

**ÉCOLE DOCTORALE MATHÉMATIQUES, INFORMATIQUE,
PHYSIQUE THÉORIQUE ET INGÉNIERIE DES SYSTÈMES**

Laboratoire Pluridisciplinaire de Recherche en Ingénierie des
Systèmes, Mécanique et Énergétique (PRISME)

THÈSE présentée par :

Mohamed NAIT MEZIANE

Soutenue le : **9 décembre 2016**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline / Spécialité : **Traitement du Signal**

**Identification d'appareils électriques par analyse des
courants de mise en marche**

THÈSE dirigée par :

Philippe RAVIER
Jean-Charles LE BUNETEL
Guy LAMARQUE

MCF–HDR, Université d'Orléans, directeur de thèse
PR, Université de Tours, co-directeur de thèse
MCF, Université d'Orléans, co-encadrant

RAPPORTEURS :

Jérôme MARS
Jean MERCKLÉ

PR, Grenoble-INP
PR, Université de Haute-Alsace

JURY :

Karim ABED-MERAIM
Bruno CHARBONNIER
Karim DROUCHE
Guy LAMARQUE
Jean-Charles LE BUNETEL
Jérôme MARS
Jean MERCKLÉ
Philippe RAVIER

PR, Université d'Orléans, président du jury
Ingénieur de Recherche, EDF R&D, examinateur
MCF–HDR, Université de Cergy-Pontoise, examinateur
MCF, Université d'Orléans, co-encadrant
PR, Université de Tours, co-directeur de thèse
PR, Grenoble-INP, rapporteur
PR, Université de Haute-Alsace, rapporteur
MCF–HDR, Université d'Orléans, directeur de thèse

Remerciements

Je tiens tout d'abord à remercier mes directeurs de thèse, Philippe Ravier et Jean-Charles Le Bunetel ainsi que mes co-encadrants Guy Lamarque et Yves Raingeaud, pour m'avoir donné l'occasion de faire une thèse dans un domaine en plein essor et assez intéressant. Je remercie particulièrement Philippe, mon directeur de thèse et Guy, mon co-encadrant de thèse sur Orléans, avec qui j'ai pris beaucoup de plaisir à travailler. Je les remercie pour leurs conseils, les discussions et réunions constructives qu'on faisait mais aussi pour leur accueil chaleureux, leur gentillesse, leur soutien, leur confiance en moi ainsi que leur disponibilité. Mes remerciements aussi à Jean-Charles Le Bunetel et à Yves Raingeaud pour leurs remarques et suggestions, surtout celles liées à la mesure et au réseau électrique, domaines ne constituant pas mon point fort au début de cette thèse. Je les remercie aussi pour leur suivi et accueil chaleureux lors de mon séjour au laboratoire GREMAN, UMR de l'université de Tours/CNRS.

Je remercie également Karim Abed-Meraim, professeur à l'université d'Orléans, pour ses directives et conseils avisés et qui ont beaucoup contribué à faire avancer ce travail surtout au moments où trouver des issues était difficile. Je pense que sa présence au laboratoire PRISME de l'université d'Orléans est un gain considérable pour ce dernier. Mes remerciements vont également aux professeurs Jérôme Mars et Jean Mercklé pour m'avoir fait l'honneur de rapporter sur mon travail ainsi que pour leurs remarques pertinentes, particulièrement Jérôme Mars qui m'a posé plein de questions intéressantes, chose que j'ai apprécié et qui, pour moi, témoigne de l'intérêt qu'il a donné à mon travail. Je remercie également Bruno Charbonnier et Karim Drouiche pour avoir accepté d'examiner mon travail ainsi que pour leurs remarques pertinentes sur le sujet ainsi que sur mon travail.

Je tiens aussi à remercier mon ami Abdenour Hacine-Gharbi, maître de conférences à l'université de Bordj Bou Arréridj, Algérie, pour m'avoir assisté dans la partie classification des appareils électrique. Je le remercie pour la collaboration qu'on a eu, ses conseils, ses encouragements, sa bonté de cœur ainsi que pour m'avoir inspiré par sa volonté infatigable et envie de continuer à faire de la recherche malgré toutes les difficultés.

Je remercie également Thomas Picon, que j'ai co-encadré durant son projet de fin d'études qu'il a effectué avec nous, pour son sérieux et son ouverture d'esprit. J'ai vraiment apprécié de travailler avec toi Thomas. Je remercie aussi Clarisse Novello pour nous avoir aidé dans la réalisation du système de mesure, design des cartes électroniques et leur réalisation ainsi que le suivi de l'assemblage de l'ensemble. J'adresse également mes remerciements à mes collègues du laboratoire PRISME, permanents, stagiaires ou visiteurs, sans pouvoir tous les citer, particulièrement Abdelbassit Boualem et Viet-Dung Nguyen, mes collègues de bureau et mes amis, avec qui j'ai vécu de très bons moments dans une ambiance très agréable. Je remercie également toute personne ayant contribué de près ou de loin à la réussite de ce modeste travail.

Finalement, je tiens à remercier ma femme qui a su m'épauler et me donner le courage aux moments difficiles, mes frères et sœurs qui contribuent à mon bonheur ainsi que mes chers parents que je ne saurais jamais remercier assez, je leur témoigne mon respect, admiration et gratitude.

Abréviations

Français		English	
ADL	Analyse Discriminante Linéaire	LDA	Linear Discriminant Analysis
ADF	Analyse Discriminante de Fisher	KFDA	Kernel Fisher Discriminant Analysis
BCR	Borne de Cramér-Rao	CRB	Cramér-Rao Bound
BE	Basé Évènements	EB	Event-Based
BFE	Basse Fréquence d'Échantillonnage	LSF	Low Sampling Frequency
CE	Calculateur d'Énergie	EM	Energy Meter
CEBC	Calculateur d'Énergie à Bas Coût	LCEM	Low Cost Energy Meter
–	–	COOLL	Controlled On/Off Loads Library
DE	Détecteur d'Enveloppe	ED	Envelope Detector
DEA	Détecteur d'Enveloppe Amélioré	EED	Enhanced Envelope Detector
ECT	Éstimation du Courant Transitoire	TCE	Transient Current Estimation
ECTA	Éstimation du Courant Transitoire Améliorée	ETCE	Enhanced Transient Current Estimation
ET	Écart-Type	STD	STandard Deviation
EQM	Erreur Quadratique Moyenne	MSE	Mean Square Error
ERDF	Électricité Réseau Distribution France	–	–
FN	Faux Négatifs	FN	False Negatives
FP	Faux Positifs	FP	False Positives
–	–	HAND	High Accuracy NILM Detector
HFE	Haute Fréquence d'Échantillonnage	HSF	High Sampling Frequency
–	–	LED	Light Emitting Diode
MAC3	Mesure Aveugle de la Consommation des Charges au Compteur	–	–
MC	Moindres Carrés	LS	Least-Squares
MCL	Moindres Carrés Linéaires	LLS	Linear Least-Squares
MCNL	Moindres Carrés Non Linéaires	NLLS	Non-Linear Least-Squares
MDE	Maîtrise de la Demande d'Énergie	–	–
MIF	Matrice d'Information de Fisher	FIM	Fisher Information Matrix
MMC	Modèles de Markov Cachés	HMM	Hidden Markov Models
MMG	Modèle de Mélange Gaussien	GMM	Gaussian Mixture Model)

Français		English	
–	–	MZC	Modified Zero-Crossing
NBE	Non Basé Évènements	NEB	Non Event-Based
PC	Pince de Courant	CT	Current Transducer
–	–	PLAID	Plug Load Appliance Identification Dataset
REQM	Racine carrée de l’Erreur Quadratique Moyenne	RMSE	Root-Mean Square Error
RSB	Rapport Signal à Bruit	SNR	Signal-to-Noise Ratio
RT	Règlementation Thermique	–	–
RTE	Réseau de Transport d’Électricité	–	–
SF	Séries de Fourier	FS	Fourier Series
SIC	Suivi Intrusif de la Charge	ILM	Intrusive Load Monitoring
SFD	Séries de Fourier Discrètes	DFS	Discrete Fourier Series
SFDCT	Séries de Fourier Discrètes à Court-Terme	STDFS	Short-Time Discrete Fourier Series
SNIC	Suivi Non-Intrusif de la Charge	NILM	Non-Intrusive Load Monitoring
SSIC	Suivi Semi-Intrusif de la Charge	SILM	Semi-Intrusive Load Monitoring
ST	Sonde de Tension	VP	Voltage Probe
TBC	Taux de Bonne Classification	CCR	Correct Classification Rate
TC	Taux de Classification	CR	Classification Rate
TF	Transformée de Fourier	FT	Fourier Transform
TFCT	Transformée de Fourier à Court-Terme	STFT	Short-Time Fourier Transform
TFD	Transformée de Fourier Discrète	DFT	Discrete Fourier Transform
TFR	Transformée de Fourier Rapide	FFT	Fast Fourier Transform
TO	Transformée en Ondelettes	WT	Wavelet Transform
–	–	TRR	Trust-Region-Reflective
VN	Vrais Négatifs	TN	True Negatives
VP	Vrais Positifs	TP	True Positives

Table des matières

Remerciements	i
Abréviations	iii
Introduction générale	1
1 Contexte de l'étude	7
1.1 Motivation	8
1.1.1 Transition énergétique	8
1.1.2 Règlementation thermique 2012 (RT 2012)	11
1.1.3 Obligation de faire du sous-comptage	13
1.1.4 Projet MDE–MAC3	15
1.2 Désagrégation des consommations des appareils électriques	17
1.2.1 État de l'art succinct	17
1.2.2 Problématique du NILM	19
1.3 Démarche retenue	22
2 Mesure de la consommation électrique	25
2.1 Introduction	26
2.2 Approches de comparaison	27
2.2.1 Mesures de puissance	28
2.2.2 Mesures d'énergie	31
2.3 Résultats expérimentaux	32
2.3.1 Puissance calculée en utilisant des mesures de laboratoire comme référence	32
2.3.2 Mesures d'énergie avec le compteur intelligent Linky comme référence	37
2.4 Conclusion	39
3 Analyse des courants de mise en marche	40
3.1 Introduction	41
3.2 Caractérisation paramétrique des transitoires de courant	42
3.2.1 Modélisation	42
3.2.2 Hypothèses de modélisation	44
3.2.3 Algorithme d'estimation des paramètres (ECT)	45
3.2.4 Calcul des performances théoriques–BCR	51
3.3 Performances expérimentales de l'algorithme	54
3.3.1 Évaluation sur données simulées	54

3.3.2	Évaluation sur données réelles	72
3.3.3	Algorithme d'estimation des paramètres amélioré (ECTA)	77
3.4	Détection d'évènements	83
3.4.1	Algorithme de détection (HAND)	83
3.4.2	Résultats expérimentaux	89
3.4.3	Conclusion	94
4	Identification des appareils électriques	95
4.1	Preuve de concept–utilité du transitoire <i>vs.</i> régime établi	96
4.1.1	Analyse harmonique	97
4.1.2	Résultats expérimentaux	99
4.2	Création d'une base de transitoires	108
4.2.1	Enjeu	108
4.2.2	Description du système de mesure réalisé et ses principales caractéristiques	110
4.2.3	Exemples de mesures	112
4.2.4	Organisation de la base de données	119
4.3	Classification des appareils	121
4.3.1	Pré-analyse des paramètres estimés	122
4.3.2	Nuages de points des paramètres estimés	126
4.3.3	Identification	133
4.3.4	Clustering	136
4.3.5	Comparaison des résultats de classification vis-à-vis de la preuve du concept	139
	Conclusion et perspectives	141
A	Algorithme Modified Zero-Crossing (MZC) pour l'estimation et le suivi de la fréquence fondamentale	145
	Bibliographie	148

Table des figures

1.1	Consommation finale d'énergie du secteur résidentiel-tertiaire en France. Source : données du SOeS, bilan de l'énergie 2014. La ligne de tirets verticaux marque la date du Grenelle 1.	9
1.2	Production d'énergie primaire en France. Source : SOeS, bilan de l'énergie 2009.	9
1.3	Évolution de la répartition des principaux secteurs dans la consommation d'électricité en France. Source : (RTE, 2016).	10
1.4	Répartition sectorielle de la consommation d'électricité en France pour l'année 2015. Source : (RTE, 2016).	11
1.5	Répartition par usage de la demande électrique résidentielle en France pour l'année 2015. Source : (RTE, 2016).	14
1.6	Citations du travail fondateur du NILM de Hart (Hart, 1992) entre 1993 à 2015. Source des données : Citations Google Scholar.	20
1.7	Schéma général d'un système NILM. En rouge les trois critères selon lesquels on peut classer les approches NILM.	20
2.1	Premier protocole de mesure de puissance utilisant un CE à la fois.	30
2.2	Deuxième protocole de mesure de puissance utilisant tous les CE simultanément.	30
2.3	Protocole de mesure d'énergie.	31
3.1	Courant de mise en marche d'un aspirateur.	42
3.2	Courant de mise en marche d'un micro-ondes.	45
3.3	Illustration du résultat de détection donné par HAND sur un signal simulé mono-appareil. En sortie HAND donne $t_{dm}, t_{fm}, t_{da}, t_{fa}$. Les instants t_{e1} et t_{e2} définissent la portion du régime établi utilisée pour l'estimation des A_i et ϕ_i	46
3.4	Courant de mise en marche simulé pour utilisation dans l'évaluation avec données simulées.	56
3.5	Biais des différents paramètres estimés $\hat{\mathbf{p}}, \hat{\mathbf{A}}, \hat{\phi}$ en considérant l'estimation du $\mathbf{p}$ indépendamment de $\mathbf{A}$ et ϕ sans phase d'optimisation. 1000 itérations Monte-Carlo.	57
3.6	Biais des différents paramètres estimés $\hat{\mathbf{p}}, \hat{\mathbf{A}}, \hat{\phi}$ en considérant l'estimation du $\mathbf{p}$ indépendamment de $\mathbf{A}$ et ϕ avec une phase optimisation. 1000 itérations Monte-Carlo.	58
3.7	Comparaison de la BCR (CRB) et variance du paramètre $\hat{\mathbf{p}}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.	60

3.8	Comparaison de la BCR (CRB) et variance du paramètre $\hat{\mathbf{A}}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.	61
3.9	Comparaison de la BCR (CRB) et variance du paramètre $\hat{\phi}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.	62
3.10	BCR des différents paramètres estimés $\hat{\mathbf{p}}, \hat{\mathbf{A}}, \hat{\phi}$ en fonction de la fréquence d'échantillonnage F_s (de 1 kHz à 100 kHz). Calculs faits à RSB = 25 dB.	64
3.11	Erreur quadratique moyenne $\frac{1}{N} \sum_{n=1}^N x(n) - s(\hat{\theta}, n) ^2$ en fonction du nombre d'itérations de l'algorithme d'estimation utilisant l'algorithme trust-region-reflective pour différentes valeurs du RSB. Pour chaque courbe, le dernier point correspond à l'arrêt de l'algorithme (convergence).	65
3.12	Δ Erreur quadratique moyenne $\frac{1}{N} \sum_{n=1}^N x(n) - s(\hat{\theta}, n) ^2$ entre deux itérations consécutives en fonction du nombre d'itérations de l'algorithme d'optimisation utilisant l'algorithme trust-region-reflective pour différentes valeurs du RSB. L'itération avec un $\Delta\text{EQM} = 0$ (minimum atteint) n'apparaissant pas sur ce graphique semi-logarithmique nous indiquons (flèches) ci-contre les itérations qui précèdent.	66
3.13	BCR du signal $s(t, \boldsymbol{\theta})$ comparée à la variance, calculée point par point, du signal estimé $\hat{s}(t, \boldsymbol{\theta})$. Calculs faits à RSB = 50 dB.	67
3.14	Différence entre la BCR du signal $s(t, \boldsymbol{\theta})$ et la variance, calculée point par point, du signal estimé $\hat{s}(t, \boldsymbol{\theta})$ (figure 3.13a).	68
3.15	Valeurs moyennes (par rapport au temps) de la BCR du signal $s(t, \boldsymbol{\theta})$ et la variance, calculée point par point, du signal estimé $\hat{s}(t, \boldsymbol{\theta})$ en fonction du RSB.	68
3.16	Comparaison de la BCR de $\mathbf{p}$ avec et sans introduction de f_0 comme inconnue.	69
3.17	Comparaison de la BCR de $\mathbf{a}$ avec et sans introduction de f_0 comme inconnue.	70
3.18	Comparaison de la BCR de ϕ avec et sans introduction de f_0 comme inconnue.	71
3.19	Fréquence fondamentale du réseau estimée sur 24h entre le 9 et le 10 mars 2016 (début de mesure à 17 :00). La méthode d'estimation est basée sur la méthode de passages par zéros simple. Sur le graphe on affiche un point par seconde.	73
3.20	Courant de mise en marche d'un ventilateur.	75
3.21	Courant de mise en marche d'un ventilateur (en bleu) avec le signal reconstruit (en rouge) avec les paramètres estimés en utilisant l'algorithme ECT.	76
3.22	Résultat d'estimation d'une fréquence simulée en utilisant l'algorithme modifié zéro-crossing (MZC). Bleu : fréquence exacte. Rouge : fréquence estimée.	79
3.23	Fréquence fondamentale estimée $\hat{f}_0(t)$ au cours de la mesure du ventilateur.	81
3.24	Courant de mise en marche d'un ventilateur (en bleu) avec le signal reconstruit (en rouge) avec les paramètres estimés en utilisant l'algorithme ECTA.	82
3.25	Amélioration de la détection de l'enveloppe du signal courant. Exemple sur un évènement de mise en marche : $e_{d1}(t)$ est détecté sans test de médiane et $e_{d2}(t)$ avec.	85

3.26	Illustration du « principe » de calcul de $\sigma_d(t)$ en utilisant une fenêtre glissante de taille L . Notons que concrètement cela est fait d'une manière équivalente mais itérative.	86
3.27	Résultats du KFDA sur données réelles. Haut : pics rouges indiquent la vérité terrain (instants de début d'évènements) et les évènements détectés sont cadrés dans des boîtes.	93
3.28	Résultats du HAND sur données réelles. Haut : pics rouges indiquent la vérité terrain (instants de début d'évènements) et les évènements détectés sont cadrés dans des boîtes.	93
4.1	Illustration de l'extraction des vecteurs de paramètres $C_{k,i}, i = 1, \dots, n$ (observations).	99
4.2	Diagramme bloc du système d'identification d'appareils électriques proposé.	101
4.3	Deux exemples de signaux de la base PLAID.	103
4.4	TC en fonction du nombre d'harmoniques d'ordre impair considérées.	107
4.5	Photo du système de mesure proposé.	110
4.6	Schéma bloc du système de mesure proposé.	111
4.7	Courant électrique d'un aspirateur. Même mesure répétée cinq fois avec un même retard de déclenchement de 0 ms (par rapport à la période secteur).	113
4.8	Mesure de courant pour une lampe halogène de 100 W et une lampe LED de 1,6 W.	114
4.9	Histogramme du bruit de courant.	115
4.10	Courant (haut) et tension (bas) d'une lampe halogène. On montre quatre exemples d'une mesure faite sur la même charge avec un retard de déclenchement variable (0 ms, 5 ms, 10 ms and 15 ms).	116
4.11	Courant (haut) et tension (bas) d'une perceuse. On montre quatre exemples d'une mesure faite sur la même charge avec un retard de déclenchement variable (0 ms, 5 ms, 10 ms and 15 ms).	117
4.12	Démarrages successifs (a) lents et (b) rapides d'une perceuse.	118
4.13	Exemple d'un fichier de configuration.	120
4.14	Exemple des phases estimées $\hat{\phi}_i$ pour la mesure du même appareil « scie » en fonction de différents délais de mise en marche.	122
4.15	Exemple des amplitudes estimées $\hat{A}_i$ pour la mesure du même appareil « scie » en fonction de différents délais de mise en marche.	123
4.16	Fonctions $g_1(t)$ et $g_2(t)$	124
4.17	Comparaison des valeurs de p_3 pour une lampe et une scie en fonction de différents délais de mise en marche.	125
4.18	Courant de mise en marche d'une lampe fluo-compacte avec zones de pré- et post-déclenchement.	125
4.19	Courant de mise en marche d'un ventilateur.	127
4.20	Nuage de points des paramètres A_0, b_3 et A_1	128
4.21	Nuage de points des paramètres A_0, b_3 et A_1 sans les lampes.	129
4.22	Nuage de points des paramètres A_0, b_3 et A_1 sans les lampes et sans sèche-cheveux.	130

4.23	Nuage de points des paramètres A_1, A_2 et A_3 .	132
4.24	Matrice de confusion pour l'identification en utilisant A_0 et b_3 . TBC = 83.33%.	134
4.25	Matrice de confusion pour l'identification en utilisant les A_i . TBC = 98.57%.	135
4.26	Matrice de confusion pour l'identification en utilisant A_0, b_3 et les A_i . TBC = 91.43%.	135
4.27	Matrice de confusion pour le clustering en utilisant A_0 et b_3 .	137
4.28	Matrice de confusion pour le clustering en utilisant les A_i .	138
4.29	Matrice de confusion pour le clustering en utilisant A_0, b_3 et les A_i .	139
A.1	Décalage du point de passage par zéro suite à l'ajout de bruit sur un signal de tension.	146

Liste des tableaux

1.1	Résumé des méthodes NILM organisées par fréquence d'échantillonnage. Source : (Clark, 2015)	24
2.1	Coût et résolution d'affichage des Calculateurs d'énergie (CE) considérés	27
2.2	Consommation intrinsèque indiquée vs. estimée des CE	31
2.3	Appareils mesurés pour le premier protocole de mesure de puissance	35
2.4	REQMs relatives des différentes mesures des CE évaluées en utilisant le premier protocole de mesure de puissance	35
2.5	REQMs relatives des différentes mesures des CE évaluées en utilisant le premier protocole de mesure de puissance–Suite	35
2.6	Appareils mesurés pour le deuxième protocole de mesure de puissance	35
2.7	REQMs relatives des différentes mesures des CE évaluées en utilisant le deuxième protocole de mesure de puissance	35
2.8	Résultat de la mesure d'énergie sur 48h–première expérience	39
2.9	Résultat de la mesure d'énergie sur 48h–seconde expérience	39
3.1	TP, FN, FP, précision et rappel du HAND <i>vs.</i> KFDA, RSB = 50 dB, Δt = 200 ms et Δt = 500 ms	91
3.2	TP, FN, FP, précision et rappel of HAND <i>vs.</i> KFDA, Δt = 500 ms, diffé- rents RSBs	92
3.3	Erreurs dans la détection des instants de début pour les cinq évènements montrés dans les figures 3.27 et 3.28	94
4.1	Résumé des appareils trouvés dans la base PLAID	102
4.2	Les différentes subdivisions de la base de données créées pour l'évaluation du système d'identification	104
4.3	Évaluation du TC pour le transitoire <i>vs.</i> le régime établi	105
4.4	Évaluation du TC pour des harmoniques d'ordre pair <i>vs.</i> impaire	106
4.5	Résumé de la base de données	119
4.6	Exemple du contenu d'un fichier en-tête « .xlsx »	120

Liste des Algorithmes

1	Estimation du Courant Transitoire (ECT).	50
2	Estimation du Courant Transitoire Amélioré (ECTA).	80
3	Détecteur d'Enveloppe Amélioré (DEA).	86
4	High Accuracy NILM Detector (HAND).	88

Introduction générale

Cette thèse a été effectuée dans le cadre du projet « Maîtrise de la Demande d’Energie– Mesure Aveugle de la Consommation des Charges au Compteur (MDE–MAC3) » financé par la région Centre-Val de Loire. Elle s’est déroulée au sein de l’équipe *Signal* du laboratoire Pluridisciplinaire de Recherche en Ingénierie des Systèmes, Mécanique et Energétique (PRISME) de l’université d’Orléans et en collaboration avec le Groupe de Recherche en Matériaux, Microélectronique, Acoustique et Nanotechnologies (GREMAN), UMR 7347 de l’université de Tours et du CNRS.

Le besoin en énergie sous toutes ses formes disponibles et exploitables ne cesse de croître pour répondre aux nécessités de la croissance que connaît l’économie mondiale. Avec toutes les avancées et avantages qu’offre cette croissance, elle impacte négativement l’environnement et par conséquent affecte notre santé et bien être.

La France, ayant conscience de ces enjeux économiques et environnementaux, a décidé de faire un saut, ou transition, énergétique qui lui permettrait de garder une économie forte tout en respectant l’environnement. Cette décision a été concrétisée au travers ce qu’on appelle le « Grenelle Environnement » qui est un ensemble de rencontres politiques organisées en septembre et décembre 2007. Le principal objectif de ces rencontres est de prendre des décisions qui vont en faveur du développement durable défini comme étant : « un développement qui répond au besoin du présent sans compromettre la capacité des générations futures à répondre aux leurs »¹.

Dans cette même optique, la réglementation thermique (RT) française vise à minimiser la consommation électrique du bâtiment neuf. Plus spécialement, la RT 2012, celle en vigueur, demande que la consommation électrique des bâtiments neufs soit donnée

1. <http://www.developpement-durable.gouv.fr/Definition-du-developpement,15067.html>

par usage selon la répartition minimale (en cinq catégories) : chauffage, refroidissement, production d'eau chaude sanitaire, réseau de prises électriques ainsi qu'une dernière catégorie rassemblant le reste de la consommation. L'intérêt d'une information détaillée sur la consommation électrique est de pouvoir l'exploiter pour des actions de réduction énergétique liées à la prise de conscience de la consommation individuelle des habitants, des actions de délestage intelligent pour le distributeur d'électricité ou même des actions en faveur de la sécurité et du bien-être des individus.

Problématique et approche considérée

Dans ce contexte d'efficacité énergétique, mon travail de thèse se focalise sur le secteur résidentiel et sur les appareils domestiques. Il cherche à contribuer à résoudre la problématique d'obtention d'une information détaillée de la consommation électrique à partir d'une mesure globale de cette dernière. Cette consommation détaillée devrait permettre d'avoir accès à trois informations essentielles, à savoir, qui est en train de consommer ? quand ? et combien consomme-t-il ?

Le travail présenté dans ce manuscrit est loin de proposer une solution complète répondant à cette problématique, mais quelques avancées sur chacune des questions posées.

Ce travail cherche principalement à répondre à la première question qui peut être reformulée, dans le cadre de la thèse, ainsi : comment exploiter les courants de mise en marche pour identifier les appareils électriques qui les génèrent ? Pour répondre à cette question, notre approche consiste tout d'abord à modéliser le courant de mise en marche avec un modèle mathématique dépendant de certains paramètres puis à utiliser ces paramètres (estimés) comme caractéristiques discriminantes pour identifier l'appareil électrique qui est à l'origine de la génération de ce courant.

En réponse à la deuxième question, nous avons proposé un détecteur amélioré d'événements de mise en marche et de mise à l'arrêt des appareils, utilisant l'enveloppe du signal de courant. De plus, nous avons été amenés à caractériser des calculateurs d'énergie grand public afin d'identifier leurs limites en termes de qualité de mesure de la puissance ou de l'énergie électriques, ce qui rentre dans le cadre de la troisième question.

Structure du manuscrit

Le manuscrit de thèse est organisé en quatre chapitres comme suit :

Chapitre 1 : Ce chapitre donne d’abord la motivation du travail de thèse par rapport au contexte de transition et d’efficacité énergétique ainsi que par rapport à l’obligation de faire du sous-comptage de l’énergie électrique consommée imposée par la réglementation thermique en vigueur et visant principalement à limiter la consommation énergétique des bâtiments neufs.

Ce chapitre présente aussi le projet MDE–MAC3 dans le cadre duquel est effectuée cette thèse et situe le travail dans le cadre plus général de la désagrégation d’énergie et plus spécialement son approche non-intrusive communément appelée NILM (Non-Intrusive Load Monitoring). Une brève description des différentes approches de désagrégation d’énergie est donnée ainsi qu’un état de l’art succinct des méthodes NILM en précisant les différentes sous-problématiques à traiter dans ce domaine.

La fin de ce chapitre décrit la démarche choisie pour ce travail de thèse et la justifie par rapport aux différentes sous-problématiques du NILM.

Chapitre 2 : Pour répondre à une des préoccupations du projet MDE–MAC3, ce chapitre présente une étude comparative (en terme de précision du résultat de mesure affiché) de différentes solutions à bas coût utilisées pour le calcul de la consommation électrique. Les différentes approches proposées pour la comparaison sont présentées ainsi que les résultats associés. Notons que ce chapitre peut être lu indépendamment du reste du manuscrit.

Chapitre 3 : Ce chapitre traite essentiellement l’analyse des courants transitoires de mise en marche et propose :

- un modèle paramétrique que nous présentons spécifiquement pour modéliser les courants de mise en marche ;
- une étude des performances théoriques, Bornes de Cramér-Rao (BCR), de ce modèle ;
- un algorithme d’estimation des paramètres pour le modèle proposé ainsi que sa version améliorée qui tient compte de la variabilité de la fréquence du réseau ;
- une évaluation des performances des estimateurs (comparaison avec les BCR)

ainsi qu’une évaluation des performances des algorithmes d’estimation sur données simulées et réelles ;

- un détecteur d’évènements amélioré des instants de mise en marche et mise à l’arrêt des appareils.

Chapitre 4 : Dans ce chapitre une étude est présentée permettant de montrer l’apport du régime transitoire (courant de mise en marche) par rapport au régime établi pour l’identification des appareils électriques. Ce chapitre présente aussi le système de mesure que nous avons créé et les enjeux de sa création ainsi que la base de données de courants transitoires de mise en marche créée afin d’être utilisée pour la classification des appareils électriques.

La problématique de classification des signaux de notre base de données en utilisant les paramètres estimés du modèle que nous avons proposé est abordée. Cette problématique est discutée avec ses deux sous-problématiques : identification et clustering. Des conclusions quant à l’utilité des paramètres de notre modèle pour la classification sont tirées.

Remarque : La plupart des figures étant renseignées en anglais, une liste d’abréviations est fournie en début de manuscrit.

Contributions principales

Les défis que j’ai rencontrés durant ma thèse sont de trois natures : théoriques, algorithmiques ainsi que des défis liés à la mesure. Les contributions issues de mon travail de thèse visent à relever ces défis et peuvent être résumés comme suit :

1. proposition d’un nouveau modèle pour le courant transitoire de mise en marche (Nait Meziane et al., 2015a) ainsi que l’étude de ses performances théoriques (**Chapitre 3**) ;
2. proposition d’un algorithme haute précision pour la détection d’évènements (Nait Meziane et al., 2017b) (**Chapitre 3**) ;
3. proposition d’un algorithme d’estimation de paramètres basé sur les moindres carrés non linéaires pour le nouveau modèle proposé. On montre que cet algorithme est efficace dans le sens où les paramètres estimés atteignent les BCR (**Chapitre 3**) ;

4. preuve de concept de l'utilité du transitoire de mise en marche pour la tâche d'identification d'appareils électriques (Nait-Meziane et al., 2016) (**Chapitre 4**);
5. création d'un système de mesure ainsi qu'une base de transitoires de mise en marche (Nait Meziane et al., 2016a) (**Chapitre 4**).

Liste de publications

Actes de conférences internationales publiés

1. Nait Meziane, M., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016c). Accuracy evaluation of low-cost energy meters for home electrical consumption assessment. In *Conference on Environment and Electrical Engineering (EEEIC), 2016 Proceedings of the 16th IEEE International*, pages 2473–2478 (**Chapitre 2**).
2. Nait Meziane, M., Ravier, P., Lamarque, G., Abed-Meraim, K., Le Bunetel, J.-C., and Raingeaud, Y. (2015a). Modeling and estimation of transient current signals. In *Signal Processing Conference (EUSIPCO), 2015 Proceedings of the 23rd European*, pages 2005–2009 (**Chapitre 3**).
3. Nait Meziane, M., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2017b). High accuracy event detection for non-intrusive load monitoring. In *42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Accepted for presentation (**Chapitre 3**).
4. Nait-Meziane, M., Hacine-Gharbi, A., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016). Hmm-based transient and steady-state current signals modeling for electrical appliances identification. In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, pages 670–677 (**Chapitre 4**).
5. Nait Meziane, M., Picon, T., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016b). A new measurement system for high frequency nilm with controlled aggregation scenarios. In *Workshop on Non-Intrusive Load Monitoring (NILM), 2016 Proceedings of the 3rd International* (**Chapitre 4**).
6. Nait Meziane, M., Picon, T., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016a). A measurement system for creating datasets of on/off-controlled electrical loads. In *Conference on Environment and Electrical Engineering (EEEIC), 2016 Proceedings of the 16th IEEE International*, pages 2579–2583 (**Chapitre 4**).

Articles preprint et soumis à des conférences

1. Picon, T., Nait Meziane, M., Ravier, P., Lamarque, G., Novello, C., Le Bunetel, J.-C., and Raingeaud, Y. (2016). COOLL: Controlled on/off loads library, a public

dataset of high-sampled electrical signals for appliance identification. *arXiv preprint arXiv:1611.05803 [cs.OH]* (**Chapitre 4**)

2. Nait Meziane, M., Hacine-Gharbi, A., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2017a). Electrical appliances identification and clustering using novel turn-on transient features. In *6th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*. Submitted for review (**Chapitre 4**)

Article revue

En préparation : modèle de transitoire de mise en marche, BCRs, algorithm(e)s d'estimation de paramètres (**Chapitre 3**) et classification d'appareils électriques (**Chapitre 4**).

Communication interne avec comité de lecture

1. Nait Meziane, M., Ravier, P., Le Bunetel, J.-C., and Lamarque, G. (2015b). Identification of consumers on an electrical network by analyzing their load profile. In *Journée des Jeunes Chercheurs (JJC) du laboratoire Prisme, juin 2015*.

Chapitre 1

Contexte de l'étude

Sommaire

1.1	Motivation	8
1.1.1	Transition énergétique	8
1.1.2	Règlementation thermique 2012 (RT 2012)	11
1.1.3	Obligation de faire du sous-comptage	13
1.1.4	Projet MDE-MAC3	15
1.2	Désagrégation des consommations des appareils électriques	17
1.2.1	État de l'art succinct	17
1.2.2	Problématique du NILM	19
1.3	Démarche retenue	22

1.1 Motivation

1.1.1 Transition énergétique

L'ensemble des rencontres appelé « Grenelle Environnement » a abouti à une première loi (n° 2009-967 du 3 août 2009 de programmation relative à la mise en œuvre du Grenelle de l'environnement), dite *Grenelle 1*, votée le 23 juillet 2009 à la quasi-unanimité et promulguée le 03 Août 2009. Elle fixe les engagements de la France dans tous les domaines retenus lors du Grenelle, des transports à l'énergie, en passant par l'eau et la biodiversité¹. Une seconde loi (n° 2010-788 du 12 juillet 2010 portant sur l'engagement national pour l'environnement), dite *Grenelle 2* détaille les modalités d'application du Grenelle 1 par objectif, chantier, et secteur.

Cette stratégie d'aller vers le développement durable a forcément affecté la production et la consommation d'énergie en France. Le bilan énergétique de la France pour 2014 publié par le Service de l'Observation et des Statistiques (SOeS) du ministère de l'Environnement, de l'Énergie et de la Mer en juillet 2015 permet de voir les variations des tendances de consommation d'énergie durant ces dernières années. La figure 1.1 donne ces tendances pour le secteur englobant le résidentiel et le tertiaire, par forme d'énergie. Les résultats de consommation sont donnés en « Million de tonnes équivalent pétrole (Mtep) » où $1 \text{ Mtep} = 41,868 \text{ gigajoules (GJ)} (10 \text{ gigacalorie (Gc)})$ et correspond au pouvoir calorifique d'une tonne de pétrole. C'est une unité d'énergie qui sert, d'un point de vue économique et industriel, à comparer différentes formes d'énergies en utilisant une équivalence au Mtep pour chacune des formes (électricité, gaz, charbon, bois, etc.).

1. <http://www.developpement-durable.gouv.fr/La-premiere-loi-du-Grenelle.html>

1.1. MOTIVATION

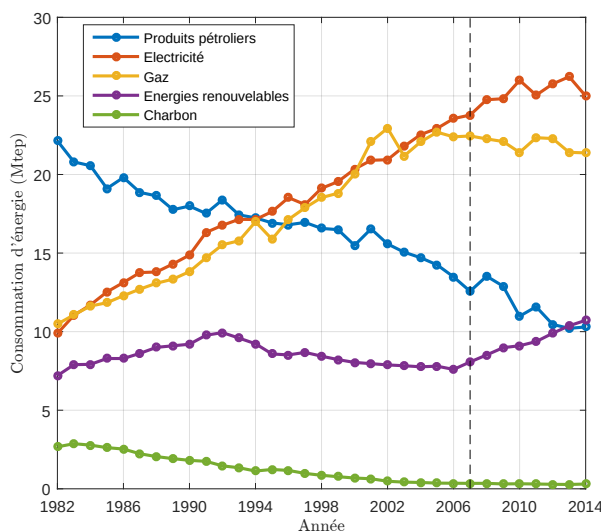


FIGURE 1.1 – Consommation finale d'énergie du secteur résidentiel-tertiaire en France. Source : données du SOeS, bilan de l'énergie 2014. La ligne de tirets verticaux marque la date du Grenelle 1.

Ces résultats montrent comment depuis 2007 la consommation en énergies renouvelables est en croissance régulière avec une consommation en gaz (naturel) qui n'augmente plus, résultats directs du Grenelle Environnement. La consommation en produits pétroliers (ne concernant que les produits raffinés : gazole, GPL, etc.) et charbon ne cesse de baisser depuis le début des années 80. Ceci s'explique en partie par une augmentation de la production électrique française liée à une politique de développement de la filière nucléaire (figure 1.2).

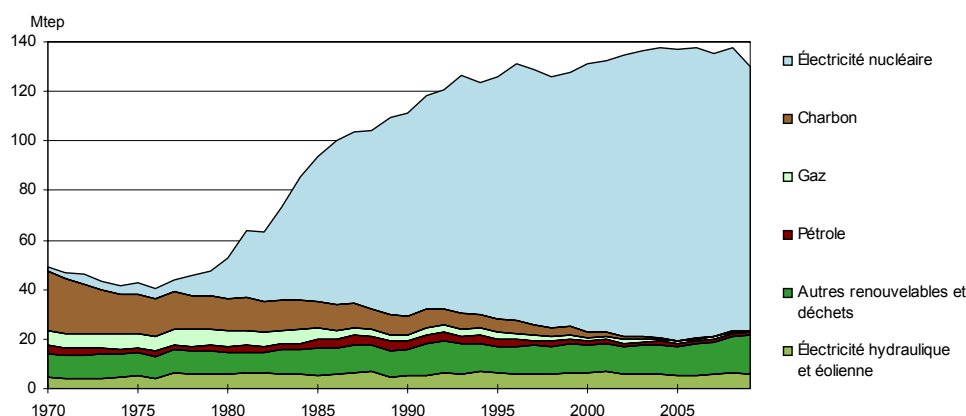


FIGURE 1.2 – Production d'énergie primaire en France. Source : SOeS, bilan de l'énergie 2009.

Une autre constatation intéressante est la croissance continue de la consommation d'électricité, même après 2007 (figure 1.1). Ceci indique qu'il faut, pour satisfaire aux

exigences du Grenelle Environnement, proposer des solutions d'**efficacité énergétique** électrique qui accompagneront cette croissance en consommation et balanceront l'équilibre offre-demande. En effet, une forte demande d'énergie électrique dépassant ponctuellement la capacité de production—on parle alors de pics de consommation—risque de causer des blackouts si aucune source de production alternative n'est disponible pour répondre à cette demande.

Le Réseau de transport d'électricité (RTE) publie annuellement un bilan prévisionnel de l'équilibre offre-demande d'électricité (RTE, 2016). On trouve dans ce bilan, entre autres, des informations sur la consommation d'électricité en France. L'évolution de la répartition des principaux secteurs dans la consommation d'électricité (figure 1.3) affiche depuis 2008 une prédominance des deux secteurs résidentiel et tertiaire, devenus les plus consommateurs d'énergie électrique comparés au secteur industriel. Pour l'année 2015 (figure 1.4) ces deux secteurs représentent presque les deux tiers de la consommation totale. Ceci montre à quel point les solutions d'efficacité énergétique peuvent être importantes pour ces deux secteurs.

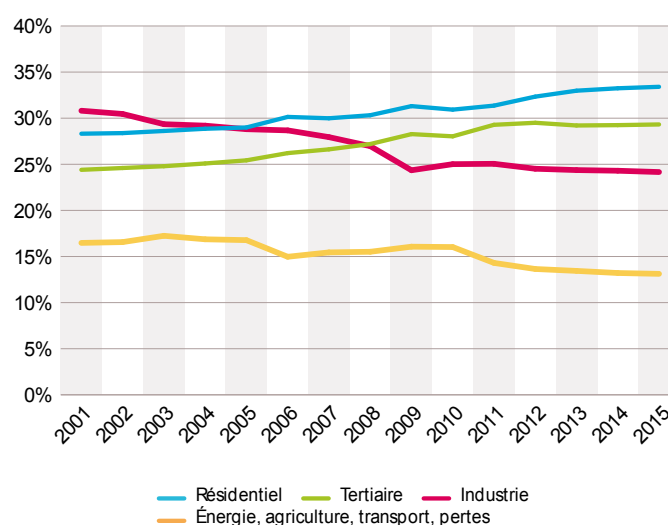


FIGURE 1.3 – Évolution de la répartition des principaux secteurs dans la consommation d'électricité en France. Source : (RTE, 2016).

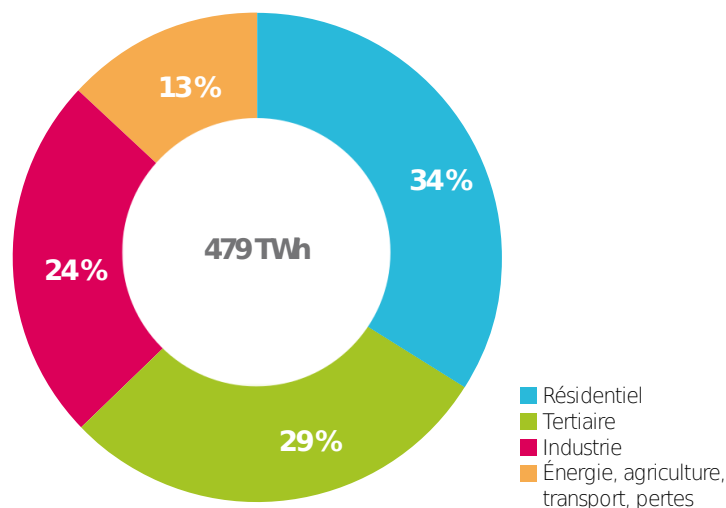


FIGURE 1.4 – Répartition sectorielle de la consommation d’électricité en France pour l’année 2015. Source : (RTE, 2016).

1.1.2 Règlementation thermique 2012 (RT 2012)

La réglementation thermique (RT) française a pour but de fixer une limite maximale à la consommation énergétique des bâtiments neufs. Avant la RT en vigueur il y en a eu quatre autres : RT 1974, RT 1988, RT 2000 et RT 2005. Ne s’appliquant au début qu’au résidentiel (RT 1974) puis s’élargissant au tertiaire (RT 1988), imposant ensuite des consommations maximales (RT 2000), puis exigeant l’amélioration de la performance thermique également et s’attaquant aussi à la rénovation (RT 2005), la RT n’a cessé d’évoluer afin de réguler au mieux la consommation du bâtiment.

Le Grenelle Environnement est le socle de la réglementation thermique en vigueur, à savoir la **RT 2012**, spécifiée par les deux textes suivants : le décret n° 2010-1269 du 26 octobre 2010 *relatif aux caractéristiques thermiques et à la performance énergétique des constructions* et l’arrêté du 26 octobre 2010 *relatif aux caractéristiques thermiques et aux exigences de performance énergétique des bâtiments nouveaux et des parties nouvelles de bâtiments*, ayant pour objet la **fixation d’exigences sur les caractéristiques thermiques et la performance énergétique des bâtiments neufs**.

Elle s’applique (décret n° 2010-1269 du 26 octobre 2010 publié le 27 octobre 2010) :

- à tous les permis de construire déposés plus d’un an après la date de publication (c.-à-d. à partir du 28 octobre 2011) du décret pour les bâtiments neufs à usage de

bureaux ou d'enseignement, les établissements d'accueil de la petite enfance et les bâtiments à usage d'habitation construits en zone ANRU².

- à tous les permis de construire déposés à partir du 1^{er} janvier 2013 pour les autres bâtiments neufs à usage d'habitation.

Les exigences de résultats imposées par la RT 2012 sont de trois types³ :

1. L'efficacité énergétique du bâti :

L'exigence d'efficacité énergétique minimale du bâti est définie par le coefficient « **Bbiomax** » (besoins bioclimatiques du bâti). Cette exigence impose une limitation simultanée du besoin en énergie pour les composantes liées à la conception du bâti (chauffage, refroidissement et éclairage), imposant ainsi son optimisation indépendamment des systèmes énergétiques mis en œuvre.

2. La consommation énergétique du bâtiment :

L'exigence de consommation conventionnelle maximale d'énergie primaire se traduit par le coefficient « **Cepmax** », portant sur les consommations de chauffage, de refroidissement, d'éclairage, de production d'eau chaude sanitaire et d'auxiliaires (pompes et ventilateurs). Conformément à l'article 4 de la loi Grenelle 1, la valeur du Cepmax s'élève à 50 kWh/(m².an) d'énergie primaire, modulé selon la localisation géographique, l'altitude, le type d'usage du bâtiment, la surface moyenne des logements et les émissions de gaz à effet de serre pour le bois énergie et les réseaux de chaleur les moins émetteurs de CO₂. Cette exigence impose, en plus de l'optimisation du bâti exprimée par le Bbio, le recours à des équipements énergétiques performants, à haut rendement.

3. Le confort d'été dans les bâtiments non climatisés :

À l'instar de la RT 2005, la RT 2012 définit des catégories de bâtiments dans lesquels il est possible d'assurer un bon niveau de confort en été sans avoir à recourir à un système actif de refroidissement. Pour ces bâtiments, la réglementation impose que le coefficient « **Tic** » (Température intérieure conventionnelle : température maximale atteinte à l'intérieur du bâti lors d'une séquence de 5 jours consécutifs de forte chaleur) n'excède pas un seuil.

2. Zone définie prioritaire par l'Agence Nationale du Renouvellement Urbain pour la réhabilitation de quartiers : ces zones sont traitées de manière globale (logements, voiries, espaces public...)

3. <http://www.rt-batiment.fr/batiments-neufs/reglementation-thermique-2012/presentation.html>

1.1. MOTIVATION

La prochaine réglementation thermique **RT 2020** vise à aller encore plus loin avec le concept du bâtiment à énergie positive (BEPOS). Cela veut dire qu'un bâtiment devra produire de l'électricité plus qu'il n'en consomme.

1.1.3 Obligation de faire du sous-comptage

L'article 23 de l'arrêté du 26 octobre 2010 demande que les bâtiments neufs soient équipés de systèmes permettant la mesure ou l'estimation de la consommation d'énergie de chaque logement et que l'information fournie par ces systèmes soit délivrée par type d'énergie, *a minima* selon la répartition suivante ([Centre Scientifique et Technique du Bâtiment \(CSTB\), 2013](#)) :

- **chauffage** :

Comprend l'énergie consommée et facturée permettant le chauffage du logement. Dans le cas d'un générateur individuel, les consommations liées aux auxiliaires sont comptabilisées dans ce poste ; elles peuvent également être comptabilisées dans le poste « autres ». Dans le cas d'un générateur collectif, les auxiliaires ne sont pas comptabilisés.

- **refroidissement** :

Comprend l'énergie consommée et facturée permettant le refroidissement du logement. Dans le cas d'un générateur individuel, les consommations liées aux auxiliaires sont comptabilisées dans ce poste ; elles peuvent également être comptabilisées dans le poste « autres ». Dans le cas d'un générateur collectif, les auxiliaires ne sont pas comptabilisés.

- **production d'eau chaude sanitaire (ECS)** :

Comprend l'énergie consommée et facturée permettant la production d'eau chaude sanitaire du logement. Dans le cas d'un générateur individuel, les consommations liées aux auxiliaires sont comptabilisées dans ce poste ; elles peuvent également être comptabilisées dans le poste « autres ». Dans le cas d'un générateur collectif, les auxiliaires ne sont pas comptabilisés.

- **réseau prises électriques** :

Comprend les consommations de tous les appareils reliés aux prises électriques : appareils électro-ménagers, informatique, hi-fi, lampes d'appoint, cuisson électriques

en pose libre (four micro-onde, plaques électriques portables, ...), etc., à l'exception des circuits spécialisés (au sens de la NFC 15-100) destinés au pôle cuisson, non reliés à une prise, dont les consommations sont comptées dans le poste « autres » (ex : four avec une ligne électrique dédiée).

– **autres :**

Comprend les consommations qui ne sont pas prises en compte dans les postes précédents : éclairage immobilier, circuits spécialisés pour plaques de cuisson et four électriques, cuisson autre qu'électrique, VMC, automatismes, etc.

Quel que soit le poste, l'énergie renouvelable gratuite (ex : énergie solaire thermique) n'est pas comptabilisée. En logement collectif, les consommations qui ne sont pas directement liées à celles du logement (éclairage extérieur et des parties communes, motorisation des portes de garage, blocs autonomes d'éclairage de sécurité, centrale de détection incendie, VMC collective, ...) ne sont pas à prendre en compte.

La consommation corrigée d'électricité du secteur résidentiel s'élève à 159,8 TWh en France continentale en 2015 (figure 1.5) dont 45 % représentent les usages « domestiques ». Les usages domestiques regroupent les usages froid et lavage (blanc), technologies de l'information et de la communication (TIC), cuisson et éclairage. Les autres usages résidentiels regroupent l'agrégation des usages diffus comme les aspirateurs, fer à repasser, piscines, aquariums, etc. (RTE, 2016).

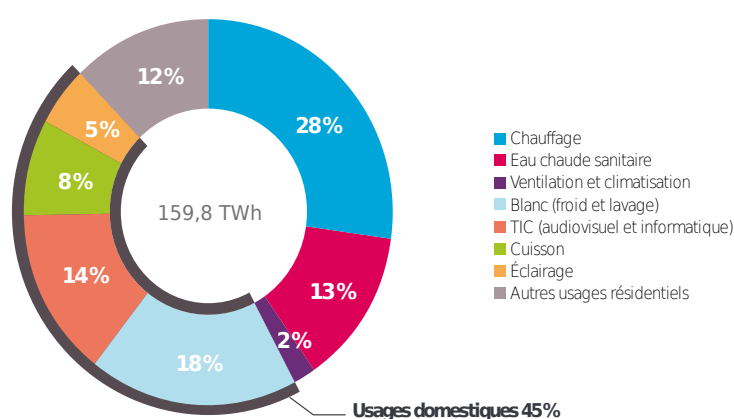


FIGURE 1.5 – Répartition par usage de la demande électrique résidentielle en France pour l'année 2015. Source : (RTE, 2016).

1.1.4 Projet MDE–MAC3

Notre projet appelé « Maîtrise de la Demande d’Energie–Mesure Aveugle de la Consommation des Charges au Compteur (MDE–MAC3) » vient s’insérer dans le contexte général du développement durable et est motivé, entre autres, par le besoin de sous-comptage émergeant de la RT 2012.

