



**HAL**  
open science

# Diagnostic par reconnaissance des formes : application à un ensemble convertisseur - machine asynchrone

Olivier Ondel

► **To cite this version:**

Olivier Ondel. Diagnostic par reconnaissance des formes : application à un ensemble convertisseur - machine asynchrone. Autre. Ecole Centrale de Lyon, 2006. Français. NNT: . tel-00113102

**HAL Id: tel-00113102**

**<https://theses.hal.science/tel-00113102>**

Submitted on 10 Nov 2006

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

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

# THESE

Présentée devant

**L'ÉCOLE CENTRALE DE LYON**

pour obtenir le grade de

**DOCTEUR**

(arrêté du 30/03/1992)

**Spécialité : Génie Électrique**

Préparée au sein de

**L'ÉCOLE DOCTORALE**

**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE**

**DE LYON**

Par

**Olivier ONDEL**

---

**DIAGNOSTIC PAR RECONNAISSANCE DES FORMES : APPLICATION A UN  
ENSEMBLE CONVERTISSEUR – MACHINE ASYNCHRONE**

---

*Soutenu le 17 Octobre 2006 devant la commission d'examen*

Jury : MM.

<b>B. DUBUISSON</b>	Professeur des Universités HEUDIASYC - Compiègne	Rapporteur
<b>M.E.H. BENBOUZID</b>	Professeur des Universités LIME - Brest	Rapporteur
<b>J. FAUCHER</b>	Professeur des Universités LEEI – Toulouse	Examineur
<b>G. CLERC</b>	Professeur des Universités CEGELY - Lyon	Examineur
<b>R. MEURET</b>	Responsable laboratoire d'électronique de puissance Hispano-Suiza – groupe SAFRAN	Examineur
<b>E. BOUTLEUX</b>	Maître de Conférences CEGELY - Lyon	Examineur





**ECOLE CENTRALE DE LYON**  
**DIRECTION DE LA RECHERCHE**

**liste des personnes Habilitées à Diriger des Recherches en poste à l'Ecole Centrale de Lyon**

Nom-Prénom	Corps grade	Laboratoire ou à défaut département ECL	Etablissement
AURIOL Philippe	professeur	CEGELY	ECL
BEROUAL Abderrahmane	professeur	CEGELY	ECL
JAFFREZIC-RENAULT Nicole	directeur de recherche	CEGELY	CNRS/ECL
KRÄHENBÜHL Laurent	directeur de recherche	CEGELY	CNRS/ECL
MARTELET Claude	professeur	CEGELY	ECL
NICOLAS Alain	professeur	CEGELY	ECL
NICOLAS Laurent	directeur de recherche	CEGELY	CNRS/ECL
SIMONET Pascal	chargé de recherche	CEGELY	CNRS/ECL
THOMAS Gérard	professeur	CEGELY	ECL
VOLLAIRE Christian	maître de conférences	CEGELY	ECL

**Nbre Cégely 10**

HELLOUIN Yves	maître de conférences	DER EEA	ECL
LE HELLEY Michel	professeur	DER EEA	ECL

**Nbre DER EEA 2**

GUIRALDENQ Pierre	professeur émérite	DER STMS	ECL
VINCENT Léo	professeur	DER STMS	ECL

**Nbre DER STMS 2**

LOHEAC Jean-Pierre	maître de conférences	ICJ	ECL
MAITRE Jean-François	professeur émérite	ICJ	ECL
MARION Martine	professeur	ICJ	ECL
MOUSSAOUI Mohand	professeur	ICJ	ECL
MUSY François	maître de conférences	ICJ	ECL
ROUY MIRONESCU Elisabeth	professeur	ICJ	ECL
ZINE Abdel-Malek	maître de conférences	ICJ	ECL

**Nbre ICJ 7**

DAVID Bertrand	professeur	ICTT	ECL
----------------	------------	------	-----

**Nbre ICTT 1**

CLOAREC Jean-Pierre	maître de conférences	LEOM	ECL
GAFFIOT Frédéric	professeur	LEOM	ECL
GAGNAIRE Alain	maître de conférences	LEOM	ECL
GARRIGUES Michel	directeur de recherche	LEOM	CNRS/ECL
GENDRY Michel	directeur de recherche	LEOM	CNRS/ECL
GRENET Geneviève	directeur de recherche	LEOM	CNRS/ECL
HOLLINGER Guy	directeur de recherche	LEOM	CNRS/ECL
JOSEPH Jacques	professeur	LEOM	ECL
KRAWCZYK Stanislas	directeur de recherche	LEOM	CNRS/ECL
MARTIN Jean-René	professeur émérite	LEOM	ECL
O'CONNOR Ian	maître de conférences	LEOM	ECL

PHANER-GOUTORBE Magali	professeur	LEOM	ECL
ROBACH Yves	professeur	LEOM	ECL
SOUTEYRAND Eliane	directeur de recherche	LEOM	CNRS/ECL
TARDY Jacques	directeur de recherche	LEOM	CNRS/ECL
VIKTOROVITCH Pierre	directeur de recherche	LEOM	CNRS/ECL

**Nbre LEOM 16**

CHEN Liming	professeur	LIRIS	ECL
-------------	------------	-------	-----

**Nbre LIRIS 1**

BAILLY Christophe	professeur	LMFA	ECL
BERTOGLIO Jean-Pierre	directeur de recherche	LMFA	CNRS/ECL
BLANC-BENON Philippe	directeur de recherche	LMFA	CNRS/ECL
CAMBON Claude	directeur de recherche	LMFA	CNRS/ECL
CARRIERE Philippe	chargé de recherche	LMFA	CNRS/ECL
CHAMPOUSSIN J-Claude	professeur émérite	LMFA	ECL
COMTE-BELLOT geneviève	professeur émérite	LMFA	ECL
FERRAND Pascal	directeur de recherche	LMFA	CNRS/ECL
GALLAND Marie-Annick	maître de conférences	LMFA	ECL
GODEFERD Fabien	chargé de recherche	LMFA	CNRS/ECL
HENRY Daniel	directeur de recherche	LMFA	CNRS/ECL
JEANDEL Denis	professeur	LMFA	ECL
JUVE Daniel	professeur	LMFA	ECL
LE RIBAUT Catherine	chargée de recherche	LMFA	CNRS/ECL
LEBOEUF Francis	professeur	LMFA	ECL
PERKINS Richard	professeur	LMFA	ECL
ROGER Michel	professeur	LMFA	ECL
SCOTT Julian	professeur	LMFA	ECL
SIMOENS Serge	chargé de recherche	LMFA	CNRS/ECL
TREBINJAC Isabelle	maître de conférences	LMFA	ECL

**Nbre LMFA 20**

BENAYOUN Stéphane	professeur	LTDS	ECL
CAMBOU Bernard	professeur	LTDS	ECL
COQUILLET Bernard	maître de conférences	LTDS	ECL
DANESCU Alexandre	maître de conférences	LTDS	ECL
FOUVRVY Siegfried	chargé de recherche	LTDS	CNRS/ECL
GEORGES Jean-Marie	professeur émérite	LTDS	ECL
GUERRET Chrystelle	chargé de recherche	LTDS	CNRS/ECL
HERTZ Dominique	past	LTDS	ECL
ICHCHOU Mohamed	maître de conférences	LTDS	ECL
JEZEQUEL Louis	professeur	LTDS	ECL
JUVE Denyse	ingénieur de recherche	LTDS	ECL
KAPSA Philippe	directeur de recherche	LTDS	CNRS/ECL
LE BOT Alain	chargé de recherche	LTDS	CNRS/ECL
LOUBET Jean-Luc	directeur de recherche	LTDS	CNRS/ECL
MARTIN Jean-Michel	professeur	LTDS	ECL
MATHIA Thomas	directeur de recherche	LTDS	CNRS/ECL
MAZUYER Denis	professeur	LTDS	ECL
PERRET-LIAUDET Joël	maître de conférences	LTDS	ECL
SALVIA Michelle	maître de conférences	LTDS	ECL
SIDOROFF François	professeur	LTDS	ECL
STREMSDOERFER Guy	professeur	LTDS	ECL
THOUVEREZ Fabrice	professeur	LTDS	ECL
TREHEUX Daniel	professeur	LTDS	ECL
VANNES André-Bernard	professeur émérite	LTDS	ECL

**Nbre LTDS 24**

Total HdR ECL

91

## **REMERCIEMENTS**

*Les travaux de recherche développés dans ce mémoire ont été réalisés au sein du Centre de Génie Electrique de LYon (CEGELY – UMR 5005 CNRS) sur le site de l'Ecole Centrale de Lyon et de l'Université Claude Bernard Lyon 1.*

*Je remercie sincèrement Messieurs Laurent NICOLAS, directeur du CEGELY, et Philippe AURIOL, directeur du département EEA, de m'avoir accueilli dans ce laboratoire.*

*J'adresse mes vifs remerciements et ma profonde gratitude envers Messieurs Guy CLERC, Professeur à l'Université Claude Bernard Lyon 1, et Emmanuel BOUTLEUX, maître de Conférences à l'Ecole Centrale de Lyon, pour l'aide précieuse qu'ils m'ont apporté durant ces travaux. Je les remercie pour leurs conseils avisés, leurs nombreuses remarques et suggestions qui ont pu faire avancées le travail et surtout leurs grandes qualités humaines.*

*Je tiens à remercier tout particulièrement Monsieur Eric BLANCO, maître de Conférences à l'Ecole Centrale de Lyon, qui, même s'il ne figurait pas dans l'encadrement administratif de ma thèse, m'a permis de part son expérience, ses conseils, la pertinence de ses critiques et la finesse de ses suggestions d'améliorer ce travail.*

*Je remercie vivement Monsieur Bernard DUBUISSON, Professeur à l'Université de Technologie de Compiègne, pour avoir rapporté avec soin ce mémoire et pour avoir accepté de présider le jury.*

*Je tiens également à remercier Monsieur Mohamed El Hachemi BENBOUZID d'avoir accepté d'examiner ce mémoire et d'en être un des rapporteurs. Qu'il trouve ici le témoignage de ma gratitude pour l'important travail d'évaluation et pour sa participation au jury.*

*J'adresse également mes remerciements à Monsieur Jean FAUCHER, Professeur à l'Institut National Polytechnique de Toulouse, et à Monsieur Régis MEURET, Responsable laboratoire d'électronique de puissance Hispano-Suiza, de m'avoir fait l'honneur de participer au jury.*

*Durant ces années de doctorat, les membres administratifs et techniques du département et du laboratoire ont toujours su répondre présent à mes nombreuses sollicitations avec disponibilité, sympathie et bien sûr toutes leurs compétences. Qu'ils reçoivent toute ma gratitude.*

*Je ne saurai oublier les jeunes doctorants et futurs docteurs du laboratoire ainsi que ceux des autres laboratoires de l'Ecole, qui durant ces années, ont entretenu une très bonne ambiance et m'ont permis d'entretenir ma forme en organisant régulièrement dans match de*

*foot. Je souhaite à tous ces doctorants du succès dans leurs recherches ainsi que dans leurs vies professionnelles.*

*Enfin, un grand merci pour ceux que je n'ai pu désignés et qui m'ont aidé de près ou de loin à parcourir ce chemin.*

## RESUME

Les progrès réalisés en matière d'électronique de puissance, de circuits de commande ont contribué à l'utilisation grandissante des machines asynchrones dans les systèmes d'entraînements électriques. Le recours aux machines asynchrones est surtout lié à leur robustesse, leur puissance massique et à leur coût de fabrication. L'apparition dans les années 1980 des variateurs permettant de faire varier la fréquence de rotation dans une large gamme a largement favorisé son développement. En effet, ils entrent dans la conception de nombreux procédés industriels associant des convertisseurs statiques et des machines électriques.

La maintenance et la surveillance de ces deux systèmes permettent de rentabiliser les installations. Il est donc important de développer des outils de diagnostic pour détecter de manière précoce les défauts pouvant apparaître aussi bien sur le convertisseur que sur la machine.

Notre approche est basée sur l'utilisation des méthodes de reconnaissance des formes. Un vecteur de paramètres, appelé vecteur forme, est extrait de chacune des mesures effectuées sur la machine. Les règles de décisions utilisées permettent de classer les observations, décrites par le vecteur forme, par rapport aux différents modes de fonctionnement connus avec ou sans défaut.

Des défauts ont été créés au rotor et au stator de la machine asynchrone, alimentée soit à partir du réseau, soit par le biais d'un onduleur de tension.

La procédure de décision, basée sur la règle des  $k$  - plus proches voisins, associée à une fonction d'appartenance, permet de détecter l'évolution des modes de fonctionnements ainsi que les défauts avérés. Par la suite, le suivi d'évolution de ces modes est réalisé par une approche de type Kalman : un estimateur récursif de Kalman est utilisé pour déterminer les paramètres du modèle dynamique rendant compte de l'évolution d'un mode et un prédicteur de Kalman pour prévoir une évolution vers de nouvelles zones de l'espace. Ces algorithmes ont montré l'efficacité de l'application de la reconnaissance des formes au diagnostic.



## ABSTRACT

Advances in power electronics, control circuits and automatic have contributed to an increasing use of induction motors in electrical drive systems. The large – scale utilization of induction motors is mainly due to their robustness, their power – weight ratio, and to their manufacturing cost. The appearance of variators making it possible to vary the rotational frequency largely supported its development. Indeed, these variators enter the design of many industrial processes associating static inverters and electric machines. The maintenance and the monitoring of these two systems allow making profitable the installations. Therefore, it is important to develop diagnosis tools in order to detect earlier the faults, which can appear in these machines.

Our approach is based on pattern recognition methods. A vector of features, named pattern vector, is obtained from the measurements made on the machine. These classifications are made according to the different operating conditions, with and without fault.

Faults have been created on both the rotor and the stator sides of the induction machine. This one was fed either from the mains, or from a three – phase voltage inverter.

Fault detection has been made with decision procedure based on the k – nearest neighbors rule, associated with a membership function. This procedure allows detecting the evolution of the operating modes as well as the proven defects. Thereafter, the evolution tracking of these modes is carried out by a Kalman approach: a recursive Kalman estimator is used to determine the parameters of the dynamic model accounting for the evolution of a mode and a Kalman predictor to envisage an evolution towards new zones of space.

The results obtained with these algorithms have proved the efficiency of pattern recognition methods for diagnosis.

# **SOMMAIRE**

<b>INTRODUCTION GENERALE -----</b>	<b>1</b>
------------------------------------	----------

**CHAPITRE I : SYNTHESE DES DEFAILLANCES POSSIBLES SUR UN ENSEMBLE  
CONVERTISSEUR - MACHINE ASYNCHRONE**

---

<b><i>1.1. Introduction -----</i></b>	<b>7</b>
<b><i>1.2. Constitution de la machine asynchrone -----</i></b>	<b>7</b>
1.2.1. Organisation de la machine -----	7
1.2.1.1. Le stator -----	8
1.2.1.2. Le rotor -----	8
1.2.1.3. Les organes mécaniques -----	10
<b><i>1.3. Constitution de l'alimentation d'un moteur asynchrone -----</i></b>	<b>10</b>
<b><i>1.4. Présentation des différentes défaillances du moteur asynchrone à cage d'écureuil -----</i></b>	<b>13</b>
1.4.1. Défaillances au rotor -----	15
1.4.1.1. Ruptures de barres -----	15
1.4.1.2. Ruptures d'anneaux -----	15
1.4.1.3. Excentricité statique et dynamique -----	16
1.4.2. Défaillances au stator -----	17
1.4.2.1. Défauts d'isolant dans un enroulement -----	17
1.4.2.2. Court-circuit entre spires -----	18
1.4.2.3. Court-circuit entre phases -----	18
1.4.2.4. Court-circuit phase/bâti -----	19
1.4.2.5. Défauts du circuit magnétique -----	19
1.4.3. Défaillances mécaniques -----	19
1.4.3.1. Défauts roulements -----	19
1.4.3.2. Autres défaillances mécaniques -----	20
<b><i>1.5. Présentation des différentes défaillances sur un convertisseur de puissance -----</i></b>	<b>20</b>

1.5.1. Défaillance du condensateur -----	21
1.5.1.1. Cas des condensateurs électrolytiques -----	21
1.5.1.2. Cas des condensateurs à films métallisés -----	23
1.5.2. Défaillance de diodes classiques -----	24
1.5.3. Défaillance d'un exemple d'interrupteur statique : l'IGBT (Insulated Gate Bipolar Transistor) -----	24
1.5.3.1. Court-circuit -----	24
1.5.3.2. Circuit ouvert -----	24
1.5.3.3. Vieillessement -----	25
1.5.4. Défaillance du circuit commande (vieillessement de l'optocoupleur) -----	25
<b>1.6. Défaillances d'autres origines -----</b>	<b>26</b>
1.6.1. Défauts de la charge -----	26
1.6.2. Défauts capteurs -----	28
<b>1.7. Synthèse des défaillances -----</b>	<b>28</b>
<b>1.8. Conclusion -----</b>	<b>33</b>

---

**CHAPITRE II :**           EXTRACTION DE SIGNATURES REPRESENTATIVES DES DEFAILLANCES :

ETAT DE L'ART

---

<b>2.1. Introduction -----</b>	<b>35</b>
<b>2.2. Extraction de signatures -----</b>	<b>36</b>
2.2.1. Extraction à partir de méthodes internes ( <i>connaissance a priori</i> ) -----	36
2.2.1.1. Méthodes de diagnostic basées sur des modèles paramétriques -----	36
2.2.1.2. Diagnostic basé sur une modélisation de la machine asynchrone en fonctionnement sain -----	37
2.2.1.2.1. Modèle triphasé-triphasé -----	37
2.2.1.2.1.1. Présentation du modèle -----	37
2.2.1.2.1.2. Discrimination des défauts -----	38
2.2.1.2.2. Modèle diphasé -----	40
2.2.1.2.2.1. Présentation du modèle -----	40
2.2.1.2.2.2. Discrimination des défauts -----	42
2.2.1.3. Diagnostic de la machine basé sur un modèle prenant en compte les défauts -----	44
2.2.1.4. Conclusion -----	46

2.2.2. Extraction à partir de méthodes externes ( <i>sans connaissance a priori</i> ) -----	47
2.2.2.1. Utilisation des différents signaux autres que le courant statorique -----	47
2.2.2.1.1. Analyse des signaux vibratoires -----	47
2.2.2.1.2. Analyse des flux -----	48
2.2.2.1.3. Analyse du couple électromagnétique -----	49
2.2.2.2. Détection des défauts par analyse des courants statoriques -----	51
2.2.2.2.1. Préconditionnement des courants statoriques pour l'analyse fréquentielle -----	52
2.2.2.2.2. Signatures fréquentielles extraites à partir des signaux préconditionnés -----	53
2.2.2.2.2.1. Ruptures de barres ou portions d'anneaux -----	53
2.2.2.2.2.2. Excentricité -----	54
2.2.2.2.2.3. Défauts roulements -----	54
2.2.2.2.2.4. Défauts au niveau du stator -----	55
2.2.2.3. Signatures statistiques -----	56
2.2.2.3.1. Extraites à partir de signaux mesurés -----	56
2.2.2.3.1. Extraites à partir de signaux préconditionnés -----	57
2.2.2.3.1.1. Calcul sur les composantes $I_{s\alpha}$ et $I_{s\beta}$ -----	58
2.2.2.3.1.2. Calcul sur les puissances -----	60
2.2.2.4. Les outils d'analyse fréquentielle liés au traitement de l'information -----	62
2.2.2.5. Signatures basées sur l'échauffement de la machine -----	66
2.2.2.6. Synthèse des signatures déterminées à partir des méthodes externes -----	67
<b>2.3. Vers des approches modernes de diagnostic -----</b>	<b>70</b>
2.3.1. Les systèmes experts -----	70
2.3.2. Les arbres de défaillances -----	71
2.3.3. Les réseaux de neurones -----	72
2.3.4. La reconnaissance des formes (RdF) -----	74
<b>2.4. Conclusion -----</b>	<b>74</b>

---

**CHAPITRE III :            APPROCHE DIAGNOSTIC PAR RECONNAISSANCE DES FORMES (RdF)**

---

<b>3.1. Introduction -----</b>	<b>77</b>
<b>3.2. Détermination de l'espace de représentation -----</b>	<b>81</b>
3.2.1. Constitution du vecteur forme -----	81
3.2.2. Réduction de l'espace de représentation -----	81
3.2.2.1. Extraction de paramètres par Analyse en Composantes Principales (ACP) -----	82
3.2.2.2. Méthodes de sélection de paramètres en mode supervisé -----	84
3.2.2.2.1. Critère de qualité -----	85
3.2.2.2.2. Techniques de recherches utilisées -----	87

3.2.2.2.3. Choix de la dimension finale (valeur de $d'$ )	88
3.2.2.3. Utilisation des algorithmes génétiques pour la sélection de paramètres	89
3.2.2.3.1. Introduction	89
3.2.2.3.2. Principe d'un algorithme génétique	89
3.2.2.3.3. Les éléments constituant un algorithme génétique	90
3.2.2.3.3.1. Codage de la population	90
3.2.2.3.3.2. Fonction d'adaptation proposée	91
3.2.2.3.3.3. Sélection	92
3.2.2.3.3.4. Croisement	92
3.2.2.3.3.5. Mutation	93
3.2.2.3.3.6. Arrêt de l'AG	93
<b>3.3. Détermination de l'espace de décision</b>	<b>94</b>
<b>3.4. Association simultanée de la coalescence et de la sélection de paramètres</b>	<b>102</b>
3.4.1. Introduction	102
3.4.2. Algorithme proposé	102
3.4.2.1. Fonction d'adaptation	103
3.4.2.2. Codage utilisé pour la représentation d'un <i>chromosome</i>	103
3.4.2.3. Croisement et mutation	104
3.4.2.4. Opérations de recherche locale	105
3.4.2.5. Problème multimodal	105
3.4.2.6. AG multi-population pour l'optimisation multimodale	106
3.4.2.6.1. Diversification	106
3.4.2.6.1.1. Migration	106
3.4.2.6.1.2. Spéciation	107
3.4.2.7. Arrêt de l'algorithme	108
<b>3.5. Détermination d'une méthode de décision</b>	<b>110</b>
3.5.1. Méthodes non paramétriques	110
3.5.1.1. Règle des $k$ – plus proches voisins ( $k - ppv$ )	111
3.5.1.1.1. Choix de la distance dans la règle des $k - ppv$	111
3.5.1.1.2. Choix du nombre de plus proches voisins $k$	114
3.5.1.2. Règle des $k$ – plus proches voisins avec rejet	114
3.5.1.2.1. Rejet d'ambiguïté	115
3.5.1.2.2. Rejet de distance	115
3.5.1.3. Réduction des temps de calcul	117
3.5.1.4. Classification non exclusive	120
3.5.1.4.1. Evaluation du degré d'appartenance d'une observation aux classes définies dans $X_a$	122
3.5.1.4.2. Extension de la règle des $k - ppv$ par association d'une fonction d'appartenance non exclusive	126
3.5.1.4.1.1. Rejet d'ambiguïté	127

3.5.1.4.1.2. Rejet d'appartenance -----	128
<b>3.6. Prise en compte de l'évolution des modes de fonctionnement -----</b>	<b>129</b>
3.6.1. Approche polynomiale -----	130
3.6.1.1. Estimation des états manquants -----	130
3.6.1.2. Vers la prédiction -----	131
3.6.2. Approche par filtre de Kalman -----	131
3.6.2.1. Estimation des états manquants -----	132
3.6.2.2. Vers la prédiction -----	137
3.6.3. Mise en place du rejet d'affectation -----	137
<b>3.7. Actualisation de la connaissance -----</b>	<b>140</b>
3.7.1. Analyse de la compacité des observations rejetées -----	142
3.7.2. Localisation des observations rejetées par maillage de l'espace de représentation -----	143
<b>3.8. Procédure de diagnostic par RdF -----</b>	<b>145</b>
<b>3.9. Conclusion -----</b>	<b>145</b>

---

**CHAPITRE IV :**                      APPLICATION AU DIAGNOSTIC DE LA MACHINE ASYNCHRONE

---

<b>4.1. Introduction -----</b>	<b>148</b>
<b>4.2. Description du procédé -----</b>	<b>149</b>
4.2.1. Présentation du banc moteur -----	149
4.2.2. Acquisitions des données -----	150
4.2.3. Modes de fonctionnement étudiés -----	151
<b>4.3. Choix du vecteur forme -----</b>	<b>153</b>
4.3.1. Analyse de la sensibilité des composantes fréquentielles -----	155
4.3.2. Recherche automatique du glissement -----	159
4.3.2.1. Détermination de la fréquence fondamentale $f_s$ -----	159
4.3.2.1.1. Analyse fréquentielle -----	162
4.3.2.1.2. Analyse temporelle -----	162
4.3.2.2. Calcul de la fréquence de rotation du rotor $f_r$ -----	164
4.3.2.3. Calcul du glissement $g$ -----	165
4.3.2.4. Calcul de la vitesse de rotation du moteur -----	166
<b>4.4. Normalisation des données -----</b>	<b>167</b>
4.4.1. Constitution des classes -----	170

4.4.2. Définition du seuil de rejet en distance pour des données normalisées et non normalisées -----	172
<b>4.5. Diagnostic par RdF d'une machine asynchrone alimentée par le réseau -----</b>	<b>174</b>
4.5.1. Réduction de l'espace de représentation (sélection des $d'$ paramètres) -----	175
4.5.1.1. Dans le cas du mode <i>supervisé</i> -----	175
4.5.1.1.1. Sélection par la méthode SBS -----	176
4.5.1.1.2. Sélection par la méthode AG -----	178
4.5.1.1.3. Comparaison avec des données non normalisées -----	180
4.5.1.1.4. Conclusion -----	182
4.5.1.2. Dans le cas du mode <i>non supervisé</i> -----	183
4.5.2. Décision par la règle des $k - ppv$ -----	186
4.5.2.1. Etude de la sensibilité sur $k$ (nombre de plus proches voisins) -----	186
4.5.2.2. Phase d'exploitation : classification pour toutes nouvelles observations $X_u$ -----	188
4.5.2.2.1. Classification par la règle des $k - ppv$ exclusive pour un ensemble d'apprentissage non normalisé -----	189
4.5.2.2.2. Classification par la règle des $k - ppv$ exclusive pour un ensemble d'apprentissage normalisé -----	190
4.5.2.2.3. Classification par la règle des $k - ppv$ non exclusive pour un ensemble d'apprentissage normalisé -----	192
4.5.2.3. Actualisation de la connaissance -----	194
4.5.2.4. Estimation de l'évolution des modes de fonctionnements -----	195
4.5.2.4.1. Classification sans l'utilisation de l'interpolation et de la prédiction -----	201
4.5.2.4.2. Classification avec l'utilisation de l'interpolation et de la prédiction -----	202
4.5.2.4.3. Conclusion -----	203
<b>4.6. Diagnostic par RdF d'une machine asynchrone alimentée par onduleur de tension -----</b>	<b>203</b>
4.6.1. Détection des défauts sur la machine asynchrone -----	204
4.6.1.1. Réduction de l'espace de représentation -----	205
4.6.1.2. Décision par la règle des $k - ppv$ non exclusive -----	207
4.6.1.2. Suivi d'évolution par approche de Kalman -----	208
4.6.2. Détection des défauts sur l'onduleur de tension -----	211
4.6.2.1. Présentation des caractéristiques de l'onduleur de tension -----	212
4.6.2.2. Sélection des paramètres -----	214
4.6.2.3. Classification par la règle des $k - ppv$ non exclusive -----	216
<b>4.7. Conclusion -----</b>	<b>217</b>
<b>CONCLUSION GENERALE -----</b>	<b>219</b>
<b>BIBLIOGRAPHIE -----</b>	<b>223</b>
<b>ANNEXE -----</b>	<b>245</b>

# NOTATIONS

## Notations liées à la machine asynchrone et au convertisseur de puissance

---

[C]	Matrice de Concordia (repère diphasé de Park fixe lié au stator)
$C_{em}$	Couple électromagnétique
$C_i$	Capacité idéale anode-cathode
$C_r$	Couple résistant imposé par la charge du moteur
$C_t$	Capacité thermique
$C_u$	Couple utile en sortie du moteur
$\Delta\theta$	Echauffement de la machine asynchrone
$\Delta\theta_m$	Echauffement maximal de la machine asynchrone
ESL	Inductance équivalente série, identique à $L_{se}$ , sa valeur est de quelques nH et est indépendante de la fréquence.
ESR	Résistance équivalente série représentant toutes les pertes dans les condensateurs (fonction de la fréquence).
$f_r$	Fréquence de rotation du rotor $f_r = \frac{\omega}{2\pi} = \frac{(1-g)f_s}{p}$
$f_s$	Fréquence des courants statoriques $f_s = \frac{\omega_s}{2\pi}$
$f_t$	Coefficient de frottement
$g$	Glissement du moteur $g = \frac{p\omega_r}{\omega_s}$
$I_{Forward}$	Courant direct de la diode
$i_{ra}, i_{rb}, i_{rc}$	Courants rotoriques suivants les axes $a_r, b_r$ et $c_r$
$i_{sa}, i_{sb}, i_{sc}$	Courants statoriques suivants les axes $a_s, b_s$ et $c_s$
$I_{s\alpha,\beta}, I_{r\alpha,\beta}$	Courant statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$i_{sd,q}, i_{rd,q}$	Courant statorique (respectivement rotorique) suivant l'axe d (ou q)



$J$	Moment d'inertie
$[L]$	Matrice inductance
$L_{fs}$	Inductance de fuite au stator
$L_m$	Inductance mutuelle entre le stator et le rotor
$l_{ms}$	Inductance magnétisante
$L_r$	Inductance propre cyclique du rotor
$L_s$	Inductance propre cyclique du stator
$L_{se}$	Inductance équivalente série des connexions et des enroulements. Elle dépend de la technologie de fabrication.
$n$	Vitesse de rotation du rotor (tr/min)
$N_r$	Nombre d'encoches (ou de barres) au rotor
$n_s$	Vitesse de synchronisme de la machine asynchrone (tr/min)
$N_s$	Nombre d'encoches au stator
$p$	Nombre de paires de pôles
$P_{cste}$	Pertes de la machine déduites à vide (regroupent les pertes mécaniques et les pertes fer au stator)
$P_{jr}$	Pertes par effet joules rotoriques
$P_{js}$	Pertes par effet joules statoriques
$P_{tot}$	Pertes totales de la machine
$\varphi_{sa}, \varphi_{sb}, \varphi_{sc}$	Flux statorique suivant les $a_s, b_s$ et $c_s$
$\varphi_{sr}, \varphi_{sr}, \varphi_{sr}$	Flux rotoriques suivant les $a_r, b_r$ et $c_r$
$\varphi_{sd,q}, \varphi_{rd,q}$	Flux statorique (respectivement rotorique) suivant l'axe d (ou q)
$\varphi_{s\alpha,\beta}, \varphi_{r\alpha,\beta}$	Flux statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$[R]$	Matrice résistive
$R_p$	Résistance parallèle représentant les pertes diélectriques et les fuites entre les deux électrodes
$R_s, R_r$	Résistance statorique, respectivement rotorique
$R_{ra}, R_{rb}, R_{rc}$	Résistance rotorique de la phase $a_r, b_r, c_r$
$R_{sa}, R_{sb}, R_{sc}$	Résistance statorique de la phase $a_s, b_s, c_s$

$R_{se}$	Résistance série des connexions et des armatures
$R_t$	Résistance thermique
$\sigma$	Coefficient de fuite total ( $\sigma = 1 - \frac{L_m^2}{L_s L_r}$ )
$\tau_{th}$	Constante de temps thermique
$\theta$	Angle électrique entre le rotor et le stator
$\theta_r$	Angle électrique entre le rotor et l'axe d
$\theta_s$	Angle électrique entre le stator et l'axe d
$T_e$	Période temporelle du signal
$T_s, T_r$	Constante de temps électrique du stator, respectivement du rotor
$[T_{dq0/abc}]$	Matrice de transformation de Park
$V_{ra}, V_{rb}, V_{rc}$	Tensions rotoriques suivants les axes $a_r, b_r$ et $c_r$
$V_{sa}, V_{sb}, V_{sc}$	Tensions statoriques suivants les axes $a_s, b_s$ et $c_s$
$V_{s\alpha,\beta}, V_{r\alpha,\beta}$	Tension statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$V_{sd,q}, V_{rd,q}$	Tension statorique (respectivement rotorique) suivant l'axe d (ou q)
$\omega$	Vitesse angulaire du rotor
$\omega_n$	Pulsation naturelle du circuit C-ESR-ESL
$\omega_r$	Pulsation des courants rotoriques $\omega_r = \frac{\omega_s}{p} - \omega$
$\omega_s$	Pulsation des courants statoriques
$[X_r]$	Notation vectorielle pour un système triphasé au rotor
$[X_s]$	Notation vectorielle pour un système triphasé au stator
$\underline{x}$	Notation complexe
$X_{ra,b,c}$	Composantes rotoriques de la grandeur $X_r$ de la phase a (respectivement b, c)
$X_{sa,b,c}$	Composantes statoriques de la grandeur $X_s$ de la phase a (respectivement b, c)
$Z_{th}$	Impédance thermique de la diode

## Notations liées à la Reconnaissance des Formes

---

$\mathbf{A}$	Notation matricielle
$\det(\mathbf{A}),  \mathbf{A} $	Déterminant de la matrice $\mathbf{A}$
$\mathbf{A}^{-1}$	Matrice inverse de la matrice $\mathbf{A}$
$\mathbf{A}^t$	Matrice transposée de la matrice $\mathbf{A}$
$trace(\mathbf{A})$	Trace de la matrice $\mathbf{A}$
$\alpha$	Un paramètre extrait de l'analyse des signaux et composant le vecteur forme initial
$\beta$	Paramètre définissant la largeur de bande de la fonction $\pi$
$C_d^{d'}$	Combinaison de $d'$ éléments parmi $d$
CA	Critère d'arrêt pour chaque branche de l'arbre de décision
$C_m$	Compacité moyenne des classes définies dans $X_a$
$C_r$	Compacité des observations rejetées
$d(\underline{X}, \underline{Y})$	Calcul de la distance entre deux vecteurs $\underline{X}$ et $\underline{Y}$
$d_E^2(\underline{X}, \underline{Y})$	Distance Euclidienne
$d_M^2(\underline{X}, \underline{Y})$	Distance de Mahalanobis
$d(\underline{X}_u, m_c)$	Distance entre une observation $\underline{X}_u$ et le centre de gravité $m_c$ de la classe $\Omega_c$
$d_{seuilc}$	paramètre qui définit la position des points d'inflexion
$d$	Dimension de l'espace de représentation avant réduction
$d'$	Dimension de l'espace de représentation après réduction
E	Ensemble composé de $d$ paramètres appelé vecteur forme (avant réduction)
E'	Ensemble composé de $d'$ paramètres appelé vecteur forme optimal (après réduction)
$f(\underline{X} \Omega_c)$	Densité de probabilité de $\underline{X}$ dans $\Omega_c$
$F(V_k)$	Fonction d'adaptation définie pour chaque individu $k$ ( $k=1, P$ )
$f(\mathbf{x})$	Densité de mélange
$\Gamma$	Pourcentage d'inertie ou qualité de la représentation dans le sous – espace de dimension $d'$
$g$	Paramètre servant à varier la pente de la fonction $\pi$
$g_{kj}$	Gène codant la présence du paramètre $\alpha_j$ dans le vecteur $V_k$
$J_1, J_2, J_3$	Critères
$k$	Nombre de plus proches voisins d'une observation $\underline{X}_u$ (règle des $k - ppv$ )

$k_c$	Nombre de plus proches voisins de $\underline{X}_u$ (parmi $k$ ) dans la classe $\Omega_c$
$\lambda_j$	$j^{\text{ième}}$ valeur propre de $\Sigma$ ( $j = 1, \dots, d$ )
$\underline{m}$	Centre de gravité de l'ensemble
$m_j$	$j^{\text{ième}}$ composante de $\underline{m}$
$\underline{m}_c$	Centre de gravité de la classe $\Omega_c$
$\underline{m}_{50}$	Centre de gravité d'une classe composée de 50 échantillons (cas de données non normalisées)
$M$	Nombre de classes
$\mu_{\Omega_c}(\underline{X}_u)$	Fonction d'appartenance de l'observation $\underline{X}_u$ à la classe $\Omega_c$
$N$	Nombre de vecteurs dans l'ensemble d'apprentissage
$N_c$	Nombre d'échantillons dans la classe $\Omega_c$
$N_{ch}$	Niveau de charge ou de sévérité d'une nouvelle observation
$N_{new}$	Nombre d'observations rejetées contenues dans la classe $\Omega_{new}$
$N_{min}$	Nombre minimal d'observations rejetées pour considérer l'apparition d'une nouvelle classe
$\eta$	Paramètre servant à fixer le plateau de la fonction $\pi$
$nb\_kppv_c$	Nombre de $k$ plus proches voisins issus de la classe $\Omega_c$
$\Omega_0$	Classe fictive des vecteurs (ou individus) rejetés en ambiguïté
$\Omega_d$	Classe fictive des vecteurs rejetés en distance
$\Omega_c$	Dénomination de la classe 'c' (avec $c = 1, \dots, M$ )
$\Omega_{new}$	Nouvelle classe créée à partir des observations rejetées
$P$	Nombre d' <i>individus</i> constituant la <i>population</i> initiale
$P_m$	Probabilité de mutation pour un enfant
$\underline{P}_s$	Points de l'évolution du mode de fonctionnement ( <i>trajectoire moyenne</i> )
$\underline{P}_{smax}$	Points du seuil d'acceptation ( <i>trajectoire maximale</i> )
$P_r$	Nombre d'observations rejetées totales
$\text{Pr}(\Omega_c)$	Probabilité a priori de la classe $\Omega_c$
$\sigma_j^2$	Variance de la composante $j$ des $N$ vecteurs de $X_a$
$\sigma_c^2(\alpha)$	Variance de la composante $\alpha$ des $N_c$ vecteurs de la classe $\Omega_c$

$\sum_{i=1}^M ()$	Somme sur l'indice $i$
$\Sigma_w$	Matrice de variance intra – classe de l'ensemble
$\Sigma_B$	Matrice de variance inter – classes de l'ensemble
$\Sigma$	Matrice de variance – covariance de l'ensemble
$\Sigma_c$	Matrice de variance – covariance de la classe $\Omega_c$
Seuil(j) <sub>c</sub>	Frontière sur l'axe 'j' de la classe $\Omega_c$
$T$	Seuil de rejet de distance
$T_c$	Seuil de rejet de distance de la classe $\Omega_c$
$T_{ne}$	Seuil de rejet en ambiguïté pour la classification non exclusive
$T_{ra}$	Seuil minimal d'appartenance pour une classe
$T_{50}$	Seuil de rejet en distance d'une classe composée de 50 échantillons (cas de données non normalisées)
$u_j (j=1,d')$	Axes factoriels formant la base orthogonal de dimension $d'$
$V_{Fnorm}$	Vecteur forme optimal pour des données normalisées par rapport à la charge
$V_{F(AG)}$	Vecteur forme optimal définit par la méthode de sélection AG
$V_{F(SBS)}$	Vecteur forme optimal définit par la méthode de sélection SBS
$V_k$	Vecteur formé de <i>genes</i> , caractérisant chaque <i>individu k</i> ( $k=1,P$ )
$x_{ij}$	$j^{ième}$ composante de $\underline{X}_i$
$x_j$	$j^{ième}$ composante du vecteur forme ( $j = 1, \dots, d$ )
$\underline{X}$	Vecteur de dimension ( $d \times 1$ )
$\underline{X}_a$	Ensemble (ou matrice) des $N$ vecteurs formant l'ensemble d'apprentissage
$\underline{X}_{cv}$	$v^{ième}$ vecteur de la classe $\Omega_c$ ( $v = 1, \dots, N_c$ et $c = 1, \dots, M$ )
$\underline{X}_i$	$i^{ième}$ vecteur (ou échantillon) de l'ensemble d'apprentissage
$\underline{X}_i^p$	Projection de $\underline{X}_i$ dans la base formé par les $u_j$
$\underline{X}_t$	Ensemble test
$\underline{X}_u$	Nouvelle observation mesurée sur le processus
$\Delta$	Matrice diagonale des variances $\sigma_j^2$
$Z_c$	Zone de restriction (frontière) autour de la classe $\Omega_c$

## Introduction générale

Dans de nombreux secteurs industriels, la sûreté de fonctionnement est un enjeu majeur pour assurer une compétitivité optimale de l'outil de production. Les gains de productivité représentent un souci quotidien pour les dirigeants des entreprises. La compétitivité de ce secteur passe par la maîtrise indispensable de la disponibilité de l'outil de production, l'amélioration de son fonctionnement, la sécurité de ses utilisateurs et la réduction des coûts d'entretien. Il existe plusieurs notions permettant une optimisation des systèmes industriels qui s'intègrent à la discipline de la sécurité de fonctionnement : fiabilité, maintenabilité, sécurité et maîtrise des risques. Le diagnostic des défaillances de systèmes industriels, lorsqu'il est réalisé avec efficacité et s'il permet de détecter de façon précoce une dégradation, représente un des moyens pour contribuer à obtenir un meilleur gain de productivité. Sa vocation première est de détecter et de localiser une possible défaillance des matériels. De plus en plus, il devient une partie intégrante de la fonction de la maintenance. Devenue une discipline majeure, relayée sur le terrain par des technologies modernes (en outre l'intelligence artificielle), le diagnostic industriel est devenu une discipline incontournable dans le domaine de la sûreté de fonctionnement.

