit :

$$\begin{aligned}\delta(t - t_1) \diamond \delta(t - t_2) &= 1 \text{ si } t = t_1 = t_2 \\ &= 0 \text{ sinon}\end{aligned}$$

- de la même manière, \diamond : opérateur de la multiplication terme à terme sur des vecteurs ;
- $f_{\text{Heaviside}}(t)$: la fonction de Heaviside, ou fonction échelon ;

$$\begin{aligned}f_{\text{Heaviside}}(t) &= 1 \text{ si } t \geq 0 \\ &= 0 \text{ sinon}\end{aligned}$$

- $\text{III}_{\Delta_t}(t)$: le peigne de Dirac de pas Δ_t :

$$\text{III}_{\Delta_t}(t) = \sum_{n=-\infty}^{\infty} \delta(t - n \Delta_t)$$

- $TF(f(t))(f)$: la transformée de Fourier

$$TF(f(t))(f) = \int f(t) \exp(-i_{\mathbb{C}} f t) dt$$

- $r_x(\tau)$: l'autocorrélation du signal temporel $x(t)$

$$r_x(\tau) = \mathbb{E}_{p(x(t))} [x(t) x(t - \tau)]$$

Unités

Les unités utilisés seront :

- la seconde s et la milliseconde ms ;
- le mètre m et le nanomètre nm ;
- le Dalton Da et le kiloDalton kDa ;
- le hertz Hz, le kilohertz kHz et le Mégahertz MHz.

Conventions statistiques

Enfin, nous utiliserons plusieurs concepts afférents aux statistiques et nous noterons :

- $\mathbf{x} \sim p(\mathbf{x})$: \mathbf{x} suit une loi de probabilité $p(\mathbf{x})$;
- $\mathbb{E}_{p(\mathbf{x})} [\mathbf{x}]$: l'espérance de \mathbf{x} sous la loi $p(\mathbf{x})$;
- $\mathbf{x} \sim \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma})$: \mathbf{x} suit une loi normale multivariée, de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Sigma}$.
- $\mathcal{U}_{[m, M]}(x)$: x suit une loi uniforme entre m et M .
- $x \sim \mathcal{Gam}(x | k, \theta)$: x suit une loi Gamma, de paramètre de forme k et de paramètre d'échelle θ ;
- $x \sim \mathcal{Beta}(x | a, b)$: x suit une loi bêta, de paramètres a et b ;
- $x \sim \mathcal{B}(x | \pi, T)$: x suit une loi de Binomiale, de paramètre de probabilité de succès π et de nombre d'épreuves T .

Les détails sur les lois de probabilité utilisées sont donnés en annexe A.

Première partie

Introduction

Introduction générale

L'étude des protéines présente un intérêt primordial en biologie. En effet, les protéines traduisent à la fois au niveau cellulaire l'expression d'un patrimoine génétique et l'influence de l'environnement. Elles sont accessibles dans de nombreux fluides organiques (sang, sérum, urine . . .) [WBMT00] au contact des cellules. De ce fait, elles sont un sujet d'étude privilégié pour l'étude de certaines maladies associées à un dysfonctionnement cellulaire, notamment le cancer [WLP03, PZK⁺02, PAH⁺02, QL12, HT11]. Un outil d'analyse pour l'étude des protéines est le spectromètre de masse [ABE⁺08, AM03, DA06]. Cet appareil a pour but d'estimer le profil en masse d'une solution donnée. Les spectromètres de masse actuels reposent sur différentes technologies qui ont en commun la présence de deux étapes de séparation et de détection pour l'analyse d'un flux de particules chargées, à savoir des ions.

D'autre part, l'électronique connaît elle aussi de profondes avancées, notamment au travers des nanotechnologies. L'idée est de produire à large échelle des composants d'une taille infime qui présenteront des fonctionnalités bien définies, [CBB12, TCL12, LLP⁺12, CYP⁺12]. Dans le panorama des nanotechnologies, les systèmes nanométriques électromécaniques (ou NEMS pour *Nano ElectroMechanical Systems*) permettent justement la détection de quantités infimes de masse [YCF⁺06, LSP⁺10, DLBK07, EHR04]. L'idée des dispositifs que nous allons étudier dans cette thèse est donc d'utiliser ces NEMS comme une nouvelle technologie de spectrométrie de masse [NHH⁺09]. Ceci conduit à un système dont le principe ne repose plus sur la mesure d'un flux de particules chargées mais sur la détection de molécules uniques et l'estimation de leur masses respectives. Un tel fonctionnement peut être qualifié de « mode comptage ». L'intérêt est notamment une réduction du bruit de mesure et une amélioration de la sensibilité.

Cette thèse a été préparée au CEA-Leti, au sein de l'équipe PROTIS (PROtéomique et Traitement de l'Information pour la Santé) en collaboration avec le Laboratoire des Signaux et Systèmes (UMR 8506 CNRS – Supélec – Paris Sud). L'objectif est de proposer des méthodes d'estimation de profils de masse adaptées aux nouveaux capteurs que sont les NEMS. Pour cela, nous nous plaçons dans le cadre des approches problèmes inverses et d'inférence bayésienne. Nous proposons de modéliser le système de mesure ainsi que le lien entre les grandeurs observables et les inconnues par un modèle hiérarchique. Afin d'estimer le profil de masse, nous proposons l'utilisation d'une méthode d'inversion statistique reposant sur le cadre de travail bayésien.

Structure du document

Ce manuscrit de thèse a pour objectif de faire le point sur la problématique du traitement de signal appliqué à la spectrométrie de masse à base de NEMS. Pour cela, nous aurons besoin d'introduire les technologies et les applications mises en jeu ici, de définir le cadre d'étude et de clarifier les choix techniques qui s'ouvrent à nous pour enfin développer des méthodes et les mettre en relation avec les méthodes de l'état de l'art.

Dans cette optique, la partie II est consacrée à l'introduction de la problématique par les prismes applicatif et technologique. Cette partie regroupe les deux premiers chapitres, au cours desquels il s'agira notamment d'introduire les deux notions suivantes.

– *Chapitre 1 : Protéomique et spectrométrie de masse.* Ce chapitre introduit l'analyse protéomique

et la problématique de la spectrométrie de masse. Nous décrivons également les principaux analyseurs de spectre de masse actuels.

- *Chapitre 2 : Les systèmes électromécaniques nanométriques.* Ce chapitre présente les capteurs utilisés dans le cadre de ce travail. Le fonctionnement physique est notamment exposé.

La partie III, quant à elle, sera consacrée à la mise en place du cadre de travail pour le traitement des signaux.

- *Chapitre 3 : Modélisation du système de mesure et formalisation du problème.* Ce chapitre présente la démarche que nous avons employée pour modéliser le problème direct ainsi que la formulation probabiliste des incertitudes.
- *Chapitre 4 : Approches suivies.* Ce chapitre introduit les outils auxquels nous avons fait appel et qui ont été nécessaires pour résoudre le problème.

Le cadre de travail étant posé, nous exposerons dans la partie IV les contributions effectuées dans le cadre de ce travail de thèse.

- *Chapitre 5 : Traitement mono-mode – Déconvolution impulsionnelle.* Ce chapitre traite d'une sous-partie du problème qui est celle de la détection d'événements temporels. Le problème est mis sous la forme de la déconvolution impulsionnelle.
- *Chapitre 6 : Traitement multi-mode – Détection et quantification.* Ce chapitre étend le précédent à la détection sur plusieurs modes simultanément. Ceci permet notamment de quantifier des valeurs de masses.
- *Chapitre 7 : Traitement multi-mode – Reconstruction de spectre de masse.* Ce chapitre est à nouveau une extension du précédent au cours duquel nous allons voir comment estimer les paramètres collectifs du mélange, c'est-à-dire reconstruire le spectre de masse de la solution analysée.

La partie V rassemblera nos conclusions et les perspectives que nous proposons pour la poursuite de ces travaux et leur valorisation.

Les annexes se trouvent en partie VI. Nous présenterons d'abord les lois de probabilité utilisées dans ce document, puis nous préciserons ensuite les pré-traitements de données utilisés. Nous présenterons ensuite une étude sur le bruit puis une étude particulière de quantification de masse, ces deux parties n'ayant pas leur place dans le corps du document. Enfin, nous détaillerons certains calculs puis nous présenterons la bibliographie résultant de ce travail de thèse.

La partie VII sera constituée de la bibliographie, que nous avons choisi de regrouper par chapitres.

Deuxième partie

Chapitres introductifs

Protéomique et spectrométrie de masse

Dans ce chapitre, nous abordons deux sujets. Le premier est celui de la protéomique, domaine de la biologie moléculaire qui utilise l'ensemble des technologies et des méthodes que nous allons voir par la suite. Nous aborderons ensuite la spectrométrie de masse, qui est un outil fréquemment utilisé en protéomique, et qui constitue en quelque sorte un modèle applicatif de notre travail.

Sommaire de ce chapitre

1.1	La protéomique	8
1.1.1	Les protéines	8
1.1.1.1	Protéines, peptides, acides aminés	8
1.1.1.2	Synthèse des protéines	8
1.1.2	Protéomique	9
1.1.2.1	Les protéines, expression des gènes et de l'environnement	9
1.1.2.2	Notion de biomarqueur	9
1.2	La spectrométrie de masse	10
1.2.1	Analyser un mélange	10
1.2.2	Enjeux liés à la spectrométrie de masse	10
1.2.3	Analyseurs existants	11
1.2.3.1	Electrophorèse et chromatographie	11
1.2.3.2	Source d'ions	12
1.2.3.3	Analyseur	12
	Analyseurs à sélection	13
	Analyseurs à rotation	13
	Analyseurs à séparation spatiale ou temporelle	13
1.2.3.4	Détecteur	13
1.2.4	Fragmentation	14
1.2.5	Quantification absolue	14
1.2.6	Traitement des données	14

1.1 La protéomique

1.1.1 Les protéines

1.1.1.1 Protéines, peptides, acides aminés

Les protéines sont de grosses molécules présentes dans les organismes vivants [WBMT00] [ML99, Ch. 2].

La structure des protéines peut être décrite de manière hiérarchisée. En effet, les protéines sont avant tout composées d'un ou plusieurs peptides. Un peptide, quant à lui, est une chaîne d'acides aminés reliés entre eux par des liaisons peptidiques. Les liaisons peptidiques sont des liaisons covalentes, c'est-à-dire avec mise en commun d'électrons, entre deux acides aminés, et plus précisément entre une fonction carboxyle d'un acide aminé et une fonction amine d'un autre acide aminé.

Il existe 22 acides aminés mis en jeu dans la synthèse des protéines. Si l'on se restreint à l'Homme, on dénombre 20 acides aminés différents pour composer l'ensemble des protéines de l'organisme. Ces acides aminés sont référencés dans une nomenclature et désignés chacun par un « code », c'est-à-dire une lettre issue de l'alphabet romain. Un récapitulatif des acides aminés ainsi que leur structure chimique peut se trouver dans [ML99, p. 1132].

Là où le peptide désigne la chaîne d'acides aminés, la protéine est formée d'un ensemble de peptides liés et repliés de sorte à donner une forme tridimensionnelle à la molécule. On distingue généralement quatre niveaux de structure pour décrire une protéine :

- structure primaire qui correspond à la structure « à plat » de la protéine, c'est-à-dire à la séquence d'acides aminés ;
- structure secondaire qui correspond aux repliements locaux de la chaîne d'acides aminés, favorisés par les liaisons hydrogènes entre les peptides ;
- structure tertiaire qui correspond aux repliements tridimensionnels de la chaîne peptidique, favorisés par les liaisons covalentes (ponts disulfures notamment) ou non-covalentes ;
- la structure quaternaire qui résulte de l'assemblage de plusieurs sous-unités.

Aussi complexe et riche que leurs structures, les fonctions remplies par les protéines au sein de l'organisme sont très diverses. On pourra citer les enzymes, les récepteurs membranaires, le transport moléculaire, les protéines structurales, la nutrition, le système immunitaire [ML99, p. 51, Tab. 2.3]. Le collagène, par exemple, est une protéine fibreuse permettant de donner aux tissus une résistance à l'étirement. Autres exemples, l'actine et la myosine permettent la contraction des cellules musculaires. Citons encore l'albumine qui permet de réguler le pH sanguin.

1.1.1.2 Synthèse des protéines

Les protéines sont synthétisées par un organisme vivant à partir de deux phénomènes, la transcription et la traduction.

La transcription consiste à transformer des brins d'Acide DésoxyriboNucléique (ADN) en séquences

d'Acide RiboNucléique (ARN). L'ARN est synthétisé à partir de la double hélice d'ADN dans le noyau d'une cellule. Une fois transcrit, l'ARN subit un ensemble de modifications dites post-transcriptionnelles. Notamment, l'ARN est épissé, c'est-à-dire découpé. Un brin d'ADN peut donner plusieurs sortes d'ARN, cette variabilité étant notamment issue de l'étape d'épissage.

La traduction consiste à transformer l'ARN en protéine dans le cytoplasme d'une cellule. Cette opération est effectuée par un ribosome. La protéine va être synthétisée avec un acide aminé par acide aminé. La protéine peut subir des modifications dites post-traductionnelles. Un exemple classique de modification post-traductionnelle est l'ajout d'un groupe fonctionnel, comme un groupe phosphate dans le cadre d'une phosphorylation.

1.1.2 Protéomique

1.1.2.1 Les protéines, expression des gènes et de l'environnement

L'ensemble des protéines d'un organisme et de leurs interactions est dénommé « protéome », et la science visant à étudier le protéome est désignée par « protéomique ».

Nous avons vu précédemment que les protéines sont synthétisées à partir de l'ADN (le patrimoine génétique) d'un organisme. Mais elles subissent également tout un ensemble de modifications au cours de leur synthèse. Certaines modifications peuvent être induites sous l'influence de l'environnement de la cellule.

L'intérêt d'étudier le protéome (les protéines) plutôt que le génome (l'ADN) est que le protéome a vocation à évoluer tout au long de l'existence de l'individu là où le patrimoine génétique est fixé dès la naissance. Ainsi, le protéome va traduire à la fois l'expression des gènes mais aussi l'influence de l'environnement sur l'organisme. De ce fait, certaines maladies non génétiques vont avoir une influence sur le protéome alors qu'elles n'affectent en rien le génome. Le cancer en est un exemple particulièrement courant [PIAH⁺02, AQD⁺02, HPIM⁺03]. Il est également possible de détecter une infection virale en détectant les protéines générées au cours de l'infection [KTM98].

1.1.2.2 Notion de biomarqueur

Lorsqu'elle recherche à détecter une maladie chez un patient, à caractériser une maladie ou encore à étudier les effets d'un traitement sur une maladie, l'analyse biologique fait souvent appel à la notion de biomarqueur. Le biomarqueur est un élément biologique pouvant être détecté et/ou quantifié et qui est caractéristique d'un phénomène interne à l'organisme, comme une pathologie, par exemple.

L'intérêt de la protéomique est de pouvoir donner une vue d'ensemble du protéome et donc de sélectionner un ou plusieurs biomarqueurs, pour caractériser une pathologie donnée, par exemple. Nous pouvons donner l'exemple des protéines AFP, CEA, PSA ou encore CA125 utilisées comme biomarqueurs dans la détection du cancer [Str08, Ch. 1].

1.2 La spectrométrie de masse

1.2.1 Analyser un mélange

Afin d'analyser le contenu en protéines d'un mélange, trois grandes catégories de techniques existent.

- La méthode immuno-enzymatique (ou ELISA pour *Enzyme-Linked ImmunoSorbent Assay*) consiste à développer des anticorps spécifiques de la protéine ciblée, l'association protéine-anticorps se traduisant par un phénomène observable, comme une coloration du mélange par exemple, ou une fluorescence après excitation par une source lumineuse ou infra-rouge [EP71].
- Les techniques de séparation des composés telles que la chromatographie ou l'électrophorèse (sur gel par exemple) consistent à séparer les protéines selon un critère d'affinité puis à détecter ces protéines [SVJ87].
- La spectrométrie de masse est une technique visant à ioniser les protéines, à les séparer selon leur rapport masse sur charge puis à détecter les ions afin de pouvoir estimer le profil en masse (ou spectre de masse) d'une solution donnée. Elle peut être combinée avec une chromatographie ou une électrophorèse.

Par la suite, nous nous intéresserons aux techniques de spectrométrie de masse.

1.2.2 Enjeux liés à la spectrométrie de masse

La chaîne d'analyse liée à la spectrométrie de masse comporte de nombreuses étapes qui adressent plusieurs domaines scientifiques. Pour schématiser, nous pouvons en dresser une liste restreinte :

- Les échantillons à analyser sont prélevés (du patient, dans le cas d'une analyse biologique) dans des conditions suffisamment maîtrisées. Cette étape nécessite des compétences de médecine et de biologie.
- Les échantillons sont pré-traités (on parle d'étape pré-analytique), par exemple par centrifugation et capture par colonne d'affinité, afin de les purifier et de les concentrer. Aussi, l'échantillon peut subir plusieurs étapes de traitements chimiques afin d'éliminer certains composés parasites (comme les lipides ou les polysaccharides, par exemple), ou encore certaines protéines présentes en grande quantité et inintéressantes pour l'étude considérée (comme l'albumine, par exemple). Aussi, certains composés de concentration connue peuvent être ajoutés au mélange afin de permettre une quantification absolue de la composition du mélange (voir section 1.2.5). Enfin, lors de cette étape, les protéines sont « digérées », c'est-à-dire fragmentées en peptides. Cette étape met en œuvre des compétences de chimie, de biologie, de microfluidique et d'automatisme.
- Les échantillons passent dans l'appareil d'analyse (on parle d'étape analytique), ou spectromètre de masse (que nous détaillerons dans la prochaine section 1.2.3) qui peut comporter, en plus de l'étape d'estimation de la masse des molécules à proprement parler, des étapes de séparation du mélange, telles qu'une électrophorèse ou le passage du mélange dans une colonne de chromatographie. Ce point nécessite des compétences en chimie, physique, électronique, électromécanique ou encore microfluidique.
- Les données brutes issues de l'appareil d'analyse sont traitées, afin d'estimer des quantités

d'intérêt. Ceci peut revenir, par exemple, à quantifier la concentration d'une protéine en particulier, ou encore à découvrir les différentes protéines présentes dans le mélange analysé, voire à sélectionner des biomarqueurs d'une certaine pathologie. Cette étape met en jeu des compétences de traitement de l'information et de statistiques qui ne peuvent être décorrélées des réalités biologiques, physiques et chimiques évoquées ci-dessus. L'équipe dans laquelle cette thèse a été effectuée, l'équipe PROTIS (PROtéomique et Traitement de l'Information pour la Santé) du CEA-Leti, axe ses recherches sur ce dernier point [Str08, SGP⁺07, Sza13, SGG11].

Les difficultés et points potentiels d'amélioration associés aux techniques de mesure par spectrométrie de masse interviennent à ces différents niveaux de la chaîne d'analyse. Sans être exhaustif, il est possible de citer :

- la maîtrise de la qualité de l'échantillon extrait, que ce soit au niveau de l'état du patient (nutrition par exemple) ou au niveau de la qualité du système de prélèvement ;
- la capacité à éliminer les composés parasites sans pour autant ajouter des éléments perturbants au mélange, et assurer une digestion maîtrisée et reproductible ;
- les capacités de sensibilité, de bruit ajouté, de gamme de fonctionnement ou encore de résolution de l'appareil utilisé ;
- la précision et la robustesse de l'estimation, ou encore la capacité à extraire une liste courte et discriminante de protéines en tant que biomarqueurs.

1.2.3 Analyseurs existants

La chaîne d'analyse incluant le spectromètre de masse décrit ci-dessus a pour objectif de fournir une image de la composition en masse d'un mélange injecté en entrée. Il est possible de découper la partie analytique de cette chaîne d'analyse en plusieurs sous-fonctions mises en cascade.

1. La première fonction, optionnelle, est celle de la pré-séparation des composés présents dans le mélange. Les deux principales techniques employées en analyse biologique sont l'électrophorèse et la chromatographie.
2. La seconde fonction concerne la transformation des composés en éléments chimiques analysables, en l'occurrence des ions. Cette fonction est appelée source d'ions.
3. La troisième fonction concerne les éléments de séparation, sélection ou fragmentation. L'objectif ici est de séparer les ions en fonction de leur rapport masse sur charge.
4. La quatrième et dernière fonction est celle du détecteur. Le détecteur permet de transformer le flux d'ions en un flux d'électrons.

Les fonctions 2, 3 et 4 sont réunies à l'intérieur du spectromètre de masse.

1.2.3.1 Electrophorèse et chromatographie

L'électrophorèse et la chromatographie sont deux techniques de séparation d'éléments chimiques de différentes natures.

L'électrophorèse base son principe de séparation sur le fait que les composés chargés (ions) ne réagissent pas de la même manière à l'application d'un champ électrique : les vitesses de migration étant dépendantes de la nature du composé, il y a séparation du mélange.

La chromatographie, quant à elle, consiste à entraîner les composés à analyser par une phase en mouvement, dite phase mobile, en contact avec une phase dite stationnaire qui recouvre les parois de la colonne. La vitesse de déplacement des composés est fonction de l'affinité entre les éléments du composé et la phase stationnaire. Cette affinité repose sur des phénomènes physiques comme les forces d'attraction ou de répulsion hydrophiles ou hydrophobes.

En association avec la spectrométrie de masse, un choix souvent employé est celui de la chromatographie à phase mobile liquide (LC pour *Liquid Chromatography*).

1.2.3.2 Source d'ions

La source a pour objectif de transformer les composés en ions. En effet, les analyseurs de masse existants séparent des gaz de composés chargés, car la technologie utilise justement le fait que ces éléments ne sont pas neutres électriquement pour les séparer.

Or, à l'issue de la phase de pré-séparation, dans le cas d'une chromatographie, les composés sont électriquement neutres, et pas forcément sous phase gazeuse ; dans le cas d'une électrophorèse, les composés ne sont pas sous phase gazeuse non plus.

Il est nécessaire de transformer le mélange en un gaz d'ion. Ceci peut se faire selon plusieurs techniques. Les deux plus courantes sont l'électronébuliseur ou électrospray et la Désorption-Ionisation Laser Assistée par Matrice (ou MALDI pour *Matrix Assisted Laser Desorption Ionisation*).

- L'électrospray consiste à soumettre le mélange à analyser (liquide stocké dans un capillaire) à un fort champ électrique. Les forces sur les charges dans le liquide sont telles que le liquide explose en micro-gouttelettes chargées. Le liquide s'évapore ensuite et les molécules chargées sont injectées dans le spectromètre de masse.
- Le MALDI consiste à émettre un rayon laser sur le mélange à analyser (déposé sur une matrice). L'énergie du laser suffit à désorber et ioniser les composés qui sont injectés dans le spectromètre de masse.

1.2.3.3 Analyseur

L'analyseur a pour objectif de séparer les composés injectés sous forme de gaz d'ions, afin de permettre l'identification et la quantification de ces derniers.

A ce dessein, l'analyseur sépare les ions du flux incident en fonction de leur rapport masse sur charge. Ceci peut s'effectuer grâce à différentes technologies :

1. les systèmes à sélection par quadripôle ou trappe à ions ;
2. les systèmes basés sur une transformée de Fourier tels que l'analyseur à résonance cyclotronique ou l'Orbitrap ;
3. les systèmes à séparation spatiale ou temporelle tels que le temps de vol.

Ces différents systèmes réalisent une séparation du flux incident en fonction du rapport masse sur charge. Ce flux séparé est ensuite directement détecté.

Analyseurs à sélection Les analyseurs à sélection ont la capacité de pouvoir filtrer les ions présentant un rapport masse sur charge situé dans une certaine gamme. La thèse de Grégory Strubel donne de nombreuses informations sur le sujet [Str08, Ch. 2].

Ainsi, le quadripôle consiste à faire passer les ions dans une cavité dont les bords sont électriquement chargés. Seuls certains ions ont une trajectoire stable dans cette cavité et la traversent.

La trappe à ion, elle, contient initialement l'ensemble des ions. Certains ions ont une trajectoire résonante à l'intérieur de la trappe et en sont éjectés suivant leur rapport masse sur charge, en fonction de la tension appliquée.

Pour ces technologies, l'analyse consiste généralement à faire varier la fréquence d'excitation du système afin d'avoir un balayage des gammes de rapport masse sur charge présentant une trajectoire stable (pour le quadripôle) ou résonante (pour la trappe à ions).

Analyseurs à rotation Les analyseurs à rotation consistent à faire tourner les ions dans une cavité avec une période de rotation dépendant du rapport masse sur charge de ces ions. En détectant le mouvement circulaire des ions dans la cavité et en appliquant à cette mesure une transformée de Fourier, il est possible d'obtenir un spectre relié au spectre de masse.

Les deux technologies associées sont l'analyseur à résonance cyclotronique (ou FT-ICR pour *Fourier Transform Ion Cyclotron Resonance*) et l'Orbitrap.

Analyseurs à séparation spatiale ou temporelle Les analyseurs à séparation spatiale ou temporelle visent, à un instant donné, à séparer dans l'espace ou dans le temps les ions en fonction de leur rapport masse sur charge.

Un premier exemple de système à séparation spatiale est la déflexion magnétique : les ions sont forcés de prendre un « virage » dont l'angle va dépendre du rapport masse sur charge de l'ion. Nous pouvons aussi évoquer certains analyseurs à rotation dans lesquels le rayon de la trajectoire de la particule analysée dépend de son rapport masse sur charge.

Un second exemple de système à séparation temporelle est celui de l'analyseur à temps de vol. Dans une cavité, les ions sont accélérés par un champ électrique qui est ensuite coupé. Ainsi, à un instant donné (ultérieur à la coupure du champ), la position des ions dépend de leur rapport masse sur charge. Les ions finissent donc leur course à l'extrémité de la cavité en un temps dépendant de leur rapport masse sur charge.

1.2.3.4 Détecteur

Le détecteur permet de transformer le flux d'ions en un flux d'électrons. Ce flux d'électrons est ensuite numérisé (échantillonné et quantifié). Le détecteur couramment employé est le multiplificateur d'électrons ou *Channeltron*.

1.2.4 Fragmentation

Lorsque la masse de la protéine ne suffit pas à identifier sa nature, ou pour obtenir une plus grande certitude d'avoir ciblé la bonne protéine, il est possible de recourir à un étage de fragmentation. Un objectif ici est de casser les peptides, de sorte que les fragments résultants aient des masses connues (fragmentation au niveau des « faiblesses » de la chaîne peptidique). Un autre objectif est de réduire la masse des particules mesurées afin de situer les masses des particules dans la dynamique du capteur, si les molécules initiales sont trop lourdes pour être détectées.

La chaîne d'analyse classique en fragmentation consiste à sélectionner avec un quadripôle un peptide, à la fragmenter dans un second quadripôle contenant un gaz inerte et à sélectionner un fragment résultant dans un troisième et dernier quadripôle avant détection. Cette technique est généralement dénommée SRM pour *Selected Reaction Monitoring*.

1.2.5 Quantification absolue

Généralement, les données en sortie du détecteur ne peuvent être exploitées que de manière relative, c'est à dire que l'on ne peut que comparer les quantités entre elles, mais qu'aucune information absolue n'est accessible.

Dans l'objectif d'une quantification absolue, certaines techniques dites « *label-free* » [WWZ⁺06] existent afin de comparer les abondances de protéines, en utilisant des molécules natives présentes naturellement dans le mélange comme calibrant.

Une autre solution consiste généralement à ajouter au mélange initial des molécules contenant des isotopes alourdis (molécules marquées) [BSS⁺07]. En effet, en ajoutant ces isotopes, de masse connue et de concentration maîtrisée, et en mesurant les quantités relatives entre les molécules inconnues et ces molécules marquées, la composition du mélange en concentration absolue peut être connue.

Il existe plusieurs techniques de marquage parmi lesquelles :

- AQUA (pour *Absolute QUAntification*) [GRS⁺03], procédure consistant à introduire des peptides marqués, c'est-à-dire à intervenir sur le mélange après digestion ; AQUA permet de quantifier les peptides ;
- PSAQ (pour *Protein Absolute Standard Quantification*) [BMGD09], procédure consistant à introduire des protéines marquées, c'est-à-dire à intervenir sur le mélange initial avant digestion ; PSAQ fait l'hypothèse que la digestion ne dépend pas du caractère marqué / non marqué de la protéine et permet de calculer les gains de digestion, donc de remonter à la quantification des protéines.

1.2.6 Traitement des données

Les données en sortie de dispositif sont numérisées. Ces données sont généralement représentées par un spectrogramme en deux dimensions.

Prenons un exemple, illustré sur la figure 1.1, sur lequel les membres du groupe PROTIS travaillent. Ici, un analyseur est associé à une colonne de chromatographie. Cette figure a pour

origine le travail de l'équipe EDyP (Étude de la Dynamique des Protéomes) de l'IRTSV (institut de Recherches en Technologies et Sciences pour le Vivant - CEA) dans le cadre du projet CAPSI (Chaîne d'Analyse Protéomique multifactorielle haute Sensibilité à composants Intégrés). Elle représente une mesure LC-MS sur du plasma après dilution de la protéine NSE.

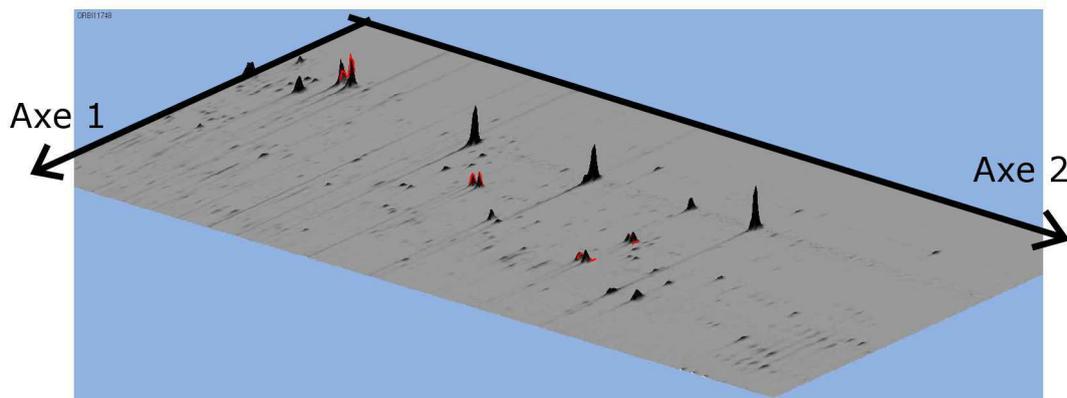


FIGURE 1.1 – Un exemple de signal de sortie d'un dispositif de spectrométrie de masse

Un axe porte l'information du temps de rétention chromatographique (axe 1), symbolisant l'affinité du composé avec la phase stationnaire de la chromatographie, et l'autre axe porte l'information spectrométrique (axe 2), en bijection avec le rapport masse-sur charge des composés.

Sur la base de ce spectrogramme, il est possible de découvrir la liste des molécules présentes (position sur les deux axes temps chromatographique - rapport masse sur charge) ou encore de quantifier certaines molécules analysées. En effet, pour connaître la liste des molécules présentes, il existe des tables permettant de connaître la position d'un peptide dans la représentation donnée ci-dessus.

L'unité de masse privilégiée en protéomique est le Dalton et le kiloDalton (symboles Da et kDa), un Dalton valant approximativement $1.66 \cdot 10^{-24}$ g et correspondant à un douzième de la masse d'un atome de Carbone 12 (soit à peu près la masse d'un atome d'hydrogène).

Afin de quantifier les molécules, il est nécessaire de quantifier l'intensité du signal donné par une molécule. Généralement, le traitement s'effectue sur un signal unidimensionnel où le rapport masse sur charge a été fixé. Ainsi, plusieurs méthodes ont été développées, comme la prise en compte du *maximum* du pic chromatographique, l'aire sous le pic, ou encore des méthodes bayésiennes reposant sur une modélisation du système d'analyse, comme étudié au sein de l'équipe PROTIS sur les thèses de Grégory Strubel [Str08, SGP⁺07] ou Pascal Szacherski [Sza13, SGG11].

Une fois ces molécules identifiées et/ou analysées, il est possible de remonter à la composition en protéine du mélange voire même d'identifier l'état clinique du patient en question. La thèse de Pascal Szacherski (équipe PROTIS) va même plus loin en proposant un modèle hiérarchique qui permet l'estimation des concentrations, l'apprentissage et la classification des patients sans passer explicitement par des étapes intermédiaires, comme la quantification des peptides dans le cas de l'estimation de la concentration des protéines [SGG11].

Les systèmes électromécaniques nanométriques

Dans ce chapitre, nous allons nous pencher sur les capteurs concernés par ce cadre de travail, à savoir les capteurs électromécaniques nanométriques. Nous présenterons le capteur à proprement parler, puis le système permettant d'en extraire une information et enfin l'utilisation de ces capteurs pour la spectrométrie de masse. Pour terminer, nous présenterons les données expérimentales sur lesquelles nous travaillerons dans le cadre de cette thèse.

Sommaire de ce chapitre

2.1	Description des capteurs	19
2.1.1	Les MEMS et les NEMS	19
2.1.2	Les NEMS capteurs de masse	19
2.1.2.1	Géométrie	19
2.1.2.2	Fréquence de résonance du capteur	20
2.1.2.3	Mesure de masse de molécules	21
2.1.2.4	Modes propres	23
2.1.3	Modèle des chutes de fréquence	23
2.1.3.1	Sensibilité du capteur	23
2.1.3.2	Valeur du gain fixe	24
2.1.3.3	Valeur du gain variable	24
	Illustration pour une configuration <i>double-clamped</i>	24
	Illustration pour une configuration <i>cantilever</i>	25
2.2	Suivi de la fréquence de résonance du capteur	27
2.2.1	Mesure de la réponse du capteur à une excitation harmonique	27
2.2.1.1	Actionnement du capteur et détection de sa position	27
2.2.1.2	Démodulation du signal de sortie	28
2.2.2	Boucle de contrôle pour le suivi de la fréquence de résonance	29
2.2.2.1	Analyse fréquentielle	29
2.2.2.2	Boucle auto-oscillante	29
2.2.2.3	Boucle à verrouillage de phase	29
2.3	La spectrométrie de masse à base de NEMS	32
2.3.1	Principe général	32

2.3.2	Hypothèses prises sur le modèle physique	32
2.3.3	Le problème de la source	33
2.4	Sessions expérimentales	35
2.4.1	Sessions du mois de Décembre 2011	35
2.4.2	Sessions des mois d'Août et de Septembre 2012	35
2.4.3	Sessions du mois de Décembre 2012	37
2.4.4	Sessions du mois de Février 2013	38

2.1 Description des capteurs

2.1.1 Les MEMS et les NEMS

Depuis quelques dizaines d'années, l'électronique numérique a pris une place considérable dans nos vies quotidiennes (appareils électro-ménagers, moyens de communication, informatique ...).

Dans la conception des puces électroniques, le silicium tient une place de choix, notamment en raison de ses propriétés de semi-conducteur qui en font un matériau sur lequel les transistors peuvent être gravés. Le transistor, et notamment le transistor à effet de champ, étant la brique élémentaire de l'électronique numérique [RRLB04], le silicium est aujourd'hui un matériau prépondérant [MVD⁺11]. Notons toutefois que certaines technologies, comme l'électronique organique [BBB⁺90, For04], ou l'utilisation de diélectriques [KMS00] existent en parallèle de la suprématie du silicium.

Les procédés de gravures sont un point d'amélioration constante, ce qui permet notamment de suivre la « loi de Moore », stipulant que le nombre de transistors sur une puce de surface fixe double tous les 18 mois.

Cette prépondérance du silicium et la maturité des technologies associées ont poussé les chercheurs et les ingénieurs à faire d'autres utilisations du silicium. Ainsi, tout en gardant une cohérence avec l'électronique numérique, se sont développés les microsystèmes électromécaniques (ou MEMS pour *Micro ElectroMechanical Systems*). Les MEMS concernent de nombreuses applications, telles que les micro-moteurs [CYP⁺12], les micro-miroirs [CBB12], les capteurs de pression [LLP⁺12] ou encore les accéléromètres [TCL12]. De par l'utilisation des technologies silicium pour mettre en place de nouvelles fonctions, on parle généralement de *more than Moore*.

Est considéré comme MEMS tout système mécanique dont au moins une des dimensions est micrométrique. Toutefois, les technologies de gravure actuelles permettent de descendre en dessous du micromètre. Ainsi sont nés les nanosystèmes électromécaniques (ou NEMS pour *Nano ElectroMechanical Systems*).

Les NEMS ne sont pas encore entrés à grande échelle dans nos vies courantes. Malgré cela, les applications potentielles de cette technologie sont très nombreuses, notamment en raison de la grande sensibilité de ces systèmes. Ainsi, des applications telles que les nano-accéléromètres ou encore les nano-bolomètres (mesure de chaleur) pourraient voir le jour. Ici, nous nous concentrons sur les nano-capteurs de masse.

2.1.2 Les NEMS capteurs de masse

2.1.2.1 Géométrie

Les nano-capteurs de masse peuvent prendre des formes diverses. Dans les cas que nous étudions ici, ces capteurs sont constitués de petites poutres de quelques micromètres de long et de quelques centaines de nanomètres de section. D'autres géométries sont aussi proposées [MPL10, LHHC10].

En particulier, nous étudierons deux configurations géométriques possibles. La première confi-

guration géométrique étudiée est la configuration « *double-clamped* », que l'on peut traduire par « doublement attaché », dans laquelle la poutre est fixée à ses deux extrémités. L'illustration présentée sur la figure 2.1 a été obtenue sur la plateforme de nanocaractérisation du CEA-Leti par Eric Sage, doctorant au CEA-Leti DCOS, en utilisant la technique de microscopie à balayage électronique. Cette figure donne une illustration de ces capteurs NEMS dans la configuration *double-clamped*.

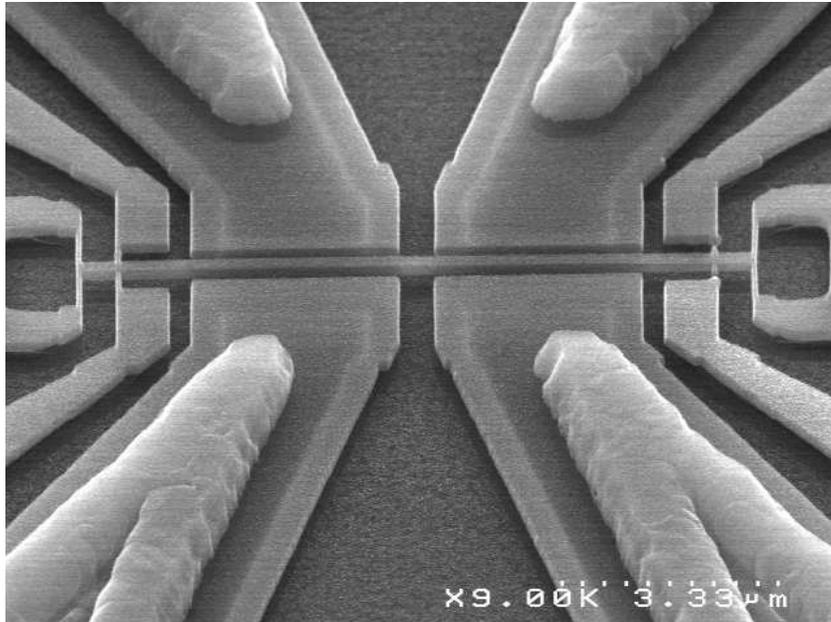


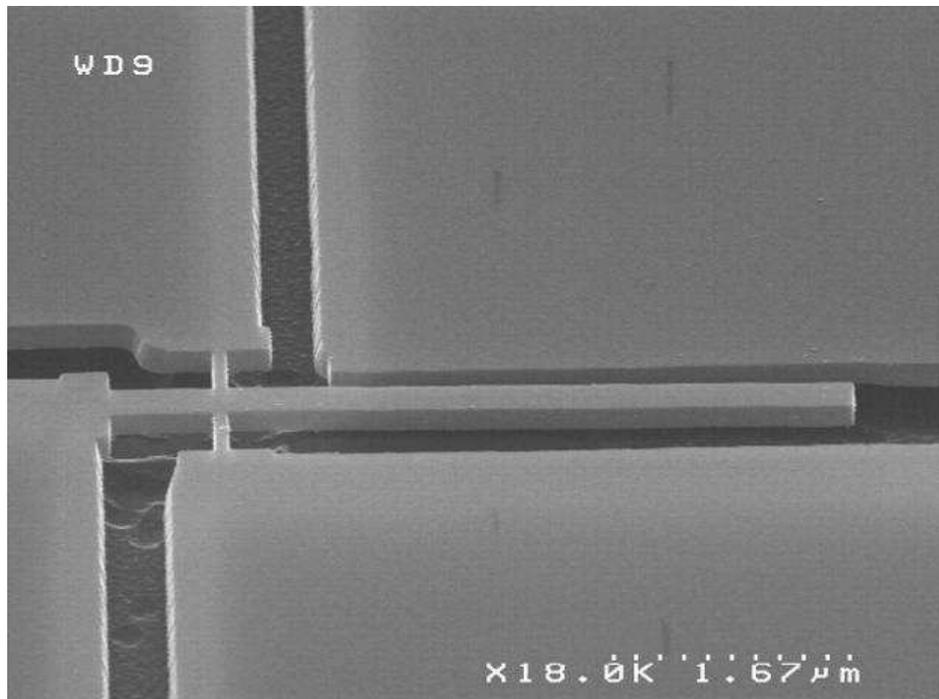
FIGURE 2.1 – Illustration d'un NEMS capteur de masse en configuration *double-clamped*

La seconde de ces configurations est la configuration « *cantilever* », que l'on peut traduire par « porte-à-faux ». Dans cette configuration, la poutre est fixée à une extrémité, la seconde extrémité étant libre. La figure 2.2, issue de l'Habilitation à Diriger des Recherches de Laurent Duraffourg [Dur08] donne une illustration de ces capteurs dans la configuration *cantilever*.

2.1.2.2 Fréquence de résonance du capteur

Pour exprimer la fréquence de résonance du capteur, on considère que la poutre est soumise à une force linéique dont l'effet ne dépend spatialement que de la position le long de la poutre [Kha09, Ch. 1]. En choisissant cette force comme oscillatoire et en écrivant l'équation différentielle résultante dans le domaine de Laplace, il apparaît que la poutre se comporte comme un système résonant amorti de fréquence de résonance f_0 [CDL⁺09] :

$$f_0 = \sqrt{\frac{k_0}{M_0}} \quad (2.1)$$

FIGURE 2.2 – Illustration d’un NEMS capteur de masse en configuration *cantilever*

Dans cette équation, k_0 est la raideur associée au système et M_0 la masse effective de la poutre. On parle ici de masse effective (et non pas de masse inertielle ou gravitationnelle) car la fréquence de résonance dépend certes de la masse totale du capteur mais aussi de la distribution spatiale de cette masse. Dans ce document, nous prendrons l’hypothèse que la masse est répartie de manière homogène au sein du capteur.

2.1.2.3 Mesure de masse de molécules

Si on augmente d’une quantité δm_{eff} la masse effective de cette poutre sans en changer la raideur, et si on suppose δm_{eff} petit devant M_0 , on peut écrire le développement au premier ordre de la fréquence de résonance du capteur [SDB10, DSBH07] :

$$\begin{aligned}
 f_{\delta m_{\text{eff}}} &= \sqrt{\frac{k_0}{M_0 + \delta m_{\text{eff}}}} \\
 &= \sqrt{\frac{k_0}{M_0}} \sqrt{\frac{1}{1 + \frac{\delta m_{\text{eff}}}{M_0}}} \\
 &\approx f_0 \left(1 - \frac{\delta m_{\text{eff}}}{2M_0}\right) \\
 &\approx f_0 - \frac{f_0}{2M_0} \cdot \delta m_{\text{eff}}
 \end{aligned}
 \tag{2.2}$$

Le principe de mesure de masse est relativement simple. On voit sur l'équation (2.2) que l'augmentation de la masse effective du capteur entraîne une chute de la fréquence de résonance de ce dernier qui peut être considérée ici comme proportionnelle à la masse effective ajoutée. Ainsi, si une molécule se « colle » au capteur, la masse de ce dernier augmentera et sa fréquence de résonance chutera.

La figure 2.3 issue du manuscrit de thèse de Chady Kharrat [Kha09, p. 118] illustre le principe de décalage en fréquence entre la fréquence de résonance initiale ω_0 et la fréquence de résonance après ajout de masse ω_{res} .

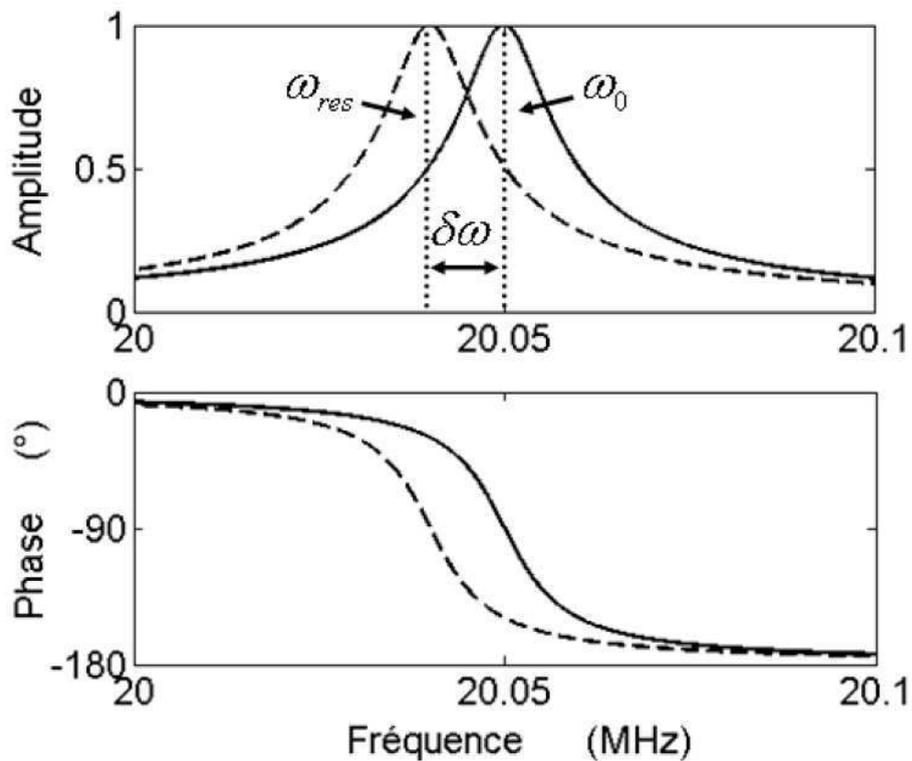


FIGURE 2.3 – Illustration du décalage en fréquence des résonateurs NEMS par ajout de masse

2.1.2.4 Modes propres

Le capteur n'a pas un seul mais plusieurs modes propres, c'est à dire plusieurs fréquences autour desquelles il est susceptible d'entrer en résonance selon différents « modes de résonances » [HKN⁺12, DSBH07]. Ces fréquences sont appelées harmoniques et classées par ordre croissant (première harmonique ou « fondamental » pour la fréquence la plus basse, puis seconde harmonique, troisième harmonique, ...). Dans ce travail, nous ne considérerons que les modes transversaux, c'est à dire les déformations qui ne dépendent que de la position le long du capteur. Notons que les fréquences harmoniques ne sont pas nécessairement multiples entre elles dans notre cas.

Dans le cas de ces résonateurs, si on note f_k la fréquence de résonance de la k -ième harmonique, M_k la masse effective pour l'harmonique k et $\Delta f_{\delta m_{\text{eff}},k}$ la chute de fréquence induite par l'ajout de la masse effective $\delta m_{\text{eff},k}$ sur la k -ième harmonique, l'équation (2.2) devient :

$$\Delta f_{\delta m_{\text{eff}},k} \approx -\frac{f_k}{2M_k} \delta m_{\text{eff},k} \quad (2.3)$$

2.1.3 Modèle des chutes de fréquence

2.1.3.1 Sensibilité du capteur

Le terme $\frac{f_k}{2M_k}$ ayant un rôle de gain dans l'équation (2.3) est d'autant plus grand que la masse du capteur est petite. On voit ici l'intérêt d'utiliser des capteurs nanométriques : étant de faible masse, ces derniers vont permettre de détecter de très faibles variations de masse effectives allant jusqu'à la molécule unique [YCF⁺06].

Cependant, de la même manière que la masse effective d'une poutre ne dépend pas que de la masse totale du capteur mais de la distribution en masse le long de la poutre, la chute de fréquence de résonance entraînée par l'ajout d'une masse ponctuelle dépend de la position à laquelle cette masse a été ajoutée. L'ajout d'une masse ponctuelle m sur la poutre à une distance z de l'extrémité de celle-ci, conduit à l'augmentation de masse effective qui est une fonction de m et de z , et qui dépend de la géométrie du capteur.

Nous proposons de normaliser les positions des molécules de sorte que celles-ci prennent des valeurs entre 0 (une extrémité du capteur) et 1 (l'autre extrémité du capteur).

Pour exprimer la sensibilité vis-à-vis de l'ajout d'une molécule, nous décomposerons le gain en deux parties. Une partie sera indépendante de la position est sera notée α_k , pour le gain fixe sur la k -ième harmonique. L'autre partie dépendra de la position de la molécule est sera notée $\phi_k(z)$. Ceci revient à développer l'équation (2.3) en donnant l'expression de la chute de fréquence $\Delta f_{m,z,k}$ sur l'harmonique k pour une masse m adsorbée à la position d'adsorption z :

$$\Delta f_{m,z,k} = -\alpha_k \phi_k(z) m \quad (2.4)$$

Nous exprimerons les gains fixes α_k en Hz.kDa^{-1} et les gains variables $\phi_k(z)$ seront exprimés sans unités.

2.1.3.2 Valeur du gain fixe

La valeur du coefficient α_k est :

$$\alpha_k = \frac{f_{0,k}}{2 M_k} \quad (2.5)$$

Dans cette équation, $f_{0,k}$ est la valeur de la fréquence de résonance du mode de vibration k , M_k est la masse effective du capteur relativement au mode de déformation k . Par rapport à la masse totale M_{tot} , cette masse vaut :

$$M_k = M_{tot} \int \phi_k(z) dz \quad (2.6)$$

Soit :

$$\alpha_k = \frac{f_{0,k}}{2 M_{tot} \int \phi_k(z) dz} \quad (2.7)$$

2.1.3.3 Valeur du gain variable

Les fonctions $\phi_k(z)$ dépendent de la géométrie du capteur. Nous proposons d'exprimer la position de la molécule sur le capteur de manière normalisée, c'est-à-dire avec le rapport entre la distance de la molécule au point d'attache de la poutre et la longueur totale de cette dernière (dans le cas d'une poutre *double-clamped*, nous choisirons arbitrairement le point de référence parmi les deux points d'attache). Aussi, nous prendrons des fonctions $\phi_k(z)$ normalisées de sorte que leur *maximum* vaille 1. Ce choix est sans conséquence dans la mesure où la fonction $\phi_k(z)$ réapparaît comme un facteur au dénominateur de l'expression de α_k .

La fonction de gain variable pour une configuration *double-clamped* a été calculée par Eric Sage, et est également donnée dans [DSBH07]. Elle a pour expression :

$$\phi_k(z) = \left(\frac{(\cosh(z\kappa_k) - \cos(z\kappa_k)) + a_k (\sinh(z\kappa_k) - \sin(z\kappa_k))}{Y_k} \right)^2 \quad (2.8)$$

Les valeurs des différents paramètres de gain κ_k , a_k , Y_k diffèrent en fonction de la géométrie du capteur. Au regard des choix effectués précédemment (position normalisée et fonctions $\phi_k(z)$ sans unité), ces paramètres seront sans unité.

Illustration pour une configuration *double-clamped* Dans le cas d'une configuration *double-clamped*, des exemples typiques de valeurs des différents paramètres de gain variable sont donnés dans le tableau 2.1. Les fonctions de gain associées sont illustrées par la figure 2.4.

a_1	a_2	a_3	κ_1	κ_2	κ_3	Y_1	Y_2	Y_3
-0.983	-1.001	-1.000	4.73	7.853	10.995	1.588	1.509	1.513

TABLE 2.1 – Valeurs de gain variable pour une configuration *double-clamped* sur les trois premières harmoniques

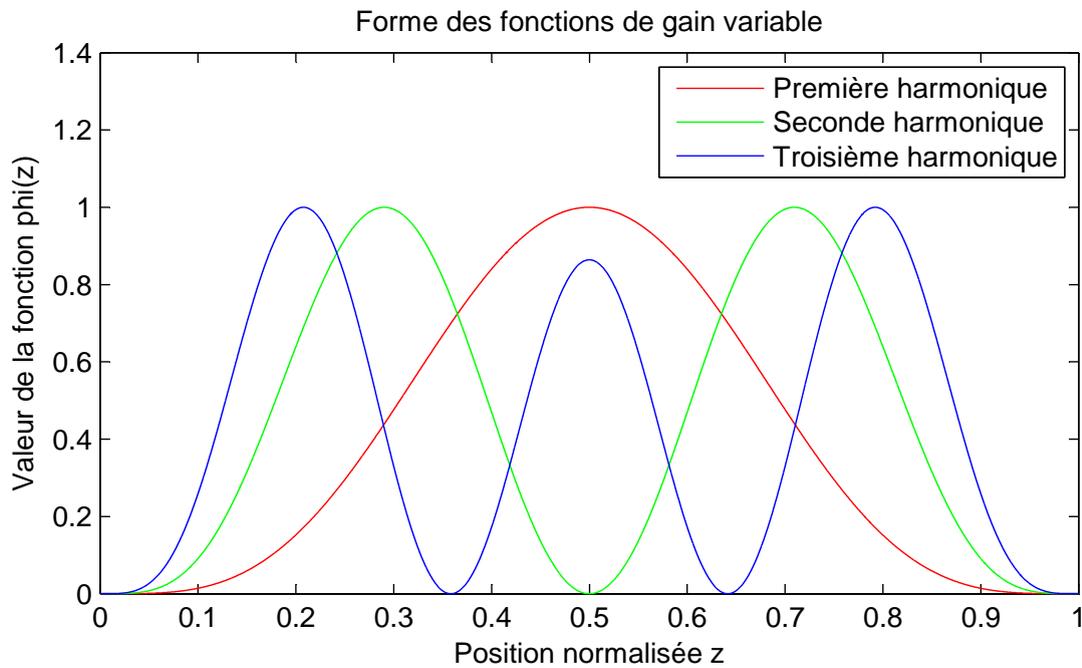


FIGURE 2.4 – Fonctions de gain variable pour une configuration *double-clamped* sur les trois premières harmoniques

Illustration pour une configuration *cantilever* Dans le cas d’une configuration *cantilever*, des exemples typiques de valeurs des différents paramètres de gain variable sont donnés dans le tableau 2.2 pour les paramètres de gain variable. Les fonctions de gain associées sont illustrées par la figure 2.5.

a_1	a_2	a_3	κ_1	κ_2	κ_3	Y_1	Y_2	Y_3
-0.734	-1.018	-0.999	1.875	4.694	7.855	2.000	1.974	2.290

TABLE 2.2 – Valeurs de gain variable pour une configuration *cantilever* sur les trois premières harmoniques

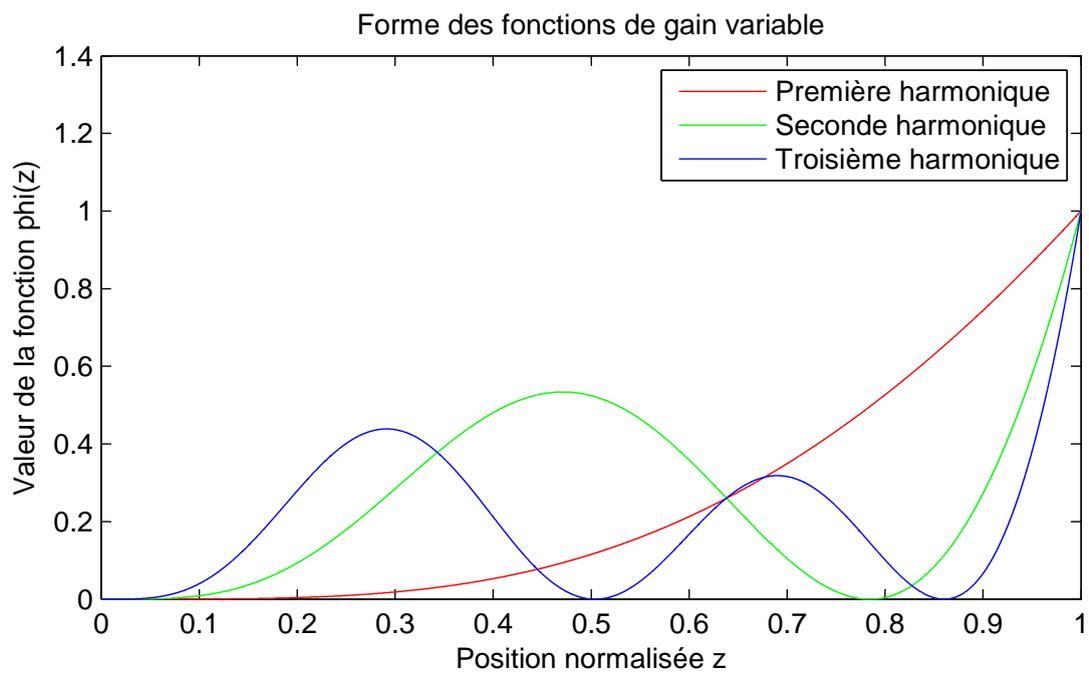


FIGURE 2.5 – Fonctions de gain variable pour une configuration *cantilever* sur les trois premières harmoniques

2.2 Suivi de la fréquence de résonance du capteur

2.2.1 Mesure de la réponse du capteur à une excitation harmonique

Nous avons vu dans la section précédente que les NEMS sont des capteurs sensibles à un ajout de masse et que ces derniers traduisent l'ajout d'une masse par une chute de fréquence de résonance.

Afin de détecter l'adsorption de masses sur la poutre et de quantifier ces masses, il est nécessaire de suivre l'évolution de cette fréquence de résonance au cours du temps. Pour cela, nous allons nous appuyer sur l'analyse harmonique. L'idée est de caractériser la réponse du capteur dans le domaine fréquentiel. Pour pouvoir connaître la réponse du capteur à une excitation sinusoïdale, les deux éléments essentiels sont un système d'actionnement du capteur associé à un système de détection de la position du capteur.

Ensuite, en commandant le système d'actionnement à une fréquence donnée et connue, et en récupérant le signal en sortie du système de détection, la réponse du capteur à une excitation harmonique peut être calculée.

2.2.1.1 Actionnement du capteur et détection de sa position

Le système d'actionnement peut reposer sur différents principes physiques [Mil09]. On peut lister les solutions techniques suivantes :

- actionnement magnétomoteur, en soumettant le capteur à un champ magnétique constant et en faisant passer dans le capteur un courant alternatif ;
- actionnement électrothermique, si le capteur est constitué de deux couches ayant des propriétés de dilatation thermique différentes, il est possible d'actionner le capteur en le chauffant par effet Joule ;
- actionnement piézoélectrique, l'effet piézoélectrique inverse étant la capacité de certains matériaux à se déformer sous l'effet d'un champ électrique (ici, alternatif) ;
- actionnement électrostatique, en plaçant une électrode à proximité du capteur et en générant un champ électrique alternatif à partir de cette électrode (des systèmes à plusieurs électrodes existent) ;
- actionnement optique à l'aide d'un dispositif LASER (*Light Amplification by Stimulated Emission of Radiation* ou amplification de la lumière par émission stimulée de rayonnement) [VDZBA⁺10].

Le système de détection, quant à lui, peut également reposer sur différents principes physiques [Mil09]. On peut lister les solutions techniques suivantes :

- détection magnétomotrice, en utilisant le champ électrique (électromoteur) généré par le capteur en mouvement soumis à un champ magnétique constant ;
- détection capacitive, en analysant la capacité du condensateur créée par le capteur et une électrode à proximité : la capacité dépend de la distance entre les deux éléments du condensateur ;
- détection piézorésistive, en utilisant le fait que certains matériaux voient leur résistivité changer sous l'effet d'une contrainte mécanique : en attachant une jauge au capteur puis en faisant

- circuler un courant dans ces jauges, la mesure de résistivité de la jauge permet de remonter à la position du capteur ;
- détection optique, en utilisant un système de détection LASER [VDZBA⁺10].

Sur les capteurs utilisés au cours de cette thèse, l'actionneur est un actionneur électrostatique et le système de détection est un système piézorésistif. Ces systèmes, pour les configurations *double-clamped* ou *cantilever*, sont illustrés par la figure 2.6. Sur cette figure, l'actionneur est représenté en bleu, le système de détection en rouge et le capteur en jaune.

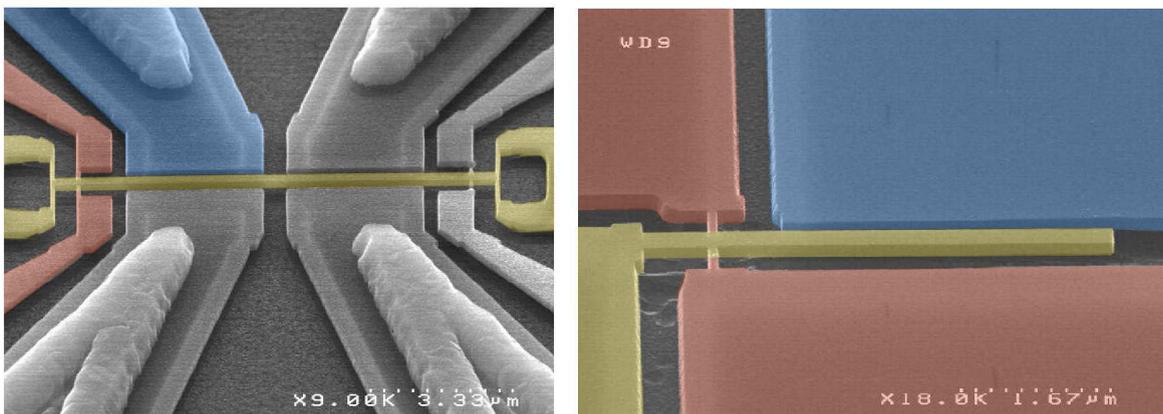


FIGURE 2.6 – Illustration du système de mesure à base de NEMS comme capteur de masse en configuration *double-clamped* (à gauche) et *cantilever* (à droite)

2.2.1.2 Démodulation du signal de sortie

Le dispositif décrit ci-dessus présente un inconvénient important. En effet, le signal d'entrée se situe aux alentours de la fréquence de résonance. Aussi, le signal de sortie du système se situe à une fréquence proche de celle du signal d'entrée (et égale lorsque l'excitation se fait à l'exacte fréquence de résonance du capteur). Or, la fréquence de résonance typique de ces capteurs est de plusieurs dizaines de MHz. Ainsi, récupérer convenablement, avec un *minimum* de pertes et de bruits, le signal de sortie s'avère compliqué, du fait de son caractère haute-fréquence. De ce fait, une technique de démodulation appelée *downmixing* est employée.

Le principe du *downmixing* est le suivant : la poutre est excitée à une fréquence f_1 , et le signal de lecture pour les jauges piézorésistives est à la fréquence $f_1 + \Delta_f$. Le signal de sortie est une modulation du signal d'entrée par le signal de lecture des jauges piézorésistives. En clair, on crée une interférence entre les signaux aux fréquences f_1 et $f_1 + \Delta_f$, ce qui génère notamment deux signaux, un étant à la fréquence Δ_f et l'autre étant à la fréquence $2f_1 + \Delta_f$. En filtrant avec un filtre passe-bas ce signal modulé, on récupère une sinusoïde à Δ_f . Ainsi, le problème de la mesure de la fréquence originale du signal de sortie est contourné.

2.2.2 Boucle de contrôle pour le suivi de la fréquence de résonance

Dans la section précédente, nous avons décrit le dispositif permettant de connaître la réponse du capteur à une fréquence donnée. Grâce à cela, il est possible de suivre la fréquence de résonance au cours du temps. Trois solutions techniques le permettent :

1. l'analyse fréquentielle ;
2. la boucle auto-oscillante [CDL⁺09] ;
3. la boucle à verrouillage de phase (PLL pour *Phase-Lock Loop*) [Kha09].

Nous allons présenter ces principes et en dresser les schémas de principe (pour simplifier ici, sans *downmixing*).

2.2.2.1 Analyse fréquentielle

L'analyse fréquentielle consiste à tracer le gain et le déphasage induit par le capteur pour une entrée sinusoïdale. Ce type d'information, lorsqu'elle est mise sous forme de courbe est appelée « diagramme de Bode ».

La figure 2.7, issue du manuscrit de thèse de Chady Kharrat [Kha09, p. 68] illustre le principe d'analyse fréquentielle.

On voit très clairement sur ce diagramme un pic de résonance se détacher. La fréquence correspondante est la fréquence de résonance du capteur. Toutefois, un tel dispositif nécessite un balayage en fréquence et n'est donc pas adapté à un système de mesure rapide.

2.2.2.2 Boucle auto-oscillante

La boucle auto-oscillante consiste à boucler l'entrée sur la sortie, c'est-à-dire à exciter le capteur avec le signal de sortie du système de détection, en appliquant une correction de gain. L'idée générale est que le gain du capteur est maximal à la résonance. Le capteur va privilégier les fréquences autour de sa fréquence de résonance, aussi appelée fréquence propre ou naturelle du système.

La figure 2.8 illustre la réalisation d'une boucle d'auto-oscillation.

2.2.2.3 Boucle à verrouillage de phase

La boucle à verrouillage de phase est un dispositif fréquemment utilisé en télécommunications afin d'asservir en phase un récepteur sur un signal.

Ici, ce dispositif repose sur un concept simple : à la résonance, le déphasage entre la sortie du dispositif et l'entrée (soit l'argument du gain complexe) vaut $-\frac{\pi}{2}$. Ainsi, en ajustant la fréquence d'excitation du capteur de sorte à maintenir un déphasage constant avec la sortie, la boucle à verrouillage de phase va suivre la fréquence de résonance du capteur NEMS.

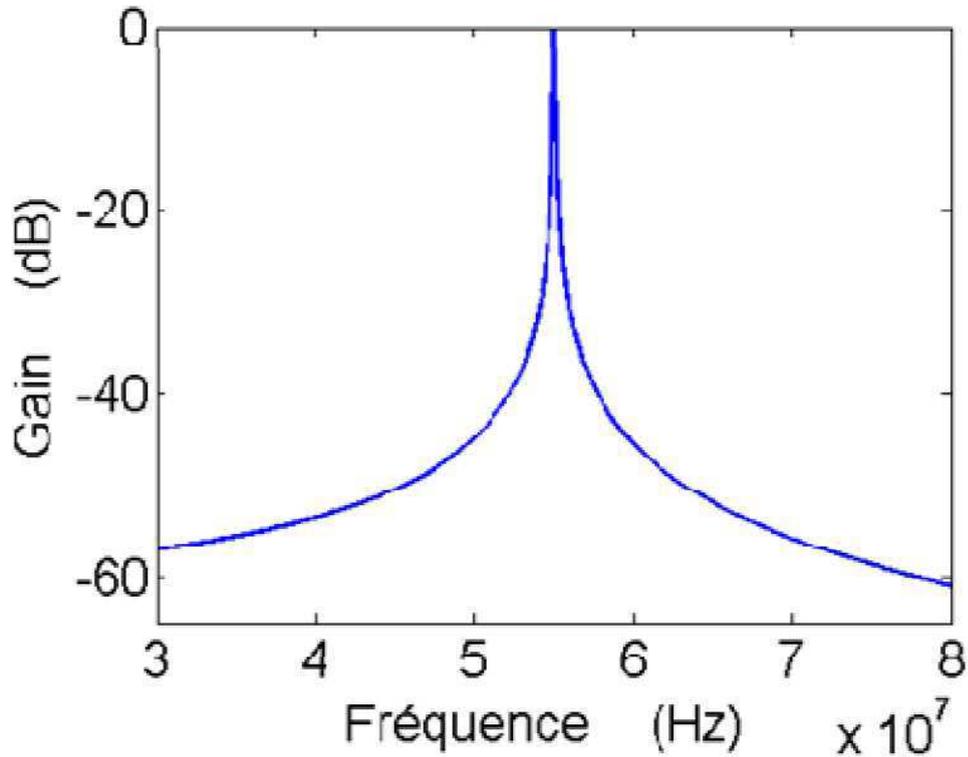


FIGURE 2.7 – Diagramme de Bode du gain du système

Notons qu'il est plus simple d'asservir la phase que le module du gain du système car la réponse en phase est linéaire au voisinage de la fréquence de résonance (voir figure 2.3 et [Kha09]).

La figure 2.9 illustre la réalisation d'une boucle à verrouillage de phase.