La maîtrise de la demande de l’énergie (MDE) a pour vocation de lisser les pics de consommation pour les producteurs et de réduire cette consommation pour les ménages et industries. Les solutions actuelles passent par les smart grids permettant de mieux gérer l’offre et la demande par les technologies de l’information et de la communication. Selon la mission d’observation smart grid aux Etats-unis qui regroupe les adhérents des quatre pôles français spécialistes de l’énergie : CAPENERGIES, DERBI, S2E2 et TENERRDIS, « la conception américaine du smart grid a évolué considérablement en l’espace d’un an. En effet, la majorité des premiers programmes américains étaient centrés sur le déploiement de compteurs intelligents ; le retour d’expérience montre que ces compteurs sont loin d’être suffisants pour garantir un réseau électrique plus fiable. L’approche américaine des smart grids, très orientée technologies de l’information et de la communication, souvent baptisée Big Data, semble évoluer vers une approche plus matérielle avec des produits permettant d’agir sur la production, la transmission et la consommation de l’énergie ».

Pour cela, il est nécessaire d’avoir une vision précise de la demande : l’instant précis de la demande, l’identification du consommateur et la connaissance du niveau de puissance. Pour accéder à ces informations, il existe aujourd’hui des solutions industrielles partielles (WATTECO, FLUIDIA, AVIDSEN, PlugWise, etc.). Les solutions complètes sont coûteuses et difficiles à mettre en œuvre au sens du plug&play. C’est la raison pour laquelle, dans le projet Feder AFFICHECO (Économies d’énergie liées à l’affichage des consommations électriques en résidentiel), les consommations ne sont relevées que sur quelques lignes. De ce fait, les résultats peuvent manquer de pertinence pour le consommateur final. Comment changer ses habitudes si on ne connaît pas précisément la part éclairage vis-à-vis de la consommation du téléviseur ou du réfrigérateur ? Ai-je intérêt à changer mon vieux réfrigérateur ? L’instrumentation des habitations retenue dans AFFICHECO illustre le défaut de solutions technologiques pour répondre à ce problème.

Le projet MDE–MAC3 est complémentaire à AFFICHECO car apporte des informations plus fines sur les habitudes de consommation dans l’habitat et ouvre la voie à des systèmes de délestage intelligents : je déleste, par exemple, mon radiateur de cuisine à l’instant où je mets en marche mon four. De plus, associé à d’autres capteurs, comme ceux de présence humaine développés par le laboratoire PRISME dans le projet CAPTHOM, on peut envisager de délester le radiateur d’une pièce non occupée ce qui évite un paramétrage toujours délicat des plages de fonctionnement des différents radiateurs. Dans le même ordre d’idée, des initiatives ont été lancées en Bretagne où le problème des pics de consommation est particulièrement critique en raison des risques de blackout. Un dispositif d’alertes Orange-écowatt, lancé par le distributeur RTE le 12 janvier 2008, avertit le particulier de ces fortes consommations par des alertes SMS. Il est du ressort de la population d’agir en ces périodes critiques. Ce principe peut être finement amélioré grâce à ce projet puisque, étant donné que les appareils connectés au réseau peuvent être identifiés, il est tout à fait possible d’arrêter ou limiter la consommation des appareils jugés non prioritaires (appareils de lavage, de chauffage, sèche-linge, évalués aux 2/3 de la consommation électrique). Dans ces situations, plutôt que de chercher à modifier les habitudes des consommateurs, la technique vient en substitut à l’action, dépendante de la bonne volonté de l’utilisateur.

Avoir une information fine de la consommation d’énergie électrique peut aussi servir d’autres finalités, mis à part celles citées ci-dessus, et être exploitée dans d’autres domaines d’application. On pourrait bien imaginer que cette information soit intégrée dans une solution pour la surveillance à distance des personnes âgées (un ménage occupé devrait forcément avoir une activité électrique différente d’un autre qui est vide).

D’un point de vue maintenance, cette information fine, enregistrée sur de longues durées, pourrait constituer une référence à utiliser pour suivre l’usure des appareils électriques. Elle peut aussi être exploitée afin de connaître les habitudes des consommateurs afin que les fournisseurs puissent proposer une tarification adaptée aux besoins de ses clients.

La partie Mesure Aveugle de la Consommation des Charges au Compteur (MAC3) précise l’approche qu’on a choisi pour avoir cette information fine de consommation. Cette approche est basée sur la mise en place d’un seul dispositif de mesure au niveau du

compteur électrique principal (d'un ménage par exemple). Ce dispositif va mesurer une mesure globale de toutes les charges se trouvant sur le point global de mesure d'où le caractère aveugle vu qu'on mesure un agrégat de consommations sans voir directement la consommation individuelle de chaque consommateur.

D'un point de vue installation électrique basse tension, il existe la norme NF C15-100 (norme d'application obligatoire, version juin 2015) qui spécifie les règles à suivre pour faire son installation électrique. En particulier, cette norme exige l'existence de circuits spécialisés pour certains usages. Par exemple, chaque appareil électroménager de forte puissance doit être alimenté par un circuit spécialisé, il doit exister un circuit spécialisé pour l'alimentation de la cuisinière ou de la plaque de cuisson seule, il faut avoir trois circuits spécialisés 16 A 2P+T au moins en prévision de l'alimentation d'appareils du type lave-linge, lave-vaisselle, sèche-linge, four, etc.

De ce fait, avoir de tels circuits spécialisés va simplifier la mesure par usage de la consommation que demande la RT 2012. On peut mettre un capteur sur chaque circuit spécialisé pour mesurer la consommation des usages qui lui sont reliés, mais cela nécessite autant de capteurs que de circuits spécialisés. Même si une grande partie des gros consommateurs d'un ménage devrait avoir un circuit spécialisé, il reste une bonne partie des appareils électriques, notamment ceux branchés sur le réseau des prises électriques, dont la consommation individuelle ne peut être obtenue que par une solution de sous-comptage telle la MAC3. De plus, l'obligation d'avoir des circuits spécialisés n'est pas en vigueur pour les bâtiments anciens, secteur dominant pour lequel la MAC3 a tout son intérêt.

1.2 Désagrégation des consommations des appareils électriques

1.2.1 État de l'art succinct

Pour décomposer, ou désagréger, la consommation globale d'énergie en ses différents éléments constitutifs (par exemple, l'ensemble de la consommation d'énergie d'un ménage

en consommation individuelle par appareil) trois approches peuvent être envisagées : approches intrusives, semi-intrusives et non intrusives. D’après (Ehrhardt-Martinez et al., 2010) les études faites sur l’utilité d’un retour (feedback) en temps réel de l’information désagrégée de consommation d’énergie montrent que cela permet de réaliser de 9 à 18 % d’économie sur la facture d’énergie.

Les approches de suivi intrusif de la charge (Intrusive Load Monitoring–ILM) introduisent une forme d’intrusion au niveau du réseau à mesurer car s’appuient sur l’utilisation de sous-compteurs distribués (Burbano Acuña, 2015), placés directement au niveau des appareils, qui envoient cette information à une unité centrale de traitement pour l’inclure dans la consommation globale d’énergie. Comme indiqué dans (Parson, 2012), différentes formes d’ILM peuvent être trouvées :

- sous-comptage électrique : un calculateur d’énergie est utilisé pour chaque appareil ;
- appareils intelligents : les appareils ont des circuits communicants qui leur permettent de reporter leur consommation à une unité centrale ;
- sondage électrique : un signal électrique est injecté au réseau domestique et le signal retourné est analysé. L’intrusivité vient ici du fait d’injecter au réseau un signal extérieur qui pourrait affecter la qualité du réseau.

La première forme de cette approche intrusive est une approche difficilement mise en place à grande échelle. Avoir à s’introduire chez les ménages pour placer des sous-compteurs partout chez eux nécessite de déployer du personnel et un effort considérable en plus du coût élevé. De plus, chaque nouvel appareil nécessite l’achat d’un nouveau sous-compteur. Il est aussi nécessaire d’associer le sous-compteur au nouvel appareil, ce qui rajoute une complication pour l’habitant surtout pour les appareils mobiles (ordinateur, aspirateur, etc.).

La deuxième forme d’intrusion est attractive. Par contre, elle nécessite d’équiper les appareils de puces à la fois pour le calcul d’énergie et l’envoi des données. À l’état actuel, il est difficile d’imaginer une telle solution. En effet, les constructeurs d’appareils ne sont pas prêts aujourd’hui à définir des protocoles et normes pour de telles puces. Il se pose également le problème des données qui peuvent porter atteinte à la vie privée des consommateurs.

La troisième forme d’intrusion est écartée du fait qu’elle interfère avec le réseau et

ajoute une perturbation qui pourra influencer la qualité de l'énergie envoyée aux appareils consommateurs.

Le suivi semi-intrusif de la charge (Semi-Intrusive Load Monitoring–SILM) est une version décontractée de l'approche intrusive, où au lieu d'utiliser des sous-compteurs d'énergie à raison de un par appareil, on utilise un sous-compteur pour mesurer un groupe d'appareils (Tang et al., 2014a). Cette approche partage les mêmes inconvénients que l'approche intrusive, mais avec moins d'intrusion. Dans la littérature, ces approches sont souvent confondues avec les approches intrusives.

L'approche non-intrusive (Non-Intrusive Load Monitoring–NILM), correspondant au MAC3, se préoccupe de proposer des solutions autorisant une décomposition de la consommation globale sans avoir recours, ou peu, à une intrusion en utilisant un seul capteur placé au niveau du compteur principal d'électricité.

1.2.2 Problématique du NILM

Le travail dans ce domaine, appelé aussi « désagrégation d'énergie » (energy disaggregation), a commencé avec Hart au milieu des années 1980 (Hart, 1985, 1989, 1992) où il a proposé d'utiliser la variation de la puissance active pour identifier les appareils ménagers.

Même si d'autres travaux ont été faits après Hart entre le début des années 1990 et la fin des années 2000 (Sultanem, 1991; Leeb et al., 1995; Cole and Albicki, 1998a; Drenker and Kader, 1999; Chan et al., 2000; Baranski and Voss, 2003; Laughman et al., 2003; Ting et al., 2005; Patel et al., 2007; Chang et al., 2008), ce n'est qu'il y a environ cinq ans que ce domaine a commencé à connaître un intérêt croissant (figure 1.6).

Un état de l'art des méthodes NILM peut être trouvé dans (Najmeddine et al., 2008; Du et al., 2010; Zeifman and Roth, 2011). Dans (Carrie Armel et al., 2013), les auteurs donnent un tableau qui résume les travaux effectués dans le domaine du NILM jusqu'à 2012. Ils ont organisés ces travaux selon un ordre chronologique et ont donné des détails sur les méthodes utilisées, le type des données et des appareils, la fréquence d'échantillonnage ainsi que d'autres caractéristiques intéressantes.

Notons aussi que quelques thèses de doctorat effectués en France se sont déjà intéressées

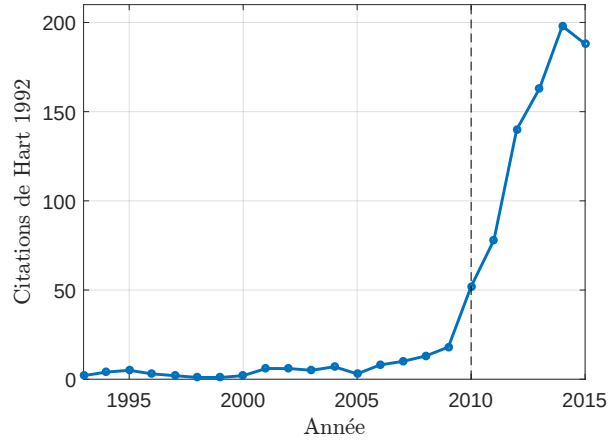


FIGURE 1.6 – Citations du travail fondateur du NILM de Hart ([Hart, 1992](#)) entre 1993 à 2015. Source des données : Citations Google Scholar.

à la problématique du NILM. Nous citons par exemple ([El Guedri, 2009](#); [Najmeddine, 2009](#); [Sanquer, 2013](#); [Bier, 2014](#); [Amirach, 2015](#); [Klein, 2015](#)).

Qui, quand et combien ?

Un système NILM essaye de répondre à trois questions principales : qui consomme ? quand ? et combien consomme-t-il ? Savoir qui consomme relève de l'*identification* du type de consommateur. Savoir quand est-ce qu'il consomme revient à faire une *détection* du moment de début et de fin de consommation. Cette étape de détection peut être nécessaire pour identifier le consommateur suivant si l'approche choisie est basée sur la détection d'évènements ou pas. Calculer ou *estimer la consommation* constitue en général le dernier étage d'un système NILM et est conditionné par les étages de détection et d'identification.

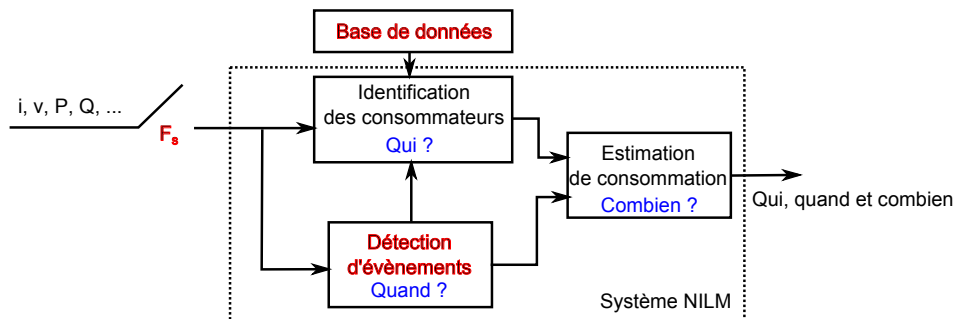


FIGURE 1.7 – Schéma général d'un système NILM. En rouge les trois critères selon lesquels on peut classer les approches NILM.

Un système NILM est donc essentiellement constitué de ces trois étages : détection,

identification et estimation de la consommation (figure 1.7).

On trouve essentiellement trois dichotomies possibles pour classer les approches NILM (se référer aux parties en rouge sur la figure 1.7) :

- approche supervisée *vs.* non supervisée (base de donnée utilisée ou pas) ;
- approche basse *vs.* haute fréquence d'échantillonnage (F_s faible ou haute) ;
- approche basée évènement *vs.* non basée évènement (détecteur d'évènement utilisé ou pas).

Les approches supervisées ont besoin d'un apprentissage ou entraînement afin que le système puisse apprendre les appareils qu'il est censé identifier avant le déploiement pour une utilisation réelle, ce qui ajoute plus de complexité pour de telles approches. Les approches non supervisées (Bonfigli et al., 2015), en contrepartie, sont exemptes de cette contrainte, mais sont plus sujettes à des erreurs liées à la variabilité de l'environnement. Une approche à mi-chemin entre ces deux dernières est l'approche semi-supervisée (Barsim and Yang, 2015; Gillis and Morsi, 2016) où une étape d'apprentissage (supervisé) vient aider la prédiction de la classe d'un nouvel appareil n'ayant pas d'étiquette (non supervisé).

D'un point de vue fréquence d'échantillonnage, les approches NILM peuvent être classées, grossièrement, en approches à basse fréquence d'échantillonnage (BFE) (1 Hz ou moins) et approches à haute fréquence d'échantillonnage (HFE) (quelques centaines de Hz au MHz). En raison du fait que la mesure basse fréquence est plus accessible et que les bases de données basse fréquence sont plus facilement disponibles, le NILM-BFE est plus répandu. Clark (Clark, 2015) a donné un tableau qui résume les travaux NILM organisés par fréquence d'échantillonnage, que nous reprenant ici (voir tableau 1.1) avec une traduction en français en mettant les références citées dans la bibliographie pour que le lecteur puisse s'y référer.

Une autre dichotomie du domaine est la subdivision en NILM basé événements (BE) *vs.* NILM non basé évènement (NBE) suivant si les méthodes dépendent de la détection des évènement ou pas (Anderson et al., 2012b; Zhao et al., 2016).

Dans les approches BE on trouve en général les étages suivants : la détection d'évènements, l'extraction des paramètres, le clustering des classes d'appareils et l'estimation

de la consommation d'énergie. L'étage de détection d'événements est particulièrement important vu qu'il se situe à l'amont du système NILM.

La plupart des travaux des approches BE utilisent les puissances réelle et réactive comme entrée pour l'étage de détection (De Baets et al., 2016; Abraha Girmay and Camarda, 2016). Néanmoins, quelques travaux utilisent d'autres signaux comme la distorsion de la tension (Cox et al., 2006) ou le courant électrique (Wild et al., 2015).

1.3 Démarche retenue

Dans le contexte général présenté ci-dessus et dans le cadre du NILM, on ne prétend pas résoudre le problème dans sa globalité (c.-à-d. proposer un système NILM complet). Nous allons nous concentrer sur la proposition d'un modèle pour les courants transitoires de mise en marche et estimer ses paramètres afin de les exploiter dans l'étape d'identification d'appareils. On propose aussi un détecteur simple et rapide à utiliser comme détecteur d'événements.

Pour ce travail de thèse on a été amené à faire certains choix afin d'aborder la problématique du NILM. Nous avons choisi d'échantillonner entre 2 kHz et 40 kHz car une telle fréquence permet d'accéder à des informations perdues en basse fréquence, notamment la dynamique de l'instant de mise en marche. Aller plus haut en fréquence compliquerait plus la tâche pour deux raisons essentielles : cela ajoute des contraintes pour la sauvegarde des données et les signatures deviennent dépendantes de l'emplacement physique de l'appareil connecté au réseau électrique domestique à cause des modifications liées au canal de transmission (de la prise au compteur) ce qui affectera négativement la robustesse de l'identification.

En terme de choix de signature, différentes possibilités s'offrent à nous : puissance active et réactive, harmoniques, tension, courant, etc. Il est fort possible de tomber sur des cas où les puissances (que ce soit actives ou réactives) sont égales pour des appareils de type différent ce qui compliquera leur identification. La tension, quant à elle, est théoriquement supposée rester invariable (une sinusoïde pure de fréquence égale à la fréquence fondamentale) et donc n'est pas adaptée pour une tâche d'identification où on cherche

1.3. DÉMARCHE RETENUE

à identifier des saillances dans le signal qui peuvent être exploitées pour distinguer et séparer les appareils entre eux.

Pour ces raisons, nous avons choisi de travailler avec le courant transitoire de mise en marche dont la forme dépend du phénomène physique responsable du fonctionnement de l'appareil, ce qui le distingue donc des autres. En se basant sur cela, l'idée est d'exploiter ce transitoire de mise en marche et essayer d'extraire des paramètres distinctifs qui reflètent ces phénomènes physiques liés au fonctionnement.

Ayant des *a priori* forts sur les courants de mise en marche tels que la périodicité, les harmoniques les plus significatives, etc. nous proposons pour l'identification un modèle mathématique paramétrique modélisant les courants de mise en marche en exploitant ces *a priori*.

Avant de présenter le modèle proposé ainsi que l'identification des appareils électriques, on aborde dans le chapitre suivant la comparaison de la qualité de mesure de quelques calculateurs d'énergie grand public.

TABLE 1.1 – Résumé des méthodes NILM organisées par fréquence d'échantillonnage.
Source : (Clark, 2015)

Nom	Fréquence d'échantillonnage	Caractéristique principale	Études
Très basse	Moins d'une fois par minute	Souvent plusieurs événements par intervalle	(Powers et al., 1991) (Prudenzi, 2002) (Kolter et al., 2010) (Huang et al., 2013) (Curtis et al., 2014)
Basse	1 minute à 1 seconde	Typiquement un événement par intervalle	(Hart, 1992) (Cole and Albicki, 1998b,a) (Drenker and Kader, 1999) (Farinaccio and Zmeureanu, 1999) (Marceau and Zmeureanu, 2000) (Baranski and Voss, 2003, 2004b,a) (Ruzzelli et al., 2010) (Kim et al., 2011) (Kelly and Knottenbelt, 2011) (Figueiredo et al., 2012) (Kolter and Jaakkola, 2012) (Parson et al., 2012) (Zoha et al., 2013) (Makonin et al., 2014) (Egarter et al., 2015) (Arberet and Hutter, 2014) (Tang et al., 2014b) (Liao et al., 2014) (Perez et al., 2014) (Elhamifar and Sastry, 2015)
Moyenne	Plus rapide que 1 Hz mais plus faible que la fréq. fondamentale	Les formes des transitoires sont analysées	(Norford and Leeb, 1996) (Berges et al., 2009) (Goncalves et al., 2011) (Barsim et al., 2014) (Du et al., 2015)
Haute	Fréq. fondamentale à 2 kHz	Harmoniques d'ordre faible sont analysées	(Sultanem, 1991) (Roos et al., 1994) (Leeb et al., 1995) (Lee et al., 2005) (Srinivasan et al., 2006) (Akbar and Khan, 2007) (Wichakool et al., 2015) (Dong et al., 2013)
Très haute	Plus rapide que 2 kHz jusqu'à 40 kHz	Forme d'onde avec haut détail ou harmoniques hautes sont en général analysées	(Chan et al., 2000) (Laughman et al., 2003) (Lee et al., 2004) (Lam et al., 2007) (Shaw et al., 2008) (Suzuki et al., 2008) (Liang et al., 2010a,b) (Inagaki et al., 2011) (Wang et al., 2012) (Popescu et al., 2014)
Ultra haute	Plus rapide que 40 kHz	Bruit électrique et très hautes fréquences sont analysés	(Patel et al., 2007) (Gupta et al., 2010)

Chapitre 2

Mesure de la consommation électrique

Sommaire

2.1	Introduction	26
2.2	Approches de comparaison	27
2.2.1	Mesures de puissance	28
2.2.2	Mesures d'énergie	31
2.3	Résultats expérimentaux	32
2.3.1	Puissance calculée en utilisant des mesures de laboratoire comme référence	32
2.3.2	Mesures d'énergie avec le compteur intelligent Linky comme référence	37
2.4	Conclusion	39

2.1 Introduction

Les Calculateurs d'Énergie à Bas Coût (CEBC) peuvent être utilisés pour obtenir la consommation électrique de chaque appareil de la maison. Ils sont connectés entre la prise de courant et l'appareil à mesurer ce qui permet d'accéder à la consommation électrique individuelle.

L'importance de la mesure et de l'information détaillée sur la consommation a été discutée dans plusieurs travaux antérieurs (Fischer, 2008; Darby, 2006, 2010; Hancke et al., 2012). Toutes ces études conviennent qu'une information sur détaillée sur la consommation électrique est nécessaire pour faire des économies d'énergie. Les avantages d'une information de consommation détaillée par appareil peuvent se classer en trois catégories (Pratt et al., 2010; Carrie Armel et al., 2013) :

- avantages pour le consommateur grâce à des économies d'énergie possible suite aux retours qu'il a sur sa consommation détaillée ;
- avantages pour la recherche et le développement puisque le retour sur la consommation permet la compréhension des profils de consommation d'énergie des appareils et ainsi revoir leurs conception pour qu'ils répondent aux exigences d'efficacité énergétique ;
- avantages pour les fournisseurs d'électricité et pour les gestionnaires des réseaux de transport et de distribution car permettant d'améliorer les prévisions de charge (Feinberg and Genethliou, 2005).

Pour utiliser les CEBC en tant que pilote d'information pour les retours sur la consommation électrique, on doit évaluer leur fiabilité de calcul de la consommation électrique. L'étude présentée dans ce chapitre donne une comparaison pour quatre différents CEBC disponibles dans le commerce.

La consommation électrique d'un appareil est donnée par les CEBC comme un affichage numérique en Watt (W) pour la puissance ou en kilowatt-heure (kWh) pour l'énergie. Les CEBC sont des calculateurs d'énergie (CE) bas coût (on confondra CE avec CEBC dans la suite), ce qui les rend accessibles aux ménages à faibles et moyens revenus.

Le tableau 2.1 donne le coût et la résolution d'affichage des quatre CE considérés dans cette étude.

TABLE 2.1 – Coût et résolution d’affichage des Calculateurs d’énergie (CE) considérés

	CE1	CE2	CE3	CE4
Coût (Euros)	30	19	13	30
Résolution d’affichage (Watts)	0.1	1	1	0.1

2.2 Approches de comparaison

Afin de comparer la précision de mesure des différents CE, on doit d’abord définir la quantité d’intérêt ainsi que la référence avec laquelle on compare. Dans cette étude, on a choisi la puissance et l’énergie comme quantités d’intérêt et on compare la précision de chaque mesure affichée, sur la CE, à une référence adéquate. Une bonne référence est une référence qui est suffisamment fiable pour être considérée comme représentant la valeur exacte.

Une référence simple de puissance pourrait être la puissance nominale des différents appareils. Une autre option peut être la puissance calculée en utilisant des mesures acquises dans le laboratoire avec une carte d’acquisition. La référence la plus fiable devrait être un analyseur de puissance, malheureusement non disponible au cours de l’étude.

Concernant la référence d’énergie, nous avons privilégié « Linky » : le compteur intelligent conçu pour remplacer la plupart des compteurs d’électricité actuellement déployés dans toute la France d’ici la fin 2021.

Les approches considérées pour la comparaison peuvent être résumées, en donnant d’abord la quantité d’intérêt puis les références choisies, comme suit :

1. mesures de puissance avec les références suivantes :
 - (a) la puissance nominale d’un appareil ;
 - (b) la puissance calculée en utilisant des mesures de laboratoire en considérant deux protocoles :
 - i. utilisant un CE à la fois,
 - ii. utilisant tous les CE simultanément,
2. mesures d’énergie avec le compteur intelligent Linky comme référence.

Dans la suite, on donne plus de détails pour les différentes approches considérées.

2.2.1 Mesures de puissance

Puissance nominale comme référence

La puissance nominale indiquée sur l'étiquette d'un appareil électrique est la première information disponible quand on veut répondre à la question : « quelle puissance consomme cet appareil ? ». La puissance nominale est une quantité très utile car elle donne une indication directe sur la consommation en puissance d'un appareil. Néanmoins, une récente enquête française¹ faite sur environ 3400 appareils ménagers trouvés soit dans les supermarchés, les magasins ou en ligne, a montré que près de 20 % des appareils des supermarchés et magasins et environ 10 % de ceux en ligne n'avaient pas d'étiquette ! La conclusion est que trouver l'indication de puissance nominale n'est pas toujours possible.

En outre, les paramètres réseau peuvent avoir une influence non négligeable sur les puissances réellement consommées. En effet, le constructeur affiche des puissances nominales sous certaines conditions standards (en général une valeur efficace de tension spécifique et une fréquence spécifique—par exemple 230 V et 50 Hz pour l'Europe et la plupart des pays du monde). Le problème est que, pour un réseau électrique ces valeurs fluctuent et donc la consommation d'un appareil peut dévier de sa puissance nominale.

Enfin, une autre raison importante est la chute de tension qui apparaît du fait des variations d'impédance du réseau, proportionnelle à l'amplitude du courant que l'appareil consomme. Ainsi, la tension appliquée aux bornes de l'appareil aura une valeur plus petite que la valeur standard, et par conséquent la consommation réelle en puissance sera généralement inférieure à la puissance nominale.

Puissance calculée en utilisant des mesures de laboratoire comme référence

Un autre choix de référence pour la consommation en puissance est un calcul en utilisant des mesures faites au laboratoire. Pour cette étude, les mesures ont été prises avec la carte d'acquisition Gage CompuScope 12400 digitizer (CompuScope 12400, 2008). Cette carte a une résolution de 12 bits, deux canaux d'acquisition, une fréquence d'échantillon-

1. <http://www.quechoisir.org/action-ufc-que-choisir-enquete-sur-l-etiquette-energie-la-reglementation-court-circuitee-n13069/>

nage variable de 50 kHz à 400 MHz et une largeur de bande élevée de 200 MHz. Le capteur de tension est une sonde différentielle Metrix MX 9030-Z ayant une bande passante de 30 MHz et le capteur de courant est une pince de courant Metrix HX0102 avec un plus faible courant mesurable de 3 mA, et une bande passante de 60 kHz. Pour le taux d'échantillonnage, on a choisi 50 kHz (la plus petite valeur possible qui peut être sélectionnée sur la carte d'acquisition) puisque on est intéressé par des mesures de puissance basse fréquence. Ces mesures de tension et de courant sont ensuite utilisées pour calculer la consommation en puissance, à des instants τ , en utilisant la formule :

$$P = \frac{1}{kT} \int_{\tau}^{\tau+kT} v \cdot i \, dt, \quad (2.1)$$

recommandée par la norme IEEE 1459-2010 ([IEEE Std 1459-2010, 2010](#)). v est la tension aux bornes de l'appareil et i le courant qu'il consomme. kT est la durée sur laquelle la moyenne est calculée et est un nombre entier de la période secteur T . Après avoir défini la référence, on va maintenant définir le protocole de mesure à utiliser pour comparer les différents CE. À cet effet, on définit deux protocoles de mesure. Dans le premier, on compare, pour chaque CE connecté à un appareil spécifique, la consommation en puissance affichée avec celle calculée à l'aide de l'équation (2.1) grâce aux mesures prises par la carte d'acquisition. Dans le second protocole, on utilise la même stratégie mais avec tous les CE connectés en série pour une mesure simultanée du même appareil.

Mesures utilisant un CE à la fois Le but de ce protocole est de comparer la consommation en puissance donnée par les différents CE pour le même appareil. Les mesures sont effectuées avec le même appareil et un des CE à chaque fois. Cela correspond à la configuration réelle de l'utilisation des CE où on prend un appareil et on le branche sur un CE pour mesurer sa consommation (figure 2.1). Ici les capteurs sont placés juste à côté de l'appareil afin d'éviter d'inclure la consommation du CE dans les mesures.

Mesures utilisant tous les CE simultanément Le protocole décrit dans la section précédente correspond à la configuration d'utilisation réelle. Le problème de ce protocole est que l'on doit supposer que la consommation de l'appareil ne varie pas trop entre deux mesures afin de comparer la même consommation. En raison de différents facteurs tels que

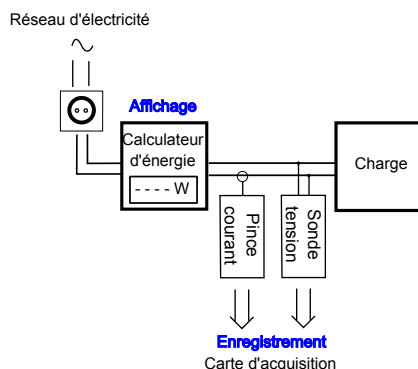


FIGURE 2.1 – Premier protocole de mesure de puissance utilisant un CE à la fois.

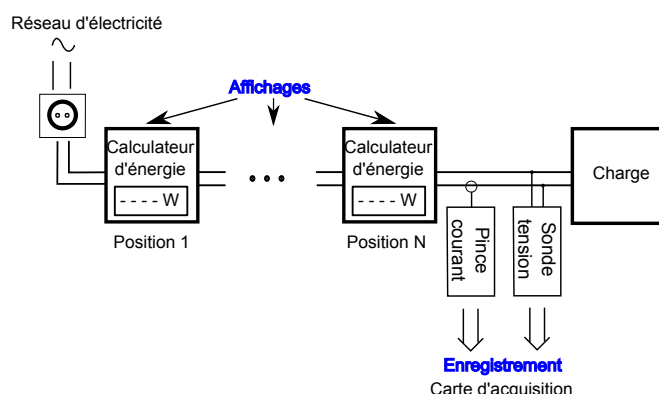


FIGURE 2.2 – Deuxième protocole de mesure de puissance utilisant tous les CE simultanément.

les variations de la tension et fréquence du réseau (discutées dans la sous-section 2.2.1), et aussi en raison des variations intrinsèques liées à l'appareil lui-même, il s'ensuit une impossibilité objective de comparaison. Pour pallier ce problème, on propose un deuxième protocole de mesure, où les CE sont tous mis ensemble pour mesurer simultanément le même appareil (figure 2.2). En faisant cela, on mesure la même consommation mais au détriment de l'introduction d'erreurs liées à la consommation intrinsèque des CE. On peut voir sur la figure 2.2 que le CE mis en position i ($i = 1, \dots, N - 1$) va mesurer la consommation de l'appareil ainsi que la consommation intrinsèque des CE mis dans les positions $i + 1$ à N . Ceci signifie qu'on doit prendre en compte ces consommations intrinsèques pour une comparaison correcte.

Le tableau 2.2 donne les valeurs indiquées (sur les notices d'utilisateur) pour la consommation intrinsèque des différents CE en comparaison avec les valeurs estimées par des mesures faites avec la carte d'acquisition. Étant donné que la seule façon d'accéder à la consommation calculée par les CE est de la lire directement à partir de l'écran d'affichage,

TABLE 2.2 – Consommation intrinsèque indiquée vs. estimée des CE

	Indiquée (notice)	Estimée
CE1	$< 0.5 \text{ W}$	0.4 W
CE2	non indiqué	1 W
CE3	$\leq 1 \text{ W}$	0.3 W
CE4	$\approx 1.8 \text{ W}$	1.63 W

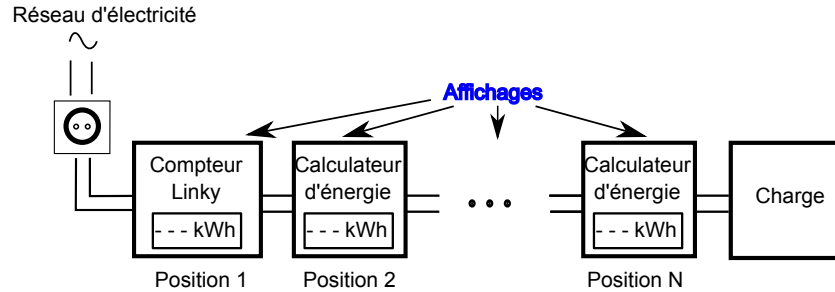


FIGURE 2.3 – Protocole de mesure d'énergie.

l'empilement des CE les uns sur les autres peut obstruer les écrans d'affichage du CE. Pour éviter cela, nous avons fabriqué des boîtes d'extension selon un principe de fonctionnement similaire aux rallonges afin de pouvoir connecter un CE à son voisin sans bloquer la lecture d'aucun d'entre eux.

2.2.2 Mesures d'énergie

La référence de comparaison dans ce cas est le compteur intelligent « Linky » (Linky, 2016). Linky est le compteur intelligent français qui remplacera les anciens compteurs d'électricité. L'administrateur français du système de distribution d'électricité Enedis (anciennement Electricité Réseau Distribution France (ERDF)), en charge du projet Linky, a annoncé que le déploiement des compteurs Linky, ayant démarré à la fin de l'année 2015, allait se poursuivre jusqu'à la fin de l'année 2021, avec un objectif de 90 % des anciens compteurs d'électricité remplacés dans les 35 millions de ménages. Le protocole de mesure d'énergie est illustré sur la figure 2.3. Ce protocole est similaire au second protocole de mesure de puissance présenté au paragraphe 2.2.1, la seule différence est que la carte d'acquisition est remplacée par le compteur Linky.

2.3 Résultats expérimentaux

Dans cette section, on donne les résultats des méthodes de mesure décrites dans la section 2.2 sauf pour l'utilisation de la puissance nominale écartée pour son inadaptation à la tâche de comparaison.

2.3.1 Puissance calculée en utilisant des mesures de laboratoire comme référence

Mesures utilisant un CE à la fois

Le tableau 2.3 donne un résumé des appareils mesurés (un total de 17 appareils) qui sont :

- lampes halogènes ;
- éléments chauffants (résistances) ;
- radiateurs électriques ;
- charge électronique programmable : des charges à puissance constante peuvent être simulées en utilisant une charge électronique programmable. Ici, on a utilisé la « Chroma Programmable ac & dc model 63800 series » ([Chroma 63800](#), 2016) qui est conçue à la fois pour la simulation de charge ac & dc (courant alternatif ou continu) ;
- aspirateurs et moteurs d'aspirateurs.

Ces différents appareils représentent différents équipements qu'il est possible de trouver dans les ménages, avec des puissances qui vont de 15 à 2000 W, ce qui nous permet aussi une comparaison du CE pour différents niveaux de puissance.

Les tableaux 2.4 et 2.5 donnent les résultats de la comparaison. Pour chaque CE et chaque appareil, on calcule la Racine carrée de l'Erreur Quadratique Moyenne (REQM) relative (à la puissance nominale) entre la consommation en puissance affichée sur le CE et le calcul fait avec les mesures de laboratoire.

La seule façon pour connaître la puissance calculée par les CE est de la lire à partir de l'écran d'affichage². Aussi simple que cela puisse paraître, une très bonne synchronisation

2. Afin de comparer avec les mesures du laboratoire on a dû prendre en vidéo l'écran d'affichage pendant la mesure, puis extraire les valeurs seconde par seconde (vu que les valeurs sur les écrans sont

2.3. RÉSULTATS EXPÉRIMENTAUX

a été nécessaire afin de comparer les mêmes valeurs mesurées. La carte d'acquisition lorsqu'elle est déclenchée pour démarrer une acquisition, commence immédiatement à mesurer le courant instantané consommé ainsi que la tension aux bornes de l'appareil à mesurer. En contrepartie, les valeurs affichées sur les CE ne sont pas les valeurs de la puissance instantanée consommée au moment actuel de la mesure, mais celles mesurées 2 ou 3 secondes avant en raison du retard de traitement nécessaire aux CE afin de calculer la puissance à partir des mesures acquises. Pour pouvoir corriger ce retard, on lance l'enregistrement avant la mise en marche de l'appareil et on le stop après sa mise à l'arrêt. Puis on aligne les deux vecteurs de valeurs à l'aide des instants de mise en marche et mise à l'arrêt.

Sur les notices d'utilisation, les quatre CE indiquent une erreur relative de mesure $\leq 2\%$ pour la consommation en puissance (sauf pour CE2 où aucune indication n'a été trouvée). Ceci est la raison pour laquelle on a choisi de classer les REQM relatives obtenues en trois intervalles ($\leq 2\%$, entre 2% et 5% , et $>5\%$). De cette façon, on pourra directement voir si les REQM relatives obtenues sont en accord avec les valeurs proclamées ou non. Les erreurs obtenues pour les différents CE ont été ordonnées dans les tableaux 2.4 et 2.5 telles que le CE avec la REQM relative la plus faible est mis en premier. Par exemple, dans la cellule correspondant à la lampe halogène 100 W, et à l'intervalle REQM (%) $\leq 2\%$, CE4/CE2/CE3 signifie que les trois CE ont des REQM relatives qui sont inférieures à 2% , mais signifie également que CE4 a la valeur la plus faible de tous et CE3 a la plus grande valeur de tous dans cet intervalle. Notons également sur ces deux tableaux qu'il n'y a pas de valeurs de REQM pour CE4 pour la charge électronique programmable, l'aspirateur et les moteurs d'aspirateurs car on ne disposait pas encore du CE4 au moment où les mesures ont été prises.

La principale conclusion tirée de ces résultats est que, pour des niveaux de puissance ≥ 100 W, la plupart des CE donnent des REQM relatives qui sont $\leq 2\%$, ce qui est en conformité avec les valeurs indiquées dans les notices d'utilisation. Cependant, pour des niveaux de puissance < 100 W, les CE donnent des REQM relatives entre 2% et 5% et parfois même $> 5\%$. En outre, à partir des résultats, on peut voir que CE3 et CE4 partagent généralement la première et la deuxième place (les mesures les plus précises)

rafraîchies chaque seconde) de la vidéo pour les mettre dans un vecteur de données.

dans le classement. Pour CE4, son prix pourrait justifier la performance, même si CE1 vendu pour un prix similaire n'a généralement pas donné de bons résultats, sauf pour les charges non résistives. Le CE3, pourtant le moins cher, a donné de bons résultats. Cette bonne performance peut être expliquée par le fait que ce CE comporte un circuit intégré dédié à la mesure de la puissance/énergie qui effectue le calcul de la puissance.

TABLE 2.3 – Appareils mesurés pour le premier protocole de mesure de puissance

Type d'appareil	Lampes halogènes	Éléments chauffants (résistances)	Radiateurs électriques	Charge électronique programmable	Aspirateurs (AS) et moteurs d'AS
Description	15, 53, 100 W 150, 250, 500 W	Résistance d'un chauffage 1000 W Résistance d'un sèche linge 1800 W	500, 1000 W 1500, 2000 W	Puissance constante 500 W Puissance constante 1000 W	Aspirateur 1100 W Moteurs d'aspirateurs 650, 1150 W

TABLE 2.4 – REQMs relatives des différentes mesures des CE évaluées en utilisant le premier protocole de mesure de puissance

Puissance nominale (W)	Lampes halogènes						Éléments chauffants (résistances)	
	15	53	100	150	250	500	1000	1800
REQM (%) $\leq 2\%$	CE2	CE4/CE2	CE4/CE2 CE3	CE4/CE3 CE2	CE4/CE3 CE2/CE1	CE3/CE4 CE1/CE2	CE3/CE4 CE2/CE1	CE4/CE3 CE2/CE1
2% < REQM (%) $\leq 5\%$	CE4/CE3	CE3/CE1	CE1	CE1	-	-	-	-
REQM (%) > 5%	CE1	-	-	-	-	-	-	-

TABLE 2.5 – REQMs relatives des différentes mesures des CE évaluées en utilisant le premier protocole de mesure de puissance–Suite

Puissance nominale (W)	Radiateurs électriques				Charge électronique programmable		Aspirateurs (AS) et moteurs d'AS		
	500	1000	1500	2000	500	1000	1100	650	1150
REQM (%) $\leq 2\%$	CE3/CE4 CE2/CE1	CE3/CE4 CE2/CE1	CE3/CE4 CE2/CE1	CE4/CE3 CE2/CE1	CE2/CE1 CE3	CE1/CE2/CE3	CE1/CE3 CE2	CE1/CE3 CE2	CE1/CE3 CE2
2% < REQM (%) $\leq 5\%$	-	-	-	-	-	-	-	-	-
REQM (%) > 5%	-	-	-	-	-	-	-	-	-

TABLE 2.6 – Appareils mesurés pour le deuxième protocole de mesure de puissance

Type d'appareil	Lampes halogènes	Bouilloire	Aspirateur	Ventilateur	Switch réseau
Description	150, 500 W	1800 W	3 niveaux de puiss. (73, 400, 615 W)	3 niveaux de puiss. (25, 31, 40 W)	3 W

TABLE 2.7 – REQMs relatives des différentes mesures des CE évaluées en utilisant le deuxième protocole de mesure de puissance

Puissance nominale (ou approchée) (W)	Lampes halogènes		Bouilloire	Aspirateur			Ventilateur			Switch réseau
	150	500	1800	73	400	615	25	31	40	3
REQM (%) $\leq 2\%$	CE1/CE2 CE3	CE1/CE2 CE3	CE2/CE1	CE1/CE2 CE3	CE1/CE2 CE3	CE1/CE2 CE3	CE1	CE1	CE1	-
2% < REQM (%) $\leq 5\%$	-	-	CE3	-	-	-	CE2/CE3	CE3/CE2	CE3/CE2	-
REQM (%) > 5%	-	-	-	-	-	-	-	-	-	CE1/CE3 CE2

Mesures utilisant tous les CE simultanément

Le tableau 2.6 donne les caractéristiques des appareils mesurés :

- lampes halogènes ;
- bouilloire ;
- aspirateur : avec un variateur de puissance qui permet de changer le niveau de la puissance consommée. On a choisi trois niveaux (puissance minimale, intermédiaire et maximale) ;
- ventilateur : ayant trois programmes, chacun fonctionnant à un niveau spécifique de puissance ;
- switch réseau : représentant les charges électroniques (différent de la charge électronique programmable), il a un très faible niveau de puissance (≈ 3 W).

Ces appareils ont notamment été choisis afin d'évaluer la réponse des CE à des appareils qui ne sont pas nécessairement résistifs. Notons que pour ce protocole de mesure on a juste comparer CE1 à CE3, puisqu'on ne disposait pas du CE4 au moment où les mesures ont été prises.

En outre, on a cherché à privilégier un agencement des CE où chacun est le moins influencé possible par la consommation des autres, en ordonnant les CE selon : réseau $\rightarrow$ CE2 $\rightarrow$ CE3 $\rightarrow$ CE1 $\rightarrow$ Appareil (figure 2.2). En mettant le CE2 avec la plus grande consommation intrinsèque en position 1 (sur la prise de courant secteur), on évite de prendre en compte sa consommation intrinsèque. Après cela, dans les positions 2 et 3, on peut mettre indifféremment soit CE1 ou CE3 qui ont des valeurs de consommation intrinsèque comparables.

Les résultats de comparaison obtenus sont présentés dans le tableau 2.7 comme REQM relatives entre les valeurs **corrigées** (en prenant en compte la consommation intrinsèque des CE) des puissances affichées sur les CE et la consommation calculée en utilisant les mesures laboratoire. Après cela, ils sont classés de la même manière que celle des tableaux 2.4 et 2.5 (soit trois intervalles différents et un ordre indiquant le classement des CE pour un même intervalle).

Tous les CE ont des REQM relatives $\leq 5\%$ pour toutes les charges mesurées à l'exception de la charge électronique (switch réseau). Ceci n'est pas surprenant puisque car ces

2.3. RÉSULTATS EXPÉRIMENTAUX

CE ne sont pas destinés à mesurer des niveaux de consommation aussi faibles. La plupart des CE mesurent des charges dans la gamme de puissance [5 - 3680] W et ne sont donc pas destinés à mesurer la consommation des charges électroniques.

Les résultats pour le ventilateur et l'aspirateur sont très intéressants car ils montrent que le CE1 est le plus adapté (toujours au premier rang) pour mesurer la consommation des appareils à moteur et c'est presque indépendamment du niveau de la consommation, car il a été capable de mesurer la consommation du ventilateur à un niveau de puissance aussi faible que 25 W avec une précision $< 2\%$.

2.3.2 Mesures d'énergie avec le compteur intelligent Linky comme référence

Pour comparer les mesures d'énergie, on a fait deux expériences différenciées par les positions respectives des CE.

Pour la première expérience, on a choisi la configuration suivante : réseau électrique $\rightarrow$ Linky $\rightarrow$ CE4 $\rightarrow$ CE2 $\rightarrow$ CE3 $\rightarrow$ CE1 $\rightarrow$ Appareil (figure 2.3), où les positions ont été sélectionnées en fonction de la consommation intrinsèque de chaque CE (tableau 2.2). Le CE ayant la plus grande consommation intrinsèque a été positionné le plus loin possible de l'appareil mesuré.

Pour la seconde expérience, on a voulu voir si la modification de positions des CE pouvait avoir un effet sur les erreurs calculées. On a ainsi inversé l'ordre choisi dans la première expérience, ce qui donne la configuration : Réseau électrique $\rightarrow$ Linky $\rightarrow$ CE1 $\rightarrow$ CE3 $\rightarrow$ CE2 $\rightarrow$ CE4 $\rightarrow$ Appareil.

Les expériences consistent à mesurer la consommation d'énergie pendant 48 heures (2 jours) d'une charge de 2650 W (lampe halogène de 500 W + lampe halogène de 150 W + radiateur électrique de 2000 W), limitée pour ne pas endommager CE4 qui ne peut supporter des puissances supérieures à 3000 W.

Les résultats de mesure des deux expériences sont donnés dans les tableaux 2.8 et 2.9. Il est à noter que le compteur Linky donne des mesures d'énergie avec une précision de

1 kWh. La consommation de 117,84 kWh affichée pour Linky sur le tableau 2.8 correspond en fait à une consommation corrigée $(118 - (1,63 + 1 + 0,3 + 0,4) * 10^{-3} * 48 = 117,84 \text{ kWh})$ qui prend en compte la consommation intrinsèque des différents CE pendant les 48h. On a également corrigé les valeurs obtenues pour les CE. On a estimé que les CE tous ensemble ont consommé environ 160 Wh pendant les 48 heures, ce qui équivaut à l'utilisation pendant deux heures d'une lampe de 80 W.

Notons tout d'abord que, pour la même durée et la même charge, les consommations d'énergie obtenues sont différentes entre les deux expériences (117,84 kWh dans la première et 115,84 kWh dans la seconde, tel que affiché sur Linky). Cela pourrait être lié aux raisons exposées dans la sous-section 2.2.1 ou éventuellement à la charge considérée. Un radiateur électrique ne fonctionne pas en général de façon continue sans interruption, mais ajuste plutôt son fonctionnement à la température ambiante. Pour cette raison, il est probable qu'au cours des deux 48 heures le radiateur électrique ait consommé différentes quantités d'énergie.

On peut également noter qu'il y a des erreurs positives et négatives correspondant, respectivement, aux CE surestimant ou sous-estimant la consommation d'énergie par rapport à la valeur Linky. Les CE2 et CE4 ont légèrement surestimé la consommation d'énergie dans les deux expériences (entre 0,17 et 0,32%). Le CE3, par contre, a sous-estimé la consommation d'énergie pour les deux expériences avec de grandes erreurs relatives (-3,91 et -4,18%), ce qui surprenant car ce CE a donné des résultats plutôt bons pour la mesure de puissance des charges résistives grande puissance (tableaux 2.4 et 2.5). Cela pourrait être lié à la façon avec laquelle la consommation d'énergie est calculée. Étant donné qu'une énergie est le produit de la puissance consommée avec la durée de consommation, une mauvaise estimation de l'énergie consommée pourrait provenir d'un mauvais mode de calcul et/ou de durée de consommation si l'on suppose que la puissance consommée est bien estimée.

Il est à noter également que la position des CE ne semble pas affecter les résultats de mesure de l'énergie puisque des erreurs relatives similaires ont été obtenues pour la majorité des CE (sauf pour CE1). Enfin, CE1, et pour une raison inconnue, a donné une consommation d'énergie sous-estimée (-1,82% d'erreur relative) dans la première expérience et une consommation surestimée (0,1% d'erreur relative) dans la seconde expérience, ce qui

2.4. CONCLUSION

TABLE 2.8 – Résultat de la mesure d'énergie sur 48h–première expérience

Calculateurs d'énergie	Linky	CE1	CE2	CE3	CE4
Énergie consommée (kWh)	117.84	115.70	118.06	112.91	118.22
Erreur par rapport à Linky (kWh)	0	-2.14	0.21	-4.93	0.38
Erreur relative par rapport à Linky (%)	0	-1.82	0.18	-4.18	0.32

TABLE 2.9 – Résultat de la mesure d'énergie sur 48h–seconde expérience

Calculateurs d'énergie	Linky	CE1	CE2	CE3	CE4
Énergie consommée (kWh)	115.84	115.96	116.03	111.31	116.10
Erreur par rapport à Linky (kWh)	0	0.12	0.19	-4.53	0.26
Erreur relative par rapport à Linky (%)	0	0.10	0.17	-3.91	0.22

montre un comportement non cohérent pour l'estimation de l'énergie consommée et que nous n'avons, malheureusement, pas su expliquer.

2.4 Conclusion

Dans ce chapitre, on a comparé la consommation électrique fournie par certains calculateurs d'énergie à bas coût (CEBC) de différents appareils ménagers, de différents types et ayant différents niveaux de puissance. Les résultats pour les mesures de puissance montrent que les CE ont donné des estimations conformes aux intervalles d'incertitude affichés pour les charges résistives ayant des niveaux de consommation élevés (plus de 100 W), mais pas pour les niveaux de consommation de puissance faibles. Pour les charges non résistives, les CEBC ont donné des résultats variables (CE1 par exemple est apparemment plus adapté à l'estimation de la puissance des appareils à moteur). Les résultats pour les mesures d'énergie montrent qu'un CEBC ayant une bonne estimation de la consommation en puissance ne donne pas nécessairement une bonne estimation de la consommation d'énergie et vice versa, et peuvent même donner des résultats incohérents.