D'après la définition du Groupement de Recherche CNRS 'Modélisation, Analyse et Conduite des Systèmes dynamiques' (GdR MACS), un système complexe sujet à un diagnostic, peut être défini comme un ensemble, susceptible d'évoluer dans le temps, composé d'éléments interagissant entre eux et avec le milieu extérieur, réunis pour assurer une mission.

Tous ces systèmes technologiques sont susceptibles d'évoluer sous divers modes de fonctionnement. Ces modes ne sont pas obligatoirement tous connus, certains pouvant n'avoir été jamais observés sur le système. D'autre part, si certains de ces modes correspondent à un fonctionnement normal, d'autres présentent la caractéristique d'apparaître en cas de défaillances et sont en général peu tolérés, voire non tolérés. Pour la surveillance automatique de tels dispositifs complexes, il convient de développer des systèmes de diagnostic disposant d'une certaine faculté d'adaptation à des situations nouvelles (apprentissage), et permettant une reconnaissance fiable (décision) du mode de fonctionnement dans lequel le système évolue. Le diagnostic peut être défini comme un processus en trois phases : la détection d'un mode de fonctionnement, son identification et la localisation de sa cause.

Les travaux présentés dans ce mémoire portent sur le diagnostic d'un ensemble « convertisseur de puissance - machine asynchrone ».

S'intéresser au diagnostic des machines asynchrones peut sembler paradoxal. Si cette machine tend effectivement à supplanter la machine à courant continu dans la plupart des entraînements à vitesse variable, c'est en raison de ses nombreuses qualités, et principalement de par sa robustesse et son rapport poids/puissance.

Certaines utilisations de ces moteurs ne tolèrent pas les pannes intempestives. Ces pannes peuvent être d'origine mécanique (excentricité du rotor, défaut sur les accouplements, usure des roulements,...), électrique ou magnétique (court circuit du bobinage statorique, rupture de barre ou d'anneau, cassure de dents,...). Les imperfections peuvent aussi être dues à d'autres éléments de l'entraînement, comme les défauts d'alimentation provenant de la source d'alimentation (réseau ou convertisseur de puissance).

Tous ces défauts qui apparaissent dans la machine électrique ont des causes variées. Ils peuvent être dus au vieillissement des constituants du moteur, aux conditions d'utilisation ou bien encore, tout simplement à un défaut de fabrication dont l'effet serait imperceptible lors de la mise en service.

Afin d'éviter des arrêts intempestifs, l'application d'une maintenance périodique durant laquelle les éléments d'un système complexe sont remplacés après une durée d'utilisation finie semble coûteuse en temps d'arrêt et en produits. De plus, la durée de vie de certains de ces éléments peut ne pas être définie au préalable.

L'efficacité de la surveillance réalisée par des observateurs réside dans le choix des paramètres significatifs. La détection des défaillances se fait en comparant l'évolution de l'écart entre le modèle de la machine et le processus réel.

Le développement d'outils ou de capteurs de surveillance et de détection de pannes paraît une solution un peu plus coûteuse à l'investissement, mais elle s'amortit sur le long terme. Les signaux mesurables tels que les courants, les tensions, la vitesse, les vibrations ou bien encore la température peuvent fournir des informations significatives sur les défauts et ainsi servir à déterminer un ensemble de paramètres représentant les signatures de défauts du moteur. A partir de ces paramètres, la mise en place de méthodes décisionnelles peut permettre de concevoir des systèmes de surveillance performants. Les performances de cette approche de supervision sont étroitement liées à la pertinence des signatures de défauts déterminées et à la finesse d'analyse des mesures.

Dans ce contexte, le mémoire comporte quatre chapitres. Le premier présente les différentes défaillances qui peuvent se produire sur un ensemble 'convertisseur - machine asynchrone' ainsi que l'occurrence de chacun de ces défauts. Ce chapitre se décompose en

deux grandes parties. La première partie fait état des différentes origines (électriques, mécaniques,...) des défaillances qui peuvent se produire sur la machine asynchrone. Quant à la seconde partie, elle présente les différentes défaillances que peut subir un convertisseur de puissance (condensateur, IGBT, diodes,...). Pour finir, une synthèse de toutes les pannes citées auparavant sur les deux systèmes est faite (AMDE) en définissant quelles sont les causes et les conséquences de ces défauts.

Le deuxième chapitre, montre comment, à partir des seules grandeurs mesurables, les informations relatives aux défauts (signatures) de l'ensemble 'convertisseur – machine' peuvent être extraites. L'association, pour chaque type de défauts, d'une signature, va permettre de disposer d'indicateurs utiles à la discrimination de ces différentes défaillances.

Ces dernières années, la surveillance et la détection des défauts des moteurs électriques se sont rapprochées des techniques dites de l'intelligence artificielle (IA). Ces techniques s'appuient sur l'utilisation des indicateurs pour faire la supervision et le diagnostic du procédé. Elles apportent un avantage certain aussi bien au niveau de la fiabilité que de l'automatisation.

De ce fait, la deuxième partie du chapitre II présente les différentes techniques de surveillance issues de l'IA et appliquées au diagnostic des entraînements électriques. On retrouve dans la littérature, quatre méthodes appliquées dans ce domaine. L'application des systèmes experts permet de modéliser l'approche des experts dans un domaine spécifique. Ils permettent de formaliser leurs connaissances et fournissent un mécanisme pour les exploiter. La logique floue permet de gérer l'incertitude dans un système d'information [Zadeh, 65]. L'utilisation des réseaux de neurones et de la reconnaissance des formes présente une propriété intéressante, qui est, la faculté d'apprentissage [Dubuisson, 90]. Ainsi le système de diagnostic conçu à partir de ces deux techniques sera assez souple et évolutif, et pourra prendre en compte les caractéristiques de toute nouvelle installation.

Dans le troisième chapitre, une introduction à la reconnaissance des formes (RdF) pour le diagnostic est présentée. Une justification du choix de la RdF parmi les méthodes citées au paragraphe précédent pour réaliser la surveillance et la supervision d'un système, y est faite. De plus, nous montrons comment les principes de la RdF peuvent être mis en œuvre pour résoudre le problème du diagnostic d'un système complexe. Bien qu'inspiré de certaines particularités de la machine asynchrone, le développement de ce système de diagnostic s'est



effectué dans un souci de non spécificité (l'intérêt étant de pouvoir l'appliquer à n'importe quel système complexe moyennant un minimum de modification). Un système de reconnaissance des formes s'élabore en deux phases. La première phase, appelée *phase d'analyse*, consiste à définir le vecteur forme (ensemble de paramètres représentant la signature des défauts du système), déterminer l'espace de représentation et structurer les modes de fonctionnement en différentes classes. Pour cela, il sera utilisé et développé différents algorithmes de sélection de paramètres associés à des méthodes de classification automatique, appelée coalescence, lorsque l'on se place dans le cas d'un système *sans connaissance a priori* (mode non supervisé). Le problème de la non exhaustivité des modes de fonctionnement est abordé en utilisant une approche par filtre de Kalman, qui permet d'obtenir une estimation de l'évolution de ces modes. D'autre part, une règle de décision est mise en place pour rattacher toute nouvelle observation à l'une des classes existantes. Les incertitudes liées aux tâches d'apprentissage et de décision sont prises en compte par l'introduction d'une fonction d'appartenance associée à cette règle de décision. La deuxième phase, couramment appelée *phase d'exploitation*, consiste à décider ou non, d'affecter une nouvelle observation à l'une des classes connues, en vertu de la règle de décision. Les observations rejetées sont ensuite analysées pour mettre en évidence l'apparition ou non d'une ou plusieurs nouvelles classes correspondant à des modes de fonctionnement encore inconnu de la base de connaissance (ensemble d'apprentissage).

Le quatrième chapitre est consacré à l'application de notre système de diagnostic aux machines asynchrones. Une description du banc expérimental ainsi que des différents modes de fonctionnement étudiés (sain, défaillants), est présentée. Les défaillances correspondent à des cassures de barres, un déséquilibre d'alimentation, des défauts de roulements, et ceci, pour différents niveaux de charge en alimentation réseau et/ou onduleur de tension. Une section est consacrée à la normalisation des données. Celle-ci a pour but d'éliminer la dispersion d'un mode de fonctionnement due à la variation de la charge. De ce fait, la discrimination des différents modes ainsi que des défaillances possibles est facilitée. La deuxième partie est consacrée au diagnostic d'une machine alimentée par le réseau. Les algorithmes de sélection de paramètres du vecteur forme présentés au chapitre III sont mis en application dans le cas, soit d'une *connaissance a priori* du système, soit *sans connaissance a priori*. Les différentes règles de décision sont alors comparées sur des essais en présence ou non de défauts et les observations rejetées sont analysées pour déterminer une éventuelle apparition d'un nouveau mode de fonctionnement. Le suivi d'évolution développé à partir de

l'approche par filtre de Kalman va permettre la prédiction des modes de fonctionnement contenus dans l'ensemble d'apprentissage et ainsi déterminer les états futurs de ces différents modes. La dernière partie est consacrée au diagnostic d'une machine alimentée par un onduleur de tension. La démarche dans cette partie sera la même que dans la partie précédente. Dans un premier temps, la discrimination se situe au niveau des différents défauts que peut subir une machine asynchrone et, dans un second temps, la discrimination se situe au niveau de la différenciation entre un fonctionnement sain de l'ensemble onduleur-machine et un fonctionnement avec une défaillance de l'onduleur (vieillesse des condensateurs de filtrage).



# CHAPITRE I

Synthèse des défaillances possibles sur un ensemble  
convertisseur – machine asynchrone



## ***1.1. Introduction***

Nos méthodes de diagnostic ont été appliquées aux entraînements électriques asynchrones. En effet, le moteur asynchrone est un des plus utilisés dans le domaine des puissances supérieures à quelques kilowatts car il présente de nombreux avantages tels que sa puissance massique, sa robustesse, sa facilité de mise en œuvre, son faible coût, etc... L'apparition dans les années 1980 des variateurs permettant de faire varier la fréquence de rotation dans une large gamme a grandement favorisé son développement. En effet, il entre dans la conception de nombreux procédés industriels associant des convertisseurs statiques et des machines électriques (traction électrique, laminoirs, levage, pompage, etc...).

Ce premier chapitre porte sur la synthèse des défauts que sont susceptibles de présenter ces systèmes.

Nous allons dans un premier temps décrire les différents constituants de l'ensemble convertisseur machine : la machine asynchrone [Hamdi, 94] et l'onduleur de tension servant à son alimentation.

Ensuite les défaillances les plus courantes pouvant apparaître aussi bien dans la machine asynchrone que dans l'onduleur seront énumérées. Ces défaillances peuvent provoquer des pannes et par conséquent des arrêts intempestifs préjudiciables au bon fonctionnement du système ainsi qu'à la sécurité des personnes.

## ***1.2. Constitution de la machine asynchrone***

Cette première partie porte sur une présentation succincte des machines asynchrones triphasées de petites et moyennes puissances (50 à 200kW). Ceci sera illustré au moyen de la machine asynchrone à cage d'écureuil utilisée lors de nos essais expérimentaux.

### **1.2.1. Constitution de la machine**

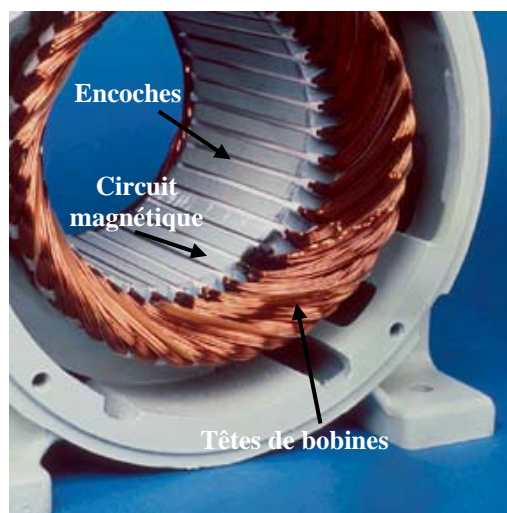
La machine asynchrone est constituée des principaux éléments suivants :

- le stator (partie fixe) constitué de disques en tôles magnétiques portant les enroulements chargés de magnétiser l'entrefer.

- le rotor (partie tournante) constitué de disques en tôles magnétiques empilés sur l'arbre de la machine portant un enroulement injecté.
- les organes mécaniques permettant la rotation du rotor et le maintien des différents sous-ensembles.

### 1.2.1.1. Le stator

Il est constitué d'un enroulement bobiné réparti dans les encoches du circuit magnétique. Ce circuit magnétique est constitué d'un empilage de tôles dans lesquelles sont découpées des encoches parallèles à l'axe de la machine (figure 1.1). Le bobinage statorique peut se décomposer en deux parties : les conducteurs d'encoches et les têtes de bobines. Les conducteurs d'encoches permettent de créer dans l'entrefer le champ magnétique à l'origine de la conversion électromagnétique. Les têtes de bobines permettent, quant à elles, la fermeture des courants en organisant la circulation judicieuse des courants d'un conducteur d'encoche à l'autre. L'objectif est d'obtenir à la surface de l'entrefer une distribution de courant la plus sinusoïdale possible, afin de limiter les ondulations du couple électromagnétique.

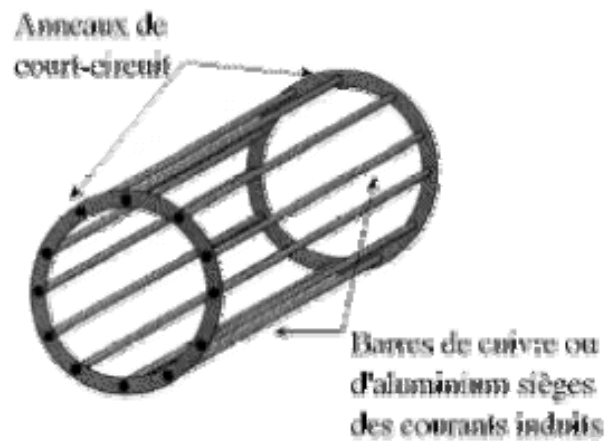


**Fig. 1.1.** Photo du stator d'une machine asynchrone

### 1.2.1.2. Le rotor

Dans le rotor à cage, les anneaux de court-circuit permettent la circulation des courants d'un conducteur d'encoche (barre rotorique) à l'autre. Ces barres conductrices sont régulièrement réparties, et constituent le circuit du rotor (figure 1.2). Cette cage est insérée à l'intérieur d'un circuit magnétique constitué de disques en tôles empilés sur l'arbre de la

machine analogue à celui du moteur à rotor bobiné (figure 1.3). Dans le cas de rotors à cage d'écureuil, les conducteurs sont réalisés par coulage d'un alliage d'aluminium, ou par des barres massives de cuivre préformées et frettés dans les tôles du rotor. Il n'y a généralement pas, ou très peu, d'isolation entre les barres rotoriques et les tôles magnétiques, mais leur résistance est suffisamment faible pour que les courants de fuite dans les tôles soient négligeables, sauf lorsqu'il y a une rupture de barre [Kerszenbaum, 84], [Muller, 03], [Walliser, 94]. Le moteur à cage d'écureuil est beaucoup plus simple à construire que le moteur à rotor bobiné et, de ce fait, son prix de revient est inférieur. De plus, il dispose d'une plus grande robustesse. Il constitue la plus grande partie du parc de moteurs asynchrones actuellement en service.



**Fig. 1.2.** Vue schématique du rotor à cage d'écureuil (conducteurs d'encoches (barres) et anneaux de court-circuit)



Tôle magnétique (ici double encoche)



Coupe de face d'un rotor complet

**Fig. 1.3.** Circuit magnétique constituant le rotor



### 1.2.1.3. Les organes mécaniques

La carcasse sert de support, elle joue le rôle d'enveloppe et assure la protection contre l'environnement extérieur.

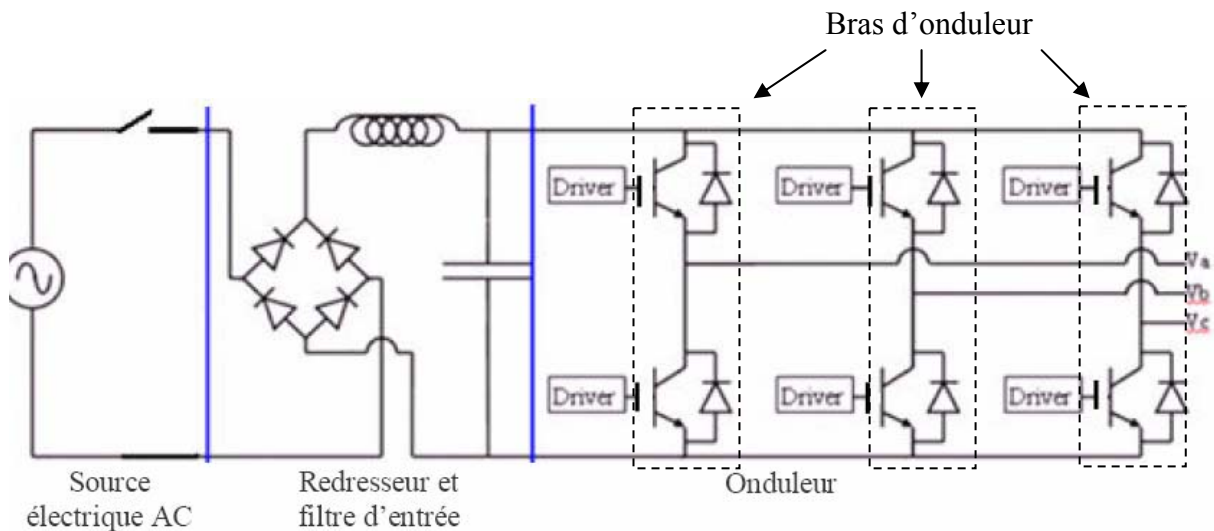
L'arbre est un organe de transmission. Il comprend une partie centrale qui sert de support au corps du rotor et un bout d'arbre sur lequel est fixé un demi accouplement. Il est généralement constitué en acier moulé ou forgé. Son dimensionnement est fonction des efforts de flexion (force centrifuge qui s'exerce sur lui, attraction magnétique radiale, etc...), des efforts radiaux et tangentiels dus aux forces centrifuges, des efforts de torsion (couple électromagnétique transmis en régime permanent, transitoire). Il est supporté par un ou plusieurs paliers. Ces paliers soutiennent le rotor et assurent la libre rotation. Le second palier est libre pour assurer les dilatations thermiques de l'arbre. Une isolation électrique de l'un des paliers assure l'élimination des courants dans l'arbre dû aux dissymétries des réluctances du circuit magnétique. Ils sont généralement à roulements pour les machines de petite et moyenne puissance.

Afin d'assurer la variation de vitesse des moteurs électriques mais aussi la conversion de l'énergie, de nombreuses machines se voient associer un convertisseur. Ainsi, de nos jours, on trouve dans un grand nombre de procédés industriels, des associations entre des convertisseurs statiques et des machines électriques (pompage, levage,...). Cela permet de fournir aux moteurs des tensions et des courants de fréquence variable servant à commander la machine en couple ou en vitesse.

## 1.3. Constitution de l'alimentation d'un moteur asynchrone

Dans le cas des moteurs asynchrones, la vitesse de rotation du rotor dépend de la fréquence statorique  $f_s$  (fréquence de la tension d'alimentation du moteur) et de la fréquence des courants rotoriques (donc de la charge). Un convertisseur statique permet de faire varier l'amplitude et la fréquence de la tension d'alimentation et donc de faire varier la vitesse de la machine.

Plusieurs stratégies de commande peuvent être utilisées. Les commandes scalaires [Bonal, 98] dont une version simplifiée assure un rapport  $\frac{U}{f_s}$  constant (avec  $f_s$ , la fréquence des courants statoriques et U, la tension efficace d'alimentation), les commandes vectorielles à flux orienté [Bonal, 98], [Bonal, 97]. Le schéma suivant présente la structure générale qui permet de réaliser la variation de vitesse.



**Fig. 1.4.** Structure de l'alimentation d'un moteur pour la variation de vitesse

Les différents éléments constituant la chaîne d'alimentation sont :

➤ **Source électrique AC**

La source électrique alternative est généralement obtenue par le réseau triphasé électrique.

➤ **Redresseur**

Il permet de transformer une tension alternative en tension continue.

➤ **Filtrage**

Il élimine les phénomènes d'ondulation de la tension en sortie du redresseur. Ce filtre, associé au pont redresseur, constitue une source de tension continue servant à alimenter le bus continu à l'entrée de l'onduleur.

### ➤ Onduleur

Il permet de transformer une tension continue en une tension alternative d'amplitude et de fréquence variables. Dans notre application, le rapport  $\frac{U}{f_s}$  est maintenu constant car on souhaite conserver les performances en régime permanent du moteur, c'est-à-dire disposer du couple maximal en permanence quelque soit la vitesse. Pour cela il faut maintenir le flux statorique  $\Phi_s$  constant et donc cela impose, en première approximation, de maintenir  $U/f_s$  constant jusqu'à la vitesse nominale. Il génère une succession d'impulsions de tension, réalisée par l'intermédiaire de la commande des interrupteurs. Sur la figure 1.4, il s'agit d'un onduleur à trois demi-ponts (ou bras), qui est la structure la plus répandue, car elle offre un compromis intéressant entre le coût et les performances.

La commande la plus répandue pour ce type de convertisseur est la modulation de largeur d'impulsions (M.L.I.) ou Pulse Width Modulation (P.W.M). Cette modulation est une technique de découpage permettant de générer des formes de courant arbitraires. Dans notre cas, elles seront quasi sinusoïdales. L'onduleur va générer des tensions ou des courants de fréquence variable riches en harmoniques. Celles-ci sont à l'origine de pertes supplémentaires (fer et Joule) s'ajoutant aux pertes 'normales' qu'on aurait si les tensions et les courants étaient sinusoïdaux. L'accroissement de l'échauffement dû à ces pertes supplémentaires et les contraintes en tension risquent de diminuer la durée de vie du moteur.

Toutes ces contraintes augmentent le risque de défaillances. De nombreux défauts peuvent apparaître sur les entraînements électriques. Ils se répartissent en trois catégories : les défauts qui se produisent dans la machine électrique (défauts de roulement, rupture de barre), ceux qui se produisent dans la chaîne d'entraînement (défaut de la charge, accouplement) et ceux intervenant au niveau de l'alimentation (déséquilibre d'alimentation, défaut convertisseur). Ces défauts sont liés car ils peuvent interagir. Dans les parties suivantes, nous nous intéresserons donc aux différentes défaillances affectant le convertisseur et le moteur.

## 1.4. Présentation des différentes défaillances du moteur asynchrone à cage d'écureuil

Les défaillances peuvent être d'origines diverses, électriques, mécaniques ou bien encore magnétiques. Leurs causes sont multiples et peuvent se classer en trois groupes [O'Donnell, 85] :

- les générateurs de pannes ou initiateurs de défauts : surchauffe du moteur, défaut électrique (court-circuit), survoltage d'alimentation, problème d'isolation électrique, usure des éléments mécaniques (roulements à billes), rupture de fixations, etc.
- les amplificateurs de défauts : surcharge fréquente, vibrations mécaniques, environnement humide, échauffement permanent, mauvais graissage, vieillissement, etc.
- les vices de fabrication et les erreurs humaines : défauts de fabrication, composants défectueux, protections inadaptées, mauvais dimensionnement de la machine, etc.

Une étude statistique, effectuée en 1988 par une compagnie d'assurance allemande de systèmes industriels [Allianz, 88] sur les pannes des machines asynchrones de moyenne puissance (de 50 kW à 200kW) a donné les résultats suivants (figure 1.5) :

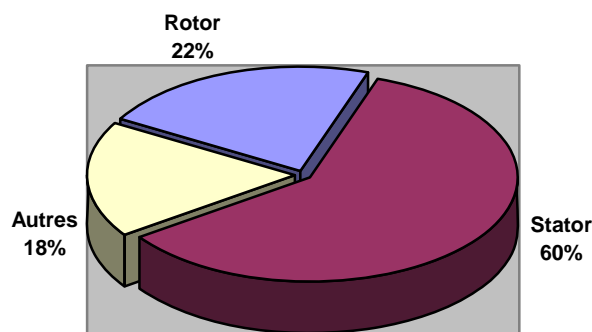
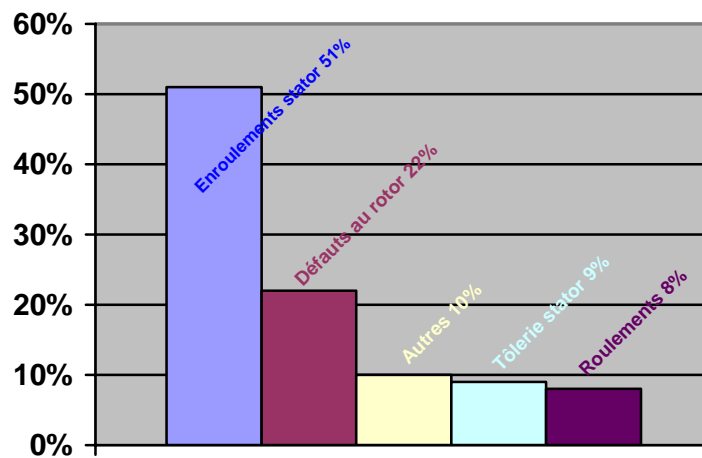


Fig. 1.5. Proportion des défauts

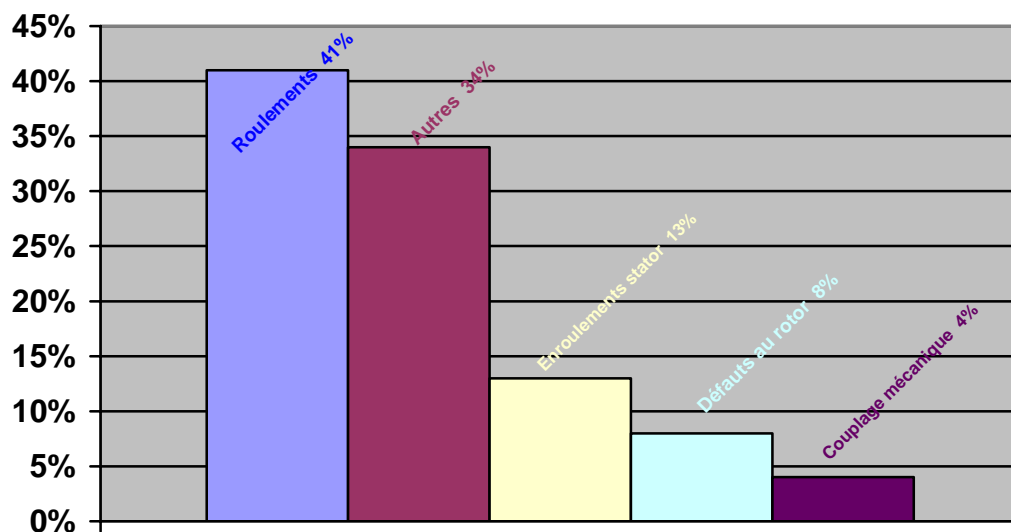
D'autre part, les mêmes études montrent qu'entre 1973 et 1988, les pannes au stator sont passées de 78% à 60% et au rotor de 12% à 22%. Ces variations sont dues à l'amélioration

des isolants sur cette période. La répartition des pannes dans les différentes parties du moteur est la suivante :



**Fig. 1.6.** Répartition des pannes sur les machines de faibles et moyennes puissances

Une autre étude statistique [Thorsen, 95] faite sur des machines de grande puissance (de 100 kW à 1 MW) donne les résultats suivants :



**Fig. 1.7.** Répartition des pannes sur les machines de fortes puissances

Les contraintes mécaniques sont plus grandes pour ces types de machines ce qui explique le taux élevé des pannes dues aux roulements. Celles-ci exigent une maintenance mécanique accrue.

### 1.4.1. Défaillances au rotor [Bonnett, 92], [Nandi, 99], [Raison, 00]

Pour le rotor, les défaillances sont essentiellement dues à un problème :

- thermique (surcharge,...)
- électromagnétique (force en  $B^2(t)$ ,...)
- résiduel (déformation,...)
- dynamique (arbre de transmission,...)
- mécanique (roulement,...)
- environnemental (agression,...)

Les défauts qui sont les plus récurrents, localisés au niveau du rotor, peuvent être définis comme suit :

- Rupture de barres
- Rupture d'une portion d'anneau de court-circuit
- Excentricité statique et dynamique

#### 1.4.1.1. Ruptures de barres [Naït-Saïd, 00], [Elkasabgy, 92]

La cassure ou rupture de barre est un des défauts les plus fréquents au rotor. Elle peut se situer soit au niveau de son encoche soit à l'extrémité qui la relie à l'anneau rotorique. La détérioration des barres réduit la valeur moyenne du couple électromagnétique et augmente l'amplitude des oscillations, qui elles-mêmes provoquent des oscillations de la vitesse de rotation, ce qui engendre des vibrations mécaniques et donc, un fonctionnement anormal de la machine. La grande amplitude de ces oscillations accélère la détérioration de la machine. Ainsi, le couple diminue sensiblement avec le nombre de barres cassées induisant un effet cumulatif de la défaillance. L'effet d'une cassure de barre croît rapidement avec le nombre de barres cassées.

#### 1.4.1.2. Ruptures d'anneaux

La rupture de portion d'anneau est un défaut qui apparaît aussi fréquemment que la cassure de barres. Ces ruptures sont dues soit à des bulles de coulées ou aux dilatations différentielles entre les barres et les anneaux.

Comme il est difficile de le détecter [Yahoui, 95a], ce défaut est généralement groupé, voir confondu, avec la rupture de barres dans les études statistiques. Ces portions d'anneaux de court-circuit véhiculent des courants plus importants que ceux des barres rotoriques. De ce

fait, un mauvais dimensionnement des anneaux, une détérioration des conditions de fonctionnement (température, humidité,...) ou une surcharge de couple et donc de courants, peuvent entraîner leur cassure.

La rupture d'une portion d'anneau déséquilibre la répartition des courants dans les barres rotoriques et de ce fait, engendre un effet de modulation d'amplitude sur les courants statoriques similaire à celui provoqué par la cassure de barres.

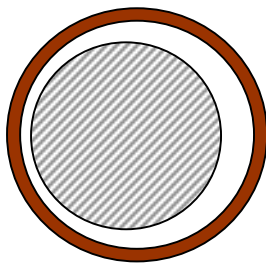
### 1.4.1.3. Excentricité statique et dynamique

Parfois, la machine électrique peut être soumise à un décentrement du rotor, se traduisant par des oscillations de couple (décalage entre le centre de rotation de l'arbre et le centre du rotor, figure 1.8). Ce phénomène est appelé excentricité (statique et dynamique) dont l'origine peut être liée à un positionnement incorrect des paliers lors de l'assemblage, à un défaut roulement (usure), à un défaut de charge, ou à un défaut de fabrication (usinage).

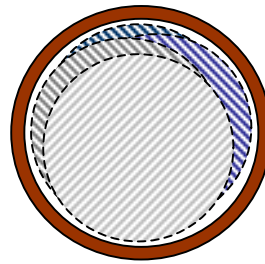
Trois cas d'excentricité [Sobczyk, 98], [Devanneaux, 2001] sont généralement distingués :

- l'excentricité statique, le rotor est déplacé du centre de l'alésage stator mais tourne toujours autour de son axe
- l'excentricité dynamique, le rotor est positionné au centre de l'alésage mais ne tourne plus autour de son axe
- l'excentricité qu'on pourrait qualifier de 'mixte', associant les deux cas précédemment cités

On peut représenter l'excentricité statique et dynamique de la manière suivante :



Excentricité statique



Excentricité dynamique

(plusieurs positions du rotor au cours de la rotation)

**Fig. 1.8.** Modélisation schématique de l'excentricité statique et dynamique

Ce défaut modifie le comportement magnétique ainsi que mécanique de la machine. En effet, l'augmentation de l'excentricité dans l'entrefer induit une augmentation des forces

électromagnétiques qui agissent directement sur le noyau statorique ainsi que l'enroulement correspondant, ce qui engendre une dégradation de son isolation.