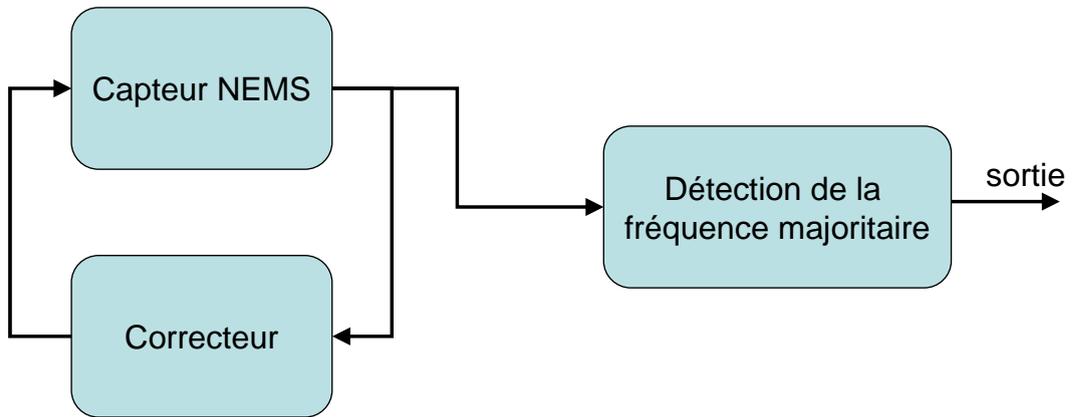


FIGURE 2.8 – Illustration de la boucle auto-oscillante

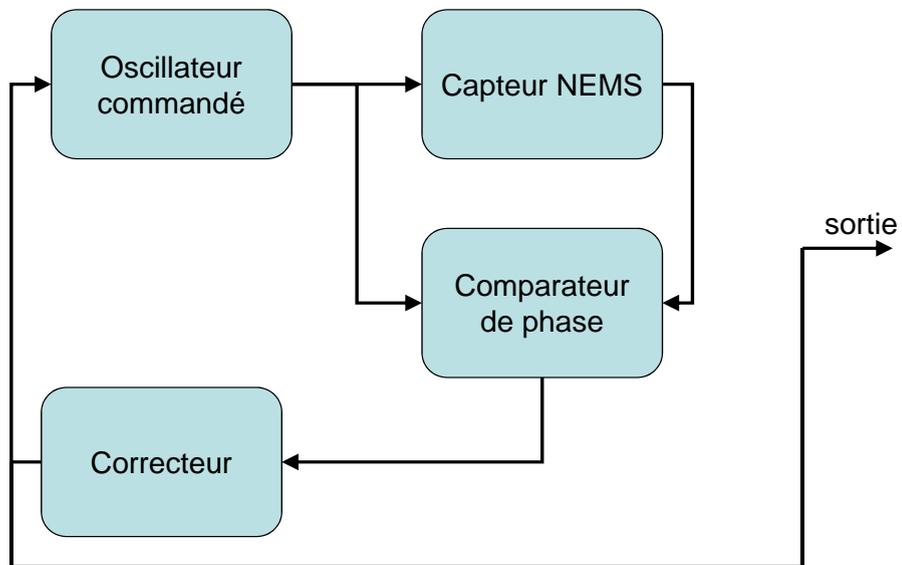


FIGURE 2.9 – Illustration de la boucle à verrouillage de phase

2.3 La spectrométrie de masse à base de NEMS

2.3.1 Principe général

Grâce aux capteurs décrits ci-dessus, il est possible d'effectuer une tâche de spectrométrie de masse. Le principe de mesure élémentaire est le suivant :

- une molécule de masse m est « projetée » sur le capteur (voir ci-dessous) ;
- cette molécule se colle à sa surface (on parle d'« adsorption ») pour des raisons d'affinité chimique (si la poutre est « fonctionnalisée ») ou de conditions de très basse température (on parle alors de « physisorption ») en une position z ;
- des systèmes électroniques permettent de suivre les fréquences de résonance sur plusieurs harmoniques au cours du temps ;
- les fréquences de résonance chutent de $\alpha_k \phi_k(z) m$ au moment de l'adsorption.

Si l'augmentation de la masse effective vient de l'ajout d'une molécule suffisamment lourde (telle une protéine) sur la poutre, et si la chute de fréquence qui en découle peut être mesurée, alors il est possible de détecter l'ajout de la molécule en question et d'en estimer la masse. On parlera alors de spectrométrie de masse à base de NEMS en mode comptage lorsque le système sera en mesure de compter une à une les molécules qui se posent sur le capteur, et à estimer leur masse.

Ainsi, le système de spectrométrie de masse à base de NEMS en mode comptage va être constitué de :

- une source permettant de générer, à partir de l'échantillon à analyser, un gaz de molécules dissociées les unes des autres ;
- un système de guidage des molécules jusqu'à l'analyseur ;
- un analyseur de masse à base de NEMS, constitué d'un (ou plusieurs) capteur(s) décrits ci-dessus, associé(s) à un système de suivi de la fréquence de résonance.

Naik *et al.* [NHH⁺09] ou Hanay *et al.* [HKN⁺12] proposent un système complet en utilisant un électrospray et en guidant les ions jusqu'au capteur *via* une électronique de focalisation à base de quadripôle.

2.3.2 Hypothèses prises sur le modèle physique

La première hypothèse que nous avons prise est celle du développement à l'ordre 1, qui vient du fait que la masse des molécules adsorbées est petite devant la masse du capteur.

On peut aussi s'interroger sur le fait que nous avons développé les calculs dans le cas où la raideur k_0 ne change pas à la suite de l'adsorption d'une molécule. Ceci se justifie par le fait que les molécules déposées sont petites en surface au regard de la poutre. Ainsi, nous supposons que les molécules ne se lient pas entre elles et ne changent pas la raideur du capteur.

Il est également possible de s'interroger sur le comportement temporel de la chute de fréquence de résonance, c'est-à-dire sur la réponse de la poutre à l'arrivée d'une molécule. Notons que la fréquence de résonance d'un capteur n'est pas une grandeur directement mesurable mais qui nécessite l'utilisation d'un transducteur comme nous l'avons vu à la section précédente. Ce

transducteur présente lui aussi un certain temps de réponse : la chute sur la fréquence mesurée n'est pas immédiate. Si l'on considère que le temps de réponse du transducteur est bien supérieur au temps de réponse de la poutre, on peut considérer la chute de fréquence physique comme immédiate et que le retard induit est principalement celui du transducteur.

Dernier point, nous avons supposé que la masse ajoutée par la molécule se posant sur le capteur était ponctuelle. Si elle ne l'était pas, le calcul de la chute de fréquence dépendrait aussi de la distribution spatiale de masse de la molécule. Nous supposons ici que les molécules sont de très petite dimension (quelques nanomètres). De ce fait, elles représentent, à l'échelle de la poutre, une masse ponctuelle.

2.3.3 Le problème de la source

Une des difficultés majeures liées à la spectrométrie de masse à base de NEMS est celle de la source. En effet, afin de permettre une analyse convenable, la source doit :

- générer un gaz de molécules ;
- préserver autant que possible les grosses molécules ;
- ne pas générer de petites molécules qui viendraient polluer les données ;
- générer suffisamment de molécules pour permettre l'analyse ;
- générer suffisamment peu de molécules pour que ces dernières arrivent une à une sur le capteur (ce point est partagé avec le système de guidage)
- avoir un comportement invariant (ou *a minima* maîtrisé) en fonction de la nature de la molécule considérée.

Ces contraintes font que la source de molécules est un des points les plus délicats aujourd'hui pour la spectrométrie de masse à base de NEMS. On dénombre plusieurs solutions techniques à l'étude pour répondre à ce problème dont des systèmes issus de la spectrométrie de masse classique (voir 1.2.3.2), à savoir le MALDI et l'électrospray, ou des systèmes inutilisés jusqu'alors dans la spectrométrie de masse, un exemple étant le système utilisé dans le cadre de cette thèse. Ce système repose sur la formation de nano-agrégats dans une chambre à très basse pression.

Le principe, illustré sur la figure 2.10, est le suivant :

- une cible (ici en Cuivre ou en Tantale) est portée à un potentiel négatifs de quelques centaines de Volts ;
- la cible est entourée d'un gaz (Argon) qui se ionise sous l'effet du champ électrique et forme un plasma ;
- les ions Argon Ar^{2+} sont attirés par la cible (de potentiel négatif) ;
- les ions Ar^{2+} frappent la cible et arrachent de la matière ;
- la matière arrachée forme des nano-agrégats, compte-tenu des conditions de pression et de température qui règnent dans l'enceinte ;
- les nano-agrégats sont portés par le gaz Argon jusqu'au détecteur NEMS.

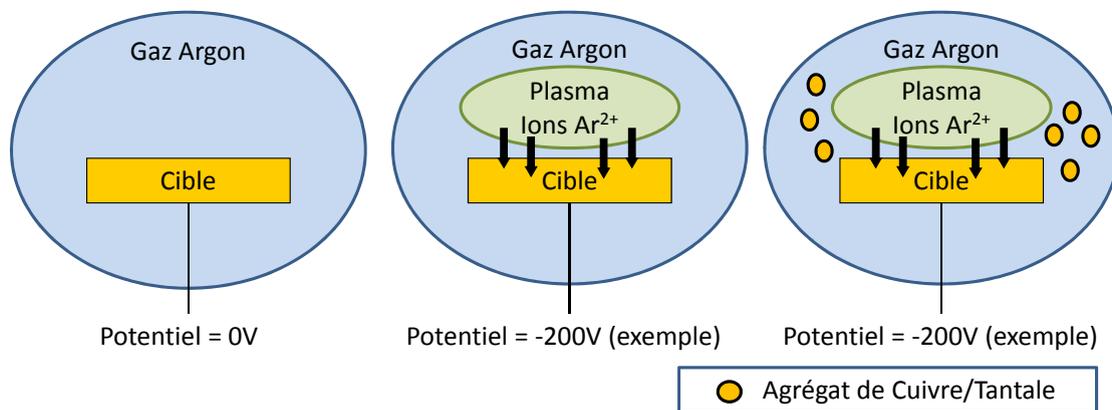


FIGURE 2.10 – Schéma de principe de génération des nano-agrégats

2.4 Sessions expérimentales

Dans le cadre de ce travail de thèse, nous travaillerons sur des jeux de données expérimentaux issus de campagnes d'acquisition effectuées par Eric Sage, ainsi que sur des données simulées. Nous allons décrire les données expérimentales que nous étudierons ici.

Ces données sont identifiées par la période à laquelle elles ont été acquises. Ces expérimentations ont été menées sur le banc de mesure présenté à la section 2.3.3, par Eric Sage, avec l'aide d'Ariel Brenac et de Robert Morel, chercheurs au CEA/INAC (Institut Nanosciences et Cryogénie). Le CEA/INAC est un institut de recherche fédératif entre le CEA (Direction des Sciences de la Matière) et l'Université Joseph Fourier.

La figure 2.11 est une photographie du banc expérimental utilisé dans le cadre de ce travail monté dans les locaux du CEA/INAC. Ce banc expérimental permet de projeter des agrégats sur la surface d'un capteur NEMS selon le principe décrit par la figure 2.10.

En aval de ce banc expérimental se trouvent un dispositif d'analyse de masse à temps de vol et une microbalance à quartz. Le dispositif à temps de vol donne une image de la répartition en masse des nano-agrégats, et la microbalance à quartz permet de contrôler la vitesse de dépôt, c'est-à-dire le débit des nano-agrégats. Grâce à un système de chambres et de positionnement des appareils dans le flux de particules, il est possible d'utiliser l'analyseur de masse à temps de vol, la microbalance à quartz ou le capteur NEMS sans avoir à ouvrir le banc et donc sans modifier les conditions expérimentales.

2.4.1 Sessions du mois de Décembre 2011

Les sessions expérimentales du mois de Décembre 2011 étaient les premières du genre au CEA. Les principales caractéristiques sont données ci-dessous.

- **Type de capteur :** *cantilever* ;
- **Matériau des agrégats :** Cuivre ;
- **Diamètre typique des agrégats :** entre 3.5 et 5 nm ;
- **Fréquence d'acquisition :** 500 Hz.

La figure 2.12 illustre un cas typique de spectre de masse acquis pendant ces sessions. Ce spectre a été obtenu à l'aide de l'analyseur à temps de vol présent lors de l'expérimentation.

Les données ont été acquises sur le mode fondamental uniquement. Les masses des capteurs n'ont pas été calculées pour cette configuration.

2.4.2 Sessions des mois d'Août et de Septembre 2012

Les principales caractéristiques des sessions expérimentales des mois d'Août et de Septembre 2012 sont données ci-dessous.

- **Type de capteur :** *double-clamped* ;
- **Matériau des agrégats :** Cuivre (Août) Tantale (Septembre) ;
- **Diamètre typique des agrégats :** entre 4 et 6 nm ;
- **Fréquence d'acquisition :** 500 Hz.



FIGURE 2.11 – Photographie du banc expérimental du CEA/INAC

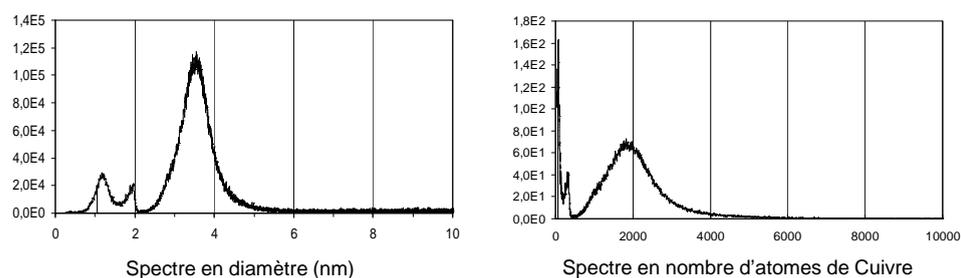


FIGURE 2.12 – Spectre de masse typique lors des sessions expérimentales de Décembre 2011

Par rapport aux sessions expérimentales du mois de Décembre 2011, le matériau des agrégats a changé. Le Tantale a notamment été utilisé en raison de sa plus grande densité que celle du Cuivre (8.96 g.cm^{-3} pour le Cuivre contre 16.4 g.cm^{-3} pour le Tantale). De ce fait, la masse d'un nano-agrégat parfaitement sphérique de Cuivre de diamètre 5 nm est d'environ 350 kDa,

alors que la masse d'un nano-agrégat parfaitement sphérique de Tantale de diamètre 6 nm est d'environ 1100 kDa. Ainsi, l'utilisation de Tantale combinée à l'augmentation du diamètre maximal des agrégats permet d'augmenter la masse des molécules adsorbées.

Ces sessions expérimentales ont été acquises sur deux modes de vibration. Le premier mode de vibration se situe aux alentours de 25 MHz et le second aux alentours de 65 MHz.

La masse des capteurs utilisés ici est de $6.20 \cdot 10^8$ kDa. Nous avons aussi calculé la valeur des masses effectives M_k en kiloDaltons pour les capteurs utilisés. Celles-ci sont données dans le tableau 2.3.

M_1	M_2
$2.46 \cdot 10^8$ kDa	$2.73 \cdot 10^8$ kDa

TABLE 2.3 – Valeurs des masses effectives sur les deux premières harmoniques pour les sessions expérimentales des mois d'Août et de Septembre 2012

Nous pouvons donc déduire les valeurs des gains fixes correspondant grâce à l'équation (2.7), ce qui sera fait lors de chaque traitement, en fonction des valeurs des fréquence de résonance.

2.4.3 Sessions du mois de Décembre 2012

Les principales caractéristiques des sessions expérimentales des mois de Décembre 2012 sont données ci-dessous.

- **Type de capteur** : *double-clamped* ;
- **Matériau des agrégats** : Tantale ;
- **Diamètre typique des agrégats** : entre 5 et 9.5 nm ;
- **Fréquence d'acquisition** : 500 Hz ou 4000 Hz.

Ces sessions expérimentales ont été acquises sur deux modes de vibration. Le premier mode de vibration se situe aux alentours de 25 MHz et le second aux alentours de 65 MHz.

Ce sont ces données que nous allons principalement utiliser dans le reste du document.

La masse des capteurs utilisés ici est de $6.06 \cdot 10^8$ kDa. Nous avons aussi calculé la valeur des masses effectives M_k en kiloDaltons pour les capteurs utilisés. Celles-ci sont données dans le tableau 2.4.

M_1	M_2
$2.41 \cdot 10^8$ kDa	$2.67 \cdot 10^8$ kDa

TABLE 2.4 – Valeurs des masses effectives sur les deux premières harmoniques pour les sessions expérimentales du mois de Décembre 2012

Nous pouvons donc déduire les valeurs des gains fixes correspondant grâce à l'équation (2.7), page 24, ce qui sera fait lors de chaque traitement, en fonction des valeurs des fréquence de résonance.

2.4.4 Sessions du mois de Février 2013

Les principales caractéristiques des sessions expérimentales des mois de Décembre 2012 sont données ci-dessous.

- **Type de capteur** : *double-clamped* ;
- **Matériau des agrégats** : Tantale ;
- **Diamètre typique des agrégats** : entre 4 et 8 nm ;
- **Fréquence d’acquisition** : 500 Hz ou 4000 Hz.

La figure 2.13 illustre un cas typique de spectre de masse acquis pendant ces sessions.

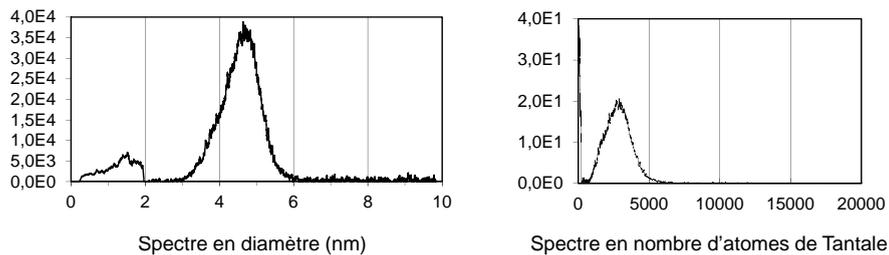


FIGURE 2.13 – Spectre de masse typique lors des sessions expérimentales de Février 2013

Ces sessions expérimentales ont été acquises sur deux modes de vibration. Le premier mode de vibration se situe aux alentours de 25 MHz et le second aux alentours de 65 MHz.

Les capteurs utilisés sont les mêmes que pour les sessions expérimentales du mois de Décembre 2012. Les masses effectives se déduisent donc du tableau 2.4.

Troisième partie

Mise en place du cadre de travail

Modélisation du système de mesure et formalisation du problème

Dans ce chapitre, nous allons modéliser le système de mesure et poser le problème sous forme d'équations. Nous présenterons dans un premier temps quelques généralités sur la manière de modéliser un problème, puis nous mettrons en place les équations qui relient les grandeurs inconnues aux grandeurs observées. Ensuite, nous verrons comment représenter les grandeurs inconnues avant de nous pencher sur la manière de prendre en compte les incertitudes dans notre modèle. Enfin, nous résumerons les problématiques liées à notre travail en trois points essentiels : détecter, quantifier et compter.

Sommaire de ce chapitre

3.1	Généralités sur la modélisation	43
3.1.1	Se donner un modèle de mesure	43
3.1.1.1	Modélisation directe du signal observé (Approche analyse)	43
3.1.1.2	Modélisation système (Approche problème inverse)	43
3.1.2	Identifier et modéliser les grandeurs d'intérêt	44
3.1.2.1	Modélisation non-paramétrique	44
3.1.2.2	Modélisation paramétrique	45
3.1.2.3	Modélisation semi-paramétrique	45
3.1.3	Prendre en compte les incertitudes	45
3.1.3.1	Différentes sources d'incertitudes	45
	Incomplétudes	45
	Erreurs de modèle	45
	Bruit	46
	Aléatoire ou pas ?	46
3.1.3.2	Utilisation de la théorie des probabilités	46
	Théorie des probabilités	47
	Vision bayésienne des probabilités	47
	Quelques exemples de lois utiles	48
	Autres méthodes non-probabilistes	48
3.2	Modélisation des comportements physiques	50
3.2.1	Schéma-bloc du système de mesure	50

3.2.1.1	Entrée du système	50
3.2.1.2	Sortie du système	51
3.2.1.3	Décomposition en blocs	51
3.2.1.4	Réseaux de capteurs	52
3.2.1.5	Indices	52
3.2.2	Modélisation des blocs	53
3.2.2.1	Modélisation de la source et de l'adsorption	53
3.2.2.2	Modélisation du comportement du capteur	53
3.2.2.3	Modélisation de la lecture de la fréquence de résonance	54
	Modèle de la réponse impulsionnelle du système de lecture	54
	Dé-synchronisation des deux modes	55
	Adsorption et échantillonnage	57
3.2.3	Résumé des équations et modèle de signal	57
3.3	Expression mathématique des grandeurs d'intérêt	59
3.3.1	Modélisation de l'arrivée des molécules	59
3.3.1.1	Écriture sous forme de processus ponctuel marqué	59
3.3.1.2	Écriture sous forme de processus à temps discret	60
3.3.1.3	Différences entre ces deux représentations	61
3.3.2	Événements élémentaires et paramètres collectifs	62
3.3.2.1	Modèles de spectre	62
3.3.2.2	Modèle hiérarchique	62
3.4	Modélisation des incertitudes	64
3.4.1	Description des incertitudes	64
3.4.1.1	Extraits de bruit	65
3.4.1.2	Densité spectrale de puissance et variance d'Allan	65
3.4.1.3	Étude de la densité spectrale de puissance	66
3.4.1.4	Égalisation de la densité spectrale de puissance	66
3.4.2	Prise en compte des perturbations par une loi normale	67
3.4.2.1	Justification basée sur l'observation	67
3.4.2.2	Écriture de la loi normale	67
3.4.2.3	Les corrélations et la matrice de covariance	68
3.4.2.4	Simplification de l'opération de blanchiment	69
	Dé-corrélation temporelle	69
	Dé-corrélation mode à mode	70
3.4.3	Synthèse	70
3.4.3.1	Pré-traitement	70
3.4.3.2	Calcul de la loi normale	71
	Gain calculatoire de la méthode proposée	73
3.5	Détecter - quantifier - compter	74
3.5.1	Détecter	74
3.5.2	Quantifier	74
3.5.3	Compter	74

3.1 Généralités sur la modélisation

3.1.1 Se donner un modèle de mesure

3.1.1.1 Modélisation directe du signal observé (Approche analyse)

Une première manière d'étudier un problème est de modéliser l'élément « observable », c'est-à-dire le signal de sortie du système de mesure, sans chercher à décrire le processus qui a servi à le générer.

Cette approche peut être qualifiée d'« approche analyse ». Elle relève aussi du domaine de la « reconnaissance des formes ». En effet, ce type de démarche revient souvent à exprimer le signal utile comme un motif qui va être ensuite recherché dans l'ensemble du signal.

Dans cette première approche, le signal observé est directement décrit par un modèle mathématique, puis les grandeurs liées à ce modèle sont reliées aux grandeurs d'intérêt.

Le principal avantage de cette approche est la généralité des méthodes qui sont transposables à plusieurs problèmes et à plusieurs instruments. Cependant, la prise en compte des incertitudes s'avère souvent difficile. Aussi ces modèles nécessitent des hypothèses souvent non satisfaites dans le cadre d'une utilisation réelle.

3.1.1.2 Modélisation système (Approche problème inverse)

Une seconde manière d'étudier un problème est de modéliser le système de mesure en identifiant clairement le signal d'entrée de ce système, qui est l'inconnue du problème. Ce même signal d'entrée peut-être modélisé d'une manière paramétrique ou non (voir section 3.1.2).

Cette approche relève du domaine de l'« assimilation de données » où l'on essaie d'identifier le modèle et les paramètres associés qui correspondent le mieux aux signaux observés.

Ce type d'approche est qualifié d'« approche système » ou approche « problème inverse ». En effet, dans ce cadre-ci, l'ensemble de la chaîne de mesure peut être pris en compte afin de décrire les incertitudes ainsi que le lien entre les grandeurs inconnues et les mesures.

On parle de « problème direct » pour décrire la relation qui lie l'entrée du système, notée f , à la sortie, notée g . Dans sa forme la plus générale, cette relation peut s'exprimer au travers de la fonctionnelle système \mathcal{H} , par la relation :

$$\mathcal{H}(f, g) = 0 \tag{3.1}$$

Il est possible d'inclure les paramètres du modèle notés θ dans cette équation :

$$\mathcal{H}(f, g, \theta) = 0 \tag{3.2}$$

Afin de ne pas alourdir les notations, nous supposons, lorsque cela n'est pas précisé dans la suite de ce document, que les paramètres θ sont connus.

Lorsque cette relation faisant intervenir la fonctionnelle système \mathcal{H} peut être remplacée par une relation linéaire explicite et lorsque cette équation est discrétisée, on obtient :

$$\mathbf{g} = \mathbf{H}\mathbf{f} \tag{3.3}$$

où \mathbf{g} contient les échantillons du signal observé, \mathbf{f} les échantillons du signal d'entrée et \mathbf{H} est un opérateur linéaire liant \mathbf{f} à \mathbf{g} , caractéristique du système étudié.

Le problème direct consiste alors à calculer le signal de sortie du système \mathbf{g} (respectivement \mathbf{g}) sachant \mathbf{f} (respectivement \mathbf{f}) et le problème inverse consiste à estimer \mathbf{f} (respectivement \mathbf{f}) à partir de \mathbf{g} (respectivement \mathbf{g}).

Les approches systèmes présentent l'intérêt d'offrir une description plus détaillée, plus précise et plus quantitative du problème. Un inconvénient de cette approche est peut-être le coût de calcul nécessaire pour la mise en œuvre des algorithmes qui en découlent.

Notons que les approche systèmes peuvent déboucher sur un modèle de signal de sortie, et permettent donc dans ce cas l'application des méthodes adaptées à l'approche analyse. Toutefois, dans certains cas, il n'est pas possible de donner une forme pré-établie à la sortie du système. Ceci peut notamment avoir lieu lorsque l'analyse du système débouche sur certaines équations aux dérivées partielles [PGBG09]. Dans une telle situation, il est particulièrement opportun d'appliquer les méthodologies problèmes inverses.

En ce qui concerne les approches système dans le domaine des nanotechnologies, le lecteur intéressé pourra se référer à [BHL10, Ch. 13].

3.1.2 Identifier et modéliser les grandeurs d'intérêt

Une fois le modèle du système de mesure posé, l'étape suivante est de se concentrer sur l'expression de l'entrée du système. Il existe plusieurs représentations plus ou moins simples et plus ou moins restrictives vis-à-vis de la réalité à modéliser. Nous avons organisé ces représentations en trois grandes familles. Il s'agit de :

1. la modélisation non-paramétrique ;
2. la modélisation paramétrique ;
3. la modélisation semi-paramétrique.

3.1.2.1 Modélisation non-paramétrique

Le cas le moins contraignant en termes de modélisation est celui de la représentation non-paramétrique. Dans ce cas, la forme du signal est totalement libre. De ce fait, chaque élément constituant le signal d'entrée est un paramètre à estimer. Le nombre d'inconnues peut être infini, notamment si le signal est un signal continu.

3.1.2.2 Modélisation paramétrique

La représentation paramétrique est une représentation dans laquelle les grandeurs d'intérêt se réduisent à un nombre restreint et fixe de paramètres desquels dépend le signal observé.

Cette représentation résulte généralement d'un modèle analytique paramétrique qui permet de transformer les paramètres en un signal modèle « observable ». La représentation paramétrique permet d'utiliser un nombre réduit de paramètres d'où une plus grande robustesse d'estimation et une facilité de mise en œuvre.

3.1.2.3 Modélisation semi-paramétrique

A mi-chemin entre la représentation paramétrique et la représentation non-paramétrique, nous proposons de définir la représentation semi-paramétrique.

Nous considérerons que la représentation semi-paramétrique est une représentation similaire à la représentation paramétrique dans laquelle le nombre de paramètres est important. Un exemple est celui de la décomposition d'un signal sur une base temporelle (l'échantillonnage).

Nous pouvons encore citer la représentation d'une image en pixels, la représentation d'un volume en voxels ou encore la représentation d'un signal temporel dans une base temps-fréquence telle qu'une base d'ondelettes, par exemple (voir section 5.2.3.3).

Cette représentation a le mérite de ramener à un nombre fini (mais élevé) la dimension de l'espace de recherche.

3.1.3 Prendre en compte les incertitudes

3.1.3.1 Différentes sources d'incertitudes

Nous venons de voir comment mettre en place une stratégie afin de modéliser un système de mesure. Cependant, les modélisations sont rarement parfaites pour différentes raisons. Nous les organisons en trois catégories :

1. les incomplétudes ;
2. les erreurs de modèles ;
3. le bruit.

Incomplétudes Les incomplétudes correspondent à des phénomènes non pris en compte par le modèle. Un exemple assez simple est celui de la géométrie. Dans le cas d'un problème volumique, supposer que l'on peut étudier le problème de manière surfacique est une source d'incomplétude : les phénomènes intervenant le long de la troisième dimension ne sont pas pris en compte.

Erreurs de modèle Les erreurs de modèle, quant à elles, sont différentes des incomplétudes dans le sens où elles correspondent à une mauvaise modélisation d'un phénomène. Par exemple,

supposer un phénomène linéaire alors qu'il ne l'est pas est une erreur de modèle. Deux exemples classiques sont la recherche d'un modèle Auto-Régressif d'ordre fixé pour modéliser un signal ou la recherche d'un polynôme de degré fixé pour approximer une courbe. Il ne s'agit pas là d'un phénomène physique non-pris en compte mais de l'approximation d'un phénomène complexe par un modèle plus simple, mais approché.

Notons également que la frontière entre incomplétudes et erreurs de modèles peut être difficile à trouver. Prenons l'exemple constitué d'un système dont la plage de fonctionnement nominale fait que l'on peut considérer son fonctionnement comme linéaire, ce qui est un cas très fréquent en électronique (modèle de l'amplificateur opérationnel, par exemple). Si le système sort de sa plage de fonctionnement nominale, le comportement du composant peut être non-linéaire (phénomènes d'ordre 2, saturations, ...). Cette source d'incertitude peut être vue comme une incomplétude (phénomène non-linéaire négligé dès le départ) ou comme une erreur de modèle (hypothèse linéaire assumée mais non valable).

Bruit La dernière source d'imperfection est celle du bruit. Derrière ce mot très fréquemment utilisé se cache une réalité assez complexe.

Un bruit usuel est le bruit thermique (ou bruit de Johnson-Nyquist). Ce bruit est généré par l'agitation thermique des porteurs de charge. D'autres bruits sont fréquemment rencontrés. En transmission, le bruit de grenaille (ou bruit de Schottky) correspond à la variance des temps d'arrivée des porteurs de charge. En électronique, le bruit de scintillation (ou bruit *flicker*) correspond à des phénomènes intervenant dans les semi-conducteurs.

Aléatoire ou pas ? Le critère que nous retiendrons pour dissocier le bruit des deux différentes sources d'imperfections citées ci-dessus est celui du caractère aléatoire de sa réalisation. Nous prendrons comme définition de l'aléatoire « qui ne pouvait être prédit auparavant ». Nous laissons aux physiciens quantiques le soin de débattre du caractère aléatoire des bruits que nous avons cités ci-dessus.

Pour illustrer la difficulté à définir le caractère aléatoire d'un phénomène, prenons un exemple extrêmement simple : le jeu du pile ou face. Avec toutes les informations sur les conditions initiales du lancer, la position initiale de la pièce dans la main, la position de la main qui rattrapera la pièce ou encore la viscosité de l'air (liste non exhaustive ...), il est possible de prédire si la pièce donnera pile ou face. Il n'y a donc pas d'aléatoire ici. Une rafale de vent peut venir perturber l'expérience : il s'agit d'une incomplétude du modèle (il faut adjoindre un modèle météorologique au modèle de prédiction). La pièce peut présenter un déséquilibre non pris en compte : il s'agit là d'une erreur sur le modèle mécanique de la pièce (pouvant venir d'une imperfection de fabrication, par exemple). L'aléatoire (le bruit) peut venir de l'opérateur, qui, si l'on considère qu'il possède un libre-arbitre, peut ne pas lancer la pièce exactement comme décrit dans le modèle. On peut aussi considérer l'opérateur comme un système extrêmement complexe mais prévisible (un robot, en quelque sorte), auquel cas son action est totalement déterministe.

3.1.3.2 Utilisation de la théorie des probabilités

Nous venons de voir dans la section précédente que de nombreuses sources d'imperfections peuvent venir perturber la bonne modélisation d'un problème. La simple étude du jeu de pile ou

face nous a amené à évoquer les modèles météorologiques, la fabrication de matériau, ou encore le libre-arbitre . . .

Considérant que dans le cas d'une étude d'un cas réel, il peut ne pas être raisonnable d'adjoindre des éléments de contrôle de tous les paramètres de l'environnement ou encore de mettre en place des modèles d'une trop grande complexité, il est nécessaire de trouver un moyen de retranscrire les incomplétudes, erreurs et bruits au travers d'une modélisation différente, en l'occurrence une modélisation non-déterministe, que nous qualifierons de stochastique (ou statistique).

La modélisation stochastique est intéressante dans le sens où elle permet une souplesse plus importante dans la mise en place du modèle, en permettant par exemple de ne modéliser que les phénomènes principaux, certains éléments d'une trop grande complexité pouvant être regroupés dans une variable aléatoire qui comprendra les incomplétudes, les erreurs de modèle et le bruit.

Théorie des probabilités Les probabilités sont une notion assez ancienne, et il prendrait une thèse entière pour retracer leur histoire. L'ouvrage [Dal91] expose l'histoire des probabilités depuis Thomas Bayes (né au début du XVIII^e siècle) à nos jours. Toutefois, les probabilités sont une notion plus ancienne que l'on retrouvait chez Aristote et qui a connu un essor durant la Renaissance. De nombreux scientifiques et philosophes de l'époque (tels que Pierre de Fermat, Christiaan Huygens ou Blaise Pascal) ont contribué à définir ce concept. Toutefois, la mise en axiome formelle des probabilités n'interviendra qu'en 1933 au travers du travail de Kolmogorov [Sap06, Ch. 1].

Il est difficile de définir la notion de probabilité en quelques lignes. L'idée générale est de fournir une mesure à un phénomène incertain. La vision la plus accessible de prime abord est la vision fréquentiste des probabilités, à savoir que la fréquence d'occurrence d'un événement sur un grand nombre de répétitions d'une même expérience s'identifie à la probabilité dudit événement.

Vision bayésienne des probabilités Contrairement à la vision fréquentiste des probabilités, la vision bayésienne des probabilités ne nécessite pas de faire appel à la fréquence statistique. En effet, pour un bayésien, la probabilité est une confiance en un événement donné [Sza13, Ch. 2].

La construction d'un système identique à celui des probabilités sans faire appel à cette notion de fréquence a été découverte par Cox [Cox61] puis par Jaynes pour donner le « théorème de Cox-Jaynes » [DT09]. Ce théorème, qui est plutôt une construction, est mis en place à l'aide de quelques règles d'inférence logique. La théorie des probabilités est ainsi reconstruite sans introduire la notion de fréquence statistique.

Notons que les résultats fréquentistes ne peuvent avoir lieu que dans le cadre d'une grande quantité de données. Les approches bayésiennes, quant à elles, sont généralement conçues pour converger vers les approches fréquentistes lorsque les données sont présentes en grande quantité [Rob05, Ch. 11].

Dans ce document, nous retiendrons cette vision des probabilités, c'est-à-dire qu'elles représentent pour nous une confiance dans la réalisation d'un phénomène.

Quelques exemples de lois utiles Nous fournissons en annexe A une description des lois de probabilités utilisées dans ce document. Sans entrer dans les détails, nous pouvons toutefois considérer plusieurs points.

La loi de probabilité la plus utilisée est la loi gaussienne ou loi normale. Plusieurs raisons « pratiques » expliquent cette popularité :

- la loi de probabilité normale est stable par un grand nombre d’opérations comme notamment les transformations d’échelle, l’addition et le produit ;
- le logarithme de l’expression de la loi de probabilité normale fait intervenir un terme quadratique par rapport à la variable considérée, ce qui permet dans de nombreux cas d’aboutir à un calcul avec une fonction à minimiser quadratique donc convexe ;
- la loi de probabilité normale prend deux paramètres explicites : la moyenne et la variance.

Bien que ces raisons pratiques soient un atout pour la loi de probabilité normale, des considérations plus théoriques viennent renforcer son attrait :

- le théorème central limite stipule que pour un grand nombre de variables aléatoires centrées réduites (vérifiant certaines conditions, une condition suffisante étant la condition de Libdeberg [PBP07, Ch. 9]), leur moyenne arithmétique converge en loi vers la loi de probabilité normale : ainsi, un phénomène aléatoire issu de plusieurs sous-phénomènes pourra être convenablement représenté au travers d’une loi de probabilité normale ;
- lorsque l’on ne connaît que la moyenne et la variance d’une loi de probabilité, le choix de la loi normale est le choix du *maximum* d’entropie [Pil10].

Toutefois, la loi normale présente quelques inconvénients, notamment :

- son support est infini, donc cette loi ne peut pas prendre une contrainte de support, comme la positivité, par exemple ;
- son asymétrie vaut 0 ;
- son kurtosis vaut 0, d’où cette loi ne peut pas facilement prendre en compte des phénomènes pouvant prendre des valeurs extrêmes.

Ainsi, la loi normale n’a pas le « monopole » des lois de probabilité, et on pourra choisir d’utiliser :

- une loi uniforme pour exprimer une limitation sur les bornes de variation, ou une loi tronquée ;
- une loi de Poisson pour représenter une variable entière positive, cette loi présentant aussi la caractéristique d’être « sans mémoire » ;
- une loi Gamma qui est en quelque sorte l’extension de la loi de Poisson sur le support \mathbb{R}^+ ;
- une loi Bêta pour représenter une variable variant entre 0 et 1 ;
- une loi de Dirichlet pour représenter des variables discrètes dont la somme vaut 1 ;
- une loi de Student-t pour exprimer un phénomène pouvant prendre des valeurs extrêmes.

Autres méthodes non-probabilistes Les probabilités ne sont pas la seule voie pour exprimer une incertitude. Sans être exhaustif, nous citerons deux alternatives.

La première est la logique floue. La logique floue étend la logique binaire [Teo99, CFM05] à des situations indéterminées. La décision à prendre par le système dépend donc de règles logiques appliquées non pas à des valeurs binaires mais à des valeurs variant entre 0 et 1. D’une certaine

manière, la logique floue peut être vue comme une simplification des probabilités.

La seconde alternative se situerait dans le contexte de la théorie de Dempster-Shafer [Ish83, Bar81]. Cette théorie fait intervenir la notion de fonction de croyance qui remplace celle de probabilité. Certains modèles comme le modèle des croyances transférables [SK94] se basent sur cette théorie.

3.2 Modélisation des comportements physiques

3.2.1 Schéma-bloc du système de mesure

Nous allons mettre en place une modélisation du système de mesure de spectrométrie de masse à base de NEMS. Pour cela, il est nécessaire de concevoir le système comme une « boîte » et d'identifier les entrées et les sorties du systèmes, ce qui est illustré par la figure 3.1.

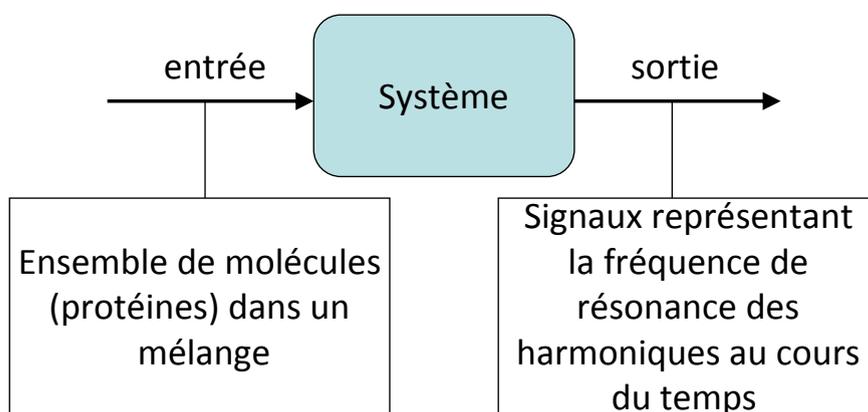


FIGURE 3.1 – Entrée et sortie du système

3.2.1.1 Entrée du système

L'entrée du système est constituée des molécules du mélange à analyser. Ce mélange peut être représenté de différentes manières :

- par identification de molécules, c'est-à-dire une liste de molécules présentes dans le mélange et les quantités associées, par exemple *deux molécules de propane, cinq molécules d'albumine, vingt molécules de silice* ;
- par identification massique, c'est-à-dire un spectre de masse, histogramme du nombre de molécules en fonction de leur masse par exemple *deux molécules de 44 Da, cinq molécules de 65 kDa, vingt molécules de 60 Da*.

Il est à noter qu'en analyse protéomique par spectrométrie de masse, la représentation choisie est la seconde. A partir de la masse d'un composé, il est possible de retrouver dans des tables la protéine correspondante (voir [Wed] pour un exemple). Dans le cas d'une analyse avec fragmentation (voir section 1.2.4), la protéine est identifiée grâce à la combinaison des masses de la molécule fragmentée et des fragments.

Nous choisirons pour cette thèse de rester au niveau masse. Nous supposons donc que la donnée de masse seule suffit à discriminer les molécules entre elles. Toutefois, comme nous l'avons présenté dans le chapitre 1, des méthodes de séparation chromatographique ou encore de fractionnement

de peptides permettent de renforcer l'identification ou la quantification de protéines, en allant plus loin que la simple notion de masse analysée. Dans ces cas-ci, la molécule est identifiée par un ensemble de grandeurs telles que « temps de sortie de la colonne chromatographique + rapport masse sur charge » dans le cadre d'une analyse par chromatographie couplée à un dispositif de spectrométrie de masse, ou encore « rapport masse sur charge de l'ion précurseur + rapports masse sur charge des fragments » dans le cas d'une analyse de type SRM.

3.2.1.2 Sortie du système

La sortie de ce système de mesure est constituée de plusieurs signaux temporels, chacun de ces signaux représentant le suivi de la fréquence de résonance du capteur au cours du temps sur une harmonique. Ce signal est constitué de sauts, correspondant chacun à l'adsorption d'une molécule sur le capteur. Un exemple de signal est donné par la figure 3.2. Ce signal est issu des sessions expérimentales du mois de Décembre 2011.

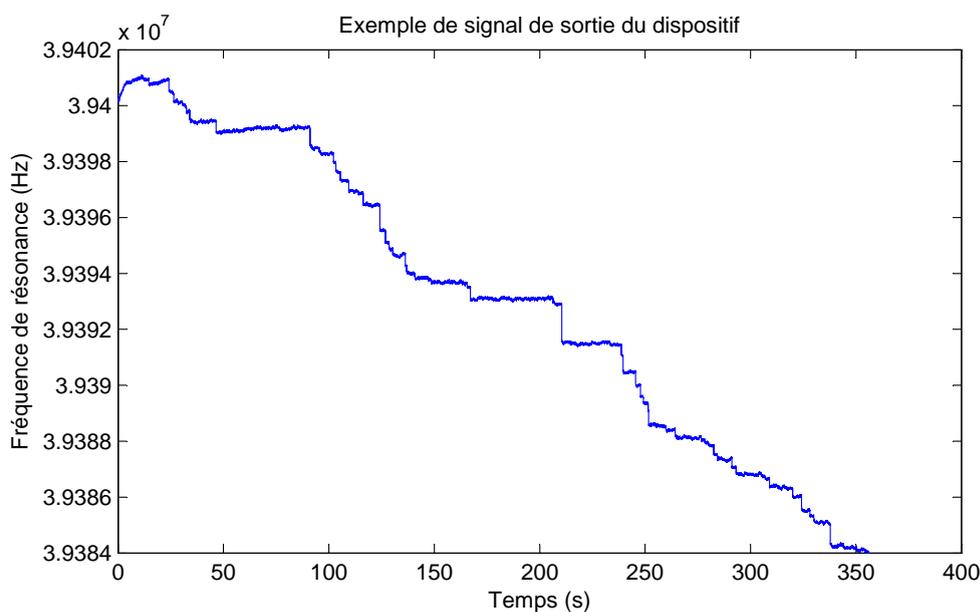


FIGURE 3.2 – Exemple de signal sur une harmonique donnée à la sortie du dispositif

3.2.1.3 Décomposition en blocs

Nous allons désormais décomposer le système de mesure en différents blocs élémentaires. Ces blocs sont :

- un système d'extraction des molécules de leur substrat, de guidage et de projection sur le capteur ;

- le capteur NEMS à proprement parler ;
- les systèmes de lectures associés à chacune des harmoniques étudiées.

Cette modélisation fonctionnelle en blocs est illustrée par la figure 3.3, sur un exemple comprenant deux harmoniques.

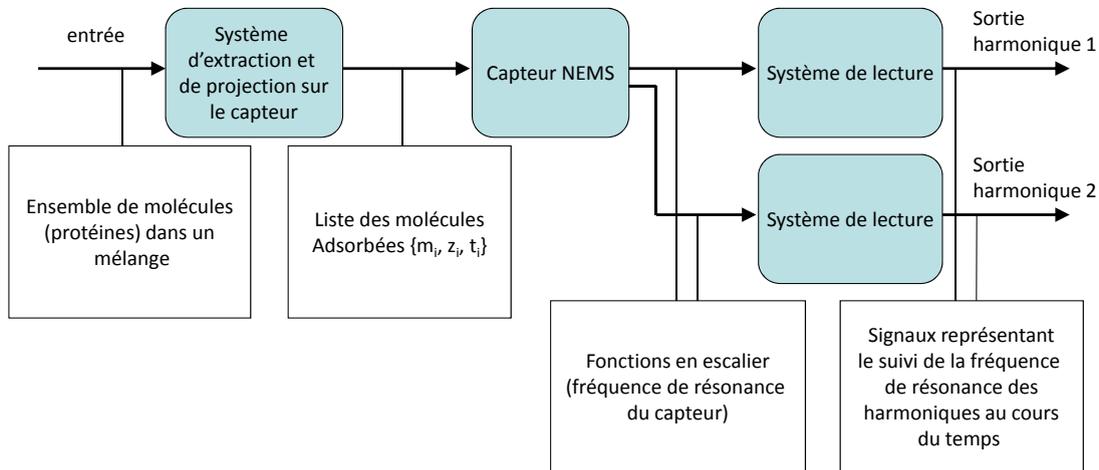


FIGURE 3.3 – Modélisation en blocs du système de mesure

3.2.1.4 Réseaux de capteurs

Jusqu'à présent, nous n'avons considéré qu'un seul capteur. Toutefois, les développements technologiques récents ont permis de mettre plusieurs capteurs en parallèle contrôlés par une seule boucle de lecture : il s'agit donc d'un système de lecture multiplexé.

Le fonctionnement du réseau de capteur est assez simple. Les capteurs du réseau ont des fréquences de résonance distinctes. La boucle de lecture balaye au fil du temps les fréquences de résonance de chacun des capteurs et excite chacun d'entre eux. La réponse la plus significative est celle du capteur effectivement en résonance.

Dans le cadre de cette thèse, nous n'étudierons que le schéma d'acquisition mono-capteur.

3.2.1.5 Indices

Nous avons défini dans les parties précédentes un certain nombre d'éléments qui forment un modèle hiérarchique. En effet, le système est constitué d'un capteur qui fonctionne sur plusieurs harmoniques et sur lequel se posent plusieurs molécules. Chacun de ces étages est repéré à l'aide d'indices. Nous proposons de les récapituler ici. Ainsi :

- nous sommes en présence d'un seul capteur ;

- ce capteur fonctionne sur K harmoniques, l'indice *harmonique* étant k , dans le cadre de cette thèse, $K = 2$;
- nous observons T échantillons temporels, l'indice de temps discret étant u ;
- sur ce capteur se posent N molécules, l'indice *molécule* étant i ; ici, N est inconnu.

3.2.2 Modélisation des blocs

3.2.2.1 Modélisation de la source et de l'adsorption

Les molécules incidentes s'adsorbent sur la poutre selon différents principes physiques décrits dans le chapitre précédent.

Nous ne possédons pas actuellement d'information sur la propension d'une molécule donnée à s'adsorber sur le capteur, ou encore sur une éventuelle affinité entre un capteur donné et une molécule donnée. De ce fait, nous prendrons les hypothèses qui nous paraissent être les plus simples et les moins engageantes.

- La première hypothèse est que dans les conditions de fonctionnement, les molécules ne se désorbent (« décollent ») pas de la poutre. Sur les données expérimentales étudiées, aucun cas de désorption spontanée n'a été observé. Nous pensons que cela est dû au caractère irréversible de l'adsorption par utilisation d'une machine de dépôt (voir section 2.3.3).
- La seconde hypothèse consiste à supposer que toutes les molécules ont la même probabilité d'adsorption sur la poutre. Cela revient à dire qu'il n'y a pas d'espèce présentant d'affinité particulière avec le matériau du capteur.
- La troisième hypothèse que nous formulons est que le débit de molécules est constant. Ainsi, la probabilité de voir une molécule se poser sur le capteur ne varie pas au cours du temps.

Par conséquent, la probabilité de voir s'adsorber à un instant donné une molécule sur le capteur sera invariante dans le temps. Elle dépendra de plusieurs paramètres comme la forme du faisceau, la section efficace du capteur, le débit de molécule ou encore la capacité d'un capteur donné à « capturer les molécules ».

Cette probabilité, ou « intensité de comptage », sera notée π .

3.2.2.2 Modélisation du comportement du capteur

Une fois qu'une molécule s'est adsorbée sur le capteur, la fréquence de résonance de ce dernier chute. Nous prendrons l'hypothèse que la réponse du capteur est immédiate et subite, c'est-à-dire que la fréquence de résonance chute immédiatement et en un temps nul. Cette hypothèse, physiquement irréaliste, se justifie par le fait que la réponse du capteur est plus rapide que celle des éléments venant en aval dans la chaîne de mesure : ainsi, les retards engendrés par le capteur peuvent être négligés.

La chute de fréquence de résonance pour la k -ième harmonique induite par l'ajout de la i -ième molécule peut se calculer à partir de m_i et de z_i au travers d'une fonction $\phi_k(z_i)$ et d'un gain α_k (voir section 2.1.2.3).

Par conséquent, la réponse élémentaire $re_k(t)$ du capteur à l'arrivée d'une unique molécule de

masse m_i , à une position z_i en un temps t_i est :

$$\begin{aligned} re_k(t) &= re_k(0) \text{ si } t < t_i \\ &re_k(0) - \alpha_k m_i \phi_k(z_i) \text{ sinon} \end{aligned} \quad (3.4)$$

Ce que l'on peut ré-écrire grâce à la fonction de Heaviside $f_{\text{Heaviside}}(t)$:

$$re_k(t) = re_k(0) - \alpha_k m_i \phi_k(z_i) f_{\text{Heaviside}}(t - t_i) \quad (3.5)$$

Intéressons-nous désormais à l'effet d'un ensemble séquentiel d'adsorptions sur le capteur. Nous supposons ainsi que l'effet de chacune des adsorptions est additif. De ce fait, la réponse $r_k(t)$ du capteur à l'arrivée d'un ensemble de molécules $\{m_i, z_i, t_i\}, i \in \{1, N\}$ est :

$$r_k(t) = re_k(0) - \sum_{i=1}^N \alpha_k m_i \phi_k(z_i) f_{\text{Heaviside}}(t - t_i) \quad (3.6)$$

3.2.2.3 Modélisation de la lecture de la fréquence de résonance

Le système électronique de lecture de la fréquence de résonance décrit dans la section 2.2 va permettre un suivi de la fréquence de résonance au cours du temps.

Sur les capteurs utilisés dans le cadre de cette thèse, le système de lecture était un système de type boucle à verrouillage de phase. Ce système peut se représenter sous forme de schéma-bloc, comme sur la figure 2.9, page 31.

Nous extrayons des sessions du mois de Décembre 2012 un exemple de réponse du système à une adsorption sur chacun des modes. Celles-ci sont données sur la figure 3.4. La période d'échantillonnage de ce signal est de 2 ms.

Modèle de la réponse impulsionnelle du système de lecture Nous modéliserons l'effet du système de lecture de fréquence de résonance comme un filtre ayant pour réponse impulsionnelle $h_{\text{PLL},k}(t)$. Ce système de lecture est attaqué par un échelon, symbolisé par la fonction de Heaviside, comme indiqué dans l'équation (3.6). La réponse du système complet à l'arrivée d'une molécule sera donc constituée de la réponse du système de lecture à un échelon, notée $u_{\text{PLL},k}(t)$, qui est la primitive de la réponse impulsionnelle.

Notons ici que la réponse impulsionnelle du système total, qui est la réponse à une adsorption, correspond à la réponse à un échelon du système de lecture, le capteur ayant un rôle d'intégrateur : les molécules restent collées au capteur et leur effet se cumule.

Nous allons approximer cette réponse impulsionnelle par une fonction exponentielle, donnée par l'équation (3.7) :

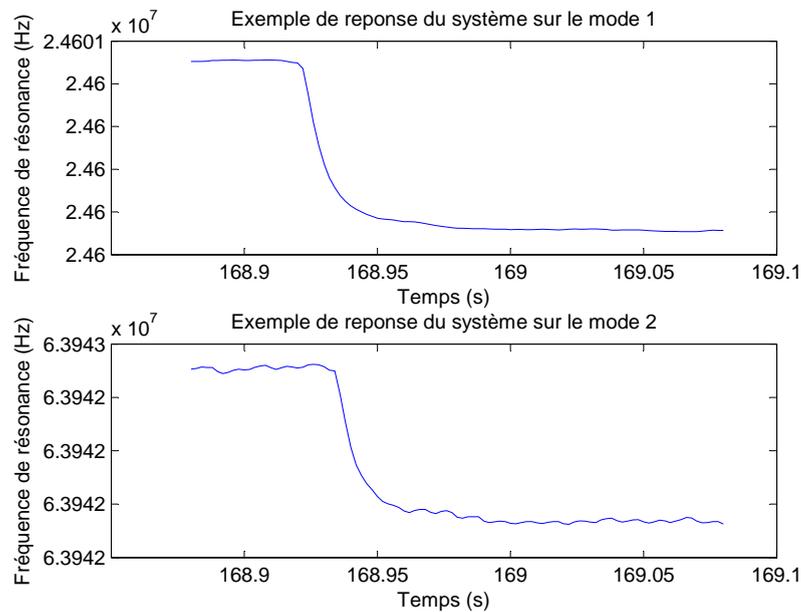


FIGURE 3.4 – Exemples de la réponse du système à une adsorption sur le mode 1 (en haut) et le mode 2 (en bas)

$$\begin{aligned}
 u_{\text{PLL}}(t) &= 1 - \exp\left(\frac{-t}{\tau}\right) \text{ si } t \geq 0 \\
 u_{\text{PLL}}(t) &= 0 \text{ sinon}
 \end{aligned}
 \tag{3.7}$$

Sur le système utilisé dans le cadre des sessions expérimentales, la constante de temps τ caractéristique du temps de réponse du système pouvait être fixée *via* une interface graphique. Cette constante de temps est choisie indépendante de l'harmonique considérée. Un ordre de grandeur typique du temps de réponse τ est de quelques millièmes de seconde.

Dé-synchronisation des deux modes Nous avons supposé que les chutes de fréquence étaient simultanées sur les différents modes. Or, il apparaît clairement sur la figure 3.4 que les chutes de fréquences sont décalées dans le temps d'un mode à l'autre.

Nous considérons que ce décalage est un problème lié à l'électronique du dispositif, électronique qui n'est pas encore définitive. Par conséquent, nous proposons une méthode simple basée sur la corrélation pour estimer ce décalage, qui sera noté κ . Nous décrivons cette méthode en annexe B.1.

Nous proposons d'écrire :

$$\begin{aligned}
 u_{\text{PLL},1}(t) &= 1 - \exp\left(\frac{-t}{\tau}\right) \text{ si } t \geq 0 \\
 u_{\text{PLL},1}(t) &= 0 \text{ sinon} \\
 u_{\text{PLL},2}(t) &= 1 - \exp\left(\frac{-(t - \kappa)}{\tau}\right) \text{ si } t \geq \kappa \\
 u_{\text{PLL},2}(t) &= 0 \text{ sinon}
 \end{aligned} \tag{3.8}$$

Représentons les deux exemples de réponse impulsionnelle de la figure 3.4 et ajustons notre modèle de réponse $-f_1 u_{\text{PLL},1}(t - \Delta_t)$ et $-f_2 u_{\text{PLL},2}(t - \Delta_t)$.

Cet ajustement est donné sur la figure 3.5. Les paramètres τ , κ de ce modèle, ainsi que les amplitudes f_1 et f_2 , ont été fixés « à la main » ici. Les valeurs utilisées pour la mise à l'échelle sont données dans le tableau 3.3.

τ	κ	f_1	f_2	Δ_t
10 ms	12 ms	380 Hz	560 Hz	168.922 s

TABLE 3.1 – Paramètres des lois normales pour l'ajustement de la figure 3.5

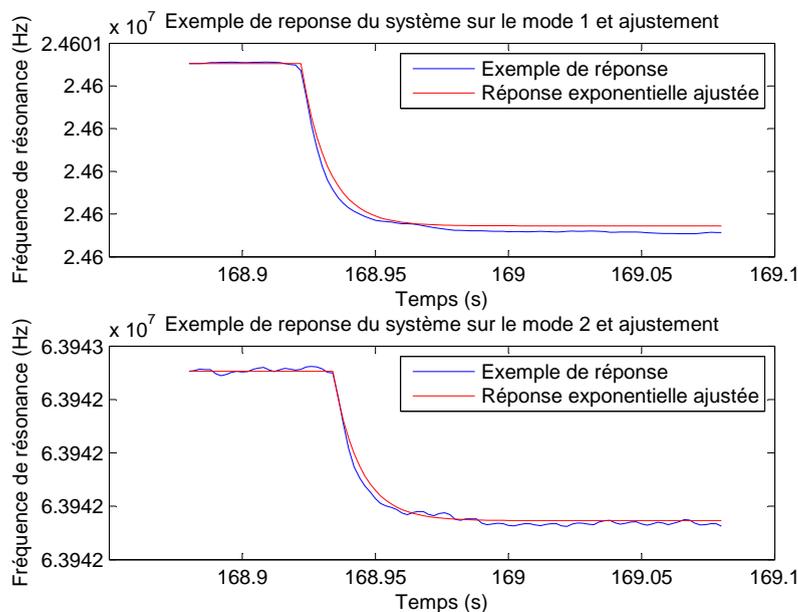


FIGURE 3.5 – Exemple de réponse du système et ajustement sur le mode 1 (en haut) et le mode 2 (en bas)

La réponse s'ajuste bien aux réponses observées. Nous pouvons donc considérer notre modèle de réponse comme étant convenable.

Adsorption et échantillonnage Observons une dernière chose sur la réponse du système : la chute de fréquence n'intervient pas exactement sur un temps d'échantillonnage mais entre deux échantillons (ce que nous appelons « instant non-entier »). Ceci peut notamment être mis en évidence sur la dérivée de la réponse du système. La figure 3.6 représente ce phénomène en exposant en haut, la dérivée de la réponse et de la réponse ajustée et en bas, la dérivée de la réponse et d'une réponse ajustée avec comme temps d'adsorption un temps aux trois-quarts d'un instant d'échantillonnage, c'est à dire $\Delta_t = 168.9235$ s.

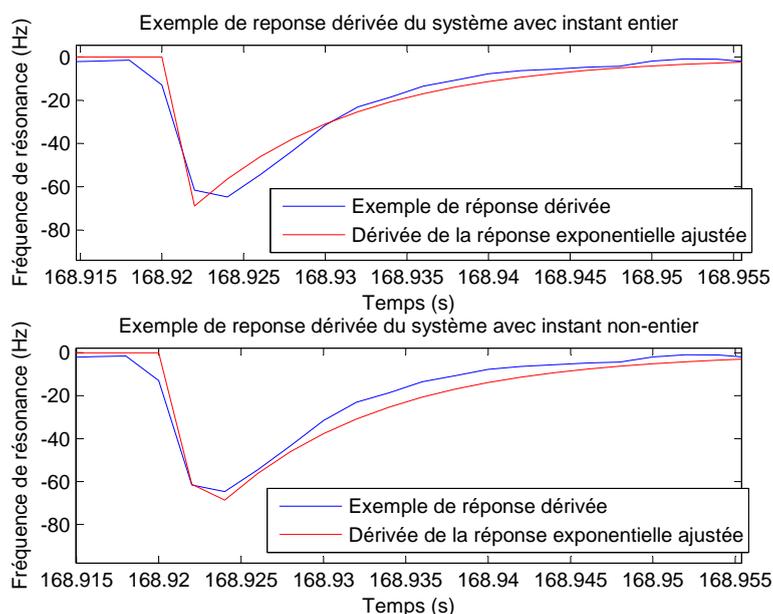


FIGURE 3.6 – Illustration du caractère "non-entier" des instants de chutes de fréquences

Nous voyons ici que la prise en compte du caractère non-entier du temps d'adsorption permet de mieux modéliser le signal observé. Ainsi, si le modèle de réponse est convenable, il est nécessaire de prévoir la possibilité de positionner des instants d'adsorption en dehors de la grille d'échantillonnage.

3.2.3 Résumé des équations et modèle de signal

Nous avons défini un ensemble de transformations et d'équations, toutes caractéristiques d'une portion du système. Ainsi, nous avons vu que les molécules étaient extraites du mélange à analyser pour être projetées puis adsorbées sur le capteur. Pour passer du mélange de protéines aux molécules adsorbées sur le capteur, nous proposons la simple utilisation d'une intensité de comptage. Ensuite, nous avons vu que les molécules faisaient chuter la fréquence de résonance du capteur, chute qui a été modélisée comme instantanée. Enfin, nous avons modélisé l'action de la boucle de lecture sur la fréquence de résonance du capteur comme une opération de filtrage.

Ainsi, pour un ensemble de N molécules de masse m_i , s'adsorbant en un temps t_i et en une

position z_i , le signal de sortie pour la k -ième harmonique $g_k(t)$ s'écrit :

$$g_k(t) = g_{0,k} - \sum_{i=1}^N \alpha_k m_i \phi_k(z_i) u_{\text{PLL},k}(t - t_i) \quad (3.9)$$

Rappelons que $u_{\text{PLL},k}$ dépend du temps de réponse τ et du temps de désynchronisation κ .

Ce que l'on peut ré-écrire à l'aide de la fonction $f_k(t)$ correspondant au train d'impulsion décrivant les chutes de fréquences découlant des adsorptions sur le NEMS et définie comme suit :

$$f_k(t) = \sum_{i=1}^N \alpha_k m_i \phi_k(z_i) \delta(t - t_i) \quad (3.10)$$

Ce qui permet de ré-écrire l'équation (3.10) :

$$g_k(t) = g_{0,k} - f_k(t) \star u_{\text{PLL},k}(t) \quad (3.11)$$

Dans ces équations, $g_{0,k}$ est la valeur initiale de la sortie pour la k -ième harmonique. Nous supposons cette valeur connue (ce sera la valeur du premier échantillon observé sur chacun des modes).

Cet ensemble d'équations, issu de la modélisation du système de mesure, est également un modèle de signal : l'effet de chacune des adsorptions $\{m_i, z_i, t_i\}$ est cumulatif et chacune des adsorptions ajoute au signal la quantité $-\alpha_k m_i \phi_k(z_i) u_{\text{PLL},k}(t - t_i)$. Ceci pourra être exploité pour le calcul du modèle direct : pour mettre à jour le modèle direct lors de l'ajout d'une adsorption, il suffira d'ajouter au signal de sortie estimé le différentiel induit par la nouvelle adsorption.

Résumons les valeurs connues du modèle :

- les paramètres de $u_{\text{PLL},k}$:
 - le temps de réponse τ ,
 - le temps de désynchronisation κ ,
- les paramètres de gain :
 - les gains fixes α_k ,
 - les fonctions de gain variable $\phi_k(\cdot)$,
- la valeur initiale $g_{0,k}$.

3.3 Expression mathématique des grandeurs d'intérêt

Pour des raisons de facilité de lecture, nous proposons d'écrire en **rouge** les grandeurs à estimer, en **bleu** les grandeurs observées et en **vert** les grandeurs connues. Ainsi, ré-écrivons l'équation (3.9) régissant le modèle, mise en place dans la partie précédente.

$$g_k(t) = g_{0,k} - \sum_{i=1}^N \alpha_k m_i \phi_k(z_i) u_{\text{PLL},k}(t - t_i) \quad (3.12)$$

Notons que la fonction $u_{\text{PLL},k}$ est déterminée par la connaissance des paramètres τ et κ , voir équation (3.8).

3.3.1 Modélisation de l'arrivée des molécules

Le flux de molécules adsorbées sur le capteur peut se modéliser au travers de l'ensemble des triplets $\{m_i, z_i, t_i\}, i \in \{1, N\}$, avec N le nombre de molécules adsorbées sur le capteur dans un intervalle de temps donné.

Cette information peut être mise sous forme de liste, pour chacun des capteurs :

$$\left\{ \begin{array}{c|c|c} m_1 & z_1 & t_1 \\ m_2 & z_2 & t_2 \\ \vdots & \vdots & \\ m_i & z_i & t_i \\ \vdots & \vdots & \\ m_N & z_N & t_N \end{array} \right\} \quad (3.13)$$

Cette représentation serait paramétrique si le nombre d'adsorptions N était connu. Ce n'est malheureusement pas le cas.

Nous allons maintenant mettre en évidence deux formulations du problème. Il s'agit :

1. de la représentation sous forme de processus ponctuel marqué ;
2. de la représentation sous forme de processus à temps discret.

3.3.1.1 Écriture sous forme de processus ponctuel marqué

L'équation (3.12), que nous rappelons ci-dessus, fait appel à une représentation non-paramétrique du signal d'entrée. En effet, chaque composante de temps peut donner lieu à un événement symbolisé par le triplet $\{m_i, z_i, t_i\}$.

Il est possible d'explicitier le caractère non-paramétrique de cette représentation en écrivant :

$$m(t) = \sum_{i=1}^N m_i \delta(t_i) \quad (3.14)$$

$$z(t) = \sum_{i=1}^N z_i \delta(t_i) \quad (3.15)$$

Alors il vient de l'équation (3.12) :

$$g_k(t) = g_{0,k} - \alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{\text{PLL},k}(t) \quad (3.16)$$

Avec l'opérateur \diamond que nous avons défini dans les notations en début de document.

Nous appellerons cette représentation processus ponctuel marqué dans la mesure où l'inconnu est un processus temporel prenant quelques valeurs non nulles (en l'occurrence aux temps t_i) ; ce processus est « marqué » car il est constitué d'un ensemble d'événements auxquels sont attribués un ou plusieurs paramètres (en l'occurrence deux paramètres m_i et z_i) aux endroits où ce processus est non nul.

Notons qu'il est possible d'introduire la fonction des chutes de fréquence de résonance $f_k(t)$, comme précédemment au travers de l'équation (3.10). Cette fonction vaut $\alpha_k (m(t) \diamond \phi_k(z(t)))$ sinon. Ainsi :

$$g(t) = g_{0,k} - f_k(t) \star u_{\text{PLL},k}(t) \quad (3.17)$$

3.3.1.2 Écriture sous forme de processus à temps discret

Il est possible de contourner la difficulté liée à la dimension infinie d'une représentation du problème sous forme non-paramétrique. Pour ce faire, une possibilité assez simple est de transformer le problème non-paramétrique en un problème semi-paramétrique en projetant sur une base finie le signal temporel.

Nous choisissons ici de projeter le signal sur une base temporelle constituée d'impulsions de Dirac séparées par un pas de discrétisation Δ_t , ce qui revient simplement à dire que nous échantillons le signal. Nous appellerons cette version discrétisée des signaux $m(t)$ et $z(t)$: \mathbf{m} et \mathbf{z} respectivement. Le signal observé sera noté \mathbf{g}_k .

$$\mathbf{m} = m(t) \text{III}_{\Delta_t}(t) \quad (3.18)$$

$$\mathbf{z} = z(t) \text{III}_{\Delta_t}(t) \quad (3.19)$$

Dans ces équations, $\text{III}_{\Delta_t}(t)$ représente le peigne de Dirac de pas Δ_t . Avec cette version discrétisée, et compte-tenu de la linéarité du modèle, il est possible d'écrire à partir de l'équation (3.12) :

$$\mathbf{g}_k = \mathbf{g}_{0,k} - \sum_{u=1}^T \alpha_k \mathbf{m}_u \phi_k(\mathbf{z}_u) \mathbf{u}_{\text{PLL},k}(u) \quad (3.20)$$

où \mathbf{m}_u et \mathbf{z}_u sont les échantillons des discrétisations des signaux $\mathbf{m}(t)$ et $\mathbf{z}(t)$, T est le nombre d'échantillons et $\mathbf{u}_{\text{PLL},k}$ est la discrétisation de $\mathbf{u}_{\text{PLL},k}(t - u\Delta_t)$.

Cette représentation permet de s'affranchir de la dimension infinie liée à la représentation non-paramétrique. De plus, si l'échantillonnage est fait conformément aux conditions de Shannon (c'est-à-dire avec une fréquence d'échantillonnage au moins deux fois supérieure à la fréquence maximale du signal), ce dernier se fait sans perte d'information. L'écueil de cette représentation est de forcer les instants d'adsorption à être des multiples de Δ_t , ce qui n'est pas vrai dans la pratique (voir figure 3.6). Toutefois, si Δ_t est suffisamment petit, cette approximation sera convenable.

Comme précédemment, notons qu'il est possible d'introduire le vecteur des chutes de fréquence de résonance \mathbf{f}_k qui vaut 0 lorsque \mathbf{m} vaut 0 et vaut $\alpha_k(\mathbf{m} \diamond \phi_k(\mathbf{z}))$ sinon. Ainsi :

$$\mathbf{g} = \mathbf{g}_{0,k} - \mathbf{f}_k \star \mathbf{u}_{\text{PLL},k}(0) \quad (3.21)$$

3.3.1.3 Différences entre ces deux représentations

Le choix d'utiliser l'une ou l'autre de ces représentations n'est pas évident. Notons que la thèse de Vincent Mazet [Maz05], à laquelle nous ferons souvent appel au cours de ce travail, comporte deux chapitres portant sur la déconvolution impulsionnelle, un prenant une modélisation sous forme de processus à temps discret et un autre prenant une forme de processus ponctuel marqué (qu'il appelle « décomposition en motifs élémentaires »).

Il est également possible de retrouver les approches type processus ponctuel marqué dans les analyses d'images aériennes, comme [DDBZ08, DBD⁺11] qui compte des populations de flamands roses ou [TBBD10] qui repositionne des bâtiments.