En outre, aussi petite qu'elle puisse être, la consommation intrinsèque de puissance/énergie des CEBC, lors du déploiement de plusieurs d'entre eux avec un fonctionnement continu, pourrait ne plus être négligeable dans la consommation globale d'énergie du ménage.

Chapitre 3

Analyse des courants de mise en marche

Sommaire

3.1	Introduction	41
3.2	Caractérisation paramétrique des transitoires de courant	42
3.2.1	Modélisation	42
3.2.2	Hypothèses de modélisation	44
3.2.3	Algorithme d'estimation des paramètres (ECT)	45
3.2.4	Calcul des performances théoriques-BCR	51
3.3	Performances expérimentales de l'algorithme	54
3.3.1	Évaluation sur données simulées	54
3.3.2	Évaluation sur données réelles	72
3.3.3	Algorithme d'estimation des paramètres amélioré (ECTA)	77
3.4	Détection d'évènements	83
3.4.1	Algorithme de détection (HAND)	83
3.4.2	Résultats expérimentaux	89
3.4.3	Conclusion	94

3.1 Introduction

Nous avons vu lors du premier chapitre que l'approche retenue pour l'identification des appareils consistait à modéliser le courant de mise en marche. Ceci passe par un modèle mathématique qui décrit l'évolution du courant à partir de son instant de mise en marche.

Dans ce chapitre on donne les contributions apportées sur certaines étapes nécessaires à l'identification des appareils. Ci-après on présente :

- un modèle paramétrique des courants de mise en marche ;
- la proposition d'algorithmes pour l'estimation des paramètres de ce modèle ;
- le calcul des performances théoriques du modèle : Bornes de Cramér-Rao (BCR) (Cramér-Rao Bounds-CRB) ; en particulier, nous discutons l'impact de certains paramètres sur la qualité d'estimation ;
- l'évaluation des performances du modèle sur des données simulées ainsi que sur des données réelles ; pour les données réelles, on montre l'amélioration des performances suite à l'ajout d'un étage qui estime la fréquence fondamentale ;
- un algorithme pour la détection des instants de mise en marche et de mise à l'arrêt des appareils.

Le détecteur d'évènements présenté en fin de ce chapitre cherche à identifier les périodes d'activité des appareils. Plus précisément, il cherche à identifier finement l'instant de démarrage d'un appareil mais aussi à segmenter le signal de courant en différentes zones : une zone de transition à la mise en marche, une zone de régime établi et une zone de transition à la mise à l'arrêt. Cette segmentation est nécessaire pour notre stratégie d'estimation.

3.2 Caractérisation paramétrique des transitoires de courant de mise en marche

3.2.1 Modélisation

Dans cette section, on propose un nouveau modèle mathématique pour les signaux de courant électrique. Plus précisément, nous nous intéressons à la modélisation des signaux transitoires de courant correspondant à la mise en marche des appareils. L'aspect transitoire du courant est caractérisé par un changement d'état dont la durée est brève par rapport à la durée de l'état antérieur et postérieur à ce changement.

Il est intéressant de souligner que le transitoire qui nous intéresse est différent du transitoire très haute fréquence (de l'ordre du MHz) généré par les dispositifs de commutation (interrupteurs mécaniques ou électroniques) qui apparaît sur les signaux de courant électrique sous forme d'impulsions (Howell, 1979) en raison de la fermeture du circuit électrique.

Le transitoire de mise en marche est caractérisé par une forte consommation d'énergie en début de consommation suivie d'une diminution (ou amortissement) de l'amplitude de l'énergie consommée jusqu'à atteindre un état stable (figure 3.1).

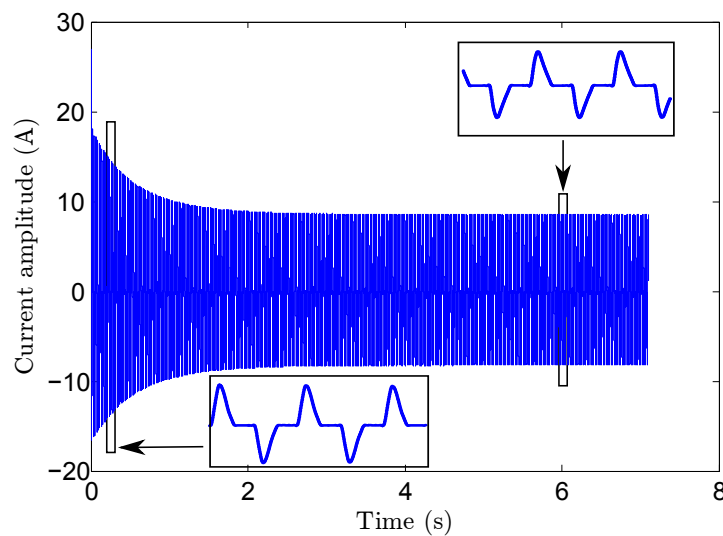


FIGURE 3.1 – Courant de mise en marche d'un aspirateur.

3.2. CARACTÉRISATION PARAMÉTRIQUE DES TRANSITOIRES DE COURANT

La forme de ce transitoire de mise en marche et son amplitude varient d'un appareil à l'autre. Nous proposons de le modéliser par une modulation en amplitude d'un signal périodique. Une version descriptive non bruitée du modèle $s(t)$ est :

$$s(t) = e(t)s_s(t), \quad (3.1)$$

où $s_s(t)$ est le courant sans modulation d'amplitude (s_s signifiant somme de sinusoides) avec $e(t)$ représentant cette modulation. La partie périodique est supposée être constituée de d harmoniques :

$$s_s(t) = \sum_{i=1}^d A_i \cos(2\pi f_i t + \phi_i), \quad (3.2)$$

$s_s(t)$ est décrite comme une somme de sinusoides non amorties et $A_i (\geq 0)$, $\phi_i \in [-\pi, \pi]$ et f_i sont leurs amplitudes, phases et fréquences respectives. Le nombre d de sinusoides (harmoniques du courant) est supposé fixe et connu *a priori*.

En raison de la symétrie demi-onde des signaux électriques (pour un signal périodique $g(t)$ de période T , une symétrie demi-onde est caractérisée par $g(t \pm T/2) = -g(t)$), les fréquences f_i des sinusoides sont censées être des harmoniques d'ordre impair de la fréquence fondamentale ($f_0 = 50$ Hz dans notre cas, aussi appelée fréquence du réseau), c.-à-d. $f_i = (2i - 1)f_0$, $i = 1, \dots, d$.

La modulation d'amplitude $e(t)$ est choisie de la forme $e^{u(t)} + 1$ telle que $e^{u(t)} \xrightarrow[t \rightarrow +\infty]{} 0$. Cette fonction exponentielle a été choisie pour son intérêt dans la description des phénomènes amortis. Un modèle d'amortissement classique est décrit par la forme $u(t) = -\alpha t$ avec $\alpha \geq 0$. Pour les signaux de courant, il a été nécessaire, après analyse expérimentale d'étendre cette forme. Nous avons fait le choix de prendre une forme polynomiale autorisant plus de flexibilité dans la description de la modulation :

$$e(t) = \begin{cases} e^{\mathbf{p}^T \mathbf{t}} + 1 & , \text{ si } t \geq 0 \\ 0 & , \text{ sinon} \end{cases} \quad (3.3)$$

L'enveloppe, ou la modulation, $e(t)$ décrit la variation de l'amplitude du début de la consommation jusqu'à atteindre le régime établi. $\mathbf{p} = [p_0, p_1, \dots, p_n]^T$ est un vecteur de $n+1$ coefficients polynomiaux et $\mathbf{t} = [1, t, \dots, t^n]^T$ est un vecteur représentant le temps tel

que $\mathbf{p}^T \mathbf{t}$ donne un polynôme de degré n permettant d’adapter le modèle aux variations du signal réel. Le polynôme est tel que $\mathbf{p}^T \mathbf{t} \xrightarrow[t \rightarrow +\infty]{} -\infty$ donnant $e(t) \xrightarrow[t \rightarrow +\infty]{} 1$ (condition vérifiée si $p_n < 0$). Le degré du polynôme est relativement faible (suivant nos simulations on a choisis $n = 3$ pour la suite).

Nous supposons que le signal de courant mesuré $x(t)$ est corrompu par un bruit blanc additif gaussien $w(t)$ ¹ ayant une moyenne nulle et une variance σ^2 :

$$x(t) = s(t) + w(t), \quad \forall t. \quad (3.4)$$

Remarque : Il est à noter que ce modèle n’est pas valide pour *tous* les types d’appareils électriques. Pour certains appareils électriques, on peut trouver un signal de courant de mise en marche qui se compose de différentes phases chacune ayant un contenu différent (harmoniques avec différentes amplitudes et phases) ainsi qu’une forme d’enveloppe différente. En l’occurrence, le courant de mise en marche du micro-ondes représenté sur la figure 3.2 comporte deux phases (quelques micro-ondes peuvent avoir plus de deux), chacune ayant ses caractéristiques propres. Néanmoins, le modèle proposé est toujours valable pour un grand nombre d’appareils électriques qui montrent un comportement à phase unique au moment de la mise en marche, tels que : les lampes à incandescence, lampes fluo-compactes, appareils de chauffage, aspirateurs, sèche-cheveux, etc.

3.2.2 Hypothèses de modélisation

Le modèle proposé modélise un signal unique où l’hypothèse implicite est celle d’un appareil fonctionnant seul sur le réseau². Pour pouvoir appliquer le modèle au cas multi-appareil, on émet les deux hypothèses suivantes :

- les instants de mise en marche des appareils sont suffisamment espacés en temps pour permettre de les traiter séparément. La durée minimale exigée entre les mises en marche dépend de la durée du transitoire qui va de quelques périodes secteur à quelques secondes suivant le type d’appareil ;

1. Pour nos mesures réelles (section 4.2.3), l’histogramme de bruit montre que ce modèle est bien justifié.

2. Réseau électrique domestique.

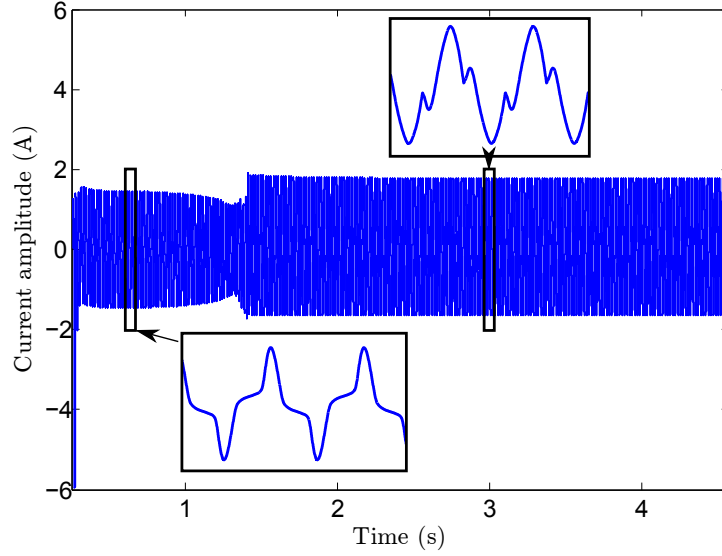


FIGURE 3.2 – Courant de mise en marche d'un micro-ondes.

- il est possible de se ramener au cas d'un appareil unique sur le réseau en soustrayant le signal courant reconstruit à partir du régime établi précédant la mise en marche.

3.2.3 Algorithme d'estimation des paramètres (ECT)

Pour cet algorithme, que nous appellerons algorithme d'Estimation du Courant Transitoire (ECT) (Transient Current Estimation–TCE), nous supposons que les fréquences sont fixes et connues telles que $f_i = (2i - 1)f_0$, $i = 1, \dots, d$ avec d : nombre de sinusoïdes supposé fixe et connu *a priori* et $f_0 = 50$ Hz.

Notons que le passage du temps continu au temps discret se fait à l'aide de la notation $t_k = t(kT_s)$, $T_s = 1/F_s$ étant la période d'échantillonnage (sampling period) et k un nombre entier.

L'algorithme d'estimation se compose de trois phases, une phase d'initialisation, une phase d'estimation et une phase d'optimisation :

1. **phase d'initialisation** : sert à fixer t_{e1} et t_{e2} : deux instants qui définissent une partie du régime établi utilisée dans la phase suivante d'estimation.
2. **phase d'estimation** : cette phase permet de faire :
 - l'estimation de $s_s(t_k)$ (c.-à-d. A_i et ϕ_i),
 - l'estimation de $e(t_k)$ (c.-à-d. $\mathbf{p}$).

3. **phase d'optimisation** : dans cette phase on optimise la solution trouvée dans la phase d'estimation.

Phase d'initialisation

Durant cette phase nous fixons deux instants t_{e1} et t_{e2} (e signifiant régime établi) qui définissent une portion du régime établi du signal courant mesuré, notée $x_e(t_k)$, utilisée dans la phase suivante pour l'estimation des A_i et ϕ_i .

Sur cette portion du régime établi $t_k \in [t_{e1}, t_{e2}]$ on peut écrire $x(t_k) = s_s(t_k) + w(t_k)$.

Pratiquement, pour retrouver ces deux instants t_{e1} et t_{e2} , nous utilisons, le détecteur d'évènements « High Accuracy NILM Detector (HAND) » que nous avons proposé et qui sera présenté en sous-section 3.4.1.

Le HAND, quand ils est appliqué sur un signal contenant un seul appareil en fonctionnement (mono-appareil), donne en sortie quatre instants définissant le début t_{dm} et la fin t_{fm} de la zone de transition à la mise en marche ainsi que le début t_{da} et la fin t_{fa} de la zone de transition à la mise à l'arrêt (figure 3.3).

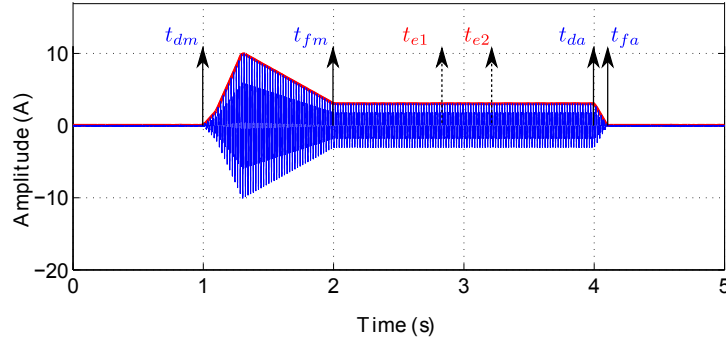


FIGURE 3.3 – Illustration du résultat de détection donné par HAND sur un signal simulé mono-appareil. En sortie HAND donne $t_{dm}, t_{fm}, t_{da}, t_{fa}$. Les instants t_{e1} et t_{e2} définissent la portion du régime établi utilisée pour l'estimation des A_i et ϕ_i .

Notons que pour l'estimation on fixe : $t_{e2} = t_{da}$ et $t_{e1} = t_{da} - 0.5$ s (une portion du régime établi de 25 périodes secteur). C'est à dire que l'on considère la dernière demi-seconde du régime établi comme représentant $x_e(t_k)$.

Phase d'estimation

1. Estimation de $s_s(t_k)$:

En utilisant le critère des Moindres Carrés Linéaires (MCL), les paramètres A_i et ϕ_i sont estimés à partir de $x_e(t_k)$. Pour cela, on pose :

$$\begin{aligned} x_e(t_k) &= \sum_{i=1}^d A_i \cos(2\pi f_i t_k + \phi_i) + w(t_k) \\ &= \sum_{i=1}^d A_i \cos \phi_i \cos(2\pi f_i t_k) - A_i \sin \phi_i \sin(2\pi f_i t_k) + w(t_k), \quad t_k \in [t_{e1}, t_{e2}] \end{aligned} \quad (3.5)$$

qui peut être écrite sous forme matricielle comme suit :

$$\mathbf{x}_e = \mathbf{M}\mathbf{c} + \mathbf{w}, \quad (3.6)$$

où $\mathbf{c} = [A_1 \cos \phi_1, A_1 \sin \phi_1, \dots, A_d \cos \phi_d, A_d \sin \phi_d]^T$, $\mathbf{x}_e = [x_e(t_{e1}), \dots, x_e(t_{e2})]^T$, $\mathbf{w}$ est un vecteur représentant le bruit et $\mathbf{M}$ est une matrice telle que :

$$\mathbf{M} = \begin{bmatrix} \cos(2\pi f_1 t_{e1}) & -\sin(2\pi f_1 t_{e1}) & \cdots & \cos(2\pi f_d t_{e1}) & -\sin(2\pi f_d t_{e1}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \cos(2\pi f_1 t_{e2}) & -\sin(2\pi f_1 t_{e2}) & \cdots & \cos(2\pi f_d t_{e2}) & -\sin(2\pi f_d t_{e2}) \end{bmatrix} \quad (3.7)$$

Le critère des moindres carrés, utilisé pour retrouver une estimée de $\mathbf{c}$, vise à minimiser le carré de la norme euclidienne ($\|\cdot\|_2$) de la différence entre la mesure et le modèle, c.-à-d. :

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \frac{1}{2} \|\mathbf{x}_e - \mathbf{M}\mathbf{c}\|_2^2 \quad (3.8)$$

Dans ce cas, la solution $\hat{\mathbf{c}}$ est donnée par :

$$\hat{\mathbf{c}} = \mathbf{M}^+ \hat{\mathbf{s}}_s \quad (3.9)$$

où $\mathbf{M}^+ = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T$ est la matrice pseudo-inverse (de Moore-Penrose) de $\mathbf{M}$. On extrait de $\hat{\mathbf{c}}$ les deux vecteurs $\hat{\mathbf{c}}_s = [\hat{A}_1 \cos \hat{\phi}_1, \dots, \hat{A}_d \cos \hat{\phi}_d]^T$ et

$\hat{\mathbf{s}}\mathbf{n} = [\hat{A}_1 \sin \hat{\phi}_1, \dots, \hat{A}_d \sin \hat{\phi}_d]^T$ et on calcule $\hat{A}_i$ et $\hat{\phi}_i$ comme suit :

$$\hat{\mathbf{a}} = \begin{bmatrix} \hat{A}_1 \\ \vdots \\ \hat{A}_d \end{bmatrix} = \sqrt{\hat{\mathbf{c}}\mathbf{s} \odot \hat{\mathbf{c}}\mathbf{s} + \hat{\mathbf{s}}\mathbf{n} \odot \hat{\mathbf{s}}\mathbf{n}} \quad (3.10)$$

$$\hat{\boldsymbol{\phi}} = \begin{bmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_d \end{bmatrix} = \arctan(\hat{\mathbf{s}}\mathbf{n} \oslash \hat{\mathbf{c}}\mathbf{s}) \quad (3.11)$$

où $\odot$ et $\oslash$ sont, respectivement, les opérateurs de produit et de division élément par élément.

2. Estimation de $e(t_k)$:

Pour estimer $e(t_k)$ on utilise un algorithme d'optimisation non linéaire de la classe d'algorithmes dite à région de confiance (trust region). Différents algorithmes peuvent être trouvés dans cette classe, un des plus connus étant l'algorithme de Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963). Cet algorithme ne permet pas d'imposer des contraintes sur les valeurs des paramètres à estimer et n'est donc pas adapté à notre estimation puisque notre modèle (équation (3.1)) suppose les contraintes suivantes sur les paramètres : $A_i (\geq 0)$, $\phi_i \in [-\pi, \pi]$ et $p_n < 0$.

Un autre algorithme de cette classe est l'algorithme appelé Trust-Region-Reflective (TRR) (Coleman and Li, 1994, 1996). Cet algorithme permet d'imposer des contraintes et c'est celui que nous utiliserons³ pour l'estimation de $e(t_k)$.

Ayant les estimées $\hat{\mathbf{a}}$, $\hat{\boldsymbol{\phi}}$ obtenues à l'étape précédente ainsi que le vecteur mesure $\mathbf{x} = [x(t_0), \dots, x(t_{N-1})]^T$, on estime $\mathbf{p}$ avec (où la seule inconnue est $\mathbf{p}$) :

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}\mathbf{c})\|_2^2, \quad (3.12)$$

$\mathbf{e} = [e(t_0), \dots, e(t_{N-1})]^T$, avec N le nombre d'échantillons du signal mesure $\mathbf{x}$.

À la fin de la phase d'estimation on obtient un vecteur de paramètres estimés $\hat{\boldsymbol{\theta}}_0 = [\hat{\mathbf{p}}, \hat{\mathbf{a}}, \hat{\boldsymbol{\phi}}]^T$.

3. Cet algorithme est disponible avec la fonction *lsqnonlin* de MATLAB.

3.2. CARACTÉRISATION PARAMÉTRIQUE DES TRANSITOIRES DE COURANT

Notons que cet algorithme a besoin d'une initialisation. On choisit pour cette étape un vecteur nul $\mathbf{p}_0 = \mathbf{0}$.

Phase d'optimisation

Comme l'estimation des A_i et ϕ_i n'est effectuée que sur une portion réduite du signal $x_e(t_k)$, l'idée de cette phase est d'affiner l'estimation de tous les paramètres en considérant le signal dans sa globalité.

On utilise le même algorithme utilisé pour l'estimation de $e(t_k)$ en prenant comme valeur initiale le vecteur $\hat{\boldsymbol{\theta}}_0$ calculé précédemment.

L'inconnue cette fois est le vecteur $\hat{\boldsymbol{\theta}}$ contenant l'ensemble des paramètres du modèle.

L'algorithme ECT est ainsi donné ci-dessous :

Algorithme 1 : Estimation du Courant Transitoire (ECT).

Initialisation :

1. signal courant : $\mathbf{x}$
2. fixer $n = 3, d = 5$ et $f_i = (2i - 1)f_0, i = 1, \dots, d$;
3. trouver t_{da} en utilisant l'algorithme HAND (algorithme 4) ;
4. fixer $t_{e2} = t_{da}$ et $t_{e1} = t_{da} - 0.5$ s.
5. déterminer $\mathbf{x}_e = \mathbf{x}(t_{e1}, \dots, t_{e2})$

Estimation :

1. calculer $\hat{\mathbf{a}}$ et $\hat{\phi}$ tels que :

$$\hat{\mathbf{a}} = \sqrt{\hat{\mathbf{c}}\hat{\mathbf{s}} \odot \hat{\mathbf{c}}\hat{\mathbf{s}} + \hat{\mathbf{s}}\hat{\mathbf{n}} \odot \hat{\mathbf{s}}\hat{\mathbf{n}}} \quad (\text{équation 3.10})$$

$$\hat{\phi} = \arctan(\hat{\mathbf{s}}\hat{\mathbf{n}} \oslash \hat{\mathbf{c}}\hat{\mathbf{s}}) \quad (\text{équation 3.11})$$

en utilisant le critère des MCL où $\hat{\mathbf{c}}\hat{\mathbf{s}}$ et $\hat{\mathbf{s}}\hat{\mathbf{n}}$ sont extraits de :

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \frac{1}{2} \|\mathbf{x}_e - \mathbf{M}\mathbf{c}\|_2^2 \quad (\text{équation 3.8})$$

avec $\mathbf{M}$ calculée en utilisant l'équation 3.7 et

$$\mathbf{c} = [A_1 \cos \phi_1, A_1 \sin \phi_1, \dots, A_d \cos \phi_d, A_d \sin \phi_d]^T ;$$

2. poser $\mathbf{p}_0 = \mathbf{0}$;
3. calculer $\hat{\mathbf{p}}$ tel que :

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}\mathbf{c})\|_2^2 \quad (\text{équation 3.12})$$

en utilisant l'algorithme TRR avec $\mathbf{p}_0$.

Optimisation :

1. poser $\boldsymbol{\theta}_0 = [\hat{\mathbf{p}}, \hat{\mathbf{a}}, \hat{\phi}]^T$;

2. calculer $\hat{\boldsymbol{\theta}}$ tel que :

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \frac{1}{2} \|\mathbf{x} - (\mathbf{e} \odot \mathbf{M}\mathbf{c})\|_2^2$$

en utilisant l'algorithme TRR avec $\boldsymbol{\theta}_0$.

3.2.4 Calcul des performances théoriques–BCR

La Borne de Cramér-Rao (BCR) (Cramér-Rao Bound–CRB) donne une borne inférieure pour la variance de tout estimateur non biaisé (Besson and Abramovich, 2013). Nous montrons, dans une des sections qui suivent, que les paramètres estimés par notre algorithme sont très faiblement biaisés ce qui permet l'utilisation des BCR pour évaluer les performances d'estimation de notre algorithme.

Évaluer la performance de l'estimation revient à comparer les variances des paramètres estimés (concrètement on compare les EQM mais comme on suppose que le biais des paramètres est nul cela revient au même) avec les BCR des paramètres correspondants.

En prenant en compte la dépendance de $s(t_k)$ du vecteur de paramètres $\boldsymbol{\theta}$ et N échantillons de la mesure $x(t_k)$, l'équation (3.4) peut être écrite sous forme vectorielle comme suit :

$$\mathbf{x} = \mathbf{s}(\boldsymbol{\theta}) + \mathbf{w}. \quad (3.13)$$

Dans ce cas, $\mathbf{x}$ suit une loi normale de moyenne $\mathbf{s}(\boldsymbol{\theta})$ et de matrice de covariance $\mathbf{C} = \sigma^2 \mathbf{I}$.

La BCR est définie comme la matrice inverse de la Matrice d'Information de Fisher (MIF) (Fisher Information Matrix–FIM). Si on suppose $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]^T$, $K = n + 1 + 2d$ et $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{R}^N$, alors pour le cas gaussien général où $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$, la MIF est donnée par (Kay, 1993) :

$$\begin{aligned} [\mathbf{F}(\boldsymbol{\theta})]_{ij} = & \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \right)^T \mathbf{C}^{-1}(\boldsymbol{\theta}) \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} \right) \\ & + \frac{1}{2} \text{tr} \left[\mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_j} \right], \quad i, j = 1, \dots, K \end{aligned} \quad (3.14)$$

avec :

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} = \begin{bmatrix} \frac{\partial [\boldsymbol{\mu}(\boldsymbol{\theta})]_1}{\partial \theta_i} \\ \vdots \\ \frac{\partial [\boldsymbol{\mu}(\boldsymbol{\theta})]_N}{\partial \theta_i} \end{bmatrix}, \quad \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} = \begin{bmatrix} \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{11}}{\partial \theta_i} & \dots & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{1N}}{\partial \theta_i} \\ \vdots & \ddots & \vdots \\ \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{N1}}{\partial \theta_i} & \dots & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{NN}}{\partial \theta_i} \end{bmatrix} \quad (3.15)$$

Le symbole $[\cdot]_i$ dénote l'élément d'indice i du vecteur correspondant et $[\cdot]_{ij}$ l'élément d'indice ij de la matrice correspondante. tr dénote l'opérateur qui permet de calculer la trace d'une matrice (somme des éléments diagonaux).

Pour notre modèle où $\boldsymbol{\mu}(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta})$ et $\mathbf{C}(\boldsymbol{\theta}) = \sigma^2 \mathbf{I}$ (matrice de covariance indépendante

de $\boldsymbol{\theta}$), la MIF devient :

$$[\mathbf{F}(\boldsymbol{\theta})]_{ij} = \frac{1}{\sigma^2} \left(\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_i} \right)^T \left(\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_j} \right), \quad i, j = 1, \dots, K. \quad (3.16)$$

BCR du vecteur de paramètres $\boldsymbol{\theta} = [\mathbf{p}, \mathbf{a}, \phi]^T$

Ici, $K = n + 1 + 2d$ et on suppose σ^2 connue. Si on suppose que σ^2 est aussi un paramètre inconnu à estimer avec $\boldsymbol{\theta}' = [\mathbf{p}, \mathbf{a}, \phi, \sigma^2]^T$, alors la MIF $\mathbf{F}(\boldsymbol{\theta}')$ est égale à la MIF $\mathbf{F}(\boldsymbol{\theta})$ augmentée d'une colonne et d'une ligne correspondantes aux dérivées partielles où σ^2 intervient. On obtient ainsi (en utilisant l'équation (3.14)) :

$$\mathbf{F}(\boldsymbol{\theta}') = \begin{bmatrix} \mathbf{F}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \frac{N}{2\sigma^4} \end{bmatrix}. \quad (3.17)$$

La BCR est alors donnée par :

$$\mathbf{F}^{-1}(\boldsymbol{\theta}') = \begin{bmatrix} \mathbf{F}^{-1}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \frac{2\sigma^4}{N} \end{bmatrix}. \quad (3.18)$$

Ce qui signifie qu'il suffit de calculer $\mathbf{F}^{-1}(\boldsymbol{\theta})$ pour retrouver $\mathbf{F}^{-1}(\boldsymbol{\theta}')$. Par conséquent, on se concentre ci-après sur le calcul de $\mathbf{F}^{-1}(\boldsymbol{\theta})$.

En tenant compte de l'équation (3.16) ainsi que de la structure du vecteur $\boldsymbol{\theta}$, la MIF peut être écrite sous forme d'une matrice à neuf blocs représentant les dérivées partielles par rapport aux éléments des trois vecteurs $\mathbf{p}$, $\mathbf{a}$, et ϕ :

$$\mathbf{F}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \begin{bmatrix} \left(\frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial p_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial A_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial \phi_m} \right) \\ \left(\frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial p_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial A_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial \phi_m} \right) \\ \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial p_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial A_m} \right) & \left(\frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial \phi_m} \right) \end{bmatrix}, \quad (3.19)$$

où $l, m = 1, \dots, n + 1$ pour les éléments de $\mathbf{p}$ et $l, m = 1, \dots, d$ pour les éléments de $\mathbf{a}$ et ϕ .

Puisque cette matrice est symétrique (à cause de la symétrie des dérivées partielles d'ordre deux) il suffit pour avoir tous les éléments de cette dernière de calculer les six

3.2. CARACTÉRISATION PARAMÉTRIQUE DES TRANSITOIRES DE COURANT

termes : $\frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial p_m}, \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial A_m}, \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial \phi_m}, \frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial A_m}, \frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial \phi_m}, \frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial \phi_m}$. Après calcul, nous trouvons :

$$\begin{aligned}
 \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial p_m} &= \sum_{k=0}^{N-1} \left(\mathbf{a}^T \mathbf{h}_k e^{\mathbf{p}^T \mathbf{t}_k} \right)^2 t_k^{l+m} \\
 \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial A_m} &= \sum_{k=0}^{N-1} \left(\mathbf{a}^T \mathbf{h}_k e^{\mathbf{p}^T \mathbf{t}_k} t_k^l \right) \cos(2\pi f_m t_k + \phi_m) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right) \\
 \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial \phi_m} &= - \sum_{k=0}^{N-1} \left(\mathbf{a}^T \mathbf{h}_k e^{\mathbf{p}^T \mathbf{t}_k} t_k^l \right) A_m \sin(2\pi f_m t_k + \phi_m) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right) \\
 \frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial A_m} &= \sum_{k=0}^{N-1} \cos(2\pi f_l t_k + \phi_l) \cos(2\pi f_m t_k + \phi_m) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2 \\
 \frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial \phi_m} &= - \sum_{k=0}^{N-1} A_m \cos(2\pi f_l t_k + \phi_l) \sin(2\pi f_m t_k + \phi_m) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2 \\
 \frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial \phi_m} &= \sum_{k=0}^{N-1} A_l A_m \sin(2\pi f_l t_k + \phi_l) \sin(2\pi f_m t_k + \phi_m) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2
 \end{aligned} \tag{3.20}$$

avec $\mathbf{h}_k = [\cos(2\pi f_1 t_k + \phi_1), \dots, \cos(2\pi f_d t_k + \phi_d)]^T$ et $\mathbf{t}_k = [1, t_k, \dots, t_k^n]^T$.

Finalement la BCR est égale à $\mathbf{F}^{-1}(\boldsymbol{\theta})$ en insérant les expressions (3.20) dans (3.19).

BCR du vecteur de paramètres $\boldsymbol{\theta}'' = [\mathbf{p}, \mathbf{a}, \phi, f_0]^T$

Nous considérons cette fois que la fréquence fondamentale f_0 n'est plus fixe et connue mais est un paramètre inconnu à estimer. Ceci permettrait de voir quels sont les paramètres qui seront susceptibles d'être affectés par l'estimation de f_0 .

Tout comme $\mathbf{F}(\boldsymbol{\theta}')$, $\mathbf{F}(\boldsymbol{\theta}'')$ est formée par la matrice $\mathbf{F}(\boldsymbol{\theta})$ à laquelle on ajoute une colonne et une ligne. Retrouver $\mathbf{F}(\boldsymbol{\theta}'')$ se résume alors au calcul des termes : $\frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial f_0}$,

$\frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial f_0}, \frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial f_0}, \frac{\partial \mathbf{s}^T}{\partial f_0} \frac{\partial \mathbf{s}}{\partial f_0}$. On trouve :

$$\begin{aligned}
 \frac{\partial \mathbf{s}^T}{\partial p_l} \frac{\partial \mathbf{s}}{\partial f_0} &= -2\pi \sum_{k=0}^{N-1} e^{\mathbf{p}^T \mathbf{t}_k} \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right) \mathbf{a}^T \mathbf{h}_k \mathbf{a}^T \mathbf{g}_k t_k^{l+1} \\
 \frac{\partial \mathbf{s}^T}{\partial A_l} \frac{\partial \mathbf{s}}{\partial f_0} &= -2\pi \sum_{k=0}^{N-1} t_k \cos(2\pi f_l t_k + \phi_l) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2 \mathbf{a}^T \mathbf{g}_k \\
 \frac{\partial \mathbf{s}^T}{\partial \phi_l} \frac{\partial \mathbf{s}}{\partial f_0} &= 2\pi \sum_{k=0}^{N-1} t_k A_l \sin(2\pi f_l t_k + \phi_l) \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2 \mathbf{a}^T \mathbf{g}_k \\
 \frac{\partial \mathbf{s}^T}{\partial f_0} \frac{\partial \mathbf{s}}{\partial f_0} &= \sum_{k=0}^{N-1} \left(2\pi t_k \left(e^{\mathbf{p}^T \mathbf{t}_k} + 1 \right)^2 \mathbf{a}^T \mathbf{g}_k \right)^2
 \end{aligned} \tag{3.21}$$

avec $\mathbf{h}_k = [\cos(2\pi f_1 t_k + \phi_1), \dots, \cos(2\pi f_d t_k + \phi_d)]^T$, $\mathbf{t}_k = [1, t_k, \dots, t_k^n]^T$
 et $\mathbf{g}_k = [\sin(2\pi f_0 t_k + \phi_1), \sin(2\pi 3f_0 t_k + \phi_2), \dots, \sin(2\pi(2d-1)f_0 t_k + \phi_d)]^T$.

3.3 Performances expérimentales de l'algorithme

3.3.1 Évaluation sur données simulées

Dans cette section, nous allons présenter les résultats de l'évaluation des performances d'estimation sur des données simulées. Ci-après, on présente :

- A le signal simulé et ses paramètres ;
- B le biais des paramètres estimés ;
- C la variance des paramètres estimés et sa comparaison avec les BCR ;
- D la variation de la BCR des paramètres en fonction de la fréquence d'échantillonnage ;
- E la convergence de l'algorithme d'estimation ;
- F la variance du signal estimé et sa comparaison avec la BCR du signal ;
- G l'influence de la connaissance de la fréquence fondamentale sur les BCR.

A Signal simulé et ses paramètres

Pour les tests sur données simulées on a choisi un jeu de paramètres fixé de sorte à avoir une signature qui se rapproche le plus possible d'une signature réelle. Pour fixer ce jeu de paramètres on a fait tourner l'algorithme d'estimation sur un signal réel et en se basant sur les valeurs des paramètres estimés on a fixé les paramètres du signal simulé qui seront utilisés tout au long de cette section d'évaluation sur données simulées.

Le modèle du signal non bruité pour rappel est le suivant :

$$s(t_k) = e(t_k)s_s(t_k) = \left(e^{\mathbf{p}^T \mathbf{t}} + 1\right) \sum_{i=1}^d A_i \cos(2\pi f_i t_k + \phi_i), \quad t_k \in [0, T] \quad (3.22)$$

avec $\mathbf{p} = [p_0, p_1, \dots, p_n]^T$, $\mathbf{t} = [1, t_k, \dots, t_k^n]^T$, $A_i (\geq 0)$, $\phi_i \in [-\pi, \pi]$ et $f_i = (2i - 1)f_0$, $i = 1, \dots, d$ avec la fréquence fondamentale f_0 qui est fixe et égale à 50 Hz.

Les valeurs choisies pour le signal simulé sont comme suit :

- $F_s = 30$ kHz : fréquence d'échantillonnage ;
- $T = 3$ s : durée du signal ;
- $n = 3$: degré du polynôme, $d = 5$: nombre d'harmoniques ;
- $\mathbf{p} = [1.9, -9, 8.5, -4]^T$;
- $\mathbf{a} = [1.8, 0.5, 0.2, 0.1, 0.05]^T$;
- $\boldsymbol{\phi} = [-3, 3, 2.5, 1.5, 1]^T$;
- $t_{e1} = 2$ s, $t_{e2} = 2.8$ s : instants définissant le régime établi.

Le signal simulé obtenu est illustré dans la figure 3.4.

Pour évaluer le biais et la variance des paramètres estimés, nous avons utilisé 1000 itérations Monte-Carlo. À des fins de comparaison, nous donnons aussi pour ces évaluations les résultats intermédiaires de la phase d'estimation ainsi que les résultats finaux après la phase d'optimisation.

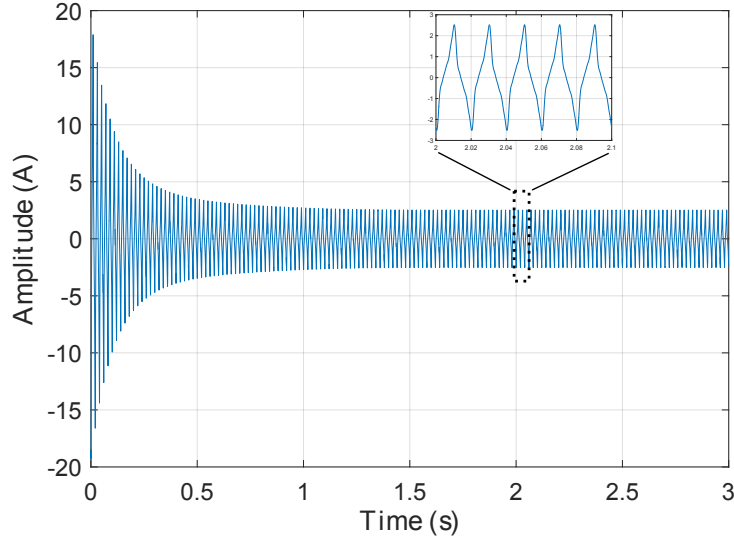


FIGURE 3.4 – Courant de mise en marche simulé pour utilisation dans l'évaluation avec données simulées.

B Biais des paramètres estimés

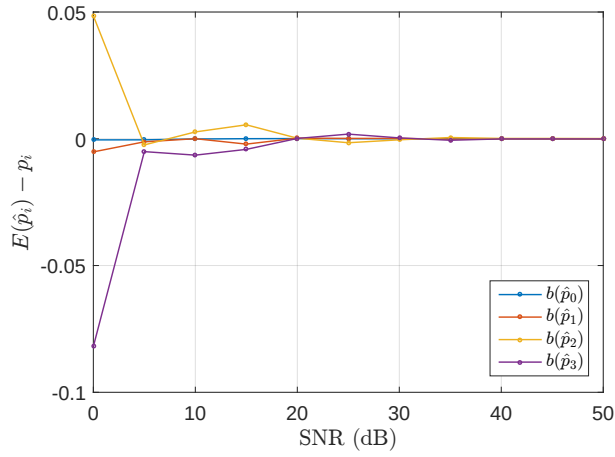
Comme mentionné précédemment, afin de pouvoir utiliser les BCR pour l'évaluation de la performance d'estimation, il faut d'abord vérifier que les paramètres estimés avec notre algorithme ECT sont non biaisés.

La figure 3.5 donne les biais calculés pour les paramètres estimés en ignorant la phase d'optimisation et ceci en fonction du Rapport Signal à Bruit (RSB) (Signal-to-Noise Ratio-SNR). La figure 3.6 donne, quant à elle, les mêmes résultats mais cette fois en considérant la phase d'optimisation.

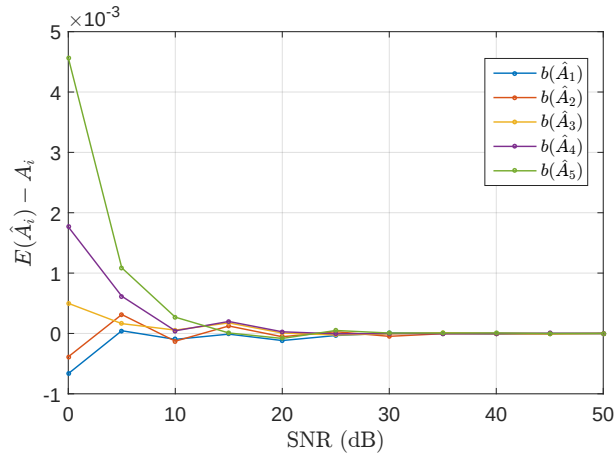
On retrouve les mêmes tendances pour les deux figures. Tous les paramètres estimés ont un biais quasi nul à partir de la valeur $RSB = 30$ dB. Entre $RSB = 10$ dB et $RSB = 30$ dB, les paramètres estimés restent très faiblement biaisés. Au dessous de $RSB = 10$ dB, on commence à avoir des biais, ce qui attendu vu le niveau de bruit. Ces biais restent, néanmoins, faibles.

On conclut suite à ces résultats que nos estimations sont **non biaisées**.

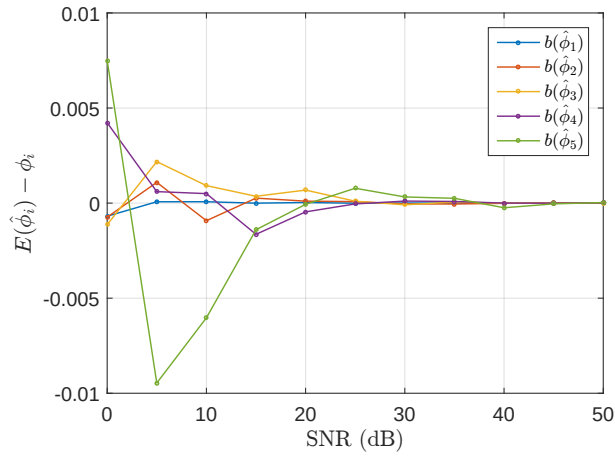
3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME



(a) Biais de $\hat{\mathbf{p}}$.



(b) Biais de $\hat{\mathbf{A}}$.



(c) Biais de $\hat{\boldsymbol{\phi}}$.

FIGURE 3.5 – Biais des différents paramètres estimés $\hat{\mathbf{p}}$, $\hat{\mathbf{A}}$, $\hat{\boldsymbol{\phi}}$ en considérant l'estimation du $\mathbf{p}$ indépendamment de $\mathbf{A}$ et $\boldsymbol{\phi}$ sans phase d'optimisation. 1000 itérations Monte-Carlo.

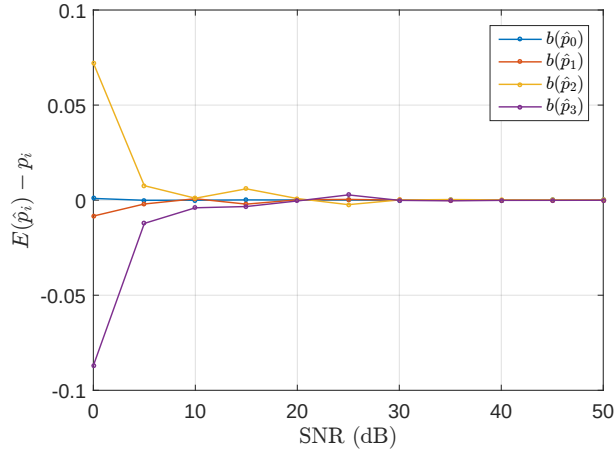
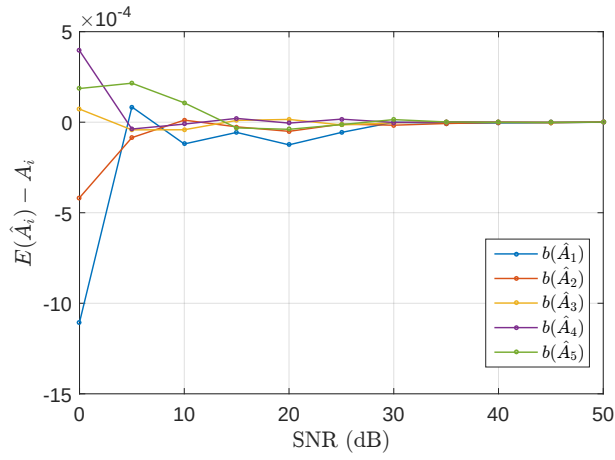
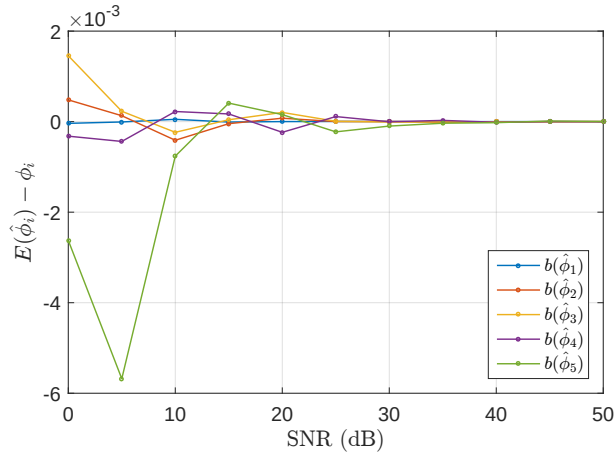

 (a) Biais de $\hat{\mathbf{p}}$.

 (b) Biais de $\hat{\mathbf{A}}$.

 (c) Biais de $\hat{\phi}$.

 FIGURE 3.6 – Biais des différents paramètres estimés $\hat{\mathbf{p}}$, $\hat{\mathbf{A}}$, $\hat{\phi}$ en considérant l'estimation du $\mathbf{p}$ indépendamment de $\mathbf{A}$ et ϕ avec une phase optimisation. 1000 itérations Monte-Carlo.

C Variance des paramètres estimés

Le calcul de la variance des paramètres estimés se fait en utilisant des itérations Monte-Carlo (ici 1000 itérations). De la même façon que pour l'étude du biais, on va chercher à quantifier le gain que ramène la phase d'optimisation par rapport à la phase estimation. La variance estimée pour chaque paramètre est comparée par la suite à la BCR correspondante.

Pour l'estimation de $\mathbf{p}$ (figure 3.7) on note déjà que pour la plupart des éléments de ce vecteur, il n'y a quasiment pas de gain entre l'estimation avec et sans phase d'optimisation mis à part pour le paramètre p_0 où on gagne à peu près 1,5 dB. Ceci est dû au fait que l'on utilise déjà tous les échantillons de mesure à notre disposition dans la phase d'estimation. Cela signifie que durant la phase d'estimation, l'algorithme exploite tout ce qu'il y avait à exploiter comme information disponible pour l'estimation et l'optimisation n'ajoute, donc, quasiment rien aux résultats.

Le gain noté pour le paramètre p_0 pourrait être dû au fait que c'est un paramètre indépendant du temps. Ce paramètre peut être vu comme un facteur d'échelle qui vient modifier l'amplitude du signal $s(t)$. Dans ce cas notre modèle peut s'écrire :

$$s(t) = \left(e^{p_0} e^{\mathbf{p}'^T \mathbf{t}} + 1 \right) s_s(t), \quad (3.23)$$

avec $\mathbf{p}' = [p_1, \dots, p_n]^T$. Ceci signifie que ce paramètre modifie l'amplitude de $s(t)$ au même titre que les amplitudes des sinusoides où chaque A_i peut être vu comme un facteur d'échelle pour la sinusoïde d'ordre i . La conséquence de ce lien est qu'une mauvaise estimation des A_i va se répercuter sur l'estimation du p_0 car l'algorithme d'estimation essaye de compenser l'erreur d'estimation liée aux A_i en modifiant la valeur de p_0 .

La deuxième chose à noter est le fait que les variances de $\hat{\mathbf{p}}$ (algorithme avec optimisation) atteignent les BCR, ce qui signifie que l'estimateur de $\mathbf{p}$ est efficace. Contrairement au cas de l'estimation de $\mathbf{p}$, on peut gagner de 3 à 10 dB avec l'optimisation sur les paramètres $\mathbf{a}$ et ϕ (figures 3.8 et 3.9). Ceci est la conséquence directe de l'utilisation d'une faible portion du régime établi du signal pour la phase d'estimation. La considération de toute la durée du signal pour l'optimisation va donc améliorer les performances.

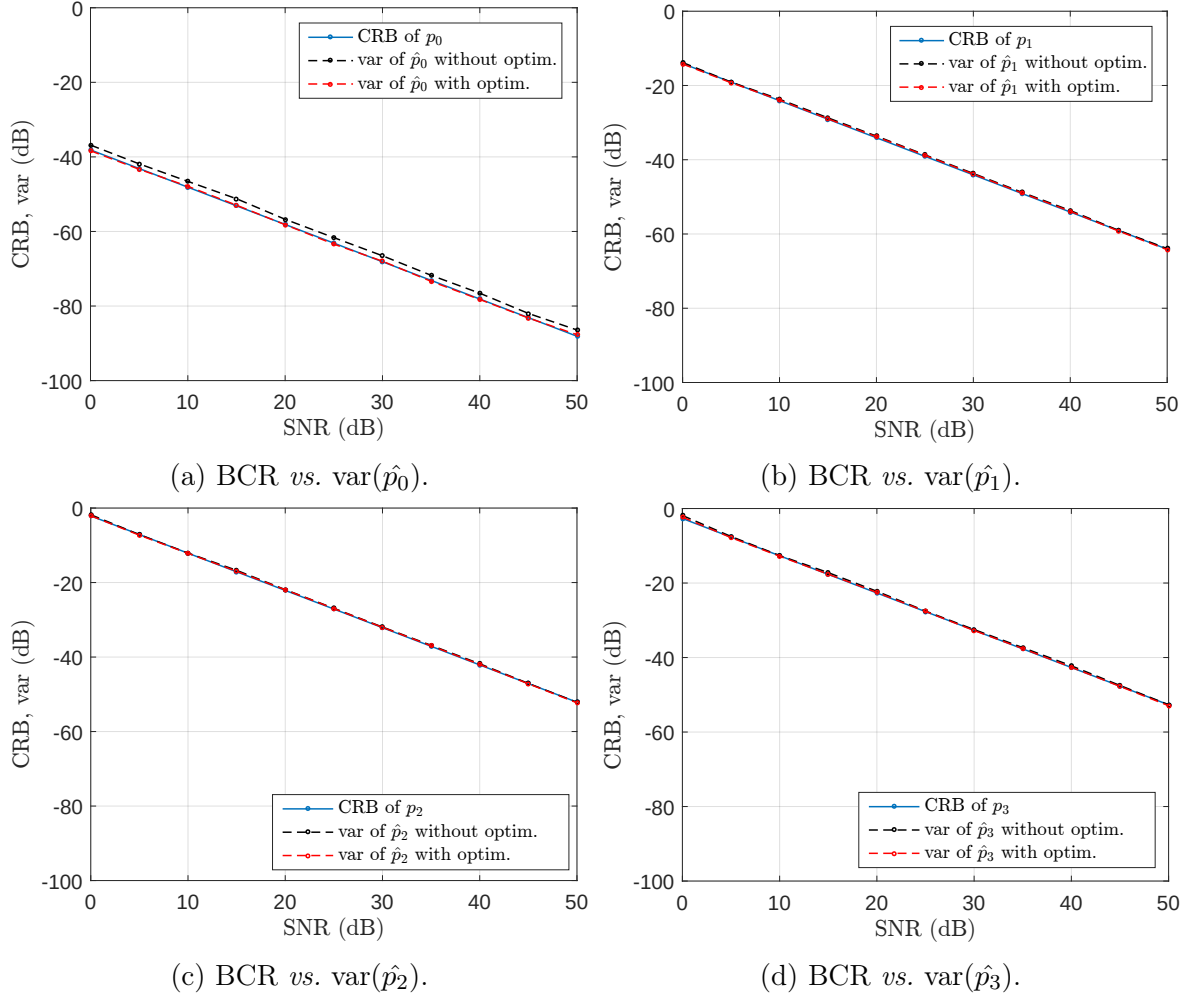


FIGURE 3.7 – Comparaison de la BCR (CRB) et variance du paramètre $\hat{\mathbf{p}}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.