D'autre part, cette augmentation peut avoir comme conséquence des frottements entre le stator et le rotor en raison des forces d'attraction magnétique qui déséquilibrent le système. Ceci donne naissance à des niveaux de vibration considérables dans les enroulements.

### **1.4.2. Défaillances au stator [Bonnett, 92]**

Pour le stator, les défaillances sont principalement dues à un problème :

- thermique (surcharge,...)
- électrique (diélectrique,...)
- mécanique (bobinage,...)
- environnemental (agression,...)

Les défauts qui sont les plus récurrents, localisés au niveau du stator, peuvent être définis comme suit :

- défaut d'isolant
- court-circuit entre spires
- court-circuit entre phases
- court-circuit phase/bâti
- déséquilibre d'alimentation
- défaut de circuit magnétique

#### **1.4.2.1. Défauts d'isolant dans un enroulement**

La dégradation des isolants dans les enroulements peut provoquer des courts-circuits. En effet, les différentes pertes (Joule, fer, mécanique,...) engendrent des phénomènes thermiques se traduisant par une augmentation de la température des différents constituants du moteur. Or les matériaux d'isolation ont une limite de température, de tension et mécanique. De ce fait, si l'environnement de travail d'un matériau d'isolation dépasse une de ces limites, ce matériau se dégrade de manière prématurée ou accélérée, puis finit par ne plus assurer sa fonction. Dans ce cas, un court-circuit peut apparaître dans l'enroulement concerné. Les différentes causes pour ce type de défaut sont :



- dégradation de l'isolant à la fabrication.
- tension de l'enroulement supérieure à la limite du matériau d'isolation.
- courant élevé dans l'enroulement dû à un court-circuit, un défaut du convertisseur, une surcharge. Ceci entraîne une élévation de la température dégradant prématurément le matériau d'isolation.
- vibrations mécaniques.
- vieillissement naturel des isolants. Tous les matériaux isolants ont une durée de vie limitée. Même dans une utilisation 'normale', l'isolant finit naturellement par se dégrader.
- fonctionnement dans un environnement sévère.

#### **1.4.2.2. Court-circuit entre spires [Joksimovic, 00]**

Un court-circuit entre spires de la même phase est un défaut assez fréquent. Cette défaillance a pour origine un ou plusieurs défauts d'isolant dans l'enroulement concerné. Il entraîne une augmentation des courants statoriques dans la phase affectée, une légère variation de l'amplitude sur les autres phases, modifie le facteur de puissance et amplifie les courants dans le circuit rotorique. Ceci a pour conséquence une augmentation de la température au niveau du bobinage et, de ce fait, une dégradation accélérée des isolants, pouvant provoquer ainsi, un défaut en chaîne (apparition d'un 2<sup>ème</sup> court-circuit). Par contre, le couple électromagnétique moyen délivré par la machine reste sensiblement identique hormis une augmentation des oscillations proportionnelle au défaut.

#### **1.4.2.3. Court-circuit entre phases**

Ce type de défaillance peut arriver en tout point du bobinage, cependant les répercussions ne seront pas les mêmes selon la localisation. Cette caractéristique rend difficile une analyse de l'incidence de ce défaut sur le système.

L'apparition d'un court-circuit proche de l'alimentation entre phases, induirait des courants très élevés qui conduiraient à la fusion des conducteurs d'alimentation et/ou à la disjonction par les protections. D'autre part, un court-circuit proche du neutre entre deux phases engendre un déséquilibre sans provoquer la fusion des conducteurs.

Les courants statoriques sont totalement déséquilibrés et ce déséquilibre est proportionnel au défaut qui apparaît. Les courants dans les barres ainsi que dans les anneaux sont augmentés

lors de l'apparition de ce défaut. La détection de ce type de défaut peut reposer sur le déséquilibre des courants de phases.

#### **1.4.2.4. Court-circuit phase/bâti**

Le bâti a généralement un potentiel flottant, mais pour des raisons de liaisons mécaniques, il est souvent relié à la masse. Si le potentiel est flottant, un court-circuit entre l'enroulement et le bâti n'a pas d'importance du point de vue matériel, excepté les effets capacitifs, le bâti prend alors le potentiel de l'enroulement à l'endroit du court-circuit. Par contre, au niveau de la sécurité des personnes, ce type de défaut peut être très dangereux et il est alors nécessaire de mettre en place des dispositifs de protection (disjoncteurs différentiels).

En présence de ce type de défaillance, la tension de la phase concernée ne change pas. Cependant le courant circulant dans cette phase augmente avec la réduction de la résistance et de l'inductance. Cette augmentation du courant se traduit par une augmentation de la température pouvant entraîner des défauts d'isolant dans l'enroulement. De plus, cette défaillance va générer une composante homopolaire entraînant l'apparition d'un couple pulsatoire. Une mesure du courant de fuite pourrait permettre de détecter ce type de défaut.

#### **1.4.2.5. Défautes de circuit magnétique**

Ces défauts aboutissent dans la plupart des cas à une dissymétrie au niveau du fonctionnement de la machine, qui à son tour peut accentuer le problème par des phénomènes de surchauffe, de surtension, d'élévation importante du courant, etc.

### **1.4.3. Défaillances mécaniques**

#### **1.4.3.1. Défautes roulements [Thomson, 99], [Salles, 97], [Cabanas, 96]**

Les roulements à billes jouent le rôle d'interface électromécanique entre le stator et le rotor. En outre, ils représentent l'élément de maintien de l'axe de la machine permettant d'assurer une bonne rotation du rotor. Dans l'article [Bonnet, 93], l'auteur présente la plupart des défauts survenant dans les roulements des moteurs à induction ainsi que les raisons de leur vieillissement.

Comme il a été présenté précédemment, ce type de défaut est le plus fréquent sur les machines de fortes puissances. Il est généralement lié à l'usure du roulement et plus précisément une dégradation des billes, ou de la bande de roulement. Ses causes possibles sont :

- l'usure due au vieillissement
- la température de fonctionnement élevée
- la perte de lubrification
- l'huile contaminée (par des paillettes métalliques issues de la dégradation des billes ou de la bande de roulement)
- le défaut de montage
- les courants d'arbres (Shaft Current)

Les conséquences directes de cette défaillance sur les roulements sont :

- des trous dans les gorges de roulement intérieures et extérieures
- l'ondulation de leur surface de roulement
- l'attaque des billes
- la corrosion due à l'eau
- défaut de graissage, problème dû à la température
- décollement, effritement de surface provoqué par une surcharge

Sur le système, ce type de défaut se traduit par des oscillations du couple de charge, une apparition de pertes supplémentaires et un jeu entre la bague interne et la bague externe du roulement entraînant des vibrations par les déplacements du rotor autour de l'axe longitudinale de la machine. Dans le cas le plus défavorable, la présence d'un roulement défectueux peut amener au blocage du moteur.

#### **1.4.3.2. Autres défaillances mécaniques**

Au stator, il n'y a pas de pièces mobiles donc a priori pas de défaillances mécaniques. Cependant, il peut apparaître des phénomènes d'oxydation liés à l'environnement de la machine et plus précisément au taux de salinité qui influe sur l'étanchéité et les contacteurs.

### ***1.5. Présentation des différentes défaillances sur un convertisseur de puissance***

A notre connaissance, il n'existe pas d'étude statistique proprement dite sur les pannes des convertisseurs de puissance. Il a été montré que les composants présentant le taux de défaillance les plus élevés sont généralement les condensateurs électrolytiques [Venet, 93].

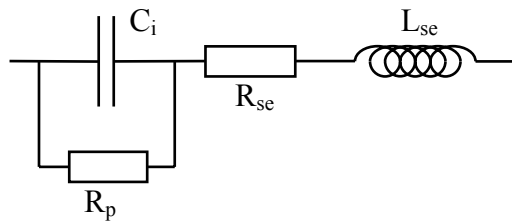
Cette étude avait été réalisée pour une alimentation à découpage de type ‘Forward unidirectionnel’. L’auteur avait montré que les condensateurs étaient responsables de 3 pannes sur 4 à température ambiante de 25°C et que, pour cette même température, le taux de défaillance des transistors était en moyenne deux fois plus élevé que celui des autres composants. Pour ces raisons et en faisant l’hypothèse que cette étude est généralisable à l’ensemble des convertisseurs de puissance, nous nous sommes limités, lors de notre application, à l’influence des défaillances de condensateurs de filtrage sur le convertisseur lui-même mais aussi sur les interactions possibles avec la machine asynchrone.

### 1.5.1. Défaillance du condensateur

Les condensateurs sont plus étudiés parce qu’ils sont dans les applications courantes les composants qui posent le plus de problèmes mais qui, inversement, sont les plus simples à analyser.

#### 1.5.1.1. Cas des condensateurs électrolytiques

Malgré leur faible fiabilité, des considérations de rendement volumique et de prix font que la grande majorité des concepteurs d’alimentations choisissent pour constituer les filtres, des condensateurs électrolytiques. Le schéma équivalent d’un condensateur réel peut être mis sous différentes formes [Sic Saftco, 93], [Besson, 90]. Nous utiliserons celle de la figure 1.9, avec les notations suivantes :



**Fig. 1.9.** Schéma équivalent d'un condensateur réel

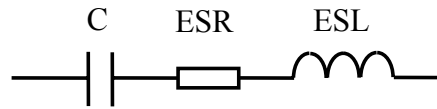
$C_i$  : capacité idéale anode-cathode, l'élément principal du condensateur

$R_p$  : résistance parallèle représentant les pertes diélectriques et les fuites entre les deux électrodes (pour les condensateurs électrolytiques à l'aluminium, celle-ci est due à la présence des couches d'alumine).

$R_{se}$  : résistance série des connexions et des armatures.

$L_{se}$  : inductance équivalente série des connexions et des enroulements. Elle dépend de la technologie de fabrication.

Le schéma précédent peut être simplifié suivant la représentation normalisée de la figure 1.10.



**Fig. 1.10.** Schéma équivalent simplifié du condensateur

Il se compose de :

$$C = C_i \left( 1 + \frac{1}{R_p^2 C_i^2 \omega_n^2} \right) \quad ESR = R_{se} + \frac{R_p}{1 + R_p^2 C_i^2 \omega_n^2} \quad ESL = L_{se}$$

$\omega_n$  : pulsation naturelle du circuit C-ESR-ESL

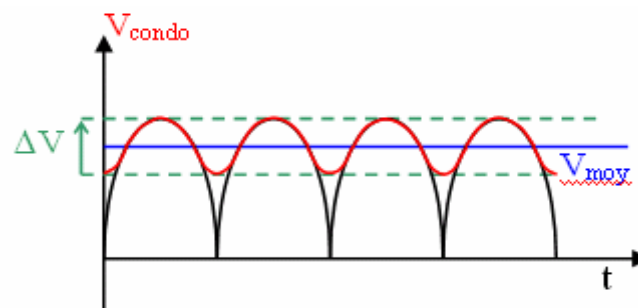
C : capacité (dépendante de la fréquence).

ESR : résistance équivalente série représentant toutes les pertes dans les condensateurs (fonction de la fréquence).

ESL : inductance équivalente série, identique à  $L_{se}$ , sa valeur est de quelques nH et est indépendante de la fréquence.

Pour ces condensateurs, la loi d'Arrhenius [Pauling, 51] nous montre que leur durée de vie est approximativement divisée par deux, chaque fois que la température est augmentée de 10°C.

Le courant qui traverse le condensateur et donc la résistance équivalente série ESR, provoque une élévation de température et influence la durée de vie. La fin de vie des condensateurs est provoquée par un phénomène de défaillance interne, évaporation de l'électrolyte ou réaction chimique qui se manifeste par la dégradation des paramètres électriques suivants : la résistance équivalente série ESR augmente et la capacité diminue [Sic Saftco, 93], [Rifa, 88], cette dernière évoluant de manière moins importante. Avec cette augmentation rapide de l'ESR et cette diminution de C, la tension aux bornes du condensateur  $V_{condo}$  présente de plus en plus d'ondulation ( $\Delta V$ ). Une surveillance de ce taux peut permettre de détecter l'état du filtre d'entrée.



**Fig. 1.11.** Tension en sortie du filtre

La température est également un facteur aggravant de l'état du condensateur et elle affecte sa capacité. Pour chaque condensateur, une plage de température de fonctionnement est définie. Si la température croît et qu'elle dépasse la limite thermique du condensateur, cela accélère son vieillissement et peut provoquer une défaillance, en cas d'élévation trop importante.

Enfin, le condensateur peut se retrouver en court-circuit lorsqu'une défaillance apparaît sur un bras (un bras est une cellule de commutation constituée de 2 interrupteurs, figure 1.4). Si on regarde le schéma équivalent du condensateur, on voit que par nature, sans être connecté à un autre composant, le condensateur se décharge (décharge naturelle de  $C_i$  dans  $R_p$ ,  $I_f$  courant de fuite).  $R_p$  représente la qualité du diélectrique et de l'oxyde. Si ce dernier est dégradé, la résistance de fuite diminue et donc le courant de fuite augmente. Ainsi, si un bras est en court-circuit, cela revient à mettre en court circuit la capacité. Dans ce cas, soit les connectiques cassent (soudure des pattes sur le condensateur) soit cela génère un courant plus intense dans le condensateur (augmentation du courant de fuite) entraînant une augmentation de la température et donc une usure accélérée de l'électrolyte.

### 1.5.1.2. Cas des condensateurs à films métallisés

Les conséquences d'une défaillance sur un condensateur à films dépendent de l'utilisation de ce condensateur.

En capacité de filtrage, comme pour les condensateurs électrolytiques, une augmentation du taux d'ondulation de la tension en sortie du filtre est observée. On observe également un échauffement dans la capacité dû à une mauvaise répartition du courant dans la métallisation du film (ondulation haute fréquence).

Dans le cas d'une utilisation dans un circuit d'aide à la commutation (CALC), un mauvais dimensionnement en courant d'une part et un fonctionnement à haute fréquence d'autre part peuvent produire des échauffements dus également à une mauvaise répartition des courants.

Dans une utilisation pour le découplage, la connectique actuelle des condensateurs ne permet pas l'élimination des parasites hautes fréquences pouvant résulter du réseau et/ou des commutations. Il n'y a pas à l'heure actuelle de solution technologique permettant de répondre complètement au problème.

## 1.5.2. Défaillance de diodes classiques

Pour les diodes classiques, l'analyse est plus complexe car il existe plusieurs modes de défaillance possibles. Seul le vieillissement, qui est la défaillance la plus importante, sera présenté ici.

Les diodes peuvent vieillir de plusieurs façons. Il peut y avoir un problème sur la puce de silicium, un problème sur le boîtier qui n'est plus ou pas hermétique. Un indicateur de l'état de la diode est l'impédance thermique  $Z_{th}$  qui devra être la plus petite possible. Elle est définie par la relation suivante :

$$Z_{th} = \frac{T_{jonction}^{\circ} - T_{boitier}^{\circ}}{P} \quad \text{avec } P \text{ puissance dissipée dans la diode en Watt}$$

Les coefficients de dilatation des différents matériaux de la diode étant différents, des craquelures peuvent apparaître. La dissipation thermique est alors de moins en moins bonne et l'impédance thermique va augmenter (se dégrader). La défaillance apparaît ainsi en fin de vie du composant, pendant la période dite 'd'usure'. Ce problème est un problème que l'on retrouve dans la plupart des semi-conducteurs de puissance.

## 1.5.3. Défaillance d'un exemple d'interrupteur statique : l'IGBT (Insulated Gate Bipolar Transistor)

### 1.5.3.1. Court-circuit

Le défaut de type court circuit se produit lorsqu'un des composants d'une cellule de commutation reste constamment fermé. Un tel défaut peut avoir deux causes : soit la mise en court-circuit physique de la puce de silicium (IGBT ou diode) par dépassement de température critique, soit la défaillance de la commande rapprochée. L'analyse du déroulement d'un court-circuit est décrite dans [Baudesson, 98].

Dans ce cas, les courants de phases sont fortement altérés. Ce type de défaillance est extrêmement préjudiciable au convertisseur et nécessiterait la mise en œuvre de sécurité permettant de déconnecter le bras défaillant dès l'apparition de cette défaillance.

### 1.5.3.2. Circuit ouvert

Il se peut, pour diverses raisons, qu'un des IGBT (supérieur ou inférieur) d'une cellule de commutation d'un bras reste constamment ouvert (défaut de type circuit ouvert). Ce type de défaut a pour principale cause une défaillance de la commande rapprochée (défaut thermique du driver ou perte d'alimentation par exemple). Selon la localisation du défaut dans la cellule,

celui-ci sera visible soit sur l'alternance positive du courant (défaut sur IGBT supérieur), soit sur l'alternance négative (défaut IGBT inférieur). Lorsque le défaut est visible, la signature se matérialise par le passage d'un rapport cyclique  $\alpha$  à 0 ou à 1, mais uniquement sur la cellule défaillante. La conséquence est que la cellule défaillante se retrouve alors en surtension et qu'un arrêt du convertisseur est obligatoire pour éviter une propagation du défaut.

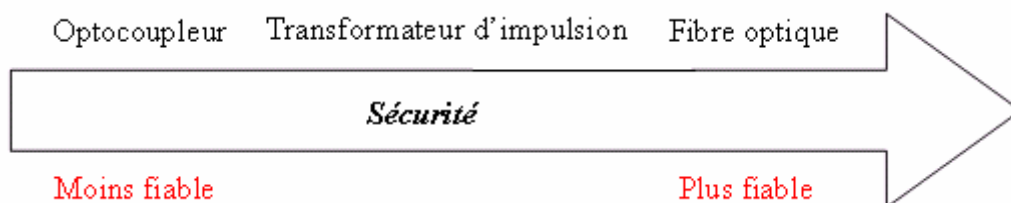
### 1.5.3.3. Vieillessement

Comme pour la diode classique, ce mode de défaillance de l'IGBT est corrélé avec l'augmentation de l'impédance thermique qui peut être entraînée par :

- Sollicitations successives (cyclage)
- Défaut au niveau de la puce de silicium
- Boîtier non hermétique, permettant à l'humidité de pénétrer à l'intérieur (plus ou moins fréquent sur les composants non militaire) et la corrosion du composant.

### 1.5.4. Défaillance du circuit commande (Vieillessement de l'optocoupleur)

Le circuit de commande impose l'état des cellules de commutation, c'est-à-dire l'ouverture ou la fermeture des IGBT. La conception du circuit de commande doit être réalisée avec soin pour réaliser au mieux les fonctions de l'interrupteur. Une isolation galvanique est présente afin de découpler le circuit porteur d'information (partie du circuit de commande générant les impulsions) du circuit porteur de puissance. Cette isolation peut être réalisée au moyen d'un optocoupleur, d'un transformateur d'impulsion ou encore par une liaison par fibre optique. Bien qu'équivalentes en terme de fonction, ces techniques ne présentent pas les mêmes niveaux de fiabilité (figure 1.12).



**Fig. 1.12.** Fiabilité des technologies dans les drivers de commande

Regardons le cas le plus défavorable. L'optocoupleur, qui est constitué d'une diode électroluminescente (LED) et d'un phototransistor, est l'élément qui présente le plus de



défaillance. Un défaut de l'optocoupleur est principalement lié au vieillissement de ce dernier. En vieillissant, la résine de l'optocoupleur s'opacifie, et, de ce fait, le niveau d'énergie du signal et la bande de fréquence couverte diminuent. Il n'y a alors plus de transmission de l'information et la sortie du circuit de commande reste inactive. Ce vieillissement est accentué par les effets de la température. Il semble que la défaillance de la LED intervient systématiquement avant la défaillance du phototransistor. Le vieillissement de la LED peut être détecté relativement simplement, en mesurant le rapport entre le courant d'entrée et le courant de sortie (Current Transfer Ratio – CTR).

$$CTR = \frac{I_s}{I_e} = \frac{I_{collecteur}}{I_{Forward}} \quad I_{Forward} = \text{courant direct de la diode}$$

Si la LED vieillit, sa puissance diminue, ce qui entraîne une variation du CTR passant de la courbe bleue (figure 1.13) pour un composant neuf à la courbe verte pour un composant usé. La surveillance du vieillissement de la LED peut donc se faire par la surveillance du CTR.

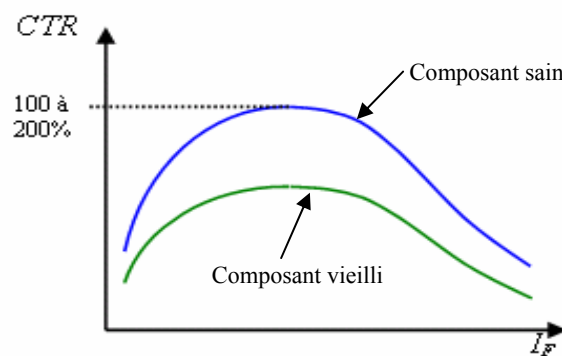


Fig. 1.13. Courbes de CTR - Composant *sain* et composant *défaillant*

## 1.6. Défaillances d'autres origines

### 1.6.1. Défauts de la charge [Bigret, 94], [Schoen, 97]

Les travaux menés dans ce domaine ont surtout concerné l'influence de la dynamique de la charge sur la détection et le diagnostic d'autres défauts tels que la rupture d'une barre ou l'excentricité du rotor. Ces études ont montré que l'introduction de perturbations ou d'oscillations par la charge pouvait masquer l'apparition des défauts du moteur [Saunders, 95], [Calis, 99].

Pour maîtriser l'étude des défauts de la charge, une connaissance parfaite de cet élément s'avère nécessaire. Pour ce faire, on peut procéder tout d'abord à une classification des différents types de charge les plus rencontrés [Saunders, 95] :

- charge à couple constant : le couple n'est pas fonction de la vitesse (ex : traction, convoyeur, tapis roulant,...),
- charge à couple variable : le couple varie en fonction de la vitesse (ex : pompe centrifuge, ventilateur,...),
- charge à puissance constante : tels que les broyeuses, les affûteuses, les bobineurs,...

La nature de l'application de la charge, quelle que soit son type, peut influencer d'une manière destructive le fonctionnement du variateur en amont. En l'occurrence, une mauvaise utilisation de l'ensemble moteur – charge peut engendrer des défaillances mécaniques du système [Beguenane, 97].

Citons quelques exemples de fonctionnements perturbateurs :

- application successive et répétitive de surcharges qui peut causer une fatigue mécanique des roulements et de l'arbre du moteur,
- des à-coups de couple et de charge qui se manifestent par des couples pulsatoires,
- des démarrages en force répétitifs.

La charge peut donc être le siège de mauvaises conditions de fonctionnement qui aboutissent à l'apparition de défauts tels que :

- déséquilibre de charge et oscillations du couple de charge,
- surcouples accidentels,
- à-coups de couple accidentels,
- perturbations mécaniques dues aux vibrations et phénomènes de résonance mécanique liés à la construction de la machine ou introduites par l'environnement extérieur,
- phénomène de balourd,
- désalignement des différents arbres en jeu pour assurer l'accouplement et la rotation,
- manque de lubrifiants,
- vitesse critique de rotation,
- rupture de la charge.

Tous ces défauts de la charge vont naturellement avoir une influence dégradante sur les grandeurs mécaniques mises en jeu au niveau de l'entraînement, à savoir le couple électromagnétique et la vitesse de rotation. En outre, et selon la stratégie de commande adoptée pour le pilotage de l'entraînement, les grandeurs électriques vont être aussi plus ou moins affectées.

### 1.6.2. Défauts capteurs

Dans le cadre de la recherche d'une méthodologie assurant un diagnostic sûr et fiable, l'essentiel de la stratégie ne réside pas uniquement dans le grand nombre de capteurs et la redondance des informations mais aussi dans la fiabilité et la précision des systèmes de mesure et d'analyse de ces informations [Thomson, 99], [Henaou, 99].

En effet, le degré de pertinence de l'information en provenance des capteurs, doit être déterminé afin d'éliminer toutes les informations et les mesures aberrantes.

Une information erronée peut être le résultat d'une panne ou d'un défaut au niveau du système de mesure lui-même. La redondance matérielle peut être considérée dans ce cas comme importante pour le diagnostic. En revanche, elle est onéreuse en termes d'équipement et de maintenance.

Afin d'assurer une bonne qualité de mesure et éviter des sources possibles d'erreurs, quelques règles doivent être respectées :

- s'assurer de la fiabilité et de la précision des instruments de mesure (étalonnage, tests, etc...) [Thomson, 99],
- bien choisir le régime de fonctionnement pendant lequel sont effectuées les mesures (régime transitoire, permanent),
- rendre compte des conditions environnementales du fonctionnement du processus,
- bien définir les marges d'incertitudes tolérées (perturbations, erreurs, etc.).

## 1.7. Synthèse des défaillances

Le tableau 1.1 présente quelques applications néfastes de la charge et les conséquences que cela peut engendrer.

Les tableaux 1.2, 1.3 et 1.4 répertorient une liste non exhaustive des défaillances pouvant se produire sur un ensemble convertisseur – machine – charge en définissant son lieu d'apparition, ses causes possibles et ses effets sur le système ainsi que les différents observables possibles.

Composants	Défaillances	Causes possibles	Effets sur le système	Observable
Charge	Déséquilibre de charge Oscillations du couple de charge Surcouples accidentels Phénomène de balourd Rupture de charge Vitesse critique de rotation Désalignement des différents arbres Perturbations mécaniques	Applications successives et répétitives de surcharges A-coups de couple et de charge Démarrages en force répétitifs	Elévation de la température ⇒ dégradation des isolants ⇒ risque de court-circuit Fatigue mécanique des roulements	Température Vibrations Couple Courants Vitesse

⇒ : entraîne

**Tableau. 1. 1.** Synthèse des défaillances sur la charge

Composants		Défaillances	Causes possibles	Effets sur le système	Observable
Machine asynchrone à cage d'écurueil	Rotor	Ruptures de barres	Surcharge	Réduction de la valeur moyenne de couple électromagnétique Oscillations de la vitesse Vibrations mécaniques	- Vibrations - Couple électromagnétique - Courant
		Ruptures d'anneaux	Bulles coulées Dilatations différentielles Surcharge du couple Détérioration des conditions de fonctionnement	Déséquilibre la répartition des courants ⇒ modulation d'amplitude des courants statoriques	- Courant - Vibrations
		Excentricité	Positionnement incorrect des paliers lors de l'assemblage Usure des roulements Défaut de charge Défaut de fabrication	Balourd (mécanique et magnétique) ⇒ oscillations de couple	- Vibrations - Courants - Couple
	Axe	Roulements	Usure Température élevée Perte de lubrification Défaut de montage Huile contaminée	oscillations du couple de charge ↗ pertes ↗ vibrations Altération de l'équilibre magnétique	- Vibrations - Courants - Puissance instantanée - Variables mécaniques

↗ : augmentation

↘ : diminution

⇒ : entraîne

**Tableau. 1. 2.** Synthèse des défaillances du rotor sur la machine asynchrone

Composants		Défaillances	Causes possibles	Effets sur le système	Observable
Machine asynchrone à cage d'écurueil	Stator	Isolant dans un enroulement	Dégradations de l'isolant à la fabrication Tension élevée Courant élevé Vibrations mécaniques Température élevée Vieillessement naturel	Court-circuit dans l'enroulement	- Température - Vibrations - Courants
		Court-circuit entre spires	Défaut d'isolants	↗ courants statoriques, variation d'amplitude sur les autres phases, modification de facteur de puissance, ↗ courants rotoriques	- Température - Vibration - Courant - Puissance instantanée
		Court-circuit entre phases	Décharge partielle	Dépend fortement de la localisation <u>Proche de l'alimentation :</u> Fusion des conducteurs Disjonction par les protections <u>Proche du neutre :</u> ↗ courants des les barres et les anneaux déséquilibre des courants statoriques	- courants (déséquilibre) - température - tension - puissance instantanée
		Court-circuit phase/bâti	Dégradation des isolants	↗ courant dans la phase ⇒ ↗ de la température ⇒ défauts d'isolants apparition d'une composante homopolaire ⇒ couple pulsatoire	- Courant de fuite
		Circuit magnétique		Dissymétrie du fonctionnement de la machine ⇒ surchauffe, surtension élévation du courant	- Température - Courants - Tensions
		Déséquilibre d'alimentation	Problème sur le réseau d'alimentation Défaillance du convertisseur	↗ des courants dans les autres phases ↗ vibrations ↗ de l'échauffement	- Vibrations - Courants - Tensions - Puissance

↗ : augmentation

↘ : diminution

⇒ : entraîne

**Tableau. 1. 3.** Synthèse des défaillances du stator sur la machine asynchrone

Composants		Défaillances	Causes possibles	Effets sur le système	Observable
Convertisseur de puissance	condensateur	Electrolytique	Elévation de température Vieillessement (usure)	$\nearrow$ ESR et $\searrow$ C $\nearrow$ $\Delta V$ aux bornes condensateurs $\Rightarrow$ altération service du filtre d'entrée ou de sortie	Tension (ondulation $\Delta V$ )
		A films métallisés	usure mauvais dimensionnement composante HF du courant non homogénéité des condensateurs	mauvaise répartition du courant dans la métallisation du film $\Rightarrow$ $\nearrow$ température et $\nearrow$ $\Delta V$	Tension (ondulation $\Delta V$ )
	Diodes classiques	Vieillessement	Température Défaut de puce Cyclage	$\nearrow$ impédance thermique et $\searrow$ capacité de dissipation thermique Perte de puissance pendant l'étape de roue libre $\Rightarrow$ $\searrow$ couple	Impédance thermique $Z_{th}$
	IGBT	Court-circuit	défaut de commande Usure (cyclage) Défaut de puce	Court circuit sur un bras $\Rightarrow$ altération importante des courants de phases et du couple	Courant Couple
		Circuit ouvert	Défaillance des oxydes de grilles	Courant quasiment nul dans la phase affectée $\Rightarrow$ oscillation du couple	Courant Couple
		vieillessement	Usure du composant	Diminution de la tension de sortie du bras (élévation de la tension drain-source)	Tension
	Circuit de commande optipcoupleur	Défaillance LED	Vieillessement Température	$\searrow$ du niveau d'émission (énergie et spectre) $\Rightarrow$ sortie toujours inactive	Rapport entre le courant de sortie $I_s$ et celui d'entrée $I_e$ (CTR)

$\nearrow$  : augmentation

$\searrow$  : diminution

$\Rightarrow$  : entraîne

**Tableau. 1. 4.** Synthèse des défaillances sur les convertisseurs de puissance

## ***1.8. Conclusion***

Ce chapitre présente une liste non exhaustive des différentes défaillances pouvant se produire sur un ensemble convertisseur – machine asynchrone.

Dans un premier temps, après avoir rappelé succinctement la constitution du moteur asynchrone et de son alimentation, nous nous sommes intéressés à répertorier les principales défaillances se produisant sur les différentes parties de la machine ainsi que les causes et les conséquences de leur apparition.

Les défauts étudiés par la suite sur la machine sont :

- Ruptures de barres ou de portions d’anneaux
- Court-circuit entre spires
- Court-circuit entre phases
- Défauts roulements

Ensuite, nous avons répertorié une partie des défaillances possibles sur l’onduleur de tension en rappelant leur cause et leur conséquence.

Les défauts étudiés par la suite sur le convertisseur se limiteront à une défaillance des condensateurs de filtrage (augmentation de l’ESR). Les autres défauts n’ont pas pu être étudiés pour des raisons pratiques. Néanmoins, il s’agit de la défaillance qui a la plus forte fréquence d’apparition, du fait de la faible fiabilité des condensateurs électrolytiques.

Dans le chapitre suivant, nous allons nous intéresser aux signatures caractéristiques de ces défaillances, déterminées à partir de modèles analytiques et/ou de mesures vibratoires, électriques et électromagnétiques. Ces signatures nous serviront pour la détection, l’identification ou la localisation des défaillances précédemment citées (chapitre IV).





## CHAPITRE II

Extraction de signatures représentatives des défaillances :  
Etat de l'art



## 2.1. Introduction

La recherche de signatures ou indicateurs de défauts a pour but de caractériser le fonctionnement du système en identifiant le type et l'origine de chacun des défauts. Ceci permet d'assurer une bonne discrimination des pannes ou anomalies survenant aux différents niveaux du processus.

L'apparition d'un défaut au niveau de l'entraînement modifie le fonctionnement de ce dernier, ce qui affecte ses performances. Etant donné la diversité de ces défauts et la complexité des relations de cause à effet, de nombreuses méthodes permettant de réaliser une démarche de diagnostic et de prévenir au mieux la dégradation du système surveillé, ont été développées.

Ainsi lorsque les observations issues du système sont de type numérique, et que l'on dispose d'un modèle mathématique du système, l'approche par modèle mathématique (méthode interne) est privilégiée pour le diagnostic de systèmes dynamiques [Franck, 92], [Isermann, 84]. En revanche lorsqu'on ne peut construire un modèle analytique du système ou si sa complexité le rend inexploitable, une alternative est l'approche sans modèle (aucune connaissance a priori).

Dans ce chapitre, nous allons faire une présentation des différentes méthodes de détection ainsi que des signatures qui en découlent. Ces méthodes sont classées en deux grandes familles, analytique ou heuristique, en fonction de la forme prise par la connaissance du comportement du système, et par conséquent, de la méthode d'analyse qui en découle. La distinction est donc faite entre, les méthodes qui effectuent l'analyse des signaux d'acquisitions *sans connaissance a priori*, qualifiées de diagnostic externe par [Zwingelstein, 95], et celles qui nécessitent la formulation d'un modèle mathématique du système, qualifiées de diagnostic interne par ce même auteur.

Toutes ces méthodes permettent de générer une information pertinente (paramètres, vecteur forme, règles, etc...) pour l'élaboration des indicateurs de défauts pour le système. Le choix d'une méthode se fera en fonction de la nature de ces indicateurs de défauts.

## ***2.2. Extraction de signatures***

### **2.2.1. Extraction à partir de méthodes internes (connaissance a priori)**

Ces méthodes sont issues principalement de l'automatique et supposent une *connaissance a priori* du système. Elles s'appuient sur le suivi d'évolution des paramètres caractéristiques du système étudié ou sur la différence entre le modèle et le processus (méthode des résidus). Une comparaison entre les paramètres mesurés ou calculés et ceux associés à un mode de fonctionnement normal (sain) nous renseigne sur la présence éventuelle d'un défaut.

#### **2.2.1.1. Méthodes de diagnostic basées sur des modèles paramétriques**

Deux approches peuvent être utilisées pour diagnostiquer les défauts à partir des modèles paramétriques.

La première repose sur le suivi des paramètres électriques caractérisant les modèles. Il s'agit de détecter un écart entre l'évolution des paramètres au cours d'un fonctionnement sain et celle suivie par le système en présence de défauts. L'identification des paramètres peut se faire hors ligne. Elle se base alors sur un algorithme d'optimisation de l'erreur entre les paramètres du modèle et ceux du système étudié [Ljung, 99] [Khelif, 94]. Les paramètres peuvent aussi être identifiés en cours de fonctionnement à l'aide d'observateurs étendus. Le plus couramment utilisé est le Filtre de Kalman [Kalman, 61], [Eykhoff, 74], [Naït-Saïd, 00].

La deuxième approche consiste à analyser les résidus générés à partir d'un modèle proche du système à surveiller. Le modèle de processus est constitué de relations de contraintes dynamiques liant deux types de variables : des variables inconnues (variables internes, perturbations, entrées inconnues, ...) et des variables connues (consignes, variables mesurées). Les résidus sont théoriquement nuls en fonctionnement normal et différents de zéro lorsqu'une défaillance survient.

Plusieurs méthodes sont utilisées pour faire ressortir un résidu significatif de la présence de défaut (gradient, Newton-Raphson, Levenbert-Marquardt). La difficulté majeure consiste à faire la liaison entre le résidu établi et un défaut particulier.

Dans la section suivante, nous présentons de façon succincte et non exhaustive, les modèles de comportement dynamique de la machine asynchrone destinés soit à la recherche des

signatures des défauts, soit à l'analyse de fonctionnement en présence de défaut, et les paramètres permettant de détecter l'apparition de défauts.