Dressons un comparatif entre ces deux approches dans le tableau 3.2. Dans chacun des cas, nous désignons par le terme générique « événement » les adsorptions. Nous noterons T le nombre d'échantillons, N le nombre d'événements et L le nombre d'attributs par événements (ici, $L = 2$, le temps ne comptant pas comme « attribut »).

	PPM	PTD
Nombre de paramètres du modèle	$N(L + 1) + 1$	TL
Position des événements	Libre	Grille d'échantillonnage
Dimension de l'espace de recherche	Variable	Fixé et grand (TL)
Nom de la représentation associée	Non-paramétrique	Semi-paramétrique

TABLE 3.2 – Comparaison des approches processus ponctuels marqués (PPM) et processus à temps discret (PTD)

3.3.2 Événements élémentaires et paramètres collectifs

3.3.2.1 Modèles de spectre

Nous venons de voir que dans le cadre de notre problématique, le signal dépendait d'une liste d'événements (voir équation (3.13)), qui symbolisent l'entrée du transducteur. N'oublions pas que nous cherchons avant tout à mesurer un spectre de masse duquel sont issues les molécules qui s'adsorbent sur le capteur. Chacun des événements donne lieu à des chutes de fréquence, qui elles-mêmes donnent lieu à des signaux observés.

Nous proposons de résumer notre problématique sous la forme de graphe illustrant un rapport hiérarchique entre les grandeurs. Nous introduisons notamment le spectre de masse Sp qui peut être soit un modèle continu de spectre (les molécules peuvent prendre des valeurs de masse totalement libres) ou un modèle discret (les molécules ne peuvent prendre qu'un nombre restreint de valeurs).

Notons que ce deuxième cas, le modèle discret, n'est pas une simple vue de l'esprit. En effet, l'analyse d'un milieu biologique peut donner un spectre discret. Prenons l'exemple d'un mélange de protéines, contenant des protéines ayant des masses bien identifiées dans une liste finie prédéfinie : ainsi son spectre de masse est discret. Qui plus est, les protéines fortement semblables (par exemple, les isotopes) présenteront des masses tellement proches qu'au regard de la précision de l'instrument, supposer ces masses égales ne sera pas une hypothèse absurde.

3.3.2.2 Modèle hiérarchique

Les informations que nous venons de voir peuvent se traduire au travers d'un modèle graphique, c'est-à-dire d'un graphe permettant de mettre en valeur les dépendances entre les différents paramètres. Ce type de hiérarchie se retrouve dans plusieurs problématiques et est particulièrement bien décrit dans la thèse de Pascal Szacherski [Sza13].

Nous distinguerons plusieurs niveaux dans notre décomposition hiérarchique.

- le signal observé ;
- la fréquence de résonance ;
- la liste des adsorptions ;
- le mélange considéré.

Nous pouvons représenter ces informations hiérarchiques dans un graphe acyclique orienté. Bien que les paramètres soient inter-dépendants, les relations de dépendance pour un paramètre donné peuvent se réduire à la dépendance envers un nombre réduit de paramètres. Cette représentation hiérarchique fait ressortir cette notion dite d'indépendance conditionnelle » : pour un niveau de hiérarchie donné, un nœud dépend uniquement du nœud père à l'étage au-dessus et des nœuds fils à l'étage en-dessous. Nous représentons donc cette hiérarchie par la figure 3.7.

Nous symboliserons par des doubles flèches (\implies) les relations déterministes et par des simples flèches (\longrightarrow) les relations statistiques. Nous symboliserons par un rectangle les grandeurs concernant plusieurs harmoniques.

A ce stade, nous ne connaissons pas encore la nature de l'influence du bruit $\epsilon_k(t)$ sur le signal

observé. Nous laissons donc les liens qui pointent vers $g_k(t)$ comme statistiques.

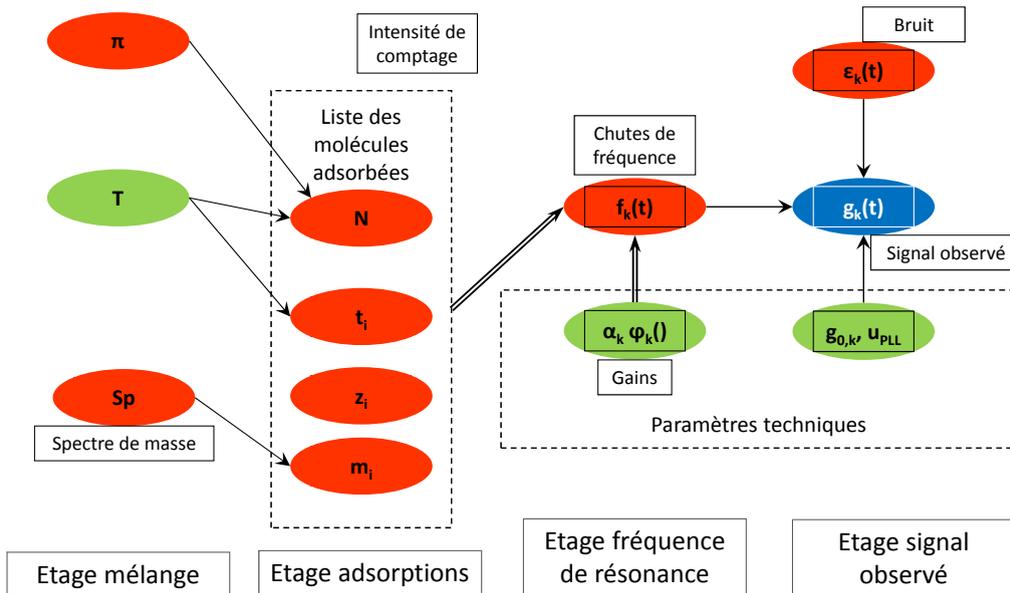


FIGURE 3.7 – Modélisation hiérarchique du problème

3.4 Modélisation des incertitudes

Nous allons désormais étudier les perturbations sur le signal observé. Pour faciliter la lecture et alléger le corps du document, nous avons placé plus de détails en annexe C.

3.4.1 Description des incertitudes

Les perturbations sur le signal de sortie sont de natures très diverses [CR02] : bruit thermomécanique agitant la poutre, bruit de l'électronique de lecture, . . . A cela, on peut éventuellement ajouter les erreurs de modèles, et les incomplétudes (phénomènes physiques non pris en compte).

A partir des données expérimentales, nous proposons de mesurer certaines caractéristiques du bruit en sortie. Au stade auquel nous écrivons ces lignes, il ne nous est pas possible de prendre explicitement en compte les erreurs de modèle et les incomplétudes.

L'hypothèse de base que nous faisons est que le bruit, les incomplétudes et les erreurs se ramènent à un bruit additif, qui agit sur le signal observé (le suivi de la fréquence de résonance au cours du temps) indépendamment de ce dernier. Ceci est illustré sur la figure 3.8, pour un système à deux harmoniques. Nous supposons aussi le bruit ergodique et stationnaire.

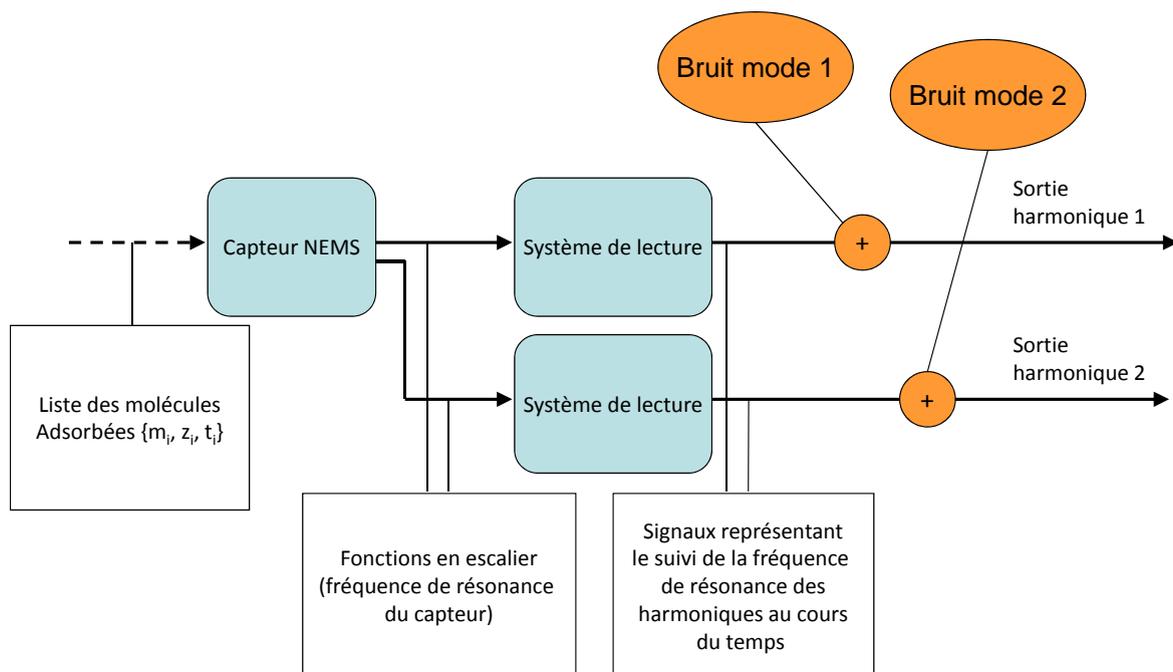


FIGURE 3.8 – Illustration du caractère additif du bruit pour un système à deux modes

Par conséquent, la connaissance du bruit $\epsilon_k(t)$, du signal des chutes de fréquences $f_k(t)$, de la réponse impulsionnelle $u_{\text{PLL},k}(t)$ et de la valeur initiale du signal $g_{0,k}$ permettent de déterminer parfaitement $g(t)$:

$$g(t) = g_{0,k} - f_k(t) \star u_{\text{PLL},k}(t) + \epsilon_k(t) \quad (3.22)$$

Ceci permet de modifier la partie « droite » de la hiérarchie initialement présentée sur la figure 3.7, comme présenté sur la figure 3.9.

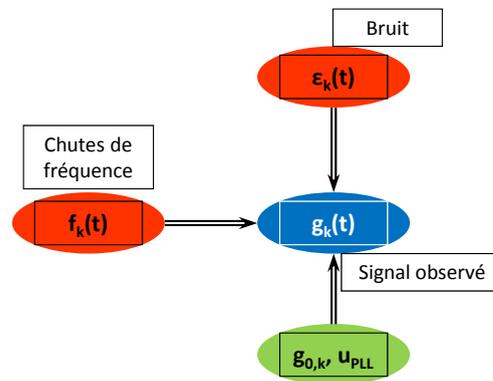


FIGURE 3.9 – Modification des relations hiérarchiques due au caractère additif et indépendant du bruit

3.4.1.1 Extraits de bruit

Étant donné les hypothèses formulées ci-dessus, et considérant que nous avons accès à des extraits de bruit dans les données expérimentales, nous proposons de mener une étude sur la structure de signaux de bruit obtenus :

- soit en isolant des parties « sans événements » dans des données expérimentales ;
- soit en utilisant des signaux de bruit pur obtenus au cours de sessions expérimentales sans génération d’agrégats.

Nous extrayons des données expérimentales des sessions de Décembre 2012 deux signaux (notés A et B). Ces signaux sont présentés en annexe C.1 sur les figures C.1 pour le signal A et C.2 pour le signal B. Le signal A est un signal de bruit seul (la source de molécule est désactivée) et le signal B est un extrait d’une partie « plate » d’un signal expérimental, que nous considérerons donc comme étant composé de bruit uniquement (dépourvue de tout signal). Ces signaux sont ré-alignés selon la procédure décrite en annexe B.1.

3.4.1.2 Densité spectrale de puissance et variance d’Allan

Un outil fréquemment utilisé dans le domaine des résonateurs est celui de la variance d’Allan [LDM10, SB08], qui est définie comme la variance d’un échantillon à l’autre d’un signal temporel

numérisé. Celle-ci est définie pour différents temps d'intégration τ , et est notée $\sigma_{\text{Allan}}^2(\tau)$.

Si l'on dispose d'un signal numérique noté \mathbf{y} , dont les échantillons ont été obtenus par intégration du signal analogique $y(t)$ sur un temps d'intégration τ , alors la variance d'Allan s'estime comme suit :

$$\widehat{\sigma_{\text{Allan}}^2(\tau)} = \frac{1}{2T} \sum_{u=1}^T (\mathbf{y}_{u+1} - \mathbf{y}_u)^2 \quad (3.23)$$

Nous préférons étudier une quantité généralement plus employée dans le domaine du traitement de signal : la densité spectrale de puissance (ou DSP). Celle-ci est estimée comme une limite faisant intervenir le carré du module de la transformée de Fourier.

$$\text{DSP}_{\mathbf{y}}(f) = \lim_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E}_{p(\mathbf{y}(t))} \left[\left| \int_0^T \mathbf{y}(t) \exp(-i_{\mathbb{C}} f t) dt \right|^2 \right] \quad (3.24)$$

Le théorème de Wiener–Khintchine permet de la relier à l'auto-corrélation du signal $r_x(\tau)$.

$$\text{DSP}_{\mathbf{y}}(f) = TF(r_x(\tau)) \quad (3.25)$$

Notons qu'il est possible de passer de la densité spectrale de puissance d'un signal à sa variance d'Allan par la formule suivante [Wit01] :

$$\widehat{\sigma_{\text{Allan}}^2(\tau)} = 2 \int_0^{+\infty} \text{DSP}_{\mathbf{y}}(f) \frac{\sin^4(\pi f \tau)}{\pi f \tau} df \quad (3.26)$$

3.4.1.3 Étude de la densité spectrale de puissance

Commençons par étudier la densité spectrale de puissance du bruit du signal A. Les densités spectrales de puissance sur chacun des modes sont présentées en annexe C.1 sur les figures C.3 pour le mode 1 et C.4 pour le mode 2.

Nous pouvons noter deux choses. La première est que l'on peut apercevoir plusieurs raies tous les 50 Hz. Celles-ci viennent certainement d'une perturbation faisant intervenir le réseau électrique qui, en France, présente une fréquence de 50 Hz. La seconde observation est que le signal est essentiellement basse fréquence. Aussi, la représentation logarithmique est assez caractéristique d'un bruit dit « en $1/f$ », caractéristique des composants électroniques actifs et des résonateurs [RGBG00].

3.4.1.4 Égalisation de la densité spectrale de puissance

Afin d'égaliser le signal de bruit, nous proposons d'effectuer deux traitements sur les signaux évoqués ci-dessus.

- dans un premier temps, un « pré-traitement » au cours duquel nous supprimons les raies tous les 50 Hz par une méthode basée sur un filtre dans le domaine de Fourier et décrit en annexe B.2;
- ensuite, nous dérivons le signal pour redresser la pente observée ci-dessus.

Recalculons les Densités Spectrales de Puissance pour chacun de ces signaux de bruit, après les opérations (filtrage et dérivation) décrites ci-dessus. Les densités spectrales de puissance après ces opérations sur chacun des modes sont présentées en annexe C.1 sur les figures C.5 pour le mode 1 et C.6 pour le mode 2.

Ces opérations ont visiblement ré-égalisé les spectres de fréquence du bruit.

3.4.2 Prise en compte des perturbations par une loi normale

3.4.2.1 Justification basée sur l'observation

Nous allons désormais essayer de modéliser le bruit par une loi de probabilité normale, ce qui se justifie par des questions à la fois théoriques et pratiques (voir section 3.1.3.2). En termes de données expérimentales, affichons sur la figure 3.10 l'histogramme des données de bruit après pré-traitement et dérivation. Nous traçons aussi une loi normale ajustée à cet histogramme. Les paramètres de ces lois normales sont donnés dans le tableau 3.3.

Moyenne mode 1	Écart-type mode 1	Moyenne mode 2	Écart-type mode 2
$5.4 \cdot 10^{-4}$	0.65	$2.6 \cdot 10^{-3}$	3.62

TABLE 3.3 – Paramètres des lois normales pour l'ajustement de la figure 3.10

Cet histogramme nous amène à penser que l'hypothèse normale est adaptée à ce cas d'étude. Nous ajouterons aussi l'hypothèse de centrage de la loi normale, compte-tenu des valeurs des moyennes dans cet exemple.

3.4.2.2 Écriture de la loi normale

Remarquons que sous les hypothèses d'additivité et de centrage du bruit décrites ci-dessus, l'étude du bruit revient à l'étude de l'écart entre un signal reconstruit et le signal observé. Prenons un exemple avec \mathbf{f}_0 le signal inconnu, \mathbf{g}_0 le signal observé et une reconstruction s'opérant par $\mathbf{H}\mathbf{f}_0$. Ici, le bruit est noté \mathbf{x} (bruit observé).

$$\begin{aligned}
 p(\mathbf{g}_0|\mathbf{f}_0, \mathbf{\Gamma}) &= \mathcal{N}(\mathbf{g}_0|\mathbf{H}\mathbf{f}_0, \mathbf{\Gamma}) \\
 \Rightarrow p(\mathbf{x}|\mathbf{\Gamma}) &= \mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Gamma})
 \end{aligned}
 \tag{3.27}$$

Nous voyons ici le lien très fort entre la loi de probabilité du bruit et la loi de probabilité du signal observé sachant une entrée donnée, venant de la relation $\mathbf{x} = \mathbf{g}_0 - \mathbf{H}\mathbf{f}_0$. Nous qualifierions donc la loi de probabilité du bruit de « vraisemblance » car c'est à travers elle que nous pourrions

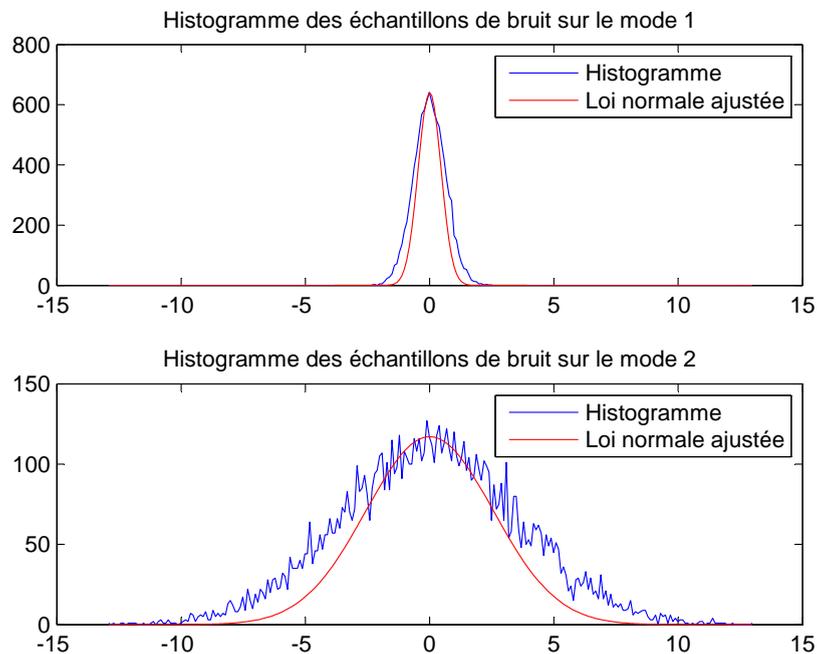


FIGURE 3.10 – Histogramme des échantillons du signal A après pré-traitement et loi normale ajustée

juger de la qualité de la reconstruction d'un signal à partir d'une entrée donnée, par rapport au signal observé. Nous remarquons aussi que l'élément central de ces lois de probabilité est la matrice de covariance $\mathbf{\Gamma}$.

En notant $\mathbf{x}_{k,u}$ la u -ième composante du vecteur du signal de bruit observé, nous considérerons par la suite le vecteur d'observation \mathbf{x} défini comme suit :

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{1,1} \\ \mathbf{x}_{1,2} \\ \vdots \\ \mathbf{x}_{1,T} \\ \mathbf{x}_{2,1} \\ \mathbf{x}_{2,2} \\ \vdots \\ \mathbf{x}_{2,T} \end{pmatrix} \quad (3.28)$$

3.4.2.3 Les corrélations et la matrice de covariance

La prise en compte des corrélations du bruit se fait au travers de la matrice de covariance $\mathbf{\Gamma}$:

$$p(\mathbf{x}|\mathbf{\Gamma}) = \mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Gamma}) \quad (3.29)$$

Il est présenté en annexe C.2 que l'utilisation de la matrice de covariance dans le calcul d'une vraisemblance normale induit une opération de blanchiment du signal, voir équation (C.5). Par conséquent, les opérations de pré-traitement et de dérivation vues précédemment font partie de ce blanchiment.

Nous proposons ici de prendre en compte dans notre loi de probabilité normale multivariée les corrélations résiduelles temporelles et inter-modes du bruit qui ne peuvent pas être retirées par le pré-traitement ou par la dérivation, au travers de la matrice de covariance $\mathbf{\Gamma}$. Par conséquent, dans notre cas, le vecteur de bruit \mathbf{x} aura au préalable subi ces deux opérations.

Le lecteur intéressé trouvera en annexe C.2 des illustrations de cette matrice par les figures C.8 et C.9 ainsi que l'auto-corrélation sur la figure C.7.

Nous voyons notamment des corrélations résiduelles qui n'ont pas été retirées par pré-traitement. Le signal reste corrélé sur des temps courts (de l'ordre de quelques millisecondes). Ceci justifie l'intérêt de faire l'étude de la matrice de covariance $\mathbf{\Gamma}$ pour prendre en compte ces corrélations dans notre modèle. Cette matrice étant estimée, nous la noterons désormais $\mathbf{\Gamma}$. Des détails sur l'estimation de la matrice de covariance sont donnés en annexe C.2.

3.4.2.4 Simplification de l'opération de blanchiment

Pour ne pas avoir à calculer l'opération matricielle caractéristique de la loi normale multivariée à chaque étape ($\mathbf{x}^T \mathbf{\Gamma}^{-1} \mathbf{x}$), nous proposons une méthode alternative qui a le même objectif : blanchir le bruit (ou le résidu). Ceci se fera en deux étapes :

1. une étape de dé-corrélation temporelle sur chacun des modes ;
2. une étape de dé-corrélation mode à mode.

Notons que les corrélations « croisées » ne seront pas prises en compte. Au regard des figures C.8 et C.9, ceci ne nous paraît pas trop impactant.

Dé-corrélation temporelle La première étape consiste à dé-corréler le signal sur chacun des modes séparément. Pour cela :

- nous estimons la matrice de covariance du bruit sur chacun des modes après pré-traitement et dérivation $\mathbf{\Gamma}_{fk}$, en prenant une matrice de dimension V ;
- nous extrayons la décomposition de Cholesky $\mathbf{\Gamma}_{fk}^{-\frac{1}{2}}$ de l'inverse de cette matrice
- de ces décompositions de Cholesky, nous extrayons la première ligne, qui forme V coefficients notés $\mathbf{a}_{u,k}$, $u \in \{1, V\}$;
- ces coefficients vont être utilisés en tant que filtre à moyenne ajustée blanchisseur pour le bruit.

A la fin de cette procédure, nous considérerons les échantillons sur chaque canal de bruit (sur chacune des harmoniques) seront des échantillons dé-corrélés temporellement, suivant une loi normale centrée.

Le bruit après filtrage sur l'harmonique k sera noté $\mathbf{x}_{f,k}$ et sa u -ième composante $\mathbf{x}_{f,k,u}$.

Dé-corrélation mode à mode Il reste encore à dé-corréler les échantillons mode à mode. Pour cela, nous construisons la matrice d'observation \mathbf{O} :

$$\mathbf{O} = \begin{pmatrix} \mathbf{x}_{f,1}^T \\ \mathbf{x}_{f,2}^T \\ \vdots \\ \mathbf{x}_{f,K}^T \end{pmatrix} \quad (3.30)$$

Les colonnes de cette matrice sont notées $\mathbf{O}^{(u)}$.

$$\mathbf{O}^{(u)} = \begin{pmatrix} \mathbf{x}_{f,1,u} \\ \mathbf{x}_{f,2,u} \\ \vdots \\ \mathbf{x}_{f,K,u} \end{pmatrix} \quad (3.31)$$

Nous considérerons $\mathbf{O}^{(u)}$ comme distribué sous une loi normale centrée de matrice de covariance $\mathbf{\Gamma}_m$, de dimension $K \times K$. La matrice $\mathbf{\Gamma}_m$ va donc permettre de blanchir le bruit au travers de l'opération $\mathbf{O}^T \mathbf{\Gamma}_m^{-1} \mathbf{O}$.

3.4.3 Synthèse

Synthétisons le modèle de bruit vu auparavant et extrayons-en le calcul simplifié de vraisemblance.

3.4.3.1 Pré-traitement

Nous proposons de pré-traiter le bruit avec une suppression du bruit de secteur avec la méthode décrite en annexe B.2. Ces opérations affectent autant le bruit que le signal. Comme les pré-traitements sont linéaires et comme le signal et le bruit s'additionnent, nous pouvons étudier l'effet des pré-traitements sur le signal seul.

Étudions donc l'effet de la suppression du bruit de secteur ces opération sur le signal. En effet, le pré-traitement affecte aussi la réponse impulsionnelle $u_{\text{PLL},k}(t)$. La figure 3.11 compare la réponse impulsionnelle à laquelle on fait subir le pré-traitement et $u_{\text{PLL}}(t)$.

Nous considérerons que $u_{\text{PLL}}(t)$ est une approximation acceptable de la réponse impulsionnelle après suppression du bruit de secteur.

Ainsi, le signal sera pré-traité pour enlever le bruit de secteur, sans que cela affecte le modèle de réponse impulsionnelle. Nous choisissons d'effectuer cette opération au préalable de tout calcul car :

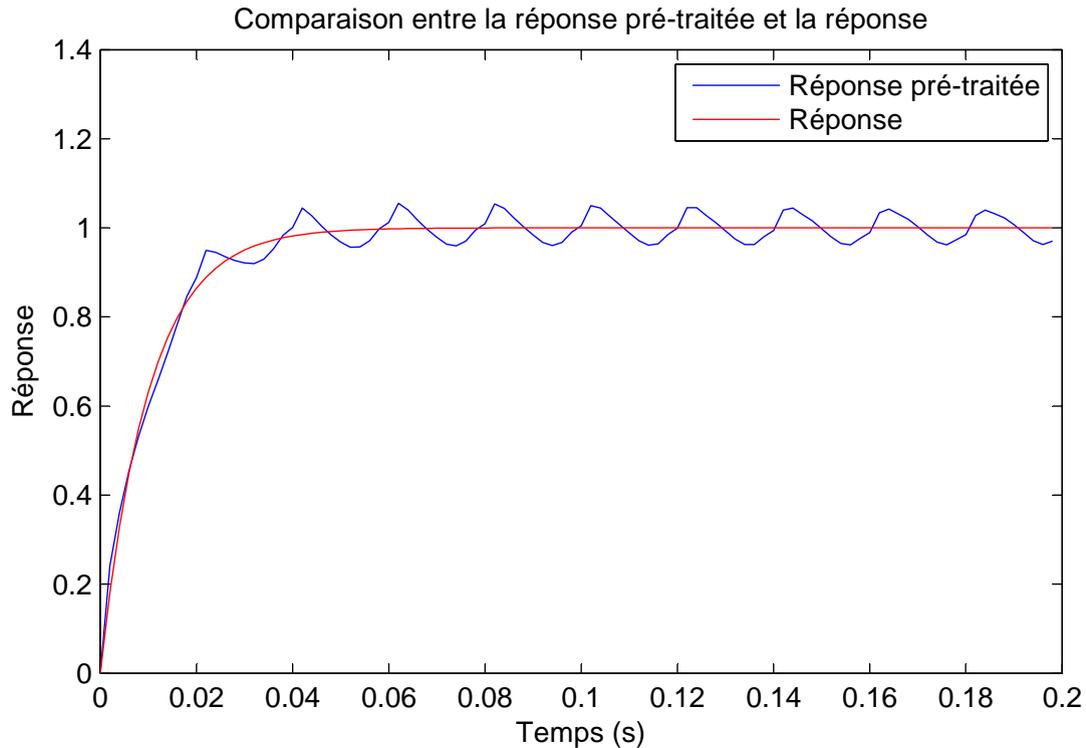


FIGURE 3.11 – Influence du pré-traitement sur la réponse impulsionnelle

- cette opération se fait une seule fois au lieu de se faire à chaque évaluation de la vraisemblance ;
- nous pensons que sortir cette opération de l'opération de filtrage intervenant par la suite permet de ne pas consommer les degrés de liberté du filtre pour supprimer ce phénomène.

3.4.3.2 Calcul de la loi normale

La loi de vraisemblance sera notée dans le cas mono-mode :

$$p(\mathbf{g}|m(t), z(t)) = \mathcal{N}\left(\mathbf{g}|\underline{f(t) \star u_{PLL}, k, \Gamma}\right) \quad (3.32)$$

et dans le cas multi-mode :

$$p(\mathbf{g}_k|m(t), z(t)) = \mathcal{N}\left(\mathbf{g}_k|\underline{\alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{PLL}(t), \Gamma}\right)_{k \in \{1;K\}} \quad (3.33)$$

Nous noterons $\cdot \star u_{\text{PLL}}$ ou $\cdot \star u_{\text{PLL}, k_{k \in \{1, K\}}}$ l'opération consistant à reconstruire, échantillonner, mettre sous forme de vecteur la construction du modèle direct.

Prenons le cas multi-mode : comme expliqué ci-dessus, nous commencerons par calculer le résidu $\mathbf{g}_k - \alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{\text{PLL}}(t)$. Le vecteur obtenu suivra une loi normale multivariée centrée de matrice de covariance $\mathbf{\Gamma}$.

Le calcul de la vraisemblance normale fait intervenir un calcul matriciel qui peut être coûteux. Nous proposons de remplacer ce calcul par deux étapes.

La première étape est un filtrage temporel en deux temps. Le premier temps est un dérivation du résidu (fonction Matlab *diff*) puis un filtrage sur chacun des modes indépendamment les uns des autres (fonction Matlab *conv*), l'apprentissage du filtre étant décrit ci-dessus. Pour les signaux utilisés ici, les coefficients estimés du filtre sont donnés par le tableau 3.4.

Mode 1	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	1.8250	-1.0502	0.0214	-0.0140	0.0946
Mode 1	$a_{6,1}$	$a_{7,1}$	$a_{8,1}$	$a_{9,1}$	$a_{10,1}$
	0.0276	0.0854	-0.0730	0.0439	0.1170
Mode 2	$a_{1,2}$	$a_{2,2}$	$a_{3,2}$	$a_{4,2}$	$a_{5,2}$
	0.2841	-0.0643	0.0055	0.0133	0.0089
Mode 2	$a_{6,2}$	$a_{7,2}$	$a_{8,2}$	$a_{9,2}$	$a_{10,2}$
	0.0190	-0.0019	-0.0030	0.0124	0.0282

TABLE 3.4 – Coefficients des filtres blanchisseurs

La seconde étape consiste à créer pour chaque instant le vecteur constitué des échantillons de chaque mode après filtrage (équation (3.30)). Nous modéliserons ce vecteur comme étant une variable aléatoire suivant une loi normale, centrée, de covariance $\mathbf{\Gamma}_m$.

Le calcul de la vraisemblance se fait donc en appliquant :

$$\mathcal{N} \left(\mathbf{g}_k | \alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{\text{PLL}}(t), \mathbf{\Gamma} \right) \approx \sum_{u=1}^T \mathbf{O}^{(u)T} \mathbf{\Gamma}_m^{-1} \mathbf{O}^{(u)} \quad (3.34)$$

ou $\mathcal{N} \left(\mathbf{g}_k | \alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{\text{PLL}}(t), \mathbf{\Gamma} \right) \approx \text{tr}(\mathbf{O}^T \mathbf{\Gamma}_m^{-1} \mathbf{O})$

Avec les signaux décrits ci-dessus, la matrice de covariance $\mathbf{\Gamma}_m$ vaut :

$$\mathbf{\Gamma}_m = \begin{pmatrix} 0.9972 & 0.1103 \\ 0.1103 & 0.9964 \end{pmatrix} \quad (3.35)$$

Une validation de cette méthode de calcul se trouve en annexe C.2.

Gain calculatoire de la méthode proposée En simplifiant le calcul comme nous le faisons ici, nous passons d'un calcul matriciel de dimension $KT \times KT$ à une convolution suivie d'un calcul matriciel en dimension $K \times K$ et d'une somme sur T échantillons.

Ceci permet de gagner en ressource mémoire (il n'y a pas de matrice de covariance de taille $T \times T$ à stocker). Aussi, cela permet d'éviter les problèmes d'inversion numérique inhérents à l'inversion de la matrice Γ . Enfin, cela permet un gain en temps de calcul substantiel, comme le montre la figure 3.12, dans le cas où le filtre est un filtre d'ordre $V = 10$ et le système travaille sur 2 harmoniques.

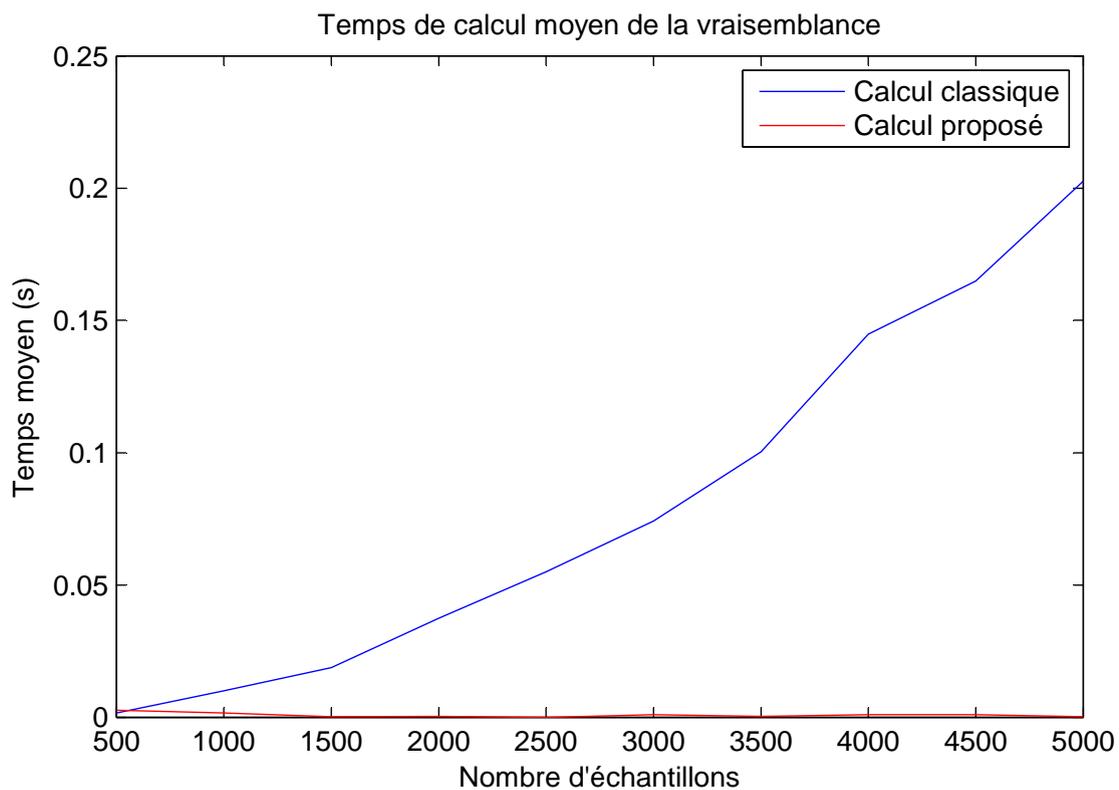


FIGURE 3.12 – Illustration du gain en temps de calcul par la méthode de calcul de vraisemblance proposée et un filtre d'ordre 10 et 2 harmoniques

3.5 Détecter - quantifier - compter

Maintenant que nous avons modélisé le problème, voyons quels sont les enjeux qui nous attendent pour la partie traitement. Ceux-ci sont au nombre de trois. Il s'agit de :

- **Détecter** les adsorptions de molécules.
- **Quantifier** chacune des masses adsorbées.
- **Compter** les molécules pour reconstruire le spectre de masse.

3.5.1 Détecter

La première tâche du traitement consiste à détecter les molécules qui s'adsorbent sur le capteur, ce qui est équivalent à retrouver les instants d'adsorption.

Concrètement, si l'on considère une seule harmonique, cette étape de détection consiste à retrouver la fonction $f(t)$ ou le vecteur \mathbf{f} . Ceci nécessite d'inverser l'opérateur de convolution $\star u_{\text{PLL},k}(t)$ ou $\star u_{\text{PLL},k}$. Ces signaux inconnus $f(t)$ ou \mathbf{f} comportent un grand nombre de valeurs nulles. On parle alors de signaux « parcimonieux » et donc de déconvolution avec contrainte de parcimonie ou déconvolution impulsionnelle.

Nous pourrions également exploiter une information très intéressante issue du modèle : la simultanéité des instants d'adsorption. Ainsi, nous aurons intérêt à chercher simultanément les chutes de fréquences sur tous les modes accessibles.

3.5.2 Quantifier

La seconde tâche qui nous attend est celle de la quantification des masses adsorbées.

Pour chaque adsorption détectée, les chutes de fréquence induites sur chacune des harmoniques doivent être estimées. Ensuite, ces chutes de fréquences peuvent être transformées en information de masse m_i et de position z_i [HKN⁺12].

Trois problèmes seront alors à résoudre :

- la fonction reliant la masse et la position d'une molécule à sa chute de fréquence est non-linéaire ;
- il peut ne pas y avoir de couple masse-position qui corresponde aux chutes de fréquences observées ;
- l'estimation des chutes de fréquence n'est pas parfaite et il peut être nécessaire de prendre en considération ces incertitudes ;
- des difficultés numériques (non convexité, *minima* locaux) peuvent rendre difficile l'estimation du couple masse-position.

3.5.3 Compter

La dernière étape consiste à reconstituer le spectre de masse à partir des masses estimées. Le problème ici est d'accéder à une information collective au mélange étudié à partir de multiples mesures qui :

- peuvent être d’une qualité différente au sein d’un même capteur, ce qui est notamment dû au fait que le gain du capteur dépend du lieu d’adsorption de la molécule ;
- ne représentent qu’une partie de l’information totale : toutes les molécules ne sont pas adsorbées et on ne connaît pas *a priori* le nombre de molécules qui se sont fixées sur le capteur.

Un élément crucial pour reconstruire convenablement le spectre de masse de la solution étudiée est de savoir propager les incertitudes tout au long de la chaîne de mesure. Ceci permet de surcroît de fournir une évaluation de l’incertitude sur les valeurs finales du spectre.

Approches suivies

Nous venons, au chapitre précédent, de modéliser le problème de la spectrométrie de masse à base de NEMS. Nous avons notamment identifié trois problématiques que nous allons devoir appréhender. Dans ce chapitre, nous introduisons les outils nécessaires pour pouvoir mettre en place une méthode de traitement adaptée aux éléments identifiés auparavant. Nous allons notamment nous pencher sur les problèmes inverses, qui sont une manière robuste et rigoureuse d'aborder notre problématique. Ensuite, nous verrons les outils bayésiens, déjà évoqués auparavant, qui vont nous permettre de retranscrire nos connaissances sous forme de loi de probabilité. Enfin, nous verrons une méthode d'estimation particulière, la méthode de Monte-Carlo par Chaîne de Markov, reposant sur une technique d'échantillonnage stochastique.

Sommaire de ce chapitre

4.1	L'approche problème inverse	79
4.1.1	Problèmes bien-posés et problèmes mal-posés	80
4.1.2	Méthodes d'inversion analytiques	81
4.1.2.1	Transformée de Radon	81
4.1.2.2	Déconvolution par transformée de Fourier	81
4.1.3	Méthodes d'inversion par décomposition	82
4.1.3.1	Décomposition en série	82
4.1.3.2	Décomposition sur une base	82
4.1.4	Méthodes d'inversion algébriques	83
4.1.4.1	Méthodes d'inversion reposant sur une expression matricielle ou tensorielle	83
4.1.4.2	Méthodes d'inversion reposant sur la définition et l'optimisation d'un critère	85
	Critère des moindres carrés	85
	Critère pénalisé	85
	Critère probabiliste	86
4.2	Le cadre bayésien	88
4.2.1	Principes généraux	88
4.2.1.1	Règle de Bayes	88
4.2.1.2	Choix de la loi <i>a priori</i>	88
	Conjugaison de lois <i>a priori</i>	89

	Lois <i>a priori</i> non informatives	89
	4.2.1.3 Marginalisation	90
4.2.2	Estimateurs bayésiens	90
	4.2.2.1 Expression d'une fonction de coût	90
	4.2.2.2 Estimateurs usuels	91
4.2.3	Choix de modèle	91
4.2.4	Modélisation bayésienne hiérarchique	92
4.2.5	Liens entre loi <i>a priori</i> et pénalisation	93
	Cas normal	94
	Pénalisation de modèle	94
4.2.6	Méthodes de calcul bayésien	94
4.3	Les méthodes MCMC	96
4.3.1	Principe du MCMC	96
	4.3.1.1 Méthode de Monte-Carlo	96
	4.3.1.2 Simulation d'une loi	96
4.3.2	L'algorithme de Metropolis-Hastings	97
	4.3.2.1 Description de l'algorithme	97
	4.3.2.2 Contrôle de convergence	98
4.3.3	Méthodes dérivées de Metropolis-Hastings	98
	4.3.3.1 L'algorithme de Gibbs	98
	4.3.3.2 <i>Reversible Jumps MCMC</i> et algorithme de Green	98

4.1 L'approche problème inverse

L'idée générale des problèmes inverses est de pouvoir fournir l'estimation de l'entrée d'un système en observant sa sortie. L'entrée du système peut prendre différentes formes comme nous en avons discuté dans la section 3.1.2.

Contrairement aux méthodes dites « directes », qui reposent sur un raisonnement visant à déduire une information directement à partir d'une observation, l'approche problème inverse offre un cadre de résolution du problème plus explicite et permet de prendre en considération à la fois une modélisation du système de mesure mais aussi une modélisation du signal d'entrée [Dem87, Ch. 1].

La première étape est la description du modèle direct, c'est-à-dire l'écriture des équations reliant l'entrée du système à la sortie, ce que nous avons fait tout au long du chapitre 3. La seconde étape est la mise en place de méthodes permettant l'estimation de l'entrée (« la solution ») à partir de la sortie (« l'observation »).

Les figures 4.1 et 4.2 illustrent cette démarche.

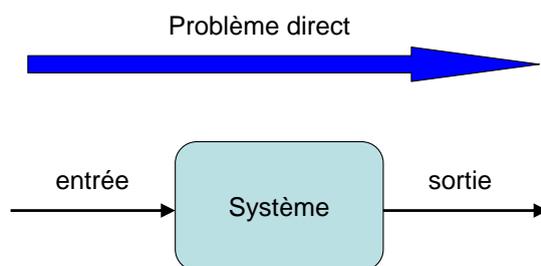


FIGURE 4.1 – Illustration du problème direct

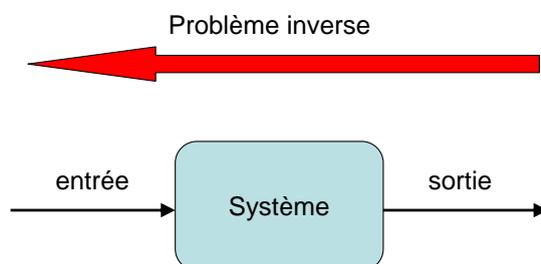


FIGURE 4.2 – Illustration du principe de l'inversion

Nous rappelons l'expression du modèle direct sous forme générale :

$$\mathcal{H}(f, g) = 0 \quad (4.1)$$

ou, si les paramètres du modèle direct sont inconnus :

$$\mathcal{H}(f, g, \theta) = 0 \quad (4.2)$$

et enfin dans le cas linéaire :

$$g = Hf \quad (4.3)$$

Dans le cas où certains paramètres sont inconnus, les problèmes inverses permettent de prendre en compte des variabilités sur ces derniers, en plus de celles sur le signal d'entrée.

Lorsque l'on rajoute au signal d'entrée inconnu des paramètres, certaines terminologies sont employées.

- Dans certains cas, l'inversion peut se faire en estimant des variables cachées (on parle d'inversion « augmentée »). Les variables cachées sont généralement des variables annexes, qui ne constituent pas l'information à retrouver, mais qui interviennent dans la mesure. Ces variables s'appellent également variables de nuisance.
- Nous pourrions avoir besoin d'estimer des paramètres du système de mesure, comme la réponse du système par exemple (inversion « myope » voire « aveugle ») lorsque le modèle direct n'est pas parfaitement connu.
- Enfin, il est possible d'introduire des informations sur l'entrée $f(t)$ ou f , et d'estimer des paramètres sur ces informations (les hyperparamètres – inversion « non-supervisée »).

4.1.1 Problèmes bien-posés et problèmes mal-posés

Au sens de Hadamard [Had02], un problème est dit « bien-posé » s'il satisfait aux trois conditions suivantes :

- la solution existe ;
- la solution est unique ;
- la solution est stable : une petite variation du signal observé doit se traduire par une petite variation sur la solution.

Les problèmes de stabilité peuvent venir des incertitudes du modèle (voir section 3.1.3), mais aussi d'une instabilité du modèle direct. Un autre problème fréquent est le manque de données, qui a tendance à rendre le problème sous-déterminé, c'est-à-dire que plusieurs solutions sont aussi valables les unes que les autres, pour un même signal mesuré (problème d'unicité). La sur-détermination (problème d'existence), quant à elle, intervient lorsque aucun signal d'entrée ne permet de retrouver les données observées, ce qui est fréquemment engendré par une perturbation sur le signal de sortie.

Pour un problème donné, si au moins une de ces conditions n'est pas satisfaite, le problème est dit « mal-posé ». Les problèmes inverses sont généralement mal-posés. Lorsqu'un problème est mal posé, il convient de changer le notion de solution, par exemple, en rajoutant de l'information sur la solution recherchée.

Il existe plusieurs méthodes de résolution de problème inverse, que nous pouvons regrouper en trois catégories [MD01] :

1. les méthodes d'inversion analytiques ;
2. les méthodes d'inversion par décomposition ;
3. les méthodes d'inversion algébriques.

Ces catégories peuvent être croisées avec le caractère déterministe ou statistique de la modélisation. Ainsi, il est possible de développer une méthode d'inversion analytique statistique, par exemple, ou encore une méthode d'inversion algébrique déterministe.

4.1.2 Méthodes d'inversion analytiques

Les méthodes d'inversion analytiques sont des méthodes pour lesquelles un opérateur inverse (au sens mathématique du terme) est recherché pour le problème direct. En clair, il est possible de trouver une opération mathématique \mathcal{I} qui inverse parfaitement le problème direct.

$$\mathcal{I}(g) = f \tag{4.4}$$

Nous illustrons ceci par les deux exemples suivants :

4.1.2.1 Transformée de Radon

Un exemple particulièrement courant est la transformée de Radon. La transformée de Radon (notée $TR(\cdot)$) est fréquemment utilisée en tomographie [Gra02] dans la mesure où c'est cette opération qui lie les éléments de l'entrée du système aux observations. On peut montrer qu'une opération dite de « rétroprojection filtrée » (notée $RF(\cdot)$) est l'opérateur inverse de la transformée de Radon, c'est-à-dire que $RF \circ TR$ est l'identité. Il suffit donc d'appliquer l'opérateur de rétro-projection filtrée à l'observation afin d'estimer l'entrée.

Ceci n'est malheureusement que théorique : l'opération de filtrage intègre une étape de dérivation rampe, combinée avec une transformée de Hilbert. Or, en pratique, les données sont bruitées, et le bruit se voit être amplifié par le filtre rampe. Ainsi l'estimation ne peut pas être parfaite. Il est néanmoins possible d'appliquer la rétro-projection filtrée, en introduisant des fenêtres d'apodisation sur le filtre rampe, par exemple, afin de régulariser l'inversion.

4.1.2.2 Déconvolution par transformée de Fourier

Dans le cadre d'un problème de déconvolution, il est possible d'exprimer le problème dans le domaine temporel :

$$g(t) = f(t) \star h(t) \tag{4.5}$$

Il est également possible d'exprimer le problème dans le domaine de Fourier. L'équation (4.5) devient, comme suit, en notant l'opérateur de la transformée de Fourier $TF(\cdot)$:

$$TF(g(t)) = TF(f(t)) TF(h(t)) \quad (4.6)$$

Ainsi, l'estimation $\widehat{f(t)}_{\text{analytique}}$ de $f(t)$ se calcule grâce à l'opérateur de la transformée de Fourier $TF(\cdot)$ ainsi que son inverse $TF^{-1}(\cdot)$ par la formule suivante :

$$\widehat{f(t)}_{\text{analytique}} = TF^{-1} \left(\frac{TF(g(t))}{TF(h(t))} \right) \quad (4.7)$$

Un premier problème numérique se pose si $TF(h(t))$ prend des valeurs nulles dans le domaine de Fourier. Ceci conduit également l'estimation à être instable : il suffit qu'une perturbation (notée $\epsilon(t)$) présente une fréquence non présente dans la réponse impulsionnelle pour que la solution tende vers l'infini :

$$\begin{aligned} g(t) &= f(t) \star h(t) + \epsilon(t) \\ \widehat{f(t)}_{\text{analytique}} &= TF^{-1} \left(\frac{TF(g(t))}{TF(h(t))} + \frac{TF(\epsilon(t))}{TF(h(t))} \right) \end{aligned} \quad (4.8)$$

Par conséquent, dans le cadre de la déconvolution, cette méthode analytique ne fonctionne pas dès lors que des perturbations affectent le signal. De plus, une telle méthode ne permet pas de prendre en compte le caractère impulsionnel de la solution recherchée ici.

Dans un état d'esprit similaire à la transformée de Fourier, nous pouvons également citer la transformée en ondelette. Sa version en temps continu se note CWT (pour *Continuous Wavelet Transform*) [Mal00].

4.1.3 Méthodes d'inversion par décomposition

4.1.3.1 Décomposition en série

Les méthodes de décomposition en série consistent à décomposer le signal observé $g(t)$ en série et à utiliser ce changement de représentation pour trouver une relation simple entre les coefficients de la décomposition et $f(t)$ ou sa décomposition.

Il est ainsi possible de reprendre l'exemple ci-dessus, dans le cas où $f(t)$ est un signal périodique, et d'effectuer une décomposition en série de Fourier.

4.1.3.2 Décomposition sur une base

Les méthodes de décomposition sur une base sont similaires aux méthodes de décomposition en série, la différence étant que les méthodes de décomposition sur une base travaillent en dimension finie.

L'idée des méthodes de décomposition sur une base peut être d'exprimer le problème sous une forme plus simple (une base bien choisie, par exemple). Aussi, ces méthodes de décomposition peuvent éventuellement projeter le signal observé dans un espace de dimension plus faible pour rendre l'estimation plus robuste.

Notons que les bases de décomposition, appelées parfois « dictionnaires », peuvent prendre différentes formes, et peuvent notamment être constituées de formes connues recherchées dans le signal. La difficulté peut alors résider dans la décomposition sur la base lorsque celle-ci n'est pas orthogonale.

Les deux exemples vus ci-dessus trouvent leur pendant avec la transformée de Fourier discrète. La transformée en ondelette existe aussi en version discrète (DWT pour *Discrete Wavelet Transform*).

Nous aborderons cette notion de décomposition sur une base pour le problème qui nous concerne dans la section suivante, 5.2.3.3.

4.1.4 Méthodes d'inversion algébriques

Une seconde catégorie de méthodes concerne les méthodes d'inversion « algébriques ». Parmi ces méthodes, on pourra distinguer deux cas :

1. les méthodes d'inversion reposant sur une inversion matricielle pour la résolution d'un système d'équations linéaires ;
2. les méthodes d'inversion reposant sur la définition et l'optimisation d'un critère.

La première étape de ces méthodes consiste à numériser le signal $g(t)$ (c'est-à-dire à l'échantillonner et à le quantifier), et de considérer le signal observé numérisé \mathbf{g} et le signal d'entrée inconnu numérisé \mathbf{f} . Observant \mathbf{g} , le but va être de mettre en place une procédure afin d'estimer \mathbf{f} .

4.1.4.1 Méthodes d'inversion reposant sur une expression matricielle ou tensorielle

Parmi les méthodes d'inversion algébriques, il existe certaines méthodes qui reposent sur une expression matricielle ou tensorielle du problème, ce qui suppose que le problème soit linéaire ou multilinéaire, et qui vont exploiter cette structure.

Pour premier exemple prenons une nouvelle fois le problème de déconvolution. Ce problème peut s'exprimer sous la forme $\mathbf{g} = \mathbf{H}\mathbf{f}$ en prenant :

$$\mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{T-1} \\ g_T \end{pmatrix} \quad \mathbf{H} = \begin{pmatrix} h_1 & 0 & \cdots & \cdots & 0 \\ h_2 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{T-1} & & \ddots & \ddots & 0 \\ h_T & h_{T-1} & \cdots & h_2 & h_1 \end{pmatrix} \quad \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{T-1} \\ f_T \end{pmatrix} \quad (4.9)$$

On dispose ainsi d'une multiplication matricielle en lieu et place de l'opération de convolution. Il est donc possible d'estimer $\widehat{\mathbf{f}}_{\text{algébrique}}$ en utilisant la matrice inverse \mathbf{H}^{-1} , si elle existe :

$$\widehat{\mathbf{f}}_{\text{algébrique}} = \mathbf{H}^{-1}\mathbf{g} \quad (4.10)$$

Si pour des raisons de dimension, ou pour des raisons de mauvais conditionnement de la matrice système \mathbf{H} , il n'est pas possible d'utiliser la matrice inverse, la matrice pseudo-inverse à gauche ou à droite \mathbf{H}^\dagger est utilisée (ou matrice d'inversion généralisée ou encore matrice de Moore-Penrose).

$$\widehat{\mathbf{f}}'_{\text{algébrique}} = \mathbf{H}^\dagger\mathbf{g} \quad (4.11)$$

avec, si le rang de \mathbf{H} est égal à son nombre de colonnes :

$$\mathbf{H}^\dagger = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \quad (4.12)$$

sinon, si le rang de \mathbf{H} est égal à son nombre de lignes :

$$\mathbf{H}^\dagger = \mathbf{H}^T(\mathbf{H}^T\mathbf{H})^{-1} \quad (4.13)$$

La forme générale consiste à décomposer \mathbf{H} en un produit de deux matrices. Si \mathbf{H} possède p_l lignes et p_c colonnes et est de rang p_0 , on peut écrire \mathbf{H} comme :

$$\mathbf{H} = \mathbf{H}_l\mathbf{H}_c \quad (4.14)$$

où \mathbf{H}_l est une matrice $p_l \times p_0$ de rang p_0 et \mathbf{H}_c est une matrice $p_0 \times p_c$ de rang p_0 .

Ainsi, la pseudo-inverse s'exprime sous la forme :

$$\mathbf{H}^\dagger = \mathbf{H}_c^T(\mathbf{H}_c\mathbf{H}_c^T)^{-1}(\mathbf{H}_l^T\mathbf{H}_l)^{-1}\mathbf{H}_l^T \quad (4.15)$$

Nous décrivons l'utilisation de la pseudo-inverse dans notre problème dans la section suivante, 5.2.3.2.

Un second exemple de méthode d'inversion matricielle serait la méthode de régression PLS (pour *Partial Least Squares*) [PBG⁺07].

Notons que dans ces méthodes, les informations sur le signal inconnu ou sur le bruit ne sont pas prises en compte.

4.1.4.2 Méthodes d'inversion reposant sur la définition et l'optimisation d'un critère

Les méthodes d'inversion que nous allons décrire dans cette section consistent en l'écriture d'une fonction appelée « critère » caractéristique de la qualité de l'estimation, puis en l'emploi d'une procédure d'optimisation, généralement en vue de minimiser cette fonction.

Il existe un grand nombre de procédures de minimisation/optimisation/recherche de solution qui ne seront pas développées dans ce document. On pourra toutefois citer les méthodes déterministes telles que la descente du gradient, du gradient conjugué, de Newton-Raphson, et les méthodes stochastiques comme le recuit simulé ou les méthodes de Monte-Carlo, que nous développerons dans la section 4.3.

La fonction critère dépend de l'entrée du système \mathbf{f} et est notée $J(\mathbf{f})$. Au plus la valeur de cette fonction est basse, au plus la qualité de l'estimation est grande.

Critère des moindres carrés Un critère fréquemment utilisé est celui des moindres carrés, qui correspond physiquement à une minimisation de l'énergie de l'erreur d'estimation en sortie, ce qui s'écrit, dans le cas d'un modèle direct linéaire :

$$\hat{\mathbf{f}}_{\text{MC}} = \arg \min_{\mathbf{f}} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 \quad (4.16)$$

La résolution de ce système donne le même estimateur de \mathbf{f} que celui utilisant la méthode de la pseudo-inverse lorsque le rang de \mathbf{H} est égal à son nombre de colonnes. La solution peut s'exprimer analytiquement en calculant le gradient du critère ∇J et en l'annulant :

$$\begin{aligned} \nabla J &= -2\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{f}) \\ \Rightarrow \nabla J = 0 &\rightarrow \hat{\mathbf{f}}_{\text{MC}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g} \end{aligned} \quad (4.17)$$

Critère pénalisé Dans la formulation précédente, les informations sur la fonction recherchée, comme le caractère impulsionnel, par exemple, ne sont pas prises en compte. Afin de prendre ces informations en considération, il est possible de regarder du côté de la « pénalisation ».

Au lieu d'utiliser comme fonction critère une simple mesure d'adéquation aux données (comme c'est le cas pour un critère des moindres carrés par exemple), à cette fonction est ajouté un terme dit de « pénalisation » qui permet de fournir des solutions plus stables.

Ainsi, en utilisant toujours un modèle linéaire associé à un terme quadratique pour l'adéquation aux données, le problème devient :

$$\hat{\mathbf{f}}_{\text{pénalisé}} = \arg \min_{\mathbf{f}} (\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \Omega(\mathbf{f})) \quad (4.18)$$

Dans cette équation, λ est un paramètre à régler (paramètre de régularisation) et $\Omega(\mathbf{f})$ est la fonction de pénalisation. De nombreuses pénalisations ont été proposées, répondant à des exigences différentes sur la forme du signal recherché.

Un exemple très simple de pénalisation est la pénalisation dite L_2 (on parle aussi de régression *ridge*). Dans ce cas, la fonction de pénalisation vaut :

$$\Omega_{L_2}(\mathbf{f}) = \|\mathbf{f}\|_2^2 \quad (4.19)$$

Cette pénalisation conduit à la solution suivante :

$$\hat{\mathbf{f}}_{L_2} = \mathbf{H}_{L_2}^\dagger \mathbf{g} \quad (4.20)$$

avec :

$$\mathbf{H}_{L_2}^\dagger = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{I})^{-1} \mathbf{H}^T \quad (4.21)$$

Cette pénalisation apporte deux choses à l'inversion. Elle permet dans un premier temps d'inverser des problèmes de minimisation dans lesquels $\mathbf{H}^T \mathbf{H}$ n'est pas inversible ou mal-conditionnée, en « chargeant » la diagonale de cette matrice. Elle permet aussi d'imposer une contrainte de minimisation d'énergie de la solution ce qui a tendance à fournir des solutions plus limitées en amplitude.

Une version plus générale de la pénalisation est donnée par Tikhonov [Tik63] :

$$\hat{\mathbf{f}}_{\text{Tikhonov}} = \arg \min_{\mathbf{f}} \left(\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_{\mathbf{W}_1}^2 \right) + \lambda \|\mathbf{f} - \mathbf{f}_0\|_{\mathbf{W}_2}^2 \quad (4.22)$$

où $\|\cdot\|_{\mathbf{W}}$ représente la distance de Mahalanobis, \mathbf{W} étant une matrice définie-positive :

$$\|\mathbf{x}\|_{\mathbf{W}}^2 = \mathbf{x}^T \mathbf{W}^{-1} \mathbf{x} \quad (4.23)$$

Dans ce cas, la solution s'écrit :

$$\hat{\mathbf{f}}_{\text{Tikhonov}} = (\mathbf{H}^T \mathbf{W}_1 \mathbf{H} + \mathbf{W}_2)^{-1} (\mathbf{H}^T \mathbf{W}_1 \mathbf{g} + \mathbf{W}_2 \mathbf{f}_0) \quad (4.24)$$

Dans le cas parcimonieux, une pénalisation fréquemment utilisée est la pénalisation utilisant une norme 1, pénalisation L_1 ou LASSO (pour *Least Absolute Shrinkage and Selection Operator*), que nous verrons plus en détails dans la section suivante 5.2.4.

Critère probabiliste Un des reproches que l'on peut faire aux méthodes par minimisation de critère est la présence d'un (voire plusieurs) paramètre(s) à régler, λ en étant l'exemple le plus visible, le choix de la pénalisation $\Omega(\cdot)$ en étant un autre.

Ce problème se retrouve aussi dans les méthodes dites directes. Généralement, ces paramètres sont appris par validation croisée, c'est-à-dire en testant différentes configurations sur des données de test et en gardant la configuration ayant donné les meilleurs résultats.

Afin de donner plus de sens à ces paramètres, et pour pouvoir construire une méthode plus explicite, il est possible de faire appel à la modélisation probabiliste. La théorie des probabilités dépasse largement le cadre de ce document et ne sera pas traitée ici. On pourra toutefois consulter [Sap06] pour de plus amples informations.

Prenons un exemple simple : supposons un modèle linéaire avec bruit additif ϵ :

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \epsilon \quad (4.25)$$

Supposons que les connaissances sur les caractéristiques du bruit permettent de le modéliser comme un bruit blanc normal de moyenne nulle et de variance σ^2 . On peut alors écrire :

$$p(\mathbf{g}|\mathbf{f}) = \mathcal{N}(\mathbf{H}\mathbf{f}|\mathbf{0}, \sigma^2\mathbf{I}) \quad (4.26)$$

Il est alors possible d'utiliser des méthodes d'inférence statistique pour estimer \mathbf{f} . Une première méthode pourrait être celle du *maximum* de vraisemblance (ou ML pour *maximum likelihood*), la vraisemblance étant la loi $p(\mathbf{g}|\mathbf{f})$.

$$\hat{\mathbf{f}}_{\text{ML}} = \arg \max_{\mathbf{f}} (p(\mathbf{g}|\mathbf{f})) \quad (4.27)$$

Dans ce cas-ci, la solution sera la même que celle obtenue avec la pseudo-inverse ou avec le critère des moindres carrés.

Afin de prendre en compte une forme particulière sur la solution recherchée, la démarche bayésienne va être privilégiée afin de mettre en place le critère à minimiser.

4.2 Le cadre bayésien

4.2.1 Principes généraux

4.2.1.1 Règle de Bayes

Nous allons désormais nous pencher sur un cadre de travail qui permet de voir les différents points évoqués ci-dessus sous un angle différent : le cadre de travail bayésien [PBP07] [Rob05]. Le cadre bayésien est un cadre d'inférence et une manière de modéliser le problème et de définir un ensemble d'estimateurs.

Le point de départ du raisonnement bayésien est de traduire toute connaissance sur une grandeur au travers d'une loi de probabilité. Ainsi, nous allons modéliser nos connaissances sur l'entrée du système \mathbf{f} par le biais d'une loi de probabilité dite loi *a priori* $p(\mathbf{f})$.

Les informations dont nous disposons sont :

- la connaissance du modèle $\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$;
- la connaissance de certaines propriétés du bruit exprimées au travers de la loi de probabilité $p(\mathbf{g}|\mathbf{f})$;
- la connaissance incomplète sur l'inconnue \mathbf{f} , décrite par une loi de probabilité *a priori*.

Ces informations vont pouvoir se combiner au travers de la règle de Bayes, pour obtenir la loi *a posteriori* $p(\mathbf{f}|\mathbf{g})$ sur la solution recherchée :

$$p(\mathbf{f}|\mathbf{g}) = \frac{p(\mathbf{g}|\mathbf{f}) p(\mathbf{f})}{p(\mathbf{g})} \quad (4.28)$$

Dans cette équation, $p(\mathbf{g})$ est appelée « loi de probabilité totale » ou « évidence » et $p(\mathbf{f}|\mathbf{g})$ est appelée « loi de probabilité *a posteriori* ». Cette loi de probabilité *a posteriori* contient à la fois l'information de cohérence de l'estimation par rapport aux données (*via* la vraisemblance $p(\mathbf{g}|\mathbf{f})$) et l'adéquation de l'estimation aux connaissances *a priori*. C'est en quelque sorte une manière rigoureuse et explicite de fusionner ces deux types d'information sur la solution. Notons que $p(\mathbf{g})$ est dans cette équation une constante de normalisation :

$$p(\mathbf{g}) = \int p(\mathbf{g}|\mathbf{f}) p(\mathbf{f}) d\mathbf{f} \quad (4.29)$$

4.2.1.2 Choix de la loi *a priori*

La grande question est le choix de la loi *a priori*. Comme l'indique Olivier Féron [F06, pp.12-13] dans sa thèse : il est quasi-systématique de posséder quelques connaissances sur la solution recherchée, et il est quasi-systématique que ces connaissances ne puissent pas être parfaitement représentées par une loi de probabilité.

Le choix de la loi *a priori* peut se faire :

- afin de retranscrire au mieux les connaissances sur la solution recherchée ;
- dans un soucis de simplification de calcul ce qui conduit à choisir des lois *a priori* dites « conjuguées » ;
- afin de n’introduire que très peu information, ce qui conduit à choisir des lois *a priori* dites « non informatives ».

Conjugaison de lois *a priori* Un choix souvent utilisé est celui de la loi *a priori* conjuguée. Cela signifie que la loi *a posteriori* est de la même famille que la loi *a priori*, ce qui va dépendre du choix de la fonction de vraisemblance.

Nous proposons de résumer dans le tableau 4.1 (extrait de [Rob05, p. 131]) les lois *a priori* conjuguées pour quelques exemples usuels.

Forme de la vraisemblance	<i>a priori</i> conjugué
Normale	Normale
Poisson	Gamma
Gamma	Gamma
Binomiale	Bêta
Normale	Normale

TABLE 4.1 – Quelques exemples de lois *a priori* conjuguées

Lois *a priori* non informatives Un choix entrepris lorsque l’utilisateur désire ne pas introduire d’information sur un paramètre inconnu est celui des lois *a priori* non informatives, qui a été décrit en profondeur dans la thèse de Pascal Szacherski [Sza13, Ch. 2.4].

Une démarche communément employée repose sur le principe de Jeffreys [Jef61, Ch. 3.10]. Ceci consiste à prendre pour loi *a priori* :

$$p(\mathbf{f}) \propto [J(\mathbf{f})]^{\frac{1}{2}} \tag{4.30}$$

avec :

$$[J(\mathbf{f})] = -\mathbb{E} \left[\frac{\partial^2 \log(p(\mathbf{g}|\mathbf{f}))}{\partial \mathbf{f}^2} \right] \tag{4.31}$$

Le problème de cette démarche vient du fait que la loi *a priori* résultante peut être impropre, c’est-à-dire ne pas satisfaire aux conditions nécessaires des lois de probabilité, et notamment l’intégrabilité. Cependant, la loi *a posteriori*, elle, peut être une vraie loi. Comme le résume Pascal Szacherski « L’utilisation des lois *a priori* impropres [...] semble dangereuse ».

4.2.1.3 Marginalisation

Un autre outil très intéressant offert par le cadre de travail bayésien est celui de la marginalisation. Ceci consiste à faire « disparaître » un paramètre inconnu en prenant en compte l'incertitude qu'il apporte sur les autres paramètres. Lorsque le paramètre \mathbf{z} est marginalisé, ceci s'écrit :

$$p(\mathbf{f}|\mathbf{g}) = \int p(\mathbf{f}, \mathbf{z}|\mathbf{g}) d\mathbf{z} \quad (4.32)$$

4.2.2 Estimateurs bayésiens

Comme exprimé précédemment, le cadre de travail bayésien permet l'écriture d'une loi *a posteriori*, qui représente notre connaissance sur l'inconnue \mathbf{f} lorsque nous avons pris en compte les informations *a priori* et les informations apportées par les données \mathbf{g} . Ce cadre de travail permet également de résumer cette information riche de loi *a posteriori* au travers d'un estimateur dont la mise en place est explicite.

4.2.2.1 Expression d'une fonction de coût

La démarche bayésienne consiste ainsi à exprimer une fonction de coût $C(\mathbf{f}, \mathbf{f}_0)$, où \mathbf{f}_0 est la vraie valeur de l'entrée. Cette fonction de coût est caractéristique de la gravité d'une erreur d'estimation.

L'estimateur bayésien est celui qui minimise le coût moyen, c'est-à-dire :

$$\hat{\mathbf{f}}_{\text{Bayes}} = \arg \min_{\mathbf{f}} (\mathbb{E} [C(\mathbf{f}, \mathbf{f}_0)]) \quad (4.33)$$

La fonction de coût la plus simple serait la fonction de Dirac appelée « coût 0-1 » :

$$C_{0-1}(\mathbf{f}, \mathbf{f}_0) = \delta(\mathbf{f} - \mathbf{f}_0) \quad (4.34)$$

Une seconde fonction de coût serait une fonction quadratique appelée « coût quadratique » :

$$C_{\text{quadratique}}(\mathbf{f}, \mathbf{f}_0) = \|\mathbf{f} - \mathbf{f}_0\|_2^2 \quad (4.35)$$

Notons que l'estimateur bayésien est celui qui minimise l'espérance du coût moyen, on parle alors d'estimateur optimal. Ceci n'est vrai que si le paramètre est effectivement distribué sous la loi *a posteriori*. Par conséquent, si l'optimalité est assurée théoriquement, dans la pratique, l'estimateur bayésien peut ne pas être optimal si les lois *a priori* sont incohérentes par rapport au problème étudié.