Notons que de même que pour le $\hat{\mathbf{p}}$, les variances de $\hat{\mathbf{a}}$ et de $\hat{\boldsymbol{\phi}}$ atteignent les BCR.

En conclusion, les estimateurs des paramètres du modèle basés sur l'algorithme ECT sont **efficaces** vu qu'ils sont non biaisés et atteignent les BCR.

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

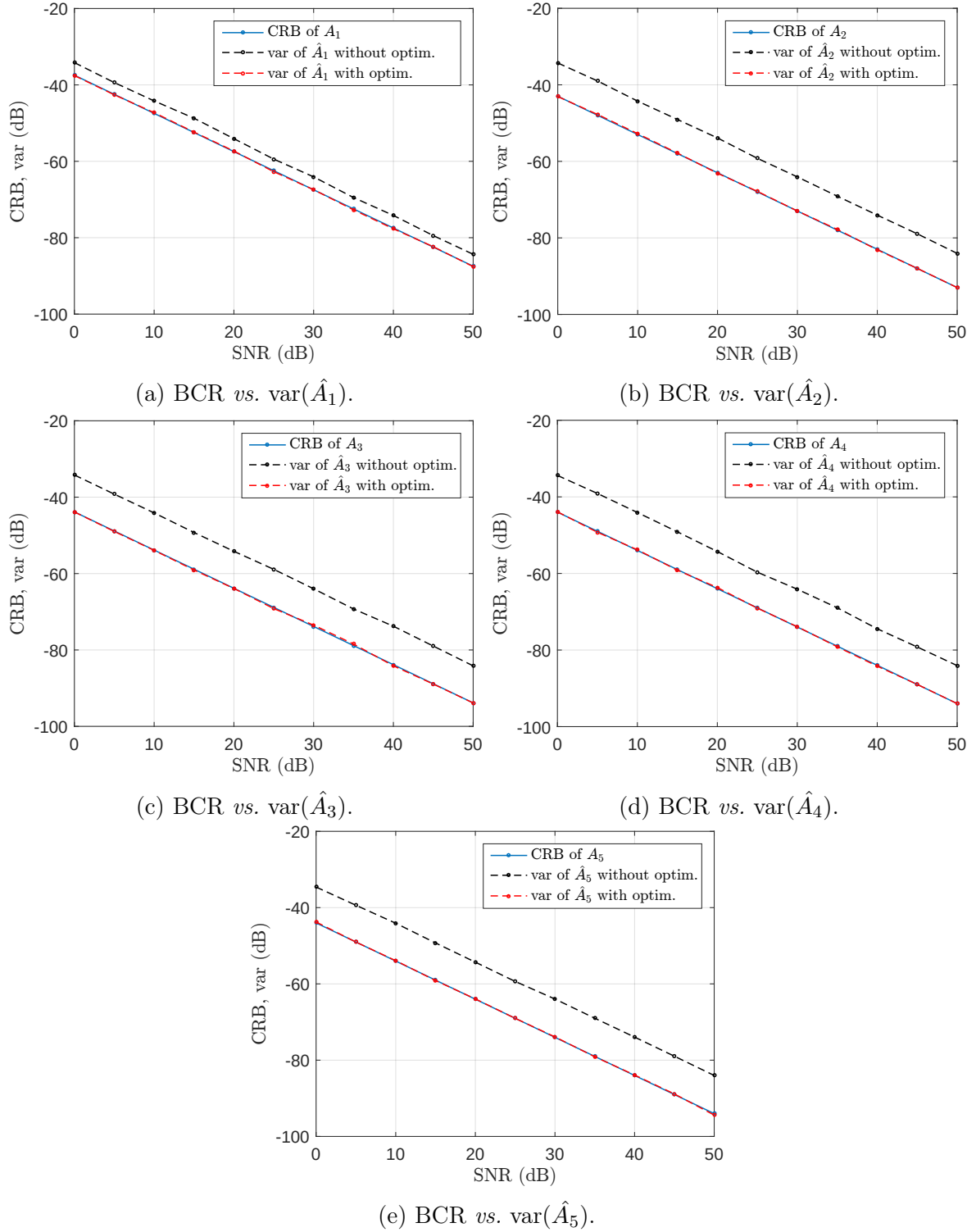


FIGURE 3.8 – Comparaison de la BCR (CRB) et variance du paramètre $\hat{\mathbf{A}}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

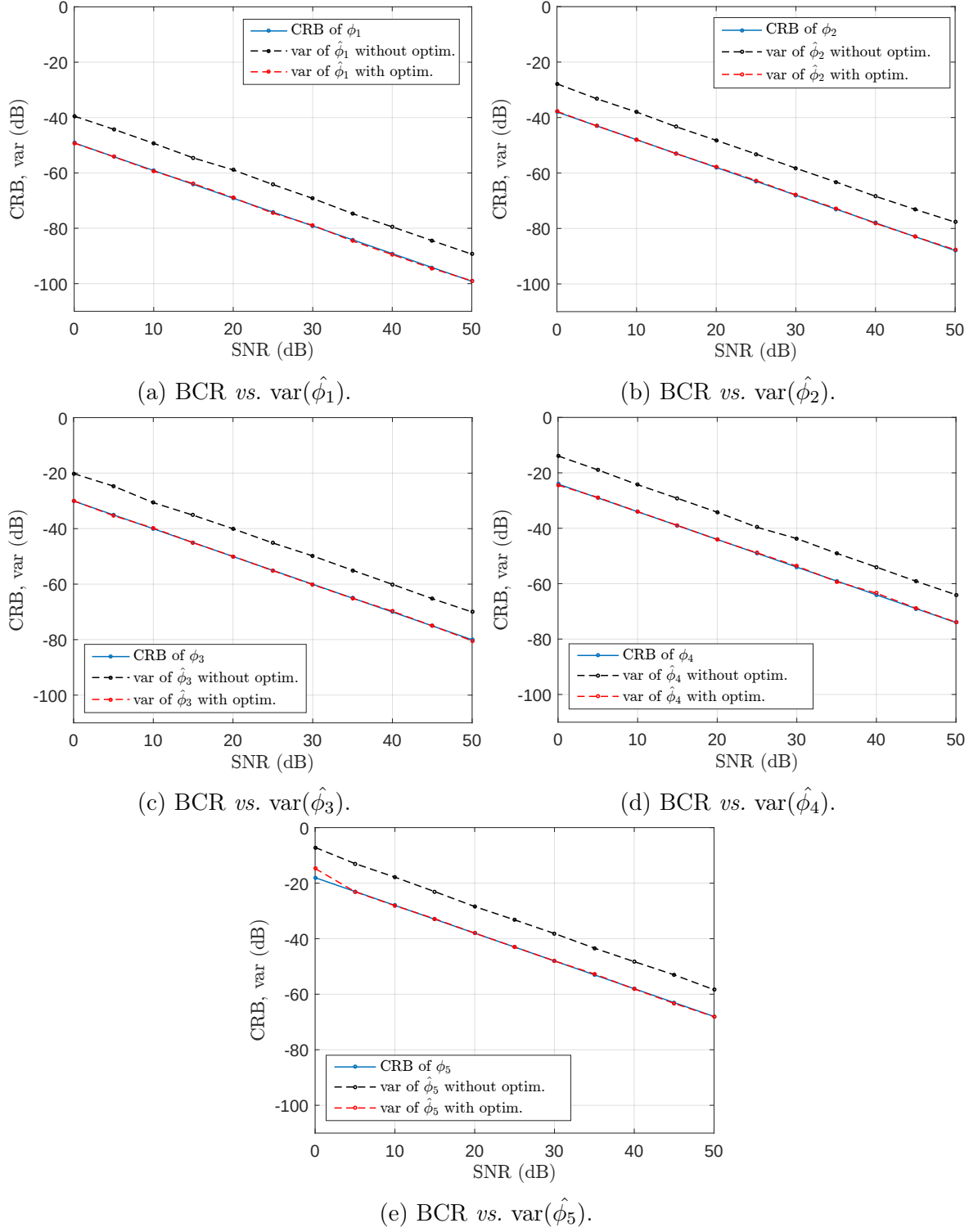


FIGURE 3.9 – Comparaison de la BCR (CRB) et variance du paramètre $\hat{\phi}$ en considérant l'estimation avec et sans optimisation. 1000 itérations Monte-Carlo.

D Variation de la BCR en fonction de la fréquence d'échantillonnage

De part la nature transitoire des phénomènes observés, il faut veiller à bien fixer la fréquence d'échantillonnage F_s pour la mesure. On cherche une fréquence d'échantillonnage suffisamment grande pour permettre de capter les phénomènes transitoires sans être trop importante pour limiter le coût de calcul. La BCR des paramètres va nous permettre d'évaluer l'impact du choix d'une fréquence donnée sur les bornes inférieures des variances des paramètres et suivant cela décider sur une fréquence d'échantillonnage avec le compromis : lourdeur de calculs *vs.* performance souhaitée.

La figure 3.10 donne la variation des BCR des paramètres en fonction de F_s sur une échelle logarithmique où F_s varie entre 1 kHz et 100 kHz.

Les résultats montrent qu'une augmentation de F_s s'accompagne d'une meilleure qualité d'estimation des paramètres. Ceci semble logique, vu qu'une augmentation de F_s signifie une augmentation du nombre de points de mesure (sur une durée fixe) et donc des informations supplémentaires susceptibles d'être captées. En situation réelle, cela n'est pas nécessairement vrai. La blancheur du bruit, initialement vérifiée pour des F_s relativement faibles, peut ne plus être valable en augmentant F_s car les échantillons deviennent de plus en plus proches et il devient difficile de supposer leur indépendance. Dans ce cas, les BCR calculées pour un bruit blanc ne peuvent plus être utilisées pour juger de la performance.

D'un point de vue pratique, répondre à la question du *bon* choix de la fréquence d'échantillonnage n'est pas une tâche facile et dépend de plusieurs paramètres. D'après notre expérience, et pour l'étude des transitoires de mise en marche, une fréquence au moins égale à 5 kHz est souhaitable et aller au delà de 100 kHz commence à alourdir le traitement des données. Une fréquence de 30 kHz semble permettre un bon compromis, d'où le choix que nous avons fait pour F_s des simulations.

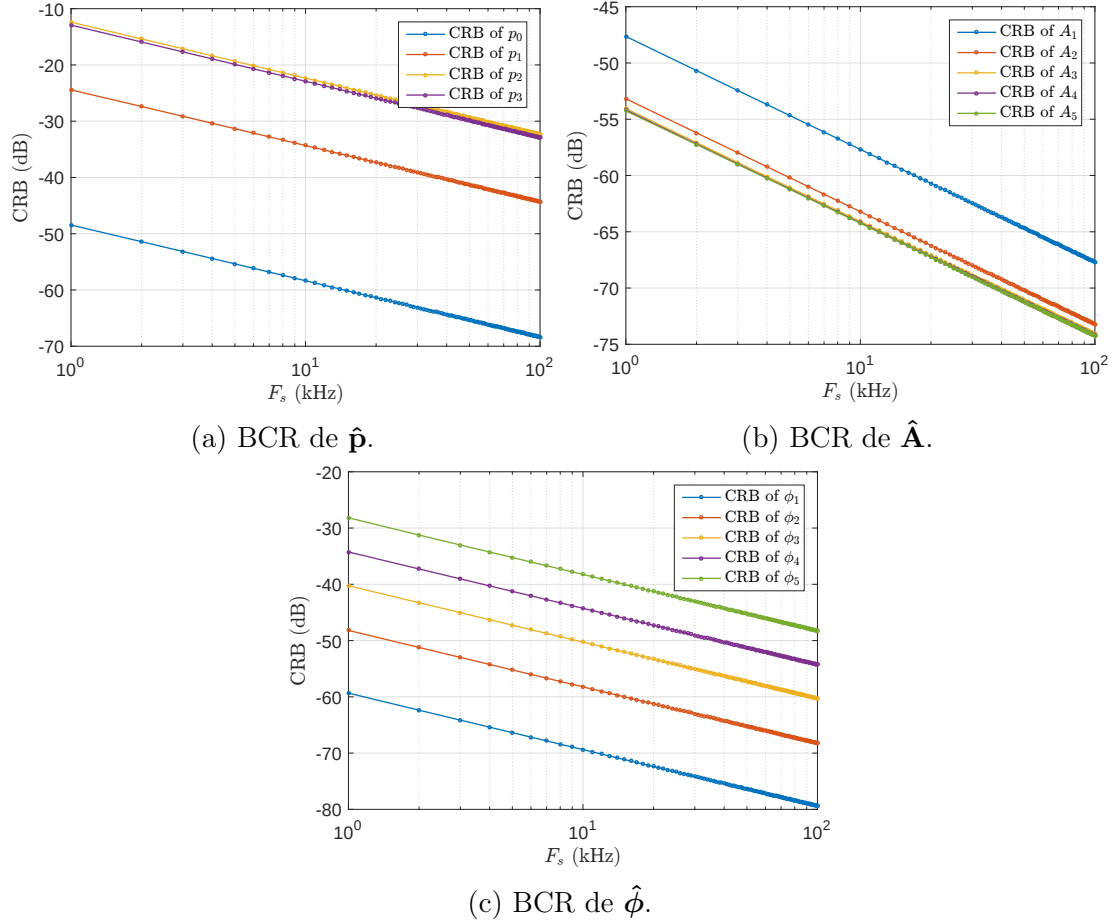


FIGURE 3.10 – BCR des différents paramètres estimés $\hat{\mathbf{p}}$, $\hat{\mathbf{A}}$, $\hat{\boldsymbol{\phi}}$ en fonction de la fréquence d'échantillonnage F_s (de 1 kHz à 100 kHz). Calculs faits à RSB = 25 dB.

E Convergence de l'algorithme d'estimation–optimisation

Du fait que l'estimation utilise un algorithme d'optimisation, il se pose le problème de la convergence de ce dernier. Le nombre d'itérations nécessaires pour la convergence d'un algorithme est un des paramètres susceptible de donner une indication sur la possibilité ou non d'opérer en temps réel.

Le même algorithme d'optimisation non linéaire est utilisé pour les deux phases d'estimation et d'optimisation. On étudie leur convergence séparément. La figure 3.11 donne, pour plusieurs RSB, la courbe de variation de la EQM en fonction du nombre d'itérations de l'algorithme en phase d'estimation. Quelque soit le RSB, l'algorithme converge en moins de onze itérations. Ce nombre d'itérations reste, cependant, dépendant d'un (ou plusieurs) critère (s) d'arrêt, paramétrés. Citons par exemple, la variation de l'EQM

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

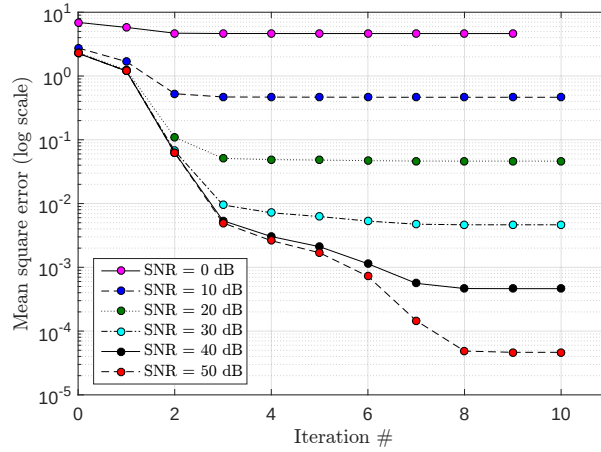


FIGURE 3.11 – Erreur quadratique moyenne $\frac{1}{N} \sum_{n=1}^N |x(n) - s(\hat{\theta}, n)|^2$ en fonction du nombre d'itérations de l'algorithme d'estimation utilisant l'algorithme trust-region-reflective pour différentes valeurs du RSB. Pour chaque courbe, le dernier point correspond à l'arrêt de l'algorithme (convergence).

comparée à un seuil donné.

La convergence de l'algorithme en phase d'optimisation est beaucoup plus rapide. L'algorithme étant mieux initialisé, il converge plus vite. La figure 3.12 donne cette fois la ΔEQM en fonction du nombre d'itérations pour plusieurs RSB. Étant plus proche du minimum, afficher la ΔEQM est plus adaptée pour voir les variations que l' EQM . L'algorithme converge au plus après trois itérations.

Ces résultats montrent que l'algorithme a une convergence plutôt rapide (treize itérations). Cette simulation permet de donner une idée grossière de la convergence de l'algorithme dans un cas simulé où le signal test suit le modèle proposé. Pratiquement, la convergence de l'algorithme dépend énormément du signal réel à estimer. Si ce dernier s'éloigne du modèle, l'algorithme va passer plus de temps à chercher le minimum et à converger.

F Variance du signal estimé

Ayant les BCR des paramètres du modèle, il est intéressant de voir la BCR du signal lui même. Ceci est possible grâce à une formule de passage qui permet d'avoir la BCR du signal en fonction de la BCR du vecteur de paramètres θ . Cette formule permet de retrouver la BCR d'une fonction $f(\theta)$ continue et différentiable (notée $BCR_{f(\theta)}$) en

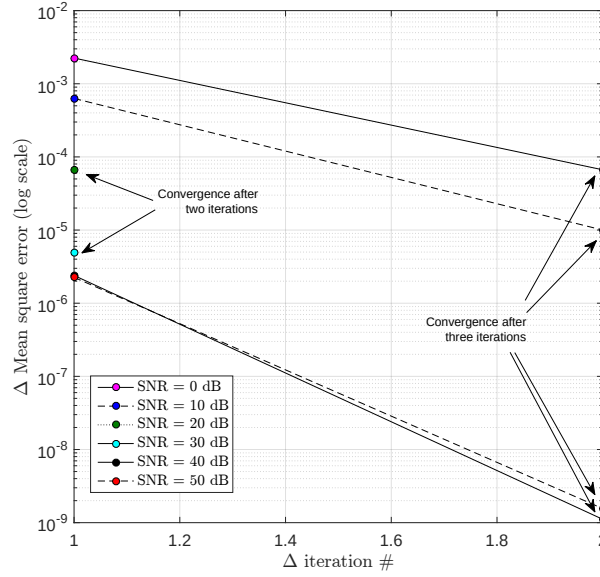


FIGURE 3.12 – Δ Erreur quadratique moyenne $\frac{1}{N} \sum_{n=1}^N |x(n) - s(\hat{\theta}, n)|^2$ entre deux itérations consécutives en fonction du nombre d'itérations de l'algorithme d'optimisation utilisant l'algorithme trust-region-reflective pour différentes valeurs du RSB. L'itération avec un $\Delta EQM = 0$ (minimum atteint) n'apparaissant pas sur ce graphique semi-logarithmique nous indiquons (flèches) ci-contre les itérations qui précèdent.

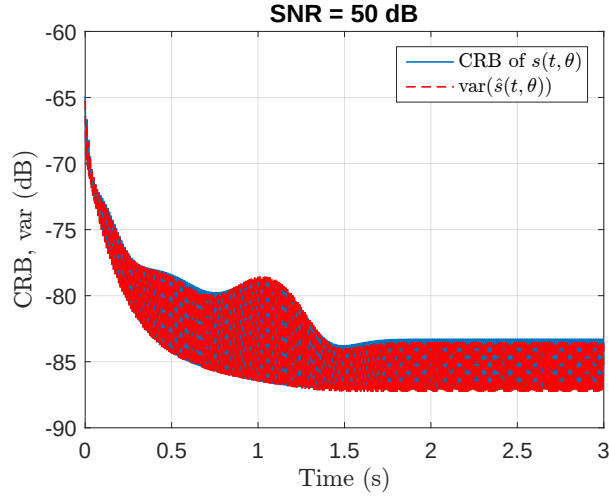
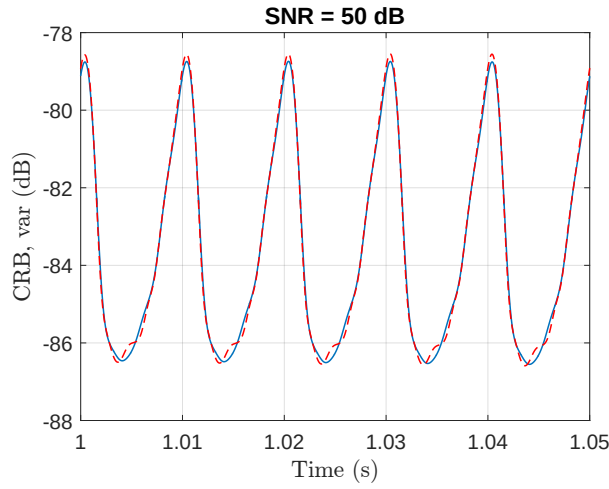
utilisant la BCR de ce paramètre θ (notée BCR_{θ}). Elle est donnée par (Friedlander and Francos, 1995) :

$$BCR_{f(\theta)} = \left(\frac{\partial f(\theta)}{\partial \theta} \right)^T BCR_{\theta} \left(\frac{\partial f(\theta)}{\partial \theta} \right) \quad (3.24)$$

Notre fonction $f(\theta)$ étant égale au modèle $s(t, \theta)$, on obtient une BCR qui dépend du temps comme suit :

$$BCR_{s(t, \theta)}(t) = \left(\frac{\partial s(t, \theta)}{\partial \theta} \right)^T BCR_{\theta} \left(\frac{\partial s(t, \theta)}{\partial \theta} \right) \quad (3.25)$$

Cette BCR est alors à comparer avec la variance estimée du signal qui est aussi une fonction du temps. Pour l'avoir, on a fait 1000 itérations Monte-Carlo d'estimation du signal simulé décrit précédemment et à chaque fois on reconstruit le signal estimé. Après cela et pour chaque point temporel, on calcul la variance en utilisant le même point temporel des 1000 itérations. La figure 3.13 donne le résultat obtenu. À noter que ces simulations se font à un RSB donné. Le résultat montre que l'on atteint la BCR encore une fois. Ce qui est la conséquence directe du fait que les paramètres eux mêmes atteignent chacun sa borne. La figure 3.14 montre plus qualitativement cet écart entre la BCR du


(a) BCR *vs.* var


(b) Zoom de 3.13a

FIGURE 3.13 – BCR du signal $s(t, \theta)$ comparée à la variance, calculée point par point, du signal estimé $\hat{s}(t, \theta)$. Calculs faits à RSB = 50 dB.

signal et la variance du signal estimé et confirme les faibles écarts qui se trouvent entre -0,5 et 0,4 dB. Un autre tracé (figure 3.15) pour plusieurs RSB, où pour chaque RSB on calcule la moyenne (temporelle) de la variance du signal estimé ainsi que la moyenne de la BCR, permet de confirmer la tendance pour différents niveaux du RSB.

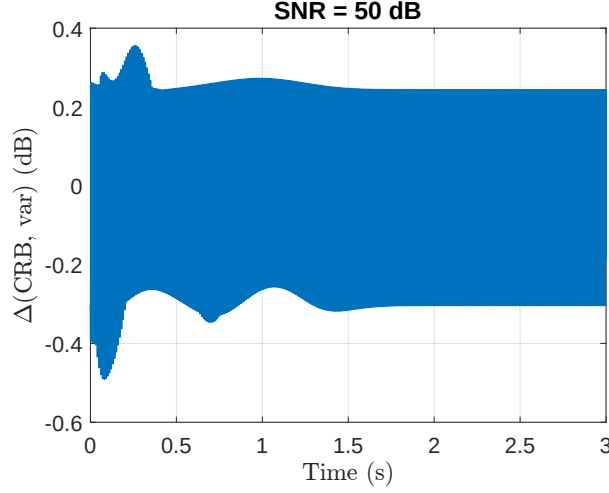


FIGURE 3.14 – Différence entre la BCR du signal $s(t, \theta)$ et la variance, calculée point par point, du signal estimé $\hat{s}(t, \theta)$ (figure 3.13a).

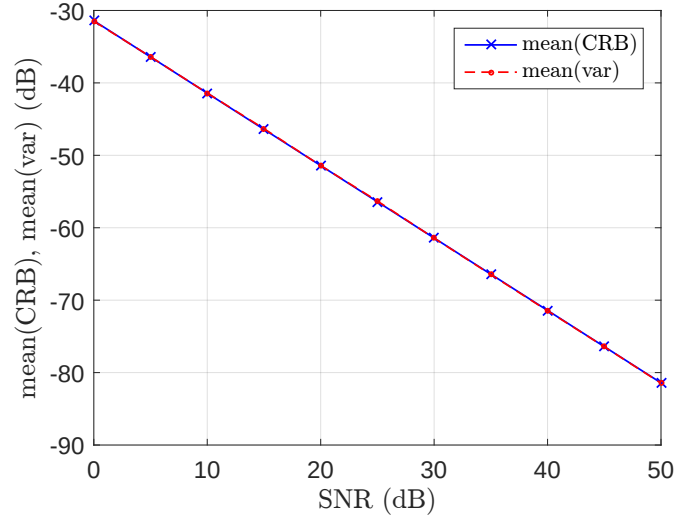


FIGURE 3.15 – Valeurs moyennes (par rapport au temps) de la BCR du signal $s(t, \theta)$ et la variance, calculée point par point, du signal estimé $\hat{s}(t, \theta)$ en fonction du RSB.

G Influence de la connaissance de la fréquence fondamentale sur les BCR

Suite à la discussion faite à la fin de la sous-section 3.2.4, on compare ici (figures 3.16, 3.17 et 3.18) la BCR du vecteur de paramètres $\theta'' = [\mathbf{p}, \mathbf{a}, \phi, f_0]^T$ avec la BCR du vecteur de paramètres $\theta = [\mathbf{p}, \mathbf{a}, \phi]^T$ pour notre signal simulé afin de voir d'éventuels changements après l'introduction de la fréquence fondamentale f_0 comme paramètre inconnu à estimer. Noter qu'en réalité ce qu'on est en train de comparer c'est les éléments diagonaux de la matrice BCR du paramètre $\theta'' = [\mathbf{p}, \mathbf{a}, \phi, f_0]^T$ avec les paramètres diagonaux de la matrice BCR du paramètre $\theta = [\mathbf{p}, \mathbf{a}, \phi]^T$.

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

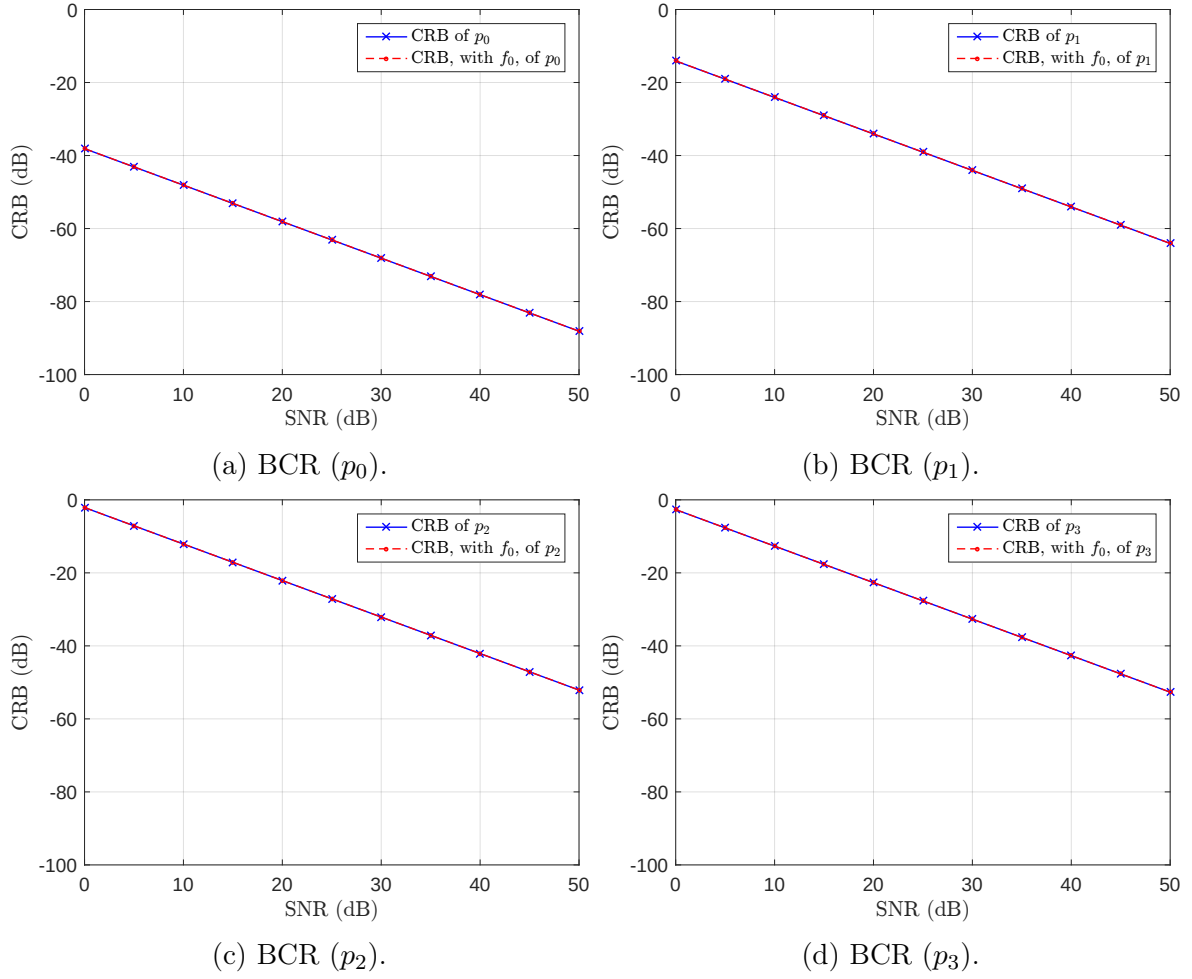


FIGURE 3.16 – Comparaison de la BCR de $\mathbf{p}$ avec et sans introduction de f_0 comme inconnue.

Ce que l'on peut tirer comme conclusion de ces résultats est que la connaissance (ou l'ignorance) de f_0 n'affecte que les performances d'estimation de ϕ . Cela est plutôt compréhensible vu que ces deux paramètres s'associent pour former les phases temporelles $\phi_i(t) = 2\pi(2i - 1)f_0t + \phi_i$ de notre modèle.

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

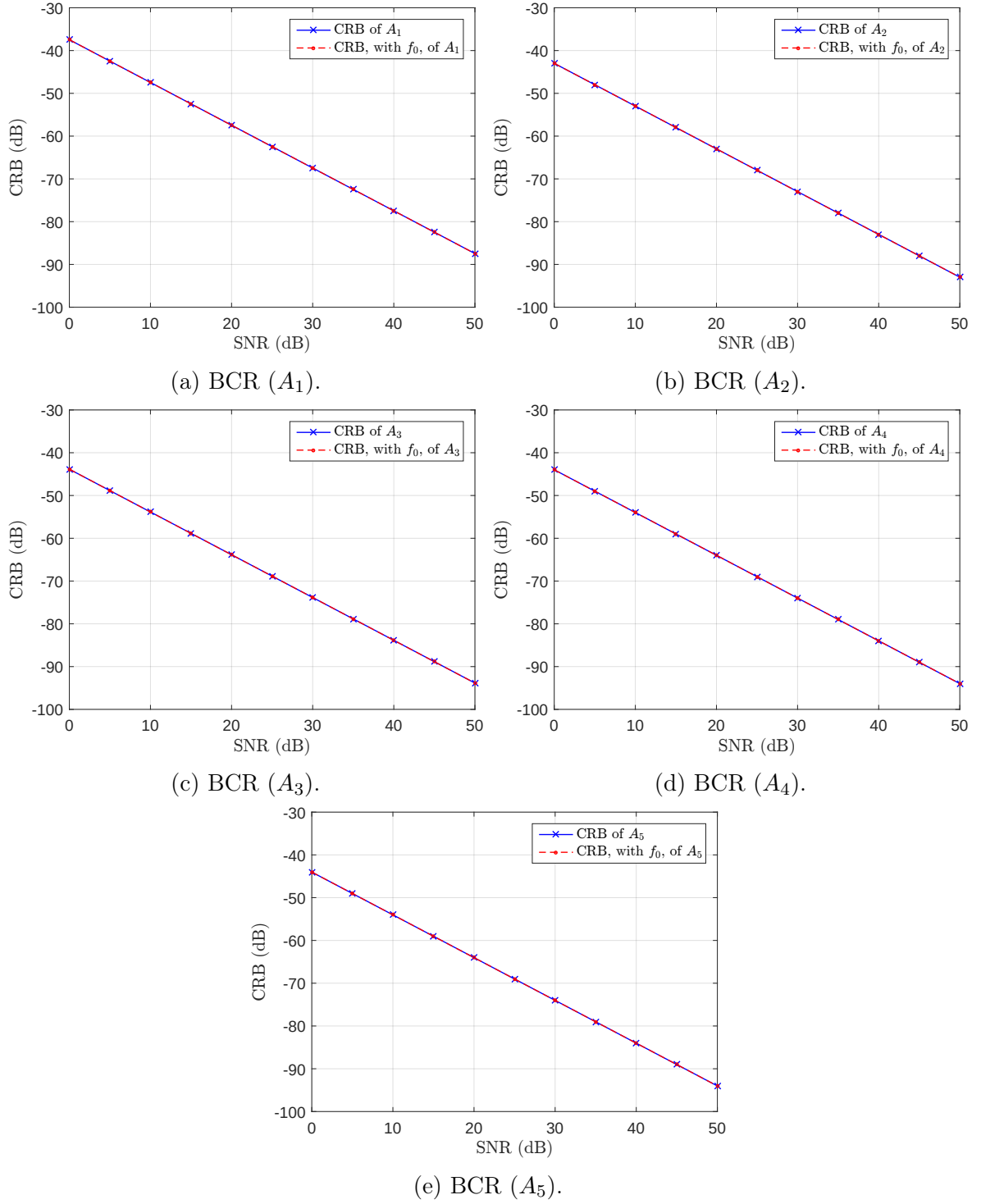


FIGURE 3.17 – Comparaison de la BCR de $\mathbf{a}$ avec et sans introduction de f_0 comme inconnue.

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

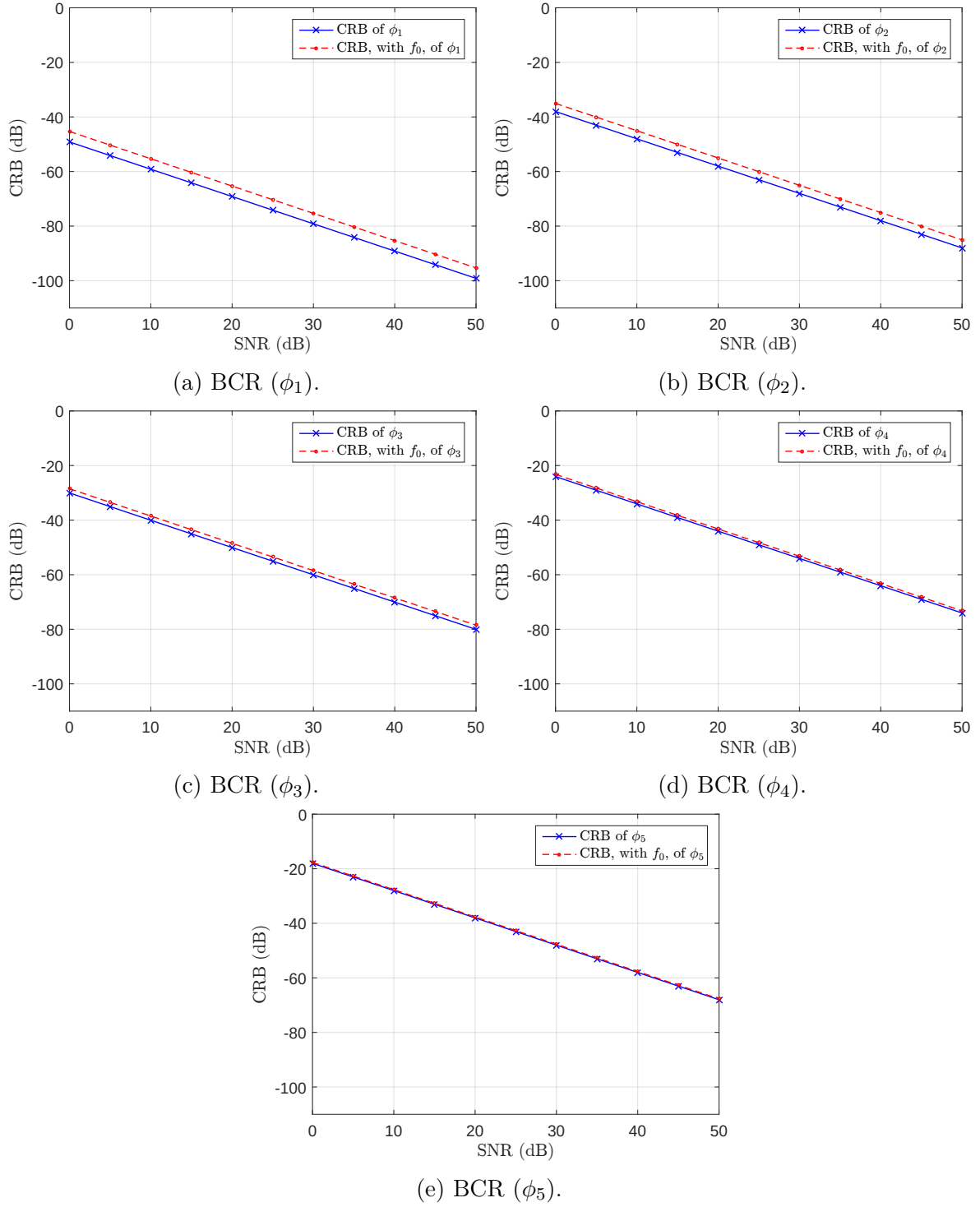


FIGURE 3.18 – Comparaison de la BCR de ϕ avec et sans introduction de f_0 comme inconnue.

3.3.2 Évaluation sur données réelles

Dans cette partie, nous donnons des exemples d'évaluation de l'algorithme ECT sur des données réelles. Jusqu'à maintenant, nous avons supposé implicitement, pour les données simulées, quelques hypothèses simplificatrices pour tester l'estimation des paramètres indépendamment de la performance des autres blocs. Ces hypothèses sont :

H1 une détection parfaite du transitoire de mise en marche ;

H2 une partie régime établi bien définie ;

H3 un transitoire qui démarre d'un maximum ;

H4 une fréquence fondamentale fixe et égale à 50 Hz.

En situation réelle, ces hypothèses ne sont pas vérifiées et ceci pour différentes raisons :

- notre précision de détection du transitoire de mise en marche n'est jamais parfaite ;
- la définition du régime établi dépend de la partie détection et est donc pas toujours bien définie ;
- physiquement, il y a toujours un temps de latence dont l'appareil a besoin avant que le signal atteigne une amplitude maximale ;
- à cause de l'équilibre offre-demande que le gestionnaire du réseau électrique doit assurer, la fréquence fondamentale n'est quasiment jamais fixe et égale à 50 Hz (figure 3.19).

La plus facile des simplifications à contourner en cas réel est **H3** où il suffit de détecter où le signal est à son maximum et de prendre ce point comme début pour le signal à estimer. Ceci suppose qu'on a un seul appareil en fonctionnement. Si ce n'est pas le cas, alors il faut d'abord détecter le transitoire de mise en marche pour limiter l'analyse et estimation à un appareil unique.

Pour les hypothèses **H1** et **H2**, on utilise le détecteur HAND (sous-section 3.4.1) que nous avons proposé et que nous avons construit en ayant à l'esprit les contraintes citées ci-dessus.

Pour **H4**, une solution serait d'estimer la fréquence fondamentale f_0 afin de l'intégrer par la suite dans l'algorithme d'estimation. Nous proposons dans cette sous-section une évaluation en supposant que l'hypothèse **H4** est vraie.

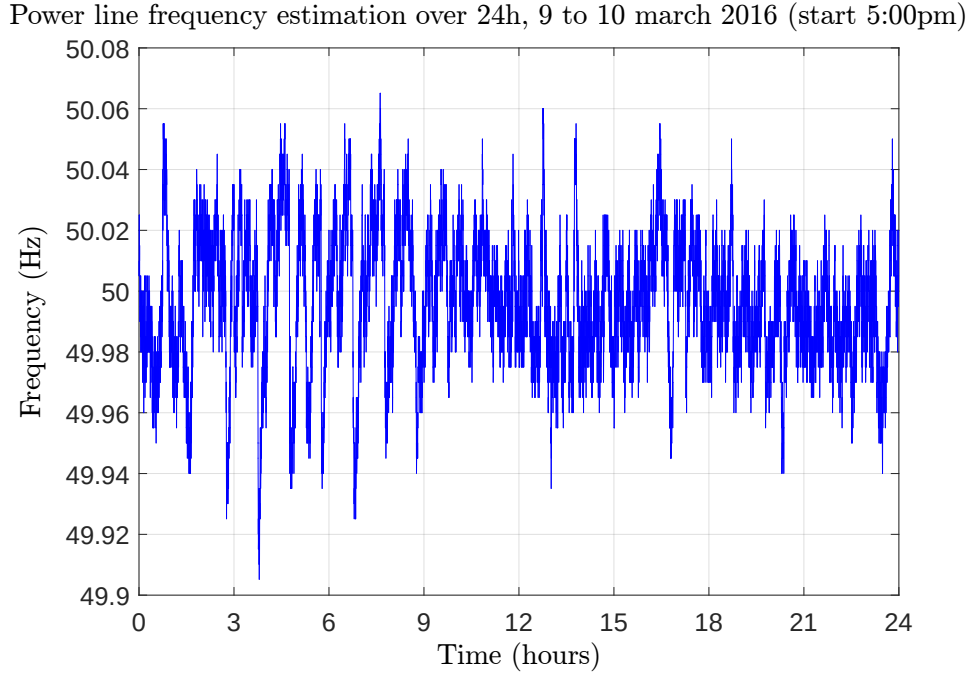


FIGURE 3.19 – Fréquence fondamentale du réseau estimée sur 24h entre le 9 et le 10 mars 2016 (début de mesure à 17 :00). La méthode d'estimation est basée sur la méthode de passages par zéros simple. Sur le graphe on affiche un point par seconde.

Nous verrons par la suite que nous serons amenés à lever cette hypothèse et la même évaluation sera proposée dans la section suivante avec estimation de f_0 .

Estimation avec ECT d'un signal réel de notre base de transitoires

Comme nous allons utiliser les signaux de notre base de transitoires contrôlés (COOLL) (sous-section 4.2.4) et où on connaît les instants de mise en marche et de mise à l'arrêt, nous supposons que les hypothèses **H1** et **H2** sont vérifiées. De plus, nous prenons le signal à partir de son maximum afin de vérifier **H3**. Cela va nous permettre de se limiter à l'étude de l'effet de **H4**.

Les signaux de COOLL sont échantillonnés à $F_s = 100$ kHz. Nous choisissons pour le test d'estimation un signal de ventilateur (figure 3.20) déclenché avec un délai de 5 ms par rapport à la sinusoïde du réseau. La durée totale de mesure est de 6 s avec 0,5 s de pré-déclenchement et 1 s de post-déclenchement. Les pré et post-déclenchement sont les durées des parties mesurées où l'appareil est hors tension. Pour la partie régime établi, nous prenons la dernière demi-seconde avant l'arrêt de fonctionnement. Nous fixons le degré du polynôme $n = 3$ et $d = 5$ sinusoïdes (algorithme 1).

$$\begin{aligned}
& \hat{\mathbf{p}} = [-1.21, 0.30, -0.43, -7.19 \times 10^{-7}]^T \\
& \hat{\mathbf{a}} = [0.20, 4.8 \times 10^{-3}, 2.3 \times 10^{-3}, 3 \times 10^{-4}, 2 \times 10^{-4}]^T \\
& \text{Résultats d'estimation avec ECT : } \hat{\boldsymbol{\phi}} = [-2.49, -1.32, -0.33, -1.96, -2.50]^T \\
& \text{REQM} = 0.0365 \text{ A}
\end{aligned}$$

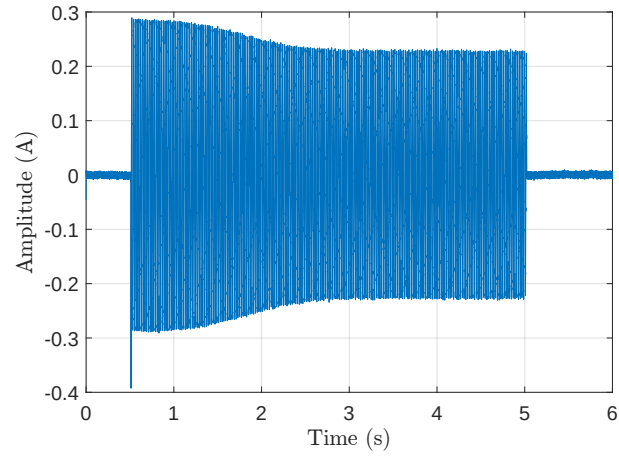
Les résultats d'estimation donnés ci-dessus indiquent peu d'information sur la qualité de l'estimation surtout que nous sommes dans un cas réel où l'on n'a pas la maîtrise la génération des signaux. La figure 3.21 nous permet d'avoir un résultat visuel afin de pouvoir évaluer la qualité d'estimation. Cette figure contient deux zooms, un sur le début du transitoire (mise en marche) et le deuxième sur une partie du régime établi. Sur la figure 3.21a on peut déjà voir que l'amplitude du signal estimé ne suit pas l'amplitude du signal réel. Les zooms des figures 3.21b et 3.21c nous permettent de voir l'écart entre les deux signaux. Notons principalement deux choses :

- le signal reconstruit a une forme très proche du signal original ;
- des décalages temporels existent entre les deux signaux.

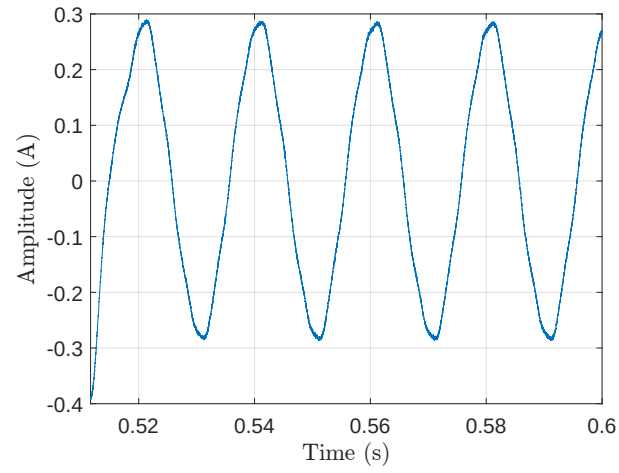
Ces constatations indiquent qu'il existe un problème dans l'estimation des phases temporelles. Noter aussi que le décalage sur la figure 3.21b est plus grand que celui constaté sur la figure 3.21c. En conclusion, ces observations nous amènent à mettre en cause la validité de la simplification **H4** en situation réelle. Il s'ensuit, de ce fait, la nécessité d'estimer la fréquence fondamentale pour prendre en compte ses variations temporelles.

Dans la sous-section suivante nous allons voir l'impact, sur les résultats d'estimation, de l'amélioration du ECT en considérant l'estimation de f_0 .

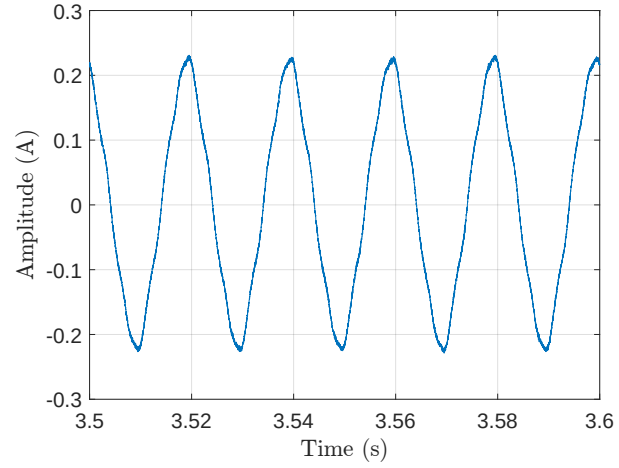
3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME



(a) Courant.

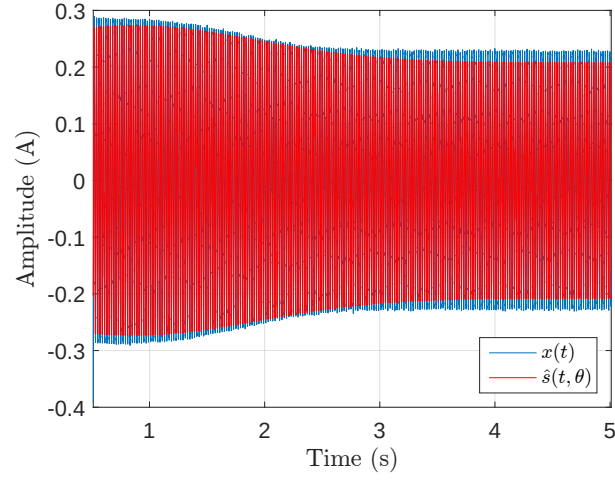


(b) Zoom de 3.20a dans l'intervalle [0,51-0,6] s.



(c) Zoom de 3.20a dans l'intervalle [3,5-3,6] s.

FIGURE 3.20 – Courant de mise en marche d'un ventilateur.



(a) Courant.

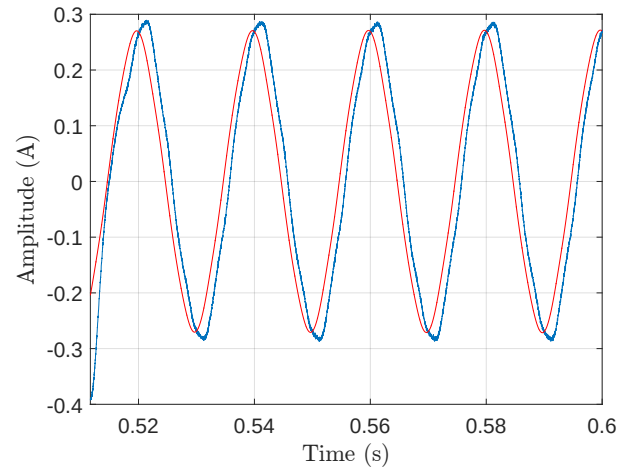
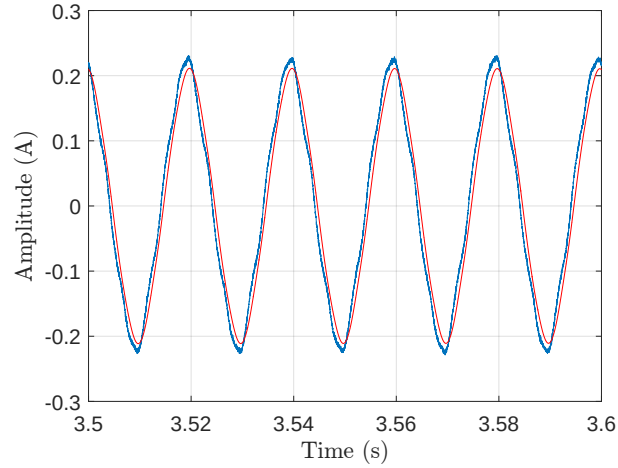
(b) Zoom de 3.21a dans l'intervalle $[0,51-0,6]$ s.(c) Zoom de 3.21a dans l'intervalle $[3,5-3,6]$ s.