## 2.2.1.2. Diagnostic basé sur une modélisation de la machine asynchrone en fonctionnement sain

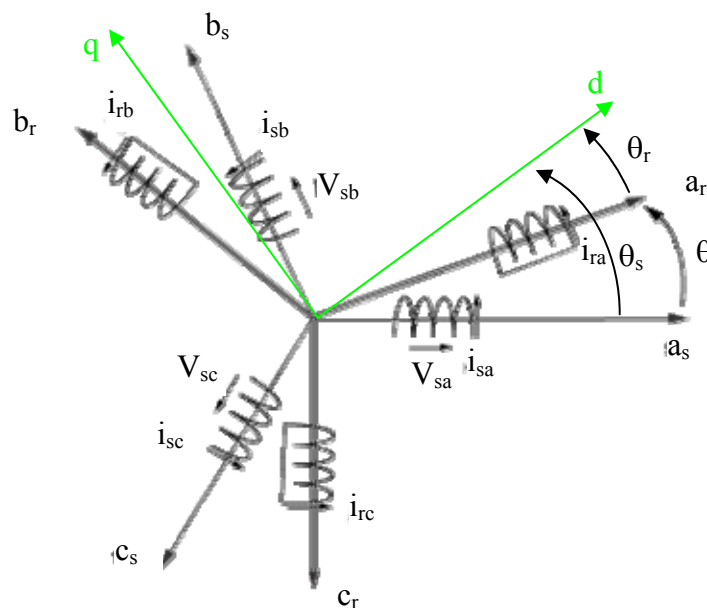
### 2.2.1.2.1. Modèle triphasé-triphasé

#### 2.2.1.2.1.1. Présentation du modèle

La machine asynchrone à cage d'écuréuil est constituée de trois enroulements logés symétriquement dans les encoches du stator et d'une cage d'écuréuil conductrice au rotor. Cette dernière est assimilable à trois enroulements en court-circuit et identiques, en fonctionnement sain.

On modélise la machine asynchrone à cage dans le repère a, b, c en émettant les hypothèses simplificatrices suivantes:

- Les pertes fer sont négligées,
- La saturation du circuit magnétique est négligée,
- L'effet d'extrémité des têtes de bobines est négligé,
- Le stator est considéré comme lisse et l'entrefer constant (l'effet des encoches est négligé),
- Les harmoniques de toute origine, autre que temporel, sont négligés.



**Fig. 2. 1.** Représentation d'une machine asynchrone triphasée au stator et au rotor

La figure 2.1 montre la représentation spatiale d'une machine asynchrone à cage comprenant un enroulement triphasé au stator ( $a_s, b_s, c_s$ ) comme au rotor ( $a_r, b_r, c_r$ ).

Les équations électriques au stator et au rotor sont alors données par :

$$\begin{aligned} [V_s] &= [R_s] \cdot [I_s] + \frac{d[\Phi_s]}{dt} \\ [V_r] &= 0 = [R_r] \cdot [I_r] + \frac{d[\Phi_r]}{dt} \end{aligned} \quad (2.1)$$

$$\text{Avec : } [I_{s,r}] = \begin{bmatrix} i_{s,ra} \\ i_{s,rb} \\ i_{s,rc} \end{bmatrix}, [V_{s,r}] = \begin{bmatrix} V_{s,ra} \\ V_{s,rb} \\ V_{s,rc} \end{bmatrix}, [\Phi_{s,r}] = \begin{bmatrix} \varphi_{s,ra} \\ \varphi_{s,rb} \\ \varphi_{s,rc} \end{bmatrix} \text{ et } [R_{s,r}] = \begin{bmatrix} R_{s,ra} \\ R_{s,rb} \\ R_{s,rc} \end{bmatrix}$$

Ce qui peut se mettre sous la forme suivante [Ah-Jaco, 97]

$$\begin{cases} \dot{[I]} = [L]^{-1} ([V] - ([R] + [G]\omega)[I]) \\ \dot{\omega} = \frac{1}{2J} ([I]^t [G][I]) - \frac{C_r}{J} - \frac{f_l}{J} \omega \\ \dot{\theta} = \omega \end{cases} \quad (2.2)$$

avec  $[V]^t = [v_{as} \ v_{bs} \ v_{cs} \ 0 \ 0 \ 0]$ ,  $[I]^t = [i_{as} \ i_{bs} \ i_{cs} \ i_{ar} \ i_{br} \ i_{cr}]$

$$[R] = \begin{bmatrix} R_{sa} & 0 & 0 & 0 & 0 & 0 \\ 0 & R_{sb} & 0 & 0 & 0 & 0 \\ 0 & 0 & R_{sc} & 0 & 0 & 0 \\ 0 & 0 & 0 & R_{ra} & 0 & 0 \\ 0 & 0 & 0 & 0 & R_{rb} & 0 \\ 0 & 0 & 0 & 0 & 0 & R_{rc} \end{bmatrix}, [L] = \begin{bmatrix} L_{sa} + L_{fsa} & M'_{sasb} & M''_{sasc} & M_{sara} & M'_{sarb} & M''_{sarc} \\ M'_{sbsa} & L_{sb} + L_{fsb} & M'_{sbsc} & M'_{sbra} & M'_{sbrb} & M'_{sbrc} \\ M'_{scsa} & M'_{scsb} & L_{sc} + L_{fsc} & M'_{scra} & M'_{scrib} & M'_{scrc} \\ M'_{rasa} & M'_{rasb} & M'_{rasc} & L_{ra} + L_{rfa} & M'_{rarb} & M'_{rarc} \\ M'_{rbsa} & M'_{rbsb} & M'_{rbsc} & M'_{rbra} & L_{rb} + L_{rfb} & M'_{rbrb} \\ M'_{rcsa} & M'_{rcsb} & M'_{rcsc} & M'_{rcra} & M'_{rcrb} & L_{rc} + L_{rfc} \end{bmatrix}$$

$$\text{et } [G] = \frac{d[L]}{d\theta}$$

En fonctionnement sain  $R_{sa} = R_{sb} = R_{sc}$  et  $R_{ra} = R_{rb} = R_{rc}$

### 2.2.1.2.1.2. Discrimination des défauts

Le modèle triphasé-triphasé modélise les trois phases statoriques ce qui permet de différencier le comportement de chaque phase. Les enroulements rotoriques (barres + anneaux) sont représentés par trois phases également.

La modélisation des enroulements stator (respectivement rotor) par des paramètres distincts permet de discriminer un comportement défectueux d'un comportement dû à un

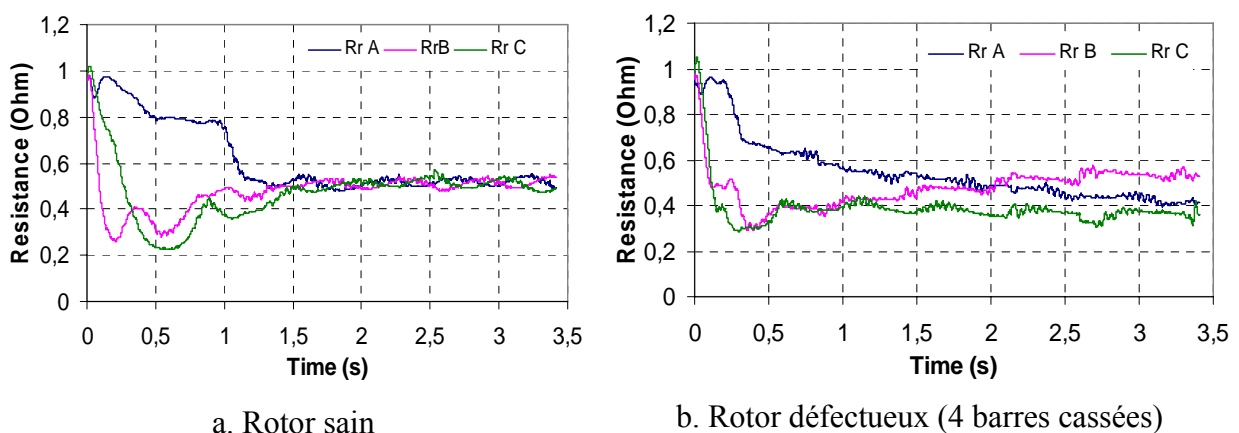
fonctionnement normal. En effet les paramètres peuvent évoluer même lors d'un régime sain. Par exemple la résistance peut varier avec la température :

$$R = R_0(1 + \alpha \cdot \Delta T), \text{ avec } R_0 : \text{valeur de } R \text{ à } T_0 = 25^\circ\text{C}$$

$\alpha$  : coefficient de température de la résistance

$\Delta T$  : variation de la température

Sur ce principe, [Boumégoura, 99] a réalisé un système de diagnostic des défauts rotoriques. Les résistances rotoriques des trois phases sont identifiées en ligne par un filtre de Kalman étendu. Cette approche permet de différencier l'effet des défauts rotoriques par rapport aux autres phénomènes qui provoquent la variation simultanée et équilibrée des résistances (température, etc.).



**Fig. 2. 2.** Résistances rotoriques observées : machine saine (a), machine avec 4 barres cassées (b) [Boumégoura, 99].

Ainsi, comme le montre la figure 2.2a, si le moteur est sain, les phases rotoriques seront équilibrées, d'où une égalité entre les trois résistances représentatives. Si par contre un défaut rotorique survient, il provoque un déséquilibre électrique et induit une variation des valeurs de résistances qui ne seront plus égales entre elles (figure 2.2b). Par conséquent, l'auteur considère que cette différence permet de diagnostiquer un défaut au rotor d'une part et d'évaluer la gravité de la détérioration rotorique d'autre part.










En ce qui concerne les défauts statoriques, [Moreau, 99] utilise le modèle triphasé pour localiser l'origine du défaut entre les trois phases. Ainsi, à l'aide d'un algorithme d'estimation des paramètres, il obtient les estimations des paramètres suivants :

$$\theta' = [R_{sa} \ R_{sb} \ R_{sc} \ L_{fsa} \ L_{fsb} \ L_{fsc} \ R_r \ L_{pr}]$$

Dans ce vecteur, seulement deux paramètres caractérisent le rotor et les six autres sont utilisés pour les enroulements stator.



En terme de localisation, l'auteur tire les enseignements suivants :

Défaut sur la phase a	Défaut sur la phase b	Défaut sur la phase c
 $R_{sa}$ et $L_{f_{sa}}$	 $R_{sb}$ et $L_{f_{sb}}$	 $R_{sc}$ et $L_{f_{sc}}$
 $(R_{sb}, R_{sc})$	 $(R_{sa}, R_{sc})$	 $(R_{sb}, R_{sa})$
 $(L_{f_{sb}}, L_{f_{sc}})$	 $(L_{f_{sa}}, L_{f_{sc}})$	 $(L_{f_{sb}}, L_{f_{sa}})$

**Tableau. 2. 1.** Localisation des défauts par identification des paramètres du modèle triphasé [Moreau, 99]

Dans le cas d'un défaut au stator (réduction de spires ou court-circuit), tous les paramètres issus du modèle triphasé changent mais seulement deux sont rigoureusement corrélés avec l'origine du défaut. Par exemple, pour une réduction du nombre de spires sur la phase b : il y a une diminution de  $R_{sb}$  et  $L_{f_{sb}}$  alors que les couples  $(R_{sa}, R_{sc})$  et  $(L_{f_{sa}}, L_{f_{sc}})$  augmentent.

### 2.2.1.2.2. Modèle diphasé

#### 2.2.1.2.2.1. Présentation du modèle

Ce modèle utilise une transformation permettant de projeter les équations électriques et magnétiques du repère triphasé dans un repère orthonormé dq0. Les variables projetées sur l'axe homopolaire 0 sont en générale nulles. Le système équivalent résultant est donc diphasé 'd-q'.

Ainsi, les grandeurs statoriques (rotoriques) liées au repère  $a_s, b_s, c_s$  ( $a_r, b_r, c_r$ ), décalées de  $120^\circ$ , sont ramenées vers les grandeurs biphasées équivalentes  $s_d, s_q$  ( $r_d, r_q$ ) en quadrature de phase comme présenté sur la figure 2.1.

On notera sur cette figure :

$\theta_s$  la position du repère 'd-q' par rapport au stator

$\theta_r$  la position du repère 'd-q' par rapport au rotor.

Ce repère peut être lié au stator (transformation de Park, le repère dq0 est noté  $\alpha\beta 0$ ), au rotor ou au champ tournant.

Mathématiquement, la transformation des courants, des tensions et des flux, est décrite par la relation matricielle suivante :

$$T_{dq0/abc} = \sqrt{\frac{2}{3}} \begin{bmatrix} \cos \theta_s & \cos(\theta_s - \frac{2\pi}{3}) & \cos(\theta_s + \frac{2\pi}{3}) \\ -\sin \theta_s & -\sin(\theta_s - \frac{2\pi}{3}) & -\sin(\theta_s + \frac{2\pi}{3}) \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (2.3)$$

Ainsi :

$$\begin{aligned} \begin{bmatrix} i_{s,r dq0} \end{bmatrix} &= \begin{bmatrix} T_{dq0/abc} \end{bmatrix} \cdot \begin{bmatrix} i_{s,r abc} \end{bmatrix} \\ \begin{bmatrix} V_{s,r dq0} \end{bmatrix} &= \begin{bmatrix} T_{dq0/abc} \end{bmatrix} \cdot \begin{bmatrix} V_{s,r abc} \end{bmatrix} \\ \begin{bmatrix} \varphi_{s,r dq0} \end{bmatrix} &= \begin{bmatrix} T_{dq0/abc} \end{bmatrix} \cdot \begin{bmatrix} \varphi_{s,r abc} \end{bmatrix} \end{aligned} \quad (2.4)$$

$\theta_s$  est l'angle électrique entre le stator et l'axe d. L'indice 0 désigne la composante homopolaire. Elle est égale à zéro dans le cas d'un système triphasé équilibré.

Si  $\theta_s$  est nul alors on obtient la matrice [C], appelée matrice de Concordia telle que :

$$[C] = \sqrt{\frac{2}{3}} \begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} & -\sqrt{\frac{3}{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (2.5)$$

Dans ce cas, les axes d et q sont appelés  $\alpha$  et  $\beta$ .

Cette transformation assure la conservation de la puissance instantanée.

La matrice de rotation  $[R(\theta_s)]$  permet le passage du vecteur courant statorique projeté dans un repère lié au stator  $[I_{s\alpha} \ I_{s\beta}]^T$  vers le vecteur projeté dans le repère 'd-q'  $[i_{sd} \ i_{sq}]^T$  :

$$\begin{aligned} i_{s,dq} &= [R(\theta_s)] I_{s,\alpha\beta} \\ \text{avec } [R(\theta_s)] &= \begin{pmatrix} \sin \theta_s & \cos \theta_s \\ -\cos \theta_s & \sin \theta_s \end{pmatrix} \end{aligned} \quad (2.6)$$

La transformation des courants rotoriques est décrite par la relation matricielle suivante :

$$i_{r,dq} = [R(\theta_s - \theta_r)] I_{r,\alpha\beta} \quad (2.7)$$

L'évolution des flux, des courants et des tensions est décrite par le système d'équations non linéaires exprimées dans le repère tournant 'd-q' :

$$\text{au stator : } \begin{cases} V_{sd} = R_s \cdot i_{sd} + \frac{d\varphi_{sd}}{dt} - \omega_s \cdot \varphi_{sq} \\ V_{sq} = R_s \cdot i_{sq} + \frac{d\varphi_{sq}}{dt} + \omega_s \cdot \varphi_{sd} \end{cases} \quad (2.8)$$

$$\text{au rotor : } \begin{cases} V_{rd} = 0 = R_r \cdot i_{rd} + \frac{d\varphi_{rd}}{dt} - (\omega_s - \omega_r) \cdot \varphi_{rq} \\ V_{rq} = 0 = R_r \cdot i_{rq} + \frac{d\varphi_{rq}}{dt} + (\omega_s - \omega_r) \cdot \varphi_{rd} \end{cases} \quad (2.9)$$

Les flux s'expriment en fonction des courants de la manière suivante :

$$\begin{bmatrix} \varphi_{sd} \\ \varphi_{rd} \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \cdot \begin{bmatrix} i_{sd} \\ i_{rd} \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} \varphi_{sq} \\ \varphi_{rq} \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \cdot \begin{bmatrix} i_{sq} \\ i_{rq} \end{bmatrix} \quad (2.10)$$

Le coefficient de fuite total et les constantes de temps statorique et rotorique sont définis respectivement par :

$$\sigma = 1 - \frac{L_m^2}{L_s \cdot L_r} \quad T_s = \frac{L_s}{R_s} \quad \text{et} \quad T_r = \frac{L_r}{R_r} \quad (2.11)$$

Pour tous les modèles de la machine décrits ci-dessus, l'expression du couple électromagnétique est donnée par l'équation (2.12) :

$$C_{em} = p \cdot (\varphi_{sd} \cdot i_{sq} - \varphi_{sq} \cdot i_{sd}) \quad (2.12)$$

Où  $p$  est le nombre de paire de pôles de la machine.

#### 2.2.1.2.2.2. Discrimination des défauts

Il est important de préciser que la représentation biphasée de la machine n'est valide physiquement que si les alimentations sont équilibrées et si la machine est saine et équilibrée. Lors de l'apparition d'un défaut, un déséquilibre s'instaure dans la machine. Les paramètres du modèle vont dériver.

Dans son étude, [Moreau, 99] utilise l'identification des paramètres issus du circuit équivalent de la machine asynchrone pour détecter l'occurrence d'un défaut.

Ces études portent sur les défauts tels que les barres cassées et la réduction de spires ou les courts-circuits dans les enroulements statoriques. Les paramètres estimés sont les suivants :

$$\theta = [R_s \ R_r \ L_{fs} \ L_r]$$

avec  $L_{fs}$  : inductance de fuite au stator

Ses résultats montrent que :

- Lors de rupture de barres au rotor, la résistance rotorique augmente et les autres paramètres, c'est à dire la résistance statorique et les inductances de fuite statorique et rotorique, diminuent.
- Lors d'un court circuit entre spires du bobinage statorique, la résistance statorique augmente et les autres paramètres diminuent.

	$R_s$	$R_r$	$L_r$	$L_f$
<i>Défaut dans la cage rotorique</i>	↘	↗	↘	↘
<i>Défaut de court circuit au bobinage statorique</i>	↗	↘	↘	↘

Tableau. 2. 2. Evolution des paramètres en présence de défauts [Moreau, 99]

Ainsi, pour déterminer la présence d'un défaut, il est nécessaire de suivre un ensemble de paramètres constitués d'une variable résistive et l'autre inductive.

Sur le même principe, [Boumégoura, 99] propose un observateur de Kalman étendu aux paramètres  $R_r$  et  $L_m$  pour détecter les défauts rotor.

Les figures suivantes montrent la convergence de la résistance et de l'inductance vers les variables imposées.

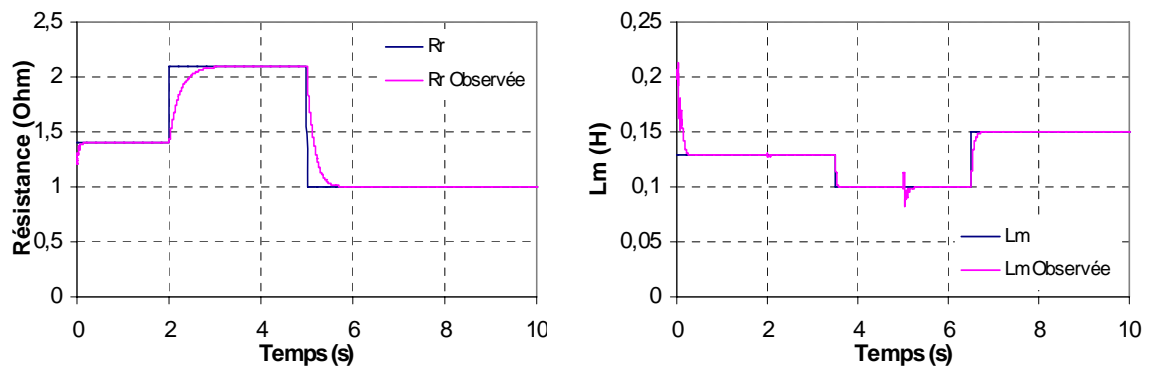
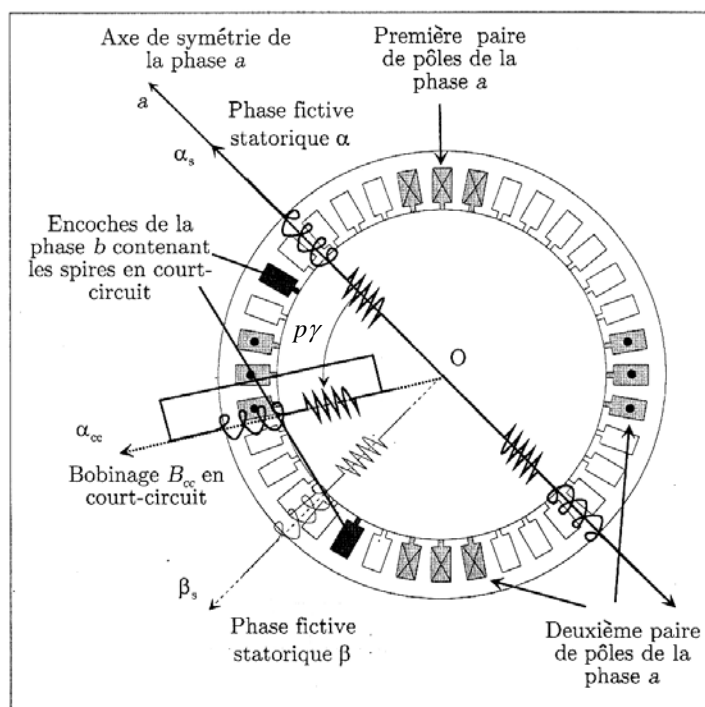


Fig. 2. 3. Les résistances rotoriques et les inductances imposées et observées

Au final, ce ou ces couples de paramètres donnent une signature du défaut. Ils ne donnent plus une représentation physique de la machine détériorée mais un indicateur de panne basé sur une évolution distincte de ces paramètres pour une machine saine ou défectueuse.

### 2.2.1.3. Diagnostic de la machine basé sur un modèle prenant en compte les défauts

Pour détecter un faible nombre de spires en court-circuit (jusqu'à 0,7% du nombre total de spire par phase), [Schaeffer, 99a] fait une approche de type circuit qui met en évidence le fait qu'un court-circuit se traduit non seulement par la création d'une bobine fictive supplémentaire notée  $B_{cc}$  (figure 2.4) mais aussi par la diminution du nombre de spires dans la phase concernée.



**Fig. 2. 4.** Représentation schématique des encoches stator [Schaeffer, 99]. Localisation des spires en court – circuit. La bobine fictive  $B_{cc}$  est représentée dans le repère  $\alpha, \beta$

Ainsi au modèle usuel de Park vient s'ajouter le paramètre  $\alpha$  (à ne pas confondre avec l'axe  $\alpha$  issu de la transformation biphasée) qui est le rapport entre le nombre  $n_{cc}$  de spires en court-circuit et le nombre total  $n_s$  de spires en série par phase du stator.

$$\alpha = \frac{n_{cc}}{n_s}$$

A la bobine  $B_{cc}$  s'associent les grandeurs vectorielles (courant et flux)  $i_{cc}(t)$  et  $\varphi_{cc}(t)$ . L'angle  $p\gamma$  est l'angle entre l'axe de  $B_{cc}$  et l'axe de référence du bobinage statorique.

L'expression (2.13) de  $\varphi_{cc}(t)$  est obtenue en considérant le schéma équivalent monophasé par phase de la machine asynchrone, et en introduisant la notion de flux magnétisant  $\phi_{ms}$  et le

flux de fuite  $\phi_{fs}$ .

$$\Phi_{cc}(t) = \sqrt{\frac{3}{2}} \mathbf{Q}(p\gamma) \left( \alpha l_{ms} \cdot \mathbf{i}_s(t) + \alpha L_m \mathbf{P}(p\theta) \cdot \mathbf{i}_r(t) + \left( l_{ms} + \frac{3}{2} l_{fs} \right) \cdot \alpha^2 \sqrt{\frac{2}{3}} \cdot \mathbf{i}_{cc}(t) \right) \quad (2.13)$$

$$\text{où } \mathbf{Q}(p\gamma) = \begin{bmatrix} \cos(p\gamma) \cos(p\gamma) & \cos(p\gamma) \sin(p\gamma) \\ \cos(p\gamma) \sin(p\gamma) & \sin(p\gamma) \sin(p\gamma) \end{bmatrix}, \quad l_{ms} \text{ et } l_{fs} \text{ sont respectivement les}$$

inductances magnétisante et de fuite que l'on aurait en l'absence de défaut,  $\theta$  est l'angle mécanique entre le stator et le rotor et  $\mathbf{P}$  la matrice de rotation, notée ici  $[\mathbf{R}(\theta_s)]$  (équation 2.6).

Le système d'équation suivant résume le comportement électrique de la machine :

$$\begin{cases} \mathbf{u}_s(t) = R_s \mathbf{i}_s(t) + \frac{d}{dt} (\phi_{ms}(t) + \phi_{fs}(t)) \\ 0 = -\frac{3}{2} \frac{R_s}{\alpha} \cdot \mathbf{i}'_{cc}(t) + \mathbf{Q}(p\gamma) \frac{d}{dt} \phi_{ms}(t) \\ 0 = -R_r \mathbf{i}'_r(t) + \frac{d}{dt} \phi_{ms}(t) - p\dot{\theta} \cdot \mathbf{P} \left( \frac{\pi}{2} \right) \cdot \phi_{ms}(t) \end{cases} \quad (2.14)$$

$\mathbf{i}'_{cc}(t)$  et  $\mathbf{i}'_r(t)$  sont les courants de  $B_{cc}$  et du rotor ramenés au primaire (en considérant le schéma équivalent monophasé par phase de la machine).

Pour la représentation d'état les paramètres estimés a, b, c et d sont :

$$a = \frac{R_s}{\sigma L_s} + \frac{1-\sigma}{\sigma T_r}, \quad b = \frac{L_s(1-\sigma)}{T_r}, \quad c = \frac{1}{\sigma L_s}, \quad d = \frac{1}{T_r} \text{ et } \alpha.$$

Ce qui donne donc le vecteur des cinq paramètres, noté  $\lambda_0 = [a, b, c, d, \alpha]$ . L'influence du défaut sur le comportement entrée/sortie de la machine vient du fait que, dans la représentation d'état, le vecteur de sortie est fonction de l'entrée de commande par l'intermédiaire de  $\alpha$ .

Avec  $\mathbf{Y}$  le vecteur des courants biphasés mesurés et  $\mathbf{U}$  le vecteur des tensions biphasées d'alimentation (obtenus par la transformation de Concordia), la représentation d'état est la suivante :

$$\begin{cases} \dot{\mathbf{X}} = \mathbf{A}(\omega, \lambda_0) \cdot \mathbf{X} + \mathbf{B}(\lambda_0) \cdot \mathbf{U} \\ \mathbf{Y} = \mathbf{C}(\lambda_0) \cdot \mathbf{X} + \frac{3}{2} \frac{\alpha}{R_s} \cdot \mathbf{P}(p\theta) \mathbf{Q}(p\gamma) \mathbf{P}(-p\theta) \mathbf{U} \end{cases} \quad (2.15)$$

Ainsi, ce paramètre évolue de manière plus importante que  $R_s$ ,  $L_s$ ,  $T_r$  et  $\sigma$ . Surtout lorsqu'il s'agit d'un faible nombre de spires en court-circuit. Dans [Schaeffer, 99], le terme caractérisant la relation entrée/sortie est simplifié et donc, dans l'équation d'état le vecteur de sortie  $\mathbf{Y}$  s'exprime de la manière suivante :

$$\mathbf{Y} = \mathbf{C}(\lambda_0) \cdot \mathbf{X} + \alpha \omega_s \frac{\mathbf{U}}{|\mathbf{U}|^2} \quad (2.16)$$

Le rapport entre le nombre de spires en court-circuit et le nombre total de spires dans une phase étudiée est compris entre 1.5 % et 16 %. Cependant, cette approche reste sensible à la présence de défauts rotor.

#### 2.2.1.4. Conclusion

Dans cette partie, nous avons présenté certains modèles permettant d'identifier les paramètres électriques de la machine asynchrone susceptibles de détecter un certain nombre de défauts. Le tableau 2.3 regroupe les résultats obtenus à partir de l'état de l'art du diagnostic interne de la machine asynchrone.

Type de défaut	Evolution des paramètres	Observations
Rupture de barres Rupture de portions d'anneaux	DETECTION $R_s$ <i>diminue</i> $R_r$ <u>augmente</u> $L_r$ <i>diminue</i> $L_s$ <i>diminue</i>	<i>Paramètres issus du modèle de Park</i>
Excentricité	$R_r$ <u>augmente</u>	<i>Augmentation possible de <math>R_r</math> si défaut au rotor</i>
Court-circuit	DETECTION $R_s$ <u>augmente</u> $R_r$ <i>diminue</i> $L_r$ <i>diminue</i> $L_f$ <i>diminue</i> LOCALISATION $R_{sa}$ et $L_{fsa}$ <u>augmentent</u> <i>si défaut sur la phase a</i>	<i><math>R_{sa}</math> et <math>L_{fsa}</math> issus du modèle triphasé</i>

**Tableau. 2. 3.** Paramètres issus de modèles représentatifs de défauts

Ces méthodes d'estimations de paramètres présentent l'avantage de permettre l'observation de grandeurs difficilement, voire même parfois non mesurables. Si les estimations sont correctement réalisées, alors les résistances, inductances, inductances mutuelles peuvent

servir de bons indicateurs de défauts. Cependant ces méthodes restent tributaires d'un compromis entre la finesse de la modélisation, la prise en compte des phénomènes physiques ainsi que du temps de calcul requis. De plus, ces méthodes effectuent un filtrage des signaux d'acquisitions, ce qui entraîne une perte d'information.

Il serait intéressant de compléter ces méthodes par des méthodes ne nécessitant aucun modèle, c'est-à-dire, *sans connaissance a priori* du processus. Cela introduit donc la seconde famille des méthodes de diagnostic, à savoir les méthodes externes, qui présentent l'avantage de ne pas filtrer l'information.

## **2.2.2. Extraction à partir de méthodes externes (sans connaissance a priori)**

Nous avons vu que la première famille de méthodes nécessite la connaissance du comportement dynamique de la machine asynchrone. Dans cette deuxième partie, nous nous intéressons aux résultats obtenus à partir du suivi direct des grandeurs telles que les courants, le couple estimé ou mesuré, les flux ou encore les vibrations.

Ces méthodes dites 'sans modèles' se basent sur l'analyse des signaux d'acquisitions. Elles ont l'avantage de l'indépendance de l'analyse par rapport aux fluctuations internes du système. D'autre part, l'information contenue dans les signaux, n'étant pas filtrée par la modélisation, elle reste intacte.

### **2.2.2.1. Utilisation des différents signaux autres que le courant statorique**

#### **2.2.2.1.1. Analyse des signaux vibratoires**

L'analyse des défaillances des moteurs électriques par les signaux vibratoires et acoustiques permet une détection de quasiment tous les défauts, notamment ceux mécaniques, pouvant se produire sur le processus [Bigret, 94], [Schoen, 94]. Cette analyse peut être réalisée à partir de capteurs, généralement des accéléromètres, placés sur les paliers dans les directions verticales, axiales et radiales.

[Cameron, 86], [Calis, 99], [Filbert, 90], [Tavner, 87] permettent de construire le tableau 2.4 résumant les caractéristiques de l'analyse vibratoire des défauts.



Types de défaut	Signatures (analyse vibratoire)	Commentaires
Rupture de barres Rupture de portions d'anneaux	$f_r \pm 2 \cdot g \cdot f_s$ (g : glissement)	Les vibrations axiales augmentent avec le nombre de barres cassées
Excentricité Statique Dynamique	Excentricité statique : $2 \cdot f_s$ ou des composantes à $N_r \cdot f_r \pm f_s$ Excentricité dynamique : $f_r \pm 2 \cdot g \cdot f_s$ ou $f_r$	L'excentricité dynamique est fonction du temps et de l'espace L'excentricité statique est fonction de l'espace seulement
Désalignement	$f_r$ et $2 \cdot f_r$ principalement	Vibrations plus importantes dans la direction axiale
Déséquilibre mécanique	$f_r$	Au niveau de la charge notamment Vibrations plus importantes dans la direction radiale
Défaillance des paliers	$n \cdot f_e$ et $n \cdot f_i$ $f_i \pm n \cdot f_r$ avec n un nombre entier :1,2,...	Bagues externes Bagues internes Ces composantes fréquentielles sont modulées par la vitesse

**Tableau. 2. 4.** Analyses vibratoires, signatures fréquentielles

Cependant, ces analyses vibratoires comportent certains inconvénients :

- problème d'accessibilité,
- difficultés rencontrées dans les connexions mécaniques des accéléromètres pour effectuer les mesures nécessaires au voisinage direct du défaut.

De plus, le coût de ces capteurs reste relativement élevé par rapport aux autres capteurs tels que les capteurs de vitesse ou de courant.

Pour s'affranchir de ces problèmes, les recherches focalisent leurs efforts pour détecter et localiser les défauts par l'analyse d'autres signaux.

Dans la section suivante, les études concernant l'analyse du spectre du flux dans l'entrefer, du flux axial et même du flux de fuite des têtes de bobines sont présentées.

#### 2.2.2.1.2. Analyse des flux

Tout déséquilibre quel qu'il soit, magnétique, mécanique, électrique ou bien encore électromagnétique, situé au rotor ou au stator, peut affecter la conversion électromécanique et la répartition de champ dans et hors de la machine. Des études ont donc été menées pour

extraire de la mesure des flux d'entrefer, axial ou de fuite des signatures caractéristiques de certains défauts. Pour cela, des bobines exploratrices sont placées à l'extérieur de la machine, parallèlement et perpendiculairement à l'axe du rotor.

Dans [Henao, 99a], l'auteur met en évidence les composantes fréquentielles du flux de fuite dans l'axe du rotor, qui permettent de détecter et d'identifier des courts-circuits dans les enroulements statoriques.

Il en déduit les fréquences caractéristiques suivantes :

$$k \cdot f_s \pm n \cdot f_r \quad (2.17)$$

Avec  $n$  et  $k$  des nombres entiers et  $k$  impair

Ces composantes présentent dans le spectre du flux pour une machine saine, vont augmenter avec l'apparition d'un défaut [Cabanas, 98].

Certaines composantes du flux axial permettent de détecter d'éventuels défauts liés aux barres rotoriques et aux paliers (roulements). Les composantes à prendre en compte pour les barres sont définies dans le tableau 2.5. En ce qui concerne les paliers, les auteurs se focalisent sur les raies situées dans la plage de fréquence 600 Hz à 900 Hz pour une alimentation 50 Hz.

$f_s$	$(5 - 4g)f_s$
$(2g - 1)f_s$	$(6g - 5)f_s$
$(3 - 2g)f_s$	$(7 - 6g)f_s$
$(4g - 3)f_s$	$(8g - 7)f_s$

**Tableau. 2. 5.** *Fréquences caractéristiques du flux axial*







### 2.2.2.1.3. Analyse du couple électromagnétique

Certains défauts mécaniques peuvent être détectés par la recherche d'harmoniques dans le spectre du couple électromagnétique mesuré, résultant d'une interaction entre le flux et le courant [Yahoui, 95]. Ce couple peut être reconstruit, soit à partir de deux des trois courants statoriques, soit en utilisant un modèle physique de la machine.

L'utilisation de ce signal peut s'avérer un bon choix pour la détection des défauts de charge. En effet, les variations du couple de charge vont induire des variations du flux et du courant dans la machine. De même la torsion de l'arbre, entraîne l'apparition d'harmoniques dans le spectre du couple.

Les oscillations de couple, provoquées par certains défauts rotoriques, peuvent servir à détecter ceux-ci. [Yahoui, 96] montre par une analyse théorique et des essais, comment le spectre du couple se trouve modifié lorsqu'un des deux anneaux d'une cage est cassé.

[Thollon, 93] analyse le couple électromagnétique d'une machine asynchrone présentant des portions d'anneaux cassées. Dans le tableau 2.6 sont présentés les résultats obtenus concernant les variations des harmoniques du couple en présence de défaut pour une alimentation 50 Hz.