4.2.2.2 Estimateurs usuels

L'utilisation de la fonction de coût de l'équation (4.34) conduit à une estimation dite « *Maximum A Posteriori* » (ou MAP) :

$$\hat{\mathbf{f}}_{\text{MAP}} = \arg \max_{\mathbf{f}} (p(\mathbf{f}|\mathbf{g})) \quad (4.36)$$

De la même manière, l'utilisation de la fonction de coût de l'équation (4.35) conduit à une estimation dite « *Espérance A Posteriori* » (ou EAP) :

$$\hat{\mathbf{f}}_{\text{EAP}} = \mathbb{E}[\mathbf{f}|\mathbf{g}] = \int \mathbf{f} p(\mathbf{f}|\mathbf{g}) d\mathbf{f} \quad (4.37)$$

4.2.3 Choix de modèle

Le cadre de travail bayésien permet également de comparer des modèles entre eux et de fournir des probabilités de modèle sachant les données [Rob05, Ch. 5]. Ce modèle peut être, par exemple, le nombre de variables inconnues dans l'entrée du système.

Écrivons un modèle \mathcal{M} dans lequel interviennent des paramètres $\boldsymbol{\theta}$. Alors, selon le principe de marginalisation :

$$\begin{aligned} p(\mathcal{M}|\mathbf{g}) &\propto p(\mathcal{M}) p(\mathbf{g}|\mathcal{M}) \\ &\propto p(\mathcal{M}) \int p(\mathbf{g}|\boldsymbol{\theta}, \mathcal{M}) d\boldsymbol{\theta} \\ &\propto p(\mathcal{M}) \int p(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathcal{M}) d\boldsymbol{\theta} \end{aligned} \quad (4.38)$$

Ainsi, lorsque l'on cherche à comparer deux modèles \mathcal{M}_1 et \mathcal{M}_2 , il vient :

$$\begin{aligned} \frac{p(\mathcal{M}_1|\mathbf{g})}{p(\mathcal{M}_2|\mathbf{g})} &= \frac{p(\mathcal{M}_1) \int p(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathcal{M}_1) d\boldsymbol{\theta}}{p(\mathcal{M}_2) \int p(\mathbf{g}|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathcal{M}_2) d\boldsymbol{\theta}} \\ &= \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \text{BF}_{1,2} \end{aligned} \quad (4.39)$$

La quantité $\text{BF}_{1,2}$ est appelée facteur de Bayes. C'est ce facteur qui détermine quel modèle explique le plus convenablement les données observées.

Les termes *a priori* $p(\mathcal{M}_1)$ et $p(\mathcal{M}_2)$ vont permettre d'introduire une information *a priori* sur les modèles, en pénalisant les modèles les plus complexes par exemple (on parle alors de « rasoir d'Ockham »), ce que nous verrons plus loin dans la section 4.2.5.

Dans le cas de la détection d'une maladie, par exemple, avec un modèle « malade » et un modèle « sain », ces termes peuvent symboliser la prévalence de la maladie dans la population étudiée [Sza13].

Deux démarches de choix de modèle se démarquent. La première consiste à retenir le modèle le plus probable. On parle alors de *Maximum A Posteriori* Marginalisé (ou MMAP pour *Marginalized Maximum A Posteriori*).

$$\widehat{\mathcal{M}}_{\text{MMAP}} = \max_{\mathcal{M}} (p(\mathcal{M}|g)) \quad (4.40)$$

Une seconde approche consiste à moyenner les différents modèles (on parle alors de *Model Averaging*) [RC04, Ch. 11].

4.2.4 Modélisation bayésienne hiérarchique

Nous avons vu dans la section 3.3.2.2 que le modèle direct peut se décrire sous la forme d'un graphe acyclique orienté relatant une dépendance hiérarchique entre les paramètres. Le terme généralement utilisé est celui de « réseaux bayésiens ». Le point essentiel des réseaux bayésiens est le principe d'indépendance conditionnelle. Pour illustrer ceci, prenons un exemple simple, illustré par la figure 4.3.

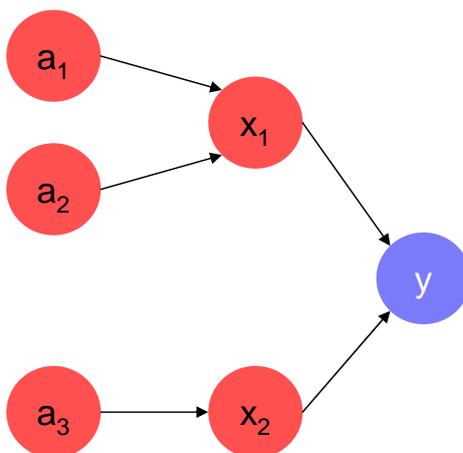


FIGURE 4.3 – Exemple de modèle hiérarchique

Nous supposons que x_1 et x_2 sont indépendants, ainsi que a_1 et a_2 . Nous modélisons les connaissances que nous avons sur le modèle par le principe d'indépendance conditionnelle :

$$\begin{aligned}
 p(y|x_1, x_2, a_1, a_2, a_3) &= p(y|x_1, x_2) \\
 p(x_1|y, x_2, a_1, a_2, a_3) &= p(x_1|y, x_2, a_1, a_2) \\
 p(x_2|y, x_1, a_1, a_2, a_3) &= p(x_2|y, x_1, a_3) \\
 p(a_1|y, x_1, x_2, a_2, a_3) &= p(a_1|x_1, a_2) \\
 p(a_2|y, x_1, x_2, a_1, a_3) &= p(a_2|x_1, a_1) \\
 p(a_3|y, x_1, x_2, a_1, a_2) &= p(a_3|x_2)
 \end{aligned} \tag{4.41}$$

La démonstration se trouve en annexe E.1.

Ceci signifie que pour une variable donnée, la loi *a posteriori* conditionnelle aux autres paramètres ne dépend que des paramètres à laquelle la variable est directement connectée et les variables du même niveau hiérarchique. Par conséquent, retranscrire nos connaissances au travers d'un modèle hiérarchique va permettre une simplification importante des calculs dans le cadre bayésien.

De plus, le cadre bayésien permet de fournir une interprétation de la relation hiérarchique entre les paramètres. En prenant l'exemple de la deuxième équation de (4.41), il est possible d'écrire [Rob05, Ch. 10] :

$$p(x_1|y, x_2, a_1, a_2, a_3) \propto p(y|x_1, x_2) p(x_1|a_1, a_2) \tag{4.42}$$

D'une certaine manière, on peut considérer que cette loi *a posteriori* conditionnelle est constituée du produit d'un terme pouvant être assimilé à une vraisemblance $p(y|x_1, x_2)$ et d'un terme pouvant être assimilé à un *a priori* $p(x_1|a_1, a_2)$.

La loi jointe, quant à elle, s'écrit :

$$p(y, x_1, x_2, a_1, a_2, a_3) = p(y|x_1, x_2) p(x_1|a_1, a_2) p(x_1|a_3) p(a_1) p(a_2) p(a_3) \tag{4.43}$$

4.2.5 Liens entre loi *a priori* et pénalisation

Nous avons vu précédemment que l'écriture d'un estimateur de type *Maximum A Posteriori* conduit à une maximisation :

$$\begin{aligned}
 \hat{\mathbf{f}}_{\text{MAP}} &= \arg \max_{\mathbf{f}} (p(\mathbf{f}|\mathbf{g})) \\
 &= \arg \max_{\mathbf{f}} (p(\mathbf{g}|\mathbf{f}) p(\mathbf{f})) \\
 &= \arg \min_{\mathbf{f}} (-\log(p(\mathbf{g}|\mathbf{f})) - \log(p(\mathbf{f})))
 \end{aligned} \tag{4.44}$$

On retrouve la forme évoquée dans l'équation (4.18). Le terme $-\log(p(\mathbf{f}))$ fait office ici de pénalisation $\lambda \Omega(\mathbf{f})$.

Par conséquent, la démarche bayésienne encadre la démarche de pénalisation de critère. De plus, comme il est possible de donner du sens aux *a priori*, il sera plus facile de régler la fonction de pénalisation ainsi que le paramètre de régularisation.

Cas normal Prenons l'exemple (courant) d'une vraisemblance normale (centrée de matrice de covariance $\sigma^2 \mathbf{I}$). Il est possible de ré-écrire l'équation 4.44 :

$$\hat{\mathbf{f}}_{\text{MAP}} = \arg \min_{\mathbf{f}} (\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 - \sigma^2 \log(p(\mathbf{f}))) \quad (4.45)$$

Pénalisation de modèle Nous avons vu dans la section 4.2.3 que le cadre de travail bayésien permettait un choix de modèle. Le rapport des probabilités *a posteriori* des modèles s'écrit comme le produit d'un rapport de lois *a priori* sur le modèle et d'un terme de vraisemblance marginalisée, le facteur de Bayes.

Compte-tenu du lien qui existe entre pénalisation et loi *a priori*, il est possible de se poser la question sur les lois *a priori* sur le modèle. Comme vu précédemment à la section 4.2.3, la démarche généralement employée est celle du rasoir d'Ockham, qui consiste à favoriser les modèles les plus simples. Ainsi, il est fréquent de pénaliser le modèle par sa complexité.

$$\begin{aligned} \log \left(\frac{p(\mathcal{M}_1|\mathbf{g})}{p(\mathcal{M}_2|\mathbf{g})} \right) &= \log \left(\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \text{BF}_{1,2} \right) \\ &= \Omega(\mathcal{M}_1) - \Omega(\mathcal{M}_2) + \log(\text{BF}_{1,2}) \end{aligned} \quad (4.46)$$

$\Omega(\mathcal{M}_1)$ symbolise la pénalisation sur le modèle. On peut citer comme exemple de pénalisation l'AIC (pour *Akaike Information Criterion*) [Shi76], le BIC (pour *Bayesian Information Criterion*) [BA04] ou encore la démarche dite du Cp de Mallows [KS90] qui fait intervenir une pénalisation.

La vision pénalisation est souvent reliée aux démarches de régression. Les bayésiens ont plutôt tendance à s'exprimer en terme de loi *a priori* sur le modèle.

4.2.6 Méthodes de calcul bayésien

Une fois que la loi *a posteriori* est écrite et que l'estimateur bayésien l'est également, il reste à calculer la solution.

Une première méthode consiste à calculer analytiquement la solution. Dans de nombreux cas, ceci n'est pas possible. L'emploi de lois *a priori* conjuguées peut permettre de mener des calculs de manière analytique.

Une seconde méthode consiste à évaluer en plusieurs points la valeur de la loi *a posteriori*. Si le calcul sur une grille fixe peut s'avérer longue et inefficace, les méthodes d'échantillonnage stochastique sont très en vogue. A ce titre, nous verrons les méthodes de Monte-Carlo dans le prochain chapitre.

Dans le cas d'une estimation au sens du *Maximum A Posteriori*, il est possible d'optimiser le critère bayésien grâce à un algorithme déterministe ou stochastique.

Il est également possible d'approximer la loi *a posteriori* par une loi normale et de calculer sa moyenne et sa variance (en d'autres termes, ses deux premiers moments statistiques) pour déterminer la loi. Ce type de démarche est employé dans [OSBE13] par exemple et permet de simplifier le calcul car elle ne nécessite que l'estimation des deux premiers moments statistiques de la loi *a posteriori*. Il est possible de retrouver cette approximation sous le terme « approximation de Laplace », terme qui peut être néanmoins ambigu par rapport à une autre méthode utilisée en statistique [Mur12, Ch. 8.4.1].

Une dernière catégorie de méthode concerne les méthodes bayésiennes variationnelles [ŠQ10]. Ces méthodes consistent à approximer la loi *a posteriori* à l'aide d'une loi de probabilité séparable. En clair, si \mathbf{f} est un vecteur à N paramètres :

$$p(\mathbf{f}|\mathbf{g}) \approx \prod_{i=1}^N \tilde{p}_i(f_i|\mathbf{g}) \quad (4.47)$$

L'approximation entre les deux lois se fait au sens de la divergence de Kullback-Leibler, notée $\text{KL}(p, q)$ et qui vaut :

$$\text{KL}(p, q) = \int p(x) \frac{p(x)}{q(x)} dx \quad (4.48)$$

L'équation résolvant la minimisation de cette divergence est :

$$\tilde{p}_i(f_i|\mathbf{g}) = \exp \left(\int \log(p(\mathbf{f}|\mathbf{g})) \prod_{j=1, j \neq i}^N \tilde{p}_j(\mathbf{f}_j, \mathbf{g}) d\mathbf{f}_j \right) \quad (4.49)$$

Soit ce système d'équation présente une solution analytique, soit il est possible d'emprunter la solution de l'algorithme *iterative VB* consistant à résoudre itérativement l'équation (4.49) pour chacun des indices.

4.3 Les méthodes MCMC

4.3.1 Principe du MCMC

Nous allons nous intéresser désormais à une méthode permettant de calculer des intégrales, en particulier dans un espace de grande dimension. Une telle méthode permet notamment de calculer un estimateur de type *Esperance A Posteriori* (équation (4.37)), ou de marginaliser certains paramètres (équation (4.32)). Cette méthode, appelée Monte-Carlo par chaîne de Markov (ou MCMC pour *Markov-Chain Monte-Carlo*), consiste à générer un ensemble d'échantillons distribués sous la loi *a posteriori*.

4.3.1.1 Méthode de Monte-Carlo

L'idée générale des méthodes de Monte-Carlo est de permettre d'intégrer une fonction *via* une méthode stochastique. Dans le cas du calcul d'une espérance, la méthode de Monte-Carlo s'écrit de la manière suivante :

$$\begin{aligned}\hat{\mathbf{f}}_{\text{EAP}} &= \mathbb{E}[\mathbf{f}|\mathbf{g}] = \int \mathbf{f} p(\mathbf{f}|\mathbf{g}) d\mathbf{f} \\ &= \frac{1}{J} \sum_{j=1}^J \mathbf{f}^{(j)} \text{ si } \mathbf{f}^{(j)} \sim p(\mathbf{f}|\mathbf{g})\end{aligned}\quad (4.50)$$

Par la méthode de Monte-Carlo, on transforme un problème de calcul d'intégrale en un problème de simulation de loi de probabilité.

4.3.1.2 Simulation d'une loi

Il peut arriver que la simulation de la loi *a posteriori* $p(\mathbf{f}|\mathbf{g})$ ne soit pas directe, notamment dans le cadre d'un modèle bayésien multidimensionnel comportant des lois *a priori* non conjuguées. Il est alors possible de faire appel à certaines méthodes de simulation.

Une des méthodes les plus simples serait celle de la fonction de répartition inverse. Une autre méthode est la méthode dite d'acceptation-rejet. Ici, nous allons nous pencher sur la simulation d'une loi par une chaîne de Markov.

Une chaîne de Markov est un ensemble d'éléments $\mathbf{f}^{(j)}$ tels que :

$$p(\mathbf{f}^{(j)}|\mathbf{f}^{(j-1)}, \mathbf{f}^{(j-2)}, \dots, \mathbf{f}^{(2)}, \mathbf{f}^{(1)}) = p(\mathbf{f}^{(j)}|\mathbf{f}^{(j-1)})\quad (4.51)$$

Sous certaines conditions, une chaîne de Markov peut converger vers une loi stationnaire. La stratégie MCMC consiste justement à faire en sorte que la chaîne de Markov converge et que la

loi stationnaire soit la loi à simuler. Si on suppose qu'après B échantillons dits « de chauffe » la chaîne de Markov $\mathbf{f}^{(j)}$ est distribuée sous $p(\mathbf{f}|\mathbf{g})$, alors on obtient :

$$\hat{\mathbf{f}}_{\text{EAP MCMC}} = \frac{1}{J} \sum_{j=1+B}^{J+B} \mathbf{f}^{(j)} \quad (4.52)$$

4.3.2 L'algorithme de Metropolis-Hastings

4.3.2.1 Description de l'algorithme

Nous avons vu précédemment que la stratégie MCMC permettait de passer d'un problème de calcul d'intégrale à un problème de génération d'une chaîne de Markov convergeant vers une loi stationnaire donnée.

Un algorithme permettant cela est l'algorithme de Metropolis [MRR⁺53] qui a été complété par Hastings [Has70] pour donner l'algorithme de Metropolis-Hastings.

Cet algorithme repose sur la génération d'un nouvel échantillon potentiel qui pourra être inclus dans la chaîne de Markov. A partir de cet échantillon, une probabilité d'acceptation est calculée puis l'échantillon en question est soit accepté soit rejeté.

Pour simuler la loi $\pi(\mathbf{f})$, l'algorithme consiste en un ensemble d'itérations sur le paramètre j qui s'écrivent chacune de cette manière :

- Générer \mathbf{f}^p selon $Q_j(\mathbf{f}^p|\mathbf{f}^{(j-1)})$: $Q_j(\cdot, \cdot)$ étant la loi de proposition pour cette itération.
- Calculer la probabilité d'acceptation α , voir ci-dessous.
- Prendre $\mathbf{f}^{(j)} = \mathbf{f}^p$ avec la probabilité α ou prendre $\mathbf{f}^{(j)} = \mathbf{f}^{(j-1)}$ sinon.

La probabilité d'acceptation α se calcule comme suit :

$$\alpha = \min \left(1, \frac{\pi(\mathbf{f}^p)}{\pi(\mathbf{f}^{(j)})} \frac{Q_j(\mathbf{f}^{(j-1)}|\mathbf{f}^p)}{Q_j(\mathbf{f}^p|\mathbf{f}^{(j-1)})} \right) \quad (4.53)$$

D'une certaine manière, cette équation traduit un équilibre entre le fait d'aller explorer des zones de probabilité forte de la loi *a posteriori* (terme $\frac{\pi(\mathbf{f}^p)}{\pi(\mathbf{f}^{(j)})}$) et le fait de ré-équilibrer cette exploration par la probabilité d'exploration de la zone, une zone ayant moins de chance d'être visitée étant favorisée (terme $\frac{Q_j(\mathbf{f}^{(j-1)}|\mathbf{f}^p)}{Q_j(\mathbf{f}^p|\mathbf{f}^{(j-1)})}$).

Si $Q_j(\mathbf{f}^p|\mathbf{f}^{(j-1)}) = Q_j(\mathbf{f}^p)$, on parle alors de « Metropolis Hastings indépendant », à opposer au « Metropolis Hastings à marche aléatoire ». Aussi l'algorithme de Metropolis initial supposait $Q_j(\mathbf{f}^p|\mathbf{f}^{(j-1)}) = Q_j(\mathbf{f}^{(j-1)}|\mathbf{f}^p)$ (loi de proposition symétrique).

4.3.2.2 Contrôle de convergence

Une question importante en ce qui concerne l'utilisation de la méthode d'estimation MCMC concerne le contrôle de convergence. En effet, nous avons supposé que la chaîne de Markov admettait la loi *a posteriori* comme loi stationnaire, mais encore faut-il s'assurer que la chaîne ait atteint cette stationnarité.

Le lecteur intéressé trouvera dans [RC04, 6.6] de nombreuses informations sur la convergence des chaînes MCMC. Une solution très simple consiste à tester l'hypothèse d'ergodicité, c'est-à-dire à vérifier que la variance au sein de la chaîne de Markov est comparable à la variance entre plusieurs chaînes de Markov lancées en parallèle.

4.3.3 Méthodes dérivées de Metropolis-Hastings

Plusieurs algorithmes ont été dérivés de l'algorithme de Metropolis-Hastings. Nous allons voir :

1. l'algorithme de Gibbs ;
2. l'algorithme de Green.

4.3.3.1 L'algorithme de Gibbs

Un des dérivés les plus connus de Metropolis-Hastings est l'algorithme de Gibbs, échantillonneur de Gibbs, ou encore structure de Gibbs.

Dans ce cas, autant de lois de propositions qu'il y a de paramètres inconnus sont mises en place, ces lois de proposition étant utilisées à tour de rôle. Chacune de ces lois de proposition sont constituées par l'échantillonnage sous les lois *a posteriori* d'un paramètre inconnu (dans notre cas, par exemple, ce pourraient être les composantes de \mathbf{f}) conditionnellement aux autres paramètres. Ainsi, toutes les variables sont séquentiellement échantillonnées.

Si on note $\mathbf{f}_{i_0}^p$ la proposition pour le paramètre échantillonné, $\mathbf{f}_{i_0}^{(j-1)}$ la valeur courante du paramètre échantillonné et \mathbf{f}_{-i_0} les valeurs des autres paramètres, alors :

$$Q_{\text{Gibbs}}(\mathbf{f}_{i_0}^p) = \pi(\mathbf{f}_{i_0}^p | \mathbf{f}_{-i_0}, \mathbf{g}) \quad (4.54)$$

Dans de telles conditions, la probabilité d'acceptation α vaut systématiquement 1. La démonstration se trouve en annexe E.2.

Si la loi *a posteriori* d'un paramètre ne peut pas être directement échantillonnée, il peut être fait recours à l'algorithme de Metropolis-Hastings (on parle alors d'algorithme de Gibbs hybride).

4.3.3.2 Reversible Jumps MCMC et algorithme de Green

Un autre algorithme dit « MCMC à sauts réversibles » (ou RJMCMC pour *Reversible Jumps MCMC*) a été mis en place, pour faire face à un problème à dimension variable. Ici, il s'agit de se ramener pour une itération donnée à un problème à dimension fixe en complétant artificiellement

les espaces pour x et y dans $Q_j(y, x)$ afin qu'ils soient de même dimension et comparables. Un exemple d'algorithme MCMC est donné dans [RC04, Ch. 11, Alg. A.48] au travers de l'algorithme de Green.

Pour passer d'un modèle \mathcal{M}_m de paramètres θ_m à un modèle \mathcal{M}_n de paramètres θ_n , l'idée est de compléter les espaces de paramètres par u_m d'une part et par v_n d'autre part, puis de trouver une bijection \mathcal{T} telle que :

$$(\theta_n, v_n) = \mathcal{T}(\theta_m, u_m) \quad (4.55)$$

Ainsi, l'algorithme va pouvoir passer d'un modèle à l'autre d'une manière très similaire à ce qui se passe lors d'une étape de Metropolis-Hastings. A l'itération j , la procédure pour passer de \mathcal{M}_m , avec des paramètres $\theta_m^{(j)}$ à \mathcal{M}_n est :

- sélectionner le modèle n avec la probabilité $\pi_{m,n}$;
- générer u_m^p selon une loi de probabilité $Q_{m,n}(u_m^p)$ définie à l'avance ;
- générer (θ_n^p, v_n^p) à partir de $\theta_m^{(j)}$ et u_m^p grâce à (4.55) ;
- calculer la probabilité d'acceptation α , voir ci-dessous ;
- sélectionner le modèle \mathcal{M}_n et prendre θ_n^p pour paramètres avec la probabilité α ou garder \mathcal{M}_m avec les paramètres $\theta_m^{(j)}$ sinon.

La probabilité d'acceptation α se calcule comme suit :

$$\alpha = \min \left(1, \frac{\pi(\mathcal{M}_n, \theta_n)}{\pi(\mathcal{M}_m, \theta_m)} \frac{Q_{n,m}(v_n^p)}{Q_{m,n}(u_m^p)} \frac{\pi_{n,m}}{\pi_{m,n}} \left| \frac{\partial \mathcal{T}(\theta_n, v_n^p)}{\partial(\theta_m, u_m^p)} \right| \right) \quad (4.56)$$

où $\pi(\mathcal{M}_m, \theta_m)$ est la probabilité *a posteriori* jointe du modèle et de ses paramètres.

Quatrième partie

Méthodologies développées

Traitement mono-mode – Déconvolution impulsionnelle

Dans ce chapitre, nous allons voir le sujet de la déconvolution impulsionnelle, qui constitue une première contribution de notre travail. Dans un premier temps, nous introduirons le cadre de travail, puis nous établirons un état de l'art sur le sujet. Ensuite, nous développerons une méthode basée sur un modèle Bernoulli avant d'en illustrer les apports en termes de résultats.

Sommaire de ce chapitre

5.1	Cadre de travail	105
5.2	Méthodes de référence	107
5.2.1	État de l'art – la méthode publiée par Caltech	107
5.2.2	Méthodes directes adaptées	107
5.2.2.1	Détection de rupture	107
	Méthodes reposant sur un test statistique	107
	Algorithme CUSUM	108
5.2.2.2	Détection de pics	109
5.2.2.3	Apprentissage	110
5.2.2.4	Discussion sur les méthodes directes	111
5.2.3	Approches problème inverse adaptées	111
5.2.3.1	Approche filtrage et théorie de la détection	111
5.2.3.2	Pseudo-inverse et filtre de Haar	112
5.2.3.3	Décomposition du signal sur une base	113
	Ondelettes	113
	<i>Matching-Pursuit</i> et <i>Orthogonal Matching Pursuit</i>	114
5.2.4	Approche pénalisation et approche bayésienne	114
5.2.4.1	Pénalisations associées à la « parcimonie »	115
5.2.4.2	Lois à queue lourde	117
5.2.4.3	Modèles de mélange	119
5.3	Développement d'une méthode d'inversion probabiliste reposant sur un choix de modèle	122
5.3.1	Justification de l'approche employée	122
5.3.1.1	Mise en évidence du caractère mal-posé du problème	122

5.3.1.2	Dérivation du modèle Bernoulli-Gaussien	122
	État de l'art sur le modèle Bernoulli-Gaussien	123
	Loi <i>a priori</i> permettant d'assurer la positivité des amplitudes de chute	124
	Permettre des instants d'adsorption non-entiers	124
	Estimer les hyperparamètres	124
5.3.2	Modélisation statistique des grandeurs	124
5.3.2.1	Expression de la fonction de vraisemblance	125
5.3.2.2	Définition des lois <i>a priori</i>	125
	Loi de probabilité sur l'ordre du modèle	125
	Loi de probabilité sur les temps d'adsorptions	125
	Loi de probabilité sur les amplitudes de chute	126
	Loi de probabilité sur l'intensité de comptage	126
5.3.2.3	Écriture des lois <i>a posteriori</i> des paramètres de la liste pour un ordre de modèle donné	126
5.3.2.4	Écriture des autres lois <i>a posteriori</i> conditionnelles	126
5.3.3	Estimation	127
5.3.3.1	Utilisation d'un algorithme hybride RJMCMC MCMC	127
	Étape RJMCMC	127
	Étape MCMC	128
5.3.3.2	Le système de vote	129
	Travail de Vincent Mazet	130
	Solution proposée	130
5.4	Résultats	132
5.4.1	Définitions des critères de performance	132
	Calcul du taux de bonne et de fausse détection	132
	Compromis et courbe ROC	132
5.4.2	Résultats sur données simulées	133
5.4.2.1	Modalités de simulation	133
5.4.2.2	Paramètres de l'algorithme	133
5.4.2.3	Algorithmes de référence	134
	Seuillage	134
	Pseudo-inverse	134
	LASSO	134
	<i>Matching Pursuit</i>	135
	Bernoulli-Gaussien et SMLR	135
5.4.2.4	Résultats avec instants d'adsorption entiers	135
5.4.2.5	Résultats avec instants d'adsorption non-entiers	136
5.4.3	Validation de la méthode sur données réelles	138
5.4.4	Discussion	138

5.1 Cadre de travail

Le cadre de travail dans lequel nous allons nous inscrire sera le cadre de la déconvolution impulsionnelle. Ceci correspond à la partie « Détecter » identifiée dans la section 3.5. Généralement, la déconvolution impulsionnelle comporte en plus une étape d'estimation (on parle de « détection estimation »). Ici, nous ne nous intéresserons qu'à la détection.

Dans ce cadre-ci, il s'agit d'inverser l'opération de convolution que l'on retrouve dans l'équation (3.11), c'est-à-dire de prendre en compte l'effet de la boucle de lecture pour détecter les chutes de fréquences correspondant à l'adsorption de molécules uniques.

Nous travaillerons avec les hypothèses suivantes :

- Une seule harmonique est observée. Ainsi, la seule grandeur estimable est la chute de fréquence engendrée par une adsorption (il n'est pas possible de remonter à la masse de la molécule). Nous chercherons donc à estimer $f_k(t)$, que nous noterons, par soucis de simplicité $f(t)$, car nous situons ce chapitre dans un contexte mono-mode.
- Nous n'exploiterons qu'une seule harmonique à la fois : nous ne chercherons pas dans cette partie à utiliser l'information de synchronisation des chutes de fréquence entre les harmoniques.
- La réponse du système à une adsorption $u_{\text{PLL}}(t)$ est connue. Cette hypothèse est cohérente car ces paramètres sont réglés par l'utilisateur au cours de l'expérience. Le paramètre dont dépend la réponse est le temps de réponse τ .
- La loi de probabilité du bruit est connue. En effet, pour chaque session expérimentale, des signaux de bruit purs ont été tirés. Nous considérons donc que les paramètres de bruit peuvent être appris sur ces signaux, ou à défaut sur des zones sans adsorption au début de certains signaux expérimentaux. Ces paramètres seront les coefficients du filtre blanchisseur \mathbf{a}_1 .

Nous décomposerons le signal temporel $f(t)$ comme suit :

$$f(t) = \sum_{i=1}^N f_i \delta(t_i) \quad (5.1)$$

avec N le nombre d'adsorptions sur la fenêtre d'observation. A partir de $f(t)$, il est possible de reconstruire g par convolution et échantillonnage. Nous nous plaçons donc dans le cadre des processus ponctuels marqués (voir section 3.3.1.1).

Nous pourrions également travailler dans le contexte des processus à temps discret (voir section 3.3.1.2) pour comparer notre algorithme à d'autres méthodes. Dans ce cas, le signal pourra être discrétisé sous sa forme \mathbf{f} .

Nous introduirons également le paramètre π , inconnu, qui représente l'intensité de comptage (ou densité temporelle d'adsorptions).

Nous représentons les connaissances *a priori* sur le bruit par le jeu de paramètres θ_ϵ , celles sur π par θ_π , celles sur les f_i par θ_f . Le nombre de molécules N dépend de π et de T , et les temps t_i ne dépendent que de T .

Ce cadre de travail peut se représenter au travers de la modélisation hiérarchique illustrée par la figure 5.1.

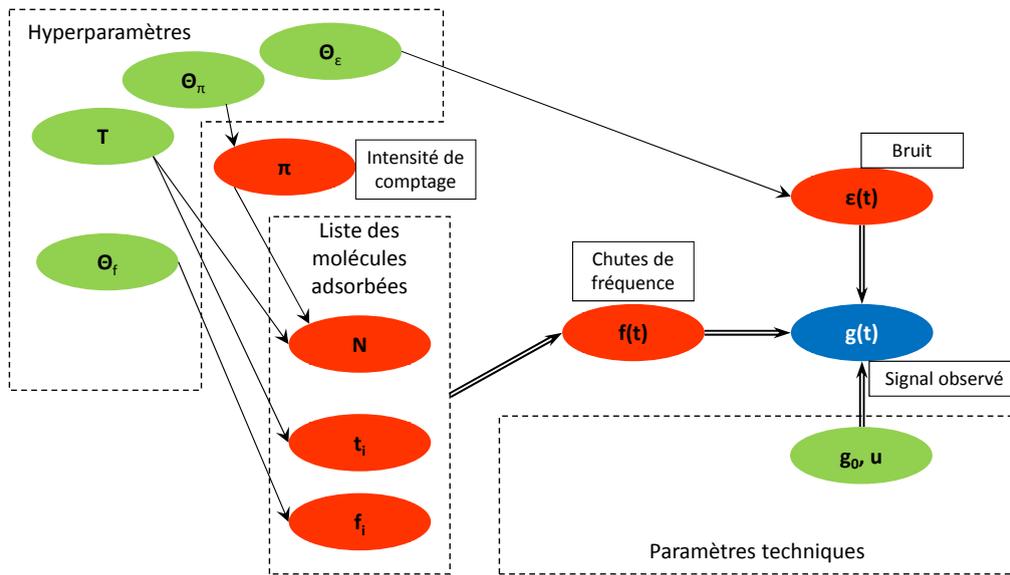


FIGURE 5.1 – Modélisation hiérarchique du problème de détection

5.2 Méthodes de référence

5.2.1 État de l'art – la méthode publiée par Caltech

Si l'on restreint l'état de l'art aux seules méthodes proposées dans le cadre de la spectrométrie de masse à base de NEMS, la seule méthode de détection publiée que nous ayons trouvée est celle proposée par l'équipe de Caltech et décrite dans les articles de Naik *et al.* [NHH⁺09], ainsi que Hanay *et al.* [HKN⁺12] que nous avons déjà évoqués auparavant. Pour les futurs chapitres, nous ferons référence à ces deux articles comme état de l'art. Notons que cette équipe travaille en collaboration avec les équipes du CEA-Leti au travers de l'alliance NanoVLSI.

En termes de détection, cette méthode consiste à :

- dériver le signal ;
- estimer l'écart-type du bruit sur le signal dérivé ;
- seuiller le signal à deux fois l'écart-type du bruit estimé.

L'ensemble des points non nuls après seuillage constitue les instants d'adsorption estimés. Reste ensuite à quantifier le décalage en fréquence (et donc les masses) pour chacune des adsorptions, ce qui est fait par une procédure reposant sur l'algorithme des moindres carrés.

5.2.2 Méthodes directes adaptées

Au-delà de la méthode de seuillage proposée par l'équipe de Caltech, il est possible de trouver d'autres méthodes directes adaptées au problème. Ces méthodes directes ne reposent pas sur une analyse du système de mesure mais sur la simple observation du signal de sortie, que l'on cherche à relier à une information utile.

Parmi ces méthodes, nous formons trois sous-ensembles :

1. les méthodes axées « Détection de rupture » ;
2. les méthodes axées « Détection de pics » ;
3. les méthodes par « Apprentissage ».

5.2.2.1 Détection de rupture

Les premières méthodes qui peuvent être mises en place afin d'estimer l'ensemble des masses adsorbées sur le capteur se basent sur le fait que le signal présente des discontinuités qui correspondent aux instants de détection (dans l'hypothèse où l'on ne dérive pas le signal). L'idée première serait alors de détecter ces ruptures dans le signal. On retrouve très fréquemment ce type de problématique en génomique [XGB07, CW09] (ou en climatologie [MWJ05, RCW⁺07]). La figure 5.2 illustre l'idée de la détection de rupture.

Méthodes reposant sur un test statistique Une première catégorie de méthodes concerne les démarches axées sur la détection de rupture par le biais d'un test statistique. Généralement,

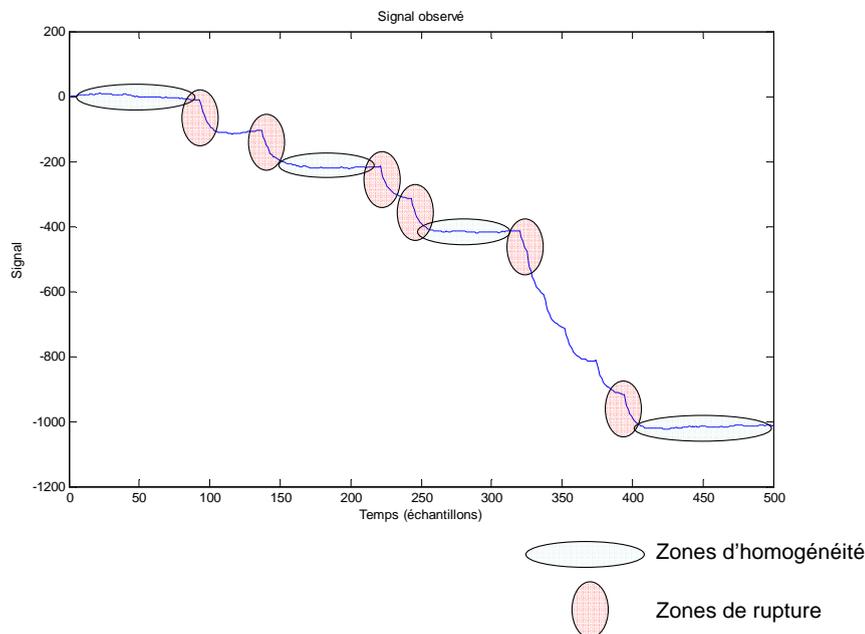


FIGURE 5.2 – Illustration de la démarche « détection de rupture »

il s'agit de formuler une hypothèse et de la tester contre son alternative. L'hypothèse « il n'y a pas de rupture » est appelée hypothèse H_0 (hypothèse nulle), l'hypothèse alternative étant appelée H_1 .

Par exemple [LYFLLC11] propose d'utiliser le test statistique de Mann-Whitney et Wilcoxon. Ce test permet de tester l'égalité des distributions de deux jeux de données *via* un test de rang.

L'idée serait alors d'extraire une fenêtre du signal. Les échantillons de la fenêtre seraient séparés en deux groupes. Nous formulerions l'hypothèse H_0 « tous les points des deux fenêtres sont indépendants et identiquement distribués » contre H_1 « les points des deux fenêtres sont distribués selon deux distributions différentes ». L'utilisation du test permettrait dans notre cas de sélectionner l'hypothèse H_0 ou l'hypothèse H_1 , et donc dans ce dernier cas de détecter qu'au moins une rupture se situe dans la fenêtre. L'article propose également une méthode afin de trouver le point de rupture le plus crédible, par maximisation.

Citons aussi [WWW07] ou [Wan08] qui sont des méthodes de détection de rupture basées sur une dérivation du test de Student ou du test de Fisher.

Notons qu'il existe des versions bayésiennes de détection de rupture [TDL03, DTD07, MDF06].

Algorithme CUSUM L'ouvrage [BN93] traite spécifiquement de la détection de rupture dans un signal et propose notamment l'algorithme CUSUM qui se base sur la comparaison d'une somme de rapports de vraisemblance.

Cet algorithme consiste à balayer le signal selon l'axe des temps. Pour chaque temps u , deux hypothèses H_0 et H_1 sont comparées, hypothèses qui donnent toutes deux lieu à des lois de

vraisemblance, dont les valeurs sont estimées sur la base des échantillons précédents.

L'algorithme CUSUM [BN93, Ch. 2] consiste à comparer la quantité x_u à un seuil γ , l'expression de x_u étant :

$$x_u = S_u - m_u \quad (5.2)$$

avec :

$$S_u = \sum_{j=1}^u \log \left(\frac{p_{H_1}(\mathbf{g}_j)}{p_{H_0}(\mathbf{g}_j)} \right) \quad (5.3)$$

et :

$$m_u = \min_{1 \leq j \leq k} S_j \quad (5.4)$$

Si un nouvel échantillon est testé et donne beaucoup de crédit à l'hypothèse H_1 , alors la quantité x_u va augmenter, ainsi que S_u . Si ce crédit à l'hypothèse H_1 est apporté par d'autres données, par effet cumulatif, cette quantité S_u va dépasser le seuil γ : la rupture sera détectée.

Dans le cas d'un test d'hypothèse entre variables normales de même variance σ^2 et de moyennes μ_1 pour l'hypothèse H_1 et μ_0 pour l'hypothèse H_0 , l'expression de S_u devient :

$$S_u = \frac{\mu_1 - \mu_0}{\sigma^2} \sum_{j=1}^u \left(\mathbf{g}_j - \frac{\mu_1 + \mu_0}{2} \right) \quad (5.5)$$

5.2.2.2 Détection de pics

Une seconde catégorie de méthodes qui peuvent être utilisées pour résoudre le problème de détection consiste à se pencher sur la détection de pics. Dans notre cas, ceci nécessiterait une dérivation du signal de sortie au préalable, ce qui serait cohérent avec la structure du bruit (voir section 3.4.1). La figure 5.3 illustre l'idée de la détection de pics.

La méthode publiée par l'équipe de Caltech et décrite ci-dessus peut être placée dans cette catégorie (dérivation-seuillage).

Un domaine scientifique qui s'intéresse à ce type d'approches est celui des neuro-sciences, dans la mesure où le signal élémentaire transitant sur les neurones est un potentiel d'action (en anglais, *spike*) qui est constitué d'une impulsion brève dans le domaine des potentiels électriques. Dans l'objectif de détecter ces potentiels d'action, un nombre important de méthodes ont été mises en place [GJM12].

On pourra citer [GJM10] qui décrit plusieurs méthodes de détection. Parmi elles, ce papier décrit :

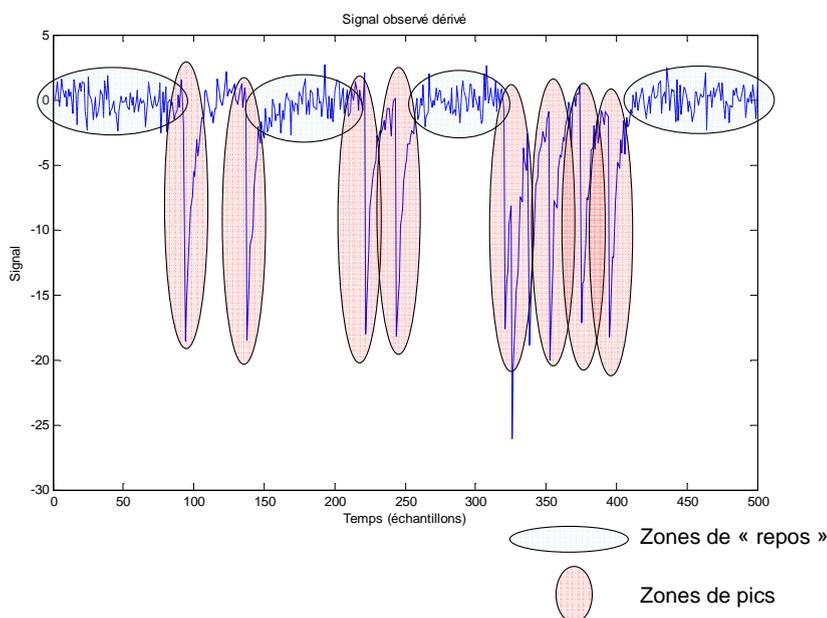


FIGURE 5.3 – Illustration de la démarche « détection de pics »

- des méthodes de seuillage avec seuillage adaptatif, dans la lignée de la méthode proposée par Caltech ;
- des méthodes basées sur un « opérateur non-linéaire de calcul de l'énergie » (NEO pour *Nonlinear Energy Operator*) qui consiste en l'estimation d'une énergie « haute-fréquence » ;
- l'utilisation d'une décomposition en ondelettes [EBG07] (voir plus loin, section 5.2.3.3).

5.2.2.3 Apprentissage

Une tierce catégorie de méthodes de traitement des données serait celles des approches de type apprentissage non basées sur une analyse du signal.

Ce type de méthode ne repose pas sur une modélisation du système de mesure et permet généralement une première approche du sujet pour un coût d'analyse qui peut être réduit. De cette manière, il est possible d'aborder le problème considéré sous l'angle de la reconnaissance de la forme générée par une adsorption à l'aide d'un système de classification.

La démarche consisterait en :

- une extraction de caractéristiques à partir du signal (elles peuvent être de natures très diverses [JDM00]) ;
- un apprentissage de la forme d'une adsorption à partir d'une base de données contenant des morceaux de signal étiquetés : classification binaire (pas d'adsorption / une adsorption) ;
- une application de l'algorithme de classification au signal observé.

Il serait ainsi possible de se ramener à un problème de classification et donc d'utiliser les méthodes de classification de l'état de l'art (plus proches voisins, réseaux de neurones, machines à vecteurs de support, etc. . .).

5.2.2.4 Discussion sur les méthodes directes

La principale critique que nous pouvons faire en ce qui concerne les méthodes directes est que ces dernières n'exploitent pas toute l'information disponible dans notre problème. En particulier, ces méthodes ne prennent pas en compte l'effet de la convolution par la réponse impulsionnelle induite par la boucle de lecture. Notons aussi que les méthodes directes présentent généralement un certain nombre de paramètres à régler (seuils, taille de fenêtres, etc. . .) auxquels il peut être délicat de donner un sens.

Par conséquent, si nous considérons que ces méthodes peuvent être pleines de sens dans des contextes où peu d'informations peuvent être issues d'une analyse *a priori* du problème, nous préférons l'emploi de méthodes issues d'une approche problème inverse dans notre cas.

5.2.3 Approches problème inverse adaptées

Compte-tenu de la rigueur du cadre de travail offert par l'approche problème inverse (voir section 4.1), nous allons nous pencher plus en détails sur des méthodologies qui en découlent.

Rappelons la forme générale du modèle direct :

$$\mathcal{H}(f, g) = 0 \tag{5.6}$$

avec f l'inconnue, g les données observées et \mathcal{H} la fonctionnelle système.

Dans le cas linéaire, en utilisant la matrice système H :

$$g = Hf \tag{5.7}$$

Nous verrons quatre catégories de méthodes permettant d'estimer l'inconnue sachant les données observées, avec une contrainte sur le caractère impulsionnel de la solution recherchée :

1. l'approche détection ;
2. l'utilisation de la pseudo-inverse ;
3. l'utilisation de méthodes de décomposition sur une base ;
4. l'utilisation d'approches pénalisation et d'approches bayésiennes.

5.2.3.1 Approche filtrage et théorie de la détection

Nous proposons de nous pencher en premier lieu sur les approches type « détection ».

Cette approche consiste en une modélisation du signal observé associée à l'expression d'hypothèses qui vont être testées. L'hypothèse H_0 est l'hypothèse d'absence de signal et l'hypothèse H_1 est celle de la présence d'un signal.

Dans notre cas, il s'agirait de détecter un adsorption, c'est-à-dire de reconnaître une forme de signal connue \mathbf{h} mais d'amplitude a inconnue dans un signal bruité \mathbf{g} , ϵ étant le bruit.

Ainsi, le problème se formule :

$$\begin{aligned} H_0 : \mathbf{g} &= \epsilon \\ H_1 : \mathbf{g} &= a\mathbf{h} + \epsilon \end{aligned} \quad (5.8)$$

Une manière de résoudre ce problème serait de mettre en place une stratégie de Neyman-Pearson afin de garantir un taux de fausse détection constant. En supposant le bruit ϵ normal centré de matrice de covariance identité, on se ramène à un problème classique de détection d'un signal connu à amplitude inconnue dans un bruit blanc. Le test statistique devient [Kay98, Ch.7] le test du GLRT (pour *Generalized Likelihood Ratio Test*) qui s'écrit ici :

$$|\mathbf{g}^T \mathbf{h}|^2 \underset{<}{\underset{>}{\geq}} \gamma \quad (5.9)$$

Ici, γ est une constante dépendant du niveau de bruit, de la réponse \mathbf{h} et du taux de fausse-détection visé.

Généralement, lorsque l'on cherche à effectuer la détection d'un signal impulsionnel \mathbf{f} filtré par une réponse impulsionnelle \mathbf{h} dans un bruit, la méthode consiste à effectuer une suite de corrélations comme indiqué sur l'équation (5.9) avec des réponses impulsionnelles \mathbf{h} se décalant dans le temps. Cette opération de corrélation peut se ré-écrire sous la forme d'un filtrage, avec un filtre dit « filtre adapté ». Le filtre adapté $\bar{\mathbf{h}}$ correspond à la réponse impulsionnelle \mathbf{h} retournée temporellement.

Ainsi, la méthode consiste généralement à calculer la sortie du filtre adapté $\hat{\mathbf{f}}$:

$$\hat{\mathbf{f}} = \mathbf{g} \star \bar{\mathbf{h}} \quad (5.10)$$

Le filtre adapté est originellement conçu pour détecter une seule impulsion. Pour effectuer une déconvolution impulsionnelle, il est nécessaire d'adjoindre un organe de décision. Ainsi, la sortie de ce filtre adapté peut être ensuite lissée, seuillée ou voir ses *maxima* extraits. L'estimation d'amplitude se fait soit à partir de la valeur de sortie du filtre, soit par une procédure externe.

Le filtre adapté peut également être utilisé au travers de l'algorithme *Matching-Pursuit*, que vous verrez dans la section 5.2.3.3.

5.2.3.2 Pseudo-inverse et filtre de Haar

Une approche algébrique simple serait celle basée sur la matrice pseudo-inverse (que nous avons décrite dans la section 4.1.4.1). En effet, la matrice pseudo-inverse est une manière simple d'inverser une opération de convolution.

Rappelons l'expression de la matrice système \mathbf{H} dans le cas d'une convolution :

$$\mathbf{H} = \begin{pmatrix} h_1 & 0 & \cdots & \cdots & 0 \\ h_2 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{T-1} & & \ddots & \ddots & 0 \\ h_T & h_{T-1} & \cdots & h_2 & h_1 \end{pmatrix} \quad (5.11)$$

Dans notre cas, la réponse impulsionnelle ne s'annule pas, donc la matrice \mathbf{H} , quoique mal conditionnée, est de rang plein. Par conséquent la matrice pseudo-inverse vaut : $\mathbf{H}^\dagger = \mathbf{H}^{-1}$.

La méthode consiste donc à appliquer :

$$\hat{\mathbf{f}}_{\text{pseudo-inverse}} = \mathbf{H}^\dagger \mathbf{g} = \mathbf{H}^{-1} \mathbf{g} \quad (5.12)$$

Le reproche principal que l'on peut formuler à l'encontre de cette solution est qu'elle ne prend pas en compte le caractère impulsionnel de la solution. Pour pallier à cela, une méthode très simple consiste à seuilier $\hat{\mathbf{f}}_{\text{pseudo-inverse}}$.

Un second reproche vient du fait que la matrice \mathbf{H} peut être mal conditionnée. L'article [MBC04] propose une alternative à la pseudo-inverse, consistant à utiliser une pénalisation L_2 (appelée parfois, comme le fait l'auteur, filtre de Haar). Ceci permet de régulariser l'inversion de \mathbf{H} en ajoutant des éléments sur la diagonale (voir la section 4.1.4.2). Le signal ainsi restauré est ensuite seuillé afin d'en retrouver le caractère impulsionnel.

5.2.3.3 Décomposition du signal sur une base

Les premières méthodes consistent à décomposer le signal sur une base connue (voir section 4.1.3.2). D'une manière plus générale, il s'agit de régulariser l'inversion en se projetant sur des sous-espaces de dimension plus petite. La décomposition sur une base peut avoir trois effets :

- réduire la dimension du signal observé, et donc limiter les problèmes d'existence de la solution ;
- réduire la dimension de l'espace de départ (variables recherchées) et donc limiter les problèmes d'unicité de la solution ;
- réduire la dimension des espaces de départ et d'arrivée, donc rendre le problème plus robuste.

Les méthodes de décomposition sont encore à l'étude au travers du stage de M. Mohammed Lamine Laoufi, étudiant en Master 2 à l'INP Grenoble. Nous en donnerons toutefois les grandes lignes.

Ondelettes Un ensemble de bases de décomposition particulièrement utilisé en traitement de signal est l'ensemble des bases d'ondelettes. Les ondelettes sont des fonctions de carré sommable qui permettent de représenter un signal dans un espace temps-fréquence. En effet, la base d'ondelettes est obtenue à partir d'une ondelette dite « ondelette-mère » qui se voit translatée

(« temps ») et changée d'échelle (« fréquence »). La base ainsi constituée permet une analyse en motifs élémentaires. Cette base est plus adaptée à l'analyse de phénomènes impulsionnels [Mal00, Ch. 1] que la base de Fourier, par exemple, qui est quant à elle adaptée à l'analyse de phénomènes périodiques.

Un exemple d'utilisation d'ondelettes (dans cet exemple, ondelettes de Haar) est donné dans [Cap06] pour un contexte applicatif très différent du notre, en l'occurrence la détection de micro-séismes.

Matching-Pursuit et Orthogonal Matching Pursuit L'algorithme *Matching-Pursuit* [PRK93, Sec. 1] est un algorithme de décomposition sur une base de type « glouton », c'est-à-dire ne remettant pas en cause les choix passés. Celui-ci fournit une représentation parcimonieuse dans la base de décomposition choisie ce qui permet soit de prendre une décision si les coefficients peuvent être reliés aux grandeurs recherchées, soit de débruiter le signal pour lui appliquer un algorithme de détection basique. *Orthogonal Matching Pursuit* [PRK93, Sec. 2] est l'extension de *Matching-Pursuit* pour une base de décomposition orthogonale (assez souvent, une base en ondelettes).

L'algorithme *Matching-Pursuit* fonctionne itérativement, une itération étant structurée comme suit :

- faire la corrélation du signal avec chacun des éléments dans la base de décomposition ;
- trouver la corrélation maximale (en valeur absolue) ;
- déduire du signal observé le produit de la corrélation maximale avec le vecteur de la base de décomposition correspondant.

Prenons le cas de l'utilisation du dictionnaire formé par les réponses impulsionnelles décalées du pas d'échantillonnage. Notons ces réponses $\mathbf{h}(u)$ pour un décalage de u . Dans ce cas, l'algorithme *Matching-Pursuit* consiste à estimer l'impulsion la plus probable au sens du filtre adapté :

$$u_{\text{OPT}} = \operatorname{argmax}_u (|\mathbf{g}^T \mathbf{h}(u)|) \quad (5.13)$$

Ensuite, l'impulsion estimée serait supprimée du signal :

$$\mathbf{g} = \mathbf{g} - \max_u (|\mathbf{g}^T \mathbf{h}(u_{\text{OPT}})|) \frac{\mathbf{h}(u_{\text{OPT}})}{|\mathbf{h}(u_{\text{OPT}})^T \mathbf{h}(u_{\text{OPT}})|} \quad (5.14)$$

Puis l'algorithme recommencerait à l'étape précédente. D'une certaine manière, l'algorithme *Matching-Pursuit* permet d'adapter le filtre adapté à la déconvolution impulsionnelle. Il est possible de trouver cet algorithme utilisant la réponse impulsionnelle sous la dénomination « CLEAN » [Hö74].

5.2.4 Approche pénalisation et approche bayésienne

Les méthodes que nous allons développer plus en détail sont les méthodes reposant sur le principe de pénalisation, que nous avons décrites précédemment à la section 4.1.4.2. Nous allons notamment exprimer les méthodes en terme d'inversion bayésienne.

Compte-tenu du lien très fort qui unit pénalisation et approche bayésienne (voir section 4.2.5), nous les avons regroupées dans une même section. Notons que là où une démarche de type pénalisation fait intervenir des paramètres auxquels il est difficile de donner un sens, la démarche bayésienne fait intervenir la notion de loi *a priori* et permet de donner plus de sens aux choix à effectuer.

Rappelons le lien entre ces deux approches dans le cas d'une vraisemblance normale centrée de variance σ^2 :

$$\begin{aligned} \text{Critère pénalisé : } & \|g - Hf\|_2^2 + \lambda \Omega(f) \\ \text{Critère bayésien : } & \|g - Hf\|_2^2 - \sigma^2 \log(p(f)) \end{aligned} \quad (5.15)$$

Il est possible de trouver des pénalisations et des lois *a priori* qui vont renforcer le caractère impulsif de la solution, et seront donc adaptées pour la déconvolution impulsionnelle [MD12a]. Nous nous plaçons ici dans le cas classique des processus à temps discret (voir section 3.3.1.2).

Nous pouvons extraire trois catégories de pénalisations / lois *a priori* :

1. les pénalisations associées à la « parcimonie », qui est une approche de type « pénalisation de critère » ;
2. les lois à « queue lourde », qui est une approche de type « critère probabiliste » ;
3. les modèles de mélange, qui sont des approches de type « critère probabiliste » également.

5.2.4.1 Pénalisations associées à la « parcimonie »

Dans le cadre de la parcimonie, le but est de trouver un signal présentant de nombreuses composantes à 0 dans un dictionnaire donné, ici une base temporelle. Le terme de pénalisation le plus naturel est donc la norme l_0 , qui est égale au nombre de composantes non nulles :

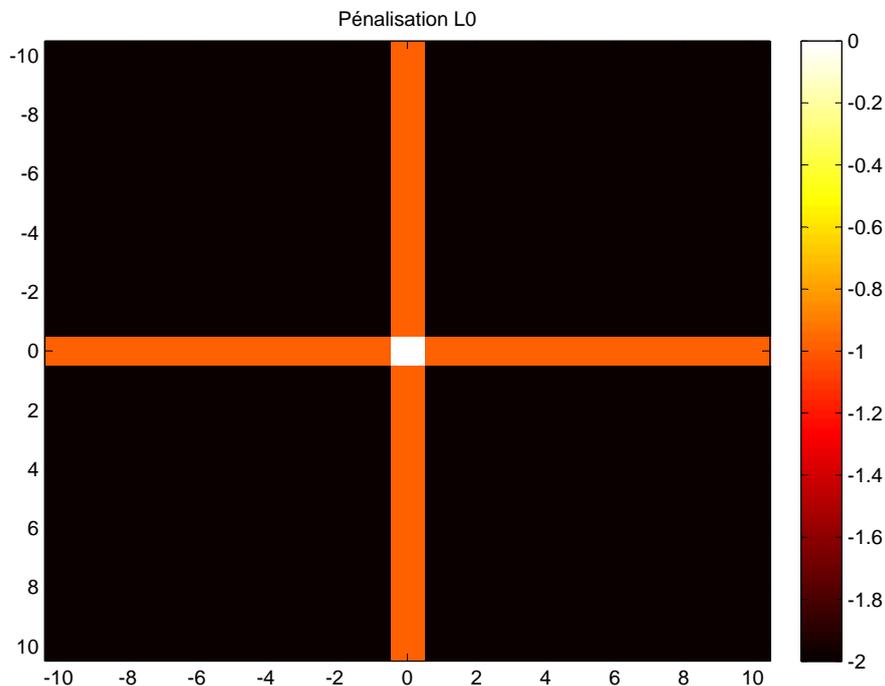
$$\Omega_{l_0}(f) = \text{card}(f \neq 0) \quad (5.16)$$

La figure 5.4 illustre cette pénalisation. Les valeurs hors des axes sont pénalisées.

Malheureusement, optimiser un critère construit avec cette pénalisation s'avère une opération délicate car combinatoire. L'idée de Tibhsirani [Tib11] a été de relaxer ce problème en un problème continu et différentiable presque partout, au travers d'une pénalisation convexe, la pénalisation L_1 .

$$\Omega_{l_1}(f) = \|f\|_1 \quad (5.17)$$

Cette pénalisation est nommée LASSO. Le pendant bayésien du LASSO est la loi double exponentielle illustrée par la figure 5.5. Nous proposons de tracer à la fois la loi de probabilité (en


 FIGURE 5.4 – Illustration de la pénalisation L_0

une dimension) et le logarithme de cette loi (c'est-à-dire la pénalisation associée, voir 4.2.5), en 2D, en supposant les composantes indépendantes et identiquement distribuées.

Cette loi prend un paramètre d'échelle a et a pour expression :

$$p(f|a) \propto e^{-\frac{|f|}{a}} \quad (5.18)$$

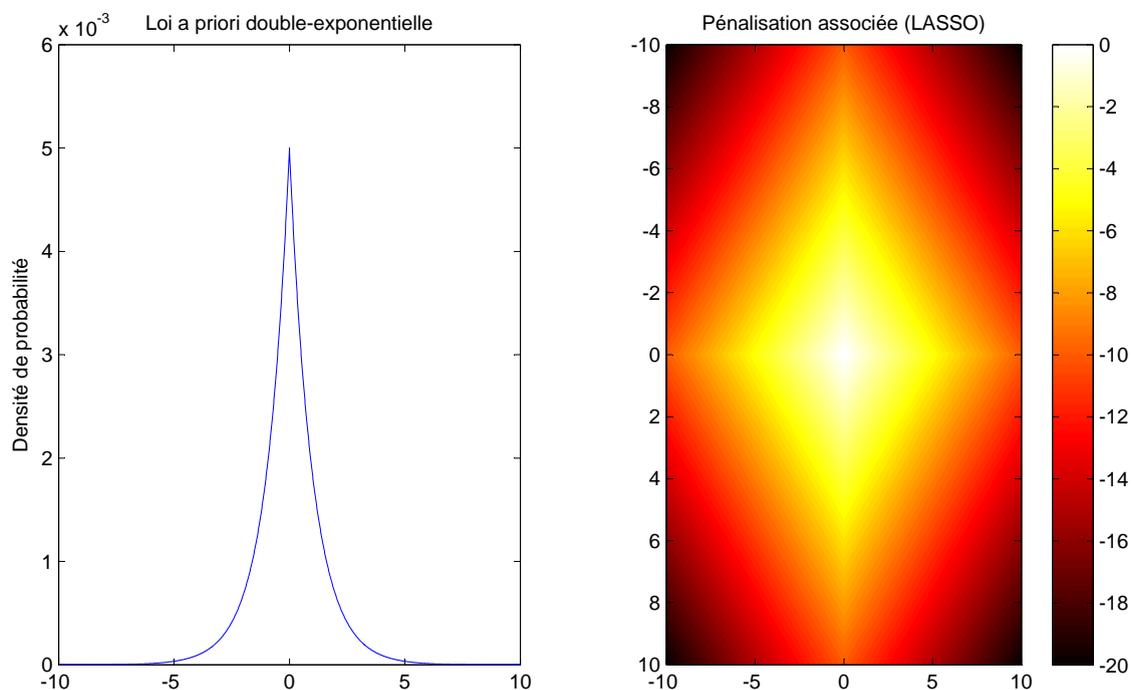
Dans ce cas, le critère devient :

$$\hat{\mathbf{f}}_{\text{LASSO}} = \arg \min_{\mathbf{f}} (\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{f}\|_1) \quad (5.19)$$

On remarquera que la vraisemblance de cette loi ne ressemble pas à la pénalisation L_0 . Qui plus est, des échantillons suivant cette loi ne seraient pas parcimonieux. En réalité ici, l'élément qui va renforcer la parcimonie ici est la singularité en 0 qui va « attraper » le critère sur les axes.

Il a été montré que sous certaines conditions, optimiser un critère pénalisé avec Ω_{l_1} donnait le même résultat que l'optimisation d'un critère avec Ω_{l_0} [Sta10].

Un algorithme permettant de calculer simplement le critère suivant pour différentes valeurs du paramètre de régularisation λ est nommé LAR (Least-Angle Regression) [EHJ⁺04].


 FIGURE 5.5 – Illustration de la loi double-exponentielle avec $a = 1$

Dans le même état d'esprit, on retrouve des pénalisations continues, différentiables et convexes : les pénalisations L_p et hyperboliques [Idi01, Ch. 5] (on parle de déconvolution « L2Lp » et « L2Hy »).

L'expression de la pénalisation dite L_p avec p variant entre 1 et 2 est :

$$\begin{aligned} \Omega_{l_p}(\mathbf{f}) &= \|\mathbf{f}\|_p \\ &= \left(\sum_i |\mathbf{f}_i|^p \right)^{\frac{1}{p}} \end{aligned} \quad (5.20)$$

L'expression de la pénalisation hyperbolique, quant à elle, est :

$$\Omega_{l_{Hy}}(\mathbf{f}) = \sum_i \sqrt{\mathbf{f}^2 + a^2} \quad (5.21)$$

5.2.4.2 Lois à queue lourde

Une deuxième solution pour renforcer la parcimonie d'une solution est l'utilisation de lois permettant des valeurs très faibles avec une probabilité assez forte et des valeurs importantes avec une probabilité ne tendant pas « rapidement » vers 0. Utilisant des lois *a priori*, nous nous

plaçons ici dans un cadre statistique bayésien, toujours avec une équivalence avec la pénalisation de critère si un critère de type MAP est utilisé.

Ces lois sont généralement qualifiées de « lois à queue lourde », « lois à longue queue », ou encore « lois leptokurtiques », dans la mesure où leur *kurtosis* est positif et important.

Ces lois ne vont pas conduire à une solution parcimonieuse *stricto sensu* mais vont renforcer le caractère piqué du signal d'entrée. Il sera donc possible de faire suivre l'inversion d'une étape de seuillage. Ces lois *a priori* vont permettre de régulariser le problème mais en préservant les « contrastes » dans le signal.

Parmi les lois à queue lourde, il est possible de citer (liste non exhaustive, voir [MD12a]) :

- la loi de Rayleigh symétrique ;
- la loi de Weibull symétrique ;
- la loi de Cauchy ;
- la loi de Student-t.

Intéressons-nous à la loi de Student-t. Cette dernière est illustrée par la figure 5.6. Cette loi prend pour paramètre $s \in \mathbb{N}^*$, dit « degrés de liberté ». Son expression est :

$$p(f|s) \propto \left(1 + \frac{f^2}{s}\right)^{-\frac{s+1}{2}} \quad (5.22)$$

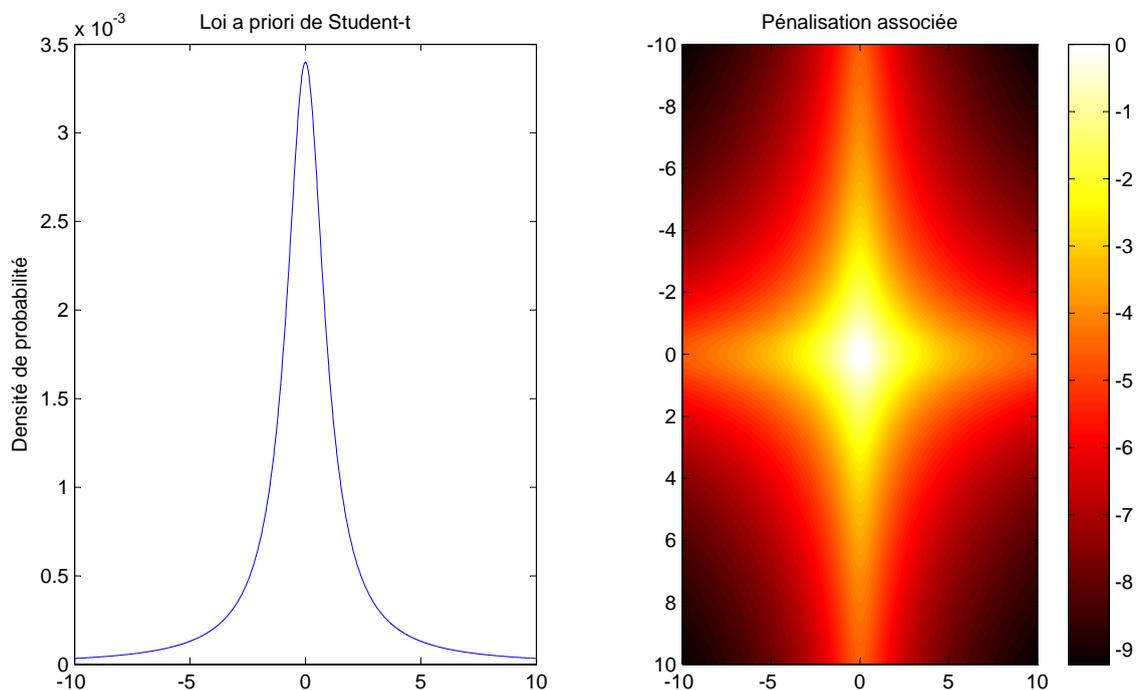


FIGURE 5.6 – Illustration de la loi de Student-t avec $s = 1$

Par rapport à la pénalisation L_1 , cette loi s'approche visuellement plus de la pénalisation L_0 .

Cette loi va renforcer le fait que la solution présentera à la fois un grand nombre de valeurs autour de 0 et quelques « grandes » valeurs (elle est « piquée »).

La loi de Student-t peut être vue comme un mélange infini de lois normales. En effet, il est possible de ré-écrire la loi de Student-t en faisant intervenir un paramètre τ :

$$\begin{aligned} p(f|s) &\propto \left(1 + \frac{f^2}{s}\right)^{-\frac{s+1}{2}} \\ &\propto \int_0^{+\infty} \mathcal{N}\left(f|0, \frac{1}{\tau}\right) \mathcal{Gam}\left(\tau|\frac{s}{2}, \frac{2}{s}\right) d\tau \end{aligned} \quad (5.23)$$

Le détail du calcul est donné en annexe E.3. On voit ici que τ s'identifie à une inverse-variance. Ainsi, la loi de Student-t est en quelque sorte la somme infinie de lois normales de variances différentes.

L'intérêt de cette expression est que l'inversion pourra se faire par un schéma itératif dans lequel f mais aussi τ seront estimés chacun à leur tour, la loi $p(f|\tau, s)$ étant une loi normale. A ce titre [MD12b, CMDPG13] proposent un algorithme de déconvolution impulsif reposant sur l'approximation bayésienne variationnelle qui exploite l'utilisation de la variable cachée τ .

5.2.4.3 Modèles de mélange

La dernière catégorie que nous avons identifiée concerne les modèles de mélange, c'est-à-dire la sommation de plusieurs lois de probabilité élémentaires. Dans ce cas également, nous nous plaçons dans un contexte statistique bayésien.

L'idée est de définir une loi *a priori* pour le bruit et une voire plusieurs lois *a priori* pour l'ensemble {signal + bruit}. La loi *a priori* pour f est la sommation de ces deux lois.

Un exemple très courant est celui du modèle Bernoulli-Gaussien, [Idi01, Ch.5], qui consiste à prendre en compte explicitement la parcimonie du signal à retrouver au cours de l'inversion.

Pour ce faire, le signal est discrétisé, et une variable binaire (appelée généralement q) valant 1 aux endroits où l'entrée est non nulle et 0 ailleurs est introduite. Cette variable q suit une loi de Bernoulli de paramètre π . Aux endroits où cette variable q vaut 1, la loi *a priori* est une loi normale centrée de variance r . Aux endroits où q vaut 0, la loi *a priori* est une loi de Dirac centrée en 0.

$$\begin{aligned} p(\mathbf{f}_i|\mathbf{q}_i = 0) &= \delta(\mathbf{f}_i) \\ p(\mathbf{f}_i|\mathbf{q}_i = 1) &= \mathcal{N}(\mathbf{f}_i|0, r) \end{aligned} \quad (5.24)$$

Ce modèle peut se mettre sous forme hiérarchique, comme illustré par la figure 5.7.

La figure 5.8 illustre cette loi. Sur cette figure, les proportions entre la loi « présence de signal » et la loi « absence de signal » n'ont pas été respectées pour faciliter la lecture.