FIGURE 3.21 – Courant de mise en marche d'un ventilateur (en bleu) avec le signal reconstruit (en rouge) avec les paramètres estimés en utilisant l'algorithme ECT.

3.3.3 Algorithme d'estimation des paramètres amélioré (ECTA)

En raison des variations de la tension du réseau dues à l'équilibre offre-demande à maintenir⁴, la fréquence fondamentale peut légèrement dévier de sa valeur nominale conduisant à $f_i = (2i - 1)(f_0 + \delta f_0)$, $i = 1, \dots, d$.

La norme européenne « EN 50160 » (CENELEC, 2010) fixe les plages de variation acceptables pour δf_0 . Ces plages, pour le réseau synchrone de l'Europe continentale (synchronous grid of Continental Europe), qui est le plus grand réseau synchrone (une seule fréquence de réseau fixée à 50 Hz) dans le monde reliant la plupart des pays européens ainsi que certains pays d'Afrique du nord, sont ± 1 % de la fréquence 50 Hz ($\delta f_0 = [-0.5, +0.5]$ Hz) pendant 99,5 % d'une année et -6 % / $+4$ % de la fréquence 50 Hz ($\delta f_0 = [-3, +2]$ Hz) pendant 100 % du temps. La dernière variation acceptable étant fixée large pour prendre en compte des variations occasionnelles élevées.

Nous proposons dans cette sous-section une version améliorée de l'algorithme ECT précédemment décrit (algorithme 1) par ajout d'une étape d'estimation de la fréquence fondamentale.

Algorithme ECT amélioré (ECTA) avec l'estimation de f_0

Nous avons supposé pour notre modèle des fréquences connues et fixes $f_i = (2i - 1)f_0$, $i = 1, \dots, d$ où d est le nombre de sinusoïdes et $f_0 = 50$ Hz. Nous proposons maintenant de considérer les fréquences comme des harmoniques impaires d'une fréquence fondamentale inconnue pouvant dévier du 50 Hz telles que $f_i = (2i - 1)(f_0 + \delta f_0)$, $i = 1, \dots, d$ ou $f_i = (2i - 1)f_0(t)$ en supposant une variation continue au cours du temps de la fréquence fondamentale.

Nous avons déjà pu vérifier pour cette modélisation, en utilisant l'algorithme « Fast Multi-frequency Estimator (FaME) » (Ye and Aboutanios, 2015), que le contenu fréquentiel est toujours constitué d'harmoniques. De ce fait, ce qui va conditionner la variation du contenu fréquentiel est la variation de la fréquence fondamentale.

4. <http://www.rte-france.com/fr/article/mettre-en-oeuvre-tous-les-leviers-pour-assurer-l-equilibre>

L'algorithme FaME permet d'obtenir de bonnes estimées du contenu fréquentiel des signaux. Néanmoins, il n'est pas adapté à ce que l'on cherche à faire. Il arrivait souvent, durant les simulations, que cet algorithme confonde des lobes secondaires d'une fréquence harmonique avec une fréquence originale du signal. Cela est une conséquence de la construction de l'algorithme nécessitant de spécifier le nombre d'harmoniques constituant le signal sans qu'aucun *a priori* n'est supposé sur ces dernières, ce qui explique son inadéquation à notre application.

À cause de la particularité de notre application et du fait qu'on travaille sur le transitoire, le problème d'estimation de fréquence se complique en ajoutant des contraintes particulières, en l'occurrence :

C1 une nécessité d'avoir des valeurs précises pour la fréquence estimée ;

C2 une nécessité d'un suivi continu et rapide de la variation de cette dernière.

Si ces deux contraintes ne sont pas respectées, nous ne pouvons pas espérer bien améliorer l'estimation des paramètres pour les signaux réels. Dans ce cadre contraint, nous avons proposé d'utiliser une méthode fondée sur la détection des passages par zéros et ceci en utilisant plutôt le signal tension que le signal courant. Les méthodes basées sur la détection de passages par zéro sont simples et permettent de suivre rapidement la variation de la fréquence. L'estimée étant obtenue à raison d'une estimée par période secteur. Cela nous permet de satisfaire à la contrainte **C2**. Le grand désavantage de ces méthodes est leur sensibilité aux harmoniques et au bruit de mesure. Sinon, à très faible bruit et n'ayant qu'un signal mono-fréquence ils offrent une très bonne précision d'estimation ce qui permettra de satisfaire à la contrainte **C1**.

Pour profiter de la précision et du suivi continu que proposent ces méthodes, nous avons proposé d'utiliser la tension filtrée (en basse fréquence) pour estimer la fréquence fondamentale afin de minimiser le plus possible l'effet des harmoniques. Afin de minimiser l'effet du bruit, nous avons choisi la version modifiée de l'algorithme simple de détection de passages par zéros « modified zero-crossing » (MZC) proposé dans (Begovic et al., 1993) (le principe de ce dernier étant expliqué en annexe A). Ainsi, on aurait respecté **C1** et **C2**.

Nous proposons le test de l'algorithme MZC sur un signal simulé de fréquence fonda-

mentale $f_0(t)$ variant suivant la loi :

$$f_0(t) = f_0 + \begin{cases} -0.04 \cos(2\pi t + \pi/3) - 0.03 \cos(2\pi 3t) & , 0 < t \leq 0.5s \\ -0.04 \cos(2\pi t + \pi/3) - 0.03 \cos(2\pi 3t) + 1/2\pi & , 0.5 < t \leq 1s \end{cases} \quad (3.26)$$

avec $f_0 = 50$ Hz. Le signal contenant aussi les harmoniques d'ordre 3 et 5 avec des amplitudes relatives (à l'amplitude de la fondamentale) égales, respectivement, à 30% et 20% ainsi qu'un bruit gaussian additif avec un RSB = 40 dB.

Le filtre passe bas choisi est un filtre à Réponse Impulsionnelle Infinie (RII) Chebyshev type I d'ordre 10 avec une fréquence de coupure à 60 Hz et une ondulation de 10^{-5} en passe bande.

La figure 3.22 donne le résultat d'estimation obtenu. Ce dernier montre que la méthode donne une bonne estimation de la fréquence fondamentale même avec la transition abrupte autour de 0,5 s. Cela montre la capacité de suivi rapide et précis des variations de la fréquence fondamentale par la méthode. Néanmoins, et à cause du filtrage passe bas IIR,

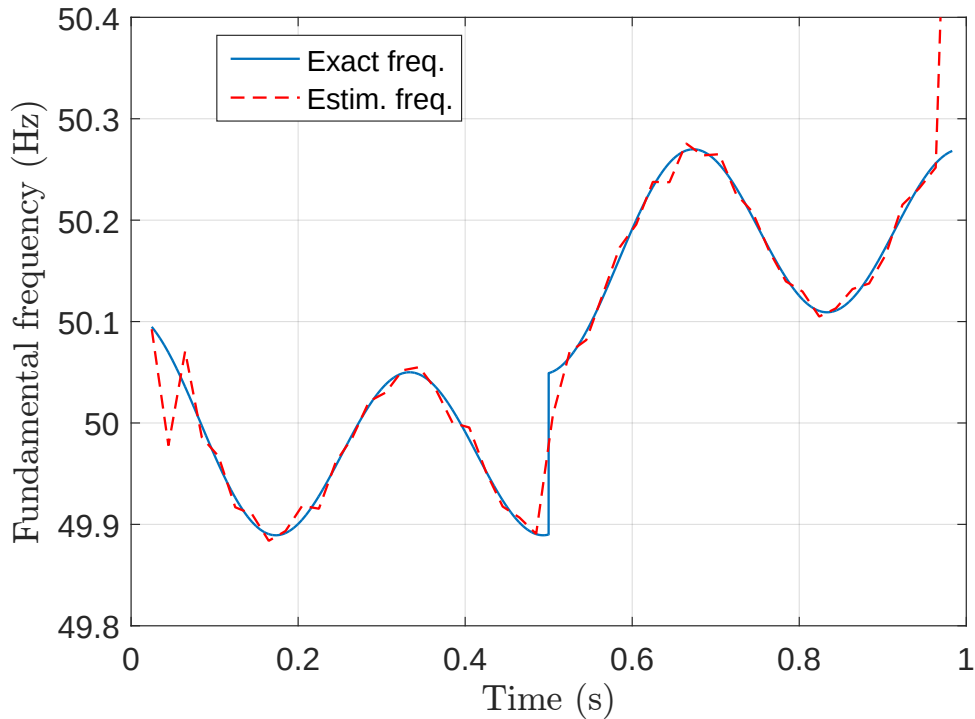


FIGURE 3.22 – Résultat d'estimation d'une fréquence simulée en utilisant l'algorithme modified zéro-crossing (MZC). Bleu : fréquence exacte. Rouge : fréquence estimée.

on remarque des artefacts qui apparaissent en début et à la fin du signal. Pour pallier ces artefacts, nous estimons la fréquence fondamentale avant de détecter le transitoire de

mise en marche. Ainsi, l'effet des artefacts va se limiter aux zones de consommation nulle qui seront éliminées par le détecteur.

En supposant que la fréquence fondamentale varie continuellement, le modèle du transitoire de mise en marche (équation 3.1) devient :

$$s(t) = \left(e^{\mathbf{P}^T \mathbf{t}} + 1 \right) \sum_{i=1}^d A_i \cos(2\pi(2i-1)f_0(t)t + \phi_i) \quad (3.27)$$

Les résultats de l'étude de l'influence de la connaissance de f_0 sur les BCR nous permettent de conclure qu'utiliser une estimée pour f_0 , en l'occurrence $\hat{f}_0(t)$, dans le modèle ne va pas affecter la qualité d'estimation de $\mathbf{a}$ et $\boldsymbol{\phi}$. Ce qui est rassurant.

L'algorithme ECT amélioré (ECTA) (Enhanced Transient Current Estimation–ETCE) est détaillée ci-dessous :

Algorithme 2 : Estimation du Courant Transitoire Amélioré (ECTA).

Estimation de $f_0(t)$:

- calculer $\hat{f}_0(t)$ en utilisant l'algorithme MZC.

Application du ECT :

- utiliser l'algorithme ECT en remplaçant f_0 (phase initialisation de l'algorithme 1) par $\hat{f}_0(t)$ pour calculer $\hat{\boldsymbol{\theta}}$.
-

Résultat de l'estimation améliorée avec estimation de f_0

Pour tester l'algorithme ECTA, nous considérons le même signal réel de ventilateur que nous avons utilisé pour l'estimation avec ECT.

$$\begin{aligned} \hat{\mathbf{p}} &= [-1.15, -0.16, 0.16, -0.32]^T \\ \hat{\mathbf{a}} &= [0.21, 9.1 \times 10^{-3}, 7 \times 10^{-3}, 5.23 \times 10^{-3}, 1.16 \times 10^{-3}]^T \\ \hat{\boldsymbol{\phi}} &= [-2.81, -2.26, -1.91, 2.48, -2.16]^T \\ \text{REQM} &= 9 \times 10^{-3} \text{ A} \end{aligned}$$

Nous remarquons déjà l'amélioration de la REQM avec une valeur environ quatre fois moindre que pour l'estimation avec ECT (où REQM = 0,0365 A). La figure 3.23 donne la fréquence fondamentale estimée $\hat{f}_0(t)$ en fonction du temps. Nous remarquons surtout

3.3. PERFORMANCES EXPÉRIMENTALES DE L'ALGORITHME

que sa valeur moyenne se retrouve autour de 50.03 Hz, donc un peu loin du 50 Hz.

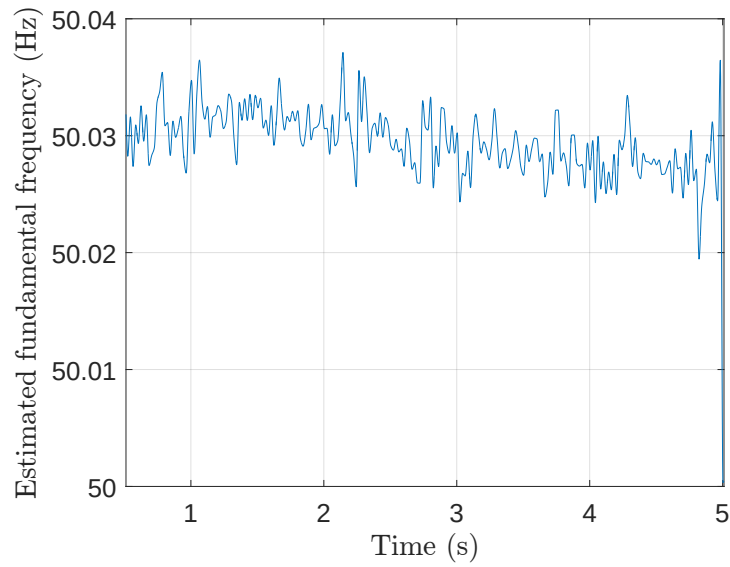
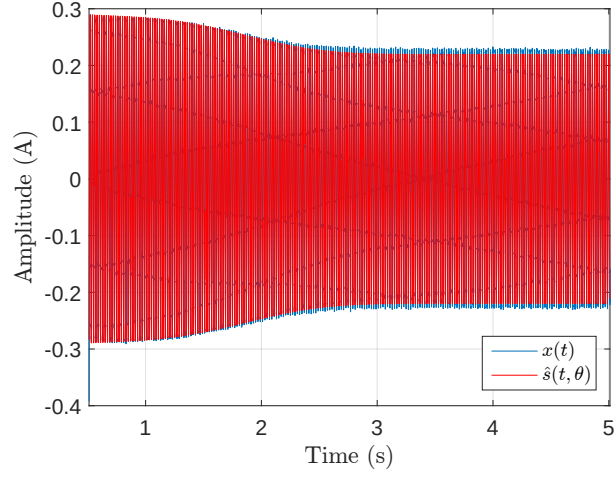


FIGURE 3.23 – Fréquence fondamentale estimée $\hat{f}_0(t)$ au cours de la mesure du ventilateur.

Sur la figure 3.24a, nous constatons que l'amplitude du signal reconstruit colle mieux au signal original que celle de la figure 3.21a. De même pour les figures 3.24b et 3.24c, les deux signaux, reconstruit et original, concordent quasi parfaitement.

En conclusion, nous avons bien vérifié les performances meilleures du ECTA comparées au ECT pour l'estimation des signaux réels. De plus, nous avons bien mis en évidence la détérioration de l'estimation suite aux variations de la fréquence fondamentale.



(a) Courant.

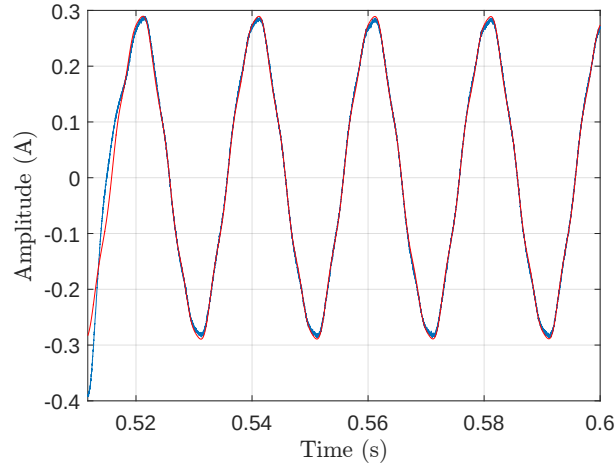
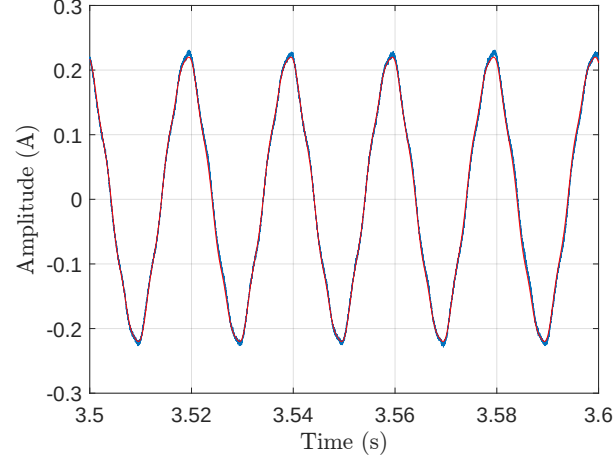
(b) Zoom de 3.24a dans l'intervalle $[0,51-0,6]$ s.(c) Zoom de 3.24a dans l'intervalle $[3,5-3,6]$ s.

FIGURE 3.24 – Courant de mise en marche d'un ventilateur (en bleu) avec le signal reconstruit (en rouge) avec les paramètres estimés en utilisant l'algorithme ECTA.

3.4 Détection d'évènements

Dans cette section, nous présentons le détecteur d'évènements ainsi que son évaluation en comparaison avec un algorithme récemment proposé dans le domaine.

L'objectif est d'améliorer la précision de la détection (localisation)⁵ des instants définissant les évènements en évitant l'utilisation de tout type de moyenne (le calcul de puissance par exemple), de filtrage, ou de transformation (FFT par exemple), etc. qui peuvent causer une perte d'information temporelle précieuse.

Rappelons que le détecteur proposé sert dans la phase d'estimation pour segmenter le signal de courant en différentes zones : une zone de transition à la mise en marche, une zone de régime établi et une zone de transition à la mise à l'arrêt.

La nouveauté de l'algorithme proposé par rapport à l'état de l'art dans le domaine est, principalement, sa *précision* de détection.

Notons que pour éviter toute confusion lors de l'utilisation du terme *précision* entre la précision de détection et le critère d'évaluation, aussi appelé précision, nous optons pour l'utilisation du terme « justesse » pour qualifier la précision de détection et nous gardons l'utilisation du terme précision pour désigner le critère d'évaluation.

Notre algorithme est le premier, à notre connaissance, capable de fournir une telle précision pour une détection d'évènements dans le cadre du NILM.

3.4.1 Algorithme de détection (HAND)

On présente dans cette section l'algorithme proposé pour la détection d'évènements, appelé ci-après High Accuracy NILM Detector (HAND). Un événement, ou *section active*, correspond à la partie du signal qui dévie du régime établi précédent et dure jusqu'à ce qu'il atteigne le régime établi suivant (Wild et al., 2015).

L'algorithme proposé est simple et rapide. Il suit la variation de l'écart type (ET)

5. Nous confondons par la suite ces deux termes, ayant en général des significations différentes dans d'autres contextes. Supposer une bonne détection des instants définissant les évènements revient à dire que nous arrivons à bien localiser ces derniers sur l'axe temporel.

(Standard deviation–STD) $\sigma_d(t)$ de l'enveloppe détectée $e_d(t)$ du signal de courant en utilisant une fenêtre glissante. Le principe du détecteur repose sur le fait que les régimes établis sont caractérisés par un faible $\sigma_d(t)$ (en théorie nul), alors qu'un évènement est caractérisé par une valeur élevée de $\sigma_d(t)$. La séparation entre le régime établi et un évènement s'effectue à l'aide d'un seuil.

L'algorithme donne en sortie les instants de début et de fin des évènements détectés (figure 3.3). L'algorithme étant composé de différentes parties, nous donnons ci-après des descriptions détaillées avant de donner l'algorithme sous format compacte (algorithme 4). Les différentes parties sont :

1. détection de l'enveloppe du signal courant $e_d(t)$;
2. choix de la taille de la fenêtre glissante L (pour nos simulations nous avons choisi $L \equiv 8$ périodes secteur [en échantillons]) ;
3. initialiser $\sigma_d(t_k)$, pour $k = 1, \dots, L$, avec l'écart-type de $e_d(t)$, $t = t_1, \dots, t_L$ et la moyenne $\mu_d(t_k)$, pour $k = 1, \dots, L$, avec la moyenne de $e_d(t)$, $t = t_1, \dots, t_L$;
4. calculer *itérativement* la moyenne $\mu_d(t_k)$ ainsi que l'estimée, non biaisée, de l'écart-type $\sigma_d(t_k)$ de $e_d(t)$, $k = L + 1, \dots, N$, en utilisant :

$$\begin{aligned}\mu_d(t_k) &= \mu_d(t_{k-1}) + \frac{1}{L} [e_d(t_k) - e_d(t_{k-L})] \\ \sigma_d^2(t_k) &= \sigma_d^2(t_{k-1}) + \frac{1}{L-1} [e_d^2(t_k) - e_d^2(t_{k-L})] + \frac{L}{L-1} [\mu_d^2(t_{k-1}) - \mu_d^2(t_k)]\end{aligned}\tag{3.28}$$

5. choix du seuil de détection γ (pour nos simulations on a pris $\gamma = 0.1$) et on trouve les instants de début et de fin pour chaque évènement tel que :
 - le temps de début t_d est défini comme étant le *premier* point d'un évènement où $\sigma_d(t) > \gamma$ et $\frac{d\sigma_d(t)}{dt} \geq 0$,
 - le temps de fin t_f est défini comme étant le *dernier* point d'un évènement où $\sigma_d(t) > \gamma$ et $\frac{d\sigma_d(t)}{dt} < 0$.
6. post-traitement :
 - approbation de l'évènement : juste si $\sum_t \sigma_d(t) > \delta$, $t \in [t_d, t_f]$,
 - correction du délai causé par la fenêtre glissante : $t_f = t_f - L$.

Pour détecter l'enveloppe du courant (étape 1), on utilise un algorithme ad-hoc que nous proposons, appelé ci-après Détecteur d'Enveloppe Amélioré (DEA) (Enhanced Envelope Detector–EED). Nous utilisons la valeur maximale sur chaque période secteur (20 ms

3.4. DÉTECTION D'ÉVÈNEMENTS

pour une fréquence réseau de 50 Hz) et nous interpolons (interpolation cubique) entre tous les pics détectés.

Cette interpolation cause le lissage des transitions abruptes du courant au début des évènements ($e_{d1}(t)$ dans la figure 3.25). Pour pallier cela, nous proposons de tester la médiane de la valeur absolue du signal courant $x_{p_k, p_{k+1}}(t)$ situé entre chaque deux pics détectés consécutifs p_k et p_{k+1} avec un seuil α (fixé empiriquement). Si la $\text{median}(|x_{p_k, p_{k+1}}(t)|) < \alpha$ alors pour cette partie $e_d(t) = x_{p_k, p_{k+1}}(t)$, sinon l'enveloppe détectée précédemment est gardée.

Ce test additionnel permet à l'enveloppe de mieux suivre les variations du signal courant, surtout quand la période secteur du régime établi précédent l'évènement a une amplitude faible ou contient une portion du début de l'évènement ($e_{d2}(t)$ dans la figure 3.25).

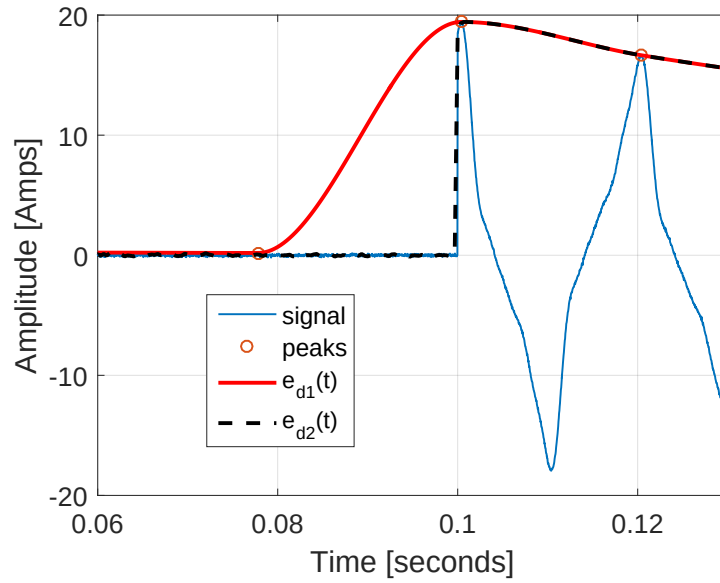


FIGURE 3.25 – Amélioration de la détection de l'enveloppe du signal courant. Exemple sur un évènement de mise en marche : $e_{d1}(t)$ est détecté sans test de médiane et $e_{d2}(t)$ avec.

L'algorithme DEA est résumé ci-dessous :

Algorithme 3 : Détecteur d'Enveloppe Amélioré (DEA).

Initialisation :

- signal courant : $x(t)$, $\alpha = 0.5$ (empirique).

Détection d'enveloppe :

- trouver les pics p_k correspondant aux max sur chaque période secteur de $x(t)$;
- interpoler (cubique) entre les p_k pour obtenir $e_d(t)$.

Test de médiane :

- $x_{p_k, p_{k+1}}(t)$: portion de $x(t)$ se trouvant entre les pics consécutifs p_k et p_{k+1} ;
- retrouver $e_d(t)$ entre p_k et p_{k+1} avec la condition :

$$e_{d(p_k, p_{k+1})}(t) = \begin{cases} x_{p_k, p_{k+1}}(t) & , \text{si } \text{mediane}(|x_{p_k, p_{k+1}}(t)|) < \alpha \\ e_{d(p_k, p_{k+1})}(t) & , \text{sinon (garder l'ancien signal)} \end{cases}$$

Après la détection de $e_d(t)$, nous calculons son écart-type $\sigma_d(t)$, temporel, représentant le paramètre caractéristique du détecteur d'évènements (étapes 2 à 4) (figure 3.26). Au même temps, nous nous fixons un seuil γ (étape 5) pour définir les instants de début t_d et de fin t_f pour chaque événement. Notons que ces dernières étapes sont faites de manière *itérative* afin d'alléger les calculs et améliorer la rapidité de l'algorithme.

Vu que notre paramètre caractéristique est l'écart-type de l'enveloppe du signal de courant, fixer γ est guidé par la variance du bruit. Par exemple, si l'écart-type du bruit = 80 mAmps, fixer $\gamma = 0,08$ correspond à considérer toutes les sections du signal ayant une variance au-dessus de la variance du bruit comme des événements. En conséquence, cela ajoute un avantage non négligeable à notre algorithme.

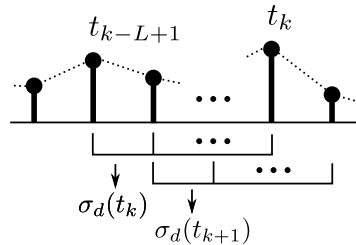


FIGURE 3.26 – Illustration du « principe » de calcul de $\sigma_d(t)$ en utilisant une fenêtre glissante de taille L . Notons que concrètement cela est fait d'une manière équivalente mais itérative.

Le post-traitement (étape 6) corrige principalement pour deux choses : (1) les sections

3.4. DÉTECTION D'ÉVÈNEMENTS

de $\sigma_d(t)$ qui dépassent à peine le seuil γ (approbation de l'événement) et (2) le délai temporel (correction du délai causé par la fenêtre glissante) qui apparaît sur l'instant de fin t_f en raison de l'utilisation, par la fenêtre glissante, des $L - 1$ points antérieurs pour calculer un point présent (figure 3.26).

Notons ici que le seuil δ utilisé pour l'approbation de l'événement est fixé de manière empirique. De plus, aucune correction n'est nécessaire pour l'instant de début t_d .

L'algorithme HAND est résumé en version compacte dans la page suivante. Suivra après les résultats expérimentaux.

Algorithme 4 : High Accuracy NILM Detector (HAND).

Initialisation :

1. signal courant : $x(t)$ [N échantillons] ;
2. fixer $L \equiv 8$ périodes secteur [en échantillons], $\gamma = 0.1$, $\delta = 3 \times 10^3$ (empirique) ;
3. vecteurs rassemblant les instants de début et de fin des évènements avant post-traitement : $\mathbf{t}_d^{(i)} = []$, $\mathbf{t}_f^{(i)} = []$;
4. vecteurs rassemblant les instants de début et de fin des évènements après post-traitement : $\mathbf{t}_d^{(f)} = []$, $\mathbf{t}_f^{(f)} = []$;

Détection d'enveloppe :

1. détecter l'enveloppe $e_d(t)$ de $x(t)$ en utilisant l'algorithme DEA (algorithme 3).

Détection de début et de fin d'évènements :

1. initialiser $\mu_d(t_k) = \text{MEAN}(e_d(t))$ pour $k = 1, \dots, L$ et $t = t_1 : t_L$;
2. initialiser $\sigma_d(t_k) = \text{STD}(e_d(t))$ pour $k = 1, \dots, L$ et $t = t_1 : t_L$;
3. $k = L + 1$;
- (*) tant que $k \leq N$ faire
 - (a)

$$\begin{aligned} \mu_d(t_k) &= \mu_d(t_{k-1}) + \frac{1}{L} [e_d(t_k) - e_d(t_{k-L})] \\ \sigma_d^2(t_k) &= \sigma_d^2(t_{k-1}) + \frac{1}{L-1} [e_d^2(t_k) - e_d^2(t_{k-L})] + \frac{L}{L-1} [\mu_d^2(t_{k-1}) - \mu_d^2(t_k)] \end{aligned}$$

- (b) si $\sigma_d(t_k) \geq \gamma$ et $\sigma_d(t_{k-1}) < \gamma$, alors $\mathbf{t}_d^{(i)} = [\mathbf{t}_d^{(i)}, t_k]$, $k = k + 1$, retour à (*),
- (c) si $\sigma_d(t_k) \leq \gamma$ et $\sigma_d(t_{k-1}) > \gamma$, alors $\mathbf{t}_f^{(i)} = [\mathbf{t}_f^{(i)}, t_k]$, $k = k + 1$, retour à (*),
- (d) sinon $k = k + 1$, retour à (*).

Post-traitement :

1. $\mathbf{t}_d^{(i)} = [t_{d1}, \dots, t_{dM}]$ et $\mathbf{t}_f^{(i)} = [t_{f1}, \dots, t_{fM}]$, le nombre d'évènement détectés $M = \text{taille}(\mathbf{t}_d^{(i)} \text{ ou } \mathbf{t}_f^{(i)})$;
 2. pour $l = 1 : M$ faire
 - (a) $\text{sum} = \sum_t \sigma_d(t)$, $t = t_{dl} : t_{fl}$,
 - (b) si $\text{sum} > \delta$ alors $\mathbf{t}_d^{(f)} = [\mathbf{t}_d^{(i)}, t_{dl}]$ et $\mathbf{t}_f^{(f)} = [\mathbf{t}_f^{(i)}, t_{fl}]$,
 3. $\mathbf{t}_f^{(f)} = \mathbf{t}_f^{(f)} - L$.
-

3.4.2 Résultats expérimentaux

Pour tester le HAND, nous proposons d'utiliser des données simulées ainsi que des réelles et de comparer les résultats avec ceux obtenus en utilisant un algorithme récemment proposé basé sur l'analyse discriminante à noyau de Fisher (Kernel Fisher Discriminant Analysis–KFDA) (Wild et al., 2015) (par souci de simplification, on va référer par la suite à cet algorithme par KFDA).

Cet algorithme utilise les harmoniques de courant I_k comme paramètres caractéristiques estimés à l'aide d'une technique de démodulation IQ (In phase, In quadrature). Il utilise la fonction discriminante linéaire de Fisher (Mika et al., 2001) comme classificateur pour distinguer entre les classes « événement » et « régime établi ». Dans le principe, il calcule un test statistique et le compare à un seuil adaptatif.

Pour obtenir des résultats fiables, les données utilisées pour le test doivent être fournies avec *une vérité terrain précise* des instants définissant les événements. Il est important de préciser que la mise en marche d'un appareil est considérée comme un événement qui a un début et une fin et la mise à l'arrêt d'un appareil est elle aussi considérée comme un autre événement ayant ses instants de début et de fin respectifs. Ceci implique qu'ils sont les instants de début d'événements qu'il est possible de connaître avec précision car ils correspondent au déclenchement de l'action (par exemple, allumer ou éteindre un appareil). L'instant correspondant à la fin de l'événement de mise en marche d'un appareil ne puisse être connu avec justesse, nous allons comparer la vérité terrain des instants de début des événements avec ceux qui sont détectés.

Tests sur données simulées

Pour les simulations, nous avons utilisé, pour le signal de courant, notre modèle (équation 3.1). Le signal simulé $x(t)$ est la somme de K différentes versions $s_k(t, \tau_k, D_k)$ retardées de $s(t)$ et d'un bruit :

$$x(t) = \sum_{k=1}^K s_k(t, \tau_k, D_k) + w(t), \quad t = 0, \dots, D$$

Le bruit est blanc gaussien avec moyenne nulle et variance σ_w^2 . τ_k et D_k sont, respectivement, des délais temporels et durées. Les délais permettent de positionner le début des différentes versions $s_k(t, \tau_k, D_k)$ à différents instants sur la durée totale du signal.

Les versions de $s(t)$ sont générées aléatoirement et les paramètres sélectionnés pour les simulations sont comme suit ($\mathcal{U}(a, b)$ étant la distribution uniforme sur l'intervalle (a, b)) :

$$\begin{aligned} A_0 &= 1, n = 1, d = 5 \\ f &= [50, 150, 250, 350, 450]^T \\ \tau_k &= \mathcal{U}(0, D) \\ D_k &= \mathcal{U}(1, 10) \\ p_1 &= \mathcal{U}(-10, 0) \\ A_i &= \mathcal{U}(0, 5) \\ \phi_i &= \mathcal{U}(0, \pi) \end{aligned}$$

La fréquence d'échantillonnage $F_s = 10$ kHz et $K = 4000$. Après avoir enlevé les versions dont les instants finaux débordaient la durée totale du signal $D = 2000$ minutes, nous nous sommes retrouvé avec moins de versions (autours de 3800 versions $\equiv$ 7600 évènements pour chaque simulation).

Comme critères d'évaluation, nous proposons d'utiliser les mesures suivantes : True Positives (TP), False Negatives (FN), False Positives (FP), précision et rappel comme suggéré par différents travaux dans le domaine (Klein et al., 2013; Makonin and Popowich, 2014). Ces mesures sont définies comme :

- TP : nombre d'évènements vrais correctement détectés ;
- FN : nombre d'évènements vrais non détectés (ratés), $FN = N - TP$ (N : nombre total d'évènements vrais) ;
- FP : nombre de faux évènements détectés par erreur ;
- précision = $TP / (TP + FP)$;
- rappel = $TP / (TP + FN)$.

Parallèlement à ces critères, nous définissons une *tolérance au retard* Δt telle que pour qu'un événement détecté et ayant un instant de début t_d soit correctement détecté, la condition $|t_d - t_v| \leq \Delta t$ doit être vérifiée, où t_v est la vérité terrain. Un faux positif (FP) est alors défini comme étant un événement détecté ne vérifiant pas cette condition. De

3.4. DÉTECTION D'ÉVÈNEMENTS

cette façon, Δt est une quantité qui définit la *justesse requise* d'un détecteur afin que ses résultats soient considérés bons.

Ci-après, nous comparons les deux algorithmes (HAND et KFDA) d'un point de vue justesse (en variant Δt) ainsi que par rapport à la sensibilité au bruit (en variant le Rapport Signal-à-Bruit (RSB)).

Comparaison de la justesse Pour tester les performances des algorithmes par rapport à une justesse requise, nous fixons le RSB à 50 dB et nous considérons deux valeurs pour Δt : 200 ms et 500 ms. Le tableau 3.1 donnent les résultats.

Ces derniers indiquent que HAND satisfait aux justesses requises et donne remarquablement une précision parfaite de 100%. D'autre part, KFDA donne un bon résultat (précision/rappel = 98,19/86,20%) pour $\Delta t = 500$ ms mais affiche une performance médiocre pour la justesse requise de $\Delta t = 200$ ms.

L'approche du KFDA est attractive, néanmoins, ce qui lui fait perdre en justesse est principalement sa dépendance de filtres passe-bas utilisés dans la démodulation IQ. Ces filtres ajoutent des artefacts qui apparaissent autour des changements abruptes. Pour enlever ces dernier KFDA utilise un lissage qui détériore ses performances de détection.

TABLE 3.1 – TP, FN, FP, précision et rappel du HAND *vs.* KFDA, RSB = 50 dB, $\Delta t = 200$ ms et $\Delta t = 500$ ms

		TP	FN	FP	précision (%)	rappel (%)
$\Delta t = 200$ ms	HAND	7417	253	0	100.00	96.70
	KFDA	1517	6153	5269	22.35	19.78
$\Delta t = 500$ ms	HAND	7399	219	0	100.00	97.13
	KFDA	6567	1051	121	98.19	86.20

Sensibilité au bruit En se basant sur les résultats du paragraphe précédent, nous fixons pour cette expérience $\Delta t = 500$ ms afin de ne pas pénaliser l'algorithme KFDA et nous faisons varier le RSB (tableau 3.2).

Les résultats montrent que KFDA est particulièrement robuste contre le bruit. Sa précision et son rappel restent autour de la même valeur pour une baisse de RSB de 20 dB (de 50 à 30 dB). Les valeurs moyennes sont, respectivement, de 98,23% et 86,83%

TABLE 3.2 – TP, FN, FP, précision et rappel of HAND *vs.* KFDA, $\Delta t = 500$ ms, différents RSBs

RSB	Algo.	TP	FN	FP	précision (%)	rappel (%)
50 dB	HAND	7399	219	0	100.00	97.13
	KFDA	6567	1051	121	98.19	86.20
40 dB	HAND	7432	202	15	99.80	97.35
	KFDA	6695	939	123	98.20	87.70
30 dB	HAND	7490	186	399	94.94	97.58
	KFDA	6801	875	117	98.31	86.60

pour la précision et pour le rappel. La précision du HAND, en revanche, est sensible au bruit. Elle baisse en allant de 100% à près de 95% pour une diminution de RSB de 20 dB. En effet, un tel résultat est attendu et est lié à l'utilisation de l'écart-type de l'enveloppe du signal courant comme paramètre caractéristique. Ainsi, plus le RSB est faible, plus de FP risquent d'être détectés.

Par ailleurs, le rappel du HAND reste élevé et n'est pas affecté (reste à environ 97.35%). Ceci est clairement un avantage pour le HAND par rapport à KFDA. Un bon rappel signifie que le détecteur ne rate pas souvent et cela est très souhaitable dans un système NILM. Si possible, on ne voudrait rater aucun événement vrai. Avoir des FP n'est aussi pas souhaitable, mais est moins désavantageux que d'avoir raté. Cela est justifié par le fait que les FP peuvent être traités lors d'une classification à une classe « inconnue », par exemple, au lieu de les affecter à une classe « appareil ».

Finalement, notant que HAND est au moins trois fois plus rapide que KFDA à cause de sa simplicité et ses calculs rapides.

Tests sur données réelles

Dans cette sous-section, nous montrons un exemple de détection sur des données réelles obtenues à l'aide de notre système de mesure décrit dans la section 4.2. Ce système permet la création de scénarios de consommation avec l'avantage de fournir un contrôle précis sur les instants de mise en marche et de mise à l'arrêt des appareils.

Le signal mesuré est échantillonné à 100 kHz et contient cinq événements (4 mises en marche et 1 mise à l'arrêt : (1) mise en marche de la partie électronique d'un aspirateur, (2) mise en marche du moteur du même aspirateur, (3) mise en marche d'une perceuse,

3.4. DÉTECTION D'ÉVÈNEMENTS

(4) mise en marche d'une lampe halogène, (5) mise à l'arrêt de la perceuse). Les résultats de détection sont présentés dans les figures 3.27 et 3.28 et les erreurs de détection des instants de début d'évènements sont données dans le tableau 3.3.

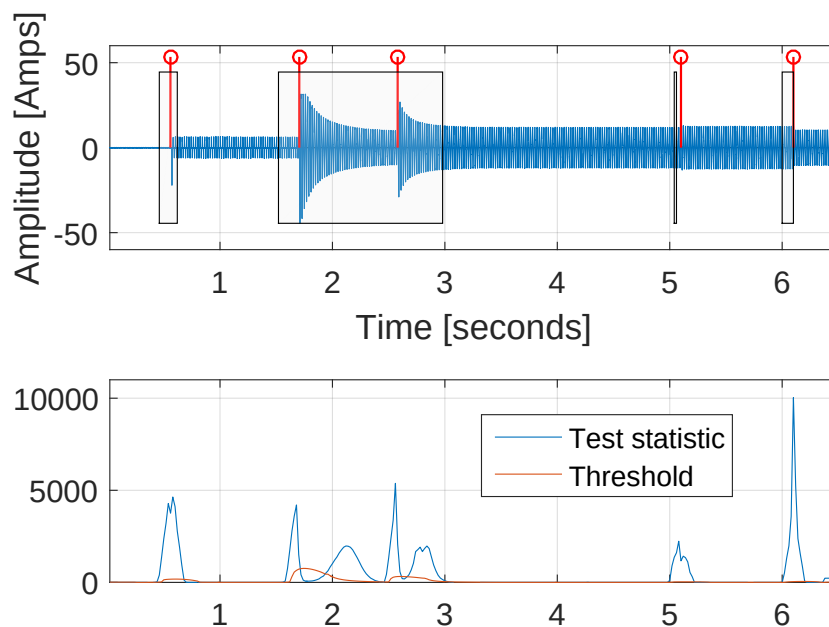


FIGURE 3.27 – Résultats du KFDA sur données réelles. Haut : pics rouges indiquent la vérité terrain (instants de début d'évènements) et les évènements détectés sont cadrés dans des boîtes.

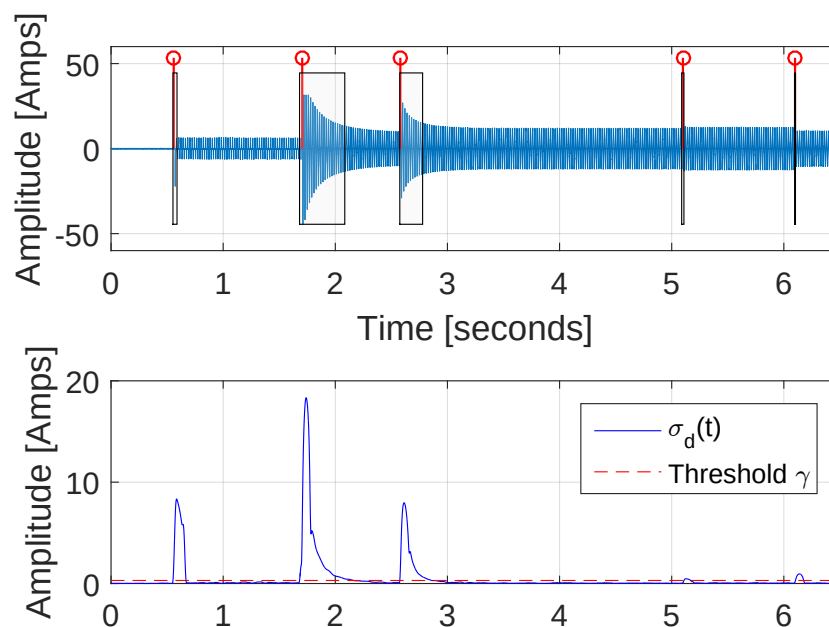


FIGURE 3.28 – Résultats du HAND sur données réelles. Haut : pics rouges indiquent la vérité terrain (instants de début d'évènements) et les évènements détectés sont cadrés dans des boîtes.

On note que la valeur moyenne de l'erreur de détection sur les cinq évènements est de 9,94 ms (presque une demi-période secteur) pour HAND et 87,25 ms pour KFDA. De plus, KFDA a raté le 3^e évènement qui a été fusionné avec le 2^e. Ceci compare sur un exemple simple la précision du HAND par rapport à celle du KFDA.

TABLE 3.3 – Erreurs dans la détection des instants de début pour les cinq évènements montrés dans les figures 3.27 et 3.28

	HAND	KFDA
Évènement 1	6.7 ms	99 ms
Évènement 2	24 ms	180 ms
Évènement 3	4.8 ms	fusionné avec l'évènement 2
Évènement 4	11.1 ms	60 ms
Évènement 5	3.1 ms	10 ms
Moyenne	9.94 ms	87.25 ms

3.4.3 Conclusion

Un nouveau détecteur d'évènements a été proposé. Il est simple car basé sur l'utilisation de l'écart-type de l'enveloppe détectée du signal courant. Il est aussi rapide grâce à des calculs itératifs. De plus, le choix du seuil de détection est guidé par le niveau du bruit dans le signal mesuré.

Les résultats montrent que l'algorithme proposé HAND est plus juste (localisation plus précise) que le KFDA (Wild et al., 2015) pour une tolérance au retard $\Delta t = 200$ ms (à RSB = 50 dB). Cet algorithme a aussi un rappel plus élevé que celui du KFDA. En contrepartie, sa précision est sensible au bruit alors que le KFDA est très robuste à ce dernier.

Chapitre 4

Identification des appareils électriques

Sommaire

4.1	Preuve de concept–utilité du transitoire <i>vs.</i> régime établi	96
4.1.1	Analyse harmonique	97
4.1.2	Résultats expérimentaux	99
4.2	Création d’une base de transitoires	108
4.2.1	Enjeu	108
4.2.2	Description du système de mesure réalisé et ses principales caractéristiques	110
4.2.3	Exemples de mesures	112
4.2.4	Organisation de la base de données	119
4.3	Classification des appareils	121
4.3.1	Pré-analyse des paramètres estimés	122
4.3.2	Nuages de points des paramètres estimés	126
4.3.3	Identification	133
4.3.4	Clustering	136
4.3.5	Comparaison des résultats de classification vis-à-vis de la preuve du concept	139

4.1 Preuve de concept–utilité du transitoire *vs.* régime établi

Afin de montrer l'utilité du régime transitoire par rapport au régime établi pour une finalité d'identification d'appareils électriques, nous proposons dans cette section de faire une étude comparative entre les deux vis-à-vis de cette tâche.

Nous proposons l'utilisation des Modèles de Markov Cachés (MMC) pour la modélisation et la classification des transitoires électriques. Il existe plusieurs travaux antérieurs utilisant les MMC pour le NILM (Zia et al., 2011; Beckel et al., 2012; Parson, 2012) et plus récemment par (Ridi and Hennebert, 2014) et (Parson et al., 2014). Dans ces travaux, les auteurs ont appliqué leurs approches sur des signaux basse fréquence. Par exemple, (Beckel et al., 2012) a utilisé la base de données REDD¹ ayant une fréquence d'échantillonnage de 1 Hz et (Ridi and Hennebert, 2014) a utilisé la base de données ACS-F1² où les signaux sont échantillonnés à 10 Hz. Dans notre étude, nous proposons l'utilisation des MMC sur des signaux échantillonnés à haute fréquence et nous utilisons à cet effet la base de données Plug Load Appliance Identification Dataset (PLAID) récemment publiée (PLAID, 2015).

Il est à noter que l'absence d'une base de données *appropriée* qui soit informative, diversifiée (diverse) et évolutive (scalable) (Lai et al., 2012) de signaux électriques échantillonnés à haute fréquence a beaucoup entravée le développement du NILM-haute fréquence. Ce fut l'une des préoccupations qui ont été discutées par la communauté NILM au cours du 2^e Workshop européen de NILM (juillet 2015) organisé à Londres³ et ce problème est encore d'actualité.

Dans la littérature, à notre connaissance, les auteurs de (Thiruvaran et al., 2013) sont les seuls qui ont fait une étude sur des signaux haute fréquence (10 kHz) en utilisant des MMC. Dans (Thiruvaran et al., 2013) les auteurs ont proposé d'utiliser les coefficients de la Transformée de Fourier à Court-Terme (TFCT) et de la Transformée en Ondelettes (TO) comme caractéristiques discriminantes pour leur système d'identification. Concernant la

1. <http://redd.csail.mit.edu/>

2. <http://www.wattict.com/web/index.php/databases/acs-f1>

3. <http://www.nilmm.eu/>

base de données, ils ont recueilli des mesures de courant et de tension pour quatre appareils différents (une lampe fluorescente, une lampe à incandescence, un écran d'ordinateur et un moteur). Pour acquérir ces signaux, les quatre appareils ont été utilisés pour former des séquences de mise en fonctionnement (24 séquences possibles au total). Par exemple, une séquence possible est : la lampe fluorescente est d'abord mise en marche, puis la lampe à incandescence, l'écran d'ordinateur puis le moteur. Suite à cette étude, les auteurs affichent un taux d'identification de 97,9% en utilisant la TFCT et 93,75% en utilisant la TO.

Dans notre étude, nous proposons de considérer non seulement le transitoire mais aussi le régime établi et nous quantifierons l'apport de la partie transitoire par rapport au régime établi pour l'identification d'appareils électriques.

Il est à noter que les taux obtenus dans (Thiruvanan et al., 2013) ne peuvent être utilisés pour quantifier la précision réelle du système d'identification à base de MMC en raison de la petite taille de la base de données utilisée. En contrepartie, l'utilisation de la base de données PLAID devrait permettre l'obtention de résultats plus fiables. De plus, nous proposons dans notre étude l'évaluation de l'ordre des harmoniques choisies sur la précision du système d'identification proposé.

Une autre différence par rapport au travail de (Thiruvanan et al., 2013) est que les signaux de la base de données PLAID sont des signaux électriques d'appareils individuels. Cela signifie que chaque signal a été acquis dans une configuration où un seul appareil est en fonctionnement. L'analyse de ces signaux devrait permettre la détermination de caractéristiques *intrinsèques* à haute fréquence des différents appareils.

4.1.1 Analyse harmonique

Il existe différentes façons d'analyser le contenu harmonique d'un signal. La méthode la plus connue est l'utilisation de la Transformée de Fourier (TF). Pour un signal à temps-discret $s[n]$ on définit la Transformée de Fourier Discrète (TFD) comme suit :

$$S[k] = \sum_{n=0}^{L-1} s[n] \exp\left(\frac{-j2\pi kn}{L}\right), \quad k = 0, \dots, L-1 \quad (4.1)$$

où L est la longueur de $s[n]$ en nombre d'échantillons. Le calcul du contenu harmonique de $s[n]$ en utilisant l'équation (4.1) devient de plus en plus lourd que la valeur de L est élevée, même en utilisant des algorithmes rapides tel que la Fast Fourier Transform (FFT). Ce problème devient important, en particulier, lorsque la base de données est grande et que un grand nombre de signaux doit être analysé. Étant donné que les signaux de courant électrique sont périodiques, nous proposons d'exploiter cette caractéristique pour améliorer le temps de calcul en utilisant la décomposition en Séries de Fourier (SF) au lieu de la TF. La SF donne des instances particulières de la TF (certaines fréquences) (Mallat, 1999). Cela signifie que pour obtenir la TF d'un signal périodique il suffit de calculer ses coefficients de SF. Pour un signal *périodique* $s[n]$, la décomposition en Séries de Fourier Discrètes (SFD) est donnée par :

$$s[n] = \sum_{k=0}^{N-1} C_k \exp\left(\frac{j2\pi kn}{N}\right), \quad (4.2)$$

où N est la période de $s[n]$ en échantillons et

$$C_k = \frac{1}{N} \sum_{n=0}^{N-1} s[n] \exp\left(\frac{-j2\pi kn}{N}\right), \quad k = 0, \dots, N-1 \quad (4.3)$$

sont les coefficients SFD de $s[n]$. Même si les équations 4.1 et 4.3 se ressemblent (à un facteur près), il existe une différence principale qui est le nombre d'échantillons sur lequel la somme est calculée. En sommant sur la période du signal N (nécessairement $< L$ pour un signal périodique) au lieu que sur toute la longueur du signal on réduit le nombre d'opérations de calcul.