50 Hz	100 Hz	200 Hz	300 Hz	700 Hz	800 Hz
					

**Tableau. 2. 6.** Evolution des différents harmoniques du couple en présence de ruptures d'anneaux [Thollon, 93]

A partir d'un modèle de la machine, [Yahoui, 96] procède à l'estimation des composantes du flux rotorique dans un repère 'd-q' lié au stator dans le but d'observer le couple électromagnétique  $C_{em}$ . Il constate qu'une dissymétrie électrique du rotor fait apparaître des harmoniques de dentures rotoriques, ce qui modifie le champ d'entrefer. Le problème peut alors être détecté en analysant les fréquences d'encoches présentées dans le spectre du couple estimé  $C_{em}$  :

$$f_{sb} = f_s \cdot \left[ \frac{Nr \cdot (1-g)}{p} \pm n \right] \pm 2 \cdot g \cdot f_s \quad (2.18)$$

Avec :

$f_{sb}$  : fréquence d'encoche au rotor

Nr : nombre d'encoches au rotor

$f_s$  : fréquence d'alimentation (fréquence des courants statoriques)

$n = 1, 2, 3, 4, \dots$

$g$  : le glissement

Pour une alimentation à 50 Hz, [Yahoui, 95] porte l'attention sur la variation d'amplitude des raies à 600 et 900 Hz en présence de barres cassées et des deux premiers harmoniques du couple qui augmentent en présence d'excentricité.

[Kral, 99] utilise le modèle triphasé pour estimer le flux statorique, ce qui permet d'estimer le couple utile  $C_u$ . Or, en présence d'excentricité, le spectre du couple  $C_u$  présente des informations aux fréquences :

$$f_{Cu} = n \cdot \left( 1 - \frac{1+g}{p} \right) \cdot f_s \quad (2.19)$$

De même que le spectre du courant présente des raies aux fréquences :

$$f_{exc} = \left( 1 \pm n \frac{1-g}{p} \right) \cdot f_s \quad (2.20)$$

Comme nous venons de le voir, les signaux évoqués précédemment permettent de détecter un grand nombre de défauts sur la machine. Cependant, ces signaux nécessitent la mise en place d'un grand nombre de capteurs (flux, couplemètre, accéléromètres, transducteurs,...) souvent onéreux, sensibles et délicats à placer dans des environnements contraignants.

Dans [Kryter, 89], l'auteur présente l'efficacité et le potentiel de l'analyse des courants du stator pour détecter les défauts (déséquilibre de phase, court-circuit de spires, ruptures de barres rotoriques, excentricité d'entrefer, etc.). En effet, il compare le spectre du courant d'alimentation du moteur avec celui du signal vibratoire. Il en déduit, bien que les spectres ne soient pas totalement identiques, que les informations présentes dans le courant correspondent à celles contenues dans les vibrations avec en plus les informations propres aux phénomènes électriques (glissement,...).

Cette approche de surveillance des entraînements électriques a été largement utilisée depuis ces dix dernières années [Nandi, 99], [Beguenane, 97], [Thomson, 01]; elle est connue sous le nom de MCSA (Motor Current Signature Analysis).

L'avantage, comparé aux signaux précédents, est que les capteurs de courants sont désormais présents dans tous les systèmes de commande des entraînements électriques et offrent donc, un accès aisé à la mesure sans une implémentation (de capteur) supplémentaire.

### 2.2.2.2. Détection des défauts par analyse des courants statoriques

La mesure des signaux puis leurs traitements dans le domaine fréquentiel ne peuvent servir, que si les composantes fréquentielles définies pour chaque défaut [Benbouzid, 99] sont connues.

Il y a plusieurs techniques pour effectuer l'analyse des courants. La plus utilisée est de réaliser une analyse fréquentielle (FFT par exemple) directement sur les courants statoriques  $i_{sa}$ ,  $i_{sb}$  ou  $i_{sc}$ . Cette méthode, connue sous le nom de MCSA, est largement employée depuis ces dernières années. Mais, bien qu'elle permette d'obtenir de bons résultats, elle présente quelques inconvénients. En effet, prenons un exemple ; la signature fréquentielle spécifique

au défaut "barres cassées" est la forte modulation des harmoniques du courant par la fréquence de glissement. Dans cette analyse fréquentielle, il est préférable de faire une étude globale sur les courants car un défaut rotor est visible sur le spectre des trois courants de ligne. Mais faire une FFT sur chaque courant risque d'être pénalisante en temps de calcul.

C'est pour cela que nous avons opté pour l'analyse spectrale de la combinaison des trois courants, réalisée à l'aide de la matrice de Fortescue (composantes symétriques).

### 2.2.2.2.1. Préconditionnement des courants statoriques pour l'analyse fréquentielle

De manière générale, tout système de trois grandeurs de même nature et de même fréquence est la superposition de trois systèmes équilibrés de même fréquence : un système triphasé direct, un système triphasé inverse et un système homopolaire.

En pratique, cela signifie que le système des trois tensions ou des trois courants, peut être représenté par les composantes  $\underline{X}_d$ ,  $\underline{X}_i$  et  $\underline{X}_0$  (tensions ou courants) telles que :

- $\underline{X}_d$  représente la composante symétrique directe
- $\underline{X}_i$  représente la composante symétrique inverse
- $\underline{X}_0$  représente la composante homopolaire

Les équations donnant les composantes symétriques des tensions et des courants (transformation de Fortescue) sont les suivantes :

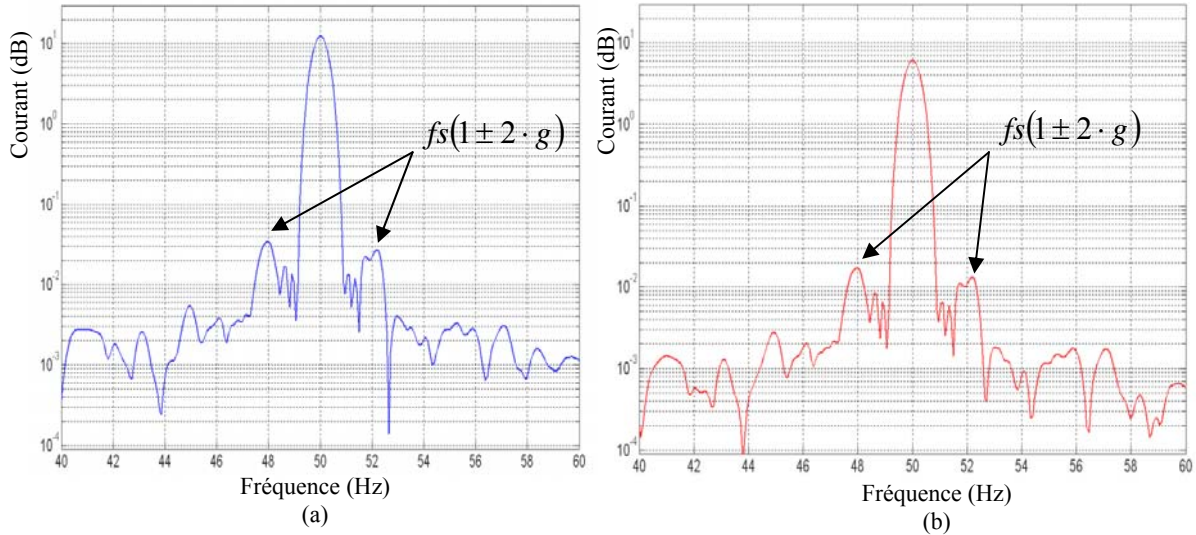
$$\begin{cases} \underline{X}_d = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{a} \cdot \underline{X}_{sb} + \underline{a}^2 \cdot \underline{X}_{sc}) \\ \underline{X}_i = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{a}^2 \cdot \underline{X}_{sb} + \underline{a} \cdot \underline{X}_{sc}) \\ \underline{X}_0 = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{X}_{sb} + \underline{X}_{sc}) \end{cases} \quad (2.21)$$

Ainsi l'expression de la composante symétrique directe (2.22), notée  $\underline{I}_{sd}$ , et obtenue à partir de la combinaison des trois courants statoriques  $i_{sa}$ ,  $i_{sb}$ ,  $i_{sc}$  et de la matrice de Fortescue est :

$$\underline{I}_{sd} = i_{sa} + \underline{a} \cdot i_{sb} + \underline{a}^2 \cdot i_{sc} \quad (2.22)$$

Où l'opérateur  $\underline{a}$  représente une rotation de  $\frac{2\pi}{3}$  :  $\underline{a} = e^{\frac{j2\pi}{3}}$ .

Pour la suite, l'analyse spectrale est réalisée sur la partie réelle de  $\underline{I}_{sd}$ , notée  $\text{Re}[\underline{I}_{sd}]$ . La figure 2.5 représente l'analyse fréquentielle réalisée sur le courant de la phase a ( $i_{sa}$ ) et celle réalisée sur la partie réelle de  $\underline{I}_{sd}$ .



**Fig. 2.5.** Analyse fréquentielle pour une machine avec 1 barre cassée en charge. FFT sur  $\text{Re}[\underline{I}_{sd}]$  (a). FFT sur le courant de phase  $i_{sa}$  (b)

Sur la figure 2.5, que ce soit pour l'analyse fréquentielle sur le courant de phase  $i_{sa}$  ou sur la partie réelle de  $\underline{I}_{sd}$ , on distingue bien l'apparition des raies caractéristiques d'un défaut 'barres cassées' aux fréquences  $f_s(1 \pm 2 \cdot g)$ . L'intérêt de réaliser, par la suite, une FFT sur  $\text{Re}[\underline{I}_{sd}]$  pour détecter l'apparition de raies caractéristiques, se situe donc au niveau du temps de calcul.

#### 2.2.2.2.2. Signatures fréquentielles extraites à partir des signaux préconditionnés

##### 2.2.2.2.2.1. Ruptures de barres ou de portions d'anneaux

La rupture de barres d'une machine asynchrone est un des défauts le plus couramment étudié en laboratoire en raison à sa simplicité de réalisation mais non nécessairement le plus probable. Ce défaut induit des modifications dans les courants du stator et entraîne donc l'apparition d'harmoniques caractéristiques dans le spectre de ce signal [Deleroi, 82].

En effet, lors de l'apparition d'une rupture de barre, des harmoniques de flux sont produits et induisent des harmoniques de courant dans l'enroulement statorique aux fréquences :

$$f = f_s(1 \pm 2 \cdot g) \quad (2.23)$$

De plus les amplitudes des raies dans ces bandes latérales du fondamental augmentent avec le nombre de barres défectueuses [Stein, 88].

### 2.2.2.2.2. Excentricité

La variation de la longueur de l'entrefer entraîne des variations dans la densité du flux d'entrefer. Les enroulements statoriques sont électriquement modifiés, ce qui entraîne des effets dans le spectre des courants.

L'identification du type d'excentricité est possible, par la visualisation dans ce spectre, des composantes fréquentielles suivantes :

Pour l'excentricité statique :

$$f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot N_r (1-g)}{p} \right] \quad (2.24)$$

Pour l'excentricité dynamique :

$$f_{exc} = f_s \cdot [1 \pm 2 \cdot n(1-g)] \quad (2.25)$$

Pour le désalignement :

$$f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot (1-g)}{p} \right] \quad (2.26)$$

L'origine du désalignement est introduite dans le paragraphe suivant.

### 2.2.2.2.3. Défauts roulements

A l'interface entre le rotor et le stator, le roulement à billes (ou palier) présente un vieillissement plus ou moins rapide. Classiquement, ce défaut est diagnostiqué à partir du spectre d'une mesure acoustique ou vibratoire.

En partant du principe que l'entrefer n'est plus constant lors d'une défaillance des roulements, [Schoen, 95a] montre que cette information est aussi présente dans le spectre du courant. Dans cet article, l'expression analytique des fréquences liées aux défauts du roulement en fonction de ses paramètres mécaniques (diamètre de bille DB, diamètre primitif du palier DP) est explicitée et vérifiée par des essais. Les raies se trouvent à des fréquences différentes de celles dues à des défauts tels que les harmoniques d'encoches ou d'excentricité dynamique ou statique.

Ces fréquences sont déduites des dimensions des paliers et des roulements à billes (DP et DB).

Les composantes fréquentielles spécifiques aux vibrations des paliers notées  $f_{pal}$  sont définies par la relation (2.27).

$$f_{pal} = f_s \pm n \cdot f_{i,e} \quad (2.27)$$

$n = 1, 2, 3, \dots$  "i, e" : interne, externe (bagues)

Les fréquences des bagues internes et externes de roulement sont définies respectivement par les expressions (2.28) et (2.29) en fonction du nombre de billes noté  $n_b$

$$f_e = \frac{n_b}{2} \cdot f_r \cdot \left[ 1 - \frac{DB}{DP} \cdot \cos \beta \right] \quad (2.28)$$

$$f_i = \frac{n_b}{2} \cdot f_r \cdot \left[ 1 + \frac{DB}{DP} \cdot \cos \beta \right] \quad (2.29)$$

Où  $f_r$  est la fréquence de rotation du rotor et  $\beta$  l'angle de contact.

Cependant, ces informations sont fonction des données du constructeur. On peut toutefois approcher les expressions de  $f_e$  et  $f_i$  pour un nombre faible de billes dans les roulements [Schoen, 95a] par :

$$\begin{aligned} f_e &= 0.4 \cdot n_b \cdot f_r \\ f_i &= 0.6 \cdot n_b \cdot f_r \end{aligned} \quad (2.30)$$

#### 2.2.2.2.4. Défauts au niveau du stator

Le stator est soumis à des contraintes électriques et mécaniques. Il présente essentiellement comme défaut (Chapitre I), la mise en court-circuit d'une ou plusieurs spires de son bobinage, le déséquilibre ou l'ouverture d'une des phases d'alimentation. Dans [Thomson, 03], l'auteur présente ces deux types de défauts ainsi que leurs conséquences dans le spectre du courant. En effet, l'harmonique de rang 3 et celui lié au passage des encoches rotoriques sont significativement augmentés lors d'une rupture de l'isolant des spires statoriques.

Pour sa part, [Casimir, 03] a utilisé, pour détecter un déséquilibre d'alimentation au stator (mais aussi pour d'autres défauts), la puissance moyenne du signal et l'impédance directe.

Les deux sont basés sur l'expression de la composante symétrique directe (2.22).

Désignons par  $I_d(f)$  la transformée en fréquence de  $\text{Re}[I_{sd}]$ . La densité de puissance spectrale du courant est :

$$S_d(f) = I_d(f) \cdot \overline{I_d(f)} = |I_d(f)|^2 \quad (2.31)$$

Le premier paramètre utilisé est la puissance moyenne du signal :

$$P_{1d} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} S_d(f) df \quad (2.32)$$



Considérant les transformées en fréquence,  $I_d(f)$  et  $V_d(f)$  de  $\text{Re}[I_{sd}]$  et  $\text{Re}[V_{sd}]$ , le deuxième paramètre, l'impédance directe  $Z_d$ , est calculée à partir du ratio entre les composantes fondamentales de  $V_d(f)$  et  $I_d(f)$  :

$$Z_d = \frac{V_d(f=f_s)}{I_d(f=f_s)} \quad (2.33)$$

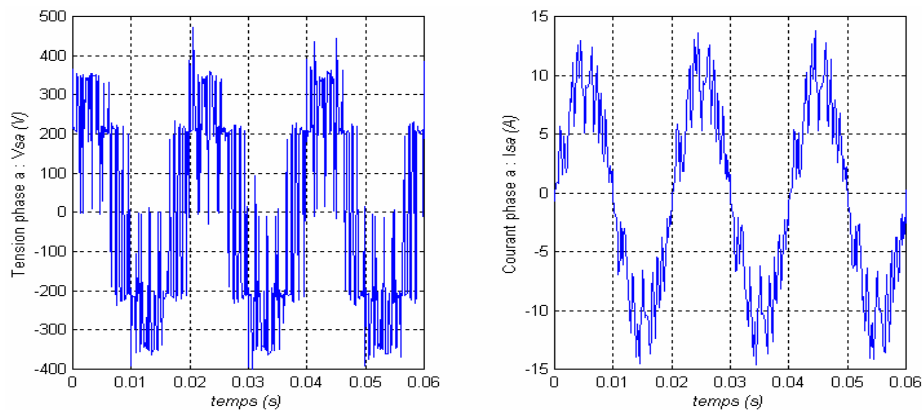
D'autres travaux [Cabanas, 98], [Assaf, 01], ont aussi montré l'importance du calcul de l'impédance directe pour la détection de spires en court-circuit ou pour le déséquilibre d'alimentation.

### 2.2.2.3. Signatures statistiques

#### 2.2.2.3.1. Extraites à partir des signaux mesurés

Nous avons vu précédemment que l'analyse spectrale des différents signaux mesurés (ou préconditionnés), et particulièrement celle des courants statoriques, permet de réaliser la détection d'une majorité des défauts d'une machine asynchrone alimentée sur le réseau.

Cependant, lorsque la machine est associée à un onduleur, les courants (ainsi que les tensions) sont affectés par les multiples harmoniques de la fréquence de commutation. Les formes d'ondes des signaux (figure 2.6) ainsi que leurs contenus fréquentiels rendent difficiles l'identification des raies caractéristiques. De plus, ces raies n'existent quasiment pas à très faible charge.



**Fig. 2.6.** Tension générée par l'onduleur sur une phase de la machine et le courant de ligne correspondant

De plus, dans son étude, [Schoen, 95b] montre que toute variation du couple interfère sur la forme du spectre d'une phase et en particulier des raies couramment utilisées pour le diagnostic des défaillances telles que la rupture de barres ou l'excentricité du rotor.

Cette étude suggère donc la décomposition des trois courants sur les axes 'αβ' (§ 2.2.1.2.2) liés au rotor ou au stator. Cette transformation autoriserait alors de mieux appréhender les composantes liées au couple mécanique comparées à celles engendrées par la variation du circuit magnétique.

Pour toutes ces raisons, d'autres paramètres caractéristiques de défauts doivent être déterminés, issus de divers types d'analyses autres que l'analyse fréquentielle.

Certains auteurs utilisent alors d'autres signatures telles que la puissance instantanée ou encore le facteur de puissance. En effet, ces signaux possèdent une information plus riche car ils contiennent aussi bien une information liée au courant, à la tension ainsi qu'une mesure du déphasage entre ces deux grandeurs. [Chang, 94] démontre ainsi que le facteur de puissance est aussi un bon indicateur pour la détection d'excentricité au rotor.

Ces signatures permettent aussi la détection de certains défauts mécaniques du fait que la puissance est le produit du couple par la vitesse [Sidahmed, 97].

### 2.2.2.3.2. Extraites à partir des signaux préconditionnés

[Casimir, 03] a développé des signatures à l'aide de la transformation triphasée-biphasée [Cruz, 00] qui est basée sur une représentation bidimensionnelle des composantes triphasées du courant stator (§ 2.2.1.2.2). Les composantes courant sont appelées  $I_{s\alpha}$  et  $I_{s\beta}$  et s'exprime selon l'équation suivante :

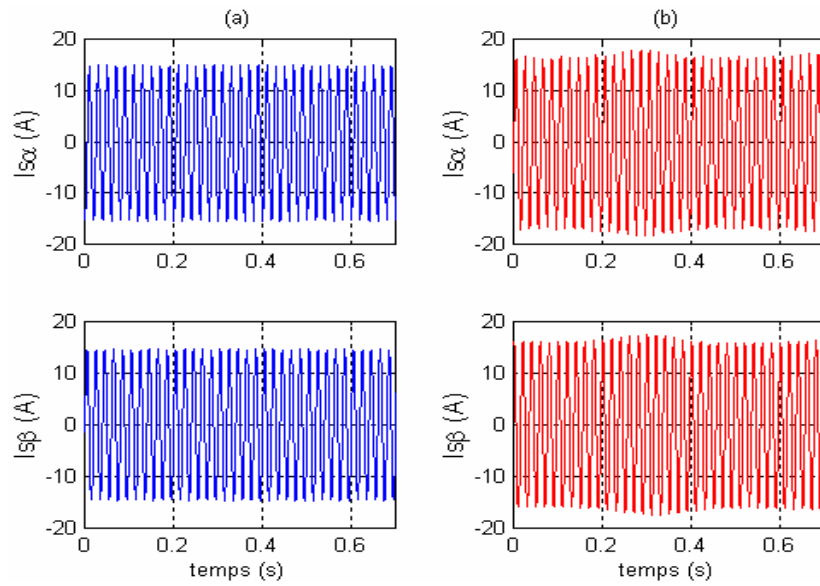
$$\begin{bmatrix} I_{s\alpha} \\ I_{s\beta} \\ I_{s0} \end{bmatrix} = [C] \cdot \begin{bmatrix} i_{sa} \\ i_{sb} \\ i_{sc} \end{bmatrix} \quad (2.34)$$

Avec [C], la matrice de Concordia définie par l'équation 2.5.

La transformation est la même pour les flux et les tensions.

Son approche étant une approche statistique, il est nécessaire de réaliser plusieurs enregistrements de signaux pour travailler. Ces signatures sont présentées dans ce qui suit.

L'étude porte sur l'évolution temporelle des composantes en quadrature  $I_{s\alpha}$  et  $I_{s\beta}$ . Celles-ci sont modulées en amplitude en présence de barres cassées tout comme les courants au stator ( $i_{sa}$ ,  $i_{sb}$  et  $i_{sc}$ ). La figure 2.7b illustre l'affectation temporelle des courants en alimentation réseau, représentée par une modulation en amplitude des courants, dans le cas du défaut rotor (4 barres cassées dans cet exemple). La figure 2.7a illustre quant à elle, l'évolution temporelle des courants en fonctionnement sain.



**Fig. 2. 7.** Evolution des composantes  $I_{s\alpha}$  et  $I_{s\beta}$  des courants stators pour un rotor sain (a) et pour un rotor ayant 4 barres cassées (b)

### 2.2.2.3.2.1. Calcul sur les composantes $I_{s\alpha}$ et $I_{s\beta}$

Les premiers paramètres déterminés sont directement liés à l'évolution temporelle des signaux et à la représentation dans le plan de phase du courant stator. Ils sont cinq au total.

Considérons le module du courant stator, noté  $I_{s\alpha,\beta}$ , en fonction des composantes en quadratures (2.35).

$$I_{s\alpha,\beta} = \left\| \overrightarrow{I_{s\alpha,\beta}} \right\| = \sqrt{I_{s\alpha}^2 + I_{s\beta}^2} \quad (2.35)$$

De la même façon, le module de la tension est donné par

$$V_{s\alpha,\beta} = \left\| \overrightarrow{V_{s\alpha,\beta}} \right\| = \sqrt{V_{s\alpha}^2 + V_{s\beta}^2} \quad (2.36)$$

Ainsi, à partir des trois signaux disponibles ( $I_{s\alpha}, I_{s\beta}, I_{s\alpha,\beta}$ ), il s'agit dans un premier temps, de calculer les écarts-types, à l'aide de leurs valeurs moyennes notées respectivement  $m_{\alpha}$ ,  $m_{\beta}$  et  $m_{\alpha,\beta}$  :

$$\begin{cases} \sigma_\alpha = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\alpha}(k) - m_\alpha) \right]^{1/2} \\ \sigma_\beta = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\beta}(k) - m_\beta) \right]^{1/2} \\ \sigma_s = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\alpha,\beta}(k) - m_{\alpha,\beta}) \right]^{1/2} \end{cases} \quad (2.37)$$

Où  $N_p$  est le nombre de points dans les signaux.

Les caractéristiques  $I_{s\beta}(I_{s\alpha})$  (figures 2.8 et 2.9), normalisées par rapport à la valeur efficace de  $I_{s\alpha,\beta}$  (2.38) ont permis de calculer deux autres paramètres.

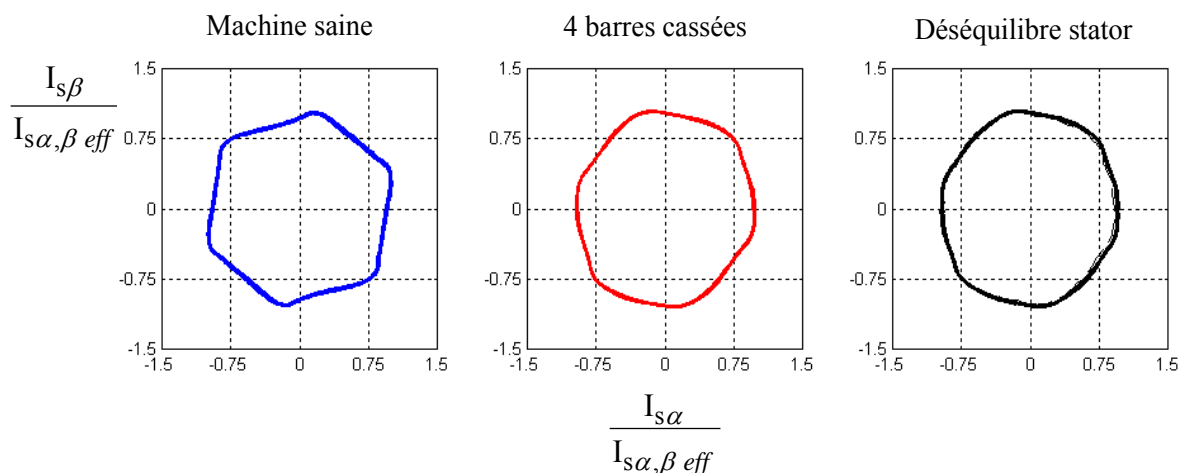
$$I_{s\alpha,\beta \text{ eff}} = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} I_{s\alpha,\beta}^2(k) \right]^{1/2} \quad (2.38)$$

La caractéristique  $I_{s\beta}(I_{s\alpha})$  traduit l'évolution du courant stator  $\underline{I}_{s\alpha,\beta}$  dans le plan complexe :

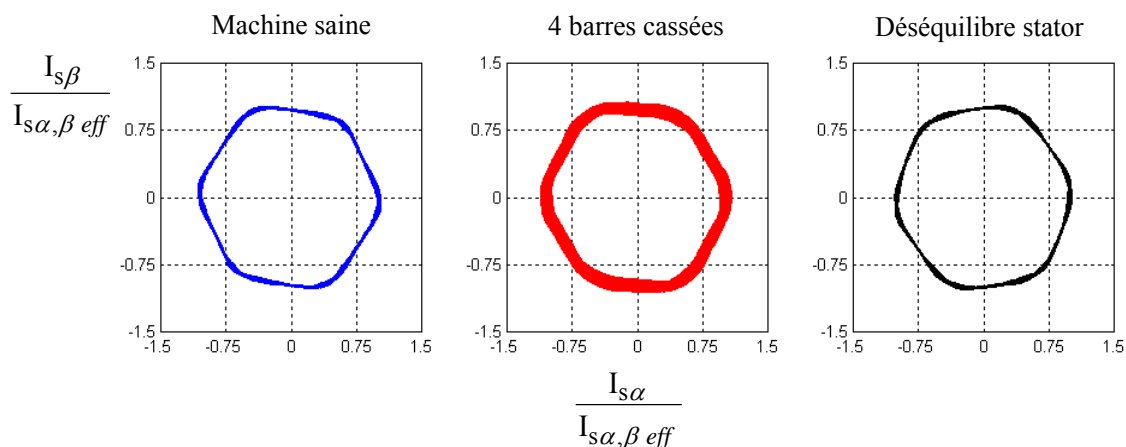
$$\underline{I}_{s\alpha,\beta} = I_{s\alpha} + jI_{s\beta} = I_{s\alpha,\beta} e^{j\varphi} \quad (2.39)$$

Les deux paramètres, notés  $\delta_\alpha$  et  $\delta_\beta$  sont les mesures de la déformation de cette caractéristique à  $\varphi = 0$  et à  $\varphi = \frac{\pi}{2}$ , autrement dit dans les axes  $\alpha$  et  $\beta$  :

$$\begin{cases} \delta_\alpha = \left| \max(I_{s\alpha,\beta}) - \min(I_{s\alpha,\beta}) \right|_{\varphi=0} \\ \delta_\beta = \left| \max(I_{s\alpha,\beta}) - \min(I_{s\alpha,\beta}) \right|_{\varphi=\frac{\pi}{2}} \end{cases} \quad (2.40)$$



**Fig. 2.8.** Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée par rapport  $I_{s\alpha, \beta eff}$ . Fonctionnement à vide.



**Fig. 2.9.** Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée par rapport  $I_{s\alpha, \beta eff}$ . Fonctionnement en charge.

Les paramètres  $\sigma_\alpha$  et  $\sigma_\beta$  fournissent une bonne distinction du fonctionnement à vide par rapport au fonctionnement en charge, mais un faible niveau de détection des défauts par rapport au fonctionnement sain. Le paramètre  $\sigma_s$  témoigne d'une relative sensibilité aux états de la machine (rotor en défaut ou déséquilibre d'alimentation au stator), que ce soit à vide ou en charge. En ce qui concerne les paramètres  $\delta_\alpha$  et  $\delta_\beta$ , on constate que la différence est délicate à faire pour les fonctionnements à vide (figure 2.8), par contre, pour le défaut rupture de barres en charge (figure 2.9), ces caractéristiques offrent des résultats intéressants.

### 2.2.2.3.2.2. Calcul sur les puissances

3 autres paramètres sont déterminés à partir des puissances active et réactive (P et Q respectivement), qui sont calculées à l'aide des composantes en quadrature des tensions et des courants [Vas, 92] :

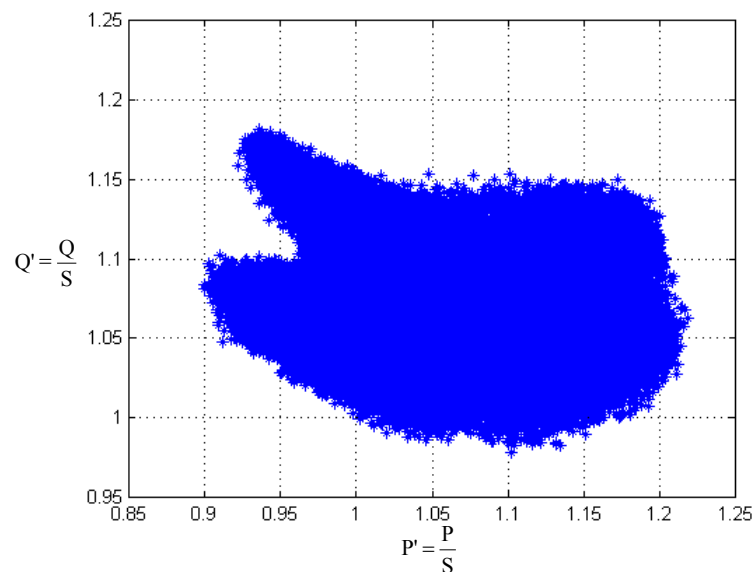
$$\begin{aligned} P &= V_{s\alpha} \cdot I_{s\alpha} + V_{s\beta} \cdot I_{s\beta} \\ Q &= V_{s\beta} \cdot I_{s\alpha} - V_{s\alpha} \cdot I_{s\beta} \end{aligned} \quad (2.41)$$

Les deux premiers sont déterminés à partir des puissances réactives et actives, normalisées par la valeur RMS du module de puissance apparente  $S = \sqrt{P^2 + Q^2}$  et notées  $Q'$  et  $P'$ . Leurs valeurs moyennes, «  $m_q$  » et «  $m_p$  », sont calculées pour chaque acquisition telle que :

$$m_p = \frac{1}{N_p} \sum_{k=1}^{N_p} P'(k) \quad (2.42)$$

$$m_q = \frac{1}{N_p} \sum_{k=1}^{N_p} Q'(k) \quad (2.43)$$

Le troisième est déterminé à partir de l'analyse de la caractéristique normalisée  $Q'(P')$  qui est composée d'un nuage de  $N_p$  points (figure 2.10). Les points de cette caractéristique représentent de manière normalisée l'évolution de la puissance apparente consommée par la machine à chaque instant.



**Fig. 2. 10.** Caractéristique normalisée  $Q'(P')$  obtenue pour la machine saine, pour une charge de 3 kW en alimentation réseau

Le paramètre représente la dispersion du nuage de points, et est noté  $\zeta$  (2.44).

$$\zeta = \sum_{k=1}^{N_p} (s_k - m_{p,q})(s_k - m_{p,q})^T \quad (2.44)$$

avec  $m_{p,q} = [m_p, m_q]$ , le centre de gravité de l'ensemble et  $s_k$  un point de coordonnées  $(P'(k), Q'(k))$ .

[Casimir, 03] montre que le paramètre  $\zeta$  est relativement sensible aux différents états de la machine tels qu'une rupture de barres ou encore un déséquilibre d'alimentation. En ce qui concerne les paramètres  $m_p$  et  $m_q$ , il montre que ces paramètres sont corrélés avec le niveau de charge, et que, associés à des paramètres sensibles du point de vue des défaillances, ces paramètres pourraient s'avérer pertinents car ils rendent compte de la puissance mise en jeu.

#### 2.2.2.4. Les outils d'analyse fréquentielle liés au traitement de l'information

Comme nous venons de le présenter, une défaillance sur le système étudié engendre l'apparition et/ou la modification d'harmoniques au niveau des spectres des signaux précédemment cités.

De nombreux outils de traitement du signal sont utilisés afin de caractériser les origines des défauts à partir des spectres.

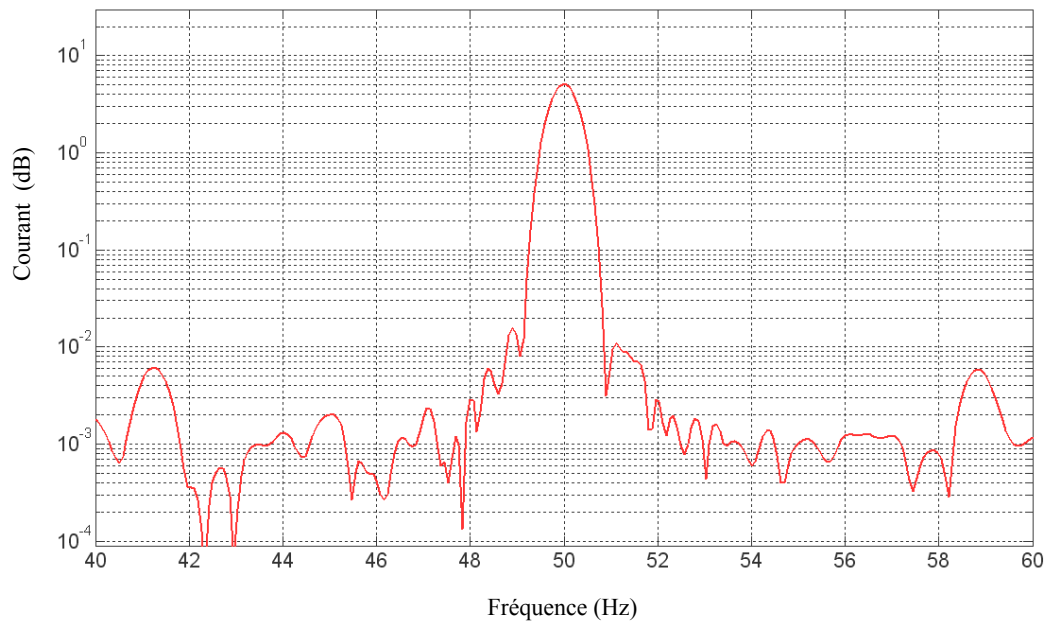
Classiquement, la méthode la plus utilisée est la Transformée de Fourier (FFT). Il s'agit d'une méthode non paramétrique, qui se prête bien à l'analyse des phénomènes stationnaires.

La composante à la fréquence  $f$  d'un signal temporel  $x(t)$  s'exprime par :

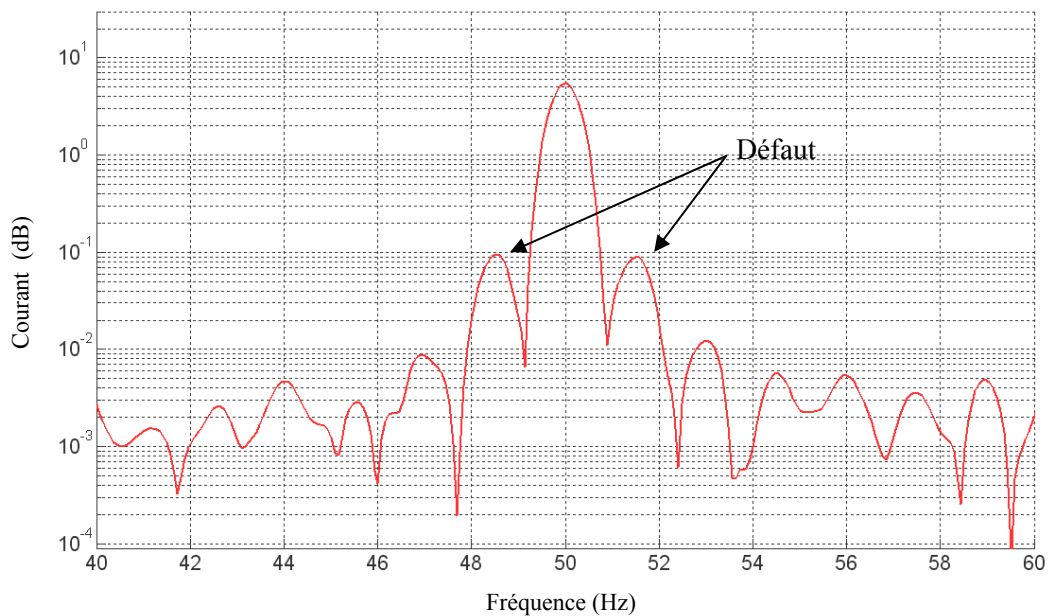
$$X(f) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-j \cdot 2\pi \cdot f \cdot t} \quad (2.45)$$

Il est impossible d'analyser le signal sur une période infinie. Il est donc nécessaire de choisir une fenêtre de pondération pour l'analyse (fenêtre rectangulaire, de Hanning, ...) pour corriger les effets d'un fenêtrage temporel fini. De plus, outre la forme, la taille de la fenêtre aura une influence sur la résolution maximale possible.