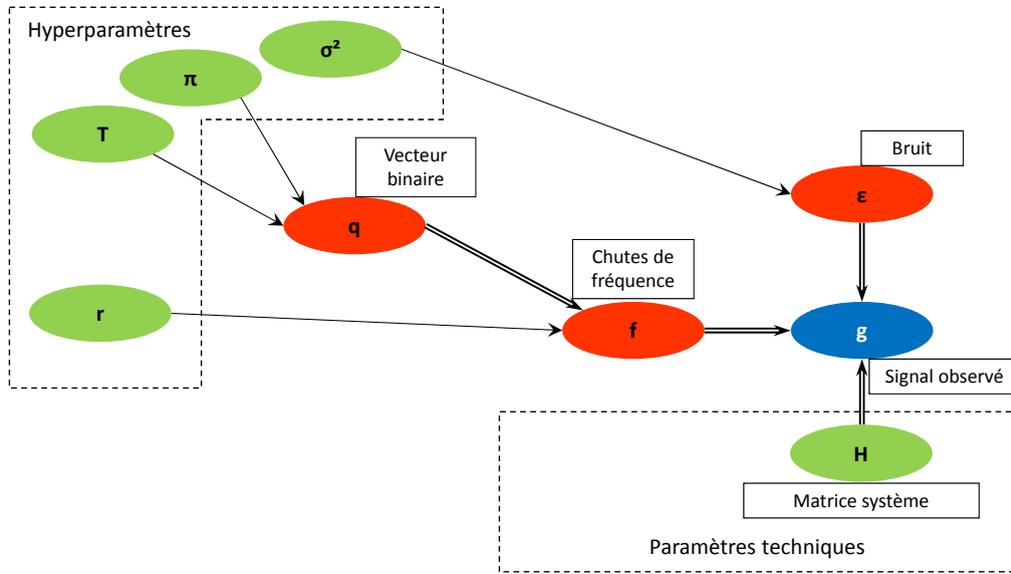


FIGURE 5.7 – Vision hiérarchique du modèle Bernoulli-Gaussien

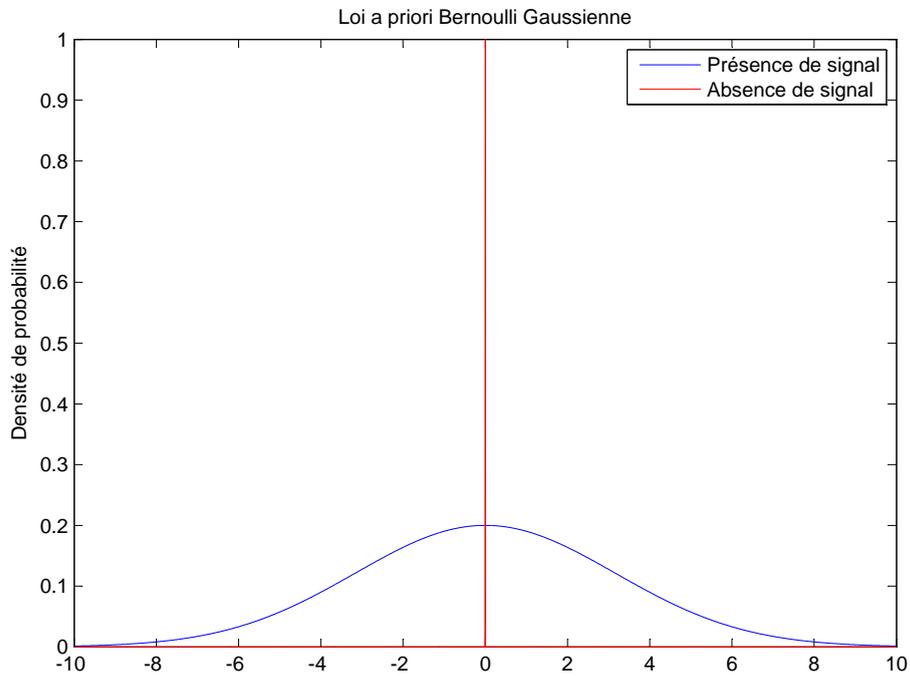


FIGURE 5.8 – Illustration du modèle Bernoulli-Gaussien avec $r = 10$

Cette loi a le mérite d'être explicitement impulsionnelle, dans la mesure où le modèle favorise les éléments nuls, alors que les solutions trouvées avec les lois *a priori* précédentes doivent être traitées (généralement par seuillage) pour retrouver un caractère réellement impulsionnel. L'autre intérêt est la marginalisation des amplitudes si la vraisemblance est modélisée avec une loi normale, ce qui permet d'écrire :

$$p(\mathbf{q}|\mathbf{g}) = p(\mathbf{q}) \mathcal{N}(\mathbf{g}|\mathbf{0}, \mathbf{R}) \quad (5.25)$$

avec :

$$\mathbf{R} = r\mathbf{H}\mathbf{\Pi}\mathbf{H}^T + \sigma^2\mathbf{I} \quad (5.26)$$

avec $\mathbf{\Pi}$ une matrice diagonale dont la diagonale est le vecteur \mathbf{q} : $\mathbf{\Pi} = \text{diag}(\mathbf{q})$.

Un reproche important que l'on peut faire au modèle Bernoulli-Gaussien est l'introduction d'une variable binaire, qui rend le calcul analytique impossible.

Pour estimer un processus Bernoulli-Gaussien, il est possible d'employer un algorithme de recherche de proche en proche, ce qui est notamment permis par le fait que seule la variable binaire \mathbf{q} est inconnue. A ce titre, l'article [KM82] décrit un algorithme d'optimisation pour les modèles Bernoulli-Gaussien appelé SMLR pour *Single Most Likely Replacement*. Cet algorithme consiste à chaque itération à tester tous les vecteurs \mathbf{q} voisins du vecteur \mathbf{q} courant, le voisinage étant défini par la modification d'une composante du vecteur. Le vecteur donnant lieu à la plus grande valeur de loi *a posteriori* est ensuite choisi.

Il est également possible de se tourner vers un modèle plus « doux », comme un mélange de deux lois normales. L'idée ici est d'avoir un mélange de deux normales, centrées, l'une ayant une très faible variance et symbolisant l'absence de signal, et l'autre ayant une variance plus importante, symbolisant le signal.

5.3 Développement d'une méthode d'inversion probabiliste reposant sur un choix de modèle

5.3.1 Justification de l'approche employée

Voyons ici les éléments qui nous ont conduit au développement de notre méthode :

1. le problème est en l'état mal-posé ;
2. le modèle Bernoulli-Gaussien nous semble explicite et facile à dériver.

5.3.1.1 Mise en évidence du caractère mal-posé du problème

Nous avons vu précédemment à la fois une modélisation du système de mesure ainsi que des méthodes permettant l'estimation d'un signal inconnu en entrée du système.

La modélisation du système de mesure aboutissait sur une équation permettant de donner une forme au signal pour un ensemble de molécules incidentes données. Traçons ce signal pour deux ensembles de molécules incidentes différentes dans la figure 5.9.

On voit ici que deux ensembles de molécules incidentes donnent quasiment le même signal observé. Ainsi, si on ajoute à ce signal une légère perturbation, il n'est plus possible de choisir l'une ou l'autre des listes de molécules données ci-dessus. Par conséquent, le caractère stable du problème n'est plus, le problème est donc en l'état mal-posé.

C'est pourquoi nous proposons de régulariser le problème au travers de l'introduction d'informations *a priori* sur la solution recherchée, et le cadre bayésien est particulièrement adapté pour cela. De plus, nous avons montré dans la section 3.3.2.2 que les différentes inconnues du système influencent les unes sur les autres sous la forme d'un ensemble de relations hiérarchiques. Ici encore, une approche bayésienne semble tout à fait adaptée pour prendre en compte ces dépendances.

Nous considérons également pouvoir aller plus loin que les méthodes de déconvolution classiques identifiées à la section 5.2.3, y compris les méthodes régularisées, pour des raisons que nous allons voir dans la section qui suit.

5.3.1.2 Dérivation du modèle Bernoulli-Gaussien

Afin de mettre en place notre méthode d'estimation, nous choisissons de nous inspirer du modèle Bernoulli-Gaussien. En effet, ce modèle nous semble adapté car :

- il s'agit d'un modèle explicitement parcimonieux ;
- les paramètres à régler (variance des impulsions, densité d'impulsions) nous paraissent explicites, et nous pourrions estimer certains de ces paramètres ;
- nous allons pouvoir facilement dériver ce modèle pour l'adapter à notre problème.

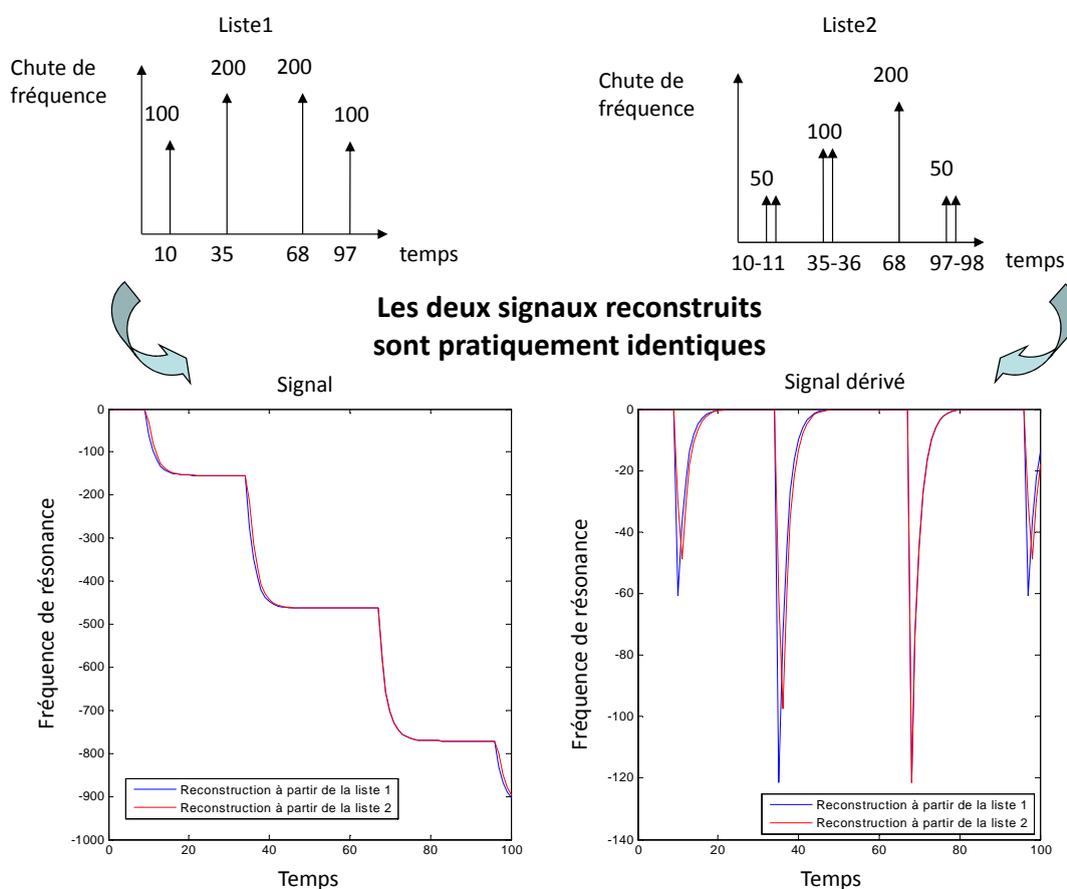


FIGURE 5.9 – Illustration du caractère mal-posé du problème

Comme précisé dans le dernier point précédent, nous n’allons pas appliquer le modèle Bernoulli-Gaussien directement. Nous allons exprimer le problème comme un choix de modèle, le modèle étant le nombre de molécules adsorbées N . Ce choix nous permettra une plus grande souplesse de modélisation, pour pouvoir adresser les spécificités du problème suivantes :

1. les chutes de fréquences sont toutes négatives (d’où $f_i \geq 0$) là où une loi *a priori* normale prend des valeurs non nulles sur tout \mathbb{R} . D’une manière plus générale, nous ne voulons pas nous restreindre à utiliser un modèle *a priori* normal sur les amplitudes ;
2. les instants d’adsorptions peuvent être non entiers (voir section 3.2.2.3) ;
3. nous souhaitons estimer les hyperparamètres de certaines lois.

État de l’art sur le modèle Bernoulli-Gaussien Notons que des variations autour du modèle Bernoulli-Gaussien existent dans l’état de l’art.

En effet, Soussen *et al.* [SIBD11] dérivent l’algorithme du SMLR en SBR (pour *Single-Best Replacement*) en faisant tendre la variance de la loi *a priori* normale du modèle Bernoulli-Gaussien vers l’infini.

Il est également possible de citer la thèse de doctorat de Sébastien Bourguignon [Bou08], qui uti-

lise l'algorithme de Gibbs pour estimer un modèle Bernoulli-Gaussien dont les hyperparamètres sont estimés conjointement.

Aussi, la thèse de Vincent Mazet [Maz05] enrichit un modèle Bernoulli-Gaussien avec contrainte de positivité et extension du modèle à des événements en dehors de la grille d'échantillonnage.

Enfin, dans un cadre légèrement différent, on peut citer la thèse de doctorat de Di Ge [Ge09], dans laquelle un modèle Bernoulli-Gaussien est utilisé dans le cadre de la déconvolution myope (réponse impulsionnelle inconnue).

Nous retirons de cette analyse de l'état de l'art que le modèle Bernoulli-Gaussien est un modèle puissant qui peut servir de base pour des développements plus complexes en fonction du problème étudié, et c'est exactement dans ce cadre que nous allons travailler.

Loi *a priori* permettant d'assurer la positivité des amplitudes de chute Le premier inconvénient que nous avons énoncé en ce qui concerne la modélisation Bernoulli-Gaussienne est celle de son support \mathbb{R} , là où nous voudrions assurer la positivité des amplitudes de $f(t)$. Pour cela, plusieurs solutions ont été proposées. On pourra notamment citer Vincent Mazet [Maz05] qui utilise une loi normale tronquée. Citons également Razul *et al.* [RFA03], article traitant des spectres d'émission nucléaire et utilisant une loi *a priori* Gamma.

Permettre des instants d'adsorption non-entiers Un point important que nous avons identifié est celui consistant à permettre des instants d'adsorption non-entier. Nous entendons par cela que la molécule s'adsorbe sur le capteur en un temps qui n'est pas un temps d'échantillonnage. Cela se traduit par l'impossibilité d'utiliser le contexte des processus à temps discret (voir section 3.3.1.2), et en particulier l'impossibilité d'écrire le problème à l'aide d'un vecteur binaire \mathbf{q} . Par conséquent, nous allons choisir d'exprimer le problème comme un processus ponctuel marqué, sans contrainte de domaine sur les paramètres recherchés, et d'estimer l'ordre du modèle.

Estimer les hyperparamètres Afin de pouvoir estimer les hyperparamètres, nous allons exploiter la structure hiérarchique du problème, comme évoqué dans la section 3.3.2.2. Notons que Vincent Mazet [Maz05] explicite également cette structure. Dans notre cas, il s'agira de retrouver le paramètre d'intensité de comptage π .

5.3.2 Modélisation statistique des grandeurs

Voyons maintenant comment nous allons modéliser les différentes grandeurs à l'aide de lois de probabilité. Les grandeurs inconnues dans notre cas sont l'intensité de comptage π , et le signal $f(t)$ qui est paramétré par le nombre d'événements N , les temps d'adsorption t_i et les amplitudes de chutes de fréquence f_i :

$$\begin{aligned}
 N &\in \mathbb{N} \\
 t_i &\in [0, T]^N \\
 f_i &\in \mathbb{R}^{+N} \\
 \pi &\in [0, 1]
 \end{aligned}
 \tag{5.27}$$

5.3.2.1 Expression de la fonction de vraisemblance

Comme nous l'avons exprimé dans la section 3.4.1, nous modélisons nos connaissances sur les incertitudes de mesure au travers d'un bruit additif, stationnaire et ergodique qui a pour loi une loi normale centrée de covariance $\mathbf{\Gamma}$.

Bien que notre signal d'entrée $f(t)$ soit non-paramétrique et ne soit pas un signal échantillonné, nous n'avons à disposition pour le calcul de la vraisemblance qu'un signal échantillonné g . Par conséquent, nous allons, pour un signal $f(t)$ donné, calculer la sortie $f(t) \star u_{\text{PLL}}(t)$ puis numériser ce signal de sortie, que nous noterons $\underline{f(t) \star u_{\text{PLL}}(t)}$.

Par conséquent :

$$p(g|\underline{f(t) \star u_{\text{PLL}}(t)}) = \mathcal{N}\left(\underline{g|\underline{f(t) \star u_{\text{PLL}}(t)}}, \mathbf{\Gamma}\right)
 \tag{5.28}$$

Ici, $\underline{\quad}$ symbolise la discrétisation du signal, voir section 3.4.2.2 pour plus de détails.

5.3.2.2 Définition des lois *a priori*

Loi de probabilité sur l'ordre du modèle Comme nous nous inspirons du modèle Bernoulli-Gaussien, nous considérons que l'ordre du modèle N suit la même loi que la somme de variables de Bernoulli indépendantes, c'est-à-dire une loi Binomiale de paramètres π et T :

$$p(N|\pi) = \mathcal{B}(N|\pi, T)
 \tag{5.29}$$

Loi de probabilité sur les temps d'adsorptions Nous modéliserons les temps d'adsorption comme uniformément répartis sur $\{1, T\}$ ou sur $[0, T]$ selon si les temps d'adsorptions sont entiers (processus à temps discret) ou libres (processus ponctuel marqué), voir sections 3.3.1.1 et 3.3.1.2 :

$$p(t_i|\pi) = \mathcal{U}_{\{1, T\}}(t_i)
 \tag{5.30}$$

$$\text{ou } p(t_i|\pi) = \mathcal{U}_{[0, T]}(t_i)
 \tag{5.31}$$

Loi de probabilité sur les amplitudes de chute Les amplitudes de chute de fréquence f_i seront distribuées sous une loi de support positif. Nous proposons de prendre la loi Gamma :

$$p(f_i) = \text{Gam}(f_i | k_m, \theta_m) \quad (5.32)$$

où k_m , et θ_m sont respectivement les paramètres de forme et d'échelle de la loi Gamma.

Loi de probabilité sur l'intensité de comptage Le paramètre d'intensité de comptage, π , sera distribué selon une loi Bêta. Cette loi, à support $[0, 1]$, est la loi *a priori* conjuguée pour le paramètre d'une loi Binomiale (voir tableau 4.1, page 89) :

$$p(\pi) = \text{Bet}(\pi | a_\pi, b_\pi) \quad (5.33)$$

où a_π et b_π sont les paramètres de la loi Bêta.

5.3.2.3 Écriture des lois *a posteriori* des paramètres de la liste pour un ordre de modèle donné

Compte-tenu de la modélisation formulée ci-dessus, nous pouvons écrire la probabilité *a posteriori* jointe de tous les paramètres de la liste sachant un modèle N donné (ces paramètres sont indépendants de π conditionnellement à N), dans le cas où les temps d'adsorption sont entiers :

$$p(t_1, \dots, t_N, f_1, \dots, f_N | N, \mathbf{g}) \propto \mathcal{N}(\mathbf{g} | \underline{f(t) \star u_{\text{PLL}}(t)}, \mathbf{\Gamma}) \prod_{i=1}^N \mathcal{U}_{\{1, T\}}(t_i) \text{Gam}(f_i | k_m, \theta_m) \quad (5.34)$$

et dans le cas où les temps d'adsorption ne sont pas entiers :

$$p(t_1, \dots, t_N, f_1, \dots, f_N | N, \mathbf{g}) \propto \mathcal{N}(\mathbf{g} | \underline{f(t) \star u_{\text{PLL}}(t)}, \mathbf{\Gamma}) \prod_{i=1}^N \mathcal{U}_{[0, T]}(t_i) \text{Gam}(f_i | k_m, \theta_m) \quad (5.35)$$

5.3.2.4 Écriture des autres lois *a posteriori* conditionnelles

Aussi, il est possible d'écrire la loi *a posteriori* conditionnelle à π du modèle N :

$$p(N | \pi, \mathbf{g}) \propto \mathcal{B}(N | \pi) \int p(t_1, \dots, t_N, f_1, \dots, f_N | N, \mathbf{g}) dt_i df_i \quad (5.36)$$

Et enfin, nous écrivons la loi *a posteriori* conditionnelle à N de l'intensité de comptage π :

$$p(\pi|N, T) = \mathcal{B}et(\pi|a_\pi + N, b_\pi + T - N) \quad (5.37)$$

5.3.3 Estimation

5.3.3.1 Utilisation d'un algorithme hybride RJMCMC MCMC

Afin d'explorer convenablement l'espace des modèles et l'espace des paramètres au sein de chaque modèle, nous proposons d'utiliser un algorithme hybride RJMCMC-MCMC (voir section 4.3).

Nous proposons ainsi d'utiliser deux structures différentes. La première sera utilisée pendant le temps de chauffe et sera une stratégie d'exploration de modèle. La seconde visera plus spécifiquement à effectuer l'étape de déconvolution impulsionnelle. Dans les deux cas, il s'agit d'une étape RJMCMC suivi d'une ou plusieurs étapes MCMC.

La structure sera la suivante :

1. à l'itération j , une étape RJMCMC est effectuée. Celle-ci permet de changer l'ordre du modèle, qui sera noté $N^{(j)}$;
2. si l'ordre du modèle a été changé par ajout d'un événement dans la liste, et que nous nous situons au-delà du temps de chauffe, nous échantillons les paramètres f_{i_0} et t_{i_0} de l'événement concerné i_0 par le changement avec B étapes MCMC (nous ne proposons pas de contrôle automatique de convergence mais nous prendrons B suffisamment grand) nous remplaçons ensuite f_{i_0} et t_{i_0} par leur valeur moyenne estimée par la chaîne ;
3. pour finir, chacun des paramètres du modèle d'ordre $N^{(j)}$ sont échantillonnés avec une étape MCMC, le paramètre π est également échantillonné ; si nous nous situons au delà du temps de chauffe, les temps d'adsorptions ne sont pas échantillonnés.

Nous notons $t_i^{(j)}$, $f_i^{(j)}$ et $\pi^{(j)}$, les valeurs des chaînes MCMC.

Les figures 5.10 et 5.11 résument ces structures respectivement pendant et après le temps de chauffe. Il sera question d'un « vote » que nous détaillerons ci-dessous.

Qui plus est, lorsque nous sortons du temps de chauffe, nous échantillons tous les paramètres avec B étapes MCMC puis nous remplaçons leur valeur par la valeur moyenne estimée par les chaînes MCMC.

Nous proposons d'utiliser un échantillonneur RJMCMC pour modifier le nombre de molécules adsorbées N , et trois échantillonneurs de Metropolis-Hastings Q_1 , pour déplacer les temps t_i , Q_2 , pour modifier les fréquences f_i , et Q_3 pour échantillonner l'intensité de comptage π .

Etape RJMCMC L'étape RJMCMC consiste à échantillonner le modèle, c'est-à-dire le nombre de molécules adsorbées N . Nous allons faire reposer cette étape RJMCMC sur l'algorithme de Green (voir section 4.3.3.2). Nous mettons en place des changements de modèle qui consistent à :

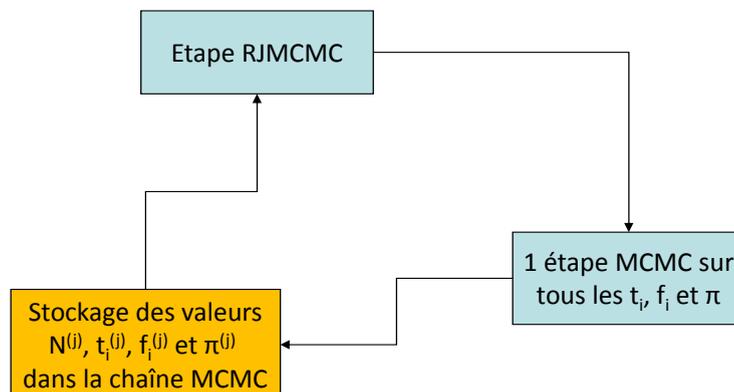


FIGURE 5.10 – Structure de l'algorithme hybride RJMCMC / MCMC mono-mode pendant le temps de chauffe

- soit ajouter un événement temporel ;
- soit supprimer un événement temporel.

Ce type d'approche est parfois dénommée BDMCMC (pour *Birth and Death Monte-Carlo Markov Chain*) [AO09]. Il existe évidemment d'autres changements de modèle qui permettent de mieux échantillonner la dimension du modèle et d'éviter de rester bloqué dans des *optima* locaux. Parmi eux, la possibilité de fusionner ou d'éclater les événements est décrite dans plusieurs articles [RFA03, MFM⁺12]. Toutefois, nous nous contenterons de ces deux mouvements.

Lors d'un changement de modèle « ajout d'un événement », nous proposons de tirer les paramètres créés avec leur loi *a priori* respective. Le détail des calculs des probabilités d'acceptation est donné en annexe E.4.

Etape MCMC Les étapes MCMC consistent à échantillonner les paramètres au sein d'un modèle de dimension $N^{(j)}$, la dimension courante, à savoir chacun des paramètres f_i et t_i , $i \in [1, N^{(j)}]$, ainsi que l'hyperparamètre sur l'intensité de comptage π .

Le premier de ces échantillonneurs sera l'échantillonneur de f_i , qui sera un échantillonneur à marche aléatoire avec une fonction de proposition normale, centrée sur le point courant et de variance δ_f :

$$Q_1(f_i^p | f_i^{(j-1)}) = \mathcal{N}(f_i^p | f_i^{(j-1)}, \delta_f) \quad (5.38)$$

Le second échantillonneur sera l'échantillonneur de t_i , qui sera :

- dans le cas où les temps d'adsorption sont situés sur la grille d'échantillonnage, un échantillonneur à marche aléatoire avec une fonction de proposition qui incrémentera ou décrémentera d'une unité les temps ;

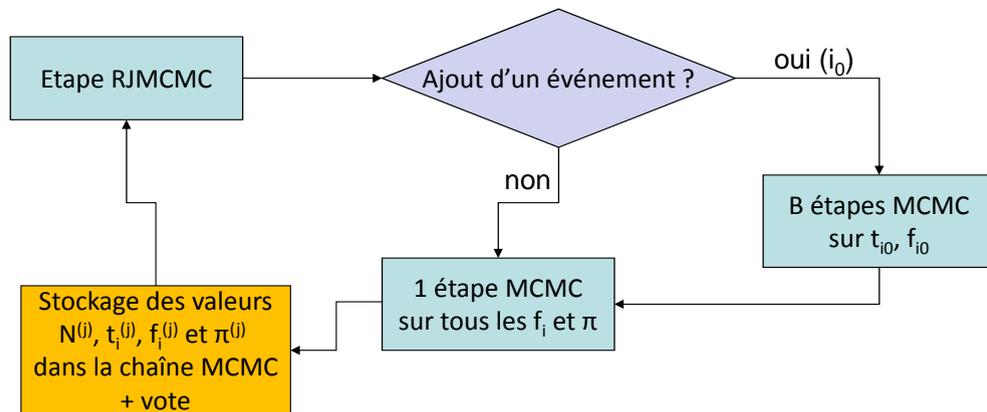


FIGURE 5.11 – Structure de l’algorithme hybride RJMCMC / MCMC mono-mode après le temps de chauffe

- dans le cas où les temps d’adsorption sont situés hors de la grille d’échantillonnage, un échantillonneur à marche aléatoire avec une fonction de proposition normale, centrée sur le point courant et de variance δ_t .

Soit, dans le premier cas :

$$Q_2(t_i^p | t_i^{(j-1)}) = 0.5 \delta(t_i^{(j-1)} + 1) + 0.5 \delta(t_i^{(j-1)} - 1) \quad (5.39)$$

Et dans le second cas :

$$Q_2(t_i^p | t_i^{(j-1)}) = \mathcal{N}(t_i^p | t_i^{(j-1)}, \delta_t) \quad (5.40)$$

Le troisième échantillonneur sera celui du paramètre π , pour lequel, compte-tenu du choix de la loi *a priori*, nous pouvons utiliser une étape de Gibbs avec échantillonnage explicite :

$$\begin{aligned} Q_3(\pi^p | \pi^{(j-1)}) &= Q_3(\pi^p) \\ &= p(\pi | N^{(j)}, T) \\ &= \mathcal{B}et(\pi | a_\pi + N^{(j)}, b_\pi + T - N^{(j)}) \end{aligned} \quad (5.41)$$

5.3.3.2 Le système de vote

Dans ce chapitre, nous visons à effectuer une étape de déconvolution impulsionnelle, c’est-à-dire à estimer un signal impulsionnel. Remarquons que cette estimation ne fait pas partie de notre

problématique de départ qui consiste à retrouver un spectre de masse. Cependant, pour pouvoir fournir un résultat intermédiaire (ce qui est l'objet de ce chapitre), nous proposons une solution pour estimer ledit signal impulsionnel.

Travail de Vincent Mazet Pour étudier ce point, il est intéressant, une fois de plus, de se pencher sur le travail de thèse de Vincent Mazet [Maz05, Ch. 3.5, Ch. 4.4]. Ce dernier distingue deux cas.

Dans le premier cas, le problème est exprimé sous forme de processus à temps discret (voir section 3.3.1.2) et à l'aide d'un vecteur \mathbf{q} . Dans ce cas, Vincent Mazet expose deux estimateurs, l'estimateur de Cheng *et al.* [CCL96] ainsi que l'estimateur de Rosec *et al.* [RBNC03], et en propose un nouveau. Dans tous les cas, l'idée générale est de sommer les vecteurs $\mathbf{q}^{(j)}$ entre eux puis de prendre une décision (par seuillage généralement) sur la détection d'un instant.

Dans le second cas, le problème est exprimé sous forme de processus ponctuel marqué (voir section 3.3.1.1), et la dimension du modèle d'entrée est fixée à un nombre suffisamment élevé. Chaque entrée de la liste formant le signal d'entrée est constituée d'un temps (l'équivalent de nos t_i), d'une amplitude (l'équivalent de nos f_i) et d'une variable binaire permettant de sélectionner ou non la ligne concernée. Ainsi, Vincent Mazet se ramène à un problème à dimension fixe (paramétrique).

Dans ce second cas, l'estimation se fait en suivant l'évolution de chacun des paramètres de la liste. Le point crucial ici est que les lignes peuvent permuter, c'est-à-dire que si à une itération j_1 le paramètre t_a vaut 13 et le paramètre t_b vaut 20, il est possible qu'à une itération j_2 le paramètre t_a vaille 20 et que le paramètre t_b vaille 13. Par conséquent, il n'est plus possible de faire une moyenne pour chaque paramètre. La solution proposée par Vincent Mazet est de ré-indexer les entrées de la liste.

Solution proposée Prenant en compte les difficultés exposées ci-dessus, nous avons proposé un système d'estimation simple, en deux étapes, décrites ci-dessus :

- une première étape de découverte ;
- une seconde étape d'estimation.

Au cours de la première étape, l'algorithme explore la loi *a posteriori*. Ceci correspond au « temps de chauffe ».

La seconde étape consiste à échantillonner le modèle et les paramètres et de faire voter à chaque itération (au sens des étapes RJMCMC) les temps estimés. Ainsi, nous créons un vecteur \mathbf{q} qui servira de vecteur de « vote ». A chaque itération du temps de chauffe, et pour chaque entrée de la liste estimée, nous notons u_0 le temps d'échantillonnage le plus proche du temps en question. Nous incrémentons de 1 le vecteur \mathbf{q} sur la composante u_0 .

La décision s'effectuera de deux manières, très simples. La première consiste à prendre comme instants d'échantillonnage les composantes de \mathbf{q} valant au moins 50% du nombre d'échantillons sur lesquels le cumul s'est effectué. La seconde consiste à calculer le nombre de molécules adsorbées selon le MMAP $\widehat{N}_{\text{MMAP}}$, c'est-à-dire en moyennant la chaîne $N^{(j)}$, puis à sélectionner les $\widehat{N}_{\text{MMAP}}$ valeurs les plus élevées de \mathbf{q} .

Notons que ce système est très largement sous-optimal. Notamment, nous ne pourrions pas prendre en considération deux événements très proches (il seront alors dans le même échantillon de q). Il est possible d'envisager alors des vecteurs q sur-échantillonnés. Nous nous contenterons d'un vecteur de dimension T .

5.4 Résultats

5.4.1 Définitions des critères de performance

Afin de pouvoir exprimer les performances de nos algorithmes en terme de détection mono-mode, nous proposons d'utiliser deux critères de performance.

La première de ces grandeurs est le taux de bonne détection T_{BD} ou « sensibilité ». Celui-ci est défini comme :

$$\text{sensibilité} = T_{BD} = \frac{\text{nombre d'événements détectés}}{\text{nombre total d'événements}} \quad (5.42)$$

La seconde de ces grandeurs sera le taux de faux positifs T_{FP} . Celui-ci est défini comme :

$$T_{FP} = \frac{\text{nombre d'événements détectés à tort}}{\text{nombre d'échantillons} - \text{nombre d'événements}} \quad (5.43)$$

Nous pouvons aussi utiliser la « spécificité » qui vaut 1 moins le taux de faux positifs.

$$\text{spécificité} = 1 - \frac{\text{nombre d'événements détectés à tort}}{\text{nombre d'échantillons} - \text{nombre d'événements}} \quad (5.44)$$

Calcul du taux de bonne et de fausse détection Pour calculer les valeurs définies ci-dessus, nous ferons l'appariement entre la liste estimée et la liste simulée selon le principe suivant :

- pour chaque instant de la vraie liste, trouver l'instant estimé le plus proche ;
- si cet instant estimé est plus proche d'un autre instant de la vraie liste, recommencer l'opération en retirant cet instant estimé ;
- si les deux instants (vrai et estimé) sont différents d'une quantité inférieure à γ_d , les instants sont appareillés puis retirés de leur liste respective (liste vraie et liste estimée) ;
- une fois une des deux listes épuisée, ou si il n'est plus possible d'appareiller deux instants :
 - le nombre d'instants appareillés est le *nombre d'événements détectés*,
 - le nombre d'éléments restant dans la liste estimée est le *nombre d'événements détectés à tort*.

Compromis et courbe ROC Ces deux grandeurs que sont la sensibilité et la spécificité sont étroitement liées. En effet, une configuration qui favorisera les détections aura tendance à augmenter la sensibilité. A l'inverse, une configuration plus sobre en terme de détection présentera une sensibilité réduite mais une spécificité plus élevée. Les cas extrêmes sont des algorithmes qui ne détectent soit rien (spécificité de 1 et sensibilité de 0) soit tout (spécificité de 0 et sensibilité de 1).

Lorsqu'une méthode présente un paramètre réglable susceptible de faire varier ces deux taux, il sera nécessaire de trouver un compromis entre sensibilité et spécificité, en fonction du cas applicatif considéré. Ainsi, les algorithmes dédiées aux applications RADAR (pour *RA*dio *D*etection

And Ranging, ou détection et télémétrie radio), par exemple, cherchent à avoir des spécificités constantes [RFKN92].

Pour visualiser ce compromis entre spécificité et sensibilité, nous allons représenter la relation entre ces deux grandeurs au travers d'une courbe ROC (pour *Receiver Operating Characteristic*), qui trace l'évolution de la sensibilité en fonction du taux de faux positifs.

Notons qu'un travail très similaire peut être trouvé dans [bou].

5.4.2 Résultats sur données simulées

5.4.2.1 Modalités de simulation

Nous allons tester notre méthode sur données simulées. Pour simuler les données, nous commençons par générer aléatoirement un signal d'observation sur T échantillons. Pour cela, nous générons N couples $\{t_i, f_i\}$, les t_i étant tirés aléatoirement entre 1 et T . Nous veillons notamment à ce qu'il n'y ait pas deux t_i identiques.

Nous générons ensuite le signal résultant de ce vecteur $f(t)$ en appliquant (3.12), page 59, puis en discrétisant le signal. Nous ajoutons à ce signal un bruit suivant une loi normale de variance σ^2 que nous intégrons (fonction Matlab *cumsum*), ce qui en fait un bruit en $1/f$ pur.

Les paramètres de simulation sont donc :

- T le nombre d'échantillons ;
- σ l'écart-type du bruit ;
- τ le temps de réponse de la boucle à verrouillage de phase (exprimé en périodes d'échantillonnage) ;
- N le nombre de molécules ;
- les valeurs des chutes de fréquence que nous préciserons plus loin.

Le tableau 5.1 donne les valeurs des paramètres de simulation.

T	σ	τ	N
200 échantillons	1	5 échantillons	10 adsorptions

TABLE 5.1 – Valeurs des paramètres de simulation mono-mode

5.4.2.2 Paramètres de l'algorithme

La méthode que nous avons proposée prend un certain nombre de paramètres, il s'agit notamment :

- du temps de chauffe, noté J_b ;
- du nombre total d'itérations J ;
- du nombre d'itérations MCMC B lorsqu'un événement est ajouté hors du temps de chauffe ;
- des paramètres des lois de proposition δ_f et δ_t ;
- des paramètres des lois *a priori* k_m et θ_m , a_π et b_π .

Les tableaux 5.2 et 5.3 donnent respectivement les paramètres des algorithmes et les paramètres des lois *a priori*.

J_b	J	B	δ_f	δ_t
3000	4000	500	20	1

TABLE 5.2 – Valeurs des paramètres de l’algorithme pour le traitement des données simulées mono-mode

k_m	θ_m	a_π	b_π
1	100	1	T

TABLE 5.3 – Valeurs des paramètres des lois *a priori* pour le traitement des données simulées mono-mode

5.4.2.3 Algorithmes de référence

Nous allons comparer notre méthode à plusieurs algorithmes de référence, qui fonctionnent comme processus à temps discret (voir section 3.3.1.2), et qui détecteront les temps d’adsorption :

1. une méthode par seuillage ;
2. une méthode par pseudo-inverse ;
3. une méthode basée sur le LASSO : l’algorithme LAR ;
4. une méthode basée sur *Matching-Pursuit* : l’algorithme CLEAN ;
5. une méthode Bernoulli-Gaussienne : l’algorithme SMLR.

Vu que le bruit est un bruit en $\ll 1/f \gg$, nous proposons que ces algorithmes traitent les signaux dérivés (fonction matlab *diff*), en utilisant comme réponse impulsionnelle la dérivée de réponse $u_{\text{PLL}}(t)$.

Seuillage La méthode par seuillage (voir section 5.2.1) est la méthode la plus basique. Celle-ci consiste à comparer le signal à un seuil, variable. Celui-ci variera entre 0 et 200 par pas de 1. Le seuillage effectué, nous proposons de détecter les « paquets » de détection, c’est-à-dire les ensembles d’instant successifs détectés, et de les réduire à un seul instant (le premier instant du paquet).

Pseudo-inverse La méthode par pseudo-inverse (voir section 5.2.3.2) consiste à appliquer la matrice pseudo-inverse sur le signal puis à le seuiller avec un seuil préalablement fixé. Celui-ci variera entre 0 et 200 par pas de 1.

LASSO La méthode LASSO (voir section 5.2.4.1) est une méthode pénalisée. Nous utilisons le code de Karl Sjöstrand accessible à l’adresse :

http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=3897

Le paramètre variable est le paramètre de régularisation, que nous ferons varier entre 0 et 1 par pas de 0.1 puis entre 1 et 100 par pas de 1.

Matching Pursuit La méthode *Matching Pursuit* (voir section 5.2.3.3) est une méthode de décomposition sur une base. Nous choisissons ici la base formée des réponses impulsionnelles décalées du pas d'échantillonnage. Nous utilisons le code de Joshua Carmichael accessible à l'adresse :

<http://www.mathworks.com/matlabcentral/fileexchange/27454-matchpurs-m>

Le paramètre variable sera le nombre de variables actives (le nombre d'instantés détectés), qui variera entre 0 et T , par pas de 1.

Bernoulli-Gaussien et SMLR Le modèle Bernoulli-Gaussien (voir section 5.2.4.3) est un modèle statistique de mélange. Nous avons implémenté le code décrit dans [Idi01, Ch. 5]. Le paramètre π (intensité de comptage) sera fixé à 0.05. Le paramètre variable sera le paramètre *a priori* de la variance des impulsions, auquel nous donnerons les valeurs 5^2 , 10^2 , 20^2 , 50^2 , 100^2 , 200^2 , 500^2 , 1000^2 , 2000^2 , 5000^2 , 10000^2 .

5.4.2.4 Résultats avec instants d'adsorption entiers

Nous allons désormais caractériser les performances dans le cas où les instants d'adsorption sont entiers. Ici, nous tirons uniformément les f_i entre 0 et 100 Hz. Nous moyennons les performances sur 50 répétitions. Le seuil pour l'appariement des instants vaut $\gamma_d = 1$ échantillon.

Nous entendons par « MCMC1 » la méthode consistant à effectuer la détection en utilisant l'estimation de l'ordre du modèle au sens du MMAP, et « MCMC2 » la méthode consistant à effectuer la détection par seuillage du vecteur \mathbf{q} .

La courbe ROC résultante est donnée par la figure 5.12 (la partie droite est un zoom sur les performances de notre méthode).

Le tableau 5.4 donne les temps de calcul moyens des différents algorithmes (*Seuil*, *PI* pour « Pseudo-inverse », *LASSO*, *MP* pour « *Matching Pursuit* », *SMLR*, *MCMC1* et *MCMC2*). Dans le cas de la pseudo-inverse et de l'algorithme LASSO, il est possible de mettre en commun le calcul pour différents paramètres de régularisation. Nous n'en avons pas tenu compte dans ce tableau.

Seuil	PI	LASSO	MP	SMLR	MCMC1	MCMC2
0.04 ms	15.3 ms	93.0 ms	104.2 ms	1.3 s	106.0 s	106.0 s

TABLE 5.4 – Temps de calcul des différents algorithmes en mono-mode avec instants d'adsorption entiers

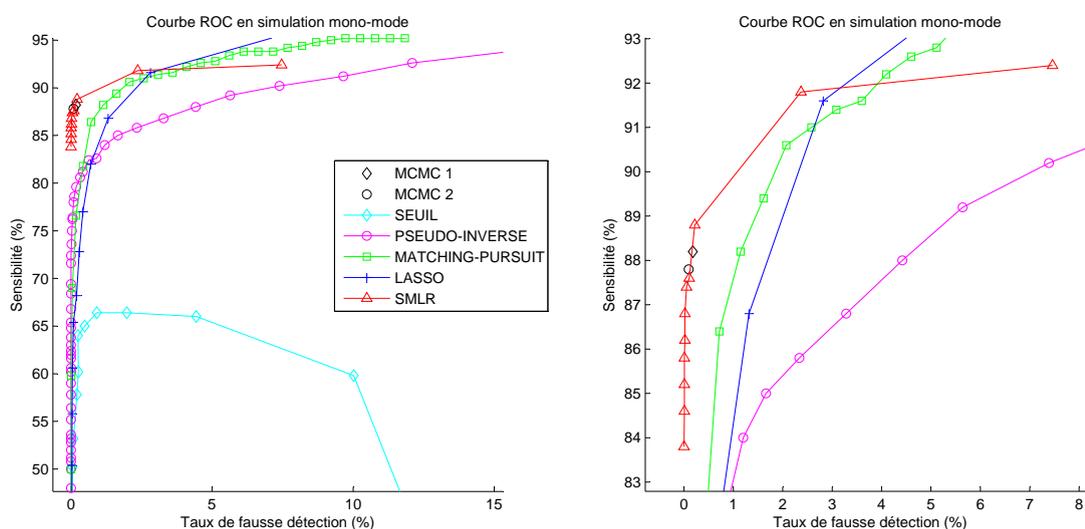


FIGURE 5.12 – Courbe ROC mono-mode pour une simulation avec $N = 10$ événements et des instants d’adsorption entiers

5.4.2.5 Résultats avec instants d’adsorption non-entiers

Nous allons désormais caractériser les performances dans le cas où les instants d’adsorption ne sont pas entiers. Ici, nous tirons uniformément les f_i entre 0 et 100 Hz. Nous moyennons les performances sur 50 répétitions. Le seuil pour l’appariement des instants vaut $\gamma_d = 2$ échantillons.

La courbe ROC résultante est donnée par la figure 5.13 (la partie droite est un zoom sur les performances de notre méthode). Le tableau 5.5 donne les temps de calcul moyens des différents algorithmes.

Seuil	PI	LASSO	MP	SMLR	MCMC1	MCMC2
0.03 ms	9.4 ms	95.2 ms	109.6 ms	1.3 s	110.6 s	110.6 s

TABLE 5.5 – Temps de calcul des différents algorithmes en mono-mode avec instants d’adsorption non-entiers

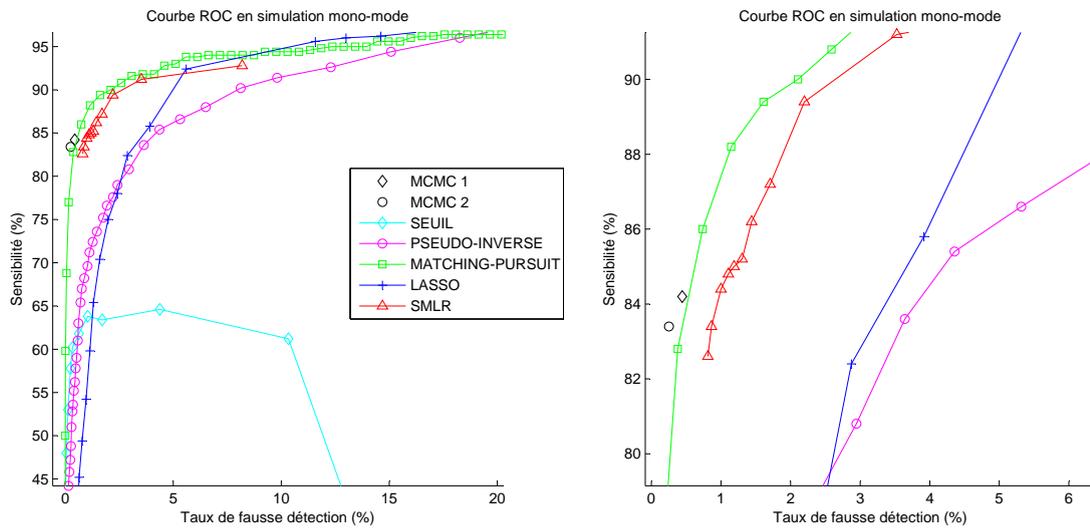


FIGURE 5.13 – Courbe ROC mono-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption non entiers

5.4.3 Validation de la méthode sur données réelles

Validons désormais notre méthode sur données réelles. Nous allons ici traiter un exemple, issu de sessions expérimentales de Décembre 2012, sur des agrégats de Tantale d'un diamètre nominal de 6.5 nm, soit une masse nominale d'environ 1420 kDa.

Le tableau 5.6 donne les valeurs des paramètres du modèle direct, à savoir le nombre d'échantillons T , la période d'échantillonnage T_e et le temps de réponse τ .

T	T_e	τ
5000 échantillons	2 ms	10 ms

TABLE 5.6 – Valeurs des paramètres du modèle direct pour les données réelles mono-mode

Nous prendrons un filtre blanchisseur d'ordre $V = 5$. Les coefficients du filtre sont donnés dans le tableau 5.7 :

Mode 1	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	1.7989	-1.1393	0.0358	0.0374	-0.1959

TABLE 5.7 – Coefficients des filtres blanchisseurs pour le traitement des données réelles mono-mode

Les tableaux 5.8 et 5.9 donnent respectivement les paramètres des algorithmes et les paramètres des lois *a priori*.

J_b	J	B	δ_f	δ_t
4000	5000	100	20	1

TABLE 5.8 – Valeurs des paramètres de l'algorithme pour le traitement des données réelles mono-mode

k_m	θ_m	a_π	b_π
1	500	1	T

TABLE 5.9 – Valeurs des paramètres des lois *a priori* pour le traitement des données réelles mono-mode

La figure 5.14 illustre la détection des instants de chute en les identifiant par un cercle rouge. La méthode de décision retenue ici est la méthode de seuillage du vecteur \mathbf{q} .

5.4.4 Discussion

Nous venons de tester notre méthode, de la confronter à des algorithmes alternatifs et de la tester sur données réelles.

Rappelons que nous voulions :

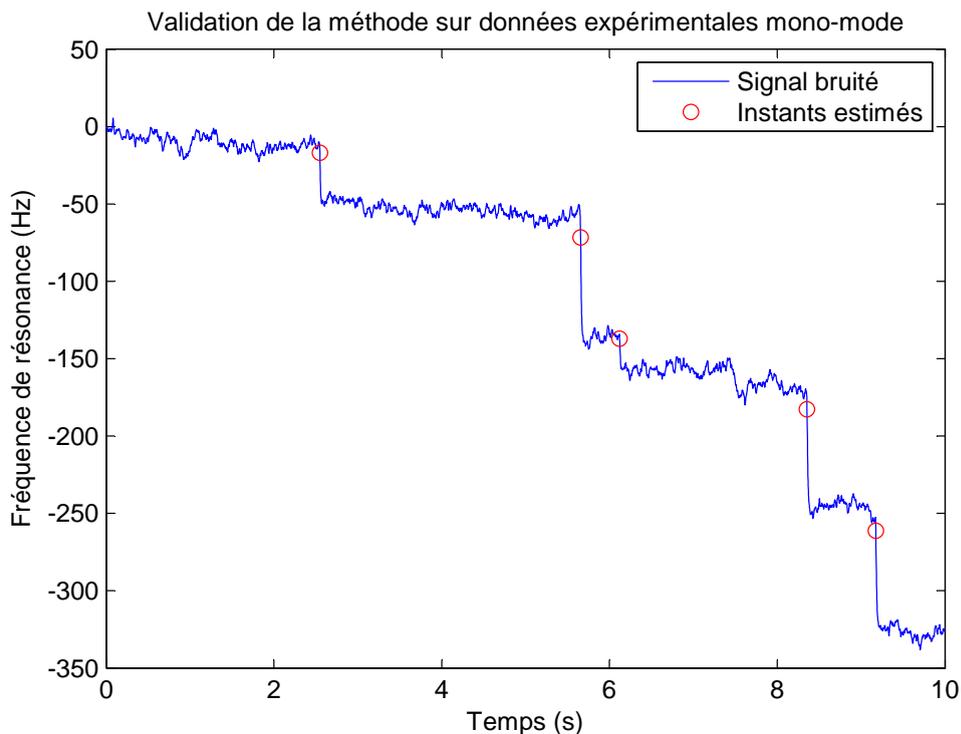


FIGURE 5.14 – Illustration de la détection mono-mode sur un signal expérimental

- ne pas nous contraindre à utiliser la loi *a priori* normale sur les amplitudes des chutes de fréquence ;
- permettre des instants d'adsorption non-entiers ;
- estimer automatiquement certains hyperparamètres.

Si la problématique peut paraître relativement simple sur données réelles, les chutes de fréquence étant espacées et de grande amplitude, nous avons voulu tester plus en profondeur nos algorithmes en augmentant significativement le taux de comptage lors des simulations.

En simulation, on note un bon comportement de l'algorithme en terme de détection, si l'on cherche à maintenir le taux de fausse détection relativement bas (inférieur à 0.5%, par exemple). L'algorithme se basant sur le seuillage de q est le plus prudent en terme de détection.

Lorsque les instants d'échantillonnage sont entiers, notre algorithme présente des performances similaires à celles de l'algorithme SMLR, qui est le meilleur algorithme ici. Un avantage de notre algorithme ici est le calcul automatique de l'intensité de comptage π . Le prix à payer est le coût calculatoire, la méthode MCMC étant presque 100 fois plus lente que la méthode SMLR,

pour des performances de détection comparables. Notons que l'algorithme SMLR exploite la possibilité de mettre à jour de manière récursive de la fonction de vraisemblance, ce que nous n'avons pas exploité dans le cadre de notre algorithme. Si nous ignorons cette amélioration, le temps de calcul moyen du SMLR (pour une variance de 10000^2 et instants d'adsorption entiers) se monte à 114.0 s. Ceci nous amène à croire que notre algorithme pourrait grandement bénéficier de cette mise à jour récursive, et pourrait, à défaut d'obtenir des temps de calcul inférieurs aux autres algorithmes, présenter des temps de calcul comparables au SMLR.

Lorsque les instants d'échantillonnage ne sont pas entiers, aucun algorithme de référence n'obtient des performances de détection comparables au notre. Ceci est tout à fait cohérent avec le fait que nous avons développé notre algorithme pour pouvoir prendre en compte ce phénomène non-entier. Les autres algorithmes pêchent dans ce cas présent car ces derniers essaient de placer deux chutes de fréquence pour un chute de fréquence vraie, ce qui est dû à l'effet du caractère non-entier des instants d'adsorption sur la réponse du système dérivée (voir figure 3.5 notamment). Notons que des améliorations du modèle Bernoulli-Gaussien dans le cas non-entier sont proposées dans l'état de l'art, comme [Car13] qui propose un sur-échantillonnage, ou [DID97] qui propose un extension continue (modèle Poisson-Gaussien). Une comparaison avec ces méthodes ne pourrait être que bénéfique.

On notera globalement un bon comportement de l'algorithme basé sur *Matching-Pursuit*, qu'il faut modérer ici. En effet, le nombre d'adsorptions sur nos simulations est constant et le paramètre variable pour cet algorithme est justement le nombre d'adsorptions. Dans un autre contexte, il serait plus cohérent d'utiliser comme paramètre variable l'erreur de reconstruction pour un ensemble de variables actives donné.

Aussi, nous ne pouvons pas ici exposer les avantages du premier point cité ci-dessus, à savoir l'utilisation d'une autre loi *a priori* sur les amplitudes des chutes de fréquence que la loi normale. Par contre, ce point prendra tout son sens au chapitre suivant.

Sur données réelles, l'algorithme détecte toutes les chutes de fréquence visibles à l'œil. Ceci a le mérite de valider notre modèle et de prouver que notre travail en simulation n'est pas déconnecté de tout sens physique.

En résumé, nous venons de mettre en place un algorithme de détection efficace en termes de performances de détection, bien que relativement lent. Notre algorithme est comparable au meilleur algorithme auquel nous nous comparons, le SMLR en l'occurrence. La méthode que nous proposons ici présente l'avantage d'offrir une grande liberté de modélisation. Ainsi, elle va nous servir de base pour mettre en place un algorithme plus spécifique à notre problème.

Le travail restant à accomplir est donc la dérivation de ce travail à la déconvolution multi-mode, ce à quoi nous adjoindrons la quantification des masses adsorbées.

Traitement multi-mode – Détection et quantification

Dans ce chapitre, nous allons étendre la déconvolution impulsionnelle vue précédemment au cas multi-mode. Comme précédemment, nous introduirons le cadre de travail, définirons des méthodes de référence puis développerons notre propre contribution que nous testerons.

Sommaire de ce chapitre

6.1	Cadre de travail	143
6.2	Méthodes de référence	145
6.2.1	État de l'art – méthodes publiées	145
	Méthode de recherche d'une solution aux équations non-linéaires	145
	Méthode probabiliste par changement de variable	145
6.2.2	Adaptation des méthodes de déconvolution mono-mode	146
	Adaptation simple des méthodes de déconvolution impulsionnelle mono-mode	146
	Adaptations plus raffinées des méthodes de déconvolution impulsionnelle mono-mode	146
6.2.3	Adaptation d'un modèle Bernoulli-Gaussien	146
6.3	Développement d'une méthode d'inversion probabiliste multi-mode	148
6.3.1	Justification de l'approche employée	148
	6.3.1.1 Un problème de plus en plus mal posé	148
	6.3.1.2 Incertitude sur les chutes de fréquence observées	148
	6.3.1.3 Structurer la distribution des chutes de fréquences	149
6.3.2	Modélisation statistique des grandeurs	149
	6.3.2.1 Expression de la fonction de vraisemblance	150
	6.3.2.2 Définition des lois <i>a priori</i>	151
	Loi de probabilité sur l'ordre du modèle	151
	Loi de probabilité sur les temps d'adsorption	151
	Loi de probabilité sur la masse	151
	Loi de probabilité sur la position	151
	Loi de probabilité sur l'intensité de comptage	152
	6.3.2.3 Écriture des lois <i>a posteriori</i> des paramètres de la liste pour un ordre de modèle donné	152

6.3.2.4	Écriture des autres lois <i>a posteriori</i> conditionnelles	152
6.3.3	Estimation	153
6.3.3.1	Utilisation d'un algorithme hybride RJMCMC MCMC	153
6.3.3.2	Le système de vote	154
6.4	Résultats	155
6.4.1	Définitions des critères de performance	155
6.4.2	Résultats sur données simulées	155
6.4.2.1	Modalités de simulation	155
6.4.2.2	Paramètres de l'algorithme	156
6.4.2.3	Algorithmes de référence	156
	Adaptation simple des méthode de déconvolution	156
	Méthode de déconvolution spécifique au cas multi-mode	157
6.4.2.4	Résultats avec instants d'adsorption entiers	157
6.4.2.5	Résultats avec instants d'adsorption non-entiers	158
6.4.3	Validation de la méthode sur données réelles	160
6.4.4	Discussion	161

6.1 Cadre de travail

Le cadre de travail dans lequel nous allons nous inscrire sera le cadre de la déconvolution impulsionnelle associée à l'estimation de la masse de chacune des molécules. Ceci correspond aux parties « Détecter » et « Quantifier » identifiées dans la section 3.5.

Dans ce cadre-ci, il s'agit non seulement d'inverser l'opération de convolution que l'on retrouve dans l'équation (3.11) et de prendre en compte le fait que l'information se répète sur plusieurs signaux, mais aussi d'inverser les équations non-linéaires $f_{i,k} = \alpha_k m_i \phi_k(z_i)$ afin de retrouver la masse m_i .

Nous travaillerons avec les hypothèses suivantes :

- $K = 2$ harmoniques sont observées. Nous exploiterons la simultanéité des chutes de fréquences sur les différents modes. Cette fois-ci, il est possible d'estimer m_i et z_i .
- Les réponses du système à une adsorption $u_{\text{PLL},k}(t)$ sont connues, comme précédemment. Ces réponses dépendent à la fois du paramètre de temps de réponse τ mais aussi du décalage entre les modes κ .
- La loi de probabilité du bruit est connue, comme précédemment. Ici, les paramètres de bruit sont les coefficients des filtres blanchisseur pour le mode 1 \mathbf{a}_1 et pour le mode 2 \mathbf{a}_2 ainsi que la matrice de covariance inter-mode $\mathbf{\Gamma}_m$.
- La géométrie des capteurs sera une géométrie de type *double-clamped*.

Nous décomposerons les signaux temporels $m_k(t)$ et $z_k(t)$ comme suit :

$$m(t) = \sum_{i=1}^N m_i \delta(t_i) \quad (6.1)$$

$$z(t) = \sum_{i=1}^N z_i \delta(t_i) \quad (6.2)$$

avec N le nombre d'adsorptions sur la fenêtre d'observation. À partir de ces signaux, il est possible de générer les signaux $f_k(t)$, voir sections 3.3.1.1 et 3.3.1.2. Il est également possible de reconstruire \mathbf{g} par convolution et échantillonnage.

Nous introduirons également le paramètre π , qui sera inconnu et qui représente l'intensité de comptage (ou densité temporelle d'adsorptions).

Comme dans le chapitre précédent, nous représentons les connaissances *a priori* sur le bruit par le jeu de paramètres θ_ϵ , celles sur π par θ_π , celles sur les m_i par θ_m et celles sur les z_i par θ_z . Le nombre de molécules N dépend de π et de T , et les temps t_i ne dépendent que de T .

Ce cadre de travail peut se représenter au travers de la modélisation hiérarchique illustrée par la figure 6.1.

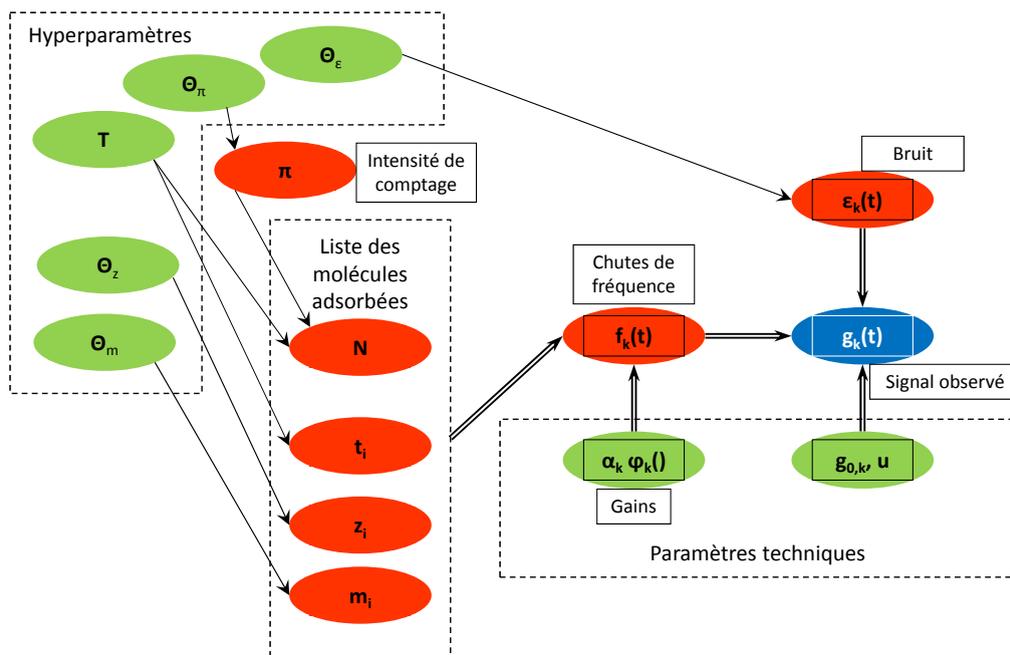


FIGURE 6.1 – Modélisation hiérarchique du problème de détection – quantification

6.2 Méthodes de référence

6.2.1 État de l'art – méthodes publiées

Dans le cadre de notre recherche bibliographique, nous n'avons pas trouvé de document traitement spécifiquement de la détection de chutes de fréquences simultanées sur un signal issu d'un capteur de masse NEMS.

Si aucune information n'est disponible en ce qui concerne la détection des chutes de fréquence, il est possible de trouver des références en ce qui concerne la quantification de la masse de la molécule adsorbée étant données les chutes de fréquence subséquentes $\Delta f_{k,m,z}$. Ce problème sera plus détaillé dans l'annexe D.

Méthode de recherche d'une solution aux équations non-linéaires La première méthode est celle publiée par Dohn *et al.* [DSBH07].

Cette méthode consiste à chercher une solution au système d'équations suivant :

$$\begin{aligned}\Delta f_{1,m,z} &= -\alpha_1 \phi_1(z) m \\ \Delta f_{2,m,z} &= -\alpha_2 \phi_2(z) m\end{aligned}\tag{6.3}$$

Sous certaines hypothèses sur les fonctions $\phi_k(z)$, le système admet une unique solution.

L'idée de l'article de Dohn *et al.* est alors de trouver le couple $\{m, z\}$ qui satisfasse à ces équations. La méthode de résolution proposée dans l'article est une méthode graphique, déclinée dans le cas de la géométrie *cantilever* en utilisant quatre harmoniques.

De plus amples détails sont donnés en annexe D.2.

Méthode probabiliste par changement de variable La seconde méthode que nous avons référencée est celle publiée par Hanay *et al.*, une équipe de Caltech [HKN⁺12].

Cette méthode consiste à voir les chutes de fréquence comme des variables aléatoires, puis à effectuer un changement de variable pour transformer les chutes de fréquence $\Delta f_{k,m,z}$ en $\{m, z\}$. Ce changement de variable nécessite notamment le calcul du déterminant d'une matrice jacobienne, qui doit être non nul.

Dans l'article, cette densité de probabilité jointe est ensuite estimée sur une grille dans l'espace des paramètres $\{m, z\}$. S'en déduisent par sommation les lois *a posteriori* marginales de chacun des paramètres.

De plus amples détails sont donnés en annexe D.3.

6.2.2 Adaptation des méthodes de déconvolution mono-mode

Comme nous l'avons exprimé dans le paragraphe précédent, l'état de l'art dans le domaine des NEMS ne contient pas d'information expliquant comment effectuer une détection multi-mode. Il est toutefois possible d'adapter des méthodes de déconvolution impulsionnelle mono-mode pour le cas étudié ici.

Adaptation simple des méthodes de déconvolution impulsionnelle mono-mode Remarquons dans un premier temps que nous pouvons facilement utiliser les méthodes de déconvolution impulsionnelle mono-mode dans notre cas pour effectuer une étape de détection. Nous identifions trois cas d'utilisation :

1. effectuer la déconvolution sur chacun des modes séparément, puis faire l'union des instants estimés (opération « OU »);
2. effectuer la déconvolution sur chacun des modes séparément, puis faire l'intersection des instants estimés (opération « ET »);
3. effectuer la déconvolution sur la moyenne des modes (éventuellement pondérés par l'inverse-variance du bruit).

Adaptations plus raffinées des méthodes de déconvolution impulsionnelle mono-mode Il est toutefois possible d'aller plus loin que cette utilisation basique des méthodes de déconvolution impulsionnelle. Des extensions de ces méthodes se retrouvent notamment sous la dénomination *Collaborative sparse coding*, que l'on peut traduire par « codage parcimonieux collaboratif ». Il existe aussi la problématique de parcimonie structurée (ou *Structured sparsity*), qui est légèrement différente dans la mesure où il s'agit d'être parcimonieux à plusieurs niveaux (impulsions et groupes d'impulsions) dans un signal.

Ainsi, des adaptations de l'algorithme *Orthogonal Matching-Pursuit*, que nous avons vu à la section 5.2.3.3, ont été proposées, au travers de l'algorithme *Simultaneous Orthogonal Matching-Pursuit* (S-OMP) [TGS05], par exemple.

L'algorithme LASSO, que nous avons vu à la section 5.2.4.1, a également été dérivé, au travers de l'algorithme *C-HiLasso* [SRSE11], par exemple. On peut également citer l'algorithme *Group LASSO* [WNF09].

Pour de plus amples détails, nous invitons les lecteurs intéressés à se référer aux deux articles de Tropp *et al.* [TGS06, Tro06] qui balaisent le sujet avec une vision large.

6.2.3 Adaptation d'un modèle Bernoulli-Gaussien

Si le LASSO et *Orthogonal Matching-Pursuit* ont été adaptées au cas multi-mode, les modèles Bernoulli-Gaussien, que nous avons vus à la section 5.2.4.3, l'ont été également [DM92].

Prenons un cas simple, que nous adopterons par la suite. Supposons notamment que le bruit sur les harmoniques est indépendant, identiquement distribué sous une loi normale centrée de variance σ^2 . Prenons le modèle *a priori* suivant, en notant $\mathbf{f}_{i,k}$ le vecteur des chutes de fréquence sur la k -ième harmonique :

$$\begin{aligned} p(\mathbf{f}_{i,k} | \mathbf{q}_i = 0) &= \delta(\mathbf{f}_{i,k}) \\ p(\mathbf{f}_{i,k} | \mathbf{q}_i = 1) &= \mathcal{N}(\mathbf{f}_{i,k} | \mathbf{0}, r_k) \end{aligned} \quad (6.4)$$

Ceci revient à dire que l'on structure les chutes de fréquences à l'aide une loi normale centrée de matrice de covariance diagonale, le k -ième élément de la diagonale valant r_k .

Ce modèle peut se mettre sous forme hiérarchique, comme illustré par la figure 6.2.

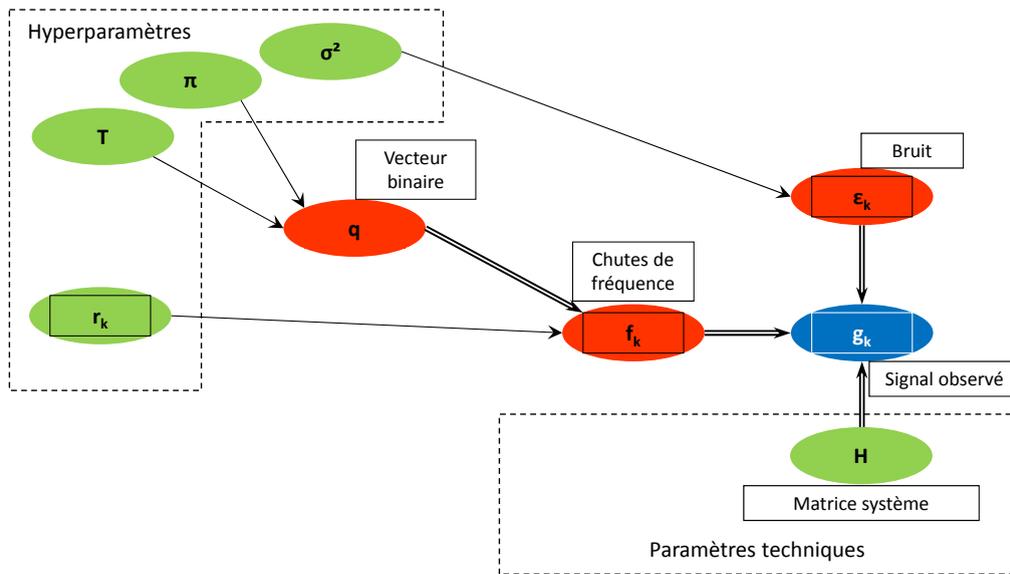


FIGURE 6.2 – Vision hiérarchique du modèle Bernoulli-Gaussien multi-mode

Dans ce cas, on a :

$$p(\mathbf{q} | \mathbf{g}_k)_{k \in 1, K} = p(\mathbf{q}) \prod_{k=1}^K \mathcal{N}(\mathbf{g}_k | \mathbf{0}, \mathbf{R}_k) \quad (6.5)$$

avec :

$$\mathbf{R}_k = r_k \mathbf{H} \mathbf{\Pi} \mathbf{H}^T + \sigma^2 \mathbf{I} \quad (6.6)$$

avec $\mathbf{\Pi}$ une matrice diagonale dont la diagonale est le vecteur \mathbf{q} : $\mathbf{\Pi} = \text{diag}(\mathbf{q})$.