Les coefficients SFD C_k , $k = 1, \dots, N-1$ (équation (4.3)) correspondent à la fréquence fondamentale du signal périodique ($k = 1$) et ses harmoniques. Pour le régime établi des signaux électriques de courant, ces fréquences semblent contenir la plupart, sinon la totalité, de l'information contenue dans le signal.

Cependant, pour le régime transitoire des signaux électriques de courant, la propriété de périodicité n'est pas nécessairement conservée, et à vrai dire, on n'est pas autorisé à utiliser les équations (4.2) et (4.3). Néanmoins, nous pensons que, même pour les signaux transitoires, l'important de l'information se trouve *autour* de la fréquence fondamentale

et ses harmoniques. Voilà pourquoi nous choisissons de représenter les signaux transitoires également en utilisant les coefficients SFD.

Du fait que notre approche est basée sur la subdivision des signaux de courant électrique en utilisant des fenêtres glissantes et que le calcul de SFD se fait sur chacune des fenêtres, les coefficients obtenus peuvent être appelés coefficients des Séries de Fourier Discrètes à Court-Terme (SFDCT). Par conséquent, pour le reste de la section, nous allons utiliser la notation SFDCT au lieu de SFD. Chaque vecteur de coefficients SFDCT (calculé sur une fenêtre) est appelée une observation. Voir la figure 4.1 pour une illustration de la méthode de construction des observations.

Ainsi, chaque signal est représenté par une séquence de vecteurs considérés comme une séquence d'observations pour l'identification utilisant les MMC.

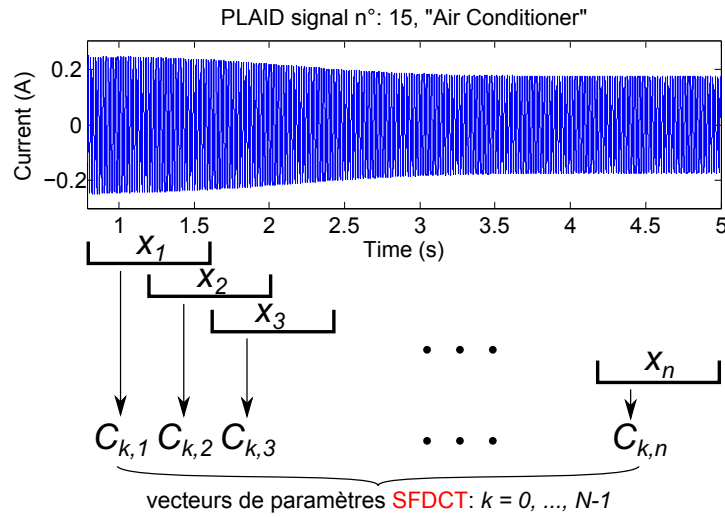


FIGURE 4.1 – Illustration de l'extraction des vecteurs de paramètres $C_{k,i}$, $i = 1, \dots, n$ (observations).

4.1.2 Résultats expérimentaux

Un système d'identification est principalement basé sur trois choses : une base de données, un algorithme d'extraction de paramètres (caractéristiques) (éventuellement avec des règles pour la sélection des paramètres pertinents) et un classifieur. Nous avons choisi la base PLAID pour laquelle nous donnons une brève description dans la sous-section suivante. Cette base de données contient des signaux de courant et de tension (parfois la puissance instantanée aussi) de différents appareils. Pour notre étude, nous avons décidé

d'utiliser uniquement les signaux de courant en tant que signature pour identifier les appareils pour deux raisons principales. La première est que la forme d'onde de la tension ne devrait pas beaucoup changer pour les différents appareils et donc ne peut ajouter beaucoup d'information au système d'identification. La deuxième raison est que l'on a trouvé sur cette base un grand nombre de signaux de tension qui sont mal calibrés (à savoir ayant des amplitudes incorrectes qui divergent beaucoup de la valeur efficace standard de la tension du réseau) et il nous était pas possible de corriger par rapport à ça car les erreurs d'amplitude constatées étaient très variables.

L'algorithme d'extraction de paramètres utilisé est basé sur les coefficients SFDCT déjà discutés dans la sous-section 4.1.1 et les paramètres considérés sont les amplitudes de ces derniers. Pour le classifieur, nous avons choisi d'utiliser les MMC pour leur capacité à modéliser les variations temporelles dynamiques des signaux.

Un MMC est une machine à états finis qui change d'état une fois pendant chaque unité de temps (la valeur de l'unité de temps dépend du phénomène observé). Pour chaque instant t où un état j est entré, un vecteur d'observation $\mathbf{o}_t$ est généré à partir d'une densité de probabilité $b_j(\mathbf{o}_t)$. La transition de l'état i à l'état j est également probabiliste et est régie par la probabilité discrète a_{ij} . La probabilité d'émission b_j , pour l'état j et l'observation $\mathbf{o}_t$ au temps t , est modélisée par un modèle de mélange gaussien (MMG) (Gaussian Mixture Model–GMM) (Hacine-Gharbi et al., 2012) :

$$b_j(\mathbf{o}_t) = \sum_{m=1}^M c_m \mathcal{N}(\mathbf{o}_t; \boldsymbol{\mu}_m, \Sigma_m), \quad (4.4)$$

où $\mathcal{N}(\mathbf{o}; \boldsymbol{\mu}, \Sigma)$ est une gaussienne multivariée à moyenne $\boldsymbol{\mu}$ et de matrice de covariance Σ . M gaussiennes sont utilisées dans le mélange, chacune pondérée par un poids c_m .

Le classifieur est basé sur la modélisation de chaque type d'appareil par un MMC en utilisant la boîte à outils Hidden Markov Model Toolkit (HTK) (Steve Young, 2009). L'apprentissage se fait sur plusieurs étapes en appliquant le embedded Baum–Welch reestimation (commande HEREST). Pour l'étape d'identification, nous utilisons l'algorithme de Viterbi (commande HVITE). Plus de détails sur les algorithmes utilisés pour l'apprentissage et l'identification peuvent être trouvés dans (Steve Young, 2009).

Pour notre système d'identification, nous avons choisi de modéliser chaque type d'appareil en utilisant 3 états, et chaque état avec 3 gaussiennes dans le mélange. Chaque signal a été subdivisé en fenêtres glissantes. De chacune de ces fenêtres N paramètres peuvent être extraits (en pratique, il est possible de choisir un nombre réduit de paramètres pour travailler avec. Voir la fin de la sous-section 4.1.2). La taille de la fenêtre a été fixée à 16,7 msec (500 échantillons à la fréquence d'échantillonnage de 30 kHz) qui correspond à la période de 60 Hz. Le chevauchement entre fenêtres est de 50%. Les résultats d'identification sont évalués en utilisant le Taux de Classification (TC) :

$$TC(\%) = \frac{T - M}{T} \times 100, \quad (4.5)$$

où T est le nombre total de signaux testés et M le nombre de signaux mal classifiés. La figure 4.2 donne le diagramme bloc du système proposé.

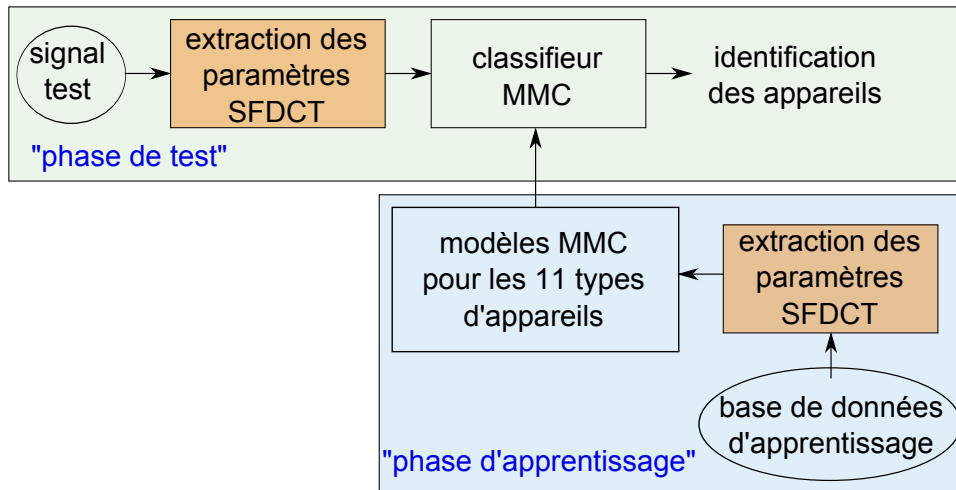


FIGURE 4.2 – Diagramme bloc du système d'identification d'appareils électriques proposé.

Résumé de la base de données PLAID

La Plug Load Appliance Identification Dataset (PLAID) est une base de données publique de signatures électriques. Ces signatures sont des mesures de courant et de tension prises au cours de l'été 2013 à Pittsburgh, Pennsylvanie, États-Unis, à partir de 55 ménages. Cette base de données contient 11 types d'appareils et pour chaque appareil de trois à six exemples (l'ensemble de données contient un total de 1074 signatures). Chaque signature de la base de données fait quelques secondes de durée, contenant le transitoire

de mise en marche (lorsqu'il est disponible, puisque pour certains appareils comme le « Réfrigérateur » fonctionnant habituellement en permanence, le transitoire de mise en marche n'est pas toujours présent) et une partie du régime établi (correspondant à la phase de la consommation régulière). Ces signaux sont échantillonnés à une fréquence de 30 kHz.

Le tableau 4.1 résume les appareils constituant la base de données : les différents types d'appareils et le nombre d'exemples existant dans la base pour chaque type. Pour plus de détails sur cette base, le lecteur est prié de se référer à (Gao et al., 2014).

TABLE 4.1 – Résumé des appareils trouvés dans la base PLAID

Type d'appareil	Nombre d'exemples
Climatiseur	66
Lampe Fluo-compact	175
Ventilateur	115
Réfrigérateur	38
Séchoir	156
Radiateur	35
Lampe à incandescence	114
Ordinateur portable	172
Micro-ondes	139
Aspirateur	38
Machine à laver	26
Total	1074

Il est intéressant de souligner que le nombre d'exemples que nous avons trouvé dans la base de données pour chaque type d'appareil était la plupart du temps différent de celui indiqué dans (Gao et al., 2014). Le nombre total d'exemples dans la base est également différent. Cela peut être dû à une mise à jour de la base de données après la publication du document (Gao et al., 2014). Deux exemples de signaux de PLAID sont donnés dans la figure 4.3.

4.1. PREUVE DE CONCEPT-UTILITÉ DU TRANSITOIRE VS. RÉGIME ÉTABLI

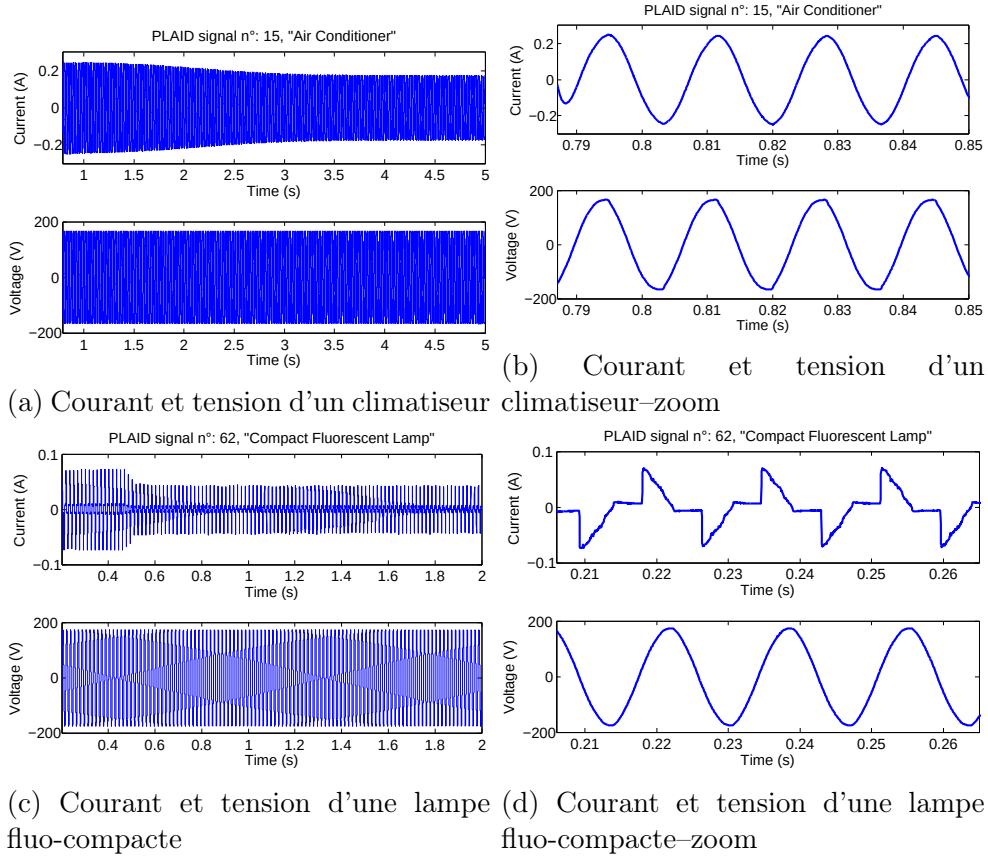


FIGURE 4.3 – Deux exemples de signaux de la base PLAID.

Subdivision de la base PLAID pour l'apprentissage et le test

Pour l'évaluation du système d'identification proposé, la base de données est subdivisée en un ensemble d'apprentissage qui permet de faire apprendre les modèles MMC des appareils au classifieur ainsi qu'un ensemble pour l'évaluation de la performance du système d'identification. Il est considéré que tous les ménages (55 au total) ont des exemples dans les deux ensembles : d'apprentissage et de test.

Afin d'étudier l'effet du nombre de versions d'un signal appartenant à un appareil donné qu'on met dans les ensembles d'apprentissage et de test, nous proposons quatre subdivisions de la base PLAID. Dans la première, nous prenons pour chaque ménage, au plus une version du signal de référence correspondant à un appareil particulier et les autres versions sont conservées pour le test. Cette subdivision nous donne 230 exemples pour l'apprentissage et 844 exemples pour le test. Les autres subdivisions sont obtenues en modifiant le nombre maximal de versions à mettre dans l'ensemble d'apprentissage de 2 à

4. Ainsi, nous obtenons les quatre subdivisions indiquées dans le tableau 4.2, où sbi indique la subdivision ayant « i » versions du signal correspondant à un appareil donné dans l'ensemble d'apprentissage. NAPP est le nombre d'exemples gardés pour l'apprentissage et NTEST est le nombre d'exemples gardés pour le test. Pour la suite, les résultats sont donnés pour chacune des subdivisions susmentionnées.

TABLE 4.2 – Les différentes subdivisions de la base de données créées pour l'évaluation du système d'identification

sb #	sb1	sb2	sb3	sb4
NAPP	230	441	648	848
NTEST	844	633	426	226

Utilité du transitoire *vs.* régime établi pour l'identification des appareils électriques

Le but de l'expérience réalisée dans cette sous-section est d'étudier l'utilité du transitoire par rapport au régime établi vis-à-vis de la tâche d'identification d'appareils électriques. Cette expérience nous a obligé à segmenter les signaux de la base de données chacun en une partie transitoire et une partie de régime établi. Cette segmentation a été faite en caractérisant la partie transitoire comme celle ayant la plus grande variation d'énergie comparée à celle du régime établi.

Pour l'évaluation de performance, le taux de classification TC a été calculé en considérant trois situations : en n'utilisant que la partie transitoire (PT), en n'utilisant que la partie régime établi (PRE) ainsi qu'en utilisant la totalité du signal (TS). Cette évaluation prend en compte les différentes subdivisions de la base de données mentionnées dans la sous-section 4.1.2.

Notons que dans cette expérience, nous n'avons utilisé que 10 coefficients SFDCT. Le tableau 4.3 donne le TC pour chaque subdivision sbi en considérant une partie différente du signal (PT, PRE et TS) à chaque fois.

TABLE 4.3 – Évaluation du TC pour le transitoire *vs.* le régime établi

sb #	sb1	sb2	sb3	sb4
TC (%), PT	88.98	91.15	90.85	92.48
TC (%), PRE	86.61	88.94	86.62	88.50
TC (%), TS	86.02	87.36	87.56	88.50

Nous pouvons conclure que l'utilisation de la partie transitoire est celle qui donne le meilleur TC . De plus, les résultats montrent que le temps de calcul lors de l'utilisation de la partie transitoire seule est beaucoup plus faible comparé aux autres cas. Par exemple, pour sb2, le temps d'exécution du test était de 218 secondes pour PT, 418 secondes pour PRE et 436 secondes pour TS. Le matériel utilisé pour l'expérience est un ordinateur muni d'un processeur Intel Core i5 et 4 Go de mémoire vive RAM.

De ce fait, l'utilisation du transitoire doit permettre la construction d'un système d'identification d'appareils qui est rapide et précis, ce qui justifie son utilité pour la tâche d'identification d'appareils électriques.

Afin d'améliorer l'identification, nous avons étudié l'influence de l'ordre des harmoniques (pair ou impair) sur l'identification ainsi que le nombre optimal de coefficients SFDCT à choisir. Les détails sont donnés dans les prochaines sous-sections.

Harmoniques d'ordre Pair *vs.* impair pour l'identification des appareils électriques

L'expérience présentée dans cette sous-section permet l'étude de la pertinence de l'ordre des harmoniques (pair ou impair) pour la tâche d'identification. Cette expérience est similaire à celle effectuée dans la sous-section 4.1.2. Suite aux résultats de cette dernière, nous ne considérons que la partie transitoire dans cette expérience. Pour l'évaluation de la performance, nous avons comparé le TC du système d'identification lors de l'utilisation des harmoniques, respectivement, d'ordre pair seulement, d'ordre impair seulement, et lors de l'utilisation des deux à la fois. Pour chacune de ces configurations, nous avons utilisé les 5 premiers coefficients SFDCT correspondants (à savoir, les 5 premières harmoniques pairs, les 5 premières harmoniques impaires et les 5 premières harmoniques indépendamment de l'ordre). Le tableau 4.4 donne le TC obtenu pour chaque configuration testée.

TABLE 4.4 – Évaluation du TC pour des harmoniques d'ordre pair *vs.* impair

sb #	sb1	sb2	sb3	sb4
TC (%), pair	63.63	64.45	64.79	63.27
TC (%), impair	90.05	89.57	89.44	90.71
TC (%), (pair+impair)	88.98	91.15	90.85	92.48

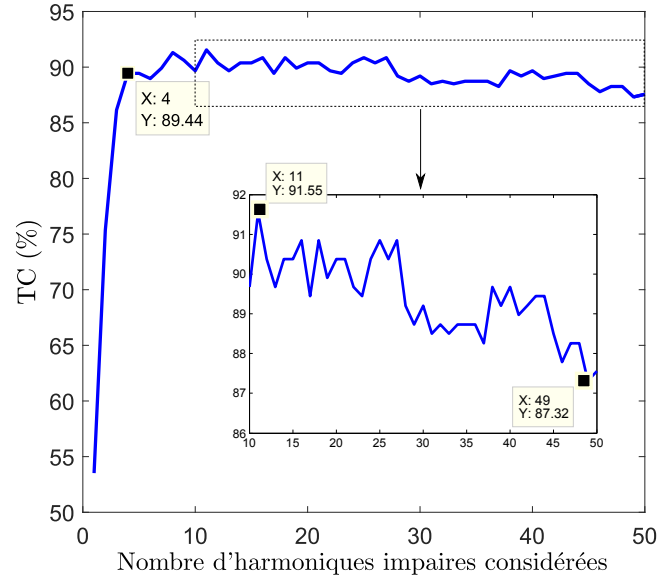
Du tableau 4.4, il est possible de déduire les conclusions suivantes :

- les harmoniques d'ordre impair donnent les meilleurs résultats de performance comparées aux harmoniques d'ordre pair, peu importe la subdivision considérée ;
- les harmoniques d'ordre pair donnent les pires des résultats de performance de tous ;
- l'ajout des harmoniques d'ordre pair aux harmoniques d'ordre impair ne semble pas trop améliorer le TC .

Pour conclure, les coefficients de la fréquence fondamentale et de ses harmoniques d'ordre impair sont les plus utiles pour la tâche d'identification d'appareils électriques. Ceci peut être expliqué par la symétrie demi-onde qu'on trouve généralement dans les signaux électriques (c.-à-d. pour un signal périodique $g(t)$ de période T , on a $g(t) = -g(t \pm T/2)$). Cette symétrie demi-onde annule les coefficients d'ordre pair.

Nombre optimal d'harmoniques pour l'identification des appareils électriques

Un paramètre très important à prendre en compte est le nombre optimal (le plus petit assurant une bonne performance) des coefficients SFDCT à considérer afin d'avoir un bon TC avec le minimum de coefficients possibles. Pour obtenir une indication sur ce nombre optimal, nous avons évalué le système d'identification en utilisant divers nombres de coefficients SFDCT. Notons que nous n'avons considéré que les coefficients d'ordre impair et allant de 1 à 50. La figure 4.4 donne les résultats obtenus. Cette figure montre le TC en fonction du nombre d'harmoniques impaires prises en compte pour l'identification. Les résultats montrent l'absence d'amélioration significative au delà de 4 harmoniques impaires.

FIGURE 4.4 – TC en fonction du nombre d'harmoniques d'ordre impair considérées.

En outre, après avoir atteint une valeur crête (pour 11 harmoniques impaires) et après avoir fluctué autour d'une valeur plus ou moins stable (jusqu'à 27 harmoniques impaires) inférieure à la valeur de crête, le TC commence à diminuer (zoom sur la figure 4.4) en prenant de plus en plus de coefficients. Ce phénomène est très connu dans le domaine de la reconnaissance des formes et est appelé la malédiction de la dimensionnalité (curse of dimensionality ou peaking phenomenon) (Jain et al., 2000). La dimension du vecteur de paramètres dépend de la dimension de la base de données utilisée. Pour éviter ce problème, et en règle générale, on a l'habitude de dire que le nombre de données d'apprentissage devrait être une fonction exponentielle de la dimension du vecteur de paramètres (Jain et al., 2000).

Finalement, nous concluons d'après cette étude que l'utilisation combinée de la partie transitoire des signaux de courant électrique avec seulement quelques harmoniques d'ordre impair permet de construire un système d'identification d'appareils électriques qui est précis, rapide et peu complexe en terme d'occupation mémoire et de calculs.

4.2 Création d'une base de transitoires

4.2.1 Enjeu

Bien que les travaux dans le domaine du NILM se sont multipliés ces dernières années, la plupart des méthodes et techniques proposées restent basées sur l'utilisation de signaux électriques échantillonnés avec une Basse Fréquence d'Échantillonnage (BFE) (période d'échantillonnage de quelques secondes à quelques minutes).

La diffusion de différentes bases de données BFE a beaucoup aidé dans le développement du NILM-BFE et ses techniques. On peut citer REDD (Kolter and Johnson, 2011) (contenant des données basse et haute fréquence et est la première base de son genre à être diffusée publiquement en 2011), Smart* (Barker et al., 2012), AMPds 2 (Makonin et al., 2013), iAWE (Batra et al., 2013), ECO (Beckel et al., 2014), GREEND (Monacchi et al., 2014), SustData (Pereira et al., 2014), Dataport (Parson et al., 2015) et DRED (Utama Nambi et al., 2015).

En contrepartie, l'absence de bases de données « adéquates » ayant des signaux échantillonnés avec une Haute Fréquence d'Échantillonnage (HFE) (quelques centaines de Hz au MHz) à utilisation dans le NILM-HFE a beaucoup freiné le développement de ce dernier.

À ce jour, les seules bases de données HFE disponibles publiquement sont :

- REDD (Kolter and Johnson, 2011) : ayant une fréquence d'échantillonnage de 15 kHz et juste des mesures globales de ménages (pas d'étiquetage ou de sous-comptage de la consommation individuelle) ;
- BLUED (Anderson et al., 2012a) : ayant une fréquence d'échantillonnage de 12 kHz. C'est une mesure globale avec des étiquettes pour l'activité des appareils individuels ;
- PLAID (Gao et al., 2014) : ayant une fréquence d'échantillonnage de 30 kHz et juste des mesures d'appareils individuels ;
- HFED (Gulati et al., 2014) : mesures sur un intervalle de fréquence allant de 10 kHz à 5 MHz. La base contient des traces de signaux d'interférence électromagnétiques (ElectroMagnetic Interference–EMI) collectées au laboratoire et dans le résidentiel ;
- UK-DALE (Kelly and Knottenbelt, 2015) : ayant une fréquence d'échantillonnage

4.2. CRÉATION D'UNE BASE DE TRANSITOIRES

de 16 kHz et juste des mesures globales de ménages ;

- WHITED (Kahl et al., 2016) : ayant une fréquence d'échantillonnage de 44.1 kHz.

La base contient des signatures individuelles de signaux transitoires (les 5 premières secondes de fonctionnement) acquises à différents endroits du monde.

Les bases de données HFE n'ayant que des mesures d'appareils individuels (comme PLAID par exemple) peuvent être utilisés pour l'*apprentissage* de certaines méthodes d'apprentissage automatique (Machine Learning–ML) afin d'identifier les types d'appareils dans le cas d'un scénario. Pour avoir un scénario plus réaliste avec un ensemble d'appareils en fonctionnement, on peut imaginer la simulation d'agrégats en superposant aléatoirement les signatures individuelles (comme nous l'avons fait pour le test du détecteur d'évènements, sous-section 3.4.2). Néanmoins, une telle approche, utilisable pour tester les performances de détection, n'est pas adaptée pour le *test* de la performance d'identification. On a besoin de scénarios de vie réelle pour vérifier l'efficacité des algorithmes HFE dans des situations réelles.

Contrairement à cela, les bases de données HFE n'ayant que des mesures globales de ménages sont adaptées pour le *test* mais le sont moins pour l'*apprentissage* puisque les signatures individuelles des appareils ne sont pas directement disponibles pour permettre cette tâche.

Une base *appropriée* pour le NILM-HFE doit être *informative* (Lai et al., 2012). Plus on a de contrôle sur la mise en place des mesures, plus on acquière d'information.

Pour toutes ces raisons, nous avons décidé de créer notre propre système de mesure qui est adapté à l'apprentissage et au test et permettrait de construire une base de données appropriée pour le NILM-HFE. Les principales caractéristiques de notre système de mesure sont :

- la capacité d'échantillonner très haut en fréquence ;
- le contrôle précis sur les instants de mise en marche et de mise à l'arrêt ;
- créer des scénarios et mettre en fonctionnement jusqu'à six appareils simultanément.

Dans ce qui suit, nous allons décrire le système de mesure et nous donnerons quelques exemples qui mettent en évidence ses capacités et son utilité.

4.2.2 Description du système de mesure réalisé et ses principales caractéristiques

Le système (figure 4.5) permet la mesure de signaux de courant et de tension d'appareils électriques avec une HFE (la carte d'acquisition est une National Instrument NI USB-6356 pouvant échantillonner jusqu'à 1,25 MHz). La fréquence d'échantillonnage de travail est choisie en prenant en compte les bandes passantes de nos capteurs : la sonde différentielle de tension (Metrix MX 9030-Z ayant une bande passante de 30 MHz) et la pince de courant (Metrix HX0102 ayant une bande passante de 60 kHz).



FIGURE 4.5 – Photo du système de mesure proposé.

Une caractéristique importante du système est que la durée de mesure est variable et varie de quelques secondes à quelques jours. La seule limite est l'espace mémoire disponible pour le stockage des données et la durée de vie des batteries des capteurs (au début, nous avons considéré alimenter les capteurs du réseau électrique en utilisant des adaptateurs AC mais cette façon de faire a généré beaucoup de bruit sur les mesures ce qui nous a conduit à la rejeter).

Le schéma bloc du système proposé est illustré sur la figure 4.6. Dans un système d'acquisition classique on trouve en général : un ordinateur, une carte d'acquisition ainsi que des capteurs de courant et de tension. L'ordinateur contrôle le lancement des mesures et l'enregistrement des données. La carte d'acquisition fonctionne comme interface entre l'ordinateur et les capteurs qui eux leur rôle est de mesurer le signal.

4.2. CRÉATION D'UNE BASE DE TRANSITOIRES

Pour notre système, nous avons ajouté à cette configuration classique une partie de *contrôle* (figure 4.6). Cette dernière permet de contrôler les instants de mise en marche et de mise à l'arrêt des appareils à mesurer et ceci par rapport à la sinusoïde réseau (tension réseau).

Le « Processeur » est le principal bloc de cette partie de contrôle, et l'ensemble du système d'ailleurs, car il contient le programme qui spécifie comment doit fonctionner le système et selon cela il donne les ordres au bloc « Triacs » (un ensemble de six interrupteurs électroniques) qui contrôle la mise en marche et la mise à l'arrêt des six sorties auxquelles les appareils électriques à mesurer sont connectés. Le « Détecteur de passage par zéro » est un circuit électronique qui détecte, de façon permanente, les points de passage par zéro (passages du positif au négatif) du signal de tension.

Suite à la détection, ce bloc détecteur envoie un signal carré cyclique, correspondant aux points de passage par zéro, au processeur qui l'utilise comme référence pour contrôler la mise en marche et la mise à l'arrêt des appareils.

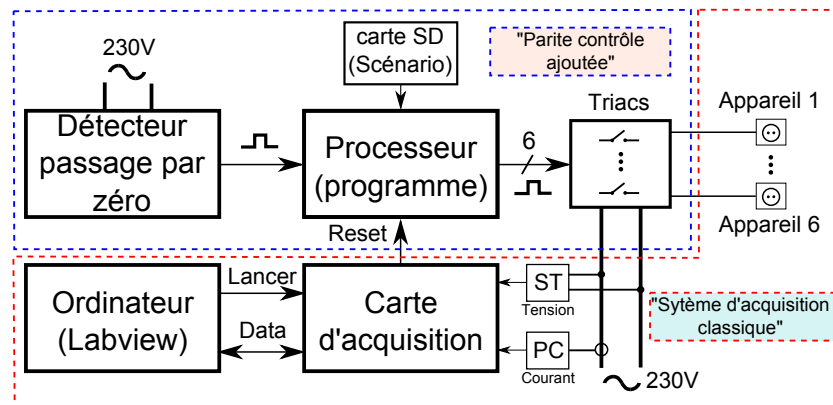


FIGURE 4.6 – Schéma bloc du système de mesure proposé.

Au cours d'une instance de mesure typique, on commence par donner l'ordre de « Lancer » la mesure sur le logiciel Labview. Cette commande permet alors à la « carte d'acquisition » d'envoyer un signal « Reset » au « Processeur » et lance la mesure. Après cela, le « Processeur » lit la « carte SD » qui contient le scénario de mesure pré-chargé et en utilisant la sortie du « Détecteur de passage par zéro » il envoie des commandes pour mettre sous/hors tension les appareils. Pendant ce temps, les capteurs de courant et de tension, respectivement, « PC » et « ST » mesurent le courant électrique consommé ainsi que la tension aux bornes du point regroupant tous les appareils de mesure. Enfin, ces mesures

sont envoyées à l'« Ordinateur » via la « Carte d'acquisition » où ils seront enregistrés.

La forme d'onde du transitoire de mise en marche dépend de l'instant où l'appareil est mis sous tension par rapport à la sinusoïde réseau. Un appareil qui est déclenché lorsque la sinusoïde réseau est à un pic maximal donne une autre forme d'onde transitoire (en terme d'amplitude maximale et aussi de temps de montée) que lorsque la sinusoïde réseau est, par exemple, à un passage par zéro. Puisque les algorithmes NILM-HFE travaillent sur de tels signaux transitoires, ces différences entre les formes d'ondes transitoires peuvent diminuer leur précision d'identification.

Par conséquent, on voit l'utilité d'un système de mesure qui peut reproduire de manière contrôlable de telles formes d'onde différentes. Le système peut choisir de mettre en marche ou à l'arrêt un appareil à différents instants par rapport à la période secteur avec une précision de $1\ \mu\text{s}$. Cela signifie qu'il peut choisir de mettre en marche ou à l'arrêt un appareil sur 2×10^4 instants différents durant les 20 ms de la période secteur.

Une autre caractéristique du système est la possibilité de créer des scénarios de mesure avec jusqu'à 6 appareils en fonctionnement simultané. Ceci permet de reproduire des situations réelles avec des mesures globales mais surtout de manière contrôlable. Cela devrait permettre le test d'algorithmes NILM-HFE pour vérifier les performances d'identification sous certains scénarios d'agrégation qui peuvent être particuliers ou difficiles.

4.2.3 Exemples de mesures

Dans cette partie, nous donnons quelques exemples de mesures prises avec le système proposé. Ces mesures montrent l'utilité du système et certaines de ses capacités. Notons également que le système proposé a été construit principalement pour la mesure d'appareils ménagers (la puissance totale des six appareils ne doit pas dépasser 4600 W limitée par un disjoncteur de 20 Ampères). Par conséquent, les exemples donnés ci-après ne consistent que de tels appareils. La fréquence d'échantillonnage pour tous les exemples est de 100 kHz.

Reproductibilité des mesures

Comme on a le contrôle sur les instants de mise en marche/arrêt des appareils, il est intéressant de voir si les résultats sont reproductibles en répétant la même mesure plusieurs fois. Pour tester cela, nous avons répété la même mesure d'un aspirateur cinq fois et nous avons choisi de mettre en marche l'appareil sur le début de la période secteur (à savoir avec 0 ms de délai de déclenchement) (figure 4.7). On arrive à voir (zoom sur la figure 4.7) que les mesures sont reproductibles puisque toutes les formes d'onde se superposent les unes les autres.

Le fait qu'on voit au début une forme d'onde qui tend vers des valeurs négatives est dû à notre choix de l'instant de mise en marche. Étant donné que le signal de tension au début de la période démarre à un passage par zéro du positif au négatif (c'est une convention que nous avons choisi, nous aurions pu prendre l'inverse bien sûr!) et son amplitude continue de diminuer jusqu'à atteindre la valeur minimale, le signal de courant essaie de suivre cette tendance et on obtient ainsi le pic négatif.

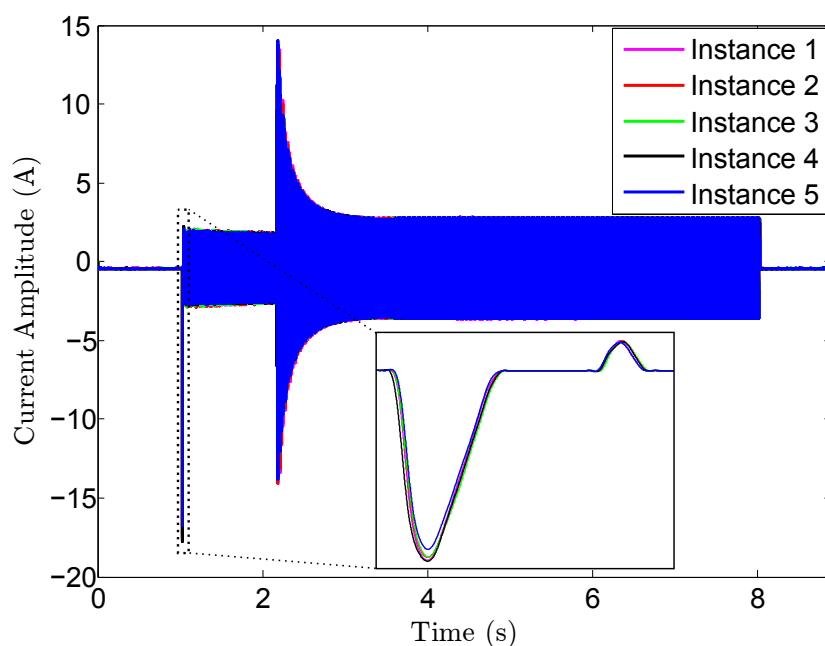


FIGURE 4.7 – Courant électrique d'un aspirateur. Même mesure répétée cinq fois avec un même retard de déclenchement de 0 ms (par rapport à la période secteur).

Mesure des appareils à faible *vs.* grande puissance

Puisque la puissance des appareils électriques habituellement trouvés dans les ménages peut varier de quelques watts (appareils électroniques) à quelques milliers de watts (radiateurs, fours, etc.), le système doit être capable de mesurer aussi bien les charges de puissance faible que les charges de puissance élevée. Pour illustrer la capacité du système proposé par rapport à cette exigence, nous proposons un scénario de mesure (figure 4.8) en utilisant deux charges : une lampe halogène de 100 W et une lampe Light Emitting Diode (LED) de 1,6 W.

Le scénario est comme suit (lu comme : Début, (temps jusqu'au premier évènement) → appareil, évènement, (temps jusqu'au prochain évènement) ... → Fin) :

- Début, (1s) → lampe halogène, ON, (1s) → lampe halogène, OFF, (1s) → lampe LED, ON, (1s) → lampe LED, OFF, (1s) → lampe halogène, ON, (1s) → lampe LED, ON, (1s) → lampes halogène & LED, OFF, (1s) → Fin.

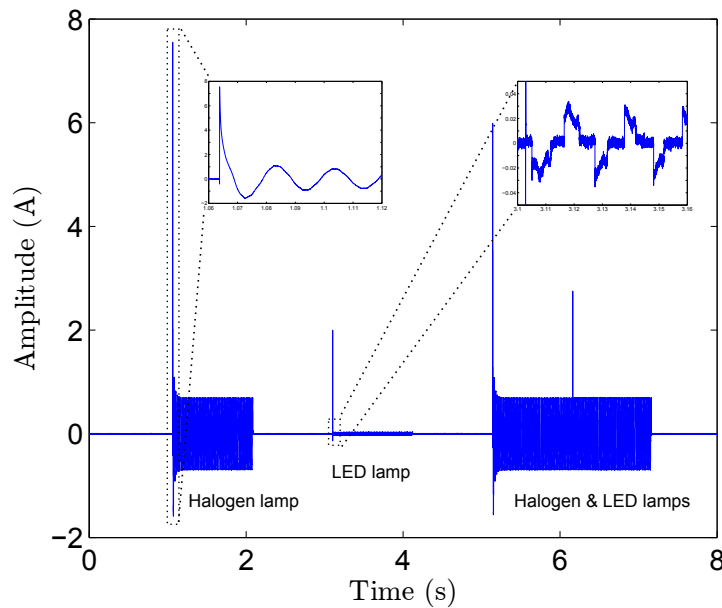


FIGURE 4.8 – Mesure de courant pour une lampe halogène de 100 W et une lampe LED de 1,6 W.

Ce scénario montre que notre système est capable de mesurer une lampe LED avec une consommation aussi faible que 1,6 W (zoom sur la figure 4.8). Ceci est le résultat des efforts déployés pour garantir un faible niveau de bruit sur les mesures. La figure 4.9 montre l'histogramme du bruit du signal courant (pris de la première seconde de la mesure) de notre système de mesure. Ce bruit est normalement distribué avec une moyenne nulle

4.2. CRÉATION D'UNE BASE DE TRANSITOIRES

(le signal de courant a été centré au préalable) et un écart-type de 2,2 mA correspondant à une résolution en puissance de 0,5 W. Cela signifie que le système peut mesurer une charge consommant 1 W avec un RSB de 3 dB.

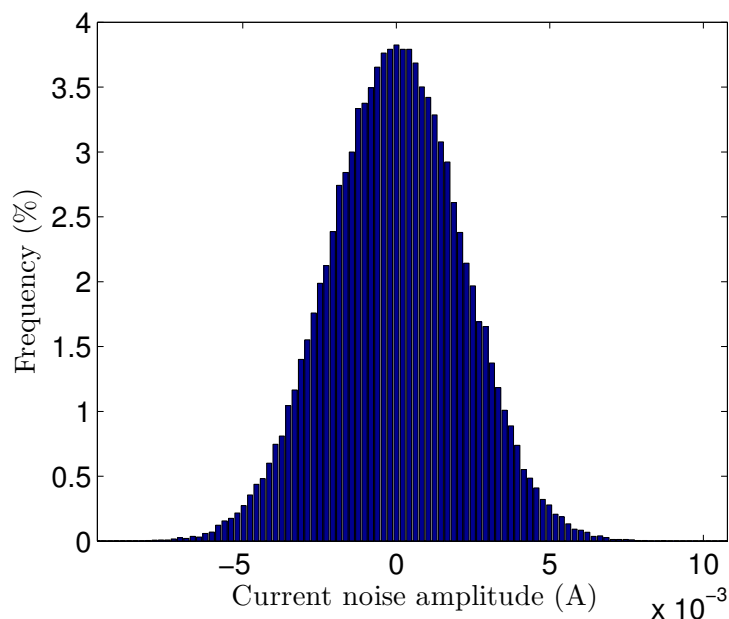


FIGURE 4.9 – Histogramme du bruit de courant.

On peut également constater que, lorsque la lampe halogène et la lampe LED fonctionnent au même temps, il est très difficile de détecter la présence de la lampe LED si l'on ne considère que le régime établi du courant. Ceci est dû au rapport des consommations relatives en puissance des deux charges (la lampe LED consomme moins de 2% de ce que la lampe halogène consomme). Néanmoins, grâce à la capacité d'échantillonnage fort du système de mesure, on peut détecter la présence de la lampe à LED indiquée par le pic apparaissant à la mise en marche de la LED.

Comportement transitoire *vs.* type d'appareil

Comme indiqué dans la section 4.2.2, le transitoire de mise en marche dépend de l'instant où l'appareil est mis sous tension par rapport à la sinusoïde réseau. Dans cette sous-section, nous allons montrer que cet instant n'est pas le seul facteur qui affecte la forme d'onde transitoire. Le type de la charge (ou appareil) joue également un rôle important dans la définition de la forme d'onde transitoire. Le transitoire de mise en marche est fortement influencé par la tâche physique que l'appareil effectue (Leeb et al., 1995).

Dans la figure 4.10, on montre le transitoire de mise en marche d'une lampe halogène en faisant varier l'instant de déclenchement de celle-ci. Ces instants sont indiqués comme des retards relatifs à un passage par zéro, du positif au négatif (dans la figure 4.10, il se trouve à 40 ms). Par exemple, 0 ms indique que la mise sous tension a été faite au passage par zéro, 5 ms indique qu'elle s'est faite 5 ms après et ainsi de suite.

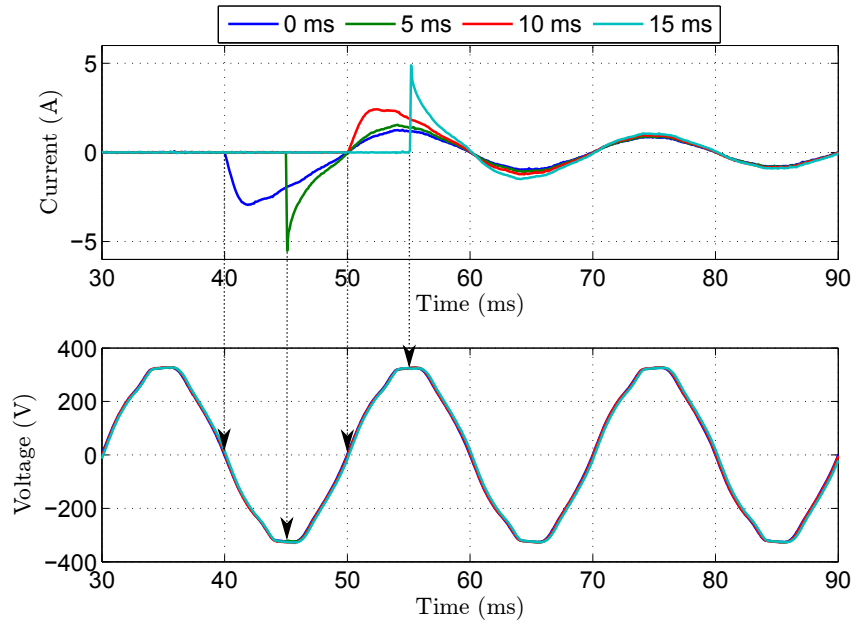


FIGURE 4.10 – Courant (haut) et tension (bas) d'une lampe halogène. On montre quatre exemples d'une mesure faite sur la même charge avec un retard de déclenchement variable (0 ms, 5 ms, 10 ms and 15 ms).

Les signaux résultants (fig 4.10) montrent que le transitoire atteint son extremum (max ou min) d'amplitude pour des retards de déclenchement de 5 ms et 15 ms. Ceci correspond aux extremums de la sinusoïde réseau. Le courant suit les variations de la tension réseau presque instantanément après la mise en marche de la lampe halogène. Notons, par exemple, la vitesse à laquelle l'amplitude du courant atteint une valeur élevée presque instantanément après la mise en marche pour le retard de 15 ms. En effet, la lampe halogène est une charge résistive et la relation courant-tension est régie par la loi d'Ohm, où le courant est directement proportionnel à la tension. Cela se traduit par le changement de la forme de transitoire de mise en marche où l'on peut remarquer que l'amplitude initiale peut même doubler en fonction de l'instant de mise en marche (notons comment l'amplitude maximale a doublé quand on compare les retard de déclenchement de 10 ms et 15 ms).

La figure 4.11 montre un résultat de mesure similaire à celui présenté sur la figure 4.10.

La différence est que la charge cette fois-ci est une perceuse. Par rapport à la figure 4.10, le transitoire de mise en marche ayant l'amplitude maximale la plus élevée n'est pas celui correspondant à la mise sous tension à un extremum de la sinusoïde réseau. En effet, la tâche principale de la perceuse est liée à la rotation d'un moteur où la relation de courant-tension n'est plus une proportionnalité simple, comme ce fut le cas pour la lampe halogène. Lorsqu'elle est mise sous tension, la perceuse ne suit pas instantanément la sinusoïde réseau. Le moteur de cette dernière ne tourne pas à la vitesse maximale instantanément après à la mise en marche, mais a besoin d'un certain temps en raison de son inertie avant qu'il ne puisse atteindre la vitesse désirée.

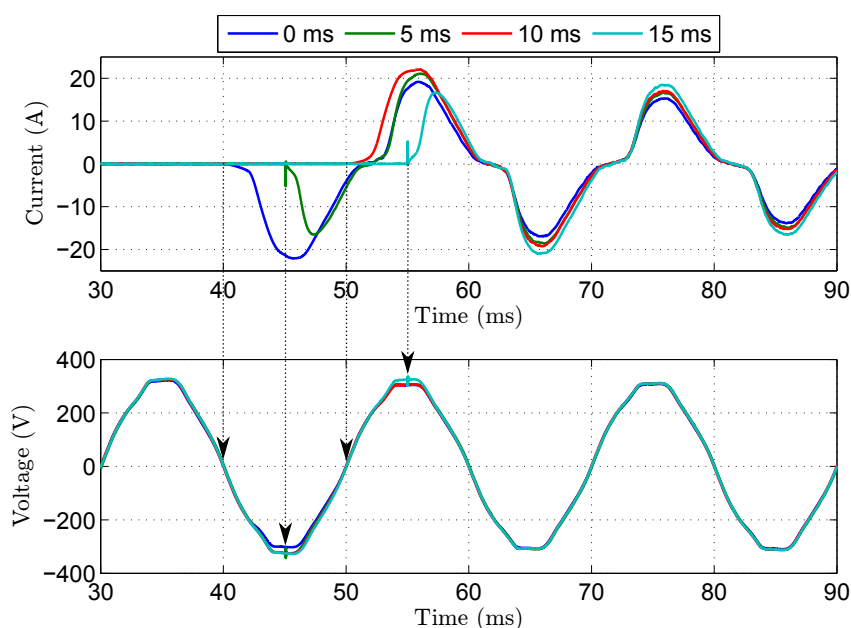


FIGURE 4.11 – Courant (haut) et tension (bas) d'une perceuse. On montre quatre exemples d'une mesure faite sur la même charge avec un retard de déclenchement variable (0 ms, 5 ms, 10 ms and 15 ms).

Il est très important de prendre en compte le principe physique qui régit le fonctionnement d'un appareil lors de l'analyse des transitoires de mise en marche afin de pouvoir comprendre comment celui-ci pourrait affecter la forme d'onde. Pour illustrer cela, nous proposons deux scénarios de mesure représentés sur la figure 4.12. Ces deux scénarios sont :

Premier scénario « lent » :

- Début, (1s) → perceuse, ON, (3s) → perceuse, OFF, (4s) → perceuse, ON, (3s) → perceuse, OFF, (1s) → Fin.

Deuxième scénario « rapide » :

- Début, (1s) $\rightarrow$ perceuse, ON, (3s) $\rightarrow$ perceuse, OFF, (**1s**) $\rightarrow$ perceuse, ON, (3s) $\rightarrow$ perceuse, OFF, (1s) $\rightarrow$ Fin.

Ces scénarios correspondent à allumer puis éteindre une perceuse deux fois consécutives. Pour le premier scénario, on garde la perceuse hors tension pendant 4 secondes avant de la ré-allumer et pour le deuxième scénario, on ne la maintient hors tension que pendant 1 seconde avant de la ré-allumer. Cela correspond à un scénario réel éventuel où on a en général des durées variables entre les utilisations successives.

Le premier et le deuxième transitoire sont identiques pour le premier scénario (figure 4.12a) alors qu'ils sont différents pour le deuxième (figure 4.12b). Ceci est le résultat de l'inertie du moteur de la perceuse. Pour le premier scénario et après la première mise hors tension de la perceuse, son moteur a le temps (4 s) pour s'arrêter complètement avant la prochaine mise en marche. Par conséquent, la deuxième mise en marche se fait à partir d'une position d'arrêt tout comme la première. Pour le deuxième scénario, le moteur de la perceuse n'a pas le temps (1 s) pour s'arrêter complètement avant de lancer la deuxième mise en marche. Par conséquent, lorsqu'il est mis en marche à nouveau, le couple nécessaire pour faire tourner le moteur est plus faible ce qui résulte en une amplitude de courant plus faible après la deuxième mise sous tension.

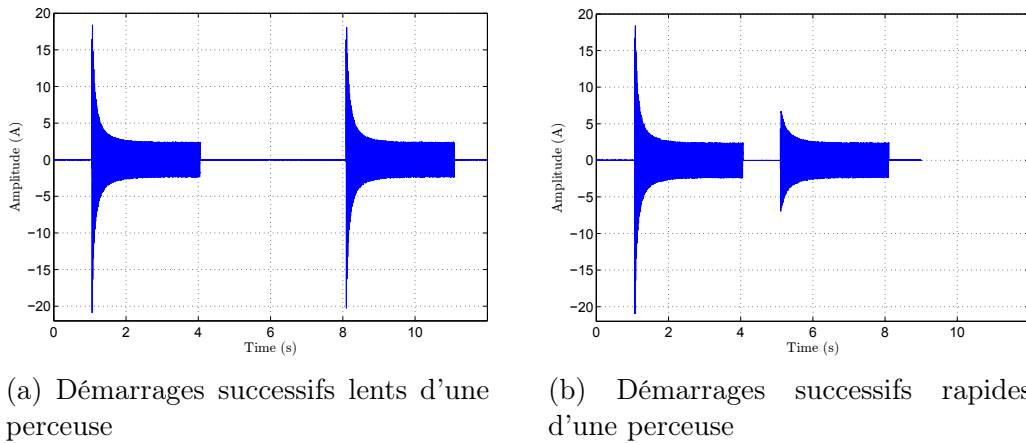


FIGURE 4.12 – Démarrages successifs (a) lents et (b) rapides d'une perceuse.