Nous avons donc appliqué l'analyse de Fourier sur plusieurs périodes du régime permanent de la machine.



**Fig. 2.11.** FFT sur la partie réelle de la composante  $I_{sd}$  pour une machine saine en charge



**Fig. 2.12.** FFT sur la partie réelle de la composante  $I_{sd}$  pour une machine avec 3 barres cassées en charge

Ceci fait apparaître deux pics aux fréquences  $f_s(1 \pm 2g)$  et permet ainsi d'extraire l'information de la présence de barres cassées au rotor.

Cependant, la FFT présente des limites quant au traitement des signaux en régime transitoire et nécessite un volume de calcul important (le pas d'analyse fréquentiel est inversement proportionnel à la durée de l'enregistrement et la bande de fréquence analysée doit être inférieure à  $f_{ech}/2$  avec  $f_{ech}$  la fréquence d'échantillonnage).



En effet la précision fréquentielle dépend de la fréquence d'échantillonnage et du nombre  $N_e$  d'échantillons de l'enregistrement :

$$\Delta f = \frac{f_{ech}}{\Delta t \cdot f_{ech}} = \frac{f_{ech}}{N_e} \quad (2.46)$$

La transformée de Fourier est surtout utilisée pour l'analyse des systèmes en régimes permanents. Cependant, il est possible de la mettre en œuvre pour l'étude de fonctionnements dynamiques (lents), par la méthode du spectrogramme [Flandrin, 93].

Certaines études se basent sur cette analyse bispectre. [Chow, 95] montre que l'intérêt de cette analyse est la prise en compte à la fois des informations sur l'amplitude et la phase.

Le spectrogramme est le calcul répétitif de la transformée de Fourier sur une fenêtre temporelle glissant sur le signal. Son calcul sera donc soumis aux paramètres suivants :

- longueur de la fenêtre,
- type de fenêtre de pondération,
- longueur totale du signal analysé,
- pas de glissement de la fenêtre ou pas temporel.

$$S_x(t, \nu) = \left| \int_{-\infty}^{+\infty} x(s) \cdot h^*(s-t) e^{-i2\pi\nu s} ds \right|^2 \quad (2.47)$$

où  $x(s)$  : le signal temporel

$h^*(s-t)$  : la forme conjuguée de la fenêtre choisie.

Le problème majeur provient du fait que la durée du régime transitoire des machines électriques est généralement très faible, ce qui oblige à prendre une fenêtre glissante très petite et engendre une précision fréquentielle réduite.

Ainsi, d'autres méthodes ont été développées pour représenter les signaux dans le plan temps -fréquence et sont connues sous le nom de 'Distribution de Wigner Ville'. [Andria, 94] illustre l'intérêt de cette méthode. Elle permet, entre autres, la détermination de la fréquence instantanée, pour l'analyse des signaux non stationnaires ou des phénomènes transitoires.

En ce qui concerne le diagnostic des machines asynchrones, [Jianguo, 92] montre que le ratio des amplitudes des composantes  $f_s(1 \pm 2g)$  sur l'amplitude du fondamental est plus important en régime transitoire qu'en régime permanent. L'analyse de Wigner Ville, qui combine l'analyse spectrale à celle temporelle, paraît plus appropriée que l'analyse de Fourier ou le spectrogramme pour le régime transitoire.

Soit  $x(t)$ , un signal complexe. La distribution continue de Wigner Ville est définie par :

$$W_x(t, \nu) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-2i\pi\nu\tau} d\tau \quad (2.48)$$

ou \* désigne la notation conjuguée.

La relation précédente est définie sur tout l'espace temporel, or d'un point de vue pratique, la plupart des signaux analysés sont de longueurs finies. De ce fait, la transformée Pseudo-Wigner Ville se définit par :

$$W_x(t, \nu) = \int_{-\infty}^{+\infty} p(\tau) \cdot x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-2i\pi\nu\tau} d\tau \quad (2.49)$$

$$\text{avec } p(\tau) = h\left(\frac{\tau}{2}\right) \cdot h\left(-\frac{\tau}{2}\right)$$

On peut voir dans [Jianguo, 92], une analyse temps-fréquence d'une machine saine et d'une machine avec deux barres cassées.

[Bourbao, 96] montre le lien étroit qu'il existe entre les mathématiques statistiques et l'analyse du signal. Il montre bien la difficulté du choix d'une transformée par rapport à une autre. En effet, ces méthodes ne sont pas utilisables durant tous les régimes de fonctionnement. En régime permanent, le nombre de points d'acquisition est suffisant pour faire tous les types d'analyse tandis qu'en régime transitoire, en général très rapide, ce nombre de points est limité et rend impossible l'utilisation de toutes ces méthodes.

Cependant, quelque soit la transformée choisie, les informations qui en découlent ne seront pertinentes que si le signal est efficacement mesuré et riche en informations.

Comme il a été présenté dans le premier chapitre, de nombreuses défaillances ont pour conséquence l'élévation de la température au sein de la machine. Ces surchauffes risquent d'entraîner une détérioration irrémédiable du moteur en faisant vieillir prématurément les isolants.

La température peut donc s'avérer, dans quelques cas de défauts, porteuse d'informations. Cependant, sa mesure reste toujours très délicate et difficile à maîtriser.

### 2.2.2.5. Signatures basées sur l'échauffement de la machine

La machine est le siège de pertes (mécaniques, électriques et magnétiques, ...) mais aussi de pertes supplémentaires dues à certaines défaillances. Ces pertes provoquent un échauffement des matériaux et en particulier des isolants.

A partir de la modélisation thermique de la machine asynchrone par un modèle approché à un corps, l'échauffement de celle-ci peut se déduire de la formulation suivante :

$$\Delta\theta(t) = R_t \cdot P_{tot} \cdot (1 - e^{-\frac{t}{R_t \cdot C_t}}) = \Delta\theta_m \cdot \left(1 - e^{-\frac{t}{\tau_{th}}}\right) \quad (2.50)$$

$P_{tot}$	Puissance dissipée par la machine (pertes totales)
$R_t$	Résistance thermique
$C_t$	Capacité thermique
$R_t \cdot P_{tot} = \Delta\theta_m$	Avec $\Delta\theta_m$ , échauffement maximal permanent qui ne devra jamais dépasser le $\theta_l$ limite des isolants
$\tau_{th} = R_t \cdot C_t$	Constante de temps thermique

Si l'on tient compte de la variation de la résistance en fonction de la température, l'expression 2.50 devient alors :

$$\Delta\theta(t) = \Delta\theta'_m \cdot (1 - e^{-\frac{t}{\tau_{th}'}}) \quad (2.51)$$

$$\Delta\theta'_m = R_t' \cdot P_{tot} = \frac{\Delta\theta_m}{1 - \alpha \cdot \Delta\theta_m} : \text{valeur finale de l'échauffement,}$$

$$R_t' = \frac{R_t}{1 - \alpha \cdot \Delta\theta_m} : \text{valeur de la résistance en fonction de la température, avec } \alpha \text{ le}$$

coefficient de température de la résistance

$$\tau_{th}' = C_t \cdot R_t' = \frac{\tau_{th}}{1 - \alpha \cdot \Delta\theta_m} : \text{constante de temps thermique}$$

Ainsi pour connaître l'échauffement présent au niveau de la machine, la puissance dissipée par celle-ci, c'est-à-dire ses pertes, doit être calculée.

A partir d'un essai à vide réalisé sur la machine, il est possible de déduire les pertes à vide, appelée généralement pertes constantes et notée  $P_{cste}$  (elles regroupent en fait les pertes mécaniques et les pertes fer au stator. Ensuite, à l'aide d'un essai en charge, il reste à déterminer les pertes par effet joules rotoriques  $P_{jr}$  et statoriques  $P_{js}$ . Au final, la somme totale des pertes vaut :

$$P_{tot} = P_{js} + P_{jr} + P_{cste} \quad (2.52)$$

A partir de l'expression (2.52), et connaissant maintenant la puissance dissipée  $P_{tot}$ , il est possible d'estimer l'échauffement au niveau de l'induit.

Pour cela, la formulation suivante, tirée de l'équation 2.51, est utilisée, prenant en compte la variation de résistance d'induit :

$$\Delta\theta(t) = R_t' \cdot P_{tot} \cdot (1 - e^{-\frac{t}{R_t' \cdot C_t}}) \quad (2.53)$$

Deux signatures sont donc extraites à partir du signal température qui sont, les pertes totales  $P_{tot}$  présentes dans la machine, ainsi que l'échauffement  $\Delta\theta$  que celles-ci engendrent.

Nous étudierons leur sensibilité par rapport aux différents états de fonctionnement de la machine dans le chapitre IV.

L'ensemble des calculs ainsi que la méthode utilisée pour déterminer les pertes de la machine et l'échauffement présent dans celle-ci, sont détaillés dans l'annexe 1.

### 2.2.2.6. Synthèse des signatures déterminées à partir des méthodes externes

Les tableaux suivants regroupent l'ensemble des signatures, aussi bien fréquentielles que statistiques extraites à partir des signaux mesurés tels que le courant, les flux, les vibrations et le couple électromagnétique.



Type de défaut	Signatures fréquentielles		Signatures 'statistiques'								
	Courants mesurés ou préconditionés	Couple, Flux, Vibrations									
Ruptures de barres ou de portions d'anneaux	$f = f_s \cdot (1 \pm 2g)$ $F_{enc} \pm 2n f_s \pm 2g f_s$	<p>Fréq. d'encoches sur <math>C_{em}</math> estimé :</p> $f_{sb} = f_s \cdot \left[ \frac{N_r(1-g)}{p} \pm n \right] \pm 2 \cdot g \cdot f_s$ <p>Flux axial :</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td><math>f_s</math></td> <td><math>(3-2g)f_s</math></td> <td><math>(2g-f_s)</math></td> <td><math>(4g-3)f_s</math></td> <td><math>(5-4g)f_s</math></td> <td><math>(6g-5)f_s</math></td> <td><math>(7-6g)f_s</math></td> <td><math>(8g-7)f_s</math></td> </tr> </table> <p>Vibrations :</p> $f_r \pm 2 \cdot g \cdot f_s$	$f_s$	$(3-2g)f_s$	$(2g-f_s)$	$(4g-3)f_s$	$(5-4g)f_s$	$(6g-5)f_s$	$(7-6g)f_s$	$(8g-7)f_s$	$\delta_\alpha$ : Déformation de la caractéristique $I_{s\beta}(I_{s\alpha})$ dans l'axe $\alpha$ $\delta_\beta$ : Déformation de la caractéristique $I_{s\alpha}(I_{s\beta})$ dans l'axe $\alpha$ $\sigma_s$ : Ecart – type du module de $I_s = I_{s\alpha} + I_{s\beta}$ $\zeta$ : Dispersion des points de la caractéristique Q(P)
$f_s$	$(3-2g)f_s$	$(2g-f_s)$	$(4g-3)f_s$	$(5-4g)f_s$	$(6g-5)f_s$	$(7-6g)f_s$	$(8g-7)f_s$				
Excentricité : - statique - dynamique - mixte	<p>Statique :</p> $f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot N_r(1-g)}{p} \right]$ <p>Dynamique :</p> $f_{exc} = (N_r \cdot f_s + f_s) \pm f_r$ <p>Mixte :</p> $f_{exc} = f_s \cdot \left[ 1 \pm n \frac{1-g}{p} \right]$	<p>Vibrations :</p> <p>Excentricité statique :</p> $2 \cdot f_s$ ou des composantes à $k \cdot f_r \pm f_s$ <p>Excentricité dynamique :</p> $f_r \pm 2 \cdot g \cdot f_s$ ou $f_r$ <p>Couple utile estimé par modèle triphasé <math>C_u</math>:</p> $f_{Cu} = n \cdot \left( 1 - \frac{1+g}{p} \right) \cdot f_s$	$\cos \varphi$ : Facteur de puissance								

Tableau. 2. 7. Signatures déterminées à partir des signaux mesurés ou estimés

Type de défaut	Signatures fréquentielles		Signatures 'statistiques'
	Courants mesurés ou préconditionnés	Couple, Flux, Vibrations	
Défauts roulements	$f_{pd} = f_s \pm n \cdot f_{i,e}$ <p>avec :</p> $f_e = 0.4 \cdot n_b \cdot f_r$ $f_i = 0.6 \cdot n_b \cdot f_r$	<p><u>Flux axial :</u></p> <p>Raies situées entre 600-900 Hz pour <math>f_s = 50\text{Hz}</math></p> <p><u>Vibrations :</u></p> <p>Bagues externes : <math>n \cdot f_e</math>                      bagues internes : <math>f_i \pm k \cdot f_r</math></p>	<p>Puissance instantanée</p> <p><math>m_p</math> : puissance active moyenne</p> <p><math>m_q</math> : puissance réactive moyenne</p>
Désalignement	$f_s \pm k \cdot f_r$	<p><u>Vibrations :</u></p> <p><math>f_r</math> et <math>2 \cdot f_r</math></p> <p>principalement</p>	
Déséquilibre d'alimentation stator	<p>Puissance moyenne de la composante directe</p> <p><math>P_{1d}</math> :</p> $P_{1d} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} S_d(f) df$ <p>Impédance directe :</p> $Z_d = \frac{V_d(f=f_s)}{I_d(f=f_s)}$		<p><math>\sigma_s</math> : Ecart – type du module de</p> <p><math>I_s = I_{s\alpha} + I_{s\beta}</math></p> <p><math>\zeta</math> : Dispersion des points de la caractéristique <math>Q'(P')</math></p> <p><math>m_p</math> : puissance active moyenne</p> <p><math>m_q</math> : puissance réactive moyenne</p>
Court-circuit entre spires au stator	<p>Harmonique de rang 3 et harmonique lié au passage des encoches rotoriques</p> <p>Impédance directe :</p> $Z_d = \frac{V_d(f=f_s)}{I_d(f=f_s)}$	<p><u>Flux axial :</u></p> <p><math>k \cdot f_s \pm n \cdot f_r</math></p>	

Tableau. 2. 8. Signatures déterminées à partir des signaux mesurés ou estimés (suite)

## 2.3. Vers des approches modernes de diagnostic

En dépit des diverses techniques mentionnées précédemment (diagnostic interne et externe), ces dernières années, la surveillance et la détection de défaut des machines électriques se sont éloignées des techniques traditionnelles pour s'orienter vers des techniques dites d'intelligence artificielle (IA) [Filippetti, 00], [Barros, 01], [Tallam, 01], [Arkan, 01], [Casimir, 03a], [Ondel, 05].

Ces techniques apportent aux systèmes de diagnostic des moteurs électriques la fiabilité, l'automatisation, la praticité et la sensibilité.

Ces méthodes ne sont pas en concurrence avec les méthodes précédemment citées. Elles exploitent les signatures avec ou sans modèle, décrites précédemment, pour réaliser la supervision et le diagnostic du système.

Elles possèdent, par ailleurs, de grandes facultés d'apprentissage. En effet, les récents développements dans le matériel et le logiciel permettent de produire un système pour la surveillance automatique des moteurs asynchrones en utilisant des techniques de traitement des signaux et des techniques de classification non supervisée pour le diagnostic des défauts.

Parmi ces méthodes, nous pouvons citer les systèmes experts, la logique floue, les réseaux neuronaux, la reconnaissance des formes,.... Dans cette partie, nous allons présenter quelques unes de ces techniques ainsi que leur application dans le domaine du diagnostic de la machine électrique. Elles sont présentées dans l'ordre chronologique de leur utilisation pour la surveillance des moteurs.

### 2.3.1. Les systèmes experts

Les systèmes experts sont des outils conçus pour pouvoir modéliser l'approche des experts dans un domaine spécifique. Ils permettent de formaliser leurs connaissances et fournissent un mécanisme pour les exploiter. Lorsque le mode de représentation de la connaissance est de type symbolique, l'approche par système expert est favorisée. Dans ce cadre, la *connaissance a priori* sur le système est représentée par un ensemble de règles et de faits (données manipulées par les règles) qui constituent ce qu'on appelle la *base de connaissance*. Cette base est construite à l'aide d'outils d'aide à la formalisation, ces outils sont fortement liés à l'application.



L'architecture générale d'un système expert se compose principalement de deux éléments : une base de connaissance et un moteur d'inférence [Farreny, 89].

Le moteur d'inférence représente l'organe de résolution, son fonctionnement est basé sur des méthodes de résolution indépendantes du contexte. Une méthode de résolution inclut en général les étapes suivantes :

1. la sélection de règles dans la base de connaissance en fonction des faits établis,
2. la résolution des conflits entre les règles sélectionnées,
3. l'exécution en indiquant les conditions de déclenchement et les conséquences jusqu'à ce que le but recherché soit atteint, par exemple le diagnostic.

En diagnostic, ce type d'approche est adapté aux problèmes nécessitant la manipulation d'une quantité importante de données non homogènes. Ils ont beaucoup été utilisés dans le diagnostic des systèmes statiques. Pour des systèmes dynamiques, ce type d'approche est peu adapté pour des problèmes de temps-réel et de gestion dynamique de l'information. Pour des applications industrielles, on pourra se référer à [Alty, 91]. On peut aussi faire référence à quelques articles traitant le problème du diagnostic des moteurs électriques à partir de ces systèmes [Leith, 88], [Filippetti, 88], [Ahrens, 93].

### **2.3.2. Les arbres de défaillances**

Il s'agit d'une méthode déductive qui vise à identifier pour chaque événement indésirable, l'ensemble des événements élémentaires qui peuvent contribuer à son apparition. Cette méthode est très employée dans les domaines de l'automobile, du nucléaire, de la chimie, de l'aéronautique, etc... pour évaluer la fiabilité d'un système [Vil, 90]. Elle consiste en une recherche exhaustive des combinaisons possibles d'événements qui conduisent à un défaut de fonctionnement ou une panne. La représentation la plus utilisée est celle qui associe à chaque événement indésirable la liste de tous les enchaînements logiques des événements jusqu'à arriver à des événements de base qui ne nécessitent pas de développements plus poussés.

Quand l'arbre est établi et les probabilités des événements élémentaires connues, la probabilité de l'événement indésirable et les combinaisons les plus probables conduisant à son apparition peuvent être déterminées.

Cette méthode présente différents avantages :

- Elle permet d'analyser toutes les causes possibles d'un défaut et donc d'améliorer la conception du système.
- La décomposition arborescente permet un diagnostic relativement rapide des défauts.

Cependant, cette méthode présente plusieurs inconvénients :

- Le principal vient de l'exhaustivité nécessaire pour établir le diagnostic : uniquement les causes prévues peuvent être déduites. La difficulté réside en la lourdeur de la mise en œuvre de la méthode pour un système complexe.
- Connaître jusqu'à quel niveau de décomposition l'analyse peut être menée.
- La méthode est difficile à utiliser pour des systèmes complexes fortement interconnectés. Une étude exhaustive de tous les défauts et de leurs combinaisons s'avère réhabilitaire pour un entraînement.

Dans le domaine des entraînements, des arbres de défaillances ont été utilisés pour analyser les pannes de cartes analogiques de commande. L'idée était d'analyser les causes de quelques combinaisons de signaux logiques issus de protections pour donner à l'utilisateur une indication sur la carte à changer après un défaut. L'expérience a montré que ces arbres étaient lourds à mettre en œuvre et difficiles à modifier. De plus, comme ils étaient basés sur un nombre restreint de signaux logiques, les indications données étaient très succinctes.

### **2.3.3. Les réseaux de neurones artificiels**

Un réseau de neurone est un modèle de calcul dont la conception est très schématiquement inspirée du fonctionnement de vrais neurones humains. Cette technique est placée dans la famille des méthodes de l'intelligence artificielle qu'ils enrichissent en permettant de prendre des décisions s'appuyant davantage sur la perception que sur le raisonnement logique formel. Dans les années 1940, les neurologues Warren Sturgis McCulloch et Walter Pitts menèrent les premiers travaux sur les réseaux de neurones. Ils constituèrent un modèle simplifié de neurone biologique communément appelé neurone formel. Ils montrèrent également théoriquement que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques complexes.

Issus de la recherche en neurophysiologie et en informatique, cette approche fait l'objet de recherches dans beaucoup de domaines tels que la compréhension et la synthétisation de la langue naturelle, la classification, le diagnostic, etc. Cette technologie comporte des caractéristiques très intéressantes comme la faculté d'apprentissage et d'auto-organisation. Elle présente des perspectives d'applications particulièrement intéressantes pour l'aide au diagnostic avec signatures externes. En effet, le calcul neuromimétique possède des propriétés

similaires à celles de la reconnaissance des formes pour la classification automatique de signatures.

Pour identifier des défauts dans un système, le diagnostic réalisé par réseaux de neurones doit disposer d'un nombre suffisant d'exemples de bon fonctionnement et de défauts pour pouvoir les apprendre. Pendant la phase d'apprentissage, les exemples sont présentés au réseau en entrée avec les diagnostics correspondants à la sortie. Le réseau s'auto-organise, apprenant à relier les exemples montrés aux diagnostics. Après l'apprentissage, le réseau ne reconnaît pas seulement les exemples appris mais également des paradigmes leur ressemblant, ce qui correspond à une certaine robustesse par rapport aux déformations de signaux par le bruit.

Pour un traitement par réseau de neurones, les exemples doivent être mis en forme pour être entrés dans le réseau. Ce sont des grandeurs caractéristiques des défauts à diagnostiquer, filtrées et prétraitées (transformée de Fourier).

Conceptuellement, un réseau de neurones peut être considéré comme une boîte noire, il n'est pas nécessaire de construire un modèle du système, contrairement aux méthodes dites internes (cf. § 2.2.1).

Cependant, l'inconvénient majeur est d'arriver à déterminer une méthodologie pour maîtriser les problèmes inhérents, qui sont principalement le choix de la structure, de la taille du réseau et des algorithmes d'apprentissage pour un problème précis.

Par contre, la principale raison de leur intérêt en diagnostic industriel est leur faculté d'apprentissage et la mémorisation d'un grand volume d'information.

D'ailleurs, en ce qui concerne la surveillance des machines électriques, cette méthode a été largement utilisée [Tallam, 01], [Filippetti, 93].

Par exemple, dans [Penman, 94] et [Schoen, 95c], le premier auteur utilise comme signal les vibrations du châssis du moteur alors que le second s'oriente vers l'emploi du courant statorique. Ces deux articles montrent l'efficacité des réseaux de neurones artificiels pour différents types de défauts. Penman expérimente la détection du déséquilibre de l'alimentation alors que Schoen s'intéresse à l'excentricité. Dans ces deux cas, leurs essais prouvent l'intérêt des réseaux de neurones pour la classification et la prise de décision.

On peut encore citer [Filippetti, 95], qui utilise avec succès un réseau de neurone artificiel pour le diagnostic en ligne des défauts rotor, en particulier la détection et le comptage du nombre de barres cassées.

### 2.3.4. La reconnaissance des formes (RdF)

Lorsque les observations sont de type numérique, l'approche par reconnaissance des formes peut être employée. Sur des problèmes complexes tels que le nucléaire, l'industrie automobile, le comportement humain, les réseaux téléphoniques, etc... où la modélisation du procédé est souvent difficile à mettre en œuvre, cette approche est privilégiée. De nombreux travaux [Pau, 74], [Peltier, 93], ont permis de montrer l'intérêt de ce type d'approche.

Dans le domaine du diagnostic des machines électriques, cette approche, quoique plus rarement utilisée, a été mise en œuvre avec succès par [Haji, 01], [Diallo, 05], [Casimir, 03b], par exemple.

Pour anticiper l'apparition d'un dysfonctionnement, soit pour les éviter, soit pour en limiter les conséquences, quelques études ont utilisées l'association de la RdF et la théorie des ensembles flous [Boutleux, 96], [Boudaoud, 97], [Frelicot, 92]. Il est parfois important de détecter l'évolution du système à diagnostiquer, notamment pour détecter la transition entre les modes de fonctionnement normal et anormal.

Dans le chapitre suivant, nous allons montrer comment la RdF peut servir de base à la conception d'un système de diagnostic, et l'avantage essentiel qu'elle présente par rapport aux autres approches. Cette méthode procure au système de diagnostic des moyens d'appréhender la connaissance a priori incomplète des modes de fonctionnement du processus et ainsi, va permettre la mise en place d'une méthode de diagnostic évolutive.

## 2.4. Conclusion

L'objectif de ce chapitre était de présenter un état de l'art des différentes méthodes de diagnostic ainsi que les différentes signatures de défaillances extraites à partir de ces méthodes.

Dans un premier temps nous nous sommes intéressés aux méthodes de diagnostic dites internes (*avec connaissance a priori*). Si le comportement électrique et dynamique de la machine est parfaitement connu, les modèles utilisés peuvent fournir une estimation fiable des grandeurs difficilement mesurables. Les signaux et paramètres de sortie sont alors utilisés pour la surveillance et le diagnostic.

Par la suite, nous avons présenté les méthodes dites externes (*sans connaissance a priori*) basées sur la mesure de signaux vibratoires, électriques, électromagnétiques rendant directement compte de l'état de la machine.

Ces deux types de méthodes donnent un large éventail de signatures plus ou moins pertinentes vis-à-vis des différents défauts pouvant se produire sur un moteur électrique. Ces méthodes ne doivent pas être mises en concurrence, mais au contraire peuvent être utilisées de manière complémentaire pour tirer partie de leurs avantages respectifs.

La liste des signatures des défauts est sans aucun doute non exhaustive, mais ces résultats peuvent servir de point de départ pour la recherche de la signature nécessaire à la reconnaissance des formes. En effet, pour être appliquée au diagnostic, la reconnaissance des formes nécessite la connaissance de l'état de fonctionnement du système étudié par le biais d'un vecteur forme.

## CHAPITRE III

Approche diagnostic par Reconnaissance des Formes (RdF)

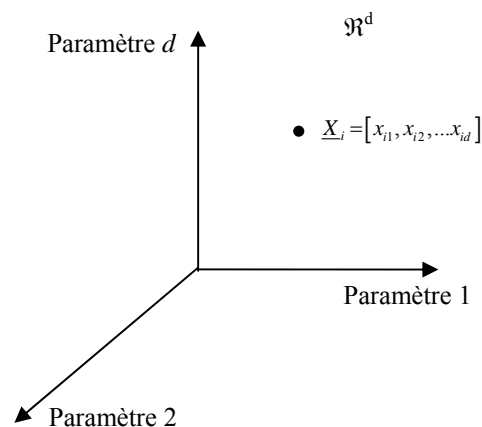


### 3.1. Introduction

La Reconnaissance des Formes (RdF) repose sur le classement des objets ou formes en les comparant à des formes-types. De manière générale, deux types de RdF se distinguent :

- la RdF structurelle qui se base sur une représentation des formes à l'aide de grammaires [Fu, 74]. Cette approche ne sera pas développée par la suite,
- la RdF statistique qui s'appuie sur une représentation numérique des formes [Dubuisson, 90].

Une forme est une observation réalisée sur le processus. Elle est caractérisée par un ensemble de  $d$  paramètres (ou caractères), et représentée par un point dans l'espace de dimension  $d$ , défini par les différents paramètres (espace de représentation). Comme les paramètres sont souvent des nombres réels, une forme ' $i$ ' peut être définie par un vecteur  $\underline{X}_i = [x_{i1}, x_{i2}, \dots, x_{id}]$  de  $\mathfrak{R}^d$  (figure 3.1).



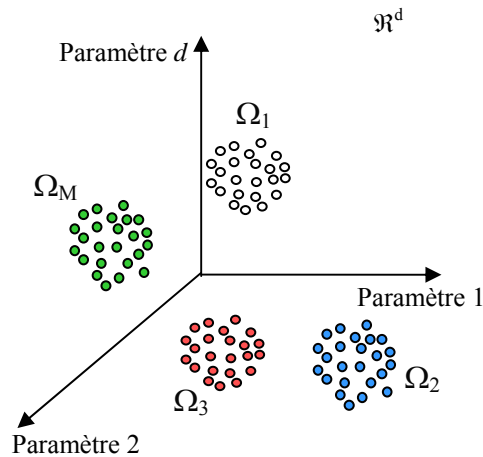
**Fig. 3. 1.** Représentation d'un objet dans l'espace de dimension  $d$

Ces formes ou objets sont donc caractérisés par ce vecteur appelé *vecteur forme*. Si l'on se place dans le contexte du diagnostic, les paramètres du vecteur forme traduisent l'état du système étudié. Ils sont issus d'analyses effectuées sur les signaux mesurés par les capteurs implantés sur le système.

Les formes types (ou prototypes) sont des points représentatifs de cet espace, et le problème de la reconnaissance consiste à associer une forme observée à une forme type connue. A cause des perturbations (bruit de mesure, précision des capteurs, ...), une nouvelle observation



sera rarement identique à l'un des prototypes. Ainsi, afin de traduire l'influence du bruit, les classes ( $\Omega_1, \Omega_2, \dots, \Omega_c, \dots, \Omega_M$ ) correspondent à des zones dans l'espace, regroupant les formes semblables (figure 3.2).



**Fig. 3. 2.** Notion de classes en Reconnaissance des Formes

Le principe de la reconnaissance est de savoir décider à quelle classe, parmi M classes connues, associer une nouvelle forme  $\underline{X}_u = [x_{u1}, x_{u2}, \dots, x_{ud}]$  (avec  $d$ , la dimension de l'espace de représentation) observée.

En terme de diagnostic, les classes correspondent aux modes de fonctionnement connus. Elles constituent notre ensemble de données initiales, appelé *ensemble d'apprentissage* et noté  $X_a$ . Le fait de classer une nouvelle observation, revient à identifier un de ces modes.

La conception d'un système de diagnostic par RdF se déroule en 2 phases : *une phase d'analyse* et *une phase d'exploitation*.

La *phase d'analyse* consiste, à partir de connaissances acquises sur le système, à :

- **Déterminer l'espace de représentation**, c'est-à-dire choisir les paramètres sensibles aux différents défauts permettant la meilleure discrimination des classes et ainsi construire le vecteur forme.
- **Réduire l'espace de représentation**, c'est-à-dire déterminer les  $d'$  ( $d' < d$ ) paramètres les plus pertinents pour la reconnaissance, afin d'améliorer les performances et de diminuer le temps de décision relatif au classement d'une nouvelle observation. Cette opération de réduction peut être réalisée soit par des méthodes d'extraction, soit par

des méthodes de sélection de caractères. L'extraction consiste à effectuer une transformation de l'espace initial par combinaison des caractères initiaux. La sélection utilise une procédure de recherche selon un critère de séparabilité entre les classes et de compacité dans les classes permettant ainsi de sélectionner un sous-ensemble des caractères initiaux.

- **Déterminer l'espace de décision**, c'est-à-dire l'ensemble de toutes les classes possibles. Si l'on dispose de N observations déjà recueillies sur le système (ensemble d'apprentissage) et pour lesquelles on connaît a priori la classe d'origine, c'est-à-dire son mode de fonctionnement correspondant, l'espace de décision est parfaitement déterminé, on parle de RdF en *mode supervisé*. Dans le cas contraire, on parle de RdF en *mode non supervisé*, il convient d'extraire cette connaissance manquante de l'ensemble d'apprentissage. Le principe est de détecter dans cet ensemble, une ou des structures de classes disjointes au sens d'un critère de similarité, de telle façon que les observations d'une même classe soient semblables et que celle de classes différentes soient dissemblables. De nombreuses méthodes existent pour mener à bien cette opération de classification automatique appelée 'coalescence' [Dubes, 79], [Jain, 88].
- **Choisir une méthode de discrimination** qui consiste à construire des frontières entre les différentes classes de l'ensemble d'apprentissage afin d'élaborer une règle de décision. Cette règle décidera d'affecter ou non une nouvelle observation à l'une des classes connues.

A la fin de la *phase d'analyse*, le système de diagnostic par reconnaissance peut être exploité. Il est alors possible d'associer une nouvelle observation  $\underline{X}_u$  recueillie sur le système à l'une des classes précédemment définies en appliquant la règle de décision. C'est la *phase d'exploitation*. La bonne exploitation du système décisionnel dépend de la pertinence du vecteur forme et des performances de la règle de décision.

Le diagnostic par reconnaissance des formes a montré son intérêt sur différentes applications [Zieba, 95], [Boutleux, 96], [Ouaghram, 99], [Casimir, 03]. Dans les parties suivantes, les différentes étapes nécessaires à la conception du diagnostic par RdF vont être détaillées en s'intéressant particulièrement à la sélection des paramètres (partie 2), et à la coalescence en mode *non supervisé* (partie 3). Dans la partie 4, une approche basée sur les algorithmes génétiques (AGs) est développée, permettant de réaliser simultanément, la

recherche des paramètres les plus discriminants ainsi que la classification automatique (coalescence). Une fois l'ensemble d'apprentissage structuré en classes, il faut choisir une règle qui réalise l'affectation d'une nouvelle forme. Dans la partie 5, une règle de décision bien connue, les  $k$  plus proches voisins, est présentée. Cette règle va être associée à une fonction d'appartenance permettant de réaliser une classification non exclusive. Cette approche va permettre de définir le degré d'appartenance d'une nouvelle forme à sa classe et de ce fait, fournir un indice de fiabilité à la classification réalisée.

L'application de la RdF au diagnostic se heurte souvent au problème de la connaissance *a priori* du système étudié. En effet, il est rare de disposer, dans l'historique du système, d'observations correspondant à tous les modes de fonctionnement possibles. En général, tout ou partie des modes de fonctionnement qualifiés d'anormaux n'ont jamais été rencontrés sur le système. Pour déterminer l'estimation de ces 'états manquants', deux approches ont été développées dans la partie 6. La première, plus classique, basée sur une interpolation linéaire à partir d'une fonction polynomiale et la deuxième, plus originale, basée sur l'utilisation du filtre de Kalman. Ces approches vont permettre de réaliser le suivi d'évolution des différents modes de fonctionnement, ce qui peut présenter l'avantage de réduire le nombre d'acquisitions nécessaires au départ pour construire la base de données. De plus, l'estimation par filtre de Kalman pourra être utilisée pour la prédiction de ces modes de fonctionnement vers des zones de l'espace (de représentation) encore inconnues. Ce suivi d'évolution peut permettre d'observer les dérives éventuelles dans le temps du fonctionnement du système étudié et, associé à la prédiction, il est alors possible d'anticiper vers quel état se dirige le processus.

En phase de décision, l'association d'une nouvelle observation à un mode répertorié, alors que celle-ci représente un mode de fonctionnement inconnu jusqu'ici, mène à un diagnostic erroné. Cette critique est valable aussi pour certaines autres approches utilisées en diagnostic (systèmes experts, méthodes à base de modèles). Il apparaît donc nécessaire de concevoir des systèmes de diagnostic adaptatifs, dans lesquels la connaissance incomplète peut être enrichie, c'est-à-dire capables d'incorporer à leur connaissance *a priori* initiale, de nouveaux modes de fonctionnement au fur et à mesure qu'ils apparaissent. C'est ce que nous présenterons dans la partie 7.

## 3.2. Détermination de l'espace de représentation

### 3.2.1. Constitution du vecteur forme

Cette étape consiste à construire le vecteur forme, c'est-à-dire, sélectionner les  $d$  paramètres à extraire à partir des signaux recueillis à l'aide des capteurs. Ces paramètres sont choisis pour optimiser la discrimination des modes de fonctionnement. Il n'existe pas d'outils algorithmiques qui permettent de régler le choix de ce vecteur, seule la connaissance *a priori* que l'on a du système, permet de retenir les paramètres les plus pertinents. Ainsi l'intervention d'un expert du processus à surveiller est souvent très utile pour orienter cette procédure.