6.3 Développement d'une méthode d'inversion probabiliste multi-mode

6.3.1 Justification de l'approche employée

Voyons ici les éléments qui nous ont conduit au développement de notre méthode :

1. le problème est en l'état (très) mal-posé ;
2. nous voulons intégrer toutes les variabilités de la chaîne de mesure ;
3. nous voudrions structurer les informations *a priori* sur les amplitudes des chutes de fréquence.

6.3.1.1 Un problème de plus en plus mal posé

Comme précédemment, nous allons nous placer dans le cadre de travail bayésien.

La justification principale reste inchangée, à savoir que nous sommes face à un problème de déconvolution impulsionnelle qui est par essence mal-posé, comme nous l'avons vu dans la section 5.3.1.1.

Nous pouvons toutefois adjoindre une justification supplémentaire à la nécessité de régulariser l'inversion, dans la mesure où l'estimation de la masse est elle aussi mal-posée. En effet, rien ne dit que la transformation non-linéaire qui à une masse et à une position associe des chutes de fréquences soit bijective. Cela dépend notamment du nombre de chutes de fréquences (harmoniques) observées et de la géométrie du capteur.

Nous avons vu précédemment une méthode de résolution, celle de Dohn *et al.* [DSBH07]. Ces auteurs ne précisent pas quelle démarche employer si la transformation n'est pas bijective (surdétermination si trop d'harmoniques sont observées, sous-détermination sinon).

Le travail publié par Hanay *et al.*, [HKN⁺12] utilise quant à lui un changement de variable. Ce changement de variable est assez simple dans le cas étudié dans l'article pour des raisons de bijectivité, qui permettent d'effectuer le changement de variable au travers de l'utilisation du jacobien. Si la transformation n'est pas bijective, par contre, le changement de variable ne peut plus s'exprimer de cette manière et le calcul peut être autrement plus complexe, avec notamment la nécessité de repasser par les fonctions de répartition.

Dans notre cas, la configuration géométrique sera la configuration *double-clamped* et nous travaillerons sur $K = 2$ harmoniques. Toutefois, si dans le futur les modalités d'acquisition venaient à changer, en ajoutant des harmoniques par exemple, nous aimerions pouvoir prendre en compte facilement cette modification.

6.3.1.2 Incertitude sur les chutes de fréquence observées

Un second point est celui des hypothèses formulées dans l'état de l'art : les articles travaillent en aval d'une observation des chutes de fréquence $\Delta f_{k,m,z}$. A ce titre, le travail publié par Hanay

et al., [HKN⁺12] permet de prendre en compte le bruit de mesure au travers d'une modélisation normale. Ceci serait exact si les perturbations n'étaient qu'une perturbation additive *via* un bruit suivant une loi normale agissant sur la chute de fréquence estimée. Malheureusement, la mesure des chutes de fréquence est certes perturbée par le bruit d'une part, que l'on peut prendre en considération, mais aussi par les erreurs liées à l'étape de déconvolution impulsionnelle qui permet d'estimer ces chutes de fréquence.

Nous pensons que notre algorithme « tout intégré », qui marginalise toutes les sources d'incertitudes, en prenant comme référence le signal observé, va permettre une quantification plus réaliste des incertitudes de notre problème, et les propager le long du modèle.

6.3.1.3 Structurer la distribution des chutes de fréquences

La dernière justification concerne les capacités de détection de notre algorithme.

En effet, la distribution des chutes de fréquence est reliée à la distribution de la masse et de la position. A l'instar de notre travail précédent en déconvolution mono-mode, au cours duquel nous voulions imposer la positivité des chutes de fréquence, nous cherchons ici à structurer la loi *a priori* des chutes de fréquence à travers notre modèle physique.

La figure 6.3 illustre l'exemple de la distribution de chutes de fréquence, pour 1000 molécules dont la masse est tirée aléatoirement entre 0 et 100 et la position tirée entre 0 et 0.5, pour une configuration géométrique *double-clamped*.

Pour prendre en compte cette structuration, deux approches s'offrent à nous. La première consiste à approcher cette distribution avec une loi de probabilité plus simple, permettant un temps de calcul moindre (comme une loi normale, ce qui permet d'utiliser le modèle Bernoulli-Gaussien multi-dimensionnel décrit précédemment dans la section 6.2.3).

La seconde approche consiste à intégrer les chutes de fréquence dans un modèle hiérarchique afin de contraindre ces dernières par d'autres variables, en l'occurrence la masse et la position. Par conséquent, les chutes de fréquence n'interviennent pas explicitement dans le modèle mais sont structurées à travers les informations *a priori* sur la masse et sur la position.

C'est cette dernière approche que nous allons retenir, car elle nous permettra de surcroît de remonter directement à la masse des molécules adsorbées.

6.3.2 Modélisation statistique des grandeurs

Voyons maintenant comment nous allons modéliser les différentes grandeurs à l'aide de lois de probabilité. Nous allons reprendre en grande partie le travail effectué au chapitre précédent, en remplaçant f_i par m_i et z_i . Les grandeurs inconnues dans notre cas sont l'intensité de comptage π , et les signaux $m(t)$ et $z(t)$ qui sont paramétrés par le nombre d'événements N , les temps d'adsorption t_i , les amplitudes des masses adsorbées m_i et les positions des molécules adsorbées z_i :

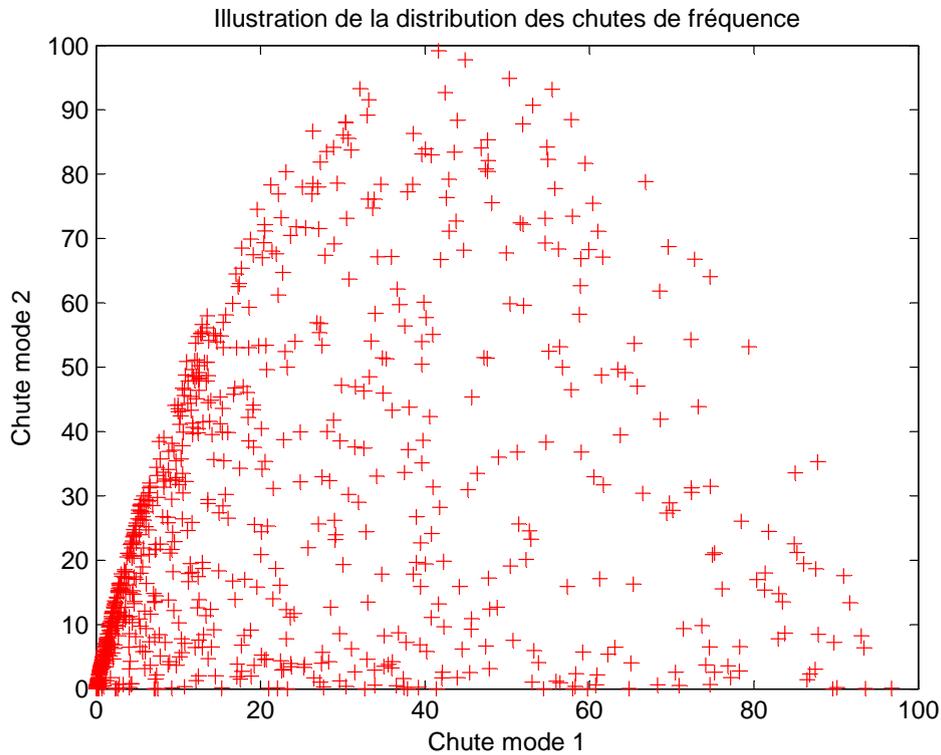


FIGURE 6.3 – Illustration de la structure des chutes de fréquence

$$\begin{aligned}
 N &\in \mathbb{N} \\
 t_i &\in [0, T]^N \\
 m_i &\in \mathbb{R}^+{}^N \\
 z_i &\in [0, 1]^N \\
 \pi &\in [0, 1]
 \end{aligned} \tag{6.7}$$

6.3.2.1 Expression de la fonction de vraisemblance

Comme nous l'avons exprimé à la section 3.4.1 puis au chapitre précédent, section 5.3.2.1, nous modélisons nos connaissances sur les incertitudes de mesure au travers d'un bruit additif, stationnaire et ergodique qui suit une loi normale centrée de covariance $\mathbf{\Gamma}$. Dans notre cas, il s'agira de concaténer les signaux observés d'une part et les signaux estimés d'autre part afin de pouvoir les inclure dans une vraisemblance normale (voir section 3.4.1).

Nous reprendrons la notation précédente afin de ne pas alourdir les équations. Nous noterons donc la vraisemblance :

$$p(\mathbf{g}_k | \mathbf{m}(t), \mathbf{z}(t)) = \mathcal{N} \left(\mathbf{g}_k | \alpha_k (\mathbf{m}(t) \diamond \phi_k(\mathbf{z}(t))) \star u_{\text{PLL}}(t), \mathbf{\Gamma} \right)_{k \in \{1;K\}} \quad (6.8)$$

Ici, $\underline{\cdot}$ symbolise la discrétisation du signal puis la mise en forme sous forme de vecteur (concaténation des différentes harmoniques), voir section 3.4.2.2 pour plus de détails.

6.3.2.2 Définition des lois *a priori*

Loi de probabilité sur l'ordre du modèle Comme précédemment, compte-tenu que nous nous inspirons du modèle Bernoulli-Gaussien, nous considérons que l'ordre du modèle N suit la même loi que la somme de variables de Bernoulli indépendantes, c'est-à-dire une loi Binomiale de paramètres π et T :

$$p(N | \pi) = \mathcal{B}(N | \pi, T) \quad (6.9)$$

Loi de probabilité sur les temps d'adsorption Comme précédemment, nous modéliserons les temps d'adsorption t_i comme uniformément répartis sur $\{1, T\}$ ou sur $[0, T]$ selon si les temps d'adsorption sont entiers (processus à temps discret) ou libres (processus ponctuel marqué), voir 3.3.1.1 et 3.3.1.2 :

$$p(t_i | \pi) = \mathcal{U}_{\{1, T\}}(t_i) \quad (6.10)$$

$$\text{ou } p(t_i | \pi) = \mathcal{U}_{[0, T]}(t_i) \quad (6.11)$$

Loi de probabilité sur la masse Les masses adsorbées m_i seront distribuées sous une loi de support positif. Nous proposons de prendre la loi Gamma :

$$p(m_i) = \mathcal{G}am(m_i | k_m, \theta_m) \quad (6.12)$$

où k_m , et θ_m sont respectivement les paramètres de forme et d'échelle de la loi Gamma.

Loi de probabilité sur la position Nous modéliserons les positions d'adsorption z_i comme uniformément réparties soit sur $[0, 1]$ dans le cas *cantilever*, soit sur $[0, 0.5]$ dans le cas *double-clamped* (compte-tenu de la symétrie du problème). Ici, nous sommes en présence d'une géométrie *double-clamped* :

$$p(z_i) = \mathcal{U}_{[0, 0.5]}(z_i) \quad (6.13)$$

Loi de probabilité sur l'intensité de comptage Comme précédemment, le paramètre d'intensité de comptage, π , sera distribué selon une loi Bêta. Cette loi, à support $[0, 1]$, est la loi *a priori* conjuguée pour le paramètre d'une loi Binomiale (voir tableau 4.1, page 89) :

$$p(\pi) = \mathcal{B}et(\pi|a_\pi, b_\pi) \quad (6.14)$$

où a_π et b_π sont les paramètres de la loi Bêta.

6.3.2.3 Écriture des lois *a posteriori* des paramètres de la liste pour un ordre de modèle donné

Compte-tenu de la modélisation formulée ci-dessus, nous pouvons écrire la probabilité *a posteriori* jointe de tous les paramètres de la liste sachant un modèle N donné (ces paramètres sont indépendants de π conditionnellement à N), dans le cas où les temps d'adsorption sont entiers :

$$p(t_1, \dots, t_N, m_1, \dots, m_N, z_1, \dots, z_N | N, \mathbf{g}) \propto \mathcal{N}\left(\mathbf{g}_k | \underline{\alpha_k(m(t) \diamond \phi_k(z(t))) \star u_{PLL}(t)}, \mathbf{\Gamma}\right)_{k \in \{1; K\}} \prod_{i=1}^N \mathcal{U}_{\{1, T\}}(t_i) \mathcal{U}_{[0, 0.5]}(z_i) \mathcal{G}am(m_i | k_m, \theta_m) \quad (6.15)$$

et dans le cas où les temps d'adsorption ne sont pas entiers :

$$p(t_1, \dots, t_N, m_1, \dots, m_N, z_1, \dots, z_N | N, \mathbf{g}) \propto \mathcal{N}\left(\mathbf{g}_k | \underline{\alpha_k(m(t) \diamond \phi_k(z(t))) \star u_{PLL}(t)}, \mathbf{\Gamma}\right)_{k \in \{1; K\}} \prod_{i=1}^N \mathcal{U}_{[0, T]}(t_i) \mathcal{U}_{[0, 0.5]}(z_i) \mathcal{G}am(m_i | k_m, \theta_m) \quad (6.16)$$

6.3.2.4 Écriture des autres lois *a posteriori* conditionnelles

Aussi, il est possible d'écrire la loi *a posteriori* conditionnelle à π du modèle N :

$$p(N | \pi, \mathbf{g}) \propto \mathcal{B}(N | \pi, T) \int p(t_1, \dots, t_N, m_1, \dots, m_N, z_1, \dots, z_N | N, \mathbf{g}) dt_i dm_i dz_i \quad (6.17)$$

Et enfin, nous écrivons la loi *a posteriori* conditionnelle à N de l'intensité de comptage π :

$$p(\pi | N, T) = \mathcal{B}et(\pi | a_\pi + N, b_\pi + T - N) \quad (6.18)$$

6.3.3 Estimation

6.3.3.1 Utilisation d'un algorithme hybride RJMCMC MCMC

Afin d'explorer convenablement l'espace des modèles et l'espace des paramètres au sein de chaque modèle, nous proposons d'utiliser comme précédemment un algorithme hybride RJMCMC-MCMC.

Nous proposons ainsi d'utiliser la même structure que précédemment. La seule différence ici concernera les étapes MCMC, au cours desquels au lieu d'échantillonner f_i , nous échantillonnerons conjointement m_i et z_i . Les figures 6.4 et 6.5 résument ces structures respectivement pendant et après le temps de chauffe. Il sera question d'un « vote » que nous détaillerons ci-dessous.

Nous notons $t_i^{(j)}$, $m_i^{(j)}$, $z_i^{(j)}$ et $\pi^{(j)}$, les valeurs des chaînes MCMC.

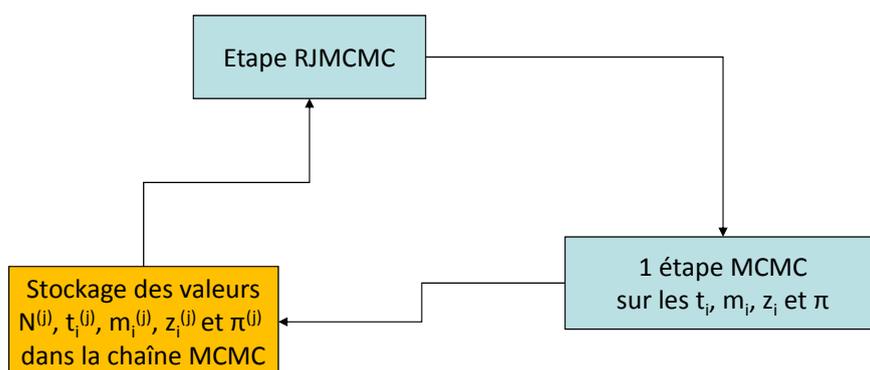


FIGURE 6.4 – Structure de l'algorithme hybride RJMCMC / MCMC multi-mode pendant le temps de chauffe

Nous proposons d'utiliser un échantillonneur RJMCMC pour modifier le nombre de molécules adsorbées N , et trois échantillonneurs de Metropolis-Hastings Q_1 , pour déplacer les temps t_i , Q_2 , pour modifier les masses m_i et les positions z_i , et Q_3 pour échantillonner l'intensité de comptage π .

Par rapport au chapitre précédent, le principe de l'échantillonneur RJMCMC reste inchangé (plus de détails en annexe E.4). Les échantillonneurs Q_1 et Q_3 restent inchangés également. Seul l'échantillonneur Q_2 est modifié, et devient ici :

$$Q_2(m_i^p, z_i^p | m_i^{(j-1)}, z_i^{(j-1)}) = \mathcal{N} \left(\begin{pmatrix} m_i^p \\ z_i^p \end{pmatrix} \middle| \begin{pmatrix} m_i^{(j-1)} \\ z_i^{(j-1)} \end{pmatrix}, \begin{pmatrix} \delta_m & 0 \\ 0 & \delta_z \end{pmatrix} \right) \quad (6.19)$$

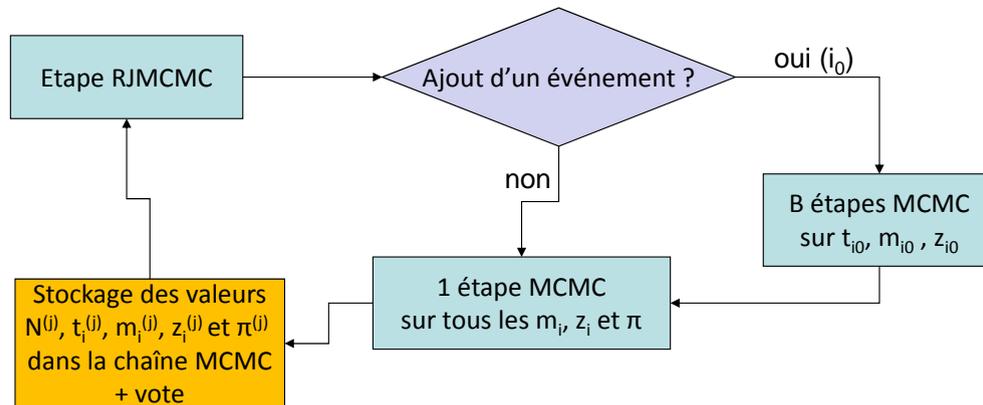


FIGURE 6.5 – Structure de l’algorithme hybride RJMCMC / MCMC multi-mode après le temps de chauffe

6.3.3.2 Le système de vote

Dans ce chapitre, nous visons, comme précédemment, à effectuer une étape de déconvolution impulsionnelle, c’est-à-dire à estimer un signal impulsionnel. Nous proposons d’utiliser le même système de vote que précédemment, avec un système d’estimation simple, en deux étapes :

- une première étape de découverte ;
- une seconde étape d’estimation.

Comme précédemment, la première étape consistera à explorer la loi *a posteriori* et la seconde étape consistera à échantillonner le modèle et les paramètres et de faire voter à chaque itération hors temps de chauffe les temps estimés.

Ainsi, nous créons un vecteur \mathbf{q} mais aussi un vecteur \mathbf{m} (vecteur des masses) qui serviront de vecteurs de « vote ». A chaque itération postérieure au temps de chauffe, et pour chaque entrée de la liste estimée, nous notons u_0 le temps d’échantillonnage le plus proche du temps en question. Nous incrémentons de 1 le vecteur \mathbf{q} sur la composante u_0 , et nous incrémentons de la valeur courante de la masse pour l’entrée de la liste considérée le vecteur \mathbf{m} sur la composante u_0 .

La décision s’effectuera de manière très simple, en prenant comme instants d’échantillonnage les composantes de \mathbf{q} valant au moins 50% du nombre d’échantillons sur lesquels le cumul s’est effectué. Ensuite, pour ces indices i_0 , nous estimerons les masses adsorbées avec $\frac{\mathbf{m}_{i_0}}{\mathbf{q}_{i_0}}$.

6.4 Résultats

6.4.1 Définitions des critères de performance

En termes de détection, nous utiliserons les mêmes critères que précédemment, à savoir la sensibilité et le taux de fausse détection (voir section 5.4.1), que nous représenterons au travers d'une courbe ROC.

Vu que nous cherchons aussi à quantifier la masse des molécules, nous proposons d'utiliser le biais normalisé Bi et le coefficient de variation de la racine de l'erreur quadratique moyenne CV_{RMSD} pour quantifier l'erreur commise sur l'estimation des masses.

$$Bi = \frac{1}{\bar{M}} \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} (\hat{m}_i - m_i) \quad (6.20)$$

$$CV_{\text{RMSD}} = \frac{1}{\bar{M}} \sqrt{\frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} (\hat{m}_i - m_i)^2} \quad (6.21)$$

où \bar{M} est la masse moyenne des molécules, \tilde{N} le nombre de molécules détectées (vrais positifs), \hat{m}_i les masses estimées et m_i les vraies masses.

Le calcul s'effectuera selon la procédure d'appariement décrite dans le chapitre précédent.

6.4.2 Résultats sur données simulées

6.4.2.1 Modalités de simulation

Nous allons tester notre méthode sur données simulées. Pour simuler les données, nous commençons par générer aléatoirement un signal d'observation sur T échantillons.

Pour cela, nous générons N triplets $\{t_i, m_i, z_i\}$, les t_i étant tirés aléatoirement entre 1 et T . Nous veillons notamment à ce qu'il n'y ait pas deux t_i identiques.

Nous générons ensuite le signal résultant de cette liste d'adorption en appliquant (3.12) puis en discrétisant le signal. Nous travaillerons sur $K = 2$ harmoniques. Nous ajoutons sur chaque harmonique k un bruit suivant une loi normale de variance σ^2 , les bruits sur chaque mode étant dé-corrélés. Ensuite, nous intégrons ce bruit (fonction Matlab *cumsum*), ce qui en fait un bruit en 1/f pur.

Les paramètres de simulation sont donc :

- T le nombre d'échantillons ;
- σ l'écart-type du bruit ;
- τ le temps de réponse de la boucle à verrouillage de phase (exprimé en périodes d'échantillonnage) ;
- κ le temps de « dé-synchronisation » entre les deux modes (exprimé en périodes d'échantillonnage) ;

- α la valeur des gains fixes ;
- $\phi(\cdot)$, les gains variables : ceux-ci seront issus de la géométrie *double-clamped* ;
- N le nombre de molécules ;
- les valeurs des chutes de fréquence que nous préciserons plus loin.

Le tableau 6.1 donne les valeurs des paramètres de simulation.

T	σ	τ	κ	α	N
200 échantillons	1	5 échantillons	0	{1, 1} Hz.kDa ⁻¹	10 adsorptions

TABLE 6.1 – Valeurs des paramètres de simulation multi-mode

6.4.2.2 Paramètres de l’algorithme

La méthode que nous avons proposée prend un certain nombre de paramètres, il s’agit notamment :

- du temps de chauffe, noté J_b ;
- du nombre total d’itérations J ;
- du nombre d’itérations B lorsqu’un événement est ajouté ;
- des paramètres des lois de proposition δ_m , δ_z et δ_t ;
- des paramètres des lois *a priori* k_m et θ_m , a_π et b_π .

Les tableaux 6.2 et 6.3 donnent respectivement les paramètres des algorithmes et les paramètres des lois *a priori*.

J_b	J	B	δ_m	δ_z	δ_t
3000	4000	500	20	0.1	1

TABLE 6.2 – Valeurs des paramètres de l’algorithme pour le traitement des données simulées multi-mode

k_m	θ_m	a_π	b_π
1	100	1	T

TABLE 6.3 – Valeurs des paramètres des lois *a priori* pour le traitement des données simulées multi-mode

6.4.2.3 Algorithmes de référence

Nous allons comparer notre méthode à plusieurs algorithmes de référence, qui détecteront les temps d’adsorption et quantifieront éventuellement la masse.

Adaptation simple des méthode de déconvolution Les méthodes évoquées précédemment, à savoir la méthode par seuillage, la pseudo-inverse, le LASSO (algorithme LAR), *Matching-Pursuit* (algorithme CLEAN) et la méthode reposant sur un modèle Bernoulli-Gaussien (algorithme SMLR) opéreront en mono-mode sur le signal formé de la moyenne de tous les modes.

Les paramètres de ces méthodes seront les mêmes que dans le chapitre précédent, à savoir : pour le seuillage et la pseudo-inverse, un seuil variant entre 0 et 200 par pas de 1, pour le LASSO, un paramètre de régularisation variant entre 0 et 1 par pas de 0.1 puis entre 1 et 100 par pas de 1, pour *Matching-Pursuit*, un nombre de variables actives qui variera entre 0 et T , par pas de 1 et pour l'algorithme SMLR, un paramètre π fixé à 0.05 et les variances *a priori* valant $5^2, 10^2, 20^2, 50^2, 100^2, 200^2, 500^2, 1000^2, 2000^2, 5000^2, 10000^2$.

Méthode de déconvolution spécifique au cas multi-mode Nous testerons aussi la méthode SMLR multi-mode décrite dans la section 6.2.3. Pour cette dernière, nous prendrons un paramètre π fixé à 0.05 et toutes les variances sur les lois *a priori* des amplitudes identiques : $r_k = r_0$. Ces variances vaudront : $5^2, 10^2, 15^2, 20^2, 30^2, 40^2, 50^2, 100^2, 200^2, 500^2, 1000^2, 2000^2, 5000^2, 10000^2$. L'algorithme SMLR multi-mode estimera les chutes de fréquence sur chacun des modes, dont nous déduirons la masse par la méthode de Dohn *et al.* (voir en annexe D.2).

6.4.2.4 Résultats avec instants d'adsorption entiers

Nous allons désormais caractériser les performances dans le cas où les instants d'adsorption sont entiers. Nous adoptons la configuration géométrique *double-clamped*. Ici, nous tirons uniformément les m_i entre 0 et 100 kDa et les z_i entre 0 et 0.5. Nous moyennons les performances sur 50 répétitions. Le seuil pour l'appariement des instants vaut $\gamma_d = 1$ échantillon.

Nous entendons par « MCMC1 » la méthode consistant à effectuer la détection en utilisant l'estimation de l'ordre du modèle au sens du MMAP, et « MCMC2 » la méthode consistant à effectuer la détection par seuillage du vecteur \mathbf{q} .

La courbe ROC résultante est donnée par la figure 6.6 (la partie droite est un zoom sur les performances de notre méthode).

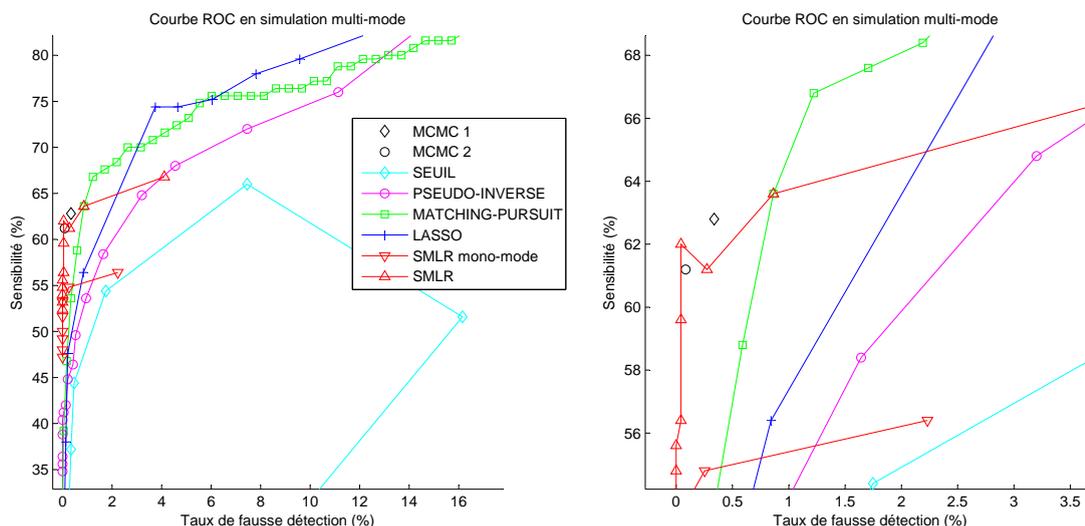


FIGURE 6.6 – Courbe ROC multi-mode pour une simulation avec $N = 10$ événements et des instants d'adsorption entiers

Sur ces données, le biais Bi et le coefficient de variation CV_{RMSD} en masse sur les molécules quantifiées valent :

- pour la méthode MCMC1 : $Bi = 0.02$ et $CV_{\text{RMSD}} = 0.42$;
- pour la méthode MCMC2 : $Bi = 0.06$ et $CV_{\text{RMSD}} = 0.37$;
- Avec l’algorithme SMLR couplé à celui de Dohn *et al.* : le biais Bi varie de $4.1 \cdot 10^8$ à $8.4 \cdot 10^8$ pour une moyenne de $5.2 \cdot 10^8$ et le coefficient de variation CV_{RMSD} varie de $1.0 \cdot 10^9$ à $3.4 \cdot 10^9$ pour une moyenne de $1.6 \cdot 10^9$.

Le tableau 6.4 donne les temps de calcul moyens des différents algorithmes (*Seuil*, *PI* pour « Pseudo-inverse », *LASSO*, *MP* pour « Matching Pursuit », *SMLR mono-mode*, *SMLR multi-mode*, *MCMC1* et *MCMC2*). Dans le cas de la pseudo-inverse et de l’algorithme LASSO, il est possible de mettre en commun le calcul pour différents paramètres de régularisation. Nous n’en avons pas tenu compte dans ce tableau.

Seuil	PI	LASSO	MP
0.1 ms	24.4 ms	114.2 ms	155.7 ms
SMLR mono-mode	SMLR multi-mode	MCMC1	MCMC2
1.2 s	31.7 s	219.5 s	219.5 s

TABLE 6.4 – Temps de calcul des différents algorithmes en multi-mode avec instants d’adsorption entiers

6.4.2.5 Résultats avec instants d’adsorption non-entiers

Nous allons désormais caractériser les performances dans le cas où les instants d’adsorption ne sont pas entiers. Nous adoptons la configuration géométrique *double-clamped*. Ici, nous tirons uniformément les m_i entre 0 et 100 kDa et les z_i entre 0 et 0.5. Nous moyennons les performances sur 50 répétitions. Le seuil pour l’appariement des instants vaut $\gamma_d = 2$ échantillons.

La courbe ROC résultante est donnée par la figure 6.7 (la partie droite est un zoom sur les performances de notre méthode).

Sur ces données, le biais Bi et le coefficient de variation CV_{RMSD} en masse sur les molécules quantifiées valent :

- pour la méthode MCMC1 : $Bi = 0.11$ et $CV_{\text{RMSD}} = 0.70$;
- pour la méthode MCMC2 : $Bi = 0.13$ et $CV_{\text{RMSD}} = 0.61$;
- Avec l’algorithme SMLR couplé à celui de Dohn *et al.* : le biais Bi varie de $-2.7 \cdot 10^9$ à $1.5 \cdot 10^8$ pour une moyenne de $-1.0 \cdot 10^9$ et le coefficient de variation CV_{RMSD} varie de $1.3 \cdot 10^9$ à $7.3 \cdot 10^9$ pour une moyenne de $3.8 \cdot 10^9$.

Le tableau 6.5 donne les temps de calcul moyens des différents algorithmes (*Seuil*, *PI* pour « Pseudo-inverse », *LASSO*, *MP* pour « Matching Pursuit », *SMLR mono-mode*, *SMLR multi-mode*, *MCMC1* et *MCMC2*). Dans le cas de la pseudo-inverse et de l’algorithme LASSO, il est possible de mettre en commun le calcul pour différents paramètres de régularisation. Nous n’en avons pas tenu compte dans ce tableau.

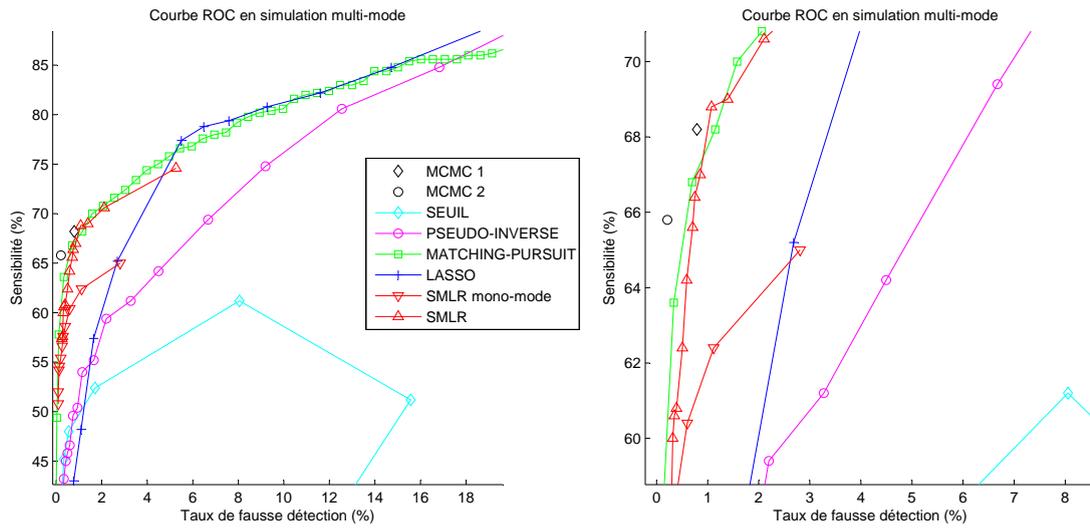


FIGURE 6.7 – Courbe ROC multi-mode pour une simulation avec $N = 10$ événements et des instants d’adsorption non entiers

Seuil	PI	LASSO	MP
0.1 ms	8.8 ms	103.6 ms	101.7 ms
SMLR mono-mode	SMLR multi-mode	MCMC1	MCMC2
1.0 s	36.8 s	264.7 s	264.7 s

TABLE 6.5 – Temps de calcul des différents algorithmes en multi-mode avec instants d’adsorption non-entiers

6.4.3 Validation de la méthode sur données réelles

Validons désormais notre méthode sur données réelles. Nous allons ici traiter un exemple, issu de sessions expérimentales de Décembre 2012 sur des agrégats de Tantale d'un diamètre nominal de 6.5 nm, soit une masse nominale d'environ 1420 kDa.

Le tableau 6.6 donne les valeurs des paramètres du modèle direct, à savoir le nombre d'échantillons T , la période d'échantillonnage T_e , le temps de réponse τ , le déphasage κ et les gains α .

T	T_e	τ	κ	α
5000 échantillons	2 ms	10 ms	0	$\{5.00 \cdot 10^{-2}, 11.73 \cdot 10^{-2}\}$ Hz.kDa ⁻¹

TABLE 6.6 – Valeurs des paramètres du modèle direct pour les données réelles multi-mode

Ici, κ vaut 0 car les signaux ont été recalés dans le temps grâce à la méthode décrite en annexe B.1. Nous prendrons un filtre blanchisseur d'ordre $V = 5$. Les coefficients du filtre sont donnés dans le tableau 6.7 :

Mode 1	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	1.7989	-1.1393	0.0358	0.0374	-0.1959
Mode 2	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	0.3187	-0.1197	-0.0409	-0.0378	-0.0346

TABLE 6.7 – Coefficients des filtres blanchisseurs pour le traitement des données réelles multi-mode

La matrice de covariance du bruit d'un mode à l'autre, quant à elle, vaut :

$$\mathbf{\Gamma}_m = \begin{pmatrix} 1.0271 & 0.1753 \\ 0.1753 & 1.0196 \end{pmatrix} \quad (6.22)$$

Les tableaux 6.8 et 6.9 donnent respectivement les paramètres des algorithmes et les paramètres des lois *a priori*.

J_b	J	B	δ_m	δ_z	δ_t
10000	15000	100	100	0.1	1

TABLE 6.8 – Valeurs des paramètres de l'algorithme pour le traitement des données réelles multi-mode

La figure 6.8 illustre la détection des instants de chute par un cercle rouge, sur les deux modes (le mode 1 est en haut et le mode 2 est en bas). La méthode de décision retenue ici est la méthode de seuillage du vecteur \mathbf{q} .

Le tableau 6.10, quant à lui, donne l'estimation des diamètres des molécules détectées (pour ce faire, on utilise la masse volumique du Tantale, 16.4 g.cm⁻³).

k_m	θ_m	a_π	b_π
1	4000	1	T

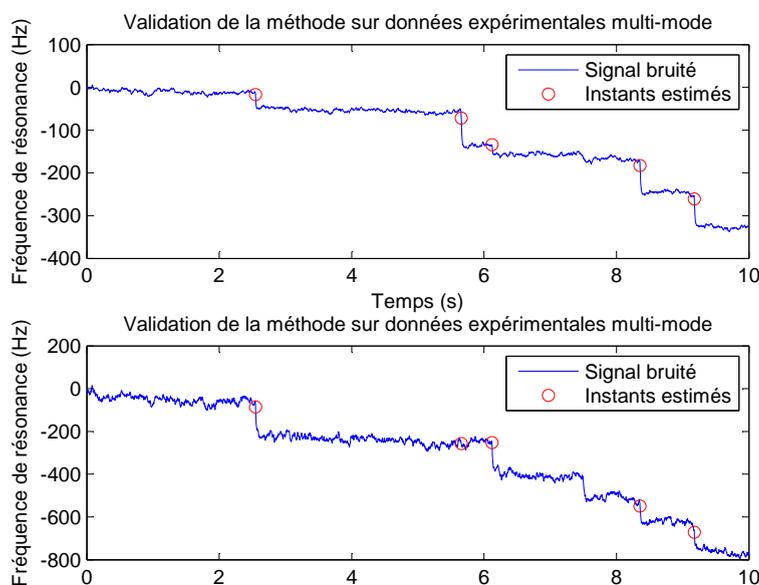
 TABLE 6.9 – Valeurs des paramètres des lois *a priori* pour le traitement des données réelles multi-mode


FIGURE 6.8 – Illustration de la détection multi-mode sur un signal expérimental

Temps (s)	2.550	5.658	6.122	8.358	9.178
Diamètre (nm)	5.94	6.38	6.35	6.72	6.72

TABLE 6.10 – Diamètre estimé des agrégats détectés sur données expérimentales

6.4.4 Discussion

Nous venons de tester notre méthode sur données multi-mode, de la confronter à des algorithmes alternatifs et de la tester sur données réelles.

Rappelons que nous voulions :

- structurer les chutes de fréquence par notre modèle physique ;
- permettre des instants d'adsorption non-entiers ;
- estimer automatiquement certains hyperparamètres.

En simulation, et en terme de détection (dans les zones où le taux de fausses détection est inférieur à 0.5%), on note que les algorithmes spécifiques à la déconvolution multi-mode fonctionnent mieux que ceux fonctionnant en moyennant tous les modes pour se ramener à un cas mono-mode. Aussi, notre algorithme offre des performances de détection comparables au SMLR multi-mode lorsque les instants d'échantillonnage sont entiers. Toutefois, nous pensons dépasser le

SMLR multi-mode grâce à la bonne structuration *a priori* de l'espace des chutes de fréquence. Il semblerait que cela n'ait pas d'impact dans le cas étudié ici.

Les avantages cités dans le chapitre précédent, à savoir la possibilité d'avoir des instants d'échantillonnage non entiers et le calcul automatique de l'intensité de comptage π , restent valables dans ce cadre-ci. La remarque sur le temps de calcul élevé de notre méthode l'est également.

En terme de quantification, l'algorithme « en deux étapes » { SMLR + méthode de Dohn *et al.* } présente de très mauvais résultats. La quantification n'étant pas régularisée, il suffit que la chute estimée sur le mode 1 soit très faible pour que l'estimation de masse diverge. On peut notamment constater des biais très forts (positivement ou négativement). Notre méthode, quant à elle, présente des valeurs de biais et de coefficients de variation « raisonnables ». Le biais est inférieur à 6% dans le cas entier et inférieur à 15% dans le cas non entier. On notera que la valeur des coefficients de variation est tout de même assez élevée (proche de 40% dans le cas entier et de 70% dans le cas non-entier). Les performances se dégradent avec le caractère non-entier des instants d'adsorption.

Sur données réelles, l'algorithme détecte toutes les chutes de fréquence visibles à l'œil, même lorsque cette chute ne se voit que sur une harmonique (voir la chute à 5.658 notamment). Seule une chute semble subsister aux alentours de 7.5 s. Est-ce une chute non détectée, un artefact de bruit bien pris en compte par notre algorithme ?

Nous pouvons valider notre modèle. En terme de quantification, les diamètres estimés par notre méthode sont cohérents avec le spectre de masse des agrégats.

Encore une fois, nous avons été obligés d'adapter notre algorithme d'exploration de modèle en lui adjoignant une phase de d'estimation. Nous pensons que si nous cherchons non plus à raisonner en termes de déconvolution mais en terme de reconstruction d'un paramètre collectif au mélange analysé (en l'occurrence, un spectre de masse), nous pouvons nous passer de cette phase d'estimation. Notamment, nous avons fait le choix du non-échantillonnage des temps d'adsorption durant cette phase, qui fait que toutes les variabilités ne sont pas intégrées.

Aussi, nous avons retenu comme critère de performance de quantification une grandeur (le coefficient de variation) reposant sur une décision (au sens de l'EAP). Or, notre algorithme peut faire beaucoup mieux : il peut évaluer l'incertitude sur la masse de chacune des molécules. Que certaines estimations soient très mauvaises peut ne pas être impactant si la méthode a la possibilité de prendre en compte cette information. Dans un contexte bayésien, cette information est portée par la loi *a posteriori*.

Enfin, même si notre structure hiérarchique ne permet pas d'améliorer les performances de détection ici par rapport à des algorithmes de l'état de l'art, c'est bien ce modèle qui va permettre la mise en place d'un algorithme de reconstruction « tout intégré ».

Le travail restant à accomplir est donc le prolongement de ce travail pour estimer non-plus des masses de molécules, mais le spectre de la solution analysée, en fusionnant les informations portées par les différentes masses.

Traitement multi-mode – Reconstruction de spectre de masse

Dans ce chapitre, nous allons utiliser le travail en déconvolution impulsioneille effectué précédemment dans le but de reconstruire un spectre de masse. Nous introduirons le cadre de travail et les motivations de nos choix, exposerons l'état de l'art, essaierons de ramener notre problème à d'autres problèmes classiques. Ensuite, nous adapterons une méthode de reconstruction de l'état de l'art à notre algorithme bayésien que nous testerons sur données réelles. Enfin, nous développerons notre propre contribution dans le cadre d'un spectre de masse parcimonieux et nous testerons notre méthode dans un cadre de simulation. Les résultats présentés ici seront des résultats préliminaires, donc à étudier avec prudence et recul.

Sommaire de ce chapitre

7.1	Cadre de travail	165
7.2	Méthodes de référence	167
7.2.1	État de l'art – la méthode publiée par Caltech	167
7.2.2	Méthodes connexes	167
7.2.2.1	Histogramme	167
7.2.2.2	Déconvolution	168
7.2.2.3	Clustering	168
7.2.2.4	Estimation d'une densité de probabilité	169
7.3	Adaptation de la méthode de Caltech à notre algorithme	170
7.3.1	Principe	170
7.3.1.1	Structure de l'algorithme	170
7.3.1.2	Reconstruction du spectre par moyenne d'histogramme	171
7.3.1.3	Reconstruction du spectre par histogramme des masses estimées	171
7.3.1.4	Reconstruction du spectre exploitant l'hypothèse de Laplace	171
7.3.2	Exemple de résultat sur données expérimentales	172
7.3.2.1	Validation de la méthode sur données réelles	172
7.3.3	Commentaires	173
7.3.3.1	Comparaison des trois méthodes	173
7.3.3.2	Critique générale	175

7.4	Développement d'une méthode de reconstruction d'un spectre de raies	176
7.4.1	Cadre de travail utilisé	176
7.4.2	Modélisation statistique des grandeurs et estimation	176
7.4.2.1	Expression de la fonction de vraisemblance	176
7.4.2.2	Définition des lois <i>a priori</i>	177
	Loi de probabilité sur l'ordre du modèle	177
	Loi de probabilité sur les temps d'adsorption	177
	Loi de probabilité sur la masse	177
	Loi de probabilité sur la position	177
	Loi de probabilité sur l'intensité de comptage	178
	Loi de probabilité sur les raies du spectre de masse	178
7.4.2.3	Écriture des lois <i>a posteriori</i> des paramètres de la liste pour un ordre de modèle et un spectre donné	178
7.4.2.4	Écriture des autres lois <i>a posteriori</i> conditionnelles	178
7.4.2.5	Utilisation d'un algorithme hybride RJMCMC-MCMC	179
	Échantillonneur de la masse et de la position	179
	Échantillonneur conjoint du spectre et des masses	180
	Estimation	180
7.4.3	Quelques résultats préliminaires sur signaux synthétiques	180
7.4.3.1	Modalités de simulation	180
7.4.3.2	Paramètres de l'algorithme	181
7.4.3.3	Résultats	182
7.4.3.4	Discussion	182

7.1 Cadre de travail

Le cadre de travail dans lequel nous allons nous inscrire sera le cadre de la reconstruction de spectre de masse d'un mélange de molécules analysé. Ceci correspond aux parties « Détecter » et « Quantifier » et « Compter » identifiées dans la section 3.5.

Dans ce cadre-ci, il s'agit non seulement d'inverser l'opération de convolution que l'on retrouve dans l'équation (3.11), de prendre en compte que l'information se répète sur plusieurs signaux, et d'inverser les équations non-linéaires $f_{i,k} = \alpha_k m_i \phi_k(z_i)$, mais aussi de reconstruire le spectre de masse associé, c'est-à-dire de retrouver les paramètres collectifs du mélange analysé.

Nous travaillerons avec les mêmes hypothèses qu'au chapitre précédent :

- $K = 2$ harmoniques sont observées. Nous exploiterons la simultanéité des chutes de fréquences sur les différents modes.
- Les réponses du système à une adsorption $u_{\text{PLL},k}(t)$ sont connues, comme précédemment. Ces réponses dépendent à la fois du paramètre de temps de réponse τ mais aussi du décalage entre les modes κ .
- La loi de probabilité du bruit est connue, comme précédemment. Ici, les paramètres de bruit sont les coefficients des filtres blanchisseur pour le mode 1 \mathbf{a}_1 et pour le mode 2 \mathbf{a}_2 ainsi que la matrice de covariance inter-mode $\mathbf{\Gamma}_m$.
- La géométrie des capteurs sera une géométrie de type *double-clamped*.

Toutes les notations définies auparavant restent valides. La seule différence vient du fait que les masses m_i ne dépendent plus des paramètres *a priori* θ_m , mais d'un spectre de masse inconnu Sp , qui dépendra lui des paramètres *a priori* θ_{Sp} .

Ce cadre de travail peut se représenter au travers de la modélisation hiérarchique illustrée par la figure 7.1.

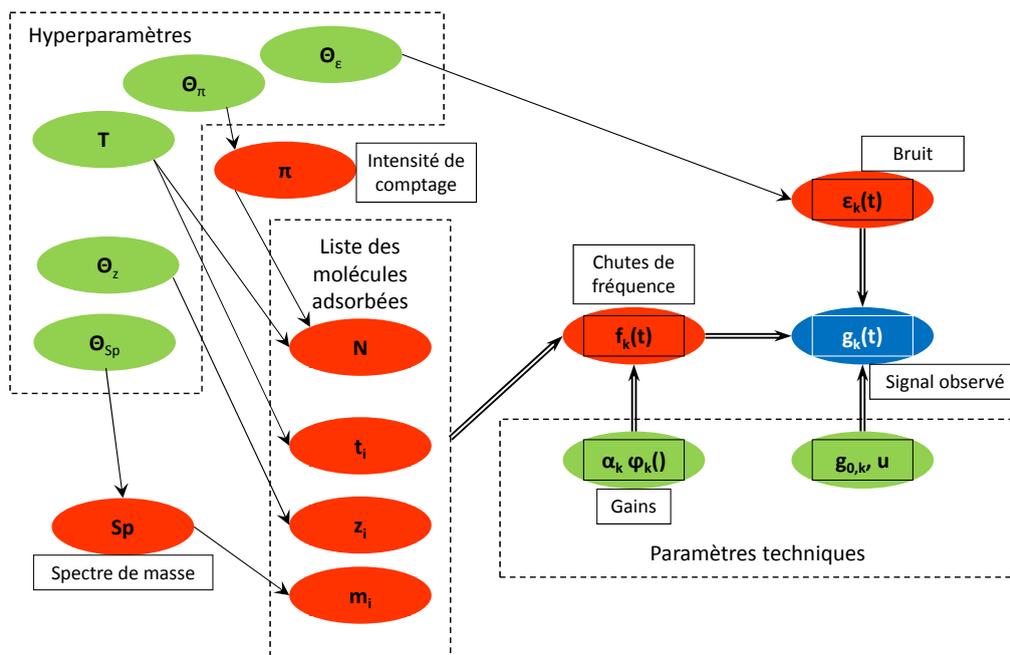


FIGURE 7.1 – Modélisation hiérarchique du problème de détection – quantification – comptage

7.2 Méthodes de référence

7.2.1 État de l'art – la méthode publiée par Caltech

La seule méthode de reconstruction de spectre de masse à partir d'une mesure issue d'un capteur NEMS que nous ayons trouvée est encore une fois décrite par Caltech [HKN⁺12].

Rappelons la procédure décrite par cet article. Dans un premier temps, les instants de chute sont détectés, à la suite de quoi les chutes de fréquences sont quantifiées (voir section 5.2.1). Dans un second temps, les lois de probabilité de la masse de chacune des molécules sont estimées par changement de variable (voir section 6.2.1).

Afin d'estimer le spectre de masse, Hanay *et al.* proposent de sommer chacune des lois de probabilité pour chaque molécule détectée.

7.2.2 Méthodes connexes

Si la méthode décrite par Caltech est la seule sur le sujet de la spectrométrie de masse à base de NEMS, nous pouvons tout de même établir une connexion avec d'autres problématiques et des solutions qui en découlent.

Nous prenons l'hypothèse qu'un système de déconvolution multi-mode et d'estimation de la masse nous fournit un ensemble de masses estimées et que nous cherchons à retrouver le spectre de masse à partir de ces estimations. Nous allons notamment aborder :

1. les méthodes par histogramme ;
2. les méthodes de déconvolution ;
3. les méthodes de *clustering* ;
4. les méthodes statistiques d'estimation d'une densité de probabilité.

7.2.2.1 Histogramme

Les méthodes de reconstruction de spectre par histogramme reposent sur un principe très simple. L'espace des masses est divisé en une somme d'ensembles disjoints, les *bins*. A chacun de ces *bins* est associé le nombre de masses observées appartenant à l'intervalle du *bin* considéré.

Un point important est le choix de la taille des *bins*. Plusieurs choix peuvent être faits.

- Il est possible de choisir la taille des *bins* en rapport avec la résolution de l'instrument. Si l'on choisit comme taille du *bin* une valeur plus grande que la résolution de l'instrument, les valeurs placées dans un *bin* ont une forte probabilité d'appartenir effectivement à ce *bin*. Si la résolution de l'instrument n'est pas constante, la taille des *bins* peut être variable.
- Il est également possible de choisir la taille des *bins* en fonction de la dynamique des données observées et du nombre d'observations. En effet, pour que les valeurs associées à chaque *bin* soient significatives, il faut que le *bin* concentre suffisamment de données, donc qu'il soit

suffisamment grand. A l'inverse, une taille de *bin* petite permettra une représentation plus fine du spectre.

7.2.2.2 Déconvolution

Pour estimer un spectre à partir des signaux observés à la sortie d'un instrument, une démarche couramment employée est celle de la déconvolution. Ici, un spectre « expérimental » est constitué (par histogramme, par exemple), et l'influence de l'instrument sur la mesure de ce spectre est représentée par une opération de convolution (voir par exemple [Kri77]).

Ce type d'approche se retrouve dans de nombreuses applications comme les applications d'imagerie par rayons X [Yua09], l'imagerie [WXW12] et même en spectrométrie de masse [MM93] (cette liste est très loin d'être exhaustive).

Vincent Mazet, que nous avons fréquemment cité tout au long de cette thèse, utilise d'ailleurs des méthodes MCMC et RJMCMC pour déconvoluer des spectres infrarouges Raman [Maz05], ou encore des spectres de photo-électrons [MFM⁺12].

Ici, toutes les méthodes de déconvolution que nous avons vues dans notre document peuvent s'appliquer.

7.2.2.3 Clustering

Nous pouvons aussi établir une connexion entre la problématique de reconstruction de spectre avec les problématiques d'apprentissage non-supervisé, et en particulier les problématiques de *clustering* (que l'on peut traduire par « partitionnement »), si l'on se place dans le cadre de l'analyse d'un mélange de molécules pouvant être regroupées en « classes ».

En effet, si l'on considère que la distribution de masse n'est pas continue mais que les molécules sont regroupées en ensembles au sein desquels les masses des molécules sont identiques ou proches, alors il est possible de chercher à regrouper les molécules entre elles. On parle alors de *clustering*.

Parmi les méthodes de *clustering*, on peut notamment citer l'algorithme des *k-means* [Dun73], qui cherche à regrouper les données en *k* classes représentées par leurs valeurs moyennes.

Il existe également des algorithmes inspirés des neuro-sciences. Les « cartes auto-organisatrices de Kohonen », par exemple, visent à regrouper des éléments entre eux afin de mieux représenter l'information, à l'instar du comportement des neurones dans le cortex. Citons l'algorithme LVQ [PBT93] (pour *Learning Vector Quantization*, que l'on peut traduire par « apprentissage de quantification vectorielle »), comme méthode de résolution de cette problématique.

Cette problématique est très étudiée en apprentissage automatique et en dresser un état de l'art exhaustif dépasserait largement le cadre de ce document.

7.2.2.4 Estimation d'une densité de probabilité

Une dernière catégorie de méthode consisterait à considérer le spectre de masse comme une loi de probabilité et les masses observées comme des tirages sous cette loi. Le but serait alors d'estimer la densité de probabilité sachant les masses mesurées.

Cette estimation peut se faire au travers d'une modélisation paramétrique de la densité de probabilité. Un exemple simple consisterait à considérer cette dernière comme une loi normale et donc à estimer son écart-type et sa variance.

Un exemple plus générique consisterait à modéliser le problème comme un mélange de lois normales [SG99] (GMM pour *Gaussian Mixture Model*). L'algorithme d'Espérance-Maximisation [McM96] (ou *Expectation-Maximization*), est bien adapté pour en estimer les paramètres.

Notons qu'il est également possible d'estimer la densité de probabilité sans pour autant en fournir un modèle paramétrique [TPJ05, ES97].

7.3 Adaptation de la méthode de Caltech à notre algorithme

7.3.1 Principe

7.3.1.1 Structure de l'algorithme

La méthode de reconstruction de spectre de Caltech repose sur la sommation de toutes les contributions de toutes les molécules, c'est-à-dire la somme des $p(m_i | \mathbf{g}_k)$.

Pour adapter cette méthode, nous gardons exactement le même modèle statistique qu'au chapitre précédent, en cherchant à développer une méthode « tout intégrée », c'est-à-dire sans prise de décision intermédiaire.

L'algorithme RJMCMC / MCMC fonctionnera lui aussi en deux phases. La première de ces phases sera une phase de chauffe, qui sera exactement la même que précédemment (voir figure 6.5). La deuxième phase sera la phase de construction du spectre (et non plus celle du vote pour effectuer une déconvolution). Cette phase est décrite par la figure 7.2. Il est fait mention d'une étape de reconstruction de spectre que nous développerons ci-dessous.

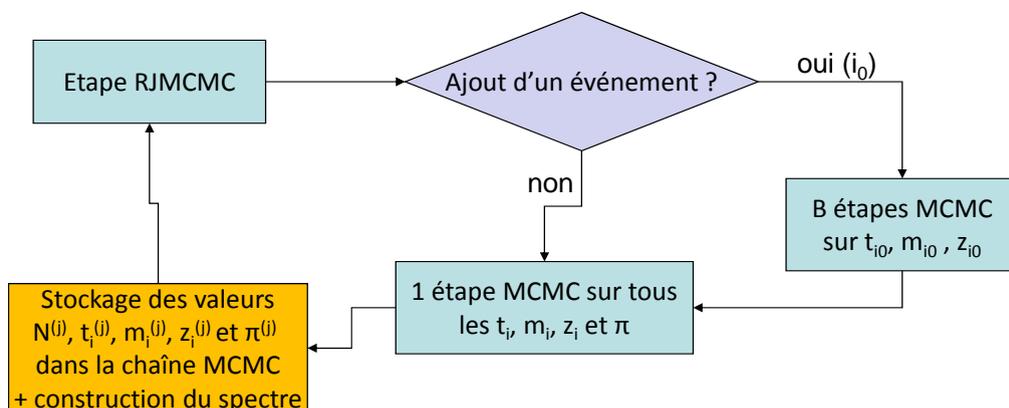


FIGURE 7.2 – Structure de l'algorithme hybride RJMCMC / MCMC en reconstruction de spectre après le temps de chauffe

Sur la base de cet algorithme, nous proposons trois algorithmes de reconstruction de spectre :

1. une méthode par moyenne d'histogrammes ;
2. une méthode par histogramme des estimations ;
3. une méthode exploitant l'hypothèse de Laplace.

7.3.1.2 Reconstruction du spectre par moyenne d'histogramme

L'étape de reconstruction du spectre consiste à exploiter les valeurs courantes des chaînes MCMC afin de construire le spectre à l'itération considérée. N'ayant à ce stade aucune information sur la forme du spectre recherchée et cherchant à adapter la méthode de Caltech, nous proposons de construire le spectre sur la base d'un histogramme. La taille des *bins* sera constante et notée ΔM .

Pour construire l'histogramme, avant chaque étape RJMCMC, nous balayons l'ensemble des masses estimées m_i et nous incrémentons de 1 la valeur des *bins* de l'histogramme dans lesquels se situent chacune des masses m_i .

D'une certaine manière, cela revient à cumuler ou à moyenniser, selon la normalisation choisie en sortie d'algorithme, tous les histogrammes des listes après le temps de chauffe.

Notons que Caltech propose de « couper » (*cut-off*) les contributions pour les positions z_i inférieures à 0.25, car la résolution en masse du capteur est mauvaise dans cette zone. Nous avons choisi de ne pas implémenter cette amélioration.

7.3.1.3 Reconstruction du spectre par histogramme des masses estimées

Nous proposons une méthode de reconstruction de spectre basée sur la méthode d'inversion multi-mode vue au chapitre précédent. Ces deux méthodes consistent à estimer les masses en sortie de l'algorithme, en seuillant le vecteur de vote \mathbf{q} et en en déduisant les masses (voir section 6.3.3.2). A la différence de la méthode du chapitre précédent, nous proposons d'échantillonner les temps d'adsorption t_i , même au-delà du temps de chauffe.

La liste des molécules adsorbées est estimée comme dans la section 6.3.3.2. Nous formons ensuite l'histogramme des masses estimées. La taille des *bins* sera constante et notée ΔM .

Notons que cette méthode est une extension naturelle du chapitre précédent et n'est pas, à proprement parlé, une adaptation de la méthode de Caltech.

7.3.1.4 Reconstruction du spectre exploitant l'hypothèse de Laplace

Nous proposons enfin une troisième méthode de reconstruction de spectre basée sur l'hypothèse de Laplace (voir section 4.2.6 page 94). Cette méthode se base sur la même algorithmique que celui décrit précédemment, à savoir l'estimation d'une liste de molécules adsorbées. La différence ici est que nous proposons d'estimer aussi l'incertitude sur les masses et d'exploiter ce résultat dans la reconstruction du spectre.

Pour estimer l'incertitude, nous proposons de fonctionner comme dans la section 6.3.3.2, à savoir de créer un vecteur \mathbf{q} et un vecteur \mathbf{m} , pour cumuler les temps et les masses au cours des itérations, et nous formons aussi le vecteur \mathbf{m}_2 pour estimer le moment d'ordre 2 des masses estimées.

A chaque itération postérieure au temps de chauffe, et pour chaque entrée de la liste estimée, nous notons u_0 le temps d'échantillonnage le plus proche du temps en question. Nous incrémentons de 1 le vecteur \mathbf{q} sur la composante u_0 , nous incrémentons de la valeur courante de la masse

pour l'entrée de la liste considérée le vecteur \mathbf{m} sur la composante u_0 , et nous incrémentons du carré de la valeur courante de la masse pour l'entrée de la liste considérée le vecteur \mathbf{m}_2 sur la composante u_0 .

Comme précédemment, nous prenons comme instants d'échantillonnage les composantes de \mathbf{q} valant au moins 50% du nombre d'échantillons sur lesquels le cumul s'est effectué. Ensuite, pour ces indices i_0 , nous estimerons les masses adsorbées avec $\frac{\mathbf{m}_{i_0}}{\mathbf{q}_{i_0}}$. Si le vecteur \mathbf{m} permet d'accéder au moment d'ordre 1 des masses adsorbées, c'est-à-dire à leur espérance, le vecteur \mathbf{m}_2 permet d'accéder au moment d'ordre 2 de la valeur des masses adsorbées. Ainsi, pour ces indices i_0 , nous estimerons la variance sur les estimations des masses adsorbées par $\frac{(\mathbf{m}_{2i_0} - \mathbf{m}_{i_0}^2)}{\mathbf{q}_{i_0}}$.

Le spectre sera reconstruit en exploitant l'hypothèse de Laplace. Nous associerons à chaque molécule adsorbée estimée une distribution normale centrée sur la masse estimée et de variance la variance estimée. Les distributions seront calculées sur une grille de pas ΔM et sommées.

7.3.2 Exemple de résultat sur données expérimentales

7.3.2.1 Validation de la méthode sur données réelles

Validons désormais notre méthode sur données réelles. Nous allons ici traiter un exemple, issu de sessions expérimentales de Décembre 2012 sur des agrégats de Tantale d'un diamètre nominal de 6.5 nm, soit une masse nominale d'environ 1420 kDa.

Nous allons travailler en découpant le signal sur des fenêtres de T échantillons.

Le tableau 7.1 donne les valeurs des paramètres du modèle direct, à savoir le nombre d'échantillons des fenêtres T , la période d'échantillonnage T_e , le temps de réponse τ et le déphasage κ . Les gains α sont recalculés pour chaque fenêtre, la valeur dans le tableau est une valeur typique (celle de la première fenêtre).

T	T_e	τ	κ	α
5000 échantillons	2 ms	10 ms	0	$\{5.10 \cdot 10^{-2}, 11.97 \cdot 10^{-2}\}$ Hz.kDa ⁻¹

TABLE 7.1 – Valeurs des paramètres du modèle direct pour les données réelles en reconstruction de spectre

Ici, κ vaut 0 car les signaux ont été recalés dans le temps grâce à la méthode décrite en annexe B.1 pour chaque fenêtre. Nous prendrons un filtre blanchisseur d'ordre $V = 5$. Les coefficients du filtre sont donnés dans le tableau 7.2.

La matrice de covariance du bruit d'un mode à l'autre, quant à elle, vaut :

$$\mathbf{\Gamma}_m = \begin{pmatrix} 1.0271 & 0.1753 \\ 0.1753 & 1.0196 \end{pmatrix} \quad (7.1)$$

Mode 1	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	1.7989	-1.1393	0.0358	0.0374	-0.1959
Mode 2	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$
	0.3187	-0.1197	-0.0409	-0.0378	-0.0346

TABLE 7.2 – Coefficients des filtres blanchisseurs pour le traitement des données réelles en reconstruction de spectre

Les tableaux 7.3 et 7.4 donnent respectivement les paramètres des algorithmes et les paramètres des lois *a priori*.

J_b	J	B	δ_m	δ_z	δ_t	ΔM
3000	5000	50	100	0.01	1	10

TABLE 7.3 – Valeurs des paramètres de l’algorithme pour le traitement des données réelles en reconstruction de spectre

k_m	θ_m	a_π	b_π
1	1000	1	T

 TABLE 7.4 – Valeurs des paramètres des lois *a priori* pour le traitement des données réelles en reconstruction de spectre

La figure 7.3 donne les spectres estimés pour les trois méthodes décrites dans les sections 7.3.1.2 (« moyennes d’histogrammes »), 7.3.1.3 (« histogramme ») et 7.3.1.4 (« Laplace »). Nous exprimons les masses en diamètres estimés.

7.3.3 Commentaires

7.3.3.1 Comparaison des trois méthodes

Globalement, nous voyons que les trois méthodes semblent bien fonctionner. En effet, nous voyons assez clairement un pic se détacher aux alentours de 6.5 nm sur les méthodes « moyennes d’histogrammes » et « Laplace », ce pic étant moins marqué sur la méthode « histogramme ». Ce pic est ici centré sur 6.7 nm. Il n’est pas possible à ce stade de savoir si cette erreur de quantification vient du modèle (valeur des gains α_k , par exemple), ou des algorithmes. Aussi, les trois méthodes donnent des résultats inégaux.

Sur la méthode « moyennes d’histogrammes », bien que nous voyons un pic se dégager, nous pouvons constater sur le spectre un « bruit » correspondant à des petites masses non censées appartenir au spectre (inférieures à 5 nm). Nous les attribuons au caractère exploratoire de notre algorithme, qui accepte l’ajout d’adsorptions qui n’en sont pas (adsorptions qui seront ensuite supprimées au cours des itérations), et donc moyenne des modèles peu probables. Nous pensons que dans cette démarche, reconstruire le spectre en faisant la moyenne de tous les modèles comme nous le faisons ici n’est pas la meilleure solution. Il y a donc un compromis à trouver entre la non-prise de décision (l’algorithme explore des modèles) et la mise à l’écart de modèles

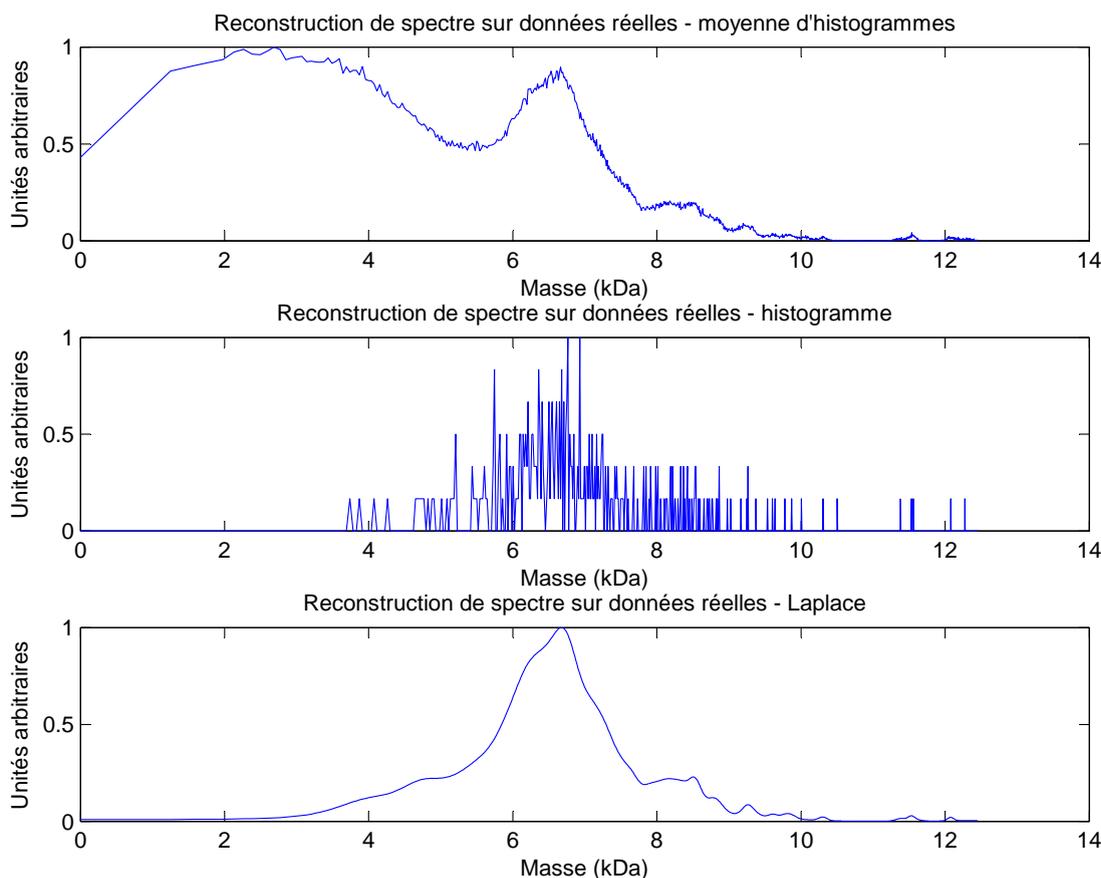


FIGURE 7.3 – Reconstruction d'un spectre sur données réelles

peu probables. Une autre solution pourrait consister à prendre $k_m = 2$ pour pénaliser les faibles valeurs de masse.

La méthode « histogramme » ne permet pas une lecture claire des résultats. En effet, la taille des *bins* est trop petite pour que le spectre ait un rendu lisse. Aussi, nous pouvons constater plusieurs masses éloignées de la valeur nominale des diamètres (de 4 à 12 nm). Ces masses, étant certainement mal estimées, devraient avoir une contribution moindre au spectre, ce que ne prend pas en compte cette méthode. Cette méthode se distingue tout de même par le fait qu'il n'y a pas de spectre de petites masses sur ce spectre.

La dernière méthode, « Laplace », semble être un bon compromis entre les deux précédentes. Le spectre est bien concentré autour de 6.7 nm et prend des valeurs faibles à la fois sur les petits et sur les grands diamètres. Cette méthode nous semble donc adaptée au problème étudié ici. Il serait bénéfique de mener une étude plus détaillée sur la qualité de l'estimation du spectre, en comparaison avec un spectre issu du spectromètre de masse à temps de vol, par exemple.