4.2.4 Organisation de la base de données

Dans cette sous-section, nous allons donner un bref descriptif du contenu de la base de données de courants transitoires de mise en marche que nous avons créé. Cette dernière est appelée Controlled On/Off Loads Library (COOLL).

Les signaux mesurés sont échantillonnés à $F_s = 100$ kHz. COOLL contient des mesures de courant et de tension. En total, elle contient 840 mesures de courant et 840 mesures de tension issues de 42 appareils différents. Les appareils sont de 12 types différents ayant chacun un certain nombre d'exemples (tableau 4.5).

TABLE 4.5 – Résumé de la base de données

N°	Type d'appareil	Type d'appareil (en)	Nombre d'appareils par type	Nombre de signaux (20 par appareil)
1	Perceuse	Drill	6	120
2	Ventilateur	Fan	2	40
3	Meuleuse	Grinder	2	40
4	Sèche-cheveux	Hair dryer	4	80
5	Taille-haie	Hedge trimmer	3	60
6	Lampe	Lamp	4	80
7	Décapeur thermique	Paint stripper	1	20
8	Rabot électrique	Planer	1	20
9	Défonceuse	Router	1	20
10	Ponceuse	Sander	3	60
11	Scie	Saw	8	160
12	Aspirateur	Vacuum cleaner	7	140
Total			42	840

Cette base contient des signatures d'appareils mesurés individuellement. Elle ne contient pas de scénario où plusieurs appareils ont été mesurés en fonctionnement simultané. En plus, les appareils choisis sont tous des appareils où le contrôle, par le système de mesure, de l'instant de mise en marche est possible. C'est à dire qu'on actionnant l'interrupteur de mise en marche de l'appareil *a priori*, l'appareil peut être commandé pour la mise en marche à l'aide des triacs.

Chaque mesure dure 6 secondes avec une durée de pré-déclenchement de 0,5 secondes (les premières mesures étant faites avec un pré-déclenchement de 1 seconde) et une durée de post-arrêt de 1 seconde correspondants, respectivement, aux instants où l'appareil était à l'arrêt avant le début de fonctionnement et après l'arrêt de l'appareil.

Pour chaque appareil, on fait 20 mesures contrôlées où chaque mesure correspond à un délai d'action (mise en marche), repéré par rapport au passage par zéro de la tension secteur, différent variant entre 0 et 19 ms avec un pas de 1 ms.

Afin de bien renseigner la base de données, des fichiers « en-tête », sous format d'un tableau .xlsx, ont été ajoutés à celle-ci. Ces fichiers contiennent des informations propres à l'appareil ainsi qu'à l'environnement où ce dernier a été mesuré.

TABLE 4.6 – Exemple du contenu d'un fichier en-tête « .xlsx »

File name	Device type	Nominal power	Date of purchase	Brand	Reference
Vacuum_cleaner_1	Vacuum cleaner	700 W		Tornado	61EKJ01
Place	Use frequency	Dimmer ?	Controllable ?	Power grid voltage	Power grid frequency
PRISME laboratory	once a week	YES	NO	230 V	50 Hz

En plus des fichiers en-tête fournis, des fichiers décrivant la configuration de la mesure pour chaque appareil sont fournis (figure 4.13). Ces fichiers représentent, en fait, le contenu de la carte SD (figure 4.6) pour chaque mesure. Le contenu de ces fichiers a été pensé pour servir deux finalités : être lu par le processeur (carte Arduino) et être facilement interprétable par l'utilisateur exploitant la base.

```

scenario1_1.txt
1 #NberOfAppliances 1
2 #Outlet1 Drill_1
3 #Config1 Power level 100% - No load
4 #Outlet2
5 #Config2
6 #Outlet3
7 #Config3
8 #Outlet4
9 #Config4
10 #Outlet5
11 #Config5
12 #Outlet6
13 #Config6
14 #Variables definition
15 #act: actions (ON=1/OFF=0)
16 #sofa: selected outlet for each action
17 #ad: action delay wrt volt. Zero-Cross. (µs)
18 #tbna: time before next action (ms)
19 #Scenario
20 tbna0=1000
21 act1=1
22 sofa1=1
23 ad1=0
24 tbna1=4000
25 act2=0
26 sofa2=1
27 ad2=0
28 tbna2=1000

```

FIGURE 4.13 – Exemple d'un fichier de configuration.

4.3 Classification des appareils

Dans cette section nous allons présenter les résultats de la classification des appareils électriques contenus dans COOLL en utilisant les paramètres extraits du modèle que nous avons proposé. Le but de cette section n'est pas de présenter une étude poussée de la problématique de classification d'appareils électrique. On cherche, plutôt, à donner des éléments de réponse par rapport à l'utilité des paramètres de notre modèle de transitoire de mise en marche pour la classification. Mais avant cela, nous allons faire une pré-analyse des paramètres estimés afin de guider la classification (faire des choix préliminaires à la classification tels que l'exclusion des paramètres de phase) et faciliter par la suite l'interprétation des résultats.

Il est possible de diviser la problématique de classification d'appareils électriques en deux sous-problématiques différentes suivant l'objectif visé :

Objectif 1 : reconnaître (identifier) les appareils par leur type ;

Objectif 2 : regrouper (clustering) les appareils par groupes ayant des caractéristiques communes.

Le premier objectif relève du domaine de l'identification et est un problème *supervisé*. Ayant un ensemble de données (appelée base d'apprentissage) labélisé (avec des types d'appareils définis au préalable), la tâche est d'identifier la classe (type d'appareil) d'appartenance d'un nouvel appareil n'appartenant pas à la base d'apprentissage.

Le deuxième objectif relève du domaine de clustering (on parle aussi de regroupement ou partitionnement de données) et est un problème *non supervisé*. Ayant un ensemble de données non labélisé (les types ou classes d'appareils sont inconnus), la tâche est de trouver des clusters, ou regroupements, permettant de définir des classes correspondants à des appareils ayant des caractéristiques communes.

Partant de ces deux objectifs, nous présentons dans cette section, à la suite de l'étape de pré-analyse, les résultats relatifs aux deux tâches d'identification et de clustering d'appareils électriques obtenus en utilisant les paramètres estimés de notre modèle sur la base COOLL.

4.3.1 Pré-analyse des paramètres estimés

L'estimation des paramètres $\mathbf{p}$, $\mathbf{a}$ et $\boldsymbol{\phi}$ des signaux (réels) de COOLL a été faite en utilisant l'algorithme ECTA (algorithme 2). L'analyse de ces estimées nous a permis d'affirmer la non adaptabilité des phases pour la tâche de classification ainsi que de découvrir certaines limitations liées au modèle mathématique que nous avons proposé quant à la tâche de classification ainsi qu'à la représentation des signaux réels.

Non adaptabilité des phases à la classification

Les phases $\phi_i, i = 1, \dots, d$ de notre modèle sont les phases initiales des sinusoides constituant la somme de sinusoides non amorties $s_s(t)$ (équation 3.1). Comme leur nom l'indique elles dépendent de l'instant initial (on réinvente la roue!) qu'on choisi comme début (référence de temps) pour le signal mesuré $x(t)$ qu'on cherche à modéliser. Ceci indique qu'elles ne sont pas adaptées pour être utilisées comme caractéristique invariable du signal qu'on pourra utiliser par la suite pour la tâche de classification (figure 4.14).

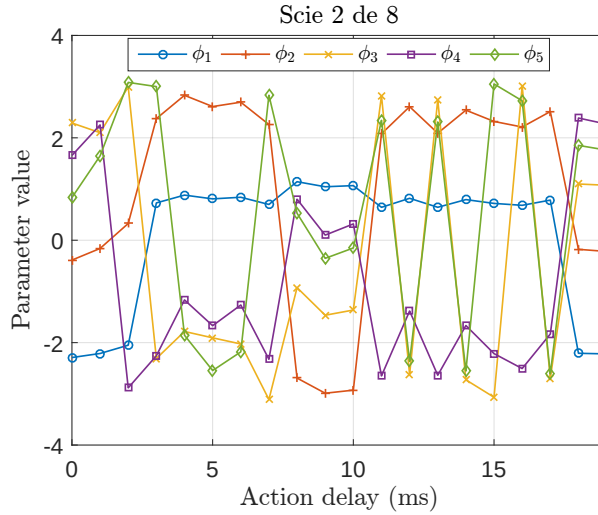


FIGURE 4.14 – Exemple des phases estimées $\hat{\phi}_i$ pour la mesure du même appareil « scie » en fonction de différents délais de mise en marche.

Contrairement aux phases, les amplitudes restent plutôt constantes par rapport au changement de la référence de temps (figure 4.15) et sont donc adaptés à être considérés pour la tâche de classification.

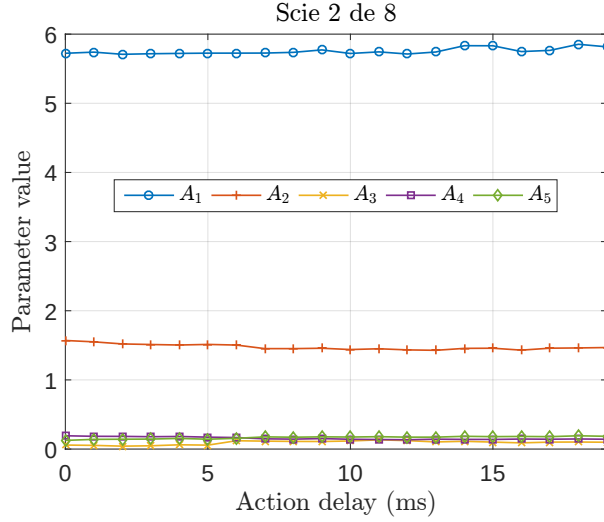


FIGURE 4.15 – Exemple des amplitudes estimées $\hat{A}_i$ pour la mesure du même appareil « scie » en fonction de différents délais de mise en marche.

Limitation du modèle par rapport à la tâche de classification

Tout comme les phases ϕ_i , les paramètres de l'enveloppe $p_j, j = 0, \dots, n$, ont eux aussi une certaine dépendance à la référence de temps représentant le début du signal $x(t)$ à modéliser.

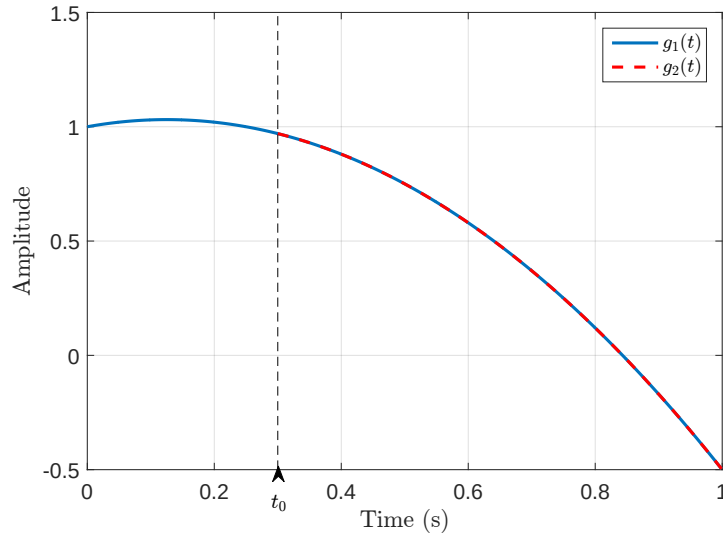
Pour illustrer cela, considérons l'exposant d'une enveloppe simulée, avec un degré $n = 2$, qui s'écrit en fonction du temps t comme suit :

$$g_1(t) = p_0 + p_1 t + p_2 t^2, \quad t \in [0, 1] \text{ s} \quad (4.6)$$

avec $p_0 = 1, p_1 = 0.5$, et $p_2 = -2$ représentant les coefficients du polynôme qui décrit la fonction $g_1(t)$. Supposons maintenant qu'on déplace notre référence de temps de 0 s à $t_0 = 0.3$ s, on obtient alors la fonction $g_2(t) = g_1(t + t_0)$. Cette fonction est juste la portion de la fonction $g_1(t)$ sur le nouvel intervalle d'analyse (figure 4.16).

En estimant les paramètres p'_0, p'_1 , et p'_2 de cette nouvelle fonction $g_2(t)$, on trouve $p'_0 = 0.97, p'_1 = -0.7$, et $p'_2 = -2$. Ceci correspond au développement :

$$\begin{aligned} g_2(t) = g_1(t + t_0) &= p_0 + p_1(t + t_0) + p_2(t + t_0)^2 \\ &= (p_0 + p_1 t_0 + p_2 t_0^2) + (p_1 + 2p_2 t_0)t + p_2 t^2 = p'_0 + p'_1 t + p'_2 t^2 \end{aligned} \quad (4.7)$$

FIGURE 4.16 – Fonctions $g_1(t)$ et $g_2(t)$.

En conclusion, le seul paramètre qui n'est pas affecté par ce changement de référence de temps est le coefficient p_2 du polynôme. On montre par un raisonnement similaire que quelque soit le degré n du polynôme considéré, c'est le dernier coefficient p_n qui a une valeur indépendante de référence de temps.

Pour cette raison, ce paramètre sera privilégié par la suite pour la classification des appareils.

Notons aussi que le dernier coefficient du polynôme reflète la rapidité de décroissance de l'enveloppe. Si on compare les valeurs de ce paramètre, par exemple, pour une lampe et une scie (figure 4.17), on voit clairement que ce paramètre reste plutôt constant pour une scie alors qu'il est très variable pour une lampe. Plus particulièrement, pour la lampe, ce paramètre atteint des valeurs élevées pour des délais de mise en fonctionnement proches des extremum de la sinusoïde secteur. Il est possible de faire un lien avec la figure 4.10 pour déduire que les valeurs élevées (négatives) correspondent à une décroissance plus rapide de l'enveloppe.

Limitation du modèle à la représentation des signaux réels

Après analyse, on s'est rendu compte que pour certains signaux réels, tels que les courants de mise en marche des lampes fluo-compactes, le régime transitoire est absent (plus exactement la signature n'affiche pas un passage transitoire vers le régime établi

4.3. CLASSIFICATION DES APPAREILS

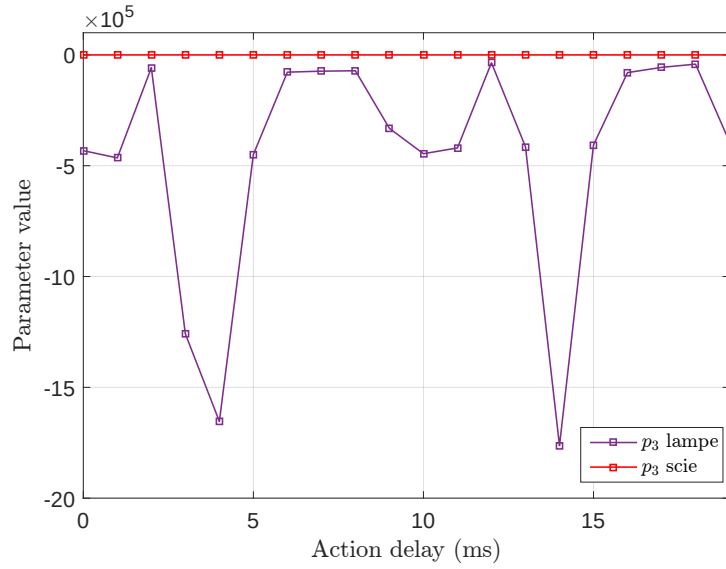


FIGURE 4.17 – Comparaison des valeurs de p_3 pour une lampe et une scie en fonction de différents délais de mise en marche.

(figure 4.18)). Pour ces signaux, les paramètres $p_j, j = 0, \dots, n$ estimés ont des valeurs négatives très élevées. La raison est que le modèle des courants transitoires tel que nous l'avons proposé ne permet pas de gérer des signaux n'ayant que le régime établi $s_s(t)$. Rappelons que notre modèle est le suivant :

$$s(t) = e(t)s_s(t) = \left(e^{\mathbf{p}^T \mathbf{t}} + 1\right) s_s(t), \quad (4.8)$$

avec $\mathbf{p} = [p_0, \dots, p_n]^T$ et $\mathbf{t} = [1, t, \dots, t^n]^T$.

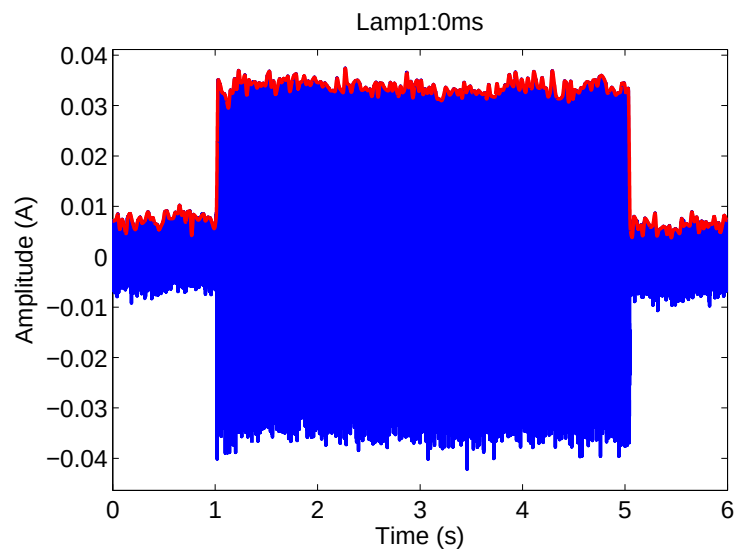


FIGURE 4.18 – Courant de mise en marche d'une lampe fluo-compacte avec zones de pré- et post-déclenchement.

Puisque $e^{\mathbf{p}^T \mathbf{t}} > 0$, $s(t)$ ne peut être égal à $s_s(t)$ que si $\mathbf{p}^T \mathbf{t} \rightarrow -\infty$, donc, pour estimer un signal sans régime transitoire, l'algorithme d'estimation essaye de vérifier cette dernière condition et génère des p_j de valeurs négatives très élevées qui n'ont pas nécessairement une signification pour nous.

Pour contourner ce problème, nous proposons de faire un « mapping » de la fonction représentant l'enveloppe $e(t)$ vers une autre fonction $e_s(t)$, qui est une version légèrement modifiée de $e(t)$ permettant de pallier ce problème. Cette nouvelle fonction est donnée par :

$$e_s(t) = A_0 e^{\mathbf{b}^T \mathbf{t}_s} + 1, \quad (4.9)$$

avec $\mathbf{b} = [b_1, \dots, b_n]^T$, $\mathbf{t}_s = [t, \dots, t^n]^T$, et $A_0 \geq 0$ ajoutant au modèle la flexibilité d'annuler le terme $e^{\mathbf{b}^T \mathbf{t}}$ afin de modéliser des signaux sans régime transitoire.

Ce mapping a les propriétés suivantes :

- * Si le transitoire existe, alors : $b_i = p_i, i = 1, \dots, n$ et $A_0 = e^{p_0}$
- * Si le transitoire est absent, alors : $b_i \neq p_i, i = 1, \dots, n$ et $A_0 = 0$

Ainsi, on obtient de nouveaux paramètres représentant l'enveloppe du signal. Notons, comme indiqué dans la section précédente, que pour la tâche de classification c'est le dernier coefficient b_n qui nous intéresse puisque sa valeur est indépendante de la référence du temps.

4.3.2 Nuages de points des paramètres estimés

Suite à la pré-analyse faite dans la sous-section précédente, on ne garde pour la classification que les paramètres $A_0, b_3, A_1, A_2, A_3, A_4$ et A_5 . Pour avoir une idée sur les relations qui existent entre les différents paramètres ainsi que pour développer une intuition et pouvoir derrière expliquer les résultats de classification, nous affichons dans cette sous-section des nuages de points pour certaines combinaisons de paramètres.

Pour la classification, identification ou clustering, nous allons faire trois tests à chaque fois : un test en n'utilisant que les paramètres de l'enveloppe A_0 et b_3 , un test en utilisant les paramètres A_i seuls ainsi qu'un troisième test en utilisant tous les paramètres.

Nous affichons d'abord le nuage de points des paramètres A_0, b_3 et A_1 en 3D (fi-

gure 4.20) avec des vues 2D correspondant aux trois combinaisons possibles des paramètres (A_0, b_3) , (A_0, A_1) et (b_3, A_1) et ceci pour tous les appareils de COOLL.

Ce que l'on remarque de suite est la distinction claire des lampes par rapport aux autres types d'appareils (vues (A_0, b_3) et (b_3, A_1)). Le paramètre qui permet cette distinction étant le b_3 . Les A_0 et A_1 permettent de retrouver les lampes sans régime transitoire (vue (A_0, A_1)) qui se trouvent proches de l'origine. À côté de ces lampes on trouve les ventilateurs, ce qui indique qu'ils ont le même comportement, c.-à-d. absence du régime transitoire (figure 4.19). Le paramètre A_1 permet de faire dégager les aspirateurs (vues (b_3, A_1) et (A_0, A_1)). Mis à part ces groupes qui se dégagent, les autres groupes se chevauchent, plus ou moins, les uns les autres.

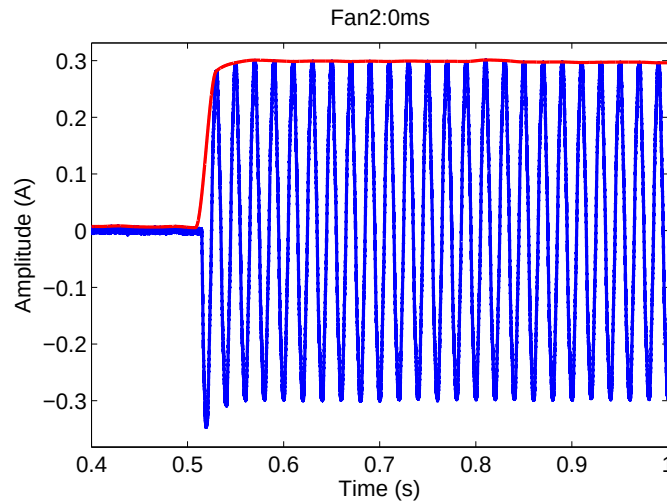
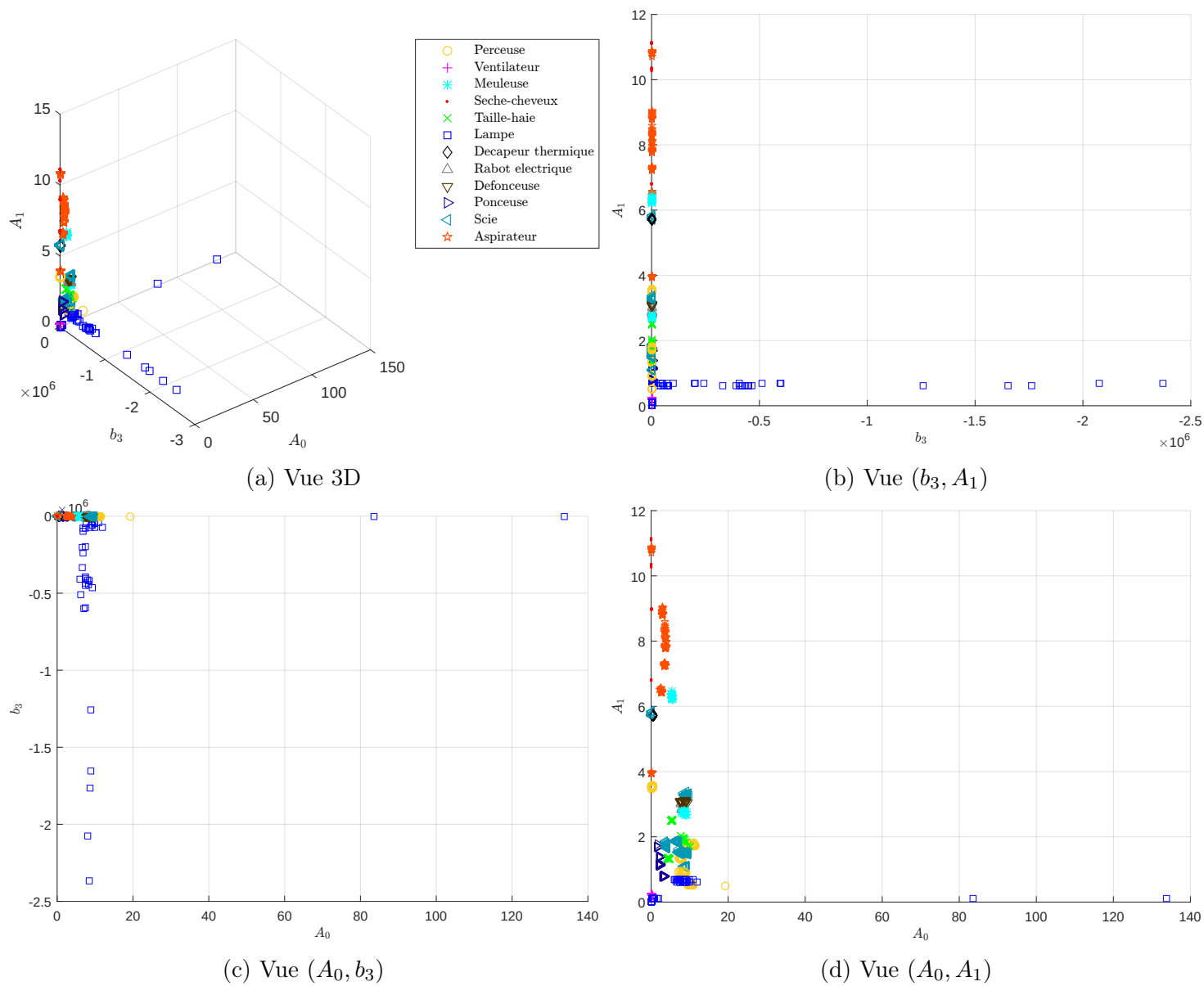
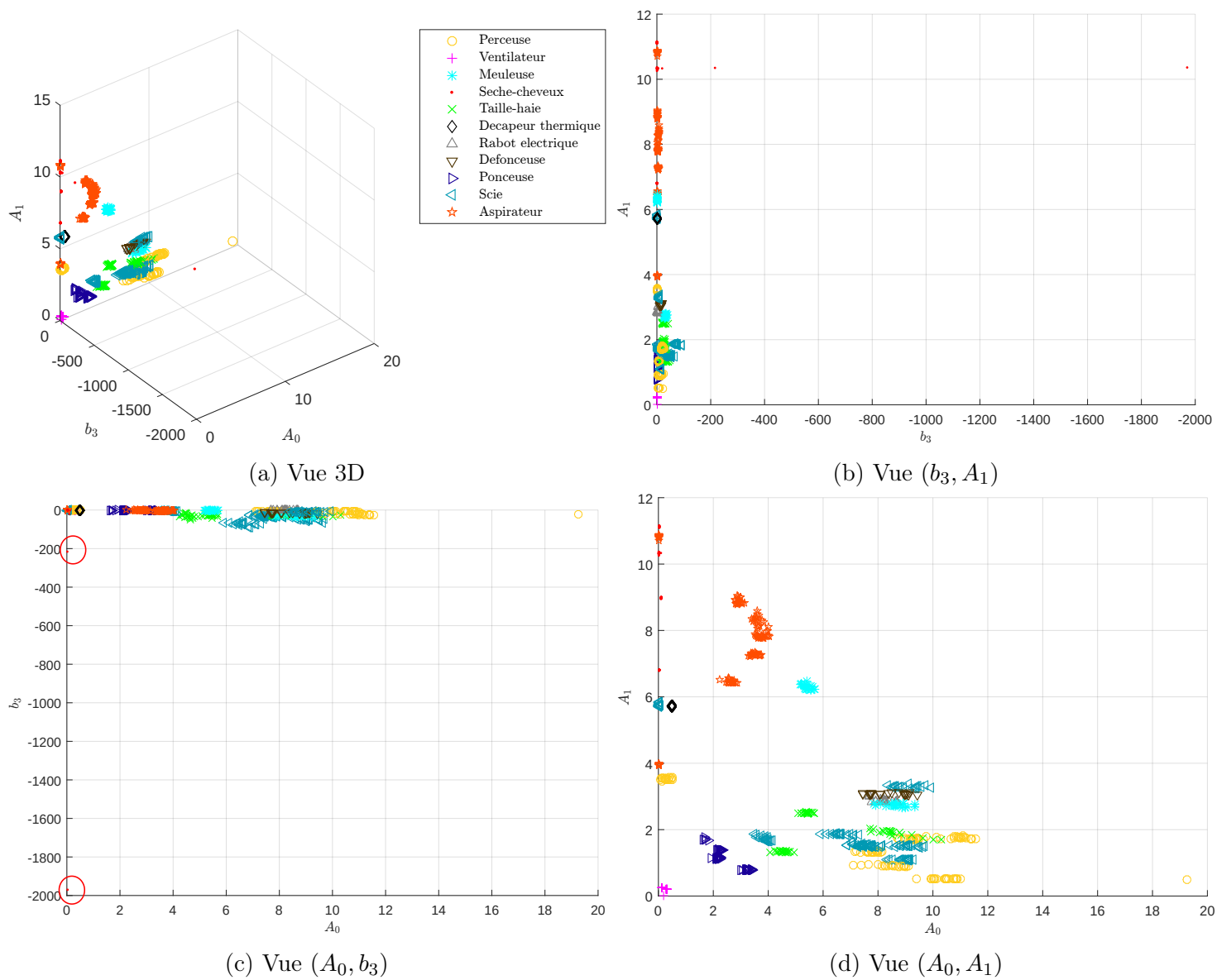


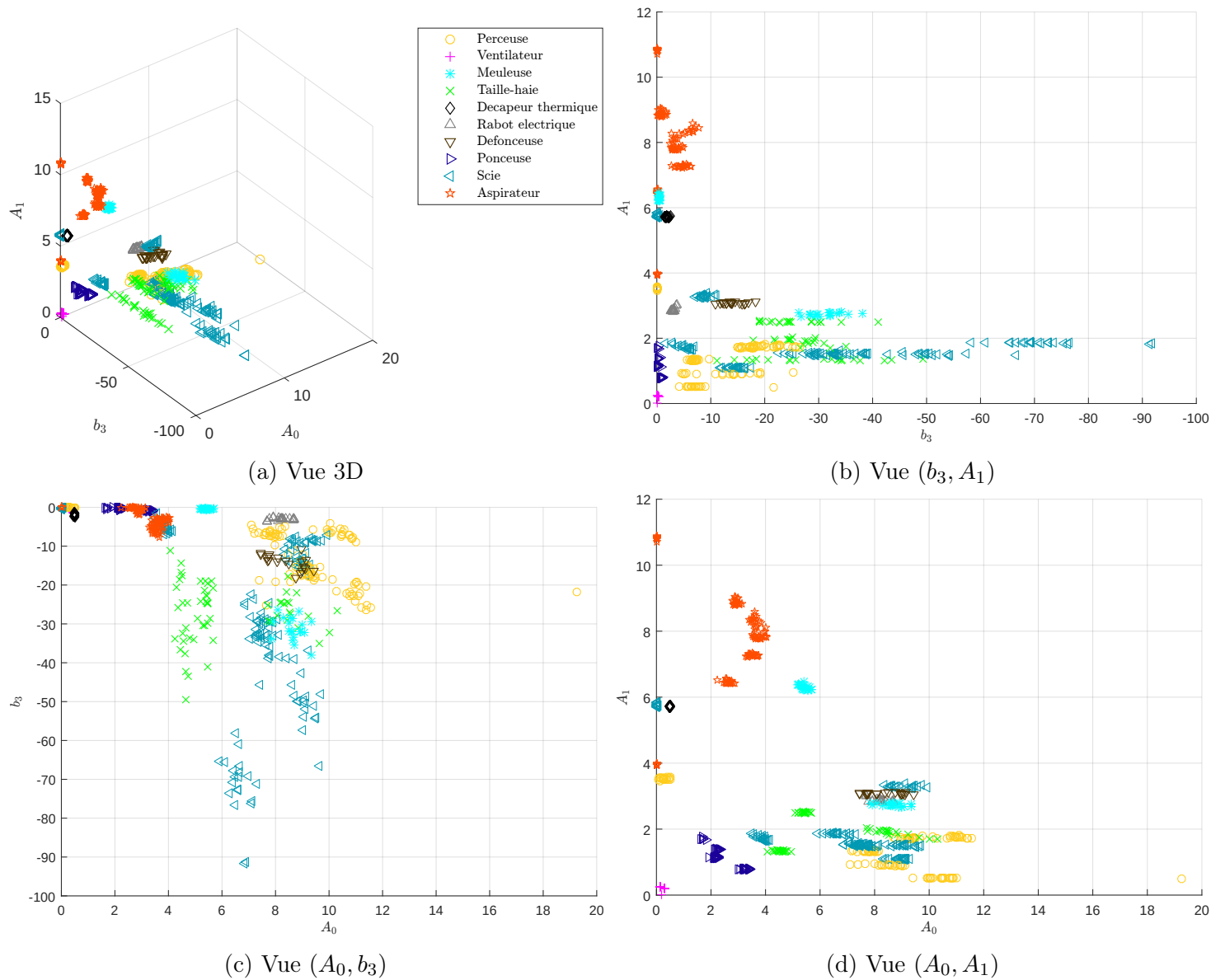
FIGURE 4.19 – Courant de mise en marche d'un ventilateur.

Pour voir plus clair, on propose d'enlever les lampes de ce nuage de points (figure 4.21). Mis à part deux sèche-cheveux singuliers que le b_3 arrive à discerner (vues (A_0, b_3) et (b_3, A_1)) ce paramètre ne permet pas plus de distinction pour le reste. A_0 et A_1 (vue (A_0, A_1)) certains groupes mais où les appareils ne sont pas bien distincts.

En enlevant encore les sèche-cheveux de ce nuage de points (figure 4.22), nous confirmons bien que ces paramètres A_0, b_3 et A_1 n'arrivent plus à distinguer plus d'appareils (mis à part la combinaison A_0, A_1 qui arrive à faire sortir quelques groupes d'appareils).

FIGURE 4.20 – Nuage de points des paramètres A_0 , b_3 et A_1 .

FIGURE 4.21 – Nuage de points des paramètres A_0 , b_3 et A_1 sans les lampes.

FIGURE 4.22 – Nuage de points des paramètres A_0 , b_3 et A_1 sans les lampes et sans sèche-cheveux.

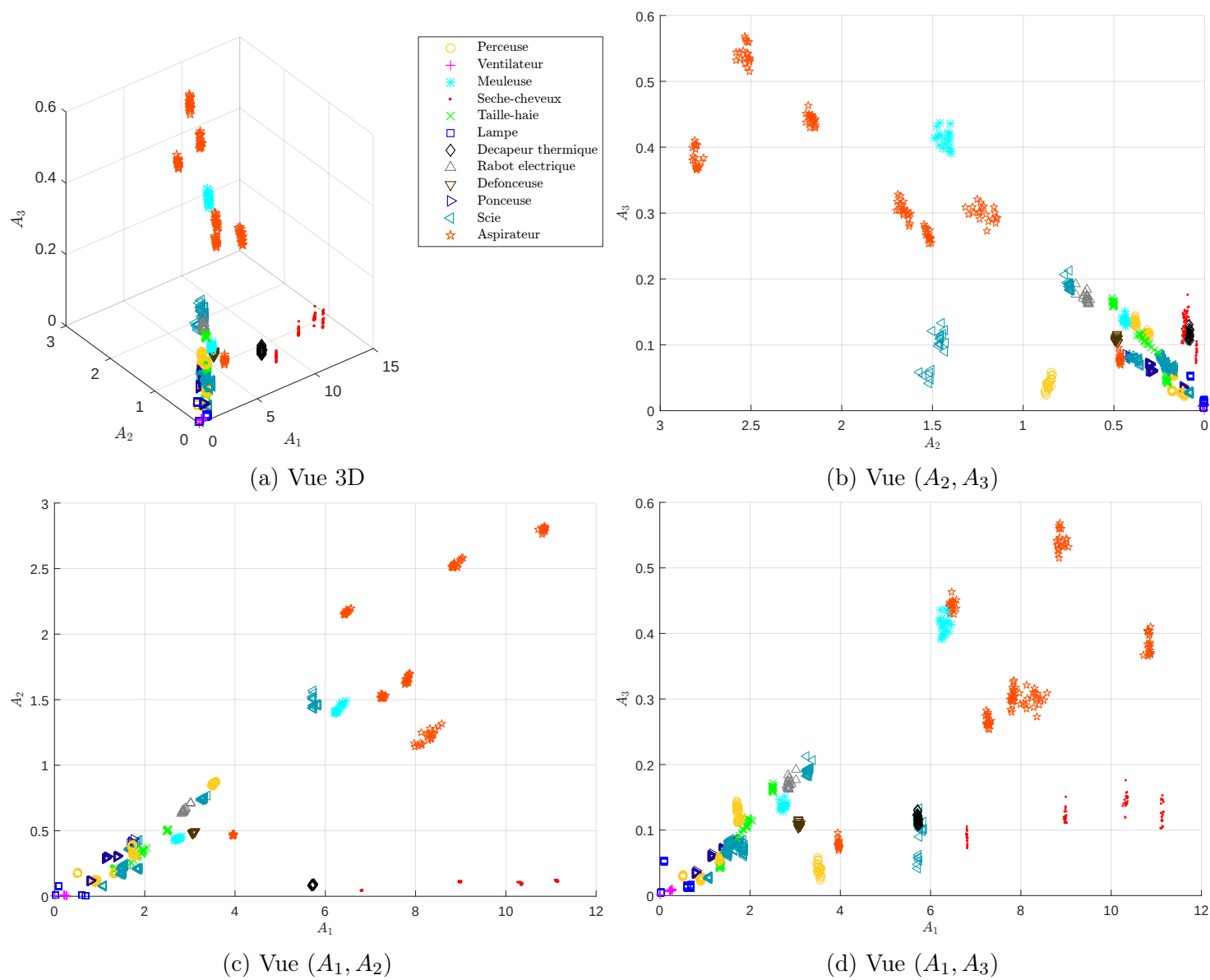
4.3. CLASSIFICATION DES APPAREILS

Le nuage de points des paramètres A_i (figure 4.23) permet clairement de retrouver de meilleures distinctions entre plusieurs ensembles de points que ce que permet d'avoir la combinaison des paramètres A_0, b_3 et A_1 . Il apparaît même que ces groupes sont liés chacun à un appareil bien précis. Nous retrouvons les 20 mesures avec différents délais de mise en marche. Nous arrivons à voir, par exemple, quatre groupes de sèche-cheveux bien distincts se trouvant en bas à droite sur la vue (A_1, A_3) .

Les observations faites ci-dessus indiquent certaines tendances. Les paramètres A_i devraient être plus adaptés à résoudre une tâche d'identification qu'à résoudre une tâche de clustering puisque ces paramètres arrivent à identifier des groupes distincts pour chaque appareil ce qui est l'objectif recherché dans une tâche d'identification.

En contrepartie, les paramètres liés à l'enveloppe, surtout le b_3 , devraient être plus adaptés à la tâche de clustering qu'à la tâche d'identification. D'après les observations, ce paramètre permet de séparer les lampes (à l'exception des lampes sans régime transitoire) de tous les autres types d'appareils. Ceci est exactement ce qui est attendu d'un algorithme de clustering : trouver des regroupements d'appareils ayant des caractéristiques communes.

Les deux sections suivantes (identification et clustering) vont permettre d'affirmer ou infirmer les observations faites dans cette sous-section.

FIGURE 4.23 – Nuage de points des paramètres A_1, A_2 et A_3 .

4.3.3 Identification

Pour l'identification, nous utilisons un algorithme supervisé basique qui est le k -NN (k Nearest Neighbors, k plus proches voisins). Cet algorithme permet de prédire la classe d'un nouveau point (dans l'espace des caractéristiques) en calculant les distances relatives entre ce dernier et les points qui l'entourent. La décision de la classe d'appartenance se fait, alors, avec un vote majoritaire entre les k plus proches voisins de la classe. Les paramètres à sélectionner sont le nombre k de voisins proches à considérer ainsi que la distance. Nous avons choisi pour nos tests la distance euclidienne et $k = 10$. Un k faible < 5 peut dégrader la robustesse du classifieur en augmentant le risque de classer par rapport à des points isolés (le cas extrême étant $k = 1$). Comme nous savons, dans notre cas, qu'un groupe bien formé devrait contenir 20 points (les 20 mesures du même appareil avec délai de mise en marche variable), nous comparons par rapport à la moitié du nombre de points a priori de la même classe, d'où le $k = 10$.

Comme indiqué précédemment, nous faisons trois tests : un test en utilisant les paramètres de l'enveloppe A_0 et b_3 , un test en utilisant les paramètres A_i de la somme de sinusoides et un troisième test en utilisant tous les paramètres.

Sachant que notre base de transitoire n'est pas grande pour permettre d'avoir plusieurs exemples et pour l'apprentissage et pour le test ainsi que pour fiabiliser le test du résultat d'identification sous ces conditions, nous faisons une validation croisée (K -fold cross-validation) avec $K = 10$. Ceci consiste à diviser la base en $K = 10$ groupes (avec des éléments aléatoirement sélectionnés) et de faire 10 itérations de test, où nous prenons pour chaque itération, un groupe pour le test et le reste pour l'apprentissage en faisant tourner le groupe de test en sélectionnant un groupe distinct parmi les dix groupes à chaque itération. Le résultat final étant la moyenne des résultats individuels.

La figure 4.24 donne la matrice de confusion pour l'identification en utilisant les paramètres A_0 et b_3 . Avec un Taux de Bonne Classification (TBC)(Correct Classification Rate-CCR) = 83.33%, l'identification fait sortir plusieurs exemples mal identifiés. Ceci rejoint les conclusions déduites des observations faites sur les nuages de points.

		Classes prédites												
		Perceuse	Ventilateur	Meuleuse	Seche-cheveux	Taille-haie	Lampe	Decapeur thermique	Rabot electrique	Defonceuse	Ponceuse	Scie	Aspirateur	
Vraies classes	Perceuse	98	4	0	0	1	0	0	2	0	0	15	0	81.67%
	Ventilateur	0	39	0	1	0	0	0	0	0	0	0	0	97.50%
	Meuleuse	0	0	26	0	3	0	0	0	0	0	11	0	65.00%
	Seche-cheveux	0	0	0	71	1	2	3	0	0	0	3	0	88.75%
	Taille-haie	4	0	5	0	34	0	0	0	2	0	15	0	56.67%
	Lampe	0	0	0	0	0	76	1	0	0	0	1	2	95.00%
	Decapeur thermique	0	0	0	0	0	0	20	0	0	0	0	0	100.00%
	Rabot electrique	0	0	0	0	0	0	0	20	0	0	0	0	100.00%
	Defonceuse	6	0	0	0	0	0	0	0	9	0	5	0	45.00%
	Ponceuse	0	0	0	0	0	0	0	0	0	58	0	2	96.67%
	Scie	9	0	1	14	5	2	0	0	3	1	115	10	71.88%
	Aspirateur	0	0	0	0	0	2	0	0	0	2	2	134	95.71%
		83.76%	90.70%	81.25%	82.56%	77.27%	92.68%	83.33%	90.91%	64.29%	95.08%	68.86%	90.54%	Tot. 83.33%

FIGURE 4.24 – Matrice de confusion pour l'identification en utilisant A_0 et b_3 . TBC = 83.33%.

La figure 4.25 donne la matrice de confusion pour l'identification en utilisant les paramètres A_i . On Note que l'utilisation de ces paramètres a donnée un très bon TBC de 98.57%. Encore une fois, comme observé sur les nuages de points, ceci est en concordance avec les conclusions tirées des nuages des poitns. La confusion entre taille-haies et perceuses, et comme pour les autres cas, est le résultat du chevauchement dans l'espace des caractéristiques entre les groupes de ces derniers (figure 4.23d autour du point $A_1 = 2, A_3 = 0.1$).

4.3. CLASSIFICATION DES APPAREILS

		Classes prédites												
		Perceuse	Ventilateur	Meuleuse	Seche-cheveux	Taille-haie	Lampe	Decapeur thermique	Rabot electrique	Defonceuse	Ponceuse	Scie	Aspirateur	
Vraies classes	Perceuse	120	0	0	0	0	0	0	0	0	0	0	100.00%	
	Ventilateur	0	39	0	0	0	1	0	0	0	0	0	97.50%	
	Meuleuse	0	0	40	0	0	0	0	0	0	0	0	100.00%	
	Seche-cheveux	0	0	0	80	0	0	0	0	0	0	0	100.00%	
	Taille-haie	6	0	0	0	54	0	0	0	0	0	0	90.00%	
	Lampe	0	0	0	0	0	80	0	0	0	0	0	100.00%	
	Decapeur thermique	0	0	0	0	0	0	20	0	0	0	0	100.00%	
	Rabot electrique	0	0	0	0	0	0	0	20	0	0	0	100.00%	
	Defonceuse	0	0	0	0	0	0	0	0	20	0	0	100.00%	
	Ponceuse	0	0	0	0	0	0	0	0	0	55	5	91.67%	
	Scie	0	0	0	0	0	0	0	0	0	0	160	100.00%	
	Aspirateur	0	0	0	0	0	0	0	0	0	0	0	140	100.00%
			95.24%	100.00%	100.00%	100.00%	100.00%	98.77%	100.00%	100.00%	100.00%	100.00%	96.97%	100.00%

FIGURE 4.25 – Matrice de confusion pour l'identification en utilisant les A_i . TBC = 98.57%.

La figure 4.26 quant à elle montre que l'utilisation de tous les paramètres pour l'identification améliore l'identification n'utilisant que les A_0 et b_3 mais dégrade le résultat d'identification n'utilisant que les A_i .

		Classes prédites												
		Perceuse	Ventilateur	Meuleuse	Seche-cheveux	Taille-haie	Lampe	Decapeur thermique	Rabot électrique	Defonceuse	Ponceuse	Scie	Aspirateur	
Vraies classes	Perceuse	104	0	0	0	2	0	0	0	0	0	14	0	86.67%
	Ventilateur	0	39	0	0	0	1	0	0	0	0	0	0	97.50%
	Meuleuse	0	0	31	0	5	0	0	0	0	0	4	0	77.50%
	Seche-cheveux	0	0	0	77	1	2	0	0	0	0	0	0	96.25%
	Taille-haie	3	0	5	0	34	0	0	0	2	0	16	0	56.67%
	Lampe	0	1	0	0	0	79	0	0	0	0	0	0	98.75%
	Decapeur thermique	0	0	0	0	0	0	20	0	0	0	0	0	100.00%
	Rabot électrique	0	0	0	0	0	0	0	20	0	0	0	0	100.00%
	Defonceuse	6	0	0	0	0	0	0	0	13	0	1	0	65.00%
	Ponceuse	0	0	0	0	0	0	0	0	0	60	0	0	100.00%
	Scie	3	0	0	0	5	0	0	0	0	1	151	0	94.38%
	Aspirateur	0	0	0	0	0	0	0	0	0	0	0	140	100.00%
			89.66%	97.50%	86.11%	100.00%	72.34%	96.34%	100.00%	100.00%	86.67%	98.36%	81.18%	100.00%

FIGURE 4.26 – Matrice de confusion pour l'identification en utilisant A_0 , b_3 et les A_i . TBC = 91.43%.

En conclusion, ces résultats d'identification confirment bien l'adéquation des paramètres A_i pour la tâche d'identification et l'inadéquation des paramètres A_0 et b_3 pour cette dernière.

4.3.4 Clustering

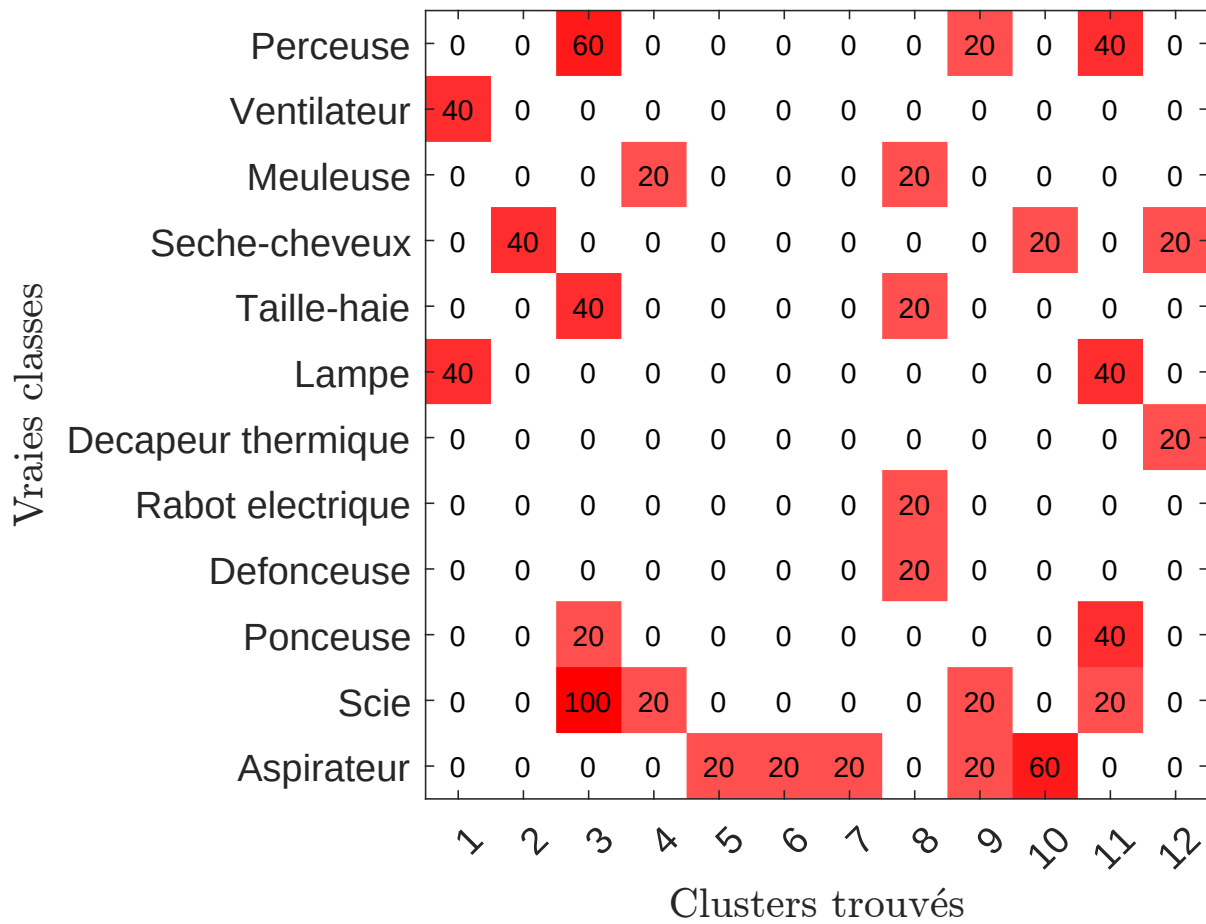
Tout comme pour l'identification, nous utilisons pour le clustering un algorithme de classification basique non supervisé, un des plus connus, le k-means. Cet algorithme appartient à la classe d'algorithmes non supervisés dits de partitionnement. Il nécessite de spécifier *a priori* le nombre de clusters qu'on cherche à former. Nous avons choisi $k = 12$ clusters par rapport au nombre de types d'appareil que nous avons dans COOLL.