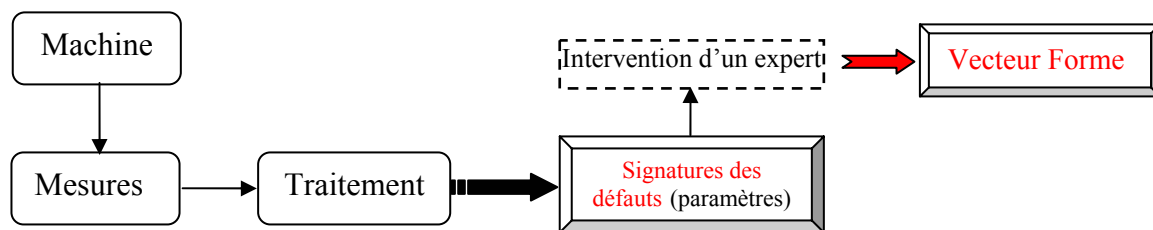


Fig. 3.3. Cheminement pour le calcul des paramètres constituant le vecteur forme initial

Les performances du système de diagnostic dépendront de la pertinence des paramètres calculés. Il est donc préférable de disposer de paramètres variant, de manière significative, en fonction des différents modes de fonctionnement du système (chapitre II).

Le nombre souvent élevé de paramètres peut être pénalisant en terme de temps de calcul. De plus, tous les paramètres calculés et constituant le vecteur forme initial ne seront pas forcément pertinents vis-à-vis des modes de fonctionnement étudiés. Des méthodes doivent donc être utilisées afin de ne conserver que les paramètres les plus représentatifs.

### 3.2.2. Réduction de l'espace de représentation

La réduction de la dimension du vecteur forme consiste à rechercher un sous-ensemble de  $d'$  paramètres ( $d' < d$ ), qui conserve le mieux possible la séparation des classes de l'ensemble d'apprentissage initial. Cette réduction de l'espace de représentation peut être réalisée soit par des méthodes d'extraction de paramètres soit par des méthodes de sélection de paramètres.

L'extraction consiste à définir de nouveaux paramètres à partir des  $d$  paramètres initiaux. Du point de vue diagnostic, les méthodes d'extraction ne diminuent pas le nombre de paramètres à calculer, qui restent au nombre de  $d$  (les nouveaux paramètres sont des combinaisons linéaires des anciens) mais en revanche, l'espace de représentation des classes

est de dimension plus faible ce qui a pour conséquence d'accélérer la phase de décision. Parmi les différentes méthodes d'extraction de paramètres, on distingue l'Analyse en Composantes Principales (ACP) ou transformée de Karhunen - Loève discrète.

Les méthodes de sélection de paramètres ont pour but de chercher  $d'$  paramètres parmi les  $d$  initiaux réalisant la meilleure discrimination possible des classes de  $X_a$ . Ainsi le nombre de mesures à effectuer diminue comme la dimension de l'espace de représentation, ce qui présente un double avantage : une réduction des temps de calcul et l'élimination des informations redondantes ou inintéressantes. Les paramètres sélectionnés seront ceux donnant la combinaison optimale au sens d'un critère de séparabilité entre classes et de compacité des classes comme par exemple le critère de Fisher [Fisher, 36].

Pour ces raisons, ces méthodes de sélection seront utilisées par la suite, pour réaliser la réduction du vecteur forme tandis que les méthodes d'extraction seront réservées aux problèmes de visualisation de l'ensemble d'apprentissage.

### 3.2.2.1. Extraction de paramètres par Analyse en Composantes Principales (ACP)

Soit  $(\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  l'ensemble d'apprentissage ( $X_a$ ) composé de  $N$  vecteurs de dimension  $d$ . L'Analyse en Composantes Principales est une méthode permettant d'identifier un sous espace  $\mathcal{R}^{d'}$  de dimension  $d' < d$  représentant au mieux l'ensemble  $(\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$ . On notera  $u_j$ , la base orthogonale de dimension  $d'$  représentant cet espace. Les vecteurs  $u_j$  sont  <sub>$j=1, d'$</sub>  appelés axes factoriels ou axes principaux.

Pour déterminer la nouvelle base, deux approches, équivalentes, mais exprimées différemment sont proposées pour traduire le critère de 'représentation optimale' de l'ensemble  $(\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  :

- Minimisation de la déformation du nuage de points projeté dans la nouvelle base, qui conduit à minimiser [Dubuisson, 90] :

$$J = \frac{1}{N} \sum_{i=1}^N d^2(\underline{X}_i, \underline{X}_i^p) \quad (3.1)$$

Avec :

$d(.,.)$  est une mesure de distance,

$N$  le nombre de vecteurs dans l'ensemble d'apprentissage,

$\underline{X}_i^p$  la projection du vecteur  $\underline{X}_i$  dans la base formée par les  $u_j$ .

- Maximisation de la variance des points projetés sur chaque axe factoriel [Poublan, 93] :

$$J' = u_j^t \left[ \sum_{i=1}^N (\underline{X}_i - \underline{m})(\underline{X}_i - \underline{m})^t \right] u_j = u_j^t \Sigma u_j \quad (3.2)$$

avec  $\Sigma$ , la matrice de variance-covariance,

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (\underline{X}_i \cdot \underline{X}_i^t) - \underline{m} \cdot \underline{m}^t \quad (3.3)$$

et  $\underline{m}$  centre de gravité de l'ensemble d'apprentissage

$$\underline{m} = \frac{1}{N} \sum_{i=1}^N \underline{X}_i \quad (3.4)$$

Dans les deux cas, les vecteurs  $u_j$  sont les vecteurs propres de la matrice de variance-covariance, classés selon l'ordre des valeurs propres décroissantes.

Si les variables ne sont pas homogènes, c'est-à-dire si l'une d'entre elles présente des valeurs beaucoup plus grandes que les autres, alors l'influence de cette variable sera prépondérante dans le calcul de la matrice  $\Sigma$  et donc sur les résultats de l'ACP. Il est alors nécessaire de normaliser les variables par leurs variances en introduisant la matrice diagonale des variances  $\sigma_j^2$  :

$$\Delta = \begin{bmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \frac{1}{\sigma_2^2} & \\ 0 & & \frac{1}{\sigma_d^2} \end{bmatrix} \quad \text{où } \sigma_j^2 = \frac{1}{N} \sum_{i=1}^N (x_{ij} - m_j)^2 \quad (3.5)$$

avec  $j = 1$  à  $d$ ,

$x_{ij}$ , la  $j^{\text{ième}}$  composante de  $\underline{X}_i$ ,

$m_j$ , la moyenne de la composante  $j$

Ainsi, à partir des relations (3.4) et (3.5), on en déduit les vecteurs formes centrés-réduits :

$$\underline{X}_i \text{ norm} = \Delta^{-1/2} (\underline{X}_i - \underline{m}) \quad (3.6)$$

avec  $\underline{X}_i$  : le  $i^{\text{ième}}$  vecteur de l'ensemble d'apprentissage

Pour la suite de l'étude, les données seront considérées comme étant centrées-réduites (c'est-à-dire de moyenne nulle et d'écart type unité).

Les résultats de l'ACP peuvent être visualisés en considérant la projection du nuage de point sur les deux ou trois premiers axes factoriels. Autrement dit, en visualisant la matrice des données dans l'espace défini par les deux ou trois premiers vecteurs propres.

Enfin la qualité de la représentation peut être quantifiée par le calcul du pourcentage d'inertie expliqué sur chaque axe factoriel ou sur le sous-espace de dimension  $d'$  obtenu. Ce pourcentage d'inertie est défini en fonction des valeurs propres.

Soit  $(\lambda_1, \lambda_2, \dots, \lambda_{d'})$  les  $d'$  valeurs propres obtenues classées par ordre décroissant, le pourcentage d'inertie expliqué par un axe  $l$  est donné par :

$$\Gamma_l = \frac{\lambda_l}{\sum_{j=1}^d \lambda_j} \quad (3.7)$$

et la qualité de la représentation du nuage de points s'exprime :

$$\Gamma = \frac{\sum_{k=1}^{d'} \lambda_k}{\sum_{j=1}^d \lambda_j} \quad (3.8)$$

Plus ce dernier rapport est proche de 1, meilleure est la représentation. En pratique, l'utilisateur fixe le seuil de qualité (80%, 90%), puis sélectionne séquentiellement les vecteurs  $u_j$  lui permettant d'atteindre ce seuil afin de définir la base du sous-espace.

Des résultats d'ACP, utilisés pour la visualisation des ensembles d'apprentissage peuvent être consultés dans le chapitre IV.

### 3.2.2.2. Méthodes de sélection de paramètres en mode supervisé

Il a été montré que l'ACP permet de construire un sous-espace à partir de l'espace initial pour visualiser et décrire l'ensemble d'apprentissage. Mais cette méthode ne diminue pas le nombre de paramètres à calculer pour chaque observation, du fait que les composantes des vecteurs projetés sont des combinaisons linéaires des composantes des vecteurs initiaux.

Pour réduire le nombre de paramètres et, de ce fait, la dimension de l'espace de représentation, il est possible d'utiliser des méthodes de sélection de paramètres.

Ces méthodes supposent que la structure en classes de l'ensemble d'apprentissage  $(\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  est connue, c'est-à-dire que l'on se situe en *mode supervisé*. Ainsi, l'objectif est de sélectionner les  $d'$  paramètres parmi les  $d$  initiaux, conservant au mieux l'information sur la structure des classes ou encore construire un sous-espace  $E'$  issu de l'espace initial  $E$ .

L'avantage de ces méthodes est de réduire le temps de calcul, ce qui peut être essentiel dans une application de diagnostic en temps réel.

Une méthode de sélection de paramètres repose sur la détermination de plusieurs éléments :

- Le critère de qualité pour quantifier la qualité d'un paramètre,
- La technique de recherche utilisée,
- Le choix de la dimension finale (valeur de  $d'$ )

### 3.2.2.2.1. Critère de qualité

Le critère utilisé pour la détermination du sous-espace  $E'$ , doit représenter au mieux les classes supposées connues de  $X_a$ , ce qui correspond intuitivement à :

- un regroupement des points d'une même classe, c'est-à-dire à une minimisation de la variance intra-classe (notion de compacité),
- une séparation des différentes classes, soit une maximisation de la variance inter-classe (notion de séparabilité).

Les matrices de variance intra-classe et inter-classe traduisent ces deux conditions. La matrice de variance intra-classe fait intervenir les échantillons de chaque classe ainsi que les centres de gravités :

$$\Sigma_w = \frac{1}{N} \sum_{c=1}^M \sum_{v=1}^{N_c} (\underline{X}_{cv} - \underline{m}_c) \cdot (\underline{X}_{cv} - \underline{m}_c)^t \quad (3.9)$$

Alors que la matrice de dispersion inter-classes caractérise la dispersion des classes entre elles.

$$\Sigma_B = \frac{1}{N} \sum_{c=1}^M (\underline{m}_c - \underline{m}) \cdot (\underline{m}_c - \underline{m})^t \quad (3.10)$$

avec :

$\underline{m}$  : centre de gravité général (toutes classes confondues)

$M$  : nombre de classes

$\underline{m}_c$  : centre de gravité de la classe  $\Omega_c$

$\underline{X}_{cv}$  :  $v^{\text{ième}}$  vecteur de la classe  $\Omega_c$

$N_c$  : nombre de vecteurs de la classe  $\Omega_c$

$N$  : nombre de vecteurs total de l'ensemble d'apprentissage

On peut alors définir la matrice de variance-covariance  $\Sigma$  (3.3) par :

$$\Sigma = \Sigma_w + \Sigma_B \quad (3.11)$$

Cette relation traduit le fait que la dispersion totale est la somme de la dispersion à l'intérieur des classes et de la dispersion des classes entre elles.



L'objectif étant de définir un nouvel espace maximisant la dispersion inter-classe en maintenant identique la dispersion intra-classe, il est possible d'utiliser les vecteurs propres de  $\Sigma_w^{-1} \cdot \Sigma_B$  dans ce but. Ces vecteurs propres sont invariants par une transformation linéaire non singulière. Ceci suppose (N-M) et M supérieurs ou égaux à  $d$ . L'idée est donc d'utiliser les valeurs propres de  $\Sigma_w^{-1} \cdot \Sigma_B$  comme indicatrices du rapport de la dispersion entre classes et de la dispersion intra-classe. Ainsi, il est possible de définir deux critères :

$$J_1 = \text{trace}(\Sigma_w^{-1} \cdot \Sigma_B) \quad \text{et} \quad J_2 = \frac{\det(\Sigma)}{\det(\Sigma_w)} \quad (3.12)$$

trace(X) est la trace de matrice X

Le rapport de Fisher peut également être appliqué comme critère, pour la sélection de paramètres. L'idée est de considérer successivement chacun des paramètres contenus dans le vecteur forme initial de dimension  $d$  et de calculer le critère  $J_3$  pour ce paramètre.

Dans le cas d'un problème à deux classes, le critère s'exprime, pour un paramètre  $\alpha$  :

$$J_3(\alpha) = \frac{m_1(\alpha) - m_2(\alpha)}{N_1\sigma_1^2(\alpha) - N_2\sigma_2^2(\alpha)} \quad (3.13)$$

avec :

- $m_c(\alpha)$  ( $c = 1, 2$ ), centre de gravité de la classe  $\Omega_c$  en ne considérant que le paramètre  $\alpha$

$$m_c(\alpha) = \frac{1}{N_c} \sum_{v=1}^{N_c} \underline{X}_{cv}(\alpha) \quad (3.14)$$

avec  $\underline{X}_{cv}$  :  $v^{\text{ième}}$  vecteur de la classe  $\Omega_c$

$N_c$  : nombre de vecteurs de la classe  $\Omega_c$

- $\sigma_c^2(\alpha)$  : variance de la composante  $\alpha$  des vecteurs de la classe  $\Omega_c$  :

$$\sigma_c^2(\alpha) = \frac{1}{N_c} \sum_{v=1}^{N_c} [\underline{X}_{cv}(\alpha) - m_c(\alpha)]^2 \quad (3.15)$$

Dans le cas général d'un problème à M classes,  $J_3(\alpha)$  s'exprime :

$$J_3(\alpha) = \sum_{c=1}^M \sum_{r=1}^{M-1} \frac{m_c(\alpha) - m_r(\alpha)}{N_c\sigma_c^2(\alpha) - N_r\sigma_r^2(\alpha)} \quad (3.16)$$

Le numérateur de l'expression 3.16 traduit la séparation des classes (variance inter – classes) alors que le dénominateur reflète la compacité des classes (par analogie avec la variance intra – classe). Plus le critère est grand, plus les classes sont séparées.

Pour la suite de l'étude, l'expression  $J_1$  (3.12) a été utilisée comme critère dans les

méthodes de sélection. Il sera noté J.

### 3.2.2.2.2. Techniques de recherche utilisées

Soit E l'espace composé de  $d$  paramètres initiaux, il est possible de rechercher un sous-ensemble E' de  $d'$  paramètres tel que :

$$J(E') = \text{Max}_{b=1, n_c} (J(E'_b)) \quad (3.17)$$

avec  $n_c$ , le nombre de combinaisons (de  $d'$  paramètres parmi  $d$ ) étudiées.

La première solution possible est de calculer le critère J sur toutes les combinaisons de  $d'$  paramètres possibles. C'est une technique exhaustive très lourde puisque le nombre de solutions à étudier est donné par :

$$C_d^{d'} = \frac{d!}{(d-d')!d!} \quad (3.18)$$

Cette solution est vite dissuasive si la dimension de départ est élevée.

Il existe aussi des méthodes de sélection optimales telles que les algorithmes basés sur une procédure d'optimisation de type 'Branch and Bound'. Ces méthodes permettent de prendre en compte implicitement toutes les combinaisons de paramètres, sans devoir effectivement calculer le critère pour chaque combinaison. Pour cela, il faut faire l'hypothèse de la monotonie du critère J, c'est-à-dire une diminution de J lorsque la dimension de l'espace décroît [Somol, 04], [Chen, 03].

Afin d'éviter une recherche exhaustive pénalisante en temps de calcul, une solution consiste à utiliser des méthodes dites sous-optimales, qui présentent l'avantage d'être rapide et de fournir des résultats réalistes. On peut citer entre autres :

- la méthode S.F.S. (Sequential Forward Selection) qui réalise une sélection de paramètres de manière successive. L'ensemble de départ est initialisé comme un ensemble vide. Soit  $E'_k$  l'ensemble des paramètres déjà sélectionnés à l'instant k. La méthode S.F.S. consiste à sélectionner parmi les  $d-k$  paramètres restants, celui qui, associé à l'ensemble  $E'_k$ , optimise le critère J.
- la méthode G.S.F.S. (Generalised Sequential Forward Selection), dont la procédure est identique à celle décrite précédemment, mais cette fois-ci, plusieurs paramètres sont sélectionnés lors d'une itération.

- La méthode S.B.S. (Sequential Backward Selection), qui comme son nom l'indique procède à l'inverse de la méthode S.F.S. Au départ, l'ensemble des paramètres sélectionnés est initialisé comme l'ensemble des paramètres. A chaque étape, on élimine le paramètre pénalisant le plus le critère  $J$ .
- La méthode G.S.B.S. (Generalised Sequential Backward Selection) est la méthode symétrique de la méthode G.S.F.S. Elle utilise la même procédure que S.B.S. mais en éliminant plus d'un paramètre à chaque itération.

Dans [Kudo, 00], les auteurs présentent les avantages et les inconvénients des algorithmes de sélection. En accord avec les auteurs, nous utiliserons, dans le cadre de nos travaux, la méthode S.B.S. pour sélectionner les  $d'$  paramètres portant le plus d'informations. En effet, bien que sous optimale, cette méthode présente l'avantage d'être rapide, facile à mettre en œuvre et performante en terme de classification.

D'une manière plus détaillée, la procédure de la méthode S.B.S. est la suivante.

A l'étape  $k$ , on calcule le critère pour les sous ensembles  $E_{k,j}$  de dimension  $d-k$ . Les sous-espaces  $E_{k,j}$  sont issus du sous-espace  $E_{k-1}$  en éliminant volontairement le paramètre  $\alpha_j$  :

$$E_{k-1} \supset E_{k,j} = \{E_{k-1} - \alpha_j\} \quad (3.19)$$

Ainsi le sous-espace  $E_k$  pris en considération pour l'étape suivante est celui pour lequel le critère est optimal :

$$J(E_k) = \underset{j=1, d-k}{Max} (J(E_{k,j})) \quad (3.20)$$

### 3.2.2.2.3. Choix de la dimension finale (valeur de $d'$ )

Comme on pourra le voir par la suite, la classification d'une nouvelle observation, c'est-à-dire le diagnostic de l'état du système, dépend fortement de la sélection des paramètres pertinents mais aussi dépend fortement de la dimension de l'espace de représentation. Ainsi, la détermination de la dimension de l'espace est aussi un point clef, car sa valeur influence le résultat de la classification finale.

Une première solution consiste à regarder le résultat de la classification obtenu pour différentes valeurs de  $d'$  et de choisir la dimension pour laquelle le diagnostic est le meilleur. Mais cette solution nécessite l'intervention de l'utilisateur.

L'idéal est de déterminer automatiquement et simultanément les meilleurs paramètres ainsi que la meilleure dimension. Pour cela, il suffit de suivre l'évolution du critère  $J$  utilisé en fonction des différentes valeurs de  $d'$ .

La figure 3.4 montre un exemple de l'évolution de la valeur du critère choisi. Cette évolution permet de déterminer la dimension de l'espace. En effet, à partir d'une certaine dimension, l'addition de paramètres n'apporte aucune information supplémentaire. Sur la figure 3.4, 'b' apparaît comme la meilleure dimension car, à ce niveau là, la pente de l'évolution du critère observe une certaine stationnarité.

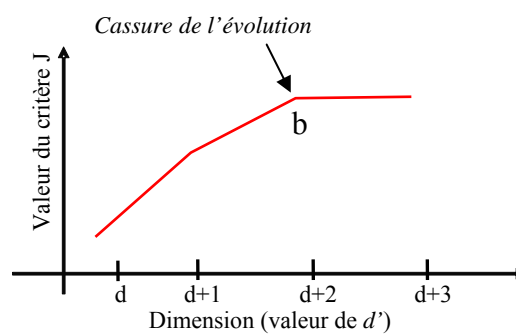


Fig. 3. 4. Choix de la dimension de l'espace de représentation

### 3.2.2.3. Utilisation des algorithmes génétiques pour la sélection de paramètres

#### 3.2.2.3.1. Introduction

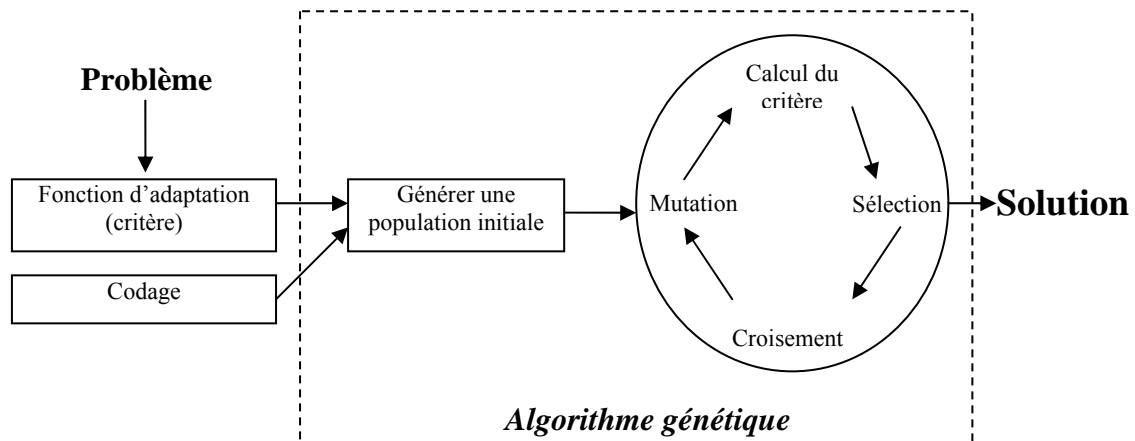
Les méthodes de sélection de paramètres présentées précédemment sont des méthodes non optimales car elles ne recherchent pas toutes les solutions possibles du problème posé. Ce type de méthode converge souvent vers un optimum local, pour des problèmes mal définis ou multimodaux.

En 1975, [Hooland, 75] a introduit de nouvelles techniques d'optimisation inspirées de l'évolution naturelle des espèces et de la génétique, connues sous le nom d'algorithmes génétiques (AGs). Depuis, l'utilisation de ces techniques n'a cessé de croître en raison de leur robustesse et surtout du fait qu'elles ne nécessitent aucune connaissance analytique du système à optimiser.

#### 3.2.2.3.2. Principe d'un algorithme génétique

Les AGs manipulent une *population* d'*individus*. Chacun d'entre eux représente le codage d'une solution du problème étudié, sous la forme d'une chaîne qui contient les paramètres du

problème, codés soit en binaire, réel, etc... Au début de l'algorithme, la *population* est générée aléatoirement, après quoi une fonction d'adaptation (définition d'un critère) est affectée à chaque *individu*. La *population* évolue de génération en génération sous l'effet de trois opérateurs de base : la sélection, le croisement et la mutation (figure 3.5) de manière à converger vers une valeur optimale du critère.



**Fig. 3.5.** Résolution d'un problème par algorithme génétique

### 3.2.2.3.3. Les éléments constituant un algorithme génétique

Dans la suite de cette partie, les différents constituants ayant servi à la réalisation d'une méthode de sélection de paramètres par AG sont présentés. Le principe général d'un AG peut être résumé de la façon suivante :

Algorithme génétique
Choix du codage Génération aléatoire de la population initiale Evaluation : calcul de la fonction d'adaptation (critère) <b>Répéter</b>  Sélection Croisement Mutation Evaluation  <b>Jusqu'à</b> satisfaction du critère d'arrêt

#### 3.2.2.3.3.1. Codage de la population

Les AGs utilisent une représentation, sous forme de chaînes, des paramètres de la fonction d'adaptation à optimiser. Il existe différents types de codage [Davis, 91] tels que le codage binaire, qui a pour avantage de coder facilement toutes sortes d'objets : des réels, des entiers,

des valeurs booléennes, des chaînes de caractères, etc..., mais aussi le codage réel où l'*individu* appelé aussi *chromosome* est représenté sous forme de vecteur de nombres réels.

Le codage binaire a été utilisé pour résoudre le problème de sélection.

Le *chromosome* représente le codage d'une solution potentielle au problème à résoudre, et est constitué d'un ensemble d'éléments appelés *gènes* pouvant prendre les valeurs 0 ou 1.

La *population* initiale est composée de  $P$  *individus*. L'initialisation de ces *individus* est faite de façon aléatoire : les valeurs des *gènes* (0 ou 1) sont choisies aléatoirement selon une distribution uniforme. La taille de la *population* (choix de  $P$ ) doit être choisie afin de réaliser un bon compromis entre le temps de calcul et la qualité du résultat.

Chaque *individu* constituant la *population* initiale est caractérisé par un vecteur  $V_k$  ( $k=1,P$ ) composé de  $d$  *gènes* représentant les  $d$  paramètres du vecteur forme.  $V_k$  est donc de la forme :

$$V_k = (g_{k1}, g_{k2}, \dots, g_{kj}, \dots, g_{kd})$$

$$\text{avec } g_{kj} = \begin{cases} 1 & \text{si } \alpha_j \text{ est présent} \\ 0 & \text{sinon} \end{cases} \quad \text{avec } P, \text{ nombre d'individus} \quad (3.21)$$

Chaque *gène*  $g_{kj}$  code la présence ou l'absence du paramètre  $\alpha_j$  dans le vecteur  $V_k$ . Par exemple, pour un vecteur forme composé de 10 paramètres, un vecteur  $V_k$  sera représenté ainsi :

$$V_k = (0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1) \quad (3.22)$$

De cette manière, seuls les paramètres 2, 3, 8 et 10 seront considérés pour calculer la fonction d'adaptation.

Le vecteur optimal donnera simultanément les meilleurs paramètres et la dimension  $d'$  de l'espace de représentation.

### 3.2.2.3.3.2. Fonction d'adaptation proposée

La fonction du problème étudié est la seule information dont se sert l'AG afin d'évaluer 'l'adaptation' de chaque *individu* à l'environnement, autrement dit sa performance. Le problème étant de choisir les  $d'$  paramètres parmi les  $d$  initiaux permettant d'obtenir la meilleure compacité et séparabilité des classes, le critère utilisé par les méthodes de sélection précédemment présentées et défini par l'expression 3.12 semble être une bonne solution.

Ainsi, a chaque vecteur  $V_k$  sera associé une fonction d'adaptation définie par :

$$F(V_k) = \text{trace}(\sum_{wk}^{-1} \sum_{Bk}) \quad (3.23)$$

Le vecteur  $V_k$  est la solution optimale au problème à résoudre si  $F(V_k)$  est maximale :

$$\text{Si } F(V_k) = \max_{\substack{r \neq k \\ r=1,P}} F(V_r) \quad \text{alors} \quad V_k \text{ est la solution optimale}$$

### 3.2.2.3.3.3. Sélection

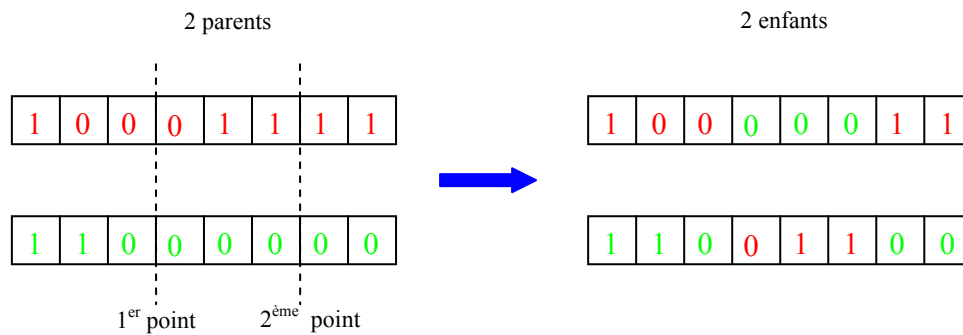
Cette première étape dans le fonctionnement d'un AG décide de la vie ou de la mort des *individus*, en fonction de leur adaptation. Plusieurs catégories d'opérateurs de sélection existent dans la littérature [Michalewicz, 94]. La technique présentée est celle utilisée dans le cas du choix des meilleurs paramètres. Il s'agit de la technique de sélection par tournoi, qui procède comme suit :

- deux *individus* sont choisis au hasard parmi les  $P$  *individus* de la *population* initiale et combattent (comparaison de leurs fonctions d'adaptation). Le plus adapté l'emporte. Cette étape est répétée jusqu'à ce que  $P/2$  *individus* soient sélectionnés. Il est tout à fait possible que certains *individus* participent à plusieurs tournois : s'ils gagnent plusieurs fois, ils auront donc droit d'être copiés plusieurs fois, ce qui favorisera la pérennité de leurs gènes.

Il se peut, de par la technique de sélection utilisée, que le meilleur *individu* ne soit pas sélectionné. Dans ce cas, cet *individu* est copié d'office à la place d'un *individu* choisi aléatoirement parmi les  $P/2$  *individus* ayant remportés le tournoi.

### 3.2.2.3.3.4. Croisement

Le rôle de l'opérateur de croisement au sein d'un AG est primordial, puisque c'est son efficacité qui détermine la bonne ou la mauvaise exploration de l'espace de recherche. Il réalise à cet effet un échange de *gènes* entre deux parents, pour créer deux enfants. Deux aspects des opérateurs de croisement peuvent être modifiés. D'abord, le choix des *individus* à croiser (appariement) et ensuite, la procédure pour réaliser le croisement. Pour la première étape, l'appariement aléatoire a été utilisé. En ce qui concerne la deuxième étape, plusieurs stratégies de croisement peuvent être utilisées [Eshelman, 93]. Le croisement en 2 points a été choisi dans un premier temps. Cet opérateur, appliqué sur chaque couple de parent, choisit 2 points de croisement aléatoirement et inverse alternativement le segment sélectionné dans chaque parent (figure 3.6).



**Fig. 3. 6.** Représentation du croisement en 2 points

Après chaque opération de croisement, l'ensemble des *gènes* de chaque enfant est considéré pour l'opération de mutation.

### 3.2.2.3.3.5. Mutation

L'opérateur de mutation a pour rôle d'apporter de la diversité dans la *population*. En effet, celle-ci a tendance à s'uniformiser de génération en génération, c'est-à-dire qu'elle contient de plus en plus d'*individus* 'clones' de son meilleur *individu*. La mutation sert aussi à générer des solutions négligées par le croisement ou la sélection. La mutation consiste à remplacer la valeur du *gène* sélectionné. Dans le cas d'un codage binaire, la valeur 1 est remplacée par 0 et vice versa. Cet opérateur est appliqué avec une probabilité  $P_m$  sur un *individu* parmi les  $P/2$  *individus* (enfants). Le *gène* de l'*individu* à muter est tiré au sort par la suite. Cette probabilité  $P_m$  est prise égale à  $1/d$ , avec  $d$  le nombre de *gène* composant un *chromosome* (égal au nombre de paramètres formant le vecteur forme).

Avec les opérations de croisement et de mutation, il est possible que des enfants soient générés sans paramètre sélectionné (ensemble des gènes à zéro). Pour éviter ce problème, les opérations sont répétées jusqu'à obtenir des enfants 'valides'.

De plus, sous l'effet de la mutation, le meilleur *individu* à la génération  $t$  peut disparaître à la génération  $t+1$ . Pour pallier cet inconvénient, on renonce à muter le meilleur *individu* trouvé jusqu'ici. Par conséquent, à chaque nouvelle génération, le meilleur *individu* est incorporé dans la *population*, ce qui garantit sa sauvegarde.

### 3.2.2.3.3.6. Arrêt de l'AG

Pour stopper le fonctionnement de l'AG, il paraît simple de regarder l'évolution de la meilleure fonction d'adaptation à chaque génération. Ainsi, lorsque de la génération  $t$  à la génération  $t+1$ , cette évolution devient nulle, cela indique que l'algorithme n'a pas trouvé de



meilleure solution au problème étudié. Cependant, pour ne pas risquer d'arrêter l'algorithme dans une solution sous-optimale, l'arrêt complet se fait lorsque aucun changement n'est apparu dans la meilleure fonction d'adaptation pendant 10 générations successives.

Afin d'évaluer les performances de cette méthode de sélection à base d'algorithme génétique, une comparaison des résultats tels que le choix des  $d'$  paramètres mais aussi le nombre de ces paramètres sera faite dans le chapitre IV avec les méthodes présentées précédemment (SBS, critère de Fisher).

### ***3.3. Détermination de l'espace de décision***

Cette étape consiste à partitionner l'ensemble d'apprentissage en classes. Soit  $(\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  un ensemble de  $N$  vecteurs constituant l'ensemble de données initiales, l'objectif est donc de structurer ces vecteurs et d'établir une correspondance entre les classes définies et les différents états du système.

Lorsque la classe d'origine de chaque observation est connue, l'espace de décision est parfaitement connu et l'apprentissage peut être effectué en mode *supervisé*. En terme de diagnostic, une telle approche permettrait de vérifier qu'on retrouve bien les différentes classes ou modes de fonctionnement.

Si au contraire aucune information sur la structuration de l'ensemble d'apprentissage en classes n'est disponible, alors la classification doit être faite en mode *non supervisé*. Il existe différentes méthodes de classification non supervisée, appelées également méthodes de coalescence.

Tout d'abord, il faut distinguer les méthodes réalisant une classification exclusive ou non. Dans le second cas, une observation peut appartenir à plusieurs sous-ensembles. En effet, chaque observation pourrait être munie d'un vecteur de fonctions d'appartenance caractérisant le degré d'appartenance de celle-ci à chacune des classes. [Boutleux, 96], [Boudaoud, 97] présentent des algorithmes de classification floue pour réaliser le diagnostic de diverses processus.

Dans le premier cas, une observation appartient à un sous-ensemble (une classe) et à un seul. Dans ce cas de figure, ces méthodes de coalescence se décomposent en deux parties :

- les méthodes de classification par partition

Ces méthodes ont pour but de créer des regroupements d'observations autour de centres de gravité recherchés. Elles sont dérivées d'une classe d'algorithmes connue sous le nom de 'méthodes des nuées dynamiques' introduites par [Diday, 76]. Parmi ces algorithmes, on distingue tout particulièrement la coalescence stricte Hard C-Means (HCM) introduite par [Duda, 73].

L'algorithme HCM fonctionne de la manière suivante :

Algorithme HCM
<ol style="list-style-type: none"> <li>1. Choix du nombre de classes M</li> <li>2. Partition aléatoire de l'ensemble des observations en M classes</li> <li>3. Calcul des centres de gravité des M classes</li> </ol> <p><b>Répéter</b></p> <ol style="list-style-type: none"> <li>4. Restructuration des classes par affectation au centre de gravité le plus proche</li> <li>5. Calcul des nouveaux centres de gravité</li> <li>6. Calcul du critère d'arrêt</li> </ol> <p><b>Jusqu'à</b> satisfaction du critère d'arrêt</p>

Généralement le critère d'arrêt suit l'évolution du partitionnement des observations en classes. Si l'affectation de l'une des observations a été modifiée entre les itérations  $t$  et  $t+1$  alors l'algorithme continue. A l'inverse lorsque la partition reste inchangée, l'algorithme s'arrête. En d'autres termes, cela signifie qu'il a défini une structure stable en classes de l'ensemble et que celle-ci n'évoluera plus.

Le résultat final de la coalescence dépend fortement de la partition aléatoire initiale. Il est donc conseillé de lancer la recherche à plusieurs reprises à partir de partitions initiales différentes et de retenir la classification qui présente le plus faible taux d'erreur.

Cet algorithme (HCM) est utilisé par la suite, pour comparer les résultats de coalescence obtenus avec une méthode hiérarchique, que nous allons présenter ci-dessous.

- les méthodes hiérarchiques (ascendantes ou descendantes)

Elles réalisent une arborescence dont les éléments à la base sont les observations elles-mêmes et l'élément au sommet est composé de l'ensemble d'apprentissage au complet. L'algorithme de classification ascendante procède à un regroupement progressif des éléments, alors que la classification descendante vise plutôt l'éclatement des sous-ensembles.