7.3.3.2 Critique générale

D'une manière plus générale, nous avons une critique importante à formuler à l'encontre de la méthode de reconstruction proposée ici : celle-ci « confond » deux choses. D'une part se trouve l'information sur la masse des molécules apportée par les données $p(m_i|g_k)$. De l'autre côté se trouve l'information apportée au spectre par l'information sur la masse des molécules $p(Sp|m_i)$. Cette méthode identifie l'une à l'autre, dans la mesure où elle consiste à sommer $p(m_i|g_k)$, sans expliciter $p(Sp|m_i)$.

D'une certaine manière, cette méthode consiste à faire l'intégration suivante (supposons N fixé) :

$$\begin{aligned} p(Sp|g_k)_{k \in \{1, K\}} &= \int p(m_i|g_k)_{k \in \{1, K\}} dm_1, \dots, dm_i \dots dm_N \\ &= \int p(g_k|m_i)_{k \in \{1, K\}} p(m_i) dm_1, \dots, dm_i \dots dm_N \end{aligned} \quad (7.2)$$

Or, en exploitant le modèle hiérarchique :

$$p(Sp|g_k)_{k \in \{1, K\}} = p(Sp) \int p(m_i|Sp) p(g_k|m_i)_{k \in \{1, K\}} dm_1, dm_2 \dots, dm_i \dots dm_N \quad (7.3)$$

Il nous paraît donc essentiel de former à la fois l'information *a priori* sur le spectre $p(Sp)$ mais aussi la loi de « tirage » $p(m_i|Sp)$ qui correspond à la question *Je prends une molécule dans le mélange, quelle est sa loi de probabilité en masse ?*. Ceci revient à dire que nous allons donner une forme au spectre, et ainsi expliciter la manière avec laquelle l'information en masse se transcrit en information sur le spectre.

7.4 Développement d'une méthode de reconstruction d'un spectre de raies

7.4.1 Cadre de travail utilisé

Ici, nous allons nous placer dans un cadre particulier de spectre : le spectre de raies. Nous supposons que le nombre de raies est connu : L . Cette hypothèse est une hypothèse très forte. Les raies ont pour valeur Sp_l .

D'une certaine manière, il est possible de mettre en relation notre problématique avec les problématiques d'apprentissage non-supervisé ou d'apprentissage-classification joints, dans lesquelles il s'agit à la fois de classer les molécules mais aussi d'estimer les paramètres des classes. En exemple en neuro-sciences utilisant le contexte bayésien est donné par [FGQ12]. Un algorithme adapté à la résolution de ce problème est l'algorithme d'Espérance-Propagation [Min] (ou *Expectation-Propagation*).

7.4.2 Modélisation statistique des grandeurs et estimation

Voyons maintenant comment nous allons modéliser les différentes grandeurs à l'aide de lois de probabilité. Nous allons reprendre en grande partie le travail effectué au chapitre précédent, en remplaçant $p(m_i)$ par $p(m_i|Sp)p(Sp)$. Les grandeurs inconnues dans notre cas sont les raies du spectre de masse Sp_l , l'intensité de comptage π , et les signaux $m(t)$ et $z(t)$ qui sont paramétrés par le nombre d'événements N , les temps d'adsorption t_i , les amplitudes des masses adsorbées m_i et les positions des molécules adsorbées z_i :

$$\begin{aligned}
 N &\in \mathbb{N} \\
 t_i &\in [0, T]^N \\
 m_i &\in \mathbb{R}^+{}^N \\
 z_i &\in [0, 1]^N \\
 \pi &\in [0, 1] \\
 Sp_l &\in \mathbb{R}^+{}^L
 \end{aligned} \tag{7.4}$$

7.4.2.1 Expression de la fonction de vraisemblance

Comme nous l'avons exprimé à la section 3.4.1 puis au chapitre précédent, section 5.3.2.1, nous modélisons nos connaissances sur les incertitudes de mesure au travers d'un bruit additif, stationnaire et ergodique qui suit une loi normale centrée de covariance $\mathbf{\Gamma}$. Dans notre cas, il s'agira de concaténer les signaux observés d'une part et les signaux estimés d'autre part afin de pouvoir les inclure dans une vraisemblance normale (voir section 3.4.1).

Nous reprendrons la notation précédente afin de ne pas alourdir les équations. Nous noterons donc la vraisemblance :

$$p(\mathbf{g}_k | \underline{m}(t), \underline{z}(t)) = \mathcal{N} \left(\underline{\mathbf{g}_k | \alpha_k (m(t) \diamond \phi_k(z(t))) \star u_{\text{PLL}}(t), \mathbf{\Gamma}} \right)_{k \in \{1; K\}} \quad (7.5)$$

Ici, $\underline{\cdot}$ symbolise la discrétisation du signal puis la mise en forme sous forme de vecteur (concaténation des différentes harmoniques), voir section 3.4.2.2 pour plus de détails.

7.4.2.2 Définition des lois *a priori*

Loi de probabilité sur l'ordre du modèle Comme précédemment, compte-tenu que nous nous inspirons du modèle Bernoulli-Gaussien, nous considérons que l'ordre du modèle N suit la même loi que la somme de variables de Bernoulli indépendantes, c'est-à-dire une loi Binomiale de paramètres π et T :

$$p(N | \pi) = \mathcal{B}(N | \pi, T) \quad (7.6)$$

Loi de probabilité sur les temps d'adsorption Nous travaillerons ici en processus ponctuel marqué uniquement :

$$p(t_i | \pi) = \mathcal{U}_{[0, T]}(t_i) \quad (7.7)$$

Loi de probabilité sur la masse Les masses adsorbées m_i seront distribuées selon le spectre de masse :

$$p(m_i) = \sum_{l=1}^L \frac{1}{L} \delta(m_i - Sp_l) \quad (7.8)$$

où Sp_l sont les raies su spectre de masse.

Loi de probabilité sur la position Nous modéliserons les positions d'adsorption z_i comme uniformément réparties soit sur $[0, 1]$ dans le cas *cantilever*, soit sur $[0, 0.5]$ dans le cas *double-clamped* (compte-tenu de la symétrie du problème). Ici, nous sommes en présence d'une géométrie *double-clamped* :

$$p(z_i) = \mathcal{U}_{[0, 0.5]}(z_i) \quad (7.9)$$

Loi de probabilité sur l'intensité de comptage Comme précédemment, le paramètre d'intensité de comptage, π , sera distribué selon une loi Bêta. Cette loi, à support $[0, 1]$, est la loi *a priori* conjuguée pour le paramètre d'une loi Binomiale (voir tableau 4.1, page 89) :

$$p(\pi) = \mathcal{B}et(\pi|a_\pi, b_\pi) \quad (7.10)$$

où a_π et b_π sont les paramètres de la loi Bêta.

Loi de probabilité sur les raies du spectre de masse Les raies sont distribuées selon une loi à support positif. Dans la lignée des chapitres précédents, nous proposons de prendre une loi Gamma :

$$p(Sp_l) = \mathcal{G}am(Sp_l|k_m, \theta_m) \quad (7.11)$$

où k_m , et θ_m sont respectivement les paramètres de forme et d'échelle de la loi Gamma.

7.4.2.3 Écriture des lois *a posteriori* des paramètres de la liste pour un ordre de modèle et un spectre donné

Compte-tenu de la modélisation formulée ci-dessus, nous pouvons écrire la probabilité *a posteriori* jointe de tous les paramètres de la liste sachant un modèle N donné (ces paramètres sont indépendants de π conditionnellement à N), ainsi qu'un spectre Sp donné :

$$p(t_1, \dots, t_N, m_1, \dots, m_N, z_1, \dots, z_N | N, Sp, \mathbf{g}) \propto \mathcal{N}\left(\mathbf{g}_k | \alpha_k \left(\mathbf{m}(t) \diamond \phi_k(z(t)) \right) \star u_{PLL}(t), \mathbf{\Gamma}\right)_{k \in \{1; K\}} \prod_{i=1}^N \mathcal{U}_{[0, T]}(t_i) \mathcal{U}_{[0, 0.5]}(z_i) \sum_{l=1}^L \frac{1}{L} \delta(m_i - Sp_l) \quad (7.12)$$

7.4.2.4 Écriture des autres lois *a posteriori* conditionnelles

Aussi, il est possible d'écrire la loi *a posteriori* du modèle N , conditionnellement à π :

$$p(N|\pi, \mathbf{g}) \propto \mathcal{B}(N|\pi) \int p(t_1, \dots, t_N, f_1, \dots, f_N | N, \mathbf{g}) dt_i df_i \quad (7.13)$$

Nous écrivons également la loi *a posteriori* conditionnelle à N de l'intensité de comptage π :

$$p(\pi|N, T) = \mathcal{B}et(\pi|a_\pi + N, b_\pi + T - N) \quad (7.14)$$

Enfin, nous écrivons la loi *a posteriori* conditionnelle du spectre Sp :

$$p(Sp|N, m_1, \dots, m_N) = \mathcal{G}am(Sp_l | k_m, \theta_m) \prod_{i=1}^N \sum_{l=1}^L \frac{1}{L} \delta(m_i - Sp_l) \quad (7.15)$$

7.4.2.5 Utilisation d'un algorithme hybride RJMCMC-MCMC

Afin d'explorer convenablement l'espace des modèles et l'espace des paramètres au sein de chaque modèle, nous proposons d'utiliser comme précédemment un algorithme hybride RJMCMC-MCMC.

Nous allons utiliser une structure inspirée de ce que nous avons fait précédemment. Cette structure, est illustrée par la figure 7.4.

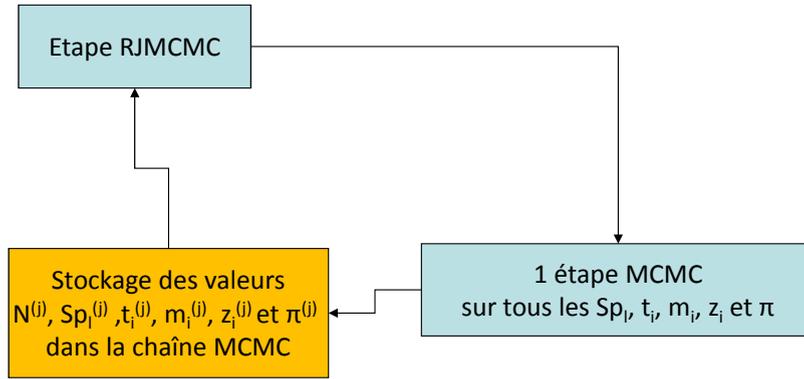


FIGURE 7.4 – Structure de l'algorithme hybride RJMCMC / MCMC en reconstruction de spectre de masse parcimonieux

Les échantillonneurs seront un échantillonneur RJMCMC pour changer l'ordre du modèle N , un échantillonneur Q_1 pour déplacer les temps d'adsorption t_i , un échantillonneur Q_2 pour modifier les masses m_i et les positions z_i , un échantillonneur Q_3 pour échantillonner l'intensité de comptage π et un échantillonneur Q_4 pour échantillonner conjointement le spectre Sp et les masses m_i .

Le principe de l'échantillonneur RJMCMC reste inchangé (plus de détails en annexe E.4). Les échantillonneurs Q_1 et Q_3 restent inchangés également. Seul l'échantillonneur Q_2 est modifié.

Échantillonneur de la masse et de la position Pour échantillonner la masse et la position, nous proposons un algorithme de Metropolis indépendant :

$$Q_2(m_i^p, z_i^p | m_i^{(j-1)}, z_i^{(j-1)}, Sp^{(j)}) = Q_{2a}(m_i^p | m_i^{(j-1)}, Sp^{(j)}) Q_{2b}(z_i^p | z_i^{(j-1)}) \quad (7.16)$$

avec :

$$\begin{aligned}
 Q_{2a}(m_i^p | m_i^{(j-1)}, Sp^{(j)}) &= p(m_i^p) = \sum_{l=1}^L \frac{1}{L} \delta(m_i^p - Sp_l^{(j)}) \\
 Q_{2b}(z_i^p | z_i^{(j-1)}) &= p(z_i^p) = \mathcal{U}_{[0,0.5]}(z_i^p)
 \end{aligned} \tag{7.17}$$

Échantillonneur conjoint du spectre et des masses Nous échantillonons conjointement chacune des raies du spectre avec les masses se situant sur cette raie. En clair nous prenons les indices $i_{0,l}$ tels que $m_{i_{0,l}}^{(j-1)} = Sp_l^{(j)}$ et nous utilisons un échantillonneur à marche aléatoire avec une fonction de proposition normale, centrée sur le point courant et de variance δ_{Sp} :

$$Q_4(Sp_l^{(j)}, m_{i_{0,l}}^{(j-1)}) = \mathcal{N} \left(\begin{pmatrix} Sp_l^p \\ m_{i_{0,l}}^p \end{pmatrix} \middle| \begin{pmatrix} Sp_l^{(j-1)} \\ m_{i_{0,l}}^{(j-1)} \end{pmatrix}, \delta_{Sp} \mathbf{I} \right) \tag{7.18}$$

Estimation il sera possible ici de suivre les chaînes MCMC des informations d'intérêt $Sp_l^{(j)}$. L'estimation pourra se faire au sens de l'Espérance *A Posteriori*, par exemple. Cette simplicité d'estimation vient du fait que nous nous sommes ramenés à un problème en dimension finie et fixée L .

7.4.3 Quelques résultats préliminaires sur signaux synthétiques

7.4.3.1 Modalités de simulation

Nous allons tester notre méthode sur données simulées. Pour simuler les données, nous commençons par générer aléatoirement un signal d'observation sur T échantillons.

Pour cela, nous générons N triplets $\{t_i, m_i, z_i\}$, les t_i étant tirés aléatoirement entre 1 et T . Nous veillons notamment à ce qu'il n'y ait pas deux t_i identiques. Nous choisissons $L = 3$. Les masses seront uniformément tirées sur $\{Sp_1, Sp_2, Sp_3\}$.

Nous générons ensuite le signal résultant de cette liste d'adoption en appliquant (3.12) puis en discrétisant le signal. Nous travaillerons sur $K = 2$ harmoniques. Nous ajoutons sur chaque harmonique k un bruit suivant une loi normale de variance σ_k^2 , les bruits sur chaque mode étant dé-corrélés. Ensuite, nous intégrons ce bruit (fonction Matlab *cumsum*), ce qui en fait un bruit en $1/f$ pur.

Les paramètres de simulation sont donc :

- T le nombre d'échantillons ;
- L le nombre de raies ;
- σ_k l'écart-type du bruit sur l'harmonique k ;
- τ le temps de réponse de la boucle à verrouillage de phase (exprimé en périodes d'échantillonnage) ;
- κ le temps de « dé-synchronisation » entre les deux modes (exprimé en périodes d'échantillonnage) ;
- α la valeur des gains fixes ;

- $\phi(\cdot)$, les gains variables : ceux-ci seront issus de la géométrie *double-clamped* ;
- N le nombre de molécules ;
- les valeurs des raies du spectre Sp .

Nous prendrons des valeurs de bruit et de gain réalistes par rapport aux données observées.

Le tableau 7.5 donne les valeurs des paramètres de simulation.

T	L	σ	τ
4000 échantillons	3 raies	{0.9, 3.9}	30 échantillons
κ	α	N	Sp
0	{0.051, 0.12} Hz.kDa ⁻¹	10 adsorptions	{700, 1000, 1500} kDa

TABLE 7.5 – Valeurs des paramètres de simulation en reconstruction de spectre de masse

7.4.3.2 Paramètres de l'algorithme

Nous lancerons deux algorithmes. Le premier sera un algorithme de reconstruction de spectre décrit dans la section précédente. Le second sera notre algorithme de reconstruction de spectre de raies. Nous initialisons le second algorithme avec le résultat du premier. Les raies, quant à elles, seront initialisées aléatoirement.

Le premier algorithme prend un certain nombre de paramètres, il s'agit notamment :

- du temps de chauffe, noté J_b ;
- du nombre total d'itérations J ;
- du nombre d'itérations B lorsqu'un événement est ajouté ;
- des paramètres des lois de proposition δ_m , δ_z et δ_t ;
- des paramètres des lois *a priori* k_m et θ_m , a_π et b_π .

Le tableau 7.6 donne les paramètres du premier algorithme.

J_b	J	B	δ_m	δ_z	δ_t
4000	5000	100	100	0.1	1

TABLE 7.6 – Valeurs des paramètres de l'algorithme de reconstruction de spectre de masse sur données simulées

Quant à notre méthode, elle aussi prend un certain nombre de paramètres, donnés dans le tableau 7.7. Il s'agit notamment :

- du temps de chauffe, noté J_b ;
- du nombre total d'itérations J ;
- des paramètres des lois de proposition δ_t , δ_z et δ_{Sp} ;

Les loi *a priori* seront les mêmes pour les deux algorithmes et données par le tableau 7.8.

J_b	J	δ_t	δ_z	δ_{Sp}
1500	2000	1	0.1	5

TABLE 7.7 – Valeurs des paramètres de l’algorithme de reconstruction de spectre de masse parcimonieux sur données simulées

k_m	θ_m	a_π	b_π
1	1000	1	T

TABLE 7.8 – Valeurs des paramètres des lois *a priori* pour les algorithmes de reconstruction de spectre de masse sur données simulées

7.4.3.3 Résultats

La figure 7.5 présente le spectre estimé par les méthodes « sans classification », décrites à la section précédente et la figure 7.6 présente le spectre estimé par notre méthode « avec classification ». Pour être exact, la figure 7.6 trace l’évolution des échantillons de Sp_1 , Sp_2 et Sp_3 après temps de chauffe.

Rappelons : $Sp_1 = 700$ kDa, $Sp_2 = 1000$ kDa et $Sp_3 = 1500$ kDa.

Si nous choisissons l’estimateur de l’Espérance *A Posteriori*, notre algorithme estime : $\widehat{Sp}_1 = 725$ kDa, $\widehat{Sp}_2 = 949$ kDa, et $\widehat{Sp}_3 = 1457$ kDa, soit des erreurs relatives de (respectivement) 3.6%, 5.1% et 2.8%.

7.4.3.4 Discussion

Nous voyons dans ces résultats préliminaires que les méthodes proposées à la section précédente ne peuvent pas gérer le problème posé ici. Le spectre reconstruit n’est absolument pas parcimonieux, et les pics ne sont pas centrés sur les vraies valeurs des masses. Ceci est particulièrement visible sur la méthode « Laplace ». Notons tout de même que le problème proposé ici est difficile, avec 30 adsorptions sur 4000 échantillons et une raie à 700 kDa, dont les molécules ne vont pas donner lieu à beaucoup de signal.

La méthode que nous proposons, quant à elle, parvient à estimer correctement les raies du spectre, avec des erreurs inférieures ou égales à 5%.

Cette méthode parvient à exploiter l’information issue de $p(m_i|Sp)$, qui est une information très forte dans ce cas-là. C’est pour cette raison que notre algorithme donne une reconstruction de spectre bien plus fidèle qu’une méthode qui n’explique pas ce lien entre les masses individuelles et la population de molécules étudiée.

Si cet exemple fonctionne plutôt bien, nous avons également été confrontés à des soucis de convergence sur d’autres. Peut-être faudrait-il initialiser l’estimation avec un modèle plus « doux », comme un mélange de lois normales, dont on estimerait les moyennes et dont les variances diminueraient au cours des itérations.

Une extension évidente est celle de la recherche automatique du nombre de raies. Ceci nécessite

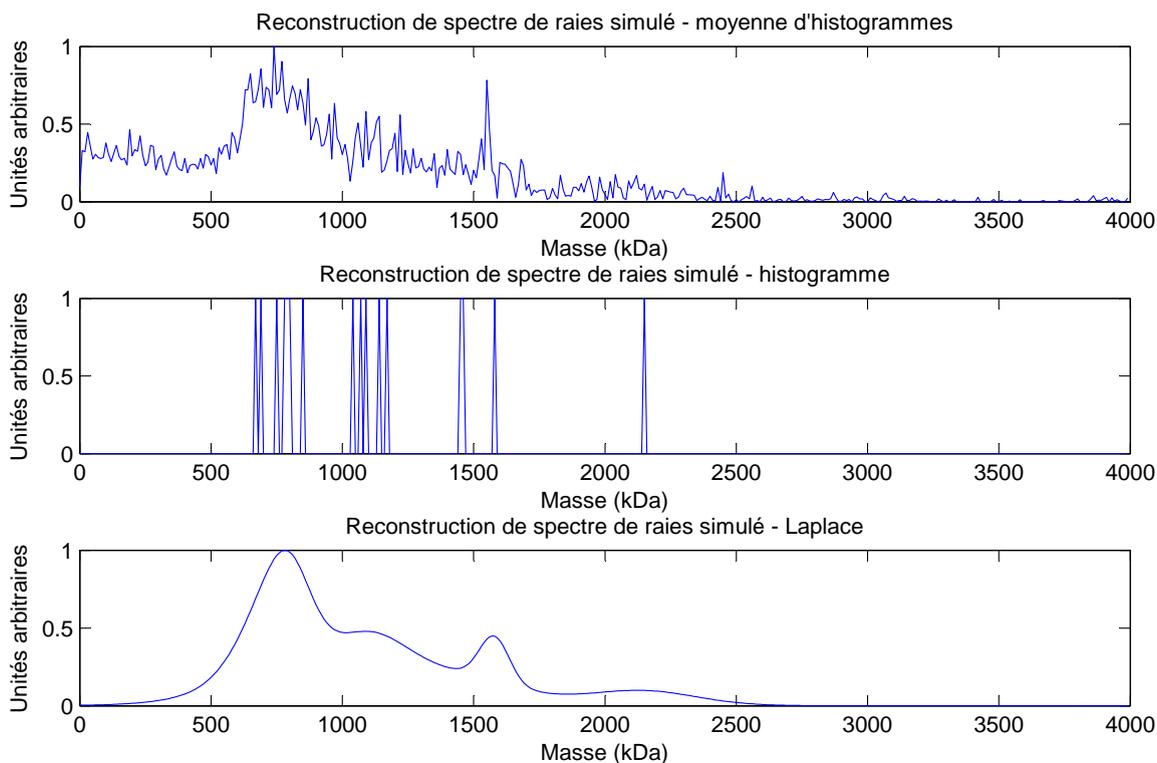


FIGURE 7.5 – Reconstruction d'un spectre parcimonieux « sans classification »

un double choix de modèle (nombre d'adsorptions d'une part et nombre de raies d'une autre part).

Enfin, nous pensons enfin que cette méthode trouverait un sens dans l'analyse de mélanges issus d'échantillons biologiques dans la mesure où dans ce cas, l'hypothèse d'un spectre de raie est une hypothèse réaliste.

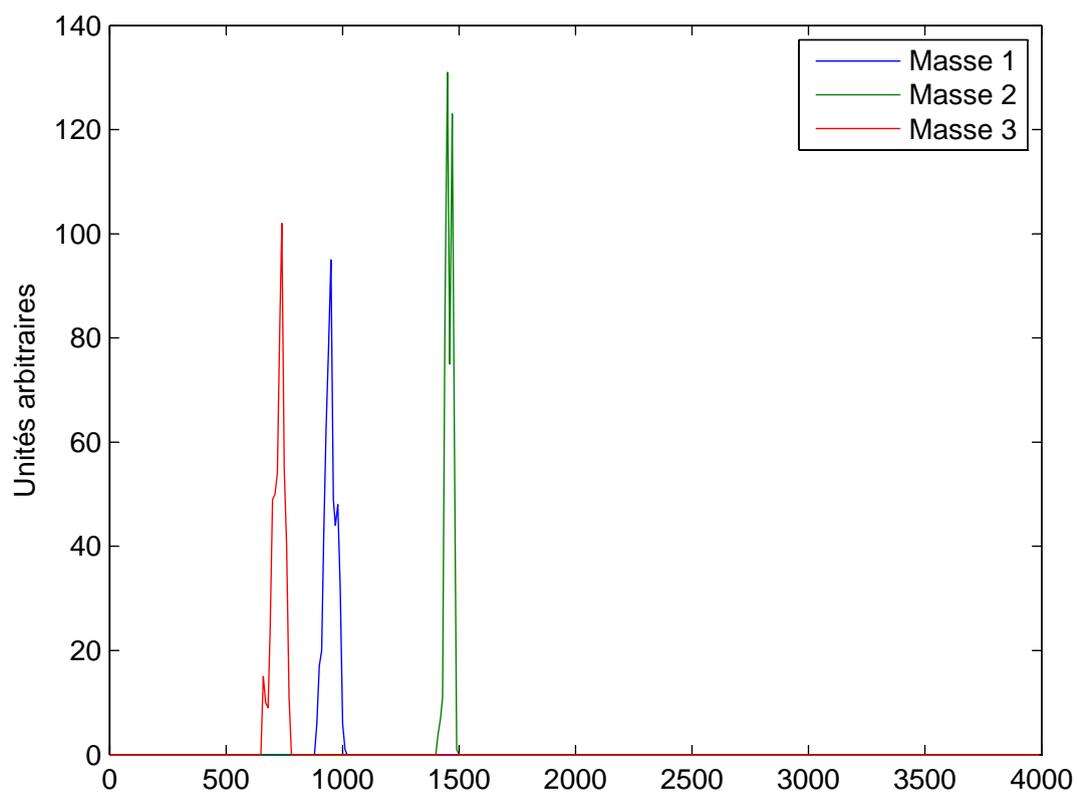


FIGURE 7.6 – Reconstruction d'un spectre parcimonieux « avec classification »

Cinquième partie

Conclusion

Résumé

Nous venons de voir dans ce rapport de thèse une étude des traitements adaptés à la spectrométrie de masse à base de NEMS. Résumons le travail.

Chapitres introductifs

Nous avons vu en premier lieu quel était le contexte applicatif général, à savoir la protéomique et son outil d'analyse qu'est la spectrométrie de masse. Celle-ci vise à identifier des molécules et est une technologie qui fait intervenir plusieurs acteurs et plusieurs briques élémentaires. Une de ces briques est l'analyseur de masse, qui est un appareil visant à transformer l'information de masse des molécules en une information plus facilement manipulable, à savoir dans de nombreux cas une information numérique. Différentes technologies existent pour l'analyseur de masse, différentes au sens qu'elles n'exploitent pas les mêmes phénomènes physiques, mais similaires au sens qu'elles reposent toutes sur la notion de flux de molécules, en l'occurrence des ions.

Nous avons ensuite introduit une nouvelle technologie issue du monde de l'électronique pouvant faire office d'analyseur de masse, les NEMS. Nous avons exposé que les NEMS sont des capteurs nanométriques susceptibles de détecter l'ajout d'une masse de l'ordre de celle d'une protéine. La grandeur à suivre afin d'obtenir une information sur la masse ajoutée au capteur est la fréquence de résonance de ce dernier. A ce titre, nous avons présenté un système permettant de suivre cette fréquence de résonance. Enfin, nous avons montré le système complet permettant d'analyser des molécules grâce à ces capteurs. Celui-ci permet de fonctionner en mode comptage, c'est-à-dire de compter chaque molécule arrivant sur le capteur, et de lui attribuer un paramètre, à savoir la masse.

Modèle et cadre de travail

Ensuite, nous avons exposé une première contribution, à savoir la modélisation du problème sous plusieurs formes (processus à temps discret et processus ponctuel marqué) qui font intervenir une modélisation soit semi-paramétrique soit non-paramétrique. Nous avons aussi proposé une modélisation qui traduit le mode comptage inhérent à notre problématique en une somme de fonctions de Dirac représentant les arrivées des molécules. Dans ce cadre, l'expression du problème direct est une convolution, et son inversion une déconvolution impulsionnelle. Nous avons proposé par la suite une modélisation du bruit ainsi que différentes façons de représenter l'information d'intérêt.

Une fois le modèle clairement exprimé, nous avons listé différentes techniques permettant de retrouver l'information d'intérêt. Si certaines techniques reposant sur une simple observation existent, nous avons tenu à exposer des techniques de déconvolution reposant sur une modélisation plus poussée du système de mesure. Ainsi, nous avons introduit le cadre des problèmes inverses, et nous avons proposé une classification des méthodes d'inversion. Nous avons aussi introduit le cadre de travail bayésien, qui offre à la fois une manière d'exprimer les informations de modélisation et des méthodes d'inversion. Parmi elles, les méthodes d'inversion MCMC ont été présentées.

Contributions méthodologiques et résultats

Ensuite, nous avons exposé notre deuxième contribution, à savoir le développement d'une méthode bayésienne pour inverser le problème en mono-mode, qui s'inscrit dans la démarche de déconvolution impulsionnelle, et qui consiste en la détection des molécules adsorbées. Nous avons expliqué pourquoi le cadre de travail bayésien convenait bien au problème étudié, et nous avons écrit la loi *a posteriori* pour tous les paramètres inconnus. Nous avons basé notre algorithme sur l'exploration de modèle. Pour calculer un estimateur, nous avons choisi une approche mixte RJMCMC-MCMC et mis en place une méthode d'échantillonnage de Metropolis-Hastings adaptée aux spécificités du problème, par le biais de fonctions de proposition adéquates.

Ce modèle a pu aisément être dérivé pour inverser le problème dans un cadre multi-mode. Ce problème, qui consiste en la détection et en la quantification de la masse des molécules adsorbées, comporte deux spécificités. La première vient du fait que les chutes de fréquence engendrées par l'adsorption d'une molécule interviennent simultanément sur plusieurs signaux temporels. La seconde réside dans la nécessité d'inverser une fonction non-linéaire faisant intervenir une variable cachée, à savoir la position de la molécule sur la capteur.

Enfin, ce travail de déconvolution par méthode d'exploration de modèle a permis la mise en place d'une méthode de reconstruction de spectre sans prise de décision intermédiaire. Ceci permet de prendre en compte toutes les sources d'incertitude afin de fournir une estimation robuste pour laquelle l'ensemble des variabilités a été intégré. Nous avons également vu une méthode « tout intégrée » bayésienne de reconstruction de spectre dans le cas d'un spectre de raies. Ces méthodes, actuellement à un degré de maturité faibles, sont très prometteuses pour le problème étudié.

En termes de résultats, nous avons vu que nos méthodes dépassaient, en terme de taux de bonne détection et de fausse détection, les méthodes classiques telles que la pseudo-inverse, le LASSO, *Matching-Pursuit* et étaient équivalentes aux méthodes SMLR les mieux paramétrées. Nos méthodes sont aussi les plus performantes dans le cas où les temps d'adsorption se situent en dehors de la grille d'échantillonnage. Nous parvenons également à obtenir des résultats de quantification plus précis qu'avec des méthodes en deux étapes. Enfin, les reconstructions de spectre de masse sont cohérentes avec les informations expérimentales. Sur données expérimentales, la méthode de reconstruction de spectre de raies permet de caractériser convenablement le spectre de masse d'une solution analysée.

L'intérêt de travailler dans un contexte statistique bayésien est multiple. Déjà, les difficultés liées au caractère mal-posé du problème ou à la présence d'une relation non-linéaire ont pu être aisément contournées. De plus, dans ce contexte, le réglage des paramètres de l'algorithme est assez simple, l'algorithme s'adaptant à la situation qu'il traite. Nous avons simplement choisi ici des lois *a priori* peu informatives relativement au problème étudié. A ce titre, nous n'avons pas fait de campagne d'optimisation paramétrique sur nos algorithmes. Enfin, le cadre de travail bayésien permet de s'inscrire dans une logique homogène et unifiée, qui nous a permis de dériver le modèle pour nous adapter aux trois cas de figure décrits ci-dessus, les modèles étant facilement enrichis lorsque nous passions d'un chapitre à l'autre.

Apport de ce travail

Nous pensons que le travail présenté ci-dessus présente un apport original en ce qui concerne :

- la modélisation du système de mesure et ses perturbations ;
- l’expression du modèle direct sous forme de processus ponctuel marqué et processus à temps discret ;
- l’utilisation d’algorithmes de l’état de l’art (LASSO, *Matching-Pursuit*, SMLR) pour résoudre le problème inverse étudié ici ;
- le développement d’une méthode de déconvolution impulsionnelle reposant sur un choix de modèle : ces méthodes existent mais elles restent assez peu explorées dans la littérature ;
- l’extension de cette méthode au cas multi-mode ;
- la reconstruction de spectre sans déconvolution explicite, et donc sans avoir besoin de choisir un estimateur pour la déconvolution.

Perspectives

Le travail présenté dans le cadre de ce rapport ne s’arrête (malheureusement / heureusement ?) pas à ces conclusions, et un certain nombre de points restent en suspens. Abordons-les en allant du plus proche de notre travail au plus éloigné.

Sur le modèle et l’algorithme

Il existe également des pistes d’amélioration concernant le modèle bayésien et l’algorithme employé.

Sur le modèle bayésien, nous avons fait le choix d’une loi *a priori* Binomiale sur le nombre d’événements car nous nous sommes inspirés du modèle Bernoulli-Gaussien. Or, que se passerait-il avec une autre loi sur le modèle ? Il est possible que l’algorithme explorerait moins de modèles et pourrait donner, avec un *a priori* bien choisi, une reconstruction de spectre plus fidèle.

En ce qui concerne l’algorithme, la mise en place d’une exploration des modèles par éclatement d’un événement nous paraît être une piste à creuser, afin de ne pas rester bloqué dans des *optima* locaux. Nous pourrions aussi trouver des fonctions de proposition plus évoluées permettant une convergence plus rapide, en particulier en ce qui concerne la fonction de proposition de la masse et de la position d’une molécule, qui pourrait être adaptée à la géométrie des fonctions $\phi(\cdot)$. Enfin, nous pensons qu’en termes de déconvolution, il pourrait être mis en place des systèmes d’estimation plus raffinés comme par exemple la mise en place d’un suivi des instants actifs, au lieu d’un simple vote par un vecteur binaire.

Enfin, les propriétés de convergence n’ont pas pu être étudiées, et nous avons d’ailleurs évoqué que si la méthode de reconstruction de spectres de raies paraissait intéressante, sa convergence n’est pas assurée pour le moment.

Sur les résultats

Le second point que nous pourrions approfondir serait celui des résultats. En particulier, de nombreuses données acquises par Eric Sage n'ont pas été traitées faute de temps. Il serait bon de prévoir un traitement automatisé de ces données et de comparer les résultats aux résultats obtenus par le spectromètre de masse à temps de vol et la micro-balance à quartz.

Un point délicat dans l'analyse de ces résultats vient du triptyque de sources d'erreurs suivant :

- le capteur NEMS n'est pas parfait, et son modèle est sûrement incomplet, nos collègues physiciens remettent notamment en cause la forme des fonctions de gain variable $\phi(\cdot)$ et la valeur des gains fixes α ;
- l'information apportée par le spectromètre de masse à temps de vol n'est pas parfaite non plus, celui-ci travaillant de surcroît dans une gamme de masse (plusieurs centaines de kDa) située sensiblement au-dessus de sa gamme habituelle ;
- l'algorithme de reconstruction de spectre peut lui aussi commettre des erreurs.

Sur les algorithmes alternatifs

Le point suivant consiste à effectuer une comparaison exhaustive de nos méthodes avec les méthodes de l'état de l'art. Si la partie déconvolution mono-mode a été assez bien testée en comparaison de performances (bien que ce genre d'étude soit extensible à merci) les aspects déconvolution multi-mode (avec les algorithmes *Simultaneous Orthogonal Matching Pursuit*, *Group LASSO* pour ne citer qu'eux) peuvent largement être approfondis. Aussi, la méthode de déconvolution multi-mode reposant sur un modèle Bernoulli-Gaussien offre des performances très proches de notre algorithme, et il est possible de les améliorer avec une loi *a priori* normale sur les amplitudes prenant mieux en compte la distribution des chutes de fréquence, c'est-à-dire une loi normale hétéroscédastique et corrélée. La comparaison avec la méthode de quantification de Hanay *et al.* devrait également être étudiée.

Il conviendrait également de tester plusieurs assemblages de chaînes de traitement, constitués d'assemblages d'algorithmes de l'état de l'art, d'algorithmes à adapter et des algorithmes proposés. Il est possible qu'un compromis performances / temps de calcul s'opère entre notre méthode et de tels assemblages, qui seraient plus rapide que notre méthode. Il serait notamment intéressant de voir l'apport de notre méthode en terme d'estimation de spectre, dans la mesure où cette dernière reconstruit les spectres sans prise de décision intermédiaire, par rapport à une méthode d'estimation de spectre classique mise en aval de notre algorithme de détection/quantification.

Les points délicats sur cette partie sont les suivants.

- Le grand nombre de chaînes de traitement possibles : avec ne serait-ce que 3 algorithmes de détection, 3 algorithmes de quantification et 3 algorithmes de reconstruction de spectre, nous arrivons déjà à un total de 27 assemblages possibles !
- Le bon choix des critères de performance, en relation avec l'application visée : notamment, le temps de calcul est-il critique pour notre application ? Est-ce que les pertes de comptage sont dommageables ou a-t-on suffisamment de données ? Quelle est la taille typique des molécules étudiées, et question corolaire : quel est le rapport signal à bruit ? Quel est le taux de comptage et quel doit être le pouvoir de séparation temporel des impulsions de l'algorithme ?

- l’information demandée par l’algorithme et l’automatisation du traitement. En effet, il peut arriver qu’une méthode directe très bien réglée en terme de paramètres présente de très bons résultats. Or, sera-t-elle aussi bonne dans d’autres conditions expérimentales ? Dans quel mesure a-t-on besoin d’apprendre des paramètres par des méthodes secondaires (calibration notamment) ?

Compromis pour optimiser les temps de calcul

Nous avons mis en place une méthode de déconvolution puissante. Malheureusement, cette dernière est assez gourmande en temps de calcul. Certaines pistes pourraient permettre de réduire ce temps de calcul, parmi eux, le calcul itératif de la fonction de vraisemblance (là où nous nous contentons ici de mettre à jour itérativement le modèle direct).

D’autres pistes permettraient de rester proches de la méthode tout en réduisant encore le temps de calcul, comme par exemple assumer l’hypothèse de blancheur du bruit (ou avec corrélation uniquement d’un mode à l’autre).

Enfin, d’autres pistes, qui nous feraient sortir du cadre purement bayésien « tout intégré » peuvent être exploitées. Il est notamment possible d’effectuer la détection des événements par une méthode externe, l’algorithme MCMC se chargeant d’estimer les autres paramètres (masse, position et spectre en l’occurrence). D’une manière plus douce, il est possible de pré-sélectionner les instants candidats pour être temps d’adsorption et d’utiliser cette information comme *a priori* sur les temps d’adsorption.

Extensions du modèle

Le modèle pourrait être étendu pour plus d’automatisation, notamment en estimant conjointement aux autres paramètres les paramètres τ , κ et les paramètres de bruit. *La question à se poser est, à notre sens : cela apporte-t-il quelques chose relativement à l’application visée ? En particulier, y-aura-t-il une phase de calibration dans le fonctionnement réel du système ?*

D’une manière plus profonde, le modèle que nous avons mis en place pourrait être enrichi des problématiques suivantes :

- gestion de réseaux de capteurs (éventuellement hétérogènes) ;
- recherche du nombre de raies lors de la reconstruction du spectre de masse, ce qui nécessiterait un double choix de modèle (nombre d’adsorption ET nombre de raies) ;
- utilisation de l’algorithme de reconstruction de spectre dans d’autres cas que le spectre de raies : les spectres de masse expérimentaux peuvent, selon nos collègues, être approximés par une loi log-normale.

Utilisation avec d’autres modalités d’acquisition

Enfin, le modèle que nous avons développé et les algorithmes qui en découlent peuvent être adaptés à bon nombre d’autres problématiques de déconvolution impulsionnelle. D’une manière plus générale, les traitements en mode comptage sont assez généraux en traitement de l’information. Nous pouvons notamment voir un domaine d’application possible dans la surveillance de

l'environnement et notamment l'exploitation d'images aériennes voire satellitaires pour compter des éléments naturels (arbres par exemple), de la faune, des piétons ou des membres d'une foule, des véhicules, etc. . .

Notons aussi que cette approche détection – quantification – comptage prend un sens très fort dans le cadre des nanotechnologies. En effet, dans ce cadre-là, il est possible de mesurer une information élémentaire à très petite échelle (dans notre cas, la molécule unique). L'idée générale est donc d'effectuer une détection-estimation (ou quantification) sur chacune des informations élémentaires, suivie de la fusion de toutes les données (ici, comptage). Notre démarche « tout intégrée » pourrait donc être une base à de futurs travaux similaires en nanotechnologies.

Sixième partie

Annexes

Table des annexes

A	Quelques lois de probabilité	197
A.1	Loi normale	197
A.2	Loi uniforme	197
A.3	Loi Gamma	197
A.4	Loi Binomiale	198
A.5	Loi Bêta	198
B	Pré-traitements	199
B.1	Recalage des horloges	199
B.2	Suppression du bruit de secteur	200
C	Étude sur le bruit	201
C.1	Présentation des signaux de bruit et analyse	201
C.2	Matrice de covariance	209
C.2.1	Estimation de la matrice de covariance	209
C.2.2	Le sens de la matrice de covariance	212
C.2.3	Validation du calcul de vraisemblance	212
D	Méthodes d'inversion du système d'équations non-linéaires	215
D.1	Contexte	215
D.2	Méthode de recherche d'une solution aux équations non-linéaires	215
D.3	Méthode de recherche d'une solution aux équations non-linéaires	216
E	Détails des calculs	219
E.1	Démonstration de la relation (4.42)	219
E.2	Calcul de la probabilité d'acceptation dans le cadre d'un échantillonneur de Gibbs	220
E.3	Marginalisation de l'inverse variance d'une loi normale	220
E.4	Calculs d'acceptation des algorithmes RJMCMC	221
E.4.1	Processus de naissance et de mort	221
E.4.2	Ajout d'une entrée dans la liste	221
E.4.3	Calcul de la probabilité d'acceptation	222

F	Bibliographie personnelle	225
F.1	Travail portant sur la déconvolution mono-mode	225
F.2	Travail portant sur la déconvolution multi-mode	226
F.3	Travail portant sur la reconstruction de spectre	226

Quelques lois de probabilité

A.1 Loi normale

La loi normale multivariée prend deux paramètres à savoir le vecteur moyenne $\boldsymbol{\mu}$ et la matrice de covariance $\boldsymbol{\Sigma}$ (en dimension N).

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \exp -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \quad (\text{A.1})$$

Sous sa forme mono-variée, la loi normale prend deux paramètres, la moyenne μ et l'écart-type σ .

$$\mathcal{N}(x|\mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \quad (\text{A.2})$$

A.2 Loi uniforme

La loi normale uniforme prend deux paramètres de borne inférieure m et supérieure M .

$$\mathcal{U}_{[m,M]}(x) = \begin{cases} \frac{1}{M - m} & \text{si } x \in [m, M] \\ =0 & \text{sinon} \end{cases} \quad (\text{A.3})$$

A.3 Loi Gamma

La loi Gamma prend deux paramètres de forme k et d'échelle θ . Cette loi prend des valeurs non nulles sur \mathbb{R}^+ .

$$\begin{aligned} \mathcal{G}am(x|k, \theta) &= \frac{x^{k-1} \exp \frac{-x}{\theta}}{\Gamma(k) \theta^k} \text{ si } x \geq 0 \\ &= 0 \text{ sinon} \end{aligned} \tag{A.4}$$

A.4 Loi Binomiale

La loi Binomiale est la loi que suit la somme de variable de Bernoulli (épreuves valant 1 avec la probabilité π et 0 avec la probabilité $1 - \pi$). Cette loi prend deux paramètres de succès π et de nombre d'épreuves T . Cette loi prend des valeurs non nulles sur \mathbb{N}^+ .

$$\begin{aligned} \mathcal{B}(x|\pi, T) &= \binom{T}{x} \pi^x (1 - \pi)^{T-x} \text{ si } x \geq 0 \\ &= 0 \text{ sinon} \end{aligned} \tag{A.5}$$

A.5 Loi Bêta

La loi Bêta prend deux paramètres a et b . Cette loi prend des valeurs non nulles sur l'intervalle $[0, 1]$.

$$\begin{aligned} \mathcal{B}eta(x|a, b) &= \frac{1}{B(a, b)} x^{a-1} (1 - x)^{b-1} \text{ si } x \in [0, 1] \\ &= 0 \text{ sinon} \end{aligned} \tag{A.6}$$

Pré-traitements

B.1 Recalage des horloges

Un problème mis en évidence à la section 3.2.2.3 est celui de la dé-synchronisation des horloges sur les différents modes. Pour pallier à ce problème, nous proposons un pré-traitement basé sur les deux idées suivantes :

- la corrélation est une grandeur caractéristique du degré de similarité de deux signaux ;
- les bruits sont corrélés de mode à mode, et les pics formés par le signal dérivé sont censés être alignés.

Ainsi, nous proposons le traitement suivant :

- dériver les signaux observés ;
- faire la corrélation de ces signaux dérivés (fonction Matlab *xcorr*) ;
- repérer le *maximum* de la corrélation M_{corr} .

Si les signaux initiaux sont de taille T , alors le *maximum* de corrélation devrait se trouver en $T - 1$. Ainsi, le déphasage est égal à l'écart entre le *maximum* et $T - 1$:

$$\hat{\kappa} = T - 1 - M_{\text{corr}} \quad (\text{B.1})$$

Il est aussi possible de chercher les deux *maxima* de la corrélation et d'estimer le déphasage par interpolation.

Une fois le déphasage estimé, il est possible de :

- soit recalculer un signal par rapport à l'autre ;
- soit utiliser ce déphasage estimé $\hat{\kappa}$ en tant que paramètre κ du modèle direct.

B.2 Suppression du bruit de secteur

Afin de supprimer le bruit de secteur mis en évidence sur les figures de densité spectrale de puissance C.3, nous employons la méthode suivante pour supprimer les bruits impulsionnels présents tous le 50 Hz.

Nous commençons par définir les bandes de fréquence à supprimer. Pour un signal échantillonné à 500 Hz, celles-ci sont [49, 51] Hz, [99, 101] Hz, [149, 151] Hz, [199, 201] Hz et [249, 250] Hz.

Le procédure consiste ensuite à :

- faire une transformée de Fourier du signal (fonction Matlab *fft*);
- pour chaque bande à retirer :
 - sélectionner quelques échantillons de la transformée de Fourier en amont et en aval de la bande considérée (soit à des fréquences légèrement supérieures ou inférieures à la bande),
 - calculer une moyenne des échantillons la zone amont et une moyenne dans la zone avale,
 - remplacer le contenu de la bande (contenu entre les deux points extrêmes) par une interpolation linéaire entre la moyenne amont et la moyenne avale,
- rendre le signal symétrique (dans le domaine de Fourier) pour sa partie réelle et antisymétrique pour sa partie imaginaire ;
- calculer la transformée de Fourier inverse du signal (fonction Matlab *ifft*).

Étude sur le bruit

C.1 Présentation des signaux de bruit et analyse

Nous allons ici illustrer l'étude sur le bruit menée dans la section 3.4.1, en illustrant les signaux temporels de bruit C.1 pour le signal A et C.2 pour le signal B, la Densité Spectrale de Puissance C.3 pour le mode 1 et C.4 pour le mode 2 et la Densité Spectrale de Puissance après pré-traitement et dérivation C.5 pour le mode 1 et C.6 pour le mode 2.

La figure C.7, quant à elle, illustre les auto-corrélations des signaux (fonction Matlab *xcorr*) sur chacun des modes, après pré-traitement et dérivation.

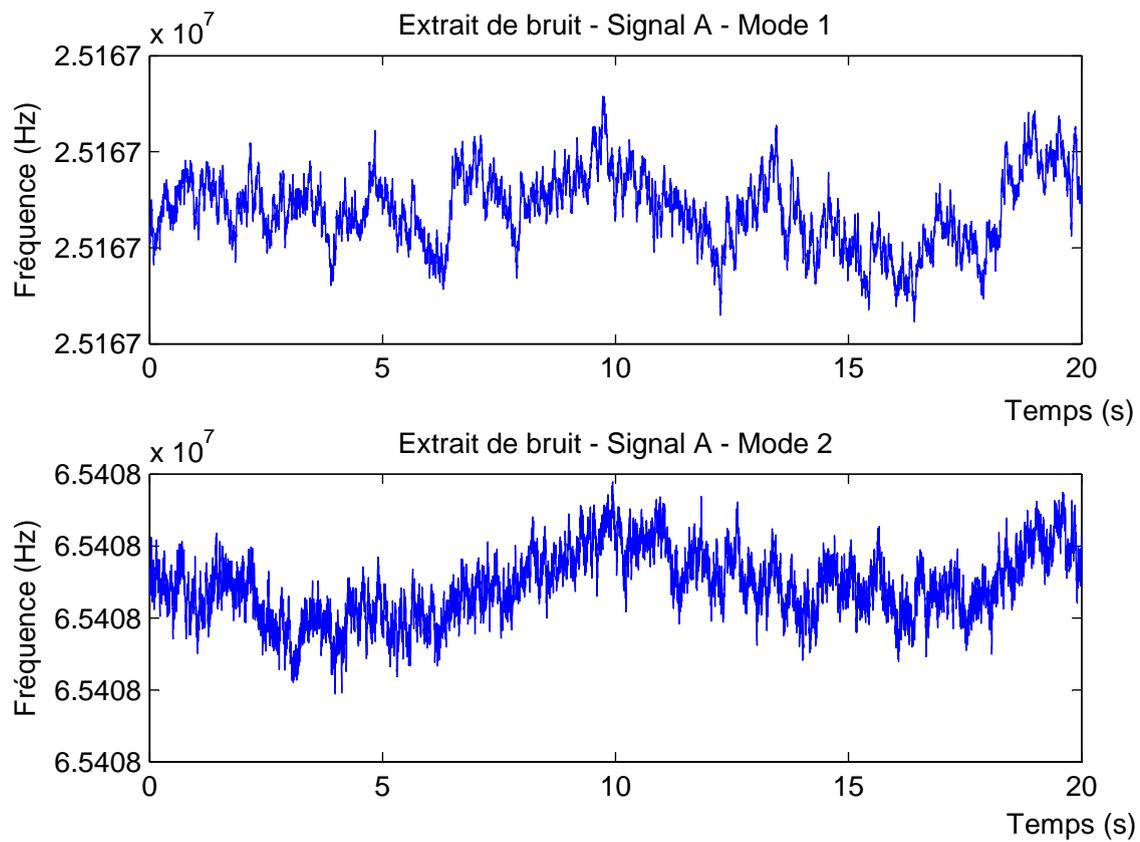


FIGURE C.1 – Illustration du signal A correspondant un bruit seul pour les modes 1 (en haut) et 2 (en bas)

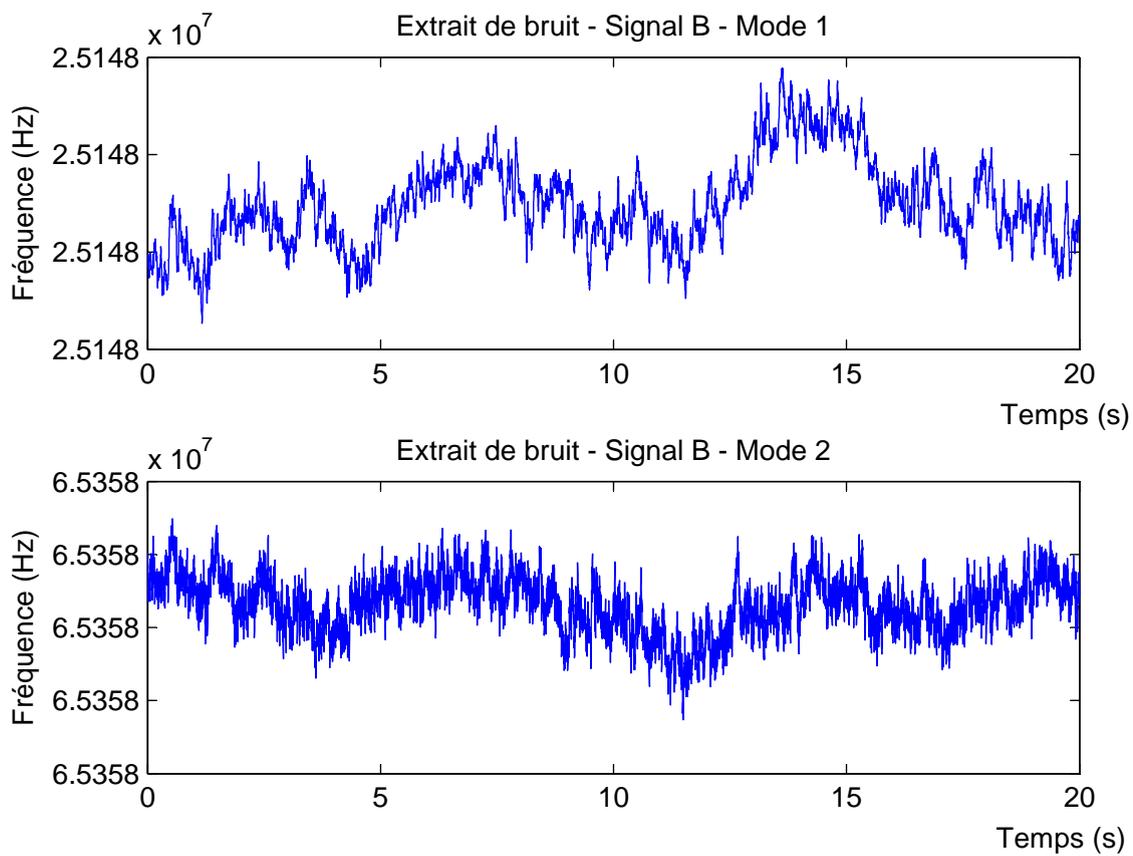


FIGURE C.2 – Illustration du signal B correspondant un bruit seul pour les modes 1 (en haut) et 2 (en bas)

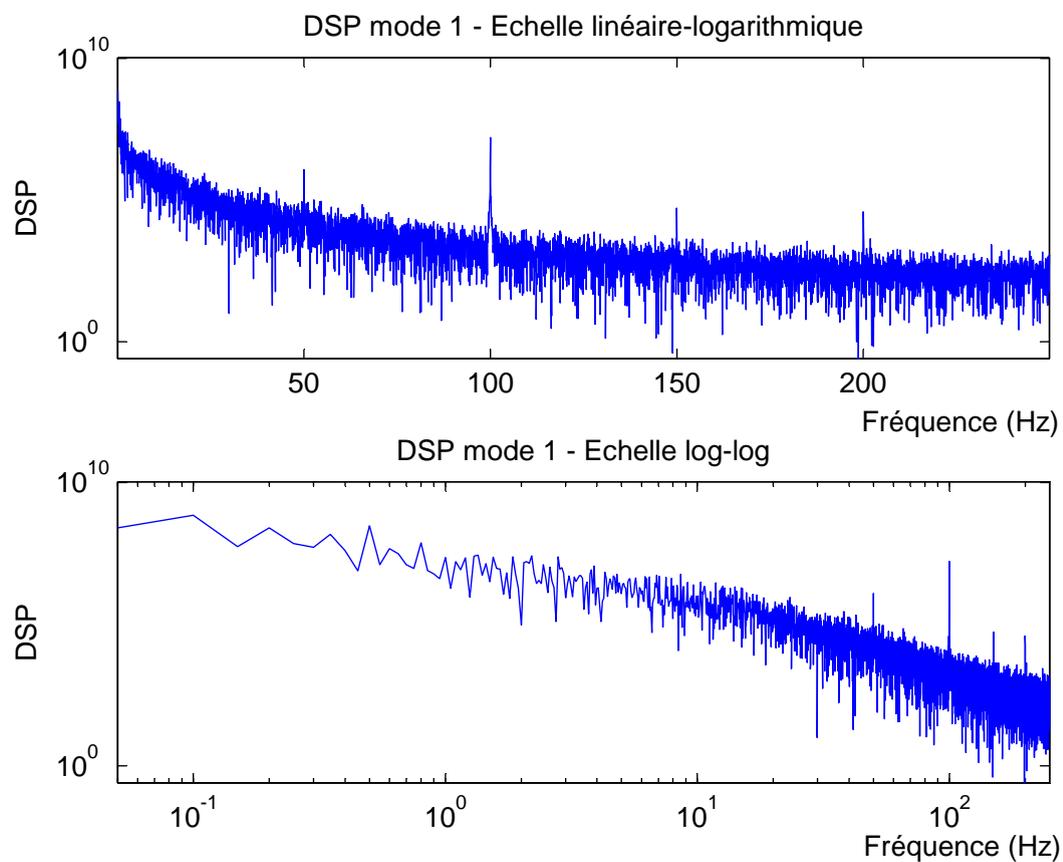


FIGURE C.3 – Densité Spectrale de Puissance (DSP) du bruit du mode 1 correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)

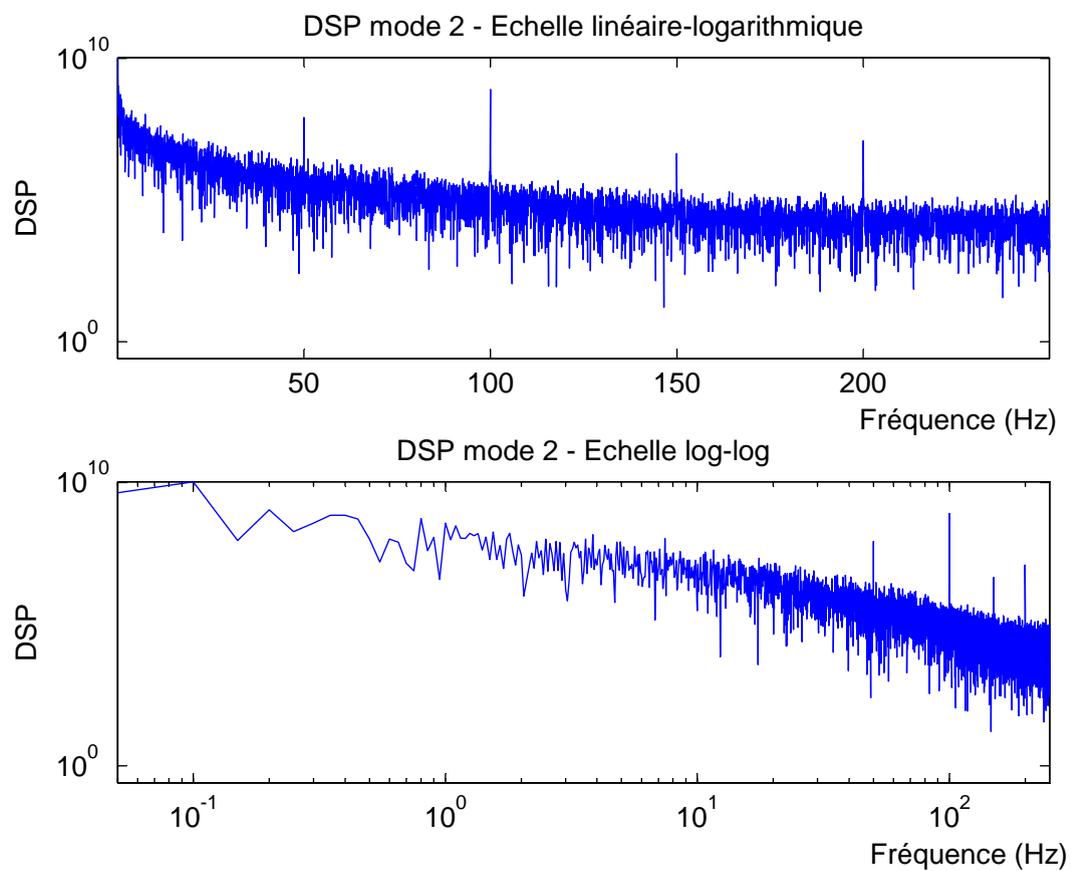


FIGURE C.4 – Densité Spectrale de Puissance (DSP) du bruit du mode 2 correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)

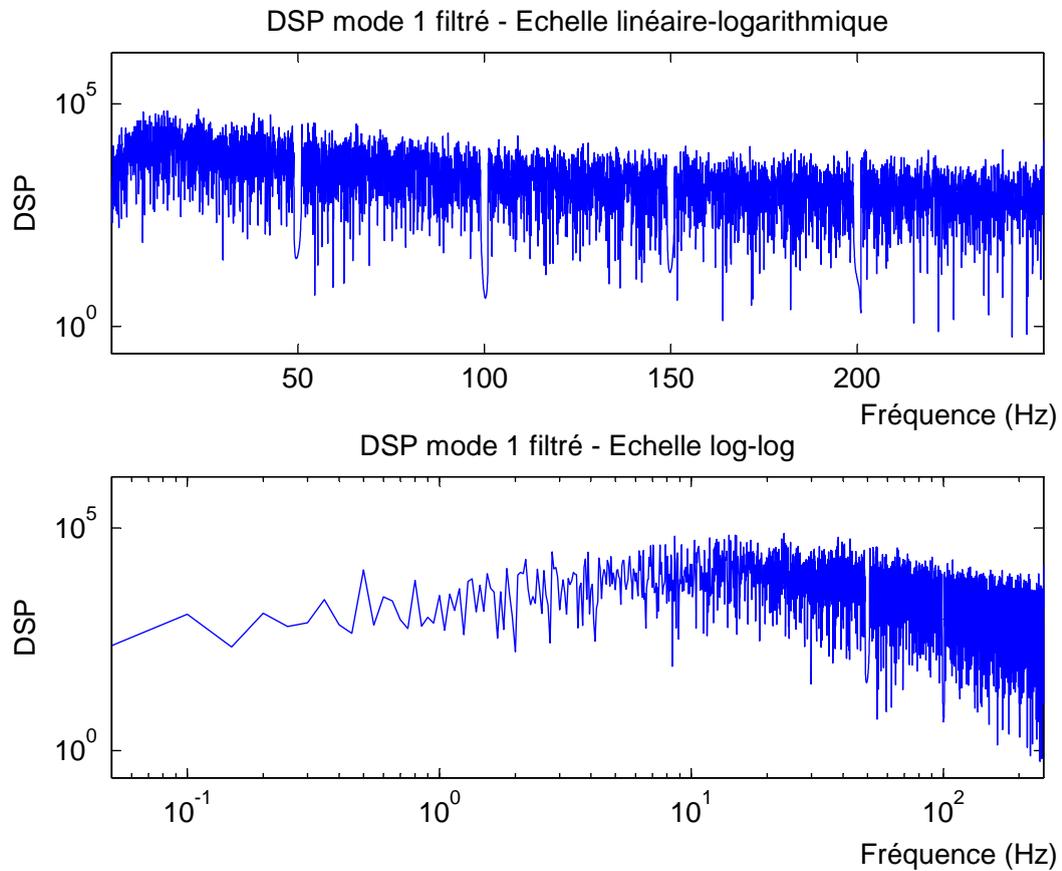


FIGURE C.5 – Densité Spectrale de Puissance (DSP) du bruit du mode 1 après pré-traitement et dérivation correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)

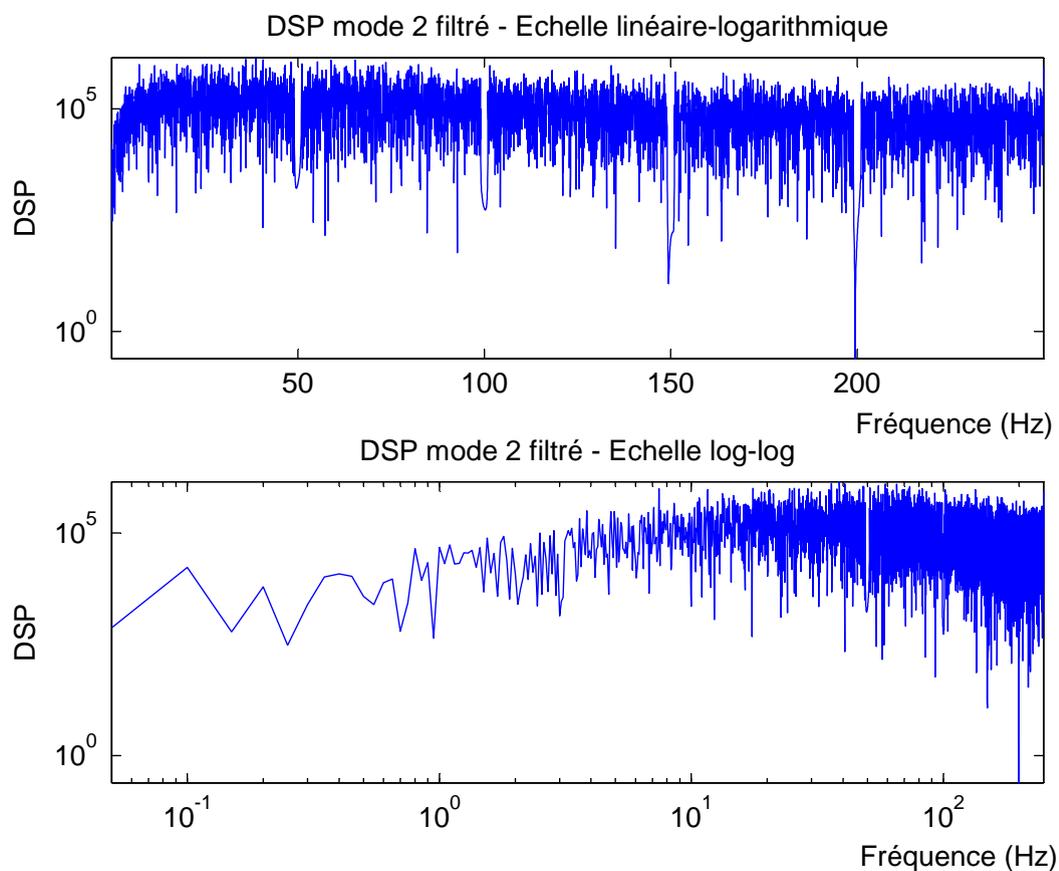


FIGURE C.6 – Densité Spectrale de Puissance (DSP) du bruit du mode 2 après pré-traitement et dérivation correspondant au signal A en échelle linéaire logarithmique (en haut) et en échelle logarithmique-logarithmique (en bas)

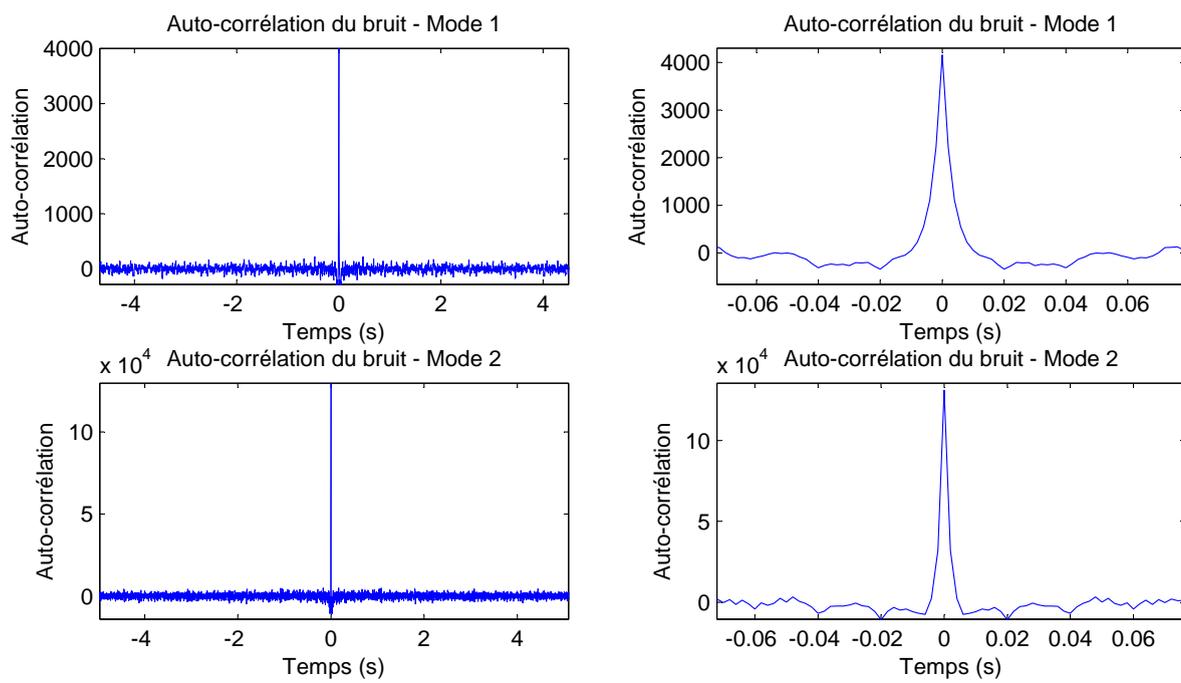


FIGURE C.7 – Auto-corrélation des signaux après-pré-traitement et dérivation pour les modes 1 (en haut) et 2 (en bas)

C.2 Matrice de covariance

Détaillons-ici les points vus à la section 3.4.2.

C.2.1 Estimation de la matrice de covariance

Nous notons $\mathbf{x}_{k,u}$ la u -ième composante du vecteur du signal de bruit sur l'harmonique k , et nous construisons le vecteur suivant :

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{1,1} \\ \mathbf{x}_{1,2} \\ \vdots \\ \mathbf{x}_{1,T} \\ \mathbf{x}_{2,1} \\ \mathbf{x}_{2,2} \\ \vdots \\ \mathbf{x}_{2,T} \end{pmatrix} \quad (\text{C.1})$$

Nous supposons ici :

$$p(\mathbf{x}|\mathbf{\Gamma}) = \mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Gamma}) \quad (\text{C.2})$$

Estimons cette matrice de covariance $\mathbf{\Gamma}$. Supposons U observations indépendantes de \mathbf{x} notées $\mathbf{x}^{(u)}$. L'estimateur du *maximum* de vraisemblance donne :

$$\hat{\mathbf{\Gamma}} = \frac{1}{U} \sum_{u=1}^U \mathbf{x}^{(u)} \mathbf{x}^{(u)T} \quad (\text{C.3})$$

Notons que cet estimateur est biaisé (un estimateur sans biais, appelé *Sampling Covariance Matrix* consiste à remplacer $\frac{1}{U}$ par $\frac{1}{U-1}$).