La figure 4.27 donne le résultat du clustering en utilisant A_0 et b_3 . L'algorithme a essentiellement fait sortir un cluster qui englobe tout les appareils à l'exception des lampes. Ceci est conforme avec les observations faites sur les nuages de points. Notons que les lampes ont été distribuées sur les autres clusters. Ceci est dû aux distances relatives entre ces lampes dans l'espace des caractéristiques (par exemple sur la figure 4.20b les cinq lampes se trouvant à droite et qui sont plutôt éloignées du reste des lampes se trouvant à gauche). Un certain nombre de lampes (46) a été regroupé dans le cluster englobant tous les appareils. Ces lampes sont les lampes sans transitoire et pour lesquelles les paramètres A_0 et b_3 ne sont pas suffisant à eux seuls pour les distinguer des autres.

Vraies classes		Clusters trouvés											
		1	2	3	4	5	6	7	8	9	10	11	12
	Perceuse	120	0	0	0	0	0	0	0	0	0	0	0
	Ventilateur	40	0	0	0	0	0	0	0	0	0	0	0
	Meuleuse	40	0	0	0	0	0	0	0	0	0	0	0
	Seche-cheveux	80	0	0	0	0	0	0	0	0	0	0	0
	Taille-haie	60	0	0	0	0	0	0	0	0	0	0	0
	Lampe	46	1	13	4	1	2	1	2	1	1	3	5
	Decapeur thermique	20	0	0	0	0	0	0	0	0	0	0	0
	Rabot electrique	20	0	0	0	0	0	0	0	0	0	0	0
	Defonceuse	20	0	0	0	0	0	0	0	0	0	0	0
	Ponceuse	60	0	0	0	0	0	0	0	0	0	0	0
	Scie	160	0	0	0	0	0	0	0	0	0	0	0
	Aspirateur	140	0	0	0	0	0	0	0	0	0	0	0

FIGURE 4.27 – Matrice de confusion pour le clustering en utilisant A_0 et b_3 .

La figure 4.28 donne le clustering en utilisant les A_i . On voit clairement que ces derniers ne sont pas du tout adaptés pour faire le clustering vu qu'ils génèrent des clusters désordonnés sans nécessairement regrouper les appareils de même type dans les même clusters.

FIGURE 4.28 – Matrice de confusion pour le clustering en utilisant les A_i .

La figure 4.29 est intéressante car elle montre que l'utilisation de tous les paramètres ne dégrade pas le clustering des A_0, b_3 et améliore le clustering des A_i contrairement à ce que nous avons remarqué pour l'identification.

Vraies classes	Perceuse	120	0	0	0	0	0	0	0	0	0	0	0
	Ventilateur	40	0	0	0	0	0	0	0	0	0	0	0
	Meuleuse	40	0	0	0	0	0	0	0	0	0	0	0
	Seche-cheveux	80	0	0	0	0	0	0	0	0	0	0	0
	Taille-haie	60	0	0	0	0	0	0	0	0	0	0	0
	Lampe	46	6	1	1	1	2	2	1	6	4	7	3
	Decapteur thermique	20	0	0	0	0	0	0	0	0	0	0	0
	Rabot electrique	20	0	0	0	0	0	0	0	0	0	0	0
	Defonceuse	20	0	0	0	0	0	0	0	0	0	0	0
	Ponceuse	60	0	0	0	0	0	0	0	0	0	0	0
	Scie	160	0	0	0	0	0	0	0	0	0	0	0
	Aspirateur	140	0	0	0	0	0	0	0	0	0	0	0
		1	2	3	4	5	6	7	8	9	10	11	12
		Clusters trouvés											

FIGURE 4.29 – Matrice de confusion pour le clustering en utilisant A_0 , b_3 et les A_i .

4.3.5 Comparaison des résultats de classification vis-à-vis de la preuve du concept

Les résultats de la sous-section 4.1.2 montrent que le régime transitoire permet une meilleure identification que le régime établi. En contrepartie, les résultats de la sous-section 4.3.3 indiquent que les paramètres de notre modèle A_0 et b_3 liés à l'enveloppe (donc le transitoire) sont moins adaptés pour l'identification que les paramètres A_i représentant le régime établi. Ces deux résultats semblent, à première vue, contradictoires. En effet, les caractéristiques discriminantes utilisées dans les deux démarches ne sont pas les mêmes. Pour la preuve du concept, les caractéristiques discriminantes sont des coefficients des Séries de Fourier Discrètes à Court-Terme (SFDCT). Ces coefficients, quand extraits de la partie transitoire du courant, permettent de suivre l'évolution des amplitudes des harmoniques. La preuve du concept indique que cette évolution est importante

pour l'identification.

Un modèle stationnaire supposant la non variabilité des amplitudes des harmoniques perd, de fait, une quantité importante d'information. Cela justifie la contradiction notée car les paramètres de notre modèle liés au transitoire ne donnent qu'une variation globale de l'amplitude du courant et omettent l'évolution des amplitudes des harmoniques.

Cela indique qu'une meilleure représentativité des données réelles devrait tenir compte de la non stationnarité de $s_s(t)$ observée dans les signaux réels. Notons que cette constatation n'enlève rien quant à l'utilité des paramètres de l'enveloppe pour le clustering.

Conclusion et perspectives

Conclusion

Pour répondre aux besoins d'efficacité énergétique en rapport avec la consommation électrique, il semble incontournable de disposer d'une information de consommation détaillée. Idéalement, on cherche à savoir sur le réseau électrique qui consomme ? quand ? et combien ?

Dans ce travail de thèse, nous avons proposé quelques avancées sur chacune des questions. Nous avons abordé la troisième question par une étude comparative de l'estimation de la consommation électrique donnée par quelques calculateurs d'énergie grand public.

Pour la deuxième question, à savoir connaître les instants de début et de fin de consommation, nous avons proposé un détecteur d'évènements haute précision. Ce dernier est basé sur l'utilisation de l'enveloppe du courant et est simple et rapide.

Les avancées liées à la première question, à savoir l'identification de l'appareil en marche, sont :

- une proposition d'un nouveau modèle pour le courant transitoire de mise en marche ;
- une proposition d'un algorithme d'estimation de paramètres pour ce modèle ;
- une preuve du concept montrant l'utilité du régime transitoire par rapport au régime établi ;
- une création d'un système de mesure de signaux électriques contrôlés en mise en marche et mise à l'arrêt ainsi que la création d'une base de données de courants de mise en marche.

En dépit de résultats montrant que le modèle de courant de mise en marche d'appareils

mériterait d'être affiné, ce travail apporte une caractérisation de ce signal au travers de paramètres informatifs pour un système plus global de classification d'appareils. Un tel système doit bien entendu prendre en compte d'autres paramètres non considérés dans notre approche tels que les puissance active, réactive, leurs variations et tous les paramètres liés au fonctionnement de certains appareils qui peuvent afficher un déterminisme lié, par exemple, au séquençage des cycles pour les appareils de lavage.

Les approches de type Machine Learning (ML) sont appropriées pour appréhender ces fonctionnements à caractère régulier ainsi que les comportements d'usage de ces appareils, par exemple utilisation du grille-pain le matin. En revanche, les appareils nomades sur lesquels nous nous sommes focalisés ont peu de régularité dans leur usage, ce qui met partiellement en défaut ces approches. Toute nouvelle information susceptible de discriminer ces appareils est importante dans un système global d'identification et de suivi de consommation des appareils. Ce travail de thèse y contribue.

Perspectives

Les perspectives peuvent se décliner selon deux axes :

- une amélioration des solutions proposées sur les problèmes traités ;
- une présentation de quelques points non abordés dans le cadre de ce travail de thèse.

Amélioration des solutions proposées :

1. *l'adaptation automatique du seuil au RSB pour le détecteur d'évènements afin d'alléger la sensibilité aux bruits de ce dernier* : le détecteur étant dépendant de la valeur du seuil, constante et fixée *a priori*, il ne s'adapte pas aux variations du RSB des signaux courant traités. Par exemple, si le seuil est fixé à 0,1 A, le détecteur ne va pas pouvoir détecter la mise en marche d'un appareil de faible puissance ayant une amplitude variant au maximum de 0,08 A. Une solution est alors d'estimer le RSB du signal traité et d'utiliser par la suite cette estimée pour définir un seuil dépendant proportionnellement de cette dernière ;
2. *augmenter la base de transitoires avec des exemples d'appareils ayant une mise en marche non nécessairement contrôlables ainsi qu'avec des scénarios de me-*

sure de plusieurs appareils fonctionnant simultanément : notre base de données étant le résultat d'une première campagne de mesure, d'autres campagnes de mesure permettront de la compléter. Les campagnes de mesure futures pourront envisager la mesure d'appareils dont l'instant de mise en marche n'est pas nécessairement bien maîtrisé. C'est le cas des appareils qui, même en gardant actionné le bouton de mise en marche, ne démarreront pas quand mis sous tension.

Un des grands intérêts du système de mesure est qu'il permet de créer des scénarios contrôlés de plusieurs appareils en fonctionnement. De tels scénarios peuvent être utilisés pour évaluer les performances des détecteurs d'évènements ainsi que celles d'un système de désagrégation. De ce fait, une base contenant de tels scénarios a tout son intérêt pour le développement du domaine de désagrégation d'énergie électrique. Néanmoins, le challenge de définir des scénarios représentant des cas réalistes de consommation reste à surmonter ;

3. *aborder la problématique de sélection de modèle pour bien définir les paramètres de ce dernier, à savoir, l'ordre du polynôme n et le nombre d'harmoniques d , supposés connus dans le cadre de cette thèse* : ces paramètres ayant été fixés empiriquement pour le travail présenté dans cette thèse, il serait intéressant de trouver un moyen automatique permettant de les définir. Différentes techniques existent permettant d'aborder une telle problématique, on peut citer, entre autres, la validation, la validation croisée, le critère d'information d'Akaike (Akaike information criterion–AIC) et le critère d'information bayésien (Bayesian Information Criterion–BIC) (Bishop, 2006) ;
4. *permettre au modèle de suivre la non stationnarité des courants transitoires de certains appareils en supposant des paramètres d'amplitude $A_i(t)$ non stationnaires* : étant donné que les signaux réels peuvent afficher beaucoup de variabilité, de part leur nature réelle variable, des simplifications supposant la stationnarité de ces derniers peuvent ne pas être valables pour certains types d'appareils ainsi que sous certaines conditions pratiques. Une amélioration possible du modèle visant à prendre en compte cette variabilité serait d'intégrer ce caractère en supposant des paramètres d'amplitude $A_i(t)$ non stationnaires.

Quelques points non abordés :

1. *étudier la possibilité de généralisation du modèle pour le cas multi-appareil* : l'étude présentée dans ce manuscrit suppose l'existence d'un seul appareil en fonctionnement sur le réseau électrique domestique. Pratiquement, tomber sur une telle situation n'est pas complètement exclu mais est très peu réaliste. Dans un cadre plus réaliste, on trouve une multitude d'appareils sur le réseau fonctionnant simultanément. Pour notre modèle, nous avons supposé certaines hypothèses permettant l'applicabilité de notre modèle au cas multi-appareil (sous-section 3.2.2). Néanmoins, pour la généralisation de ce dernier, il est nécessaire de vérifier la validité de ces hypothèses pour trouver dans quels cas peuvent-elles être facilement vérifiées pour se ramener au modèle mono-appareil. Sinon, il faudrait trouver un moyen pour généraliser le modèle mono-appareil à un cas multi-appareil ;
2. *faire une étude plus poussée et aboutie de la classification d'appareils électriques et des problématiques sous-jacentes* : le travail de cette thèse s'étant plus focalisé sur la modélisation des courants transitoires de mise en marche, l'étude faite pour la partie classification reste très primitive. De ce fait, il serait intéressant de pousser l'étude dans un travail futur vers la problématique de classification d'appareils électriques. D'autres problématiques sous-jacentes à la classification peuvent aussi être adressées. Nous pensons notamment aux problématiques d'extraction, construction et sélection de caractéristiques (Liu and Motoda, 1998; Motoda and Liu, 2002). Dans ce cadre, il serait particulièrement intéressant de pouvoir trouver d'autres caractéristiques à extraire des signaux transitoires de courant qui soient discriminantes pour la tâche de classification.

Annexe A

Algorithme Modified Zero-Crossing (MZC) pour l'estimation et le suivi de la fréquence fondamentale

La détection de passages par zéros (zero-crossings) ZC permet d'estimer ainsi que de suivre les variations de la fréquence fondamentale du réseau électrique. Un algorithme basé sur un tel principe détecte les instants t_i où la tension réseau passe par zéro, soit du positif au négatif ou bien du négatif au positif suivant une convention prédéfinie. En utilisant ces instants t_i il calcule la fréquence fondamentale $f_0(t)$ (variable en fonction du temps) pour chaque t_i comme suit : $f_0(t_i) = \frac{1}{t_i - t_{i-2}}$ (sachant qu'il existe deux passages par zéro dans une période).

Un des plus gros problèmes de cet algorithme est sa sensibilité au bruit (figure [A.1](#)). Pour alléger ce problème, les auteurs de ([Begovic et al., 1993](#)) ont proposé une version modifiée de l'algorithme, qu'ils ont appelée « Modified Zero-Crossing (MZC) technique ».

La modification consiste à approximer la tension autour du point de passage par zéro par une fonction polynômiale. Cela a pour effet de lisser le bruit.

Le signal de tension $v(t)$ peut s'écrire sous sa forme échantillonnée $v[n]$, en considérons

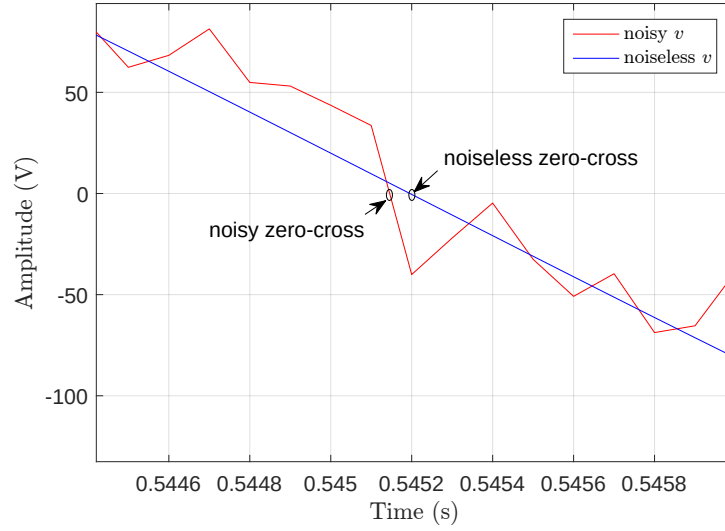


FIGURE A.1 – Décalage du point de passage par zéro suite à l’ajout de bruit sur un signal de tension.

d harmoniques, comme :

$$v[n] = v(nT_s) = \sum_{k=1}^d V_k \cos(k\omega_0 n + \phi_k), \quad n = 1, \dots, N, \quad (\text{A.1})$$

$\omega_0 = \frac{2\pi f_0}{F_s}$ (en rd) étant la fréquence fondamentale normalisée par la fréquence d’échantillonnage $F_s = \frac{1}{T_s}$ et ϕ_k les phases initiales des différentes harmoniques du signal de tension. N est le nombre total de points considérés.

D’abord, on définit une fenêtre de mesure de taille M points (nombre pair)¹ autour du point de passage par zéro i détecté par ZC :

$$\mathbf{v}_i = [v[i - M/2 + 1], \dots, v[i + M/2]]^T. \quad (\text{A.2})$$

Ensuite, on va fitter à cette fenêtre de valeurs, en utilisant les moindres carrés linéaires (MCL), un polynôme $p_l(t)$ de degré l tel que² :

$$p_l(t) = a_0 + a_1 t + \dots + a_l t^l = \sum_{j=0}^l a_j t^j. \quad (\text{A.3})$$

1. Pour nos simulations nous avons pris $M = 10$. Nous supposons aussi l’existence d’un seul passage par zéro dans chaque fenêtre.

2. Nous avons choisi pour nos simulations $l = 1$ donnant un fit acceptable pour une fenêtre $\mathbf{v}_i$ de $M = 10$ points.

La solution est obtenue en résolvant un système surdéterminé ($M > l + 1$) d'équations linéaires :

$$\mathbf{v}_i = \begin{bmatrix} v \left[i - M/2 + 1 \right] \\ v \left[i - M/2 + 2 \right] \\ \vdots \\ v \left[i + M/2 \right] \end{bmatrix} = \begin{bmatrix} 1 & \left(i - M/2 + 1 \right) T_s & \cdots & \left(i - M/2 + 1 \right)^l T_s^l \\ 1 & \left(i - M/2 + 2 \right) T_s & \cdots & \left(i - M/2 + 2 \right)^l T_s^l \\ 1 & \vdots & \ddots & \vdots \\ 1 & \left(i + M/2 \right) T_s & \cdots & \left(i + M/2 \right)^l T_s^l \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_l \end{bmatrix}, \quad (\text{A.4})$$

pouvant s'écrire :

$$\mathbf{v}_i = \mathbf{K} \mathbf{a} \quad (\text{A.5})$$

La solution des MCL est alors :

$$\hat{\mathbf{a}} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T \mathbf{v}_i. \quad (\text{A.6})$$

Ayant le vecteur de coefficients $\hat{\mathbf{a}}$, il suffit alors pour trouver l'instant estimé de passage par zéro $\hat{t}_i$ de retrouver les racines du polynôme :

$$p_l(\hat{t}_i) = \hat{a}_0 + \hat{a}_1 \hat{t}_i + \cdots + \hat{a}_l \hat{t}_i^l = 0. \quad (\text{A.7})$$

Enfin, en considérant différentes fenêtres glissantes $\mathbf{v}_i$ permettant chacune d'estimer le $\hat{t}_i$ correspondant, on calcule une estimée de la fréquence fondamentale aux instants t_i comme suit :

$$\hat{f}_0(\hat{t}_i) = \frac{1}{\hat{t}_i - \hat{t}_{i-2}}. \quad (\text{A.8})$$

L'estimée de la fréquence fondamentale $\hat{f}_0(t)$, pour tous les instants t , s'obtient en faisant une interpolation (cubique) entre les valeurs $\hat{f}_0(\hat{t}_i)$.

Notons que pour notre estimation nous ajoutant à l'algorithme MZC tel que proposé par (Begovic et al., 1993) un étage de pré-traitement consistant à filtrer passe-bas le signal de tension. Cela permettra, de plus d'une réduction supplémentaire du bruit, de réduire l'effet négatif des harmoniques sur l'estimation.

Bibliographie

- Abraha Girmay, A. and Camarda, C. (2016). Simple event detection and disaggregation approach for residential energy estimation. In *Workshop on Non-Intrusive Load Monitoring (NILM), 2016 Proceedings of the 3rd International*. (Cit  en page 22).
- Akbar, M. and Khan, D. Z. A. (2007). Modified nonintrusive appliance load monitoring for nonlinear devices. In *Multitopic Conference, 2007. INMIC 2007. IEEE International*, pages 1–5. IEEE. (Cit  en page 24).
- Amirach, N. (2015). *D tection d’ v nements simples   partir de mesures sur courant alternatif*. Th se de doctorat, Universit  de Toulon. (Cit  en page 20).
- Anderson, K., Ocneanu, A., Benitez, D., Carlson, D., Rowe, A., and Berges, M. (2012a). BLUED : a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research. In *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, Beijing, China. (Cit  en page 108).
- Anderson, K. D., Berg s, M. E., Ocneanu, A., Benitez, D., and Moura, J. M. (2012b). Event detection for non intrusive load monitoring. In *IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society*, pages 3312–3317. IEEE. (Cit  en page 21).
- Arberet, S. and Hutter, A. (2014). Non-intrusive load curve disaggregation using sparse decomposition with a translation-invariant boxcar dictionary. In *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES*, pages 1–6. IEEE. (Cit  en page 24).
- Baranski, M. and Voss, J. (2003). Nonintrusive appliance load monitoring based on an optical sensor. In *Power Tech Conference Proceedings, 2003 IEEE Bologna*, volume 4, pages 8–pp. IEEE. (Cit  en pages 19 et 24).
- Baranski, M. and Voss, J. (2004a). Detecting patterns of appliances from total load data using a dynamic programming approach. In *Data Mining, 2004. ICDM’04. Fourth IEEE International Conference on*, pages 327–330. IEEE. (Cit  en page 24).
- Baranski, M. and Voss, J. (2004b). Genetic algorithm for pattern detection in nialm systems. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 4, pages 3462–3468. IEEE. (Cit  en page 24).
- Barker, S., Mishra, A., Irwin, D., Cecchet, E., Shenoy, P., and Albrecht, J. (2012). Smart* : An open data set and tools for enabling research in sustainable homes. *SustKDD, August*, 111 :112. (Cit  en page 108).

- Barsim, K. S., Streubel, R., and Yang, B. (2014). An approach for unsupervised non-intrusive load monitoring of residential appliances. In *Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring*. (Cité en page 24).
- Barsim, K. S. and Yang, B. (2015). Toward a semi-supervised non-intrusive load monitoring system for event-based energy disaggregation. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 58–62. IEEE. (Cité en page 21).
- Batra, N., Gulati, M., Singh, A., and Srivastava, M. B. (2013). It’s different : Insights into home energy consumption in india. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, BuildSys’13, pages 3 :1–3 :8, New York, NY, USA. ACM. (Cité en page 108).
- Beckel, C., Kleiminger, W., Cicchetti, R., Staake, T., and Santini, S. (2014). The eco data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 80–89. ACM. (Cité en page 108).
- Beckel, C., Kleiminger, W., Staake, T., and Santini, S. (2012). Improving device-level electricity consumption breakdowns in private households using on/off events. *ACM SIGBED Review*, 9(3) :32–38. (Cité en page 96).
- Begovic, M. M., Djuric, P. M., Dunlap, S., and Phadke, A. G. (1993). Frequency tracking in power networks in the presence of harmonics. *IEEE Transactions on Power Delivery*, 8(2) :480–486. (Cité en pages 78, 145 et 147).
- Berges, M., Goldman, E., Matthews, H. S., and Soibelman, L. (2009). Learning systems for electric consumption of buildings. In *ASCE International Workshop on Computing in Civil Engineering, Austin, Texas*. (Cité en page 24).
- Besson, O. and Abramovich, Y. I. (2013). On the fisher information matrix for multivariate elliptically contoured distributions. *IEEE Signal Processing Letters*, 20(11) :1130–1133. (Cité en page 51).
- Bier, T. (2014). *Disaggregation of Electrical Appliances using Non-Intrusive Load Monitoring*. Thèse de doctorat, Université de Haute-Alsace. (Cité en page 20).
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc. (Cité en page 143).
- Bonfigli, R., Squartini, S., Fagiani, M., and Piazza, F. (2015). Unsupervised algorithms for non-intrusive load monitoring : An up-to-date overview. In *Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on*, pages 1175–1180. (Cité en page 21).
- Burbano Acuña, M. D. (2015). Intrusive and non-intrusive load monitoring (a survey). *Latin-American Journal of Computing, Systems Engineering, National Polytechnic School, Ecuador*, 2(1). (Cité en page 18).
- Carrie Armel, K., Gupta, A., Shrimali, G., and Albert, A. (2013). Is disaggregation the holy grail of energy efficiency ? the case of electricity. *Energy Policy*, 52 :213–234. (Cité en pages 19 et 26).

- CENELEC (2010). Voltage characteristics of electricity supplied by public electricity networks. *European Standard EN 50160*. (Cité en page 77).
- Centre Scientifique et Technique du Batiment (CSTB) (2013). Fiche d'application : Systèmes de mesure ou d'estimation des consommations en logement. Technical report, ADEME, Ministère de l'égalité des territoires et du logement, Ministère de l'écologie, du développement durable et de l'énergie. (Cité en page 13).
- Chan, W., So, A. T., and Lai, L. (2000). Harmonics load signature recognition by wavelets transforms. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2000. Proceedings. DRPT 2000. International Conference on*, pages 666–671. IEEE. (Cité en pages 19 et 24).
- Chang, H.-H., Lin, C.-L., and Yang, H.-T. (2008). Load recognition for different loads with the same real power and reactive power in a non-intrusive load-monitoring system. In *Computer Supported Cooperative Work in Design, 2008. CSCWD 2008. 12th International Conference on*, pages 1122–1127. IEEE. (Cité en page 19).
- Chroma 63800 (2016). Programmable ac and dc electronic load model 63800 series. http://www.chromaate.com/product/63800_series_programmable_ac_and_dc_electronic_load.htm [Visité le : 03/10/2016]. (Cité en page 32).
- Clark, M. S. (2015). Improving the feasibility of energy disaggregation in very high-and low-rate sampling scenarios. Master's thesis, University Of British Columbia. (Cité en pages xi, 21 et 24).
- Cole, A. I. and Albicki, A. (1998a). Algorithm for nonintrusive identification of residential appliances. In *Circuits and Systems, 1998. ISCAS'98. Proceedings of the 1998 IEEE International Symposium on*, volume 3, pages 338–341. IEEE. (Cité en pages 19 et 24).
- Cole, A. I. and Albicki, A. (1998b). Data extraction for effective non-intrusive identification of residential power loads. In *Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings. IEEE*, volume 2, pages 812–815. IEEE. (Cité en page 24).
- Coleman, T. F. and Li, Y. (1994). On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds. *Mathematical Programming*, 67(1) :189–224. (Cité en page 48).
- Coleman, T. F. and Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2) :418–445. (Cité en page 48).
- CompuScope 12400 (2008). Gage CompuScope 12400 – 12 bit, 400MS/s digitiser – PCI bus. www.egmont.com.pl/gage/katalog/compuScope_12400.pdf [Visité le : 03/10/2016]. (Cité en page 28).
- Cox, R., Leeb, S. B., Shaw, S. R., and Norford, L. K. (2006). Transient event detection for nonintrusive load monitoring and demand side management using voltage distortion. In *Applied Power Electronics Conference and Exposition, 2006. APEC '06. Twenty-First Annual IEEE*, pages 7 pp.–. (Cité en page 22).

- Curtis, R., Copeland, D., and Yates, D. J. (2014). Method and system for disaggregating heating and cooling energy use from other building energy use. US Patent 8,660,813. (Cité en page 24).
- Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, 486 :2006. (Cité en page 26).
- Darby, S. (2010). Smart metering : what potential for householder engagement ? *Building Research & Information*, 38(5) :442–457. (Cité en page 26).
- De Baets, L., Ruyssinck, J., Deschrijver, D., and Dhaene, T. (2016). Event detection in nilm using cepstrum smoothing. In *3rd International Workshop on Non-Intrusive Load Monitoring*, pages 1–4. (Cité en page 22).
- Dong, M., Meira, P. C., Xu, W., and Chung, C. (2013). Non-intrusive signature extraction for major residential loads. *IEEE Transactions on Smart Grid*, 4(3) :1421–1430. (Cité en page 24).
- Drenker, S. and Kader, A. (1999). Nonintrusive monitoring of electric loads. *Computer Applications in Power, IEEE*, 12(4) :47–51. (Cité en pages 19 et 24).
- Du, L., Yang, Y., He, D., Harley, R. G., and Habetler, T. G. (2015). Feature extraction for load identification using long-term operating waveforms. *IEEE Transactions on Smart Grid*, 6(2) :819–826. (Cité en page 24).
- Du, Y., Du, L., Lu, B., Harley, R., and Habetler, T. (2010). A review of identification and monitoring methods for electric loads in commercial and residential buildings. In *Energy Conversion Congress and Exposition (ECCE), 2010 IEEE*, pages 4527–4533. IEEE. (Cité en page 19).
- Egarter, D., Bhuvana, V. P., and Elmenreich, W. (2015). Paldi : Online load disaggregation via particle filtering. *Instrumentation and Measurement, IEEE Transactions on*, 64(2) :467–477. (Cité en page 24).
- Ehrhardt-Martinez, K., Donnelly, K. A., Laitner, S., et al. (2010). Advanced metering initiatives and residential feedback programs : a meta-review for household electricity-saving opportunities. American Council for an Energy-Efficient Economy Washington, DC. (Cité en page 18).
- El Guedri, M. (2009). *Caractérisation aveugle de la courbe de charge électrique : Détection, classification et estimation des usages dans les secteurs résidentiel et tertiaire*. Thèse de doctorat, Université Paris Sud-Paris XI. (Cité en page 20).
- Elhamifar, E. and Sastry, S. (2015). Energy disaggregation via learning powerlets and sparse coding. In *AAAI*, pages 629–635. (Cité en page 24).
- Farinaccio, L. and Zmeureanu, R. (1999). Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, 30(3) :245–259. (Cité en page 24).
- Feinberg, E. A. and Genethliou, D. (2005). Load forecasting. In *Applied math. for re-*

- structured electric power systems*, pages 269–285. Springer. (Cité en page 26).
- Figueiredo, M., De Almeida, A., and Ribeiro, B. (2012). Home electrical signal disaggregation for non-intrusive load monitoring (nilm) systems. *Neurocomputing*, 96 :66–73. (Cité en page 24).
- Fischer, C. (2008). Feedback on household electricity consumption : a tool for saving energy? *Energy efficiency*, 1(1) :79–104. (Cité en page 26).
- Friedlander, B. and Francos, J. M. (1995). Estimation of amplitude and phase parameters of multicomponent signals. *IEEE Transactions on Signal Processing*, 43(4) :917–926. (Cité en page 66).
- Gao, J., Giri, S., Kara, E. C., and Bergés, M. (2014). Plaid : A public dataset of high-resolution electrical appliance measurements for load identification research : Demo abstract. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, BuildSys ’14, pages 198–199, New York, NY, USA. ACM. (Cité en pages 102 et 108).
- Gillis, J. M. and Morsi, W. G. (2016). Non-intrusive load monitoring using semi-supervised machine learning and wavelet design. *IEEE Transactions on Smart Grid*, PP(99) :1–8. (Cité en page 21).
- Goncalves, H., Ocneanu, A., Berges, M., and Fan, R. (2011). Unsupervised disaggregation of appliances using aggregated consumption data. In *The 1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*. (Cité en page 24).
- Gulati, M., Ram, S. S., and Singh, A. (2014). An in depth study into using emi signatures for appliance identification. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 70–79. ACM. (Cité en page 108).
- Gupta, S., Reynolds, M. S., and Patel, S. N. (2010). Electrisense : single-point sensing using emi for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 139–148. ACM. (Cité en page 24).
- Hacine-Gharbi, A., Ravier, P., Harba, R., and Mohamadi, T. (2012). Low bias histogram-based estimation of mutual information for feature selection. *Pattern Recognition Letters*, 33(10) :1302 – 1308. (Cité en page 100).
- Hancke, G. P., Hancke Jr, G. P., et al. (2012). The role of advanced sensing in smart cities. *Sensors*, 13(1) :393–425. (Cité en page 26).
- Hart, G. W. (1985). Prototype nonintrusive appliance load monitor. In *MIT Energy Laboratory Technical Report, and Electric Power Research Institute Technical Report*. (Cité en page 19).
- Hart, G. W. (1989). Residential energy monitoring and computerized surveillance via utility power flows. *Technology and Society Magazine, IEEE*, 8(2) :12–16. (Cité en page 19).
- Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*,

- 80(12) :1870–1891. (Cité en pages [vii](#), [19](#), [20](#) et [24](#)).
- Howell, E. K. (1979). How switches produce electrical noise. *Electromagnetic Compatibility, IEEE Transactions on*, (3) :162–170. (Cité en page [42](#)).
- Huang, D., Thottan, M., and Feather, F. (2013). Designing customized energy services based on disaggregation of heating usage. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pages 1–6. IEEE. (Cité en page [24](#)).
- IEEE Std 1459-2010 (2010). Ieee standard definitions for the measurement of electric power quantities under sinusoidal, nonsinusoidal, balanced, or unbalanced conditions. *IEEE Std 1459-2010 (Rev. of IEEE Std 1459-2000)*. (Cité en page [29](#)).
- Inagaki, S., Egami, T., Suzuki, T., Nakamura, H., and Ito, K. (2011). Nonintrusive appliance load monitoring based on integer programming. *Electrical Engineering in Japan*, 174(2) :18–25. (Cité en page [24](#)).
- Jain, A. K., Duin, R. P. W., and Mao, J. (2000). Statistical pattern recognition : A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(1) :4–37. (Cité en page [107](#)).
- Kahl, M., Haq, A. U., Kriechbaumer, T., and Jacobsen, H.-A. (2016). Whited-a worldwide household and industry transient energy data set. In *Workshop on Non-Intrusive Load Monitoring (NILM), 2016 Proceedings of the 3rd International*. (Cité en page [109](#)).
- Kay, S. M. (1993). *Fundamentals of statistical signal processing, volume I : estimation theory*. Prentice Hall. (Cité en page [51](#)).
- Kelly, D. and Knottenbelt, W. (2011). Disaggregating smart meter readings using device signatures. *Imperial Computing Science MSc Individual Project*. (Cité en page [24](#)).
- Kelly, J. and Knottenbelt, W. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific Data*, 2(150007). (Cité en page [108](#)).
- Kim, H., Marwah, M., Arlitt, M. F., Lyon, G., and Han, J. (2011). Unsupervised disaggregation of low frequency power measurements. In *SDM*, volume 11, pages 747–758. SIAM. (Cité en page [24](#)).
- Klein, P. (2015). *Non-Intrusive Information Sources for Activity Analysis in Ambient Assisted Living Scenarios*. Thèse de doctorat, Université de Haute-Alsace. (Cité en page [20](#)).
- Klein, P., Merckle, J., Benyoucef, D., and Bier, T. (2013). Test bench and quality measures for non-intrusive load monitoring algorithms. In *Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE*, pages 5006–5011. IEEE. (Cité en page [90](#)).
- Kolter, J. Z., Batra, S., and Ng, A. (2010). Energy disaggregation via discriminative sparse coding. In *Advances in Neural Information Processing Systems*, pages 1153–1161. (Cité en page [24](#)).
- Kolter, J. Z. and Jaakkola (2012). Approximate inference in additive factorial hmms with

- application to energy disaggregation. In *AISTATS*, volume 22, pages 1472–1482. (Cité en page 24).
- Kolter, J. Z. and Johnson, M. J. (2011). Redd : A public data set for energy disaggregation research. In *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, pages 1–6. (Cité en page 108).
- Lai, P.-h., Trayer, M., Ramakrishna, S., and Li, Y. (2012). Database establishment for machine learning in nilm. In *Proceedings of the 1st International Non-Intrusive Load Monitoring Workshop*. (Cité en pages 96 et 109).
- Lam, H., Fung, G., and Lee, W. (2007). A novel method to construct taxonomy electrical appliances based on load signatures. *Consumer Electronics, IEEE Transactions on*, 53(2) :653–660. (Cité en page 24).
- Laughman, C., Lee, K., Cox, R., Shaw, S., Leeb, S., Norford, L., and Armstrong, P. (2003). Power signature analysis. *Power and Energy Magazine, IEEE*, 1(2) :56–63. (Cité en pages 19 et 24).
- Lee, K. D., Leeb, S. B., Norford, L. K., Armstrong, P. R., Holloway, J., and Shaw, S. R. (2005). Estimation of variable-speed-drive power consumption from harmonic content. *Energy Conversion, IEEE Transactions on*, 20(3) :566–574. (Cité en page 24).
- Lee, W., Fung, G., Lam, H., Chan, F., and Lucente, M. (2004). Exploration on load signatures. In *International conference on electrical Engineering (ICEE)*, volume 2, pages 690–694. (Cité en page 24).
- Leeb, S. B., Shaw, S. R., and Kirtley Jr, J. L. (1995). Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *Power Delivery, IEEE Transactions on*, 10(3) :1200–1210. (Cité en pages 19, 24 et 115).
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2) :164–168. (Cité en page 48).
- Liang, J., Ng, S., Kendall, G., and Cheng, J. (2010a). Load signature study–part i : Basic concept, structure, and methodology. *Power Delivery, IEEE Transactions on*, 25(2) :551–560. (Cité en page 24).
- Liang, J., Ng, S. K., Kendall, G., and Cheng, J. W. (2010b). Load signature study–part ii : Disaggregation framework, simulation, and applications. *Power Delivery, IEEE Transactions on*, 25(2) :561–569. (Cité en page 24).
- Liao, J., Elafoudi, G., Stankovic, L., and Stankovic, V. (2014). Non-intrusive appliance load monitoring using low-resolution smart meter data. In *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, pages 535–540. IEEE. (Cité en page 24).
- Linky (2016). Linky, the communicating meter. <http://www.erdf.fr/linky-communicating-meter> [Visité le : 03/10/2016]. (Cité en page 31).
- Liu, H. and Motoda, H. (1998). *Feature extraction, construction and selection : A data mining perspective*. Springer Science & Business Media. (Cité en page 144).

- Makonin, S., Bajic, I. V., and Popowich, F. (2014). Efficient sparse matrix processing for nonintrusive load monitoring (nilm). In *2nd International Workshop on Non-Intrusive Load Monitoring*. (Cité en page 24).
- Makonin, S. and Popowich, F. (2014). Nonintrusive load monitoring (nilm) performance evaluation. *Energy Efficiency*, pages 1–6. (Cité en page 90).
- Makonin, S., Popowich, F., Bartram, L., Gill, B., and Bajic, I. V. (2013). Ampds : A public dataset for load disaggregation and eco-feedback research. In *Electrical Power & Energy Conference (EPEC), 2013 IEEE*, pages 1–6. IEEE. (Cité en page 108).
- Mallat, S. (1999). *A wavelet tour of signal processing*. Academic press. (Cité en page 98).
- Marceau, M. L. and Zmeureanu, R. (2000). Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. *Energy Conversion and Management*, 41(13) :1389–1403. (Cité en page 24).
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2) :431–441. (Cité en page 48).
- Mika, S., Smola, A., and Schölkopf, B. (2001). An improved training algorithm for kernel fisher discriminants. pages 98–104, San Francisco, CA. Max-Planck-Gesellschaft, Morgan Kaufman. (Cité en page 89).
- Monacchi, A., Egarter, D., Elmenreich, W., D’Alessandro, S., and Tonello, A. M. (2014). GREEND : an energy consumption dataset of households in italy and austria. *CoRR*, abs/1405.3100. (Cité en page 108).
- Motoda, H. and Liu, H. (2002). Feature selection, extraction and construction. *Communication of IICM (Institute of Information and Computing Machinery, Taiwan) Vol, 5* :67–72. (Cité en page 144).
- Nait-Meziane, M., Hacine-Gharbi, A., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016). Hmm-based transient and steady-state current signals modeling for electrical appliances identification. In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, pages 670–677. (Cité en page 5).
- Nait Meziane, M., Hacine-Gharbi, A., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2017a). Electrical appliances identification and clustering using novel turn-on transient features. In *6th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*. Submitted for review. (Voir publications).
- Nait Meziane, M., Picon, T., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016a). A measurement system for creating datasets of on/off-controlled electrical loads. In *Conference on Environment and Electrical Engineering (EEEIC), 2016 Proceedings of the 16th IEEE International*, pages 2579–2583. (Cité en page 5).
- Nait Meziane, M., Picon, T., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016b). A new measurement system for high frequency nilm with controlled aggregation scenarios. In *Workshop on Non-Intrusive Load Monitoring (NILM), 2016*

- Proceedings of the 3rd International*. (Voir publications).
- Nait Meziane, M., Ravier, P., Lamarque, G., Abed-Meraim, K., Le Bunetel, J.-C., and Raingeaud, Y. (2015a). Modeling and estimation of transient current signals. In *Signal Processing Conference (EUSIPCO), 2015 Proceedings of the 23rd European*, pages 2005–2009. (Cité en page 4).
- Nait Meziane, M., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2016c). Accuracy evaluation of low-cost energy meters for home electrical consumption assessment. In *Conference on Environment and Electrical Engineering (EEEIC), 2016 Proceedings of the 16th IEEE International*, pages 2473–2478. (Voir publications).
- Nait Meziane, M., Ravier, P., Lamarque, G., Le Bunetel, J.-C., and Raingeaud, Y. (2017b). High accuracy event detection for non-intrusive load monitoring. In *42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Accepted for presentation. (Cité en page 4).
- Nait Meziane, M., Ravier, P., Le Bunetel, J.-C., and Lamarque, G. (2015b). Identification of consumers on an electrical network by analyzing their load profile. In *Journée des Jeunes Chercheurs (JJC) du laboratoire Prisme, juin 2015*. (Voir publications).
- Najmeddine, H. (2009). *Méthode d'identification et de classification de la consommation d'énergie par usages en vue de l'intégration dans un compteur d'énergie électrique*. Thèse de doctorat, Université Blaise Pascal-Clermont-Ferrand II. (Cité en page 20).
- Najmeddine, H., El Khamlichi Drissi, K., Pasquier, C., Faure, C., Kerroum, K., Diop, A., Jouannet, T., and Michou, M. (2008). State of art on load monitoring methods. In *Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International*, pages 1256–1258. IEEE. (Cité en page 19).
- Norford, L. K. and Leeb, S. B. (1996). Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings*, 24(1) :51–64. (Cité en page 24).
- Parson, O. (2012). Using hidden markov model variants for non-intrusive appliance load monitoring from smart meter data. (Cité en pages 18 et 96).
- Parson, O., Fisher, G., Hersey, A., Batra, N., Kelly, J., Singh, A., Knottenbelt, W., and Rogers, A. (2015). Dataport and nilmtk : A building data set designed for non-intrusive load monitoring. (Cité en page 108).
- Parson, O., Ghosh, S., Weal, M., and Rogers, A. (2012). Non-intrusive load monitoring using prior models of general appliance types. In *AAAI*. (Cité en page 24).
- Parson, O., Ghosh, S., Weal, M., and Rogers, A. (2014). An unsupervised training method for non-intrusive appliance load monitoring. *Artificial Intelligence*. (Cité en page 96).
- Patel, S. N., Robertson, T., Kientz, J. A., Reynolds, M. S., and Abowd, G. D. (2007). At the flick of a switch : Detecting and classifying unique electrical events on the residential power line (nominated for the best paper award). In *UbiComp 2007 : ubiquitous computing*, pages 271–288. Springer. (Cité en pages 19 et 24).

- Pereira, L., Quintal, F., Gonçalves, R., and Nunes, N. J. (2014). Sustdata : A public dataset for ict4s electric energy research. In *ICT4S*. (Cité en page 108).
- Perez, K. X., Cole, W. J., Rhodes, J. D., Ondeck, A., Webber, M., Baldea, M., and Edgar, T. F. (2014). Nonintrusive disaggregation of residential air-conditioning loads from sub-hourly smart meter data. *Energy and Buildings*, 81 :316–325. (Cité en page 24).
- Picon, T., Nait Meziane, M., Ravier, P., Lamarque, G., Novello, C., Le Bunetel, J.-C., and Raingeaud, Y. (2016). COOLL : Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. *arXiv preprint arXiv :1611.05803 [cs.OH]*. (Voir publications).
- PLAID (2015). *PLAID : the Plug Load Appliance Identification Dataset*. [Online] Available from : <http://plaidplug.com>. [Accessed : 22nd October 2015]. (Cité en page 96).
- Popescu, F., Enache, F., Vizitiu, I. C., and Ciotirnae, P. (2014). Recurrence plot analysis for characterization of appliance load signature. In *Communications (COMM), 2014 10th International Conference on*, pages 1–4. (Cité en page 24).
- Powers, J., Margossian, B., and Smith, B. (1991). Using a rule-based algorithm to disaggregate end-use load profiles from premise-level data. *Computer Applications in Power, IEEE*, 4(2) :42–47. (Cité en page 24).
- Pratt, R. G., Balducci, P., Gerkenmeyer, C., Katipamula, S., Kintner-Meyer, M. C., Sanquist, T. F., Schneider, K. P., and Secrets, T. (2010). *The smart grid : An estimation of the energy and CO2 benefits*. Pacific Northwest National Laboratory. (Cité en page 26).
- Prudenzi, A. (2002). A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel. In *Power Engineering Society Winter Meeting, 2002. IEEE*, volume 2, pages 941–946. IEEE. (Cité en page 24).
- Ridi, A. and Hennebert, J. (2014). Hidden markov models for ilm appliance identification. *Procedia Computer Science*, 32 :1010–1015. (Cité en page 96).
- Roos, J., Lane, I., Botha, E., and Hancke, G. (1994). Using neural networks for non-intrusive monitoring of industrial electrical loads. In *Instrumentation and Measurement Technology Conference, 1994. IMTC/94. Conference Proceedings. 10th Anniversary. Advanced Technologies in I & M., 1994 IEEE*, pages 1115–1118. IEEE. (Cité en page 24).
- RTE (2016). Bilan prévisionnel de l’équilibre offre-demande d’électricité en france. Technical report, Réseau de transport d’électricité. (Cité en pages vii, 10, 11 et 14).
- Ruzzelli, A. G., Nicolas, C., Schoofs, A., and O’Hare, G. M. (2010). Real-time recognition and profiling of appliances through a single electricity sensor. In *2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 1–9. IEEE. (Cité en page 24).
- Sanquer, M. (2013). *Détection et caractérisation de signaux transitoires : Application à la surveillance de courbes de charge*. Thèse de doctorat, Université de Grenoble. (Cité en page 20).

- Shaw, S. R., Leeb, S. B., Norford, L. K., and Cox, R. W. (2008). Nonintrusive load monitoring and diagnostics in power systems. *Instrumentation and Measurement, IEEE Transactions on*, 57(7) :1445–1454. (Cité en page 24).
- Srinivasan, D., Ng, W., and Liew, A. (2006). Neural-network-based signature recognition for harmonic source identification. *Power Delivery, IEEE Transactions on*, 21(1) :398–405. (Cité en page 24).
- Steve Young, Gunnar Evermann, M. G. T. H. D. K. X. L. G. M. J. O. D. O. D. P. V. V. P. W. (2009). *The HTK book (for HTK Version 3.4)*. (Cité en page 100).
- Sultanem, F. (1991). Using appliance signatures for monitoring residential loads at meter panel level. *Power Delivery, IEEE Transactions on*, 6(4) :1380–1385. (Cité en pages 19 et 24).
- Suzuki, K., Inagaki, S., Suzuki, T., Nakamura, H., and Ito, K. (2008). Nonintrusive appliance load monitoring based on integer programming. In *SICE Annual Conference, 2008*, pages 2742–2747. IEEE. (Cité en page 24).
- Tang, G., Wu, K., and Lei, J. (2014a). Semi-intrusive load monitoring for large-scale appliances. In *International Workshop on Non-Intrusive Load Monitoring (NILM 2014)*. (Cité en page 19).
- Tang, G., Wu, K., Lei, J., and Tang, J. (2014b). A simple model-driven approach to energy disaggregation. In *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, pages 566–571. IEEE. (Cité en page 24).
- Thiruvaran, T., Phung, T., and Ambikairajah, E. (2013). Automatic identification of electric loads using switching transient current signals. In *TENCON Spring Conference, 2013 IEEE*, pages 252–256. IEEE. (Cité en pages 96 et 97).
- Ting, K., Lucente, M., Fung, G. S., Lee, W., and Hui, S. (2005). A taxonomy of load signatures for single-phase electric appliances. In *IEEE PESC (Power Electronics Specialist Conference)*, pages 12–18. (Cité en page 19).
- Uttama Nambi, A. S., Reyes Lua, A., and Prasad, V. R. (2015). Loded : Location-aware energy disaggregation framework. In *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, BuildSys '15*, pages 45–54, New York, NY, USA. ACM. (Cité en page 108).
- Wang, Y., Filippi, A., Rietman, R., and Leus, G. (2012). Compressive sampling for non-intrusive appliance load monitoring (nalm) using current waveforms. In *Proc. IASTED Int. Conf. Signal Process., Pattern Recognition Applicat., Crete, Greece*, page 8. (Cité en page 24).
- Wichakool, W., Remscrim, Z., Orji, U. A., and Leeb, S. B. (2015). Smart metering of variable power loads. *IEEE Transactions on Smart Grid*, 6(1) :189–198. (Cité en page 24).
- Wild, B., Barsim, K. S., and Yang, B. (2015). A new unsupervised event detector for non-intrusive load monitoring. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 73–77. (Cité en pages 22, 83, 89 et 94).

BIBLIOGRAPHIE

- Ye, S. and Aboutanios, E. (2015). An algorithm for the parameter estimation of multiple superimposed exponentials in noise. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3457–3461. IEEE. (Cité en page [77](#)).
- Zeifman, M. and Roth, K. (2011). Nonintrusive appliance load monitoring : Review and outlook. *Consumer Electronics, IEEE Transactions on*, 57(1) :76–84. (Cité en page [19](#)).
- Zhao, B., Stankovic, L., and Stankovic, V. (2016). On a training-less solution for non-intrusive appliance load monitoring using graph signal processing. *IEEE Access*, 4 :1784–1799. (Cité en page [21](#)).
- Zia, T., Bruckner, D., and Zaidi, A. (2011). A hidden markov model based procedure for identifying household electric loads. In *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, pages 3218–3223. IEEE. (Cité en page [96](#)).
- Zoha, A., Gluhak, A., Nati, M., and Imran, M. A. (2013). Low-power appliance monitoring using factorial hidden markov models. In *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, pages 527–532. IEEE. (Cité en page [24](#)).

Mohamed NAIT MEZIANE

Identification d'appareils électriques par analyse des courants de mise en marche

Résumé :

Le domaine lié à ce travail est appelé « désagrégation d'énergie », où la principale préoccupation est de décomposer, ou désagréger, la consommation globale d'énergie électrique (par exemple, la consommation de tout un ménage) en une consommation détaillée donnée comme information de consommation par usage (par exemple, par appareil). Cette dernière permet d'avoir un retour sur la consommation pour les consommateurs ainsi que pour les fournisseurs et est utile pour permettre des économies d'énergie. Dans ce domaine de désagrégation d'énergie, il existe trois grandes questions auxquelles il faut répondre : qui consomme ? quand ? et combien ? Les recherches menées dans cette thèse se concentrent sur l'identification des appareils électriques, c'est-à-dire la réponse à la première question, en considérant particulièrement des appareils ménagers. À cet effet, nous utilisons le courant transitoire de mise en marche que nous modélisons en utilisant un nouveau modèle que nous avons proposé. De plus, nous utilisons les paramètres estimés de ce dernier pour la tâche d'identification.

Mots clés : Bornes de Cramér-Rao (BCR), courant transitoire de mise en marche, désagrégation d'énergie, détection d'événements, identification d'appareils électriques, modélisation et estimation, Non-Intrusive Load Monitoring (NILM)

Analysis of turn-on transient currents for electrical appliances identification

Abstract :

The related field to this work is called "energy disaggregation" where the main concern is to break down, or disaggregate, the global electrical energy consumption (e.g. whole-house consumption) into a detailed consumption given as end-use (e.g. appliance-level) consumption information. This latter gives consumption feedback to consumers and electricity providers and is helpful for energy savings. Three main questions have to be answered in the energy disaggregation field : who is consuming ? when ? and how much ? The research conducted in this thesis focuses on electrical appliances identification, i.e. the who question, considering particularly home appliances. For this purpose, we use the turn-on transient current signal which we model using a new model we proposed and use its estimated model parameters for the identification task.

Keywords : Cramér-Rao Bounds (CRB), electrical appliances identification, event detection, energy disaggregation, modeling and estimation, Non-Intrusive Load Monitoring (NILM), turn-on transient current

Laboratoire Pluridisciplinaire de Recherche en Ingénierie des
Systèmes, Mécanique et Energétique (PRISME)
12 rue de Blois, 45067 Orléans, France