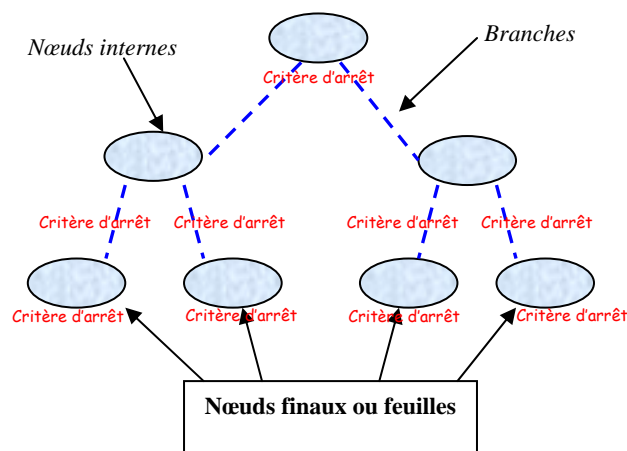
Ces différentes méthodes ont été largement développées dans la littérature [Poublan, 93].

Nous ne présentons ici qu'une méthode *non supervisée* à classification exclusive, les arbres de décision, qui est une méthode hiérarchique descendante. Pour améliorer les performances de cet algorithme, il est couplé à une méthode par partition. Ainsi, l'algorithme principal est constitué par la méthode des arbres de décision et celui-ci fera appel à un sous-programme qui utilise l'algorithme HCM. Cette technique est exploitée par la suite, dans le développement du diagnostic des machines asynchrones.

Les arbres de décision constituent des outils de classification très employés dans des problèmes de RdF variés tels que la télédétection [Bartolucci, 76], la reconnaissance de caractères [Gu, 83], ou encore le pronostic et le diagnostic médical [Anderson, 79]. [Didelet, 92] a, quant à elle, utilisé les arbres de neurones, qui associent un ensemble de neurones et un arbre de décision binaire pour le diagnostic du pilotage en temps réel du réseau téléphonique français.

Un arbre de décision est composé de plusieurs *nœuds internes* reliés entre eux par des *branches*. Ces *nœuds internes* représentent les états (classes) intermédiaires lors de la construction de l'arbre tandis que les *branches* représentent le processus de construction. Il possède également un certain nombre de *nœuds terminaux* ou *feuilles*, correspondant aux classes finales, auxquels on associe un indicateur de classe (étiquette correspondant à son mode de fonctionnement), parmi les M classes du problème à traiter.

L'intérêt de la classification par arbres de décision est de diviser les données contenues dans un ensemble d'apprentissage, par des tests définis à l'aide d'attributs (calcul de distance) jusqu'à ce que l'on arrive à des sous-ensembles de points, obtenus en suivant l'évolution d'un critère, ne contenant presque que des données appartenant toutes à la même classe.



**Fig. 3. 7.** Représentation d'un arbre de décision

L'intérêt de ces méthodes en général réside dans le fait qu'il est possible de déterminer automatiquement le partitionnement de l'ensemble d'apprentissage, c'est-à-dire le nombre de classes. Pour cela, la construction de l'arbre doit pouvoir s'interrompre à tout instant. Cette construction est donc un problème crucial pour ce type d'outil. Bien entendu, la structure de l'arbre doit être choisie de façon à obtenir un outil de classification optimal. Plusieurs méthodes de construction optimale peuvent être proposées [Kulkarni, 76], [Payne, 77], mais elles demandent souvent un large temps d'exécution et une importante place mémoire. C'est pourquoi, il est préférable de construire des arbres de décision sous optimaux en utilisant des méthodes heuristiques efficaces [Safavian, 91].

Il existe de nombreuses méthodes heuristiques pour la construction de ces arbres, dont certaines sont des approches de type *Top-Down* et d'autres des approches de type *Growing-Pruning* [Breiman, 84]. Il existe également des méthodes de type *Bottom-Up* [Landeweerd, 83].

Les méthodes de type *Growing-Pruning* consistent à laisser pousser l'arbre jusqu'à ce que tous les sous-ensembles soient homogènes, puis à réaliser un élagage de l'arbre, afin d'obtenir un sous-arbre de classification moins important. Cette approche permet d'éliminer le problème du choix d'un critère d'arrêt, qui joue un rôle dans les performances de l'arbre. Mais a contrario, cette technique impose des temps de calcul beaucoup plus longs, qui peuvent être un inconvénient pour la réalisation d'un diagnostic en ligne.

Par conséquent, une approche de type *Top-down* est, en général, utilisée pour la construction de l'arbre.

Les méthodes de type *Top-down* ont la particularité de construire l'arbre de décision en commençant par la cime de l'arbre et en descendant en profondeur vers les feuilles. Pour cette approche, la construction d'un arbre se réduit aux trois tâches suivantes :

1. Trouver une règle ou un critère pour créer chaque nœud,
2. Choisir un critère pour arrêter la construction, c'est-à-dire déterminer à quel moment un nœud devient un nœud terminal,
3. Assigner une étiquette de classe aux feuilles.

Parmi ces trois tâches, la première est celle qui est la plus difficile à effectuer, ce qui pourrait expliquer la concentration des recherches sur ce problème [Li, 86], [Wu, 75], [Swain, 77], [Quinlan, 86]. Pour cette tâche, l'algorithme HCM a été utilisé. En effet, à partir de l'ensemble d'apprentissage initial, deux attributs sont déterminés aléatoirement aux alentours

du centre de gravité de l'ensemble, en lui ajoutant et lui retranchant une faible valeur  $\varepsilon$  sur chaque dimension (cette valeur peut être choisie en prenant un faible pourcentage de l'écart type de chaque composante), et ensuite, l'algorithme HCM est utilisé pour associer chaque vecteur de  $X_a$  à l'attribut le plus proche en terme de distance (§ 3.5.1.1.1). Cette étape s'arrête lorsque le critère d'arrêt de HCM est satisfait. Ainsi, l'ensemble d'apprentissage a été partitionné en deux classes (nœuds) distinctes. La décomposition de ces classes en sous-classes se poursuit, développant la construction des branches de l'arbre.

Pour arrêter cette construction, il est nécessaire d'implanter sur chaque branche de l'arbre, un critère d'arrêt (CA) (figure 3.7). C'est la deuxième tâche.

De manière intuitive, à la fin de la coalescence, sur chaque branche, la variance intra-classe est minimisée. On peut la quantifier grâce au calcul d'un critère d'arrêt CA qui correspond à la trace de la matrice de dispersion intra-classe :

$$CA = \text{trace}(\Sigma_w) \quad (3.24)$$

Ce critère est affecté à chaque nœud interne de l'arbre. En suivant l'évolution de ce critère, il est possible de définir l'arrêt d'une branche. Cet arrêt peut être dû à une rupture ou cassure dans cette évolution qui correspondrait à une diminution peu significative de la compacité de la classe (figure 3.4).

La troisième tâche est réalisée par un expert, qui associe à chaque classe un mode de fonctionnement du processus.

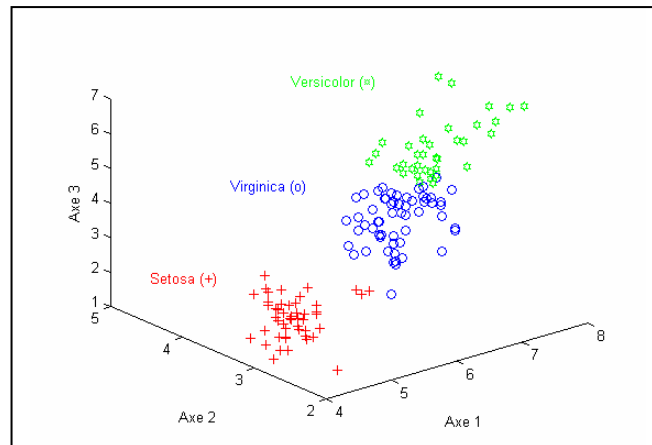
Le fonctionnement peut se mettre sous la forme algorithmique :

Algorithme d'un arbre de décision
<ol style="list-style-type: none"> <li>1. Calcul du centre de gravité de l'ensemble d'apprentissage</li> <li>2. Détermination de 2 attributs (2 échantillons) à partir de ce centre de gravité</li> </ol> <p><b>Recherche des classes :</b></p> <ol style="list-style-type: none"> <li>3. Associer les vecteurs de l'ensemble d'apprentissage à l'attribut le plus proche en terme de distance (algorithme HCM)</li> <li>4. Calcul des centres de gravité des deux classes obtenues</li> <li>5. Répéter les étapes 3 et 4 jusqu'à obtenir un extremum pour le critère retenu (nombre d'itérations, stationnarité de la partition)</li> <li>6. Déterminer 2 attributs à partir de chaque centre de gravité des deux nouvelles classes (soit 4 attributs)</li> <li>7. Répéter les étapes 3, 4, 5 et 6 jusqu'à obtenir une séparation en M classes satisfaisant un critère d'arrêt basé sur la compacité des classes (équation 3.24).</li> </ol>

Cet algorithme est illustré à travers l'exemple suivant utilisant les *données iris* [Fisher, 36]. Ces données sont relatives à 3 types de fleur, appelées Iris Setosa, Iris Versicolor et Iris

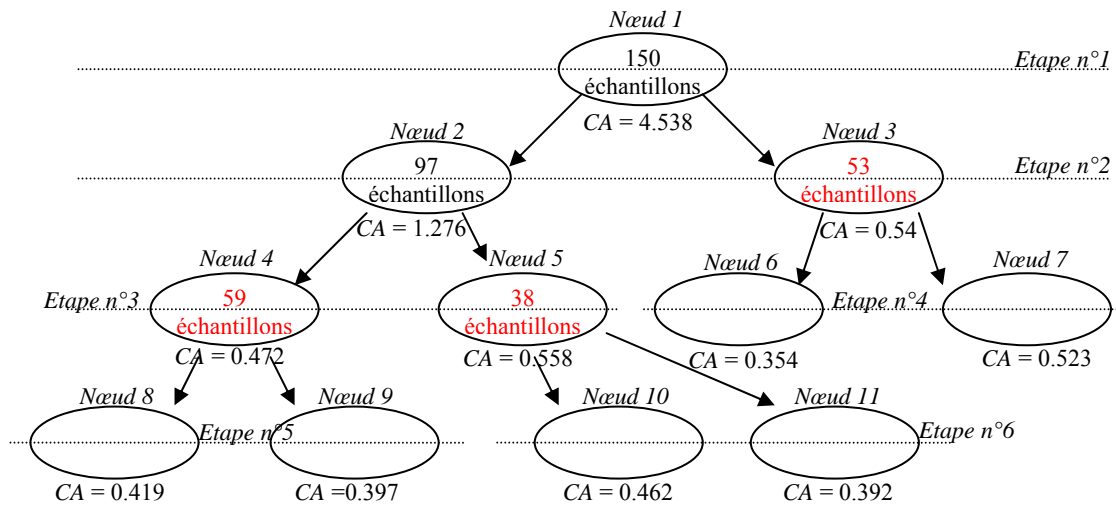
Virginia. Afin de différencier ces 3 familles, quatre grandeurs caractéristiques (paramètres) ont été mesurées sur un ensemble de 50 fleurs par famille : largeur et longueur du pétale, largeur et longueur du sépale. On dispose donc d'un ensemble de 150 données dans  $\mathcal{R}^4$ .

La figure 3.8 présente la classification des *données iris* obtenue. Cette visualisation, obtenue à l'aide d'une ACP, est réalisée dans un espace tridimensionnel constitué par les trois premiers axes factoriels.

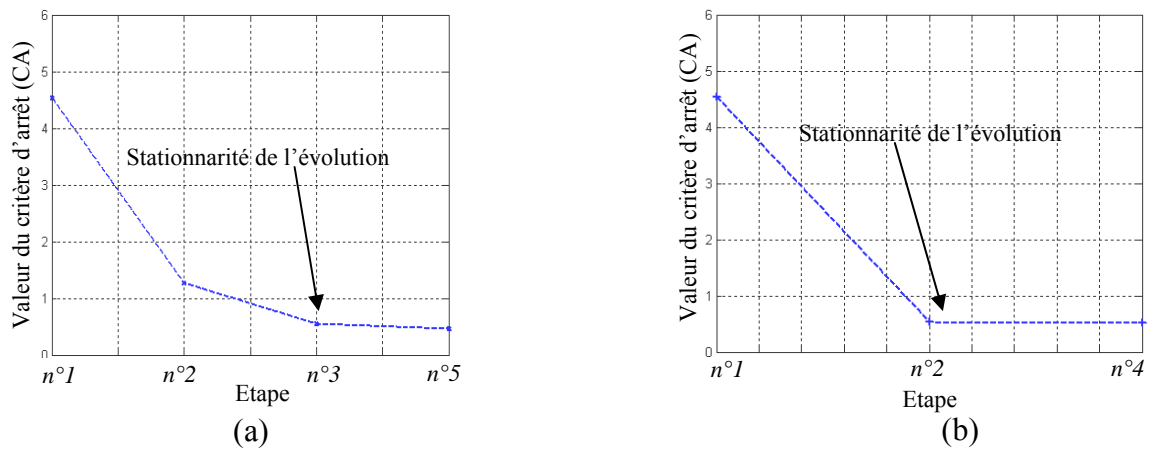


**Fig. 3. 8.** Classification par arbre de décision des 3 types d'iris (Setosa : +) (Virginica : o) (Versicolor : \*)

A l'issue de la convergence de l'algorithme précédent, l'ensemble des données a été séparé en trois classes de 53, 59 et 38 échantillons. Sur la figure 3.9, les nœuds 3, 4 et 5 représentent ces classes. Ils sont aussi appelés *nœuds terminaux*. On peut voir, qu'à chaque nœud est associé la valeur du critère d'arrêt, qui n'est autre que la trace de la matrice de dispersion intra-classe (équation 3.24), permettant ainsi de stopper la construction des branches de l'arbre. En effet, sur la figure 3.10, sont représentées deux évolutions de ce critère pour les différentes partitions. Lorsqu'une rupture (stationnarité) apparaît dans cette évolution, cela signifie que le partitionnement de l'étape  $t$  à l'étape  $t+1$  n'a pas permis de diminuer de façon significative la compacité des classes. Au final, la structure en classes a bien été trouvée.



**Fig. 3. 9.** Evolution de l'arbre de décision dans le cas des données Iris



**Fig. 3. 10.** Evolution de la valeur du critère d'arrêt. (a) Noeud 1, 2, 4 et 8, (b) Noeud 1, 3 et 7

Les résultats de la classification sont résumés dans le tableau suivant :

	$\Omega_1$	$\Omega_2$	$\Omega_3$
Setosa	50	3	0
Virginica	0	47	12
Versicolor	0	0	38

**Tableau. 3. 1.** Résultats de la classification par la méthode des arbres de décision

En comparaison, la méthode des HCM donne les résultats de classification suivants. Ces résultats ont été obtenus après avoir testé différentes initialisations. Dans le tableau 3.2 est donc présenté le meilleur résultat obtenu :

	$\Omega_1$	$\Omega_2$	$\Omega_3$
Setosa	49	0	1
Virginica	0	37	13
Versicolor	0	10	40

**Tableau. 3. 2.** Résultats de la classification par la méthode des HCM

Pour évaluer les performances de la classification choisie par rapport au problème posé, il est nécessaire d'utiliser des critères de qualité. Ainsi, deux critères peuvent être retenus :

- Le taux d'observations bien classées :

$$T_B = \frac{\text{Nombre de bien classées}}{N} = \frac{N_B}{N} \times 100 \quad (3.25)$$

- Et le taux de mal classées :

$$T_M = \frac{\text{Nombre de mal classées}}{N} = \frac{N_M}{N} \times 100 = 100\% - T_B \quad (3.26)$$

Ces deux quantités ont beaucoup d'importance. Si le taux d'erreur est proche de 0 %, alors le vecteur forme caractérise bien l'ensemble d'apprentissage, et la méthode de classification est adaptée.

De ce fait, il est possible de réaliser une comparaison entre les deux méthodes de classification employées ci-dessus. Le critère de qualité correspondant à ces deux méthodes est :

	Arbre de décision	HCM
$T_B$	90 %	84 %

**Tableau. 3. 3.** Critère de qualité correspondant aux deux méthodes de classification

Le taux d'erreur obtenu avec la méthode des arbres de décision est meilleur que celui obtenu avec la méthode des HCM. Ce taux, pour la classification est de 10 %, ce qui, compte tenu de la nature mélangée des données s'avère être un bon résultat.

Signalons à titre indicatif que Bezdek a constaté aussi bien en utilisant des algorithmes de classification flous que non flous des taux d'erreur variant de 5 à 40% avec une moyenne de 15% [Bezdek, 74].

Les différentes étapes de l'apprentissage précédemment citées dans les parties 3.2 et 3.3



peuvent être combinées simultanément pour rechercher d'une part, la meilleure représentation possible des observations (détermination du vecteur forme optimal) et, d'autre part, la meilleure structure de l'ensemble d'apprentissage en classes. C'est ce qui va être présenté dans la partie suivante.

### ***3.4. Association simultanée de la coalescence et de la sélection de paramètres***

#### **3.4.1. Introduction**

Dans la littérature, de nombreuses méthodes ont été rapportées au cas de la sélection des paramètres et de la coalescence en utilisant soit le recuit simulé [Brown, 92], [Selim, 91] ou les algorithmes évolutionnaires [Brill, 92], [Krishna, 99], [Hall, 99], [Raymer, 00], [Yang, 98]. Ces méthodes entreprennent souvent l'étape de coalescence et de sélection de paramètres séparément. En effet, ces méthodes dédiées à la coalescence, [Brown, 92], [Selim, 91], [Hall, 99] supposent que tous les paramètres sont d'importances égales tandis que les méthodes dédiées à la sélection des paramètres [Raymer, 00], [Yang, 98] sont développées en mode *supervisé*. Il semble alors intéressant de réunir ces deux tâches simultanément, c'est-à-dire de développer une méthode de sélection de paramètres en mode *non supervisé* (sans connaissance sur la structure en classes de  $X_a$ ).

Pour cela, la méthode de sélection de paramètres, développée dans 3.2.2.3, et basée sur l'application des algorithmes génétiques, servira de base au développement de cette double tâche.

#### **3.4.2. Algorithme proposé**

Récemment, beaucoup de méthodes hybrides 'algorithme génétique – recherche locale', qui incorporent les procédures d'amélioration locale aux AGs traditionnels, ont été proposées et appliquées avec succès pour résoudre une grande variété de problèmes d'optimisation. Nous utiliserons ce type de méthodes pour faire de la sélection de paramètres une partie intégrante de la procédure de structuration en M classes de  $X_a$  (étape de coalescence). Pour réaliser ceci, une représentation composée d'éléments binaires et de nombres réels, codant respectivement, les paramètres et les centres de gravité, est créée. Cette représentation forme un *chromosome*, qui constitue la *population* initiale. Le nombre de classes étant inconnu, la taille des centres

de gravité, codée dans un *chromosome*, peut varier. Par conséquent, les opérations de croisement et de mutations sont modifiées afin d'aborder le concept des *chromosomes* composés avec des tailles variables. De plus, les opérateurs de recherche locale sont conçus pour raffiner la sélection des paramètres et des centres de gravité codés dans les *chromosomes*. Ces opérateurs déplacent des solutions vers des optima locaux, et les améliorations de ces opérateurs sont accumulées au fil des générations, ayant pour résultat une amélioration importante des performances globales.

### 3.4.2.1. Fonction d'adaptation

Le problème est toujours du même ordre, c'est-à-dire obtenir la meilleure séparation entre les différentes classes de  $X_a$  en choisissant les meilleurs paramètres. Par contre, cette fois-ci, aucune connaissance n'est disponible sur le nombre de classes *a priori*. Toutefois, la fonction d'adaptation utilisée en 3.2.2.3.3.2 (équation 3.23) semble parfaitement convenir à ce genre de problème.

### 3.4.2.2. Codage utilisé pour la représentation d'un *chromosome*

Chaque *chromosome*  $k$  ( $k=1,P$ ) de la *population* initiale est composé de variables représentant la sélection des paramètres et les centres de gravité des classes. La partie 'sélection de paramètres' est une chaîne de  $d$  éléments binaires ou chacun représente un paramètre, les valeurs 1 et 0 signifiant respectivement 'sélectionné' ou 'éliminé'. La partie 'centre de gravité' est composée de  $M_k \times d$  nombres réels, ou  $M_k$  est le nombre de classes codé dans le *chromosome*  $k$ . Les  $d$  premières positions représentent les  $d$  dimensions d'un premier centre de gravité, les  $d$  suivantes positions représentent ceux du second centre de gravité, et ainsi de suite. Par exemple, dans un espace de dimension quatre avec 3 classes, le *chromosome*  $k$ , représenté par le vecteur  $V_k$ ,

$$V_k = \left[ \begin{array}{c|cccccccccccc} \text{paramètres} & & & & & & & & & & & & & & \\ \hline 1 & 0 & 1 & 0 & 1.5 & 2.1 & 1.7 & 3.5 & 23.2 & 20.9 & 15.8 & 16.7 & 7.8 & 6.9 & 9.5 & 7.6 \end{array} \right] \quad (3.27)$$

code trois centres de gravité, (1.5, 2.1, 1.7, 3.5), (23.2, 20.9, 15.8, 16.7), et (7.8, 6.9, 9.5, 7.6) avec le premier et le troisième paramètres sélectionnés. Seuls ces paramètres sont utilisés pour évaluer les centres de gravité.

Chaque *chromosome*  $k$  parmi les  $P$  initiaux formant la *population* est construit, en ce qui concerne la partie 'sélection de paramètres', par l'attribution aléatoire d'éléments binaires pour chacun des  $d$  paramètres. Pour la partie 'centre des classes', les valeurs initiales de ces

centres sont déterminées de façon aléatoire dans un intervalle correspondant à la gamme de variation de chacun des paramètres. Le nombre de classes initiales  $M_k$  est choisi lui aussi de manière aléatoire entre l'intervalle  $[2, M_{kmax}]$  avec  $M_{kmax}$  le nombre maximal de classes. Ce nombre est pris égal à  $\sqrt{N}$  ( $N$ , le nombre de vecteurs contenu dans  $X_a$ ). Cette règle est souvent utilisée dans la littérature [Pal, 95].

### 3.4.2.3. Croisement et mutation

En ce qui concerne la partie 'sélection de paramètres' codée dans un *chromosome*, l'opérateur croisement est le même que celui utilisé dans 3.2.2.3.3.4. Le croisement en 2 points, appliqué sur chaque couple de parents, choisit 2 points de croisement de façon aléatoire et inverse alternativement le segment sélectionné dans chaque parent (figure 3.6). Après chaque croisement, les  $d$  éléments binaires des  $P$  individus sont considérés pour la mutation. La valeur 1 est remplacée par la valeur 0 et vice versa. Cet opérateur est appliqué sur un individu avec une probabilité  $P_m$  égale à  $1/d$ . L'élément binaire à muter est choisi aléatoirement. De même que dans 3.2.2.3, ces opérateurs peuvent générer des enfants 'non valides' (sans paramètres sélectionnés). De ce fait, les opérateurs sont répétés jusqu'à avoir au moins un paramètre sélectionné.

Concernant la partie 'centre des classes', le même opérateur de croisement en 2 points est utilisé. Durant le croisement, les centres des classes sont considérés comme indivisibles, c'est-à-dire que les points de croisement ne peuvent se trouver qu'entre deux centres. Dans ce but, l'opérateur de croisement est défini comme suit :

- Deux *chromosomes*  $k_1$  et  $k_2$  codent respectivement  $M_1$  et  $M_2$  centres de gravité.  $x_1$  et  $x_2$ , les points de croisement de  $k_1$  sont générés aléatoirement dans l'intervalle  $[0, M_1]$ . Si  $x_1$  est plus grand que  $x_2$ , il est nécessaire de permuter ces deux valeurs pour s'assurer que  $x_2 > x_1$ . Les points de croisement de  $k_2$  sont ensuite générés toujours de façon aléatoire dans l'intervalle  $[0, M_2 - |x_2 - x_1|]$  pour  $x_3$  et  $x_4 = x_3 + |x_2 - x_1|$ , où  $|x_2 - x_1|$  est la taille du segment entre les points de croisement  $x_1$  et  $x_2$ . Après cela, l'information du segment entre  $x_1$  et  $x_2$  de  $k_1$  est échangée avec l'information du segment entre  $x_3$  et  $x_4$  de  $k_2$ . Avec la règle ci-dessus, le nombre de classes des enfants sera le même que leurs parents  $k_1$  et  $k_2$ . L'opérateur de croisement est exécuté sur chaque couple de parents. Après le croisement, une faible probabilité de mutation gaussienne est appliquée sur les enfants. La mutation gaussienne additionne une valeur aléatoire de distribution gaussienne sur une des dimensions

d'un centre de gravité. La nouvelle valeur de la dimension est coupée si elle se situe en dehors de limites inférieures ou supérieures de cette dimension.

#### 3.4.2.4. Opérations de recherche locale

Les algorithmes génétiques convergent vers l'optimum global grâce aux opérateurs de croisement et de mutation. Cependant, pour des problèmes complexes, il se peut que le temps de calcul pour converger vers la solution optimale soit assez long. Par conséquent, l'incorporation d'un algorithme de recherche locale améliore notablement les performances d'AG.

Ainsi, deux opérateurs supplémentaires, *ajout* et *élimination*, sont introduits. Ces opérateurs opèrent sur la partie 'sélection de paramètres'. Ils permettent la recherche vers des optima locaux durant le processus de sélection de paramètres.

- *Ajout* : choisir le paramètre le plus discriminant à partir du sous-ensemble de paramètres non sélectionnés, et changer son statut en 'sélectionné'.
- *Élimination* : choisir le paramètre le moins discriminant à partir du sous-ensemble de paramètres sélectionnés, et changer son statut en 'éliminé'.

Les opérateurs travaillent avec une solution courante de sélection de paramètres, codée dans un *chromosome*, et ils exécutent des recherches locales autour de cette solution en enlevant le paramètre le moins discriminant et en ajoutant le plus discriminant. De ce fait, il y a toujours le même nombre de paramètres sélectionnés. Cette procédure est appliquée sur tous les nouveaux enfants durant chaque génération.

Afin d'améliorer l'efficacité de calcul, l'algorithme des HCM est appliqué sur les centres des classes codés dans tous les nouveaux enfants durant chaque génération, après la procédure d'*ajout* et d'*élimination* de paramètres. Cette étape est réalisée en affectant les vecteurs de  $X_a$  aux classes dont le centre de gravité, codé dans le *chromosome*, est le plus proche (en terme de distance soit euclidienne, soit de Mahalanobis). Après cela, les centres de gravité, codés dans le *chromosome*, sont remplacés par la moyenne des vecteurs de leurs classes respectives.

#### 3.4.2.5. Problème multimodal

Pour un AG standard, la *population* converge rapidement vers une région prometteuse de l'espace de recherche. Cependant, cette caractéristique, attrayante pour les problèmes unimodaux, peut ne pas être intéressante pour beaucoup de problèmes. En effet dans le cas de fonctions multimodales telles que la coalescence et la sélection de paramètres, l'algorithme

risque de converger vers un optimum local, du fait de la perte de la diversité au sein de la *population*. De plus, le couplage avec les opérations de recherche locale peut accentuer cet effet puisque la recherche locale accélère la convergence des algorithmes génétiques. Par conséquent, une certaine diversité doit être maintenue, tout au long des générations, au sein de la *population*, afin d'éviter que tous ces *individus* ne soient des clones du meilleur d'entre eux. Le principe consiste à empêcher l'accumulation d'*individus* dans une région de l'espace de recherche, en considérant des sous-populations, autour des meilleurs points de la fonction, au sein de la *population*. Plusieurs méthodes peuvent être utilisées comme le *sharing* [Goldberg, 94], le *clearing* ou le *crowding* [Mahfoud, 92].

### 3.4.2.6. AG multi-population pour l'optimisation multimodale

L'algorithme proposé permet d'identifier plusieurs optima de la fonction, en générant des sous-populations centrées sur ces optima. Cette caractéristique est très intéressante du fait qu'en général, les problèmes d'optimisation réels, tels que la coalescence, mènent souvent à plusieurs optima locaux. La recherche des optima d'une fonction est réalisée à travers un processus d'*exploration*, qui a pour tâche de localiser les zones prometteuses de la fonction (pics pour le cas d'une maximisation). Durant cette phase, dite de '*diversification*' (exploration), l'algorithme génétique multi-population détecte les 'niches' potentielles, en regroupant les *individus* dans différentes sous-populations, en fonction de leur similarité (*spéciation*).

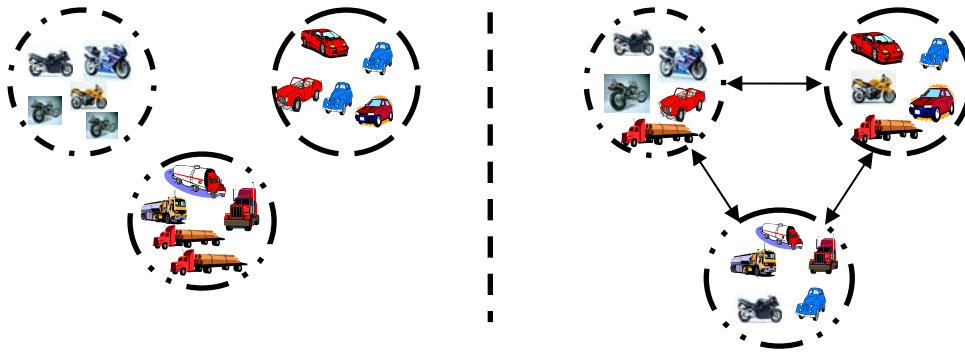
#### 3.4.2.6.1. Diversification

Afin de maintenir une diversité suffisante au cours de l'évolution, une organisation de la *population* en sous-populations a été proposée. Ce scindement de la *population* est réalisé lors de l'étape dite de '*spéciation*'. Celle-ci a pour fonction de maintenir une certaine diversité entre les sous-populations, en regroupant des *individus* exploitant une même région au sein d'une même sous-population.

Cependant, pour encourager la prolifération de bons *individus* dans l'ensemble des sous-populations, un échange d'*individus* se fait périodiquement entre celles-ci, à travers l'opérateur de '*migration*'. En outre, l'échange d'*individus* issus de régions différentes permet de générer (par le biais de l'opérateur de croisement) des *individus* de nouvelles régions, encore non explorées, et peut être prometteuses.

### 3.4.2.6.1.1. Migration

Cet opérateur consiste en un échange d'*individus* entre les sous-populations. Ce mécanisme, représenté sur la figure 3.11, permet de garder une certaine diversité au sein des sous-populations et de mieux explorer l'espace de recherche. La sélection des *individus* à migrer et de ceux qui seront remplacés peut être faite de différentes manières, chacune d'elles apportant un niveau de diversité plus ou moins accentué.



**Fig. 3. 11.** Représentation symbolique de la migration dans un AG multi-population

Ainsi, la *migration* choisie, se fait entre toutes les sous-populations, et de façon aléatoire. En ce qui concerne la stratégie de *sélection/remplacement*, qui détermine pour la sous-population réceptrice, l'*individu* à remplacer et celui qui le remplace, il a été choisi de remplacer le plus mauvais *individu* par un *individu* au hasard.

### 3.4.2.6.1.2. Spéciation

La spéciation est le regroupement des *individus* selon leurs ressemblances génétiques. De ce fait, après cette étape, chaque sous-population regroupe des *individus* 'semblables' exploitant un même pic de la fonction d'adaptation. Le nombre  $n$  de sous-populations retenu dépend du paysage de la fonction d'adaptation. Idéalement  $n$  doit être égal au nombre de pics de cette fonction. Pour notre problème de coalescence associée à la recherche du sous-ensemble de paramètres le plus discriminant, on distingue autant de sous-populations (niches) qu'il y a de nombre de classes, c'est-à-dire de 2 à  $\sqrt{N}$  ( $N$ , nombre de vecteurs dans  $X_a$ ). Ceci peut être résumé de la façon suivante : les niches principales regroupent les *individus* qui présentent le même nombre de classes mais un nombre de paramètres sélectionnés différent.

Dans ses niches principales, il est encore possible de diviser en sous-populations les *individus* en y regroupant ceux avec le même nombre de paramètres.

Prenons un exemple : on crée une *population* initiale composée de 60 *individus*. Tout d'abord on regroupe les *individus* constitués du même nombre de classes. Ceci forme les niches principales. Ensuite dans chaque niche, les *individus* sont constitués d'un nombre de paramètres différents. Il est donc possible de regrouper les *individus* possédant le même nombre de classes et un même nombre de paramètres de la façon suivante :

Niche 1				Niche 2			Niche 3			
20 individus avec 2 classes				10 individus avec 3 classes			30 individus avec 4 classes			
Niche 11	Niche 12	Niche 13	Niche 14	Niche 21	Niche 22	Niche 23	Niche 31	Niche 32	Niche 33	Niche 34
5 individus avec 3 paramètres	2 individus avec 5 paramètres	7 individus avec 7 paramètres	6 individus avec 2 paramètres	3 individus avec 3 paramètres	5 individus avec 2 paramètres	2 individus avec 6 paramètres	9 individus avec 3 paramètres	2 individus avec 2 paramètres	3 individus avec 5 paramètres	6 individus avec 7 paramètres

**Tableau. 3. 4.** Formation des niches

Au final, il y a 11 niches de créées : les 4 premières, les niches 11 à 14, regroupent les *individus* constitués de 2 classes et respectivement de 3, 5, 7 et 2 paramètres. Les 3 niches suivantes, 21, 22 et 23, regroupent les *individus* formés de 3 classes et respectivement de 3, 2 et 6 paramètres. Les quatre dernières niches, 31, 32, 33 et 34, regroupent les *individus* constitués de 4 classes et respectivement 3, 2, 5 et 7 paramètres.

Une fois que les *individus* ‘semblables’ sont regroupés en sous-populations et que l’opération de migration est effectuée, les *individus* parents vont être sélectionnés et ensuite couplés pour donner naissance, grâce aux opérateurs de croisement et de mutation, à des *individus* enfants. Concernant l’opération de sélection, la stratégie utilisée est la même que dans le paragraphe 3.2.2.3.3.3, c’est-à-dire une sélection par tournoi avec en plus l’insertion du meilleur individu si celui-ci n’a pas été sélectionné lors du tournoi. Concernant l’accouplement des parents, ce choix est fait de façon aléatoire en tirant deux parents et en les associant. Bien sur, cette procédure de sélection et d’accouplement est réalisée dans chaque sous-population.

### 3.4.2.7. Arrêt de l’algorithme

L’arrêt de l’algorithme multi-population se fait lorsque, de la génération  $t$  à la génération  $t+10$ , les meilleures fonctions d’adaptation présentent dans chaque sous-population n’ont pas variées. Par conséquent, si pendant 10 générations successives, chaque fonction présente une

stationnarité dans son évolution, l'algorithme compare la meilleure fonction entre chaque sous-population et choisit celle qui est la plus grande. Cette fonction est associée à un *chromosome*, qui lui, va donner le nombre de classes (et leurs centres de gravité) optimal ainsi que le sous-ensemble de paramètres maximisant la séparabilité et la compacité des classes.

Au final, l'algorithme permettant de rechercher simultanément le nombre de classes optimal structurant l'ensemble d'apprentissage et le sous-ensemble de paramètres maximisant la séparabilité et la compacité de ces classes, peut se résumer comme suit

#### Algorithme génétique multi-population

- 1) Initialisation aléatoire de  $k$  ( $k=1,P$ ) *chromosomes* qui codent à la fois, la présence ou non des paramètres et les centres des classes avec différents nombre de classes, en utilisant respectivement un codage binaire et réel (§ 3.4.2.2).
- 2) Calcul de la fonction d'adaptation en fonction de (3.22) pour chaque *individu* dans la *population* initiale (§ 3.4.2.1).
- 3) Répéter de (a) à (f) jusqu'à ce que le critère d'arrêt soit rencontré (§ 3.4.2.7).
  - a) Définir l'ensemble des sous-populations (niches) regroupant les *individus* 'semblables' et réaliser l'opération de migration (§ 3.4.2.6)
  - b) Sélectionner les couples de parents dans chaque sous-population à partir de l'opérateur de sélection (sélection par tournoi + élitisme) (§ 3.4.2.6). Cette procédure se répète jusqu'à ce que  $P/2$  parents soient sélectionnés.
  - c) Générer les enfants en appliquant les différents opérateurs génétiques (croisement et mutation) sur les différentes parties (sélection des paramètres et centre des classes) des couples de parents (§ 3.4.2.3).
  - d) Appliquer les opérateurs *ajout* et *élimination* sur la partie 'sélection des paramètres' codée dans les enfants et ensuite mettre à jour ces enfants (§ 3.4.2.4).
  - e) Lancer une étape de l'algorithme HCM sur la partie 'centre des classes' codée dans les enfants et ensuite mettre à jour