N'ayant pas accès à U observations indépendantes du bruit, nous proposons de considérer des vecteurs de bruit plus courts. Nous construisons le vecteur d'observation suivant :

$$\mathbf{x}^{(u)} = \begin{pmatrix} \mathbf{x}_{1,u} \\ \mathbf{x}_{1,u+1} \\ \vdots \\ \mathbf{x}_{1,u+V-1} \\ \mathbf{x}_{2,u} \\ \mathbf{x}_{2,u+1} \\ \vdots \\ \mathbf{x}_{2,u+V-1} \end{pmatrix} \quad (\text{C.4})$$

Ici, V est un ordre de représentation (jusqu'auquel nous considérons que les échantillons peuvent être corrélés temporellement). Nous avons donc accès à $U' = \frac{U}{V}$ observations indépendantes.

Prenons $V = 10$. Dans ce cas, nous bénéficions de 1000 échantillons pour estimer la matrice de covariance d'ordre V après dérivation nommée \mathbf{F}_{2V} . Nous prendrons l'estimateur du *maximum* de vraisemblance pour estimer cette matrice de covariance. La figure C.8 illustre la matrice de covariance estimée.

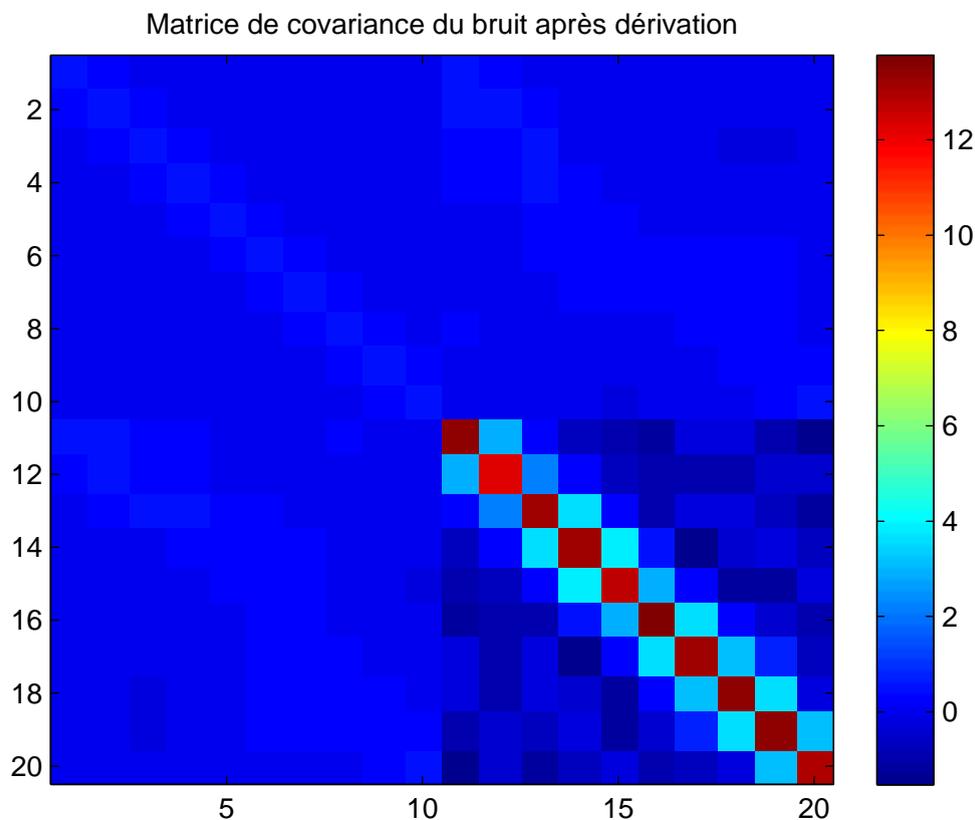


FIGURE C.8 – Estimation de la matrice de covariance du bruit pour $V = 10$

La corrélation temporelle du mode 1 est représentée dans le quart en haut à gauche. La corrélation temporelle du mode 2 est représentée dans le quart en bas à droite. Les corrélations temporelles inter-modes sont dans les deux autres quarts.

Nous voyons une forte asymétrie entre les éléments des deux modes. Le mode 1 apparaît en bleu tant la dynamique sur celui-ci est faible par rapport au mode 2. Pour plus de clarté, nous allons ré-égaliser les échantillons. Recalculons la matrice de covariance après avoir normalisé les échantillons de bruit par leur écart-type estimé sur les deux modes. Les valeurs de ces écarts-types ont été données dans le tableau 3.3. La figure C.9 illustre cette matrice après égalisation.

Nous voyons que si les pré-traitements ont redressé la Densité Spectrale de Puissance, il reste de corrélations au sein de la matrice de covariance (termes non-diagonaux). Si le filtrage de raies à 50 Hz, la dérivation et l'égalisation suppriment des corrélations dans le bruit, nous proposons

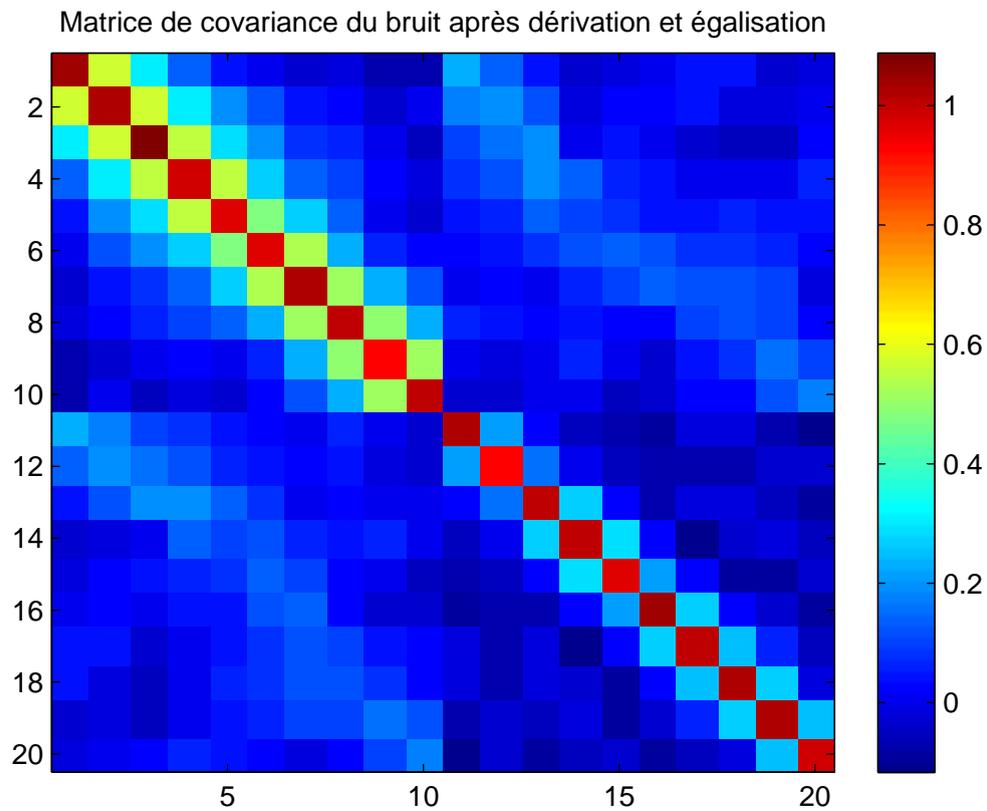


FIGURE C.9 – Estimation de la matrice de covariance du bruit pour $V = 10$ après égalisation
une méthode afin de prendre simplement en compte les corrélations résiduelles.

C.2.2 Le sens de la matrice de covariance

Notons que la matrice de covariance a le rôle d'une matrice de blanchiment par rapport au bruit. En effet, lorsque l'on écrit la log-vraisemblance d'une loi normale multivariée, nous obtenons :

$$\begin{aligned} -\log(\mathbf{x} \sim \mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Gamma})) &= \eta + \mathbf{x}^T \mathbf{\Gamma}^{-1} \mathbf{x} \\ &= \eta + \left(\mathbf{\Gamma}^{-\frac{1}{2}} \mathbf{x}\right)^T \left(\mathbf{\Gamma}^{-\frac{1}{2}} \mathbf{x}\right) \end{aligned} \quad (\text{C.5})$$

où η est un terme de normalisation, et $\mathbf{\Gamma}^{-\frac{1}{2}}$ une matrice telle que $\mathbf{\Gamma}^{-\frac{1}{2} T} \mathbf{\Gamma}^{-\frac{1}{2}} = \mathbf{\Gamma}^{-1}$.

On retrouve dans l'équation (C.5) la vraisemblance normale d'un bruit blanc avec le vecteur de bruit : $\left(\mathbf{\Gamma}^{-\frac{1}{2}} \mathbf{x}\right)$. Cela signifie qu'il est équivalent de considérer \mathbf{x} comme un bruit suivant une loi normale centrée de matrice de covariance $\mathbf{\Gamma}$ et de considérer $\left(\mathbf{\Gamma}^{-\frac{1}{2}} \mathbf{x}\right)$ comme un bruit suivant une loi normale centrée de matrice de covariance \mathbf{I} (soit un bruit blanc). Ainsi, nous pouvons considérer que la matrice $\mathbf{\Gamma}^{-\frac{1}{2}}$ permet de blanchir le bruit \mathbf{x} . Cela revient à dire que le bruit correspond à un bruit blanc filtré par la matrice $\mathbf{\Gamma}^{\frac{1}{2}}$.

L'opération de dérivation, qui peut s'écrire $\mathbf{D}\mathbf{x}$ (étant linéaire), ne gêne en rien le calcul de matrice de covariance. En effet, il est possible d'écrire la relation suivante entre la matrice de covariance calculée après dérivation $\mathbf{\Gamma}_2$ et la matrice calculée sans dérivation $\mathbf{\Gamma}$, si l'on note \mathbf{x}_2 le signal dérivé et \mathbf{x} le signal non dérivé :

$$\begin{aligned} \mathbf{\Gamma}_2 &= \frac{1}{U} \mathbb{E} [\mathbf{x}_2 \mathbf{x}_2^T] \\ &= \frac{1}{U} \mathbb{E} [\mathbf{D} \mathbf{x} \mathbf{x}^T \mathbf{D}^T] \\ &= \mathbf{D} \mathbf{\Gamma} \mathbf{D}^T \end{aligned} \quad (\text{C.6})$$

Notons que les pré-traitements n'étaient finalement qu'une étape de blanchiment supplémentaire.

C.2.3 Validation du calcul de vraisemblance

Nous allons désormais valider notre méthode de calcul de vraisemblance en vérifiant le caractère blanchisseur de notre méthode. Concrètement, nous vérifions que notre procédure a aussi bien blanchi le bruit que l'opération $\left(\mathbf{\Gamma}^{-\frac{1}{2}} \mathbf{x}\right)$ induite par le calcul de la vraisemblance normale. Pour cela, la figure C.10 compare les matrices de covariances à l'ordre 10 obtenues sur le signal A et sur le signal B, tous deux blanchis avec $\mathbf{\Gamma}^{-\frac{1}{2}}$, où $\mathbf{\Gamma}$ a été apprise sur le signal A. Nous présentons aussi la différence entre les deux matrices de covariance.

La figure C.11, quant à elle, compare les matrices de covariances à l'ordre 10 obtenues sur le signal A et sur le signal B, tous deux blanchis avec notre méthode, les paramètres étant appris sur le signal A. Nous présentons aussi la différence entre les deux matrices de covariance.

Nous voyons sur ces figures que si une procédure classique permet de blanchir le bruit plus efficacement sur le signal d'apprentissage, sur un signal test (le signal B), les deux procédures sont équivalentes. A ce titre, la norme de Frobenius (définie comme la norme 2 du vecteur formé des valeurs singulières d'une matrice) de la matrice de différence vaut, pour la procédure classique : 1.06, et dans notre cas : 0.98.

Il serait intéressant de savoir si cette différence est significative, auquel cas cela signifierait que nous avons extrait des propriétés du bruit plus génériques dans notre cas que dans le cas classique. Faute de temps, nous n'avons pas mené cette étude.

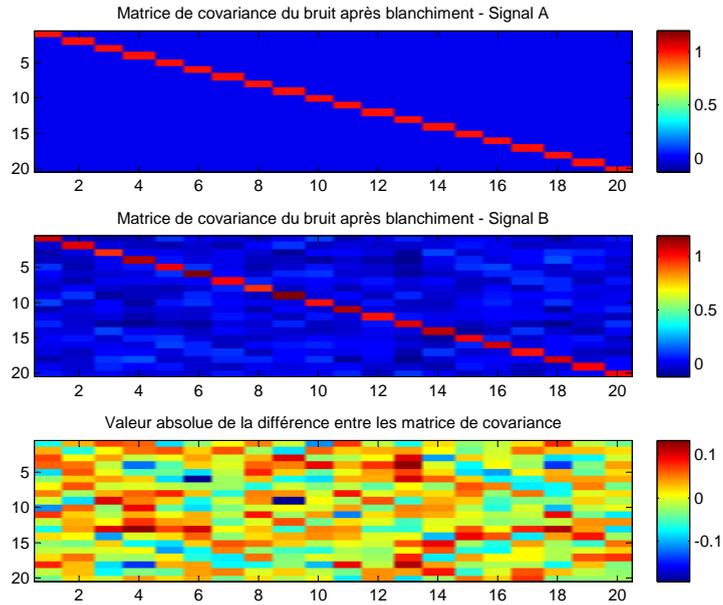


FIGURE C.10 – Illustration du blanchiment du bruit par $\Gamma^{-\frac{1}{2}}$

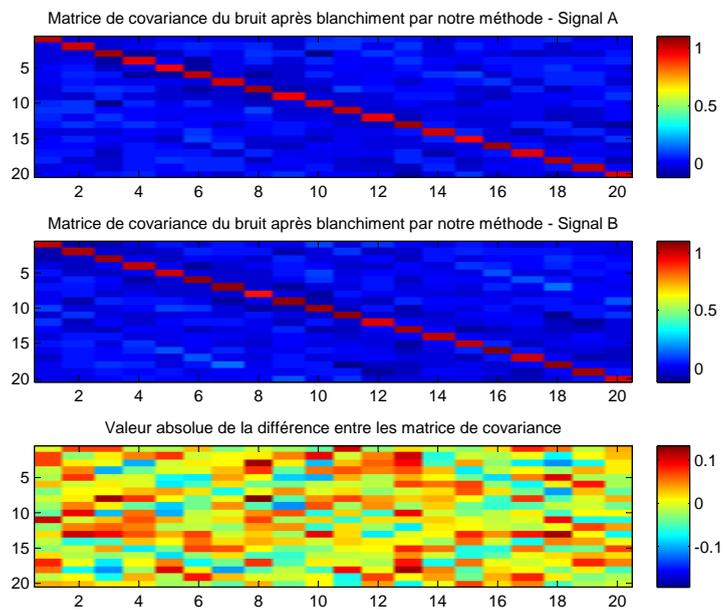


FIGURE C.11 – Illustration du blanchiment du bruit par notre méthode

Méthodes d'inversion du système d'équations non-linéaires

D.1 Contexte

Nous allons étudier désormais le cas de l'estimation de la masse et de la position d'une molécule dans un contexte particulier, celui où les chutes de fréquences ont été au préalable détectées et quantifiées. Nous noterons ces chutes de fréquence $\Delta f_{k,m,z}$.

D.2 Méthode de recherche d'une solution aux équations non-linéaires

La première méthode est celle publiée par Dohn *et al.*. Cette méthode consiste à écrire les équations (2.4) :

$$\Delta f_{k,m,z} = -\alpha_k \phi_k(z) m \quad (\text{D.1})$$

L'idée de l'article de Dohn *et al.* est alors de trouver un couple $\{m, z\}$ qui satisfasse à ces équations. La méthode de résolution proposée dans l'article est une méthode graphique.

L'article de Dohn *et al.* prend l'exemple d'une configuration *cantilever*, qui nécessite quatre harmoniques afin de garantir que le problème présente une solution unique.

Dans le cas *double-clamped*, la situation est plus simple. En effet, deux harmoniques suffisent. Pour démontrer cela, ré-écrivons l'équation (D.1) :

$$\begin{aligned} \Delta f_{1,m,z} &= -\alpha_1 \phi_1(z) m \\ \Delta f_{2,m,z} &= -\alpha_2 \phi_2(z) m \end{aligned} \quad (\text{D.2})$$

En faisant le rapport des deux quantités, étant donné que $\phi_1(z)$ ne s'annule pas sur $z \in]0, 0.5]$, et que $\frac{\phi_2(z)}{\phi_1(z)}$ admet une limite en $z = 0$, nous proposons d'étudier le rapport de ces deux quantités sur $z \in [0, 0.5]$.

$$\frac{\Delta f_{2,m,z}}{\Delta f_{1,m,z}} = \frac{\alpha_2 \phi_2(z)}{\alpha_1 \phi_1(z)} \quad (\text{D.3})$$

Traçons la quantité $\frac{\phi_2(z)}{\phi_1(z)}$ sur la figure D.1.

Nous voyons que la fonction qui à $z \in [0, 0.5]$ associe $\frac{\phi_2(z)}{\phi_1(z)} \in \mathbb{R}^+$ est bijective. Ainsi, observant le rapport des chutes de fréquence, il est possible de retrouver la position z indépendamment de m . Une fois z estimée, il est facile de retrouver m avec l'une ou l'autre de ces équations.

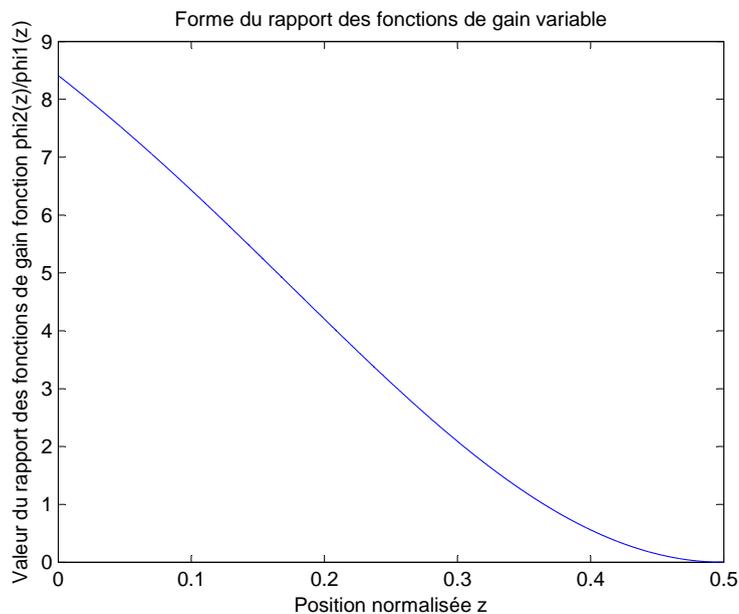


FIGURE D.1 – Illustration du rapport des fonctions de gain variables en fonction de la position

D.3 Méthode de recherche d'une solution aux équations non-linéaires

La seconde méthode que nous avons référencée est celle publiée par Hanay *et al.*, une équipe de Caltech.

Cette méthode consiste à voir les chutes de fréquence comme des variables aléatoires. L'article

considère la loi de probabilité jointe des paramètres $\Delta f_{1,m,z}$ et $\Delta f_{2,m,z}$ (de l'article en question, nous reprenons la notation *JPDF*, pour *Joint Probability Distribution Function*, ou loi de probabilité jointe) :

$$JPDF(\Delta f_{1,m,z}, \Delta f_{2,m,z}) = \frac{1}{2 \pi \sigma_1 \sigma_2 \sqrt{(1 - \rho^2)}} \exp \frac{-w}{2 (1 - \rho^2)} \quad (D.4)$$

où σ_1 est l'écart-type du bruit sur le mode 1, σ_2 l'écart-type du bruit sur le mode 2, ρ la corrélation du bruit d'un mode à l'autre et w vaut :

$$w = \frac{(\Delta f_{1,m,z} - \mu_1)^2}{\sigma_1^2} + \frac{(\Delta f_{2,m,z} - \mu_2)^2}{\sigma_2^2} - \frac{2 \rho (\Delta f_{1,m,z} - \mu_1) (\Delta f_{2,m,z} - \mu_2)}{\sigma_1 \sigma_2} \quad (D.5)$$

μ_1 et μ_2 sont les moyennes du processus de bruit, que les auteurs qualifient de négligeables (« *negligible* »). Afin d'estimer m et z , l'idée est d'opérer un changement de variable J en considérant le jacobien de la transformation. La matrice de transformation vaut :

$$\begin{pmatrix} \frac{\partial \Delta f_{1,m,z}}{\partial m} & \frac{\partial \Delta f_{1,m,z}}{\partial z} \\ \frac{\partial \Delta f_{2,m,z}}{\partial m} & \frac{\partial \Delta f_{2,m,z}}{\partial z} \end{pmatrix} \quad (D.6)$$

Dans la configuration *cantilever*, ce jacobien est positif et non nul. Par conséquent, les auteurs écrivent :

$$JPDF(m, z) = |J| \frac{1}{2 \pi \sigma_1 \sigma_2 \sqrt{(1 - \rho^2)}} \exp \frac{-W}{2 (1 - \rho^2)} \quad (D.7)$$

Dans cette équation, $|J|$ est le déterminant de la matrice jacobienne, et le terme W , quant à lui, vaut :

$$W = \frac{(\frac{m \phi_1(z)}{\alpha_1} + \Delta f_{1,m,z})^2}{\sigma_1^2} + \frac{(\frac{m \phi_2(z)}{\alpha_2} + \Delta f_{2,m,z})^2}{\sigma_2^2} - \frac{2 \rho (\frac{m \phi_1(z)}{\alpha_1} + \Delta f_{1,m,z})(\frac{m \phi_2(z)}{\alpha_2} + \Delta f_{2,m,z})}{\sigma_1 \sigma_2} \quad (D.8)$$

Dans l'article, cette densité de probabilité jointe est ensuite estimée sur une grille dans l'espace des paramètres m, z . S'en déduisent par sommation les lois *a posteriori* marginales de chacun des paramètres.

Détails des calculs

E.1 Démonstration de la relation (4.42)

Nous avons vu précédemment un exemple d'utilisation du cadre bayésien hiérarchique au travers de la figure 4.3. Nous en avons déduit la relation (4.42). Démontrons-là.

Par indépendance conditionnelle :

$$p(x_1|y, x_2, a_1, a_2, a_3) = p(x_1|y, a_1, a_2) \quad (\text{E.1})$$

Puis, selon le théorème de la probabilité conditionnelle :

$$\begin{aligned} p(x_1|y, a_1, a_2) &= \frac{p(y, x_1, a_1, a_2)}{p(y, a_1, a_2)} \\ &\propto p(y, x_1, a_1, a_2) \end{aligned} \quad (\text{E.2})$$

Toujours selon le même théorème :

$$p(y, x_1, a_1, a_2) = p(y|x_1, a_1, a_2) p(x_1, a_1, a_2) \quad (\text{E.3})$$

Et également :

$$\begin{aligned} p(x_1, a_1, a_2) &= p(x_1|a_1, a_2) p(a_1, a_2) \\ &\propto p(x_1|a_1, a_2) \end{aligned} \quad (\text{E.4})$$

De plus, par indépendance conditionnelle :

$$p(y|x_1, a_1, a_2) = p(y|x_1) \quad (\text{E.5})$$

Nous pouvons donc écrire :

$$p(x_1|y, x_2, a_1, a_2, a_3) \propto p(y|x_1) p(x_1|a_1, a_2) \quad (\text{E.6})$$

CQFD.

E.2 Calcul de la probabilité d'acceptation dans le cadre d'un échantillonneur de Gibbs

Dans la section 4.3.3.1, nous avons prétendu que la probabilité d'acceptation d'une étape de Gibbs valait systématiquement 1. Démontrons cela.

Si on note $\mathbf{f}_{i_0}^p$ la proposition pour le paramètre échantillonné, $\mathbf{f}_{i_0}^{(j-1)}$ la valeur courante du paramètre échantillonné et \mathbf{f}_{-i_0} les valeurs des autres paramètres, alors :

$$Q_{\text{Gibbs}}(\mathbf{f}_{i_0}^p) = \pi(\mathbf{f}_{i_0}^p | \mathbf{f}_{-i_0}, \mathbf{g}) \quad (\text{E.7})$$

Ici, il vient :

$$\begin{aligned} \alpha &= \min \left(1, \frac{\pi(\mathbf{f}_{i_0}^p, \mathbf{f}_{-i_0} | \mathbf{g})}{\pi(\mathbf{f}_{i_0}^{(j-1)}, \mathbf{f}_{-i_0} | \mathbf{g})} \frac{\pi(\mathbf{f}_{i_0}^{(j-1)} | \mathbf{f}_{-i_0}, \mathbf{g})}{\pi(\mathbf{f}_{i_0}^p | \mathbf{f}_{-i_0}, \mathbf{g})} \right) \\ &= \min \left(1, \frac{\pi(\mathbf{f}_{i_0}^p, \mathbf{f}_{-i_0} | \mathbf{g})}{\pi(\mathbf{f}_{i_0}^{(j-1)}, \mathbf{f}_{-i_0} | \mathbf{g})} \frac{\frac{\pi(\mathbf{f}_{i_0}^{(j-1)}, \mathbf{f}_{-i_0} | \mathbf{g})}{p(\mathbf{f}_{-i_0} | \mathbf{g})}}{\frac{\pi(\mathbf{f}_{i_0}^p, \mathbf{f}_{-i_0} | \mathbf{g})}{p(\mathbf{f}_{-i_0} | \mathbf{g})}} \right) \\ &= 1 \end{aligned} \quad (\text{E.8})$$

CQFD.

E.3 Marginalisation de l'inverse variance d'une loi normale

$$\begin{aligned} \int_0^{+\infty} \mathcal{N}\left(f|0, \frac{1}{\tau}\right) \mathcal{G}am\left(\tau \left| \frac{s}{2}, \frac{2}{s}\right.\right) d\tau &\propto \int_0^{+\infty} \tau^{\frac{1}{2}} \exp\left(-\frac{\tau f^2}{2}\right) \tau^{\frac{s}{2}-1} \exp\left(\frac{-\tau}{\frac{2}{s}}\right) d\tau \\ &\propto \int_0^{+\infty} \tau^{\frac{s-1}{2}} \exp\left(-\tau \left(\frac{f^2+s}{2}\right)\right) d\tau \end{aligned} \quad (\text{E.9})$$

Alors, en prenant le changement de variable $\tau' = \tau \left(\frac{f^2+s}{2}\right)$, et en notant $\Gamma(\cdot)$ la fonction Gamma :

$$\begin{aligned}
\int_0^{+\infty} \mathcal{N}\left(f|0, \frac{1}{\tau}\right) \mathcal{G}am\left(\tau|\frac{s}{2}, \frac{2}{s}\right) d\tau &\propto \left(\frac{f^2+s}{2}\right)^{-1} \left(\frac{f^2+s}{2}\right)^{\frac{1-s}{2}} \int_0^{+\infty} \tau'^{\frac{s-1}{2}} \exp(-\tau') d\tau' \\
&\propto (s+f^2)^{-\frac{s+1}{2}} \Gamma\left(\frac{s+1}{2}\right) \\
&\propto \left(1+\frac{f^2}{s}\right)^{-\frac{s+1}{2}}
\end{aligned} \tag{E.10}$$

CQFD.

E.4 Calculs d'acceptation des algorithmes RJMCMC

Détaillons l'échantillonneur RJMCMC évoqué dans les section 5.3.3.1 (mono-mode), 6.3.3.1 (multi-mode) et 7.4.2.5 (reconstruction de spectre).

E.4.1 Processus de naissance et de mort

L'idée ici est d'ajouter ou de retirer un événement (*Birth and Death MCMC*).

Supposons être en présence de $N^{(j-1)}$ événements. Nous choisissons de manière équiprobable de supprimer ou d'ajouter un événement. Si nous décidons un ajout, alors la position de la nouvelle entrée est choisie uniformément parmi les $N^{(j-1)} + 1$ possibilités. Si une suppression est décidée, l'entrée à supprimer est choisie uniformément parmi les $N^{(j-1)}$ entrées.

E.4.2 Ajout d'une entrée dans la liste

Lors d'un ajout, le nouvel élément de la liste $\{t^p, f^p\}$ ou $\{t^p, m^p, z^p\}$ est généré à partir des lois *a priori* soit :

- dans le cas mono-mode :
 - pour t^p : $\mathcal{U}_{\{1,T\}}(t^p)$ ou $\mathcal{U}_{[0,T]}(t^p)$
 - pour f^p : $\mathcal{G}am(f^p|k_m, \theta_m)$
- dans le cas multi-mode :
 - pour t^p : $\mathcal{U}_{\{1,T\}}(t^p)$ ou $\mathcal{U}_{[0,T]}(t^p)$
 - pour m^p : $\mathcal{G}am(m^p|k_m, \theta_m)$
 - pour z^p : $\mathcal{U}_{[0,0.5]}(z^p)$
- dans le cas de reconstruction de spectre :
 - pour t^p : $\mathcal{U}_{\{1,T\}}(t^p)$ ou $\mathcal{U}_{[0,T]}(t^p)$
 - pour m^p : $\sum_{l=1}^L \frac{1}{L} \delta(m^p - Sp_l)$
 - pour z^p : $\mathcal{U}_{[0,0.5]}(z^p)$

E.4.3 Calcul de la probabilité d'acceptation

Rappelons la valeur de la probabilité d'acceptation α RJMCMC :

$$\alpha = \min \left(1, \frac{\pi(\mathcal{M}_n, \boldsymbol{\theta}_n)}{\pi(\mathcal{M}_m, \boldsymbol{\theta}_m)} \frac{Q_{n,m}(v_n^p)}{Q_{m,n}(u_m^p)} \frac{\pi_{n,m}}{\pi_{m,n}} \left| \frac{\partial \mathcal{T}(\theta_n, v_n^p)}{\partial(\theta_n, v_n^p)} \right| \right) \quad (\text{E.11})$$

Dans notre cas $\frac{\pi_{n,m}}{\pi_{m,n}} = 1$. En effet, si un événement a été supprimé sur une ligne (avec un probabilité $\frac{1}{N^{(j-1)}}$), nous avons la probabilité $\frac{1}{N^{(j-1)} - 1 + 1}$ de voir une entrée ré-apparaître sur cette ligne.

Aussi, vu qu'il ne s'agit ici que d'une insertion / suppression de paramètres $\left| \frac{\partial \mathcal{T}(\theta_n, v_n^p)}{\partial(\theta_n, v_n^p)} \right| = 1$.

Donc :

$$\alpha = \min \left(1, \frac{\pi(\mathcal{M}_n, \boldsymbol{\theta}_n)}{\pi(\mathcal{M}_m, \boldsymbol{\theta}_m)} \frac{Q_{n,m}(v_n^p)}{Q_{m,n}(u_m^p)} \right) \quad (\text{E.12})$$

E.4.3.1 Dans le cas d'un ajout

Si on note $\boldsymbol{\theta}^p$ les paramètres ajoutés : $Q_{m,n}(u_m^p) = p(\boldsymbol{\theta}^p)$, les lois *a priori*. Aussi, $Q_{n,m}(v_n^p) = 1$.

Si on note $\boldsymbol{\theta}_{-p}$ les autres paramètres (ceux qui n'ont pas été modifiés), alors, en notant $\mathcal{L}(\boldsymbol{\theta})$ la vraisemblance :

$$\frac{\pi(\mathcal{M}_n, \boldsymbol{\theta}_n)}{\pi(\mathcal{M}_m, \boldsymbol{\theta}_m)} = \frac{p(N^{(j-1)} + 1)}{p(N^{(j-1)})} p(\boldsymbol{\theta}^p) \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \quad (\text{E.13})$$

Ainsi :

$$\begin{aligned} \alpha &= \min \left(1, \frac{p(N^{(j-1)} + 1)}{p(N^{(j-1)})} \frac{p(\boldsymbol{\theta}^p)}{p(\boldsymbol{\theta}^p)} \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \right) \\ &= \min \left(1, \frac{\mathcal{B}(N^{(j-1)} + 1 | \pi^{(j-1)}, T)}{\mathcal{B}(N^{(j-1)} | \pi^{(j-1)}, T)} \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \right) \\ &= \min \left(1, \frac{(T - N^{(j-1)}) \pi^{(j-1)}}{(1 + N^{(j-1)}) (1 - \pi^{(j-1)})} \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \right) \end{aligned} \quad (\text{E.14})$$

E.4.3.2 Dans le cas d'une suppression

Ici, le calcul est le même et il vient :

$$\begin{aligned}\alpha &= \min \left(1, \frac{\mathcal{B}(N^{(j-1)} - 1 | \pi^{(j-1)}, T)}{\mathcal{B}(N^{(j-1)} | \pi^{(j-1)}, T)} \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \right) \\ &= \min \left(1, \frac{(N^{(j-1)})(1 - \pi^{(j-1)})}{(T - N^{(j-1)} + 1) \pi^{(j-1)}} \frac{\mathcal{L}(\boldsymbol{\theta}^p, \boldsymbol{\theta}_{-p})}{\mathcal{L}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}_{-p})} \right)\end{aligned}\tag{E.15}$$

Bibliographie personnelle

F.1 Travail portant sur la déconvolution mono-mode

Communication GRETSI

R. PERENON, A. MOHAMMAD-DJAFARI, L. DURAFFOURG, P. GRANGEAT. Quantification moléculaire par spectrométrie de masse à base de NEMS : modélisation et inversion du problème. « 23ème Colloque GRETSI sur le Traitement du Signal et des Images » - GRETSI 2011, Bordeaux, France, 2011.

Présentation GdR Isis

R. PERENON. Méthodes MCMC dédiées à la spectrométrie de masse à base de NEMS. Réunion du GdR ISIS thème A - Méthodes de Monte Carlo pour les problèmes inverses bayésiens en traitement des signaux et des images, 2011.

Brevet

R. PERENON, A. MOHAMMAD-DJAFARI, P. GRANGEAT. Procédé et dispositif d'estimation d'un paramètre de masse moléculaire dans un échantillon. Demande de brevet en France - Numéro de dépôt 11 57857 Extension aux USA sous la dénomination *Method and apparatus for estimating a molecular mass parameter in a sample*.

Communication MaxEnt

R. PERENON, A. MOHAMMAD-DJAFARI, E. SAGE, L. DURAFFOURG, S. HENTZ, A. BRENAC, R. MOREL, P. GRANGEAT. MCMC-Based Bayesian estimation algorithm dedicated to NEMS Mass Spectrometry. « 32nd International Workshop on Bayesian Inference and

Maximum Entropy Methods in Science Engineering » – MAXENT 2012 - Garching bei Munchen, Allemagne, 2012 (article à paraître).

F.2 Travail portant sur la déconvolution multi-mode

Communication EUSIPCO

R. PERENON, E. SAGE, A. MOHAMMAD-DJAFARI, L. DURAFFOURG, S. HENTZ, A. BRENAC, R. MOREL, P. GRANGEAT. Bayesian Inversion of Multi-Mode NEMS Mass Spectrometry Signal. « 21st European Signal Processing Conference 2013 » – EUSIPCO 2013 – Marrakech, Maroc, 2013 (article accepté).

F.3 Travail portant sur la reconstruction de spectre

Un projet d'article est en cours de rédaction

Septième partie

Bibliographies

Bibliographie pour l'introduction

- [ABE⁺08] B. L. Ackermann, M. J. Berna, J. A. Eckstein, L. W. Ott, and A. K. Chaudhary. *Current applications of liquid chromatography/mass spectrometry in pharmaceutical discovery after a decade of innovation*, volume 1 of *Annual Review of Analytical Chemistry*. 2008.
- [AM03] R. Aebersold and M. Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928) :198–207, 2003.
- [CBB12] S. A. Cornelissen, T. G. Bifano, and P. A. Bierden. Mems deformable mirror actuators with enhanced reliability. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 8253, 2012.
- [CYP⁺12] M. L. Chan, B. Yoxall, H. Park, Z. Kang, I. Izyumin, J. Chou, M. M. Megens, M. C. Wu, B. E. Boser, and D. A. Horsley. Design and characterization of mems micromotor supported on low friction liquid bearing. *Sensors and Actuators, A : Physical*, 177 :1–9, 2012.
- [DA06] B. Domon and R. Aebersold. Mass spectrometry and protein analysis. *Science*, 312(5771) :212–217, 2006.
- [DLBK07] V. Dauksaite, M. Lorentzen, F. Besenbacher, and J. Kjems. Antibody-based protein detection using piezoresistive cantilever arrays. *Nanotechnology*, 18(12), 2007.
- [EHR04] K. L. Ekinici, X. M. H. Huang, and M. L. Roukes. Ultrasensitive nanoelectromechanical mass detection. *Applied Physics Letters*, 84(22) :4469–4471, 2004.
- [HT11] S. Hanash and A. Taguchi. Application of proteomics to cancer early detection. *Cancer Journal*, 17(6) :423–428, 2011.
- [LLP⁺12] X. Li, Q. Liu, S. Pang, K. Xu, H. Tang, and C. Sun. High-temperature piezoresistive pressure sensor based on implantation of oxygen into silicon wafer. *Sensors and Actuators, A : Physical*, 179 :277–282, 2012.
- [LSP⁺10] J. Lee, W. Shen, K. Payer, T. P. Burg, and S. R. Manalis. Toward attogram mass measurements in solution with suspended nanochannel resonators. *Nano Letters*, 10(7) :2537–2542, 2010.
- [NHH⁺09] A. K. Naik, M. S. Hanay, W. K. Hiebert, X. L. Feng, and M. L. Roukes. Towards single-molecule nanomechanical mass spectrometry. *Nature Nanotechnology*, 4(7) :445–450, 2009.
- [PAH⁺02] E. F. Petricoin, A. M. Ardekani, B. A. Hitt, P. J. Levine, V. A. Fusaro, S. M. Steinberg, G. B. Mills, C. Simone, D. A. Fishman, E. C. Kohn, and L. A. Liotta. Use

- of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 359(9306) :572–577, 2002.
- [PZK⁺02] E. F. Petricoin, K. C. Zoon, E. C. Kohn, J. C. Barrett, and L. A. Liotta. Clinical proteomics : Translating benchside promise into bedside reality. *Nature Reviews Drug Discovery*, 1(9) :683–695, 2002.
- [QL12] X. J. Qin and B. X. Ling. Proteomic studies in breast cancer (review). *Oncology Letters*, 3(4) :735–743, 2012.
- [TCL12] A. Tocchio, A. Caspani, and G. Langfelder. Mechanical and electronic amplitude-limiting techniques in a mems resonant accelerometer. *IEEE Sensors Journal*, 12(6) :1719–1725, 2012.
- [WBMT00] M.-R.-H. White, A.-D. Bates, A.-G. McLennan, and P.-C. Turner. *Biologie moléculaire*. Broché, 2000. ISBN 2-911808-10-X.
- [WLP03] J. D. Wulfkühle, L. A. Liotta, and E. F. Petricoin. Proteomic applications for the early detection of cancer. *Nature Reviews Cancer*, 3(4) :267–275, 2003.
- [YCF⁺06] Y. T. Yang, C. Callegari, X. L. Feng, K. L. Ekinici, and M. L. Roukes. Zeptogram-scale nanomechanical mass sensing. *Nano Letters*, 6(4) :583–586, 2006.

Bibliographie pour le chapitre 1

- [AQD⁺02] B. . Adam, Y. Qu, J. W. Davis, M. D. Ward, M. A. Clements, L. H. Cazares, O. J. Semmes, P. F. Schellhammer, Y. Yasui, Z. Feng, and G. L. Wright Jr. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer research*, 62(13) :3609–3614, 2002.
- [BMGD09] V. Brun, C. Masselon, J. Garin, and A. Dupuis. Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, 72(5) :740–749, 2009.
- [BSS⁺07] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, and B. Kuster. Quantitative mass spectrometry in proteomics : A critical review. *Analytical and Bioanalytical Chemistry*, 389(4) :1017–1031, 2007.
- [EP71] E. Engvall and P. Perlmann. Enzyme-linked immunosorbent assay (elisa) quantitative assay of immunoglobulin g. *Immunochemistry*, 8(9) :871–874, 1971.
- [GRS⁺03] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, and S. P. Gygi. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem ms. *Proceedings of the National Academy of Sciences of the United States of America*, 100(12) :6940–6945, 2003.
- [HPIM⁺03] S. R. Hingorani, E. F. Petricoin III, A. Maitra, V. Rajapakse, C. King, M. A. Jacobetz, S. Ross, T. P. Conrads, T. D. Veenstra, B. A. Hitt, Y. Kawaguchi, D. Johann, L. A. Liotta, H. C. Crawford, M. E. Putt, T. Jacks, C. V. E. Wright, R. H. Hruban, A. M. Lowy, and D. A. Tuveson. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell*, 4(6) :437–450, 2003.
- [KTM98] M. König, H. . Thiel, and G. Meyers. Detection of viral proteins after infection of cultured hepatocytes with rabbit hemorrhagic disease virus. *Journal of virology*, 72(5) :4492–4497, 1998.
- [ML99] E.N. Marieb and R. Lachaîne. *Anatomie et physiologie humaines*. de Boeck Université, 1999.
- [PIAH⁺02] E. F. Petricoin III, A. M. Ardekani, B. A. Hitt, P. J. Levine, V. A. Fusaro, S. M. Steinberg, G. B. Mills, C. Simone, D. A. Fishman, E. C. Kohn, and L. A. Liotta. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 359(9306) :572–577, 2002.

- [SGG11] P. Szacherski, J. F. Giovannelli, and P. Grangeat. Joint bayesian hierarchical inversion-classification and application in proteomics. In *IEEE Workshop on Statistical Signal Processing Proceedings*, pages 121–124, 2011.
- [SGP⁺07] G. Strubel, J. F. Giovannelli, C. Paulus, L. Gerfault, and P. Grangeat. Bayesian estimation for molecular profile reconstruction in proteomics based on liquid chromatography and mass spectrometry. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pages 5979–5982, 2007.
- [Str08] G. Strubel. *Reconstruction de profils moléculaires : modélisation et inversion d'une chaîne de mesure protéomique*. PhD thesis, Institut Polytechnique de Grenoble, 2008.
- [SVJ87] H. Schagger and G. Von Jagow. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kda. *Analytical Biochemistry*, 166(2) :368–379, 1987.
- [Sza13] P. Szacherski. *Reconstruction de profils protéiques pour la recherche de biomarqueurs*. PhD thesis, Université de Bordeaux 1, 2013.
- [WBMT00] M.-R.-H. White, A.-D. Bates, A.-G. McLennan, and P.-C. Turner. *Biologie moléculaire*. Broché, 2000. ISBN 2-911808-10-X.
- [WWZ⁺06] G. Wang, W. W. Wu, W. Zeng, C. . Chou, and R. . Shen. Label-free protein quantification using lc-coupled ion trap or ft mass spectrometry : Reproducibility, linearity, and application with complex proteomes. *Journal of Proteome Research*, 5(5) :1214–1223, 2006.

Bibliographie pour le chapitre 2

- [BBB⁺90] J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes. Light-emitting diodes based on conjugated polymers. *Nature*, 347(6293) :539–541, 1990.
- [CBB12] S. A. Cornelissen, T. G. Bifano, and P. A. Bierden. Mems deformable mirror actuators with enhanced reliability. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 8253, 2012.
- [CDL⁺09] E. Colinet, L. Duraffourg, S. Labarthe, S. Hentz, P. Robert, and P. Andreucci. Self-oscillation conditions of a resonant nanoelectromechanical mass sensor. *Journal of Applied Physics*, 105(12), 2009.
- [CYP⁺12] M. L. Chan, B. Yoxall, H. Park, Z. Kang, I. Izyumin, J. Chou, M. M. Megens, M. C. Wu, B. E. Boser, and D. A. Horsley. Design and characterization of mems micromotor supported on low friction liquid bearing. *Sensors and Actuators, A : Physical*, 177 :1–9, 2012.
- [DSBH07] S. Dohn, W. Svendsen, A. Boisen, and O. Hansen. Mass and position determination of attached particles on cantilever based mass sensors. *Review of Scientific Instruments*, 78(10), 2007.
- [Dur08] L. Duraffourg. Nano systèmes électromécaniques sur silicium (silicon-nems). Mémoire pour l’obtention de l’Habilitation à Diriger des Recherches, 2008.
- [For04] S. R. Forrest. The path to ubiquitous and low-cost organic electronic appliances on plastic. *Nature*, 428(6986) :911–918, 2004.
- [HKN⁺12] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9) :602–608, 2012.
- [Kha09] C. Kharrat. *Application des techniques de contrôle sur les réseaux de micro et nanostructures*. PhD thesis, Institut Polytechnique de Grenoble, 2009.
- [KMS00] A. I. Kingon, J. . Maria, and S. K. Streiffer. Alternative dielectrics to silicon dioxide for memory and logic devices. *Nature*, 406(6799) :1032–1038, 2000.
- [LHHC10] H. . Liao, K. . Huang, E. . Hwu, and C. . Chang. Resonance-enhanced micromechanical cantilever for mass sensing. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, pages 437–441, 2010.

- [LLP⁺12] X. Li, Q. Liu, S. Pang, K. Xu, H. Tang, and C. Sun. High-temperature piezoresistive pressure sensor based on implantation of oxygen into silicon wafer. *Sensors and Actuators, A : Physical*, 179 :277–282, 2012.
- [Mil09] E. Milé. *Systèmes électromécaniques nanométriques base de nanofils de silicium et nanotubes de carbone*. PhD thesis, École Polytechnique, 2009.
- [MPL10] M. Melli, A. Pozzato, and M. Lazzarino. Inverted tapered pillars for mass sensing. *Microelectronic Engineering*, 87(5-8) :730–733, 2010.
- [MVD⁺11] L. Malier, H. Vatouyas, J.-B. David, J. Marty, O. Peyret, D. Louis, and F. Vacherand. Rapport d’activité 2011. Technical report, CEA-Leti, 2011.
- [NHH⁺09] A. K. Naik, M. S. Hanay, W. K. Hiebert, X. L. Feng, and M. L. Roukes. Towards single-molecule nanomechanical mass spectrometry. *Nature Nanotechnology*, 4(7) :445–450, 2009.
- [RRLB04] C. Reese, M. Roberts, M. . Ling, and Z. Bao. Organic thin film transistors. *Materials Today*, 7(8) :20–27, 2004.
- [SDB10] S. Schmid, S. Dohn, and A. Boisen. Real-time particle mass spectrometry based on resonant micro strings. *Sensors*, 10(9) :8192–8100, 2010.
- [TCL12] A. Tocchio, A. Caspani, and G. Langfelder. Mechanical and electronic amplitude-limiting techniques in a mems resonant accelerometer. *IEEE Sensors Journal*, 12(6) :1719–1725, 2012.
- [VDZBA⁺10] A. M. Van Der Zande, R. A. Barton, J. S. Alden, C. S. Ruiz-Vargas, W. S. Whitney, P. H. Q. Pham, J. Park, J. M. Parpia, H. G. Craighead, and P. L. McEuen. Large-scale arrays of single-layer graphene resonators. *Nano Letters*, 10(12) :4869–4873, 2010.
- [YCF⁺06] Y. T. Yang, C. Callegari, X. L. Feng, K. L. Ekinici, and M. L. Roukes. Zeptogram-scale nanomechanical mass sensing. *Nano Letters*, 6(4) :583–586, 2006.

Bibliographie pour le chapitre 3

- [Bar81] Jeffrey A. Barnett. Computational methods for a mathematical theory of evidence. volume 2, pages 868–875, 1981.
- [BHL10] P. Boisseau, P. Houdy, and M. Lahmani. *Nanoscience : Nanobiotechnology and Nanobiology*. SpringerLink : Springer e-Books. Springer-Verlag Berlin Heidelberg, 2010.
- [CFM05] A. Ciabattini, C. G. Fermüller, and G. Metcalfe. *Uniform rules and dialogue games for fuzzy logics*, volume 3452 LNAI of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2005.
- [Cox61] R.T. Cox. *Algebra of Probable Inference*. Algebra of Probable Inference. Johns Hopkins University Press, 1961.
- [CR02] A. N. Cleland and M. L. Roukes. Noise processes in nanomechanical resonators. *Journal of Applied Physics*, 92(5) :2758, 2002.
- [Dal91] A.I. Dale. *A history of inverse probability : from Thomas Bayes to Karl Pearson*. Studies in the history of mathematics and physical sciences. Springer-Verlag, 1991.
- [DBD⁺11] S. Descamps, A. Béchet, X. Descombes, A. Arnaud, and J. Zerubia. An automatic counter for aerial images of aggregations of large birds. *Bird Study*, 58(3) :302–308, 2011.
- [DDBZ08] S. Descamps, X. Descombes, A. Béchet, and J. Zerubia. Automatic flamingo detection using a multiple birth and death process. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 1113–1116, 2008.
- [DT09] M. J. Dupré and F. J. Tiplery. New axioms for rigorous bayesian probability. *Bayesian Analysis*, 4(3) :599–606, 2009.
- [HKN⁺12] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9) :602–608, 2012.
- [Ish83] M. Ishizuka. Inference methods based on extended dempster & shafer’s theory for problems with uncertainty/fuzziness. *New Generation Computing*, 1(2) :159–168, 1983.
- [LDM10] R. Levy, A. Dupret, and H. Mathias. Phase noise analysis of the frequency tracking oscillator. In *2010 IEEE International Frequency Control Symposium, FCS 2010*, pages 624–628, 2010.

- [Maz05] V. Mazet. *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. PhD thesis, Université Henri Poincaré, Nancy 1, 2005.
- [PBP07] I. Parent, J. Bernier, and R. Parent. *Le raisonnement bayésien : Modélisation et inférence*. Statistique et probabilités appliquées. Springer, 2007.
- [PGBG09] C. Paulus, L. Gerfault, V. Brun, and P. Grangeat. Inversion of a nano-liquid chromatography dynamical model for quantitative proteomics. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, volume 15, 2009.
- [Pil10] J. Pillow. Maximum entropy distributions & models. Cours à L'Université du Texas, Austin – Methods in Computational Neuroscience. Consulté le 13 Juin 2013, 2010.
- [RGBG00] E. Rubiola, J. Gros Lambert, M. Brunet, and V. Giordano. Flicker noise measurement of hf quartz resonators. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 47(2) :361–368, 2000.
- [Rob05] C.P. Robert. *Le Choix Bayésien : Principes et pratique*. Statistique et probabilités appliquées. Springer, 2005.
- [Sap06] G. Saporta. *Probabilités, analyse des données et statistique*. Editions Technip, 2006.
- [SB08] J. H. Seo and O. Brand. High q-factor in-plane-mode resonant microsensor platform for gaseous/ liquid environment. *Journal of Microelectromechanical Systems*, 17(2) :483–493, 2008.
- [SK94] P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66(2) :191–234, 1994.
- [Sza13] P. Szacherski. *Reconstruction de profils protéiques pour la recherche de biomarqueurs*. PhD thesis, Université de Bordeaux 1, 2013.
- [TBBD10] O. Tournaire, M. Brédif, D. Boldo, and M. Durupt. An efficient stochastic approach for building footprint extraction from digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(4) :317–327, 2010.
- [Teo99] D. Teodorovic. Fuzzy logic systems for transportation engineering : The state of the art. *Transportation Research Part A : Policy and Practice*, 33(5) :337–364, 1999.
- [Wed] N.S. Wedd. Peptide mass spec (consultée le 16 juin 2013).
- [Wit01] T. J. Witt. Using the allan variance and power spectral density to characterize dc nanovoltmeters. *IEEE Transactions on Instrumentation and Measurement*, 50(2) :445–448, 2001.

Bibliographie pour le chapitre 4

- [BA04] K. P. Burnham and D. R. Anderson. Multimodel inference : Understanding aic and bic in model selection. *Sociological Methods and Research*, 33(2) :261–304, 2004.
- [Dem87] G. Demoment. Déconvolution des signaux. Cours à l’Ecole Supérieure d’Electricité, Gif-sur-Yvette, 1987.
- [F06] O. Féron. *Champs de Markov cachés pour les problèmes inverses. Application à la fusion de données et à la reconstruction d’images en tomographie micro-ondes*. PhD thesis, Université Paris-Sud 11, 2006.
- [Gra02] P. Grangeat. *La tomographie : fondements mathématiques, imagerie microscopique et imagerie industrielle*. IC2 : Série Traitement du signal et de l’image. Hermes Science Publications, 2002.
- [Had02] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13 :49–52, 1902.
- [Has70] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.
- [Jef61] H. Jeffreys. *Theory of Probability ... Third Edition*. International series of monographs on physics. Oxford, 1961.
- [KS90] M. Kobayashi and S. Sakata. Mallows’ cp criterion and unbiasedness of model selection. *Journal of Econometrics*, 45(3) :385–395, 1990.
- [Mal00] S. Mallat. *Une exploration des signaux en ondelettes*. Éditions de l’Ecole Polytechnique, 2000.
- [MD01] A. Mohammad-Djafari. Problèmes inverses. Cours du DEA Automatique et traitement de Signal – Ecole Supérieure d’Electricité – Laboratoire des Signaux et Systèmes, 2001.
- [MRR⁺53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of chemical physics*, 21(6) :1087–1092, 1953.
- [Mur12] K.P. Murphy. *Machine Learning : A Probabilistic Perspective*. Adaptive computation and machine learning series. Mit Press, 2012.
- [OSBE13] N. Oppermann, M. Selig, M. R. Bell, and T. A. Enßlin. Reconstruction of gaussian and log-normal fields with spectral smoothness. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(3), 2013.

- [PBG⁺07] C. Paulus, S. Bonnet, L. Gerfault, E. Mery, G. Strubel, F. Ricoul, and P. Grangeat. Chromatographic alignment combined with chemometrics profile reconstruction approaches applied to lc-ms data. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pages 5983–5986, 2007.
- [PBP07] I. Parent, J. Bernier, and R. Parent. *Le raisonnement bayésien : Modélisation et inférence*. Statistique et probabilités appliquées. Springer, 2007.
- [RC04] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 2004.
- [Rob05] C.P. Robert. *Le Choix Bayésien : Principes et pratique*. Statistique et probabilités appliquées. Springer, 2005.
- [Sap06] G. Saporta. *Probabilités, analyse des données et statistique*. Editions Technip, 2006.
- [Shi76] R. Shibata. Selection of the order of an autoregressive model by akaike’s information criterion. *Biometrika*, 63(1) :117–126, 1976.
- [ŠQ10] V. Šmídl and A. Quinn. *The Variational Bayes Method in Signal Processing*. Signals and Communication Technology. Springer, 2010.
- [Sza13] P. Szacherski. *Reconstruction de profils protéiques pour la recherche de biomarqueurs*. PhD thesis, Université de Bordeaux 1, 2013.
- [Tik63] A. Tikhonov. Regularization of incorrectly posed problems. *Soviet. Math. Dokl.*, 4 :1624–1627, 1963.

Bibliographie pour le chapitre 5

- [AO09] S. E. Adlouni and T. B. M. J. Ouarda. Joint bayesian model selection and parameter estimation of the generalized extreme value model with covariates using birth-death markov chain monte carlo. *Water Resources Research*, 45(6), 2009.
- [BN93] M. Basseville and I.V. Nikiforov. *Detection of abrupt changes : theory and application*. Prentice-Hall information and system sciences series. Prentice Hall, 1993.
- [bou]
- [Bou08] S. Bourguignon. *Analyse spectrale à haute résolution de signaux à échantillonnage irrégulier : application à l'astrophysique*. PhD thesis, Université Paul Sabatier Toulouse 3, 2008.
- [Cap06] C. Capilla. Application of the haar wavelet transform to detect microseismic signal arrivals. *Journal of Applied Geophysics*, 59(1) :36–46, 2006.
- [Car13] Bourguignon S. Idier J. ans Simon L. Carcreff, E. Resolution enhancement of ultrasonic signals by up-sampled sparse deconvolution. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013.
- [CCL96] Q. Cheng, R. Chen, and T. . Li. Simultaneous wavelet estimation and deconvolution of reflection seismic signals. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2) :377–384, 1996.
- [CMDPG13] N. Chu, A. Mohammad-Djafari, J. Picheral, and N. Gac. An efficient variational bayesian inference approach via student's-t priors for acoustic imaging in colored noises. In *Proceedings of Meetings on Acoustics*, volume 19, 2013.
- [CW09] J. Chen and Y. P. Wang. A statistical change point model approach for the detection of dna copy number variations in array cgh data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 6(4) :529–541, 2009.
- [DID97] Frederic Dublanquet, Jerome Idier, and Patrick Duwaut. Direction-of-arrival and frequency estimation using poisson-gaussian modeling. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, volume 5, pages 3501–3504, 1997.
- [DTD07] N. Dobigeon, J. . Tourneret, and M. Davy. Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a bayesian sampling approach. *IEEE Transactions on Signal Processing*, 55(4) :1251–1263, 2007.
- [EBG07] R. Escola, S. Bonnet, and R. Guillemaud. Détection de potentiels d'action neuroaux par ondelettes. 2007.

- [EHJ⁺04] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, H. Ishwaran, K. Knight, J. . Loubes, P. Massart, D. Madigan, G. Ridgeway, S. Rosset, J. I. Zhu, R. A. Stine, B. A. Turlach, S. Weisberg, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Annals of Statistics*, 32(2) :407–499, 2004.
- [Ge09] D. Ge. *Déconvolution impulsionnelle multi-source. Application aux signaux électromyographiques*. PhD thesis, Ecole Centrale de Nantes, 2009.
- [GJM10] S. Gibson, J. W. Judy, and D. Marković. Technology-aware algorithm design for neural spike detection, feature extraction, and dimensionality reduction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(5) :469–478, 2010.
- [GJM12] S. Gibson, J. W. Judy, and D. Marković. Spike sorting : The first step in decoding the brain : The first step in decoding the brain. *IEEE Signal Processing Magazine*, 29(1) :124–143, 2012.
- [Hö74] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement, volume=15, pages=417,*, 1974.
- [HKN⁺12] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9) :602–608, 2012.
- [Idi01] Jérôme Idier. *Approche bayésienne pour les problèmes inverses*. Traité IC2, Série traitement du signal et de l’image, Hermès, Paris, nov. 2001.
- [JDM00] A. K. Jain, R. P. W. Duin, and J. Mao. Statistical pattern recognition : A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1) :4–37, 2000.
- [Kay98] S.M. Kay. *Fundamentals of Statistical Signal Processing : Detection theory*. Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1998.
- [KM82] John J. Kormylo and Jerry M. Mendel. Maximum likelihood detection and estimation of bernoulli-gaussian processes. *IEEE Transactions on Information Theory*, IT-28(3) :482–488, 1982.
- [LYFLLC11] A. Lung-Yut-Fong, C. Lévy-Leduc, and O. Cappé. Robust changepoint detection based on multivariate rank statistics. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 3608–3611, 2011.
- [Mal00] S. Mallat. *Une exploration des signaux en ondelettes*. Éditions de l’Ecole Polytechnique, 2000.
- [Maz05] V. Mazet. *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. PhD thesis, Université Henri Poincaré, Nancy 1, 2005.
- [MBC04] V. Mazet, D. Brie, and C. Caironi. Sparse spike train deconvolution using the hunt filter and a thresholding method. *IEEE Signal Processing Letters*, 11(5) :486–489, 2004.
- [MD12a] A. Mohammad-Djafari. Bayesian approach with prior models which enforce sparsity in signal and image processing. *EURASIP Journal on Advances in Signal Processing*, 2012.
- [MD12b] A. Mohammad-Djafari. Variational bayesian approximation methods for inverse problems. *Journal of Physics : Conference Series*, 386(1), 2012.

- [MDF06] A. Mohammad-Djafari and O. Féron. Bayesian approach to change points detection in time series. *International Journal of Imaging Systems and Technology*, 16(5) :215–221, 2006.
- [MFM⁺12] V. Mazet, S. Faisan, A. Massony, M. . Gaveau, L. Poisson, and J. . Mestdagh. Joint bayesian decomposition of a spectroscopic signal sequence with rjmc. In *2012 IEEE Statistical Signal Processing Workshop, SSP 2012*, pages 253–256, 2012.
- [MWJ05] M. J. Menne and C. N. Williams Jr. Detection of undocumented changepoints using multiple test statistics and composite reference series. *Journal of Climate*, 18(20) :4271–4286, 2005.
- [NHH⁺09] A. K. Naik, M. S. Hanay, W. K. Hiebert, X. L. Feng, and M. L. Roukes. Towards single-molecule nanomechanical mass spectrometry. *Nature Nanotechnology*, 4(7) :445–450, 2009.
- [PRK93] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad. Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition. In *Conference Record of the Asilomar Conference on Signals, Systems & Computers*, volume 1, pages 40–44, 1993.
- [RBNC03] O. Rosec, J. . Boucher, B. Nsiri, and T. Chonavel. Blind marine seismic deconvolution using statistical mcmc methods. *IEEE Journal of Oceanic Engineering*, 28(3) :502–512, 2003.
- [RCW⁺07] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu. A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, 46(6) :900–915, 2007.
- [RFA03] S. G. Razul, W. J. Fitzgerald, and C. Andrieu. Bayesian model selection and parameter estimation of nuclear emission spectra using rjmc. *Nuclear Instruments and Methods in Physics Research, Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 497(2-3) :492–510, 2003.
- [RFKN92] Frank C. Robey, Daniel R. Fuhrmann, Edward J. Kelly, and Ramon Nitzberg. A cfar adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1) :208–216, 1992.
- [SIBD11] C. Soussen, J. Idier, D. Brie, and J. Duan. From bernoulli-gaussian deconvolution to sparse signal restoration. *IEEE Transactions on Signal Processing*, 59(10) :4572–4584, 2011.
- [Sta10] J.-L. Starck. Sparsity and astronomical data analysis. In *Séminaire DRT sur la Parcimonie*, Octobre 2010.
- [TDL03] J. Y. Tourneret, M. Doisy, and M. Lavielle. Bayesian off-line detection of multiple change-points corrupted by multiplicative noise : Application to sar image edge detection. *Signal Processing*, 83(9) :1871–1887, 2003.
- [Tib11] R. Tibshirani. Regression shrinkage and selection via the lasso : A retrospective. *Journal of the Royal Statistical Society. Series B : Statistical Methodology*, 73(3) :273–282, 2011.
- [Wan08] X. L. Wang. Penalized maximal f test for detecting undocumented mean shift without trend change. *Journal of Atmospheric and Oceanic Technology*, 25(3) :368–384, 2008.
- [WWW07] X. L. Wang, Q. H. Wen, and Y. Wu. Penalized maximal t test for detecting undocumented mean change in climate data series. *Journal of Applied Meteorology and Climatology*, 46(6) :916–931, 2007.

- [XGB07] B. Xing, C. M. T. Greenwood, and S. B. Bull. A hierarchical clustering method for estimating copy number variation. *Biostatistics*, 8(3) :632–653, 2007.

Bibliographie pour le chapitre 6

- [DM92] Guan-Zhong Dai and Jerry M. Mendel. Maximum a posteriori estimation of multichannel bernoulli-gaussian sequences. *IEEE Transactions on Information Theory*, v(n) :181–183, 1992.
- [DSBH07] S. Dohn, W. Svendsen, A. Boisen, and O. Hansen. Mass and position determination of attached particles on cantilever based mass sensors. *Review of Scientific Instruments*, 78(10), 2007.
- [HKN⁺12] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9) :602–608, 2012.
- [SRSE11] P. Sprechmann, I. Ramírez, G. Sapiro, and Y. C. Eldar. C-hilasso : A collaborative hierarchical sparse modeling framework. *IEEE Transactions on Signal Processing*, 59(9) :4183–4198, 2011.
- [TGS05] J. A. Tropp, A. C. Gilbert, and M. J. Strauss. Simultaneous sparse approximation via greedy pursuit. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, volume V, pages V721–V724, 2005.
- [TGS06] J. A. Tropp, A. C. Gilbert, and M. J. Strauss. Algorithms for simultaneous sparse approximation. part i : Greedy pursuit. *Signal Processing*, 86(3) :572–588, 2006.
- [Tro06] J. A. Tropp. Algorithms for simultaneous sparse approximation. part ii : Convex relaxation. *Signal Processing*, 86(3) :589–602, 2006.
- [WNF09] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7) :2479–2493, 2009.

Bibliographie pour le chapitre 7

- [Dun73] J. C. Dunn. Fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *J Cybern*, 3(3) :32–57, 1973.
- [ES97] O. Egecioglu and A. Srinivasan. A fast non-parametric density estimation algorithm. *Communications in Numerical Methods in Engineering*, 13(10) :755–763, 1997.
- [FGQ12] J. Fleureau, P. Guillotel, and H. . Quan. Physiological-based affect event detector for entertainment video applications. *IEEE Transactions on Affective Computing*, 3(3) :379–385, 2012.
- [HKN⁺12] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9) :602–608, 2012.
- [Kri77] Allen S. Krieger. Analysis of photographic x-ray images. *X-Ray Imaging*, 106 :24–33, 1977.
- [Maz05] V. Mazet. *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. PhD thesis, Université Henri Poincaré, Nancy 1, 2005.
- [McM96] Daniel W. McMichael. Estimating gaussian mixture models from data with missing features. In *Proceedings of the International Symposium on Signal Processing and its Applications, ISSPA*, volume 1, pages 377–378, 1996.
- [MFM⁺12] V. Mazet, S. Faisan, A. Massony, M. . Gaveau, L. Poisson, and J. . Mestdagh. Joint bayesian decomposition of a spectroscopic signal sequence with rjcmc. In *2012 IEEE Statistical Signal Processing Workshop, SSP 2012*, pages 253–256, 2012.
- [Min] T. Minka.
- [MM93] A. A. Marchetti and A. C. Mignerey. Deconvolution of mass spectra. *Nuclear Inst. and Methods in Physics Research, A*, 324(1-2) :288–296, 1993.
- [PBT93] Nikhil R. Pal, James C. Bezdek, and Eric C. . Tsao. Generalized clustering networks and kohonen’s self-organizing scheme. *IEEE Transactions on Neural Networks*, 4(4) :549–557, 1993.
- [SG99] Chris Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2 :246–252, 1999.

- [TPJ05] M. Taron, N. Paragios, and M. . Jolly. Modelling shapes with uncertainties : Higher order polynomials, variable bandwidth kernels and non parametric density estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, volume II, pages 1659–1666, 2005.
- [WXW12] Q. . Wen, S. Xu, and S. . Wan. Medical image deconvolution in besov space based on sparse decomposition. *Moshi Shibie yu Rengong Zhineng/Pattern Recognition and Artificial Intelligence*, 25(3) :550–556, 2012.
- [Yua09] J. Yuan. Blind deconvolution of x-ray diffraction profiles by using high-order statistics. *Optical Engineering*, 48(7), 2009.

Résumé : La miniaturisation des composants électroniques conduit aujourd'hui au développement de capteurs ultra-sensibles. En particulier, les capteurs NEMS (systèmes électromécaniques nanométriques) ont maintenant une sensibilité suffisante pour détecter des molécules uniques. Ceci permet d'intégrer ces capteurs dans des dispositifs de spectrométrie de masse dont la particularité sera d'opérer en mode comptage de molécules uniques. Notre travail consiste à reconstruire le spectre de masse de la solution analysée à partir des signaux fréquentiels délivrés par les NEMS. Nous nous plaçons dans le cadre des approches problèmes inverses et des méthodes d'inférence bayésienne. Nous modélisons le système de mesure qui lie les inconnues aux signaux observés par un modèle graphique hiérarchique et nous introduisons un modèle de signal de type processus ponctuel marqué. Nous le comparons à un modèle de type processus à temps discret.

Nous mettons en place un algorithme de déconvolution impulsionnelle intégrant une exploration de modèles qui réalise la détection des molécules analysées, l'estimation de leur masse et le comptage, afin de reconstruire le spectre de masse de la solution analysée.

Nous présentons des résultats sur données simulées et sur des données expérimentales acquises au CEA/INAC sur des agrégats de Tantale en utilisant des capteurs NEMS développés au CEA-Leti/DCOS. Relativement aux méthodes de l'état de l'art, la méthode que nous proposons améliore le taux de comptage tout en gardant un taux de fausses détections suffisamment bas. Notre méthode délivre également les incertitudes sur les paramètres reconstruits. Enfin, nous développons le cas particulier de la reconstruction de spectres de masse discrets.

Mots-clés : Protéomique, NEMS, nanotechnologies, spectrométrie de masse, traitement de l'information, traitements spectrométriques, profil moléculaire, traitement en mode comptage, problèmes inverses, approche bayésienne, déconvolution impulsionnelle, détection, quantification.

Abstract : The miniaturization of electronic components drives the development of very sensitive sensors. In particular, NEMS (Nano ElectroMechanical Systems) are now sensitive enough to detect single molecules. This enables to use these sensors in order to design mass spectrometry devices, in an individual molecules counting mode. Our objective is to reconstruct the mass spectrum of the analyzed solution, based on the NEMS output signals.

We use inverse problems approach and Bayesian framework. We model the acquisition system linking the unknown parameters to the observable signals with a hierarchical graphical model. We propose a marked-point process model of signal that we compare with discrete-time process one. We develop an impulse deconvolution algorithm which relies on a model exploration scheme. This enables us to detect the molecules, to quantify their mass and to count them in order to estimate the mass spectrum of the analyzed solution.

We show results on simulated data and on experimental ones acquired in CEA/INAC using Tantalum nano-aggregates and devices developed in CEA-Leti/DCOS. Compared to state-of-the-art, our method offers high counting rate and keeps a low false detection rate. It also permits the computation of uncertainties on estimated values. Finally, we propose a derivation of the method to deal with the reconstruction of discrete mass spectra.

Key words : Proteomics, NEMS, nanotechnologies, mass spectrometry, information processing, spectrometry, molecular profile, counting-mode processing, inverse problems, Bayesian framework, impulse deconvolution, detection, quantification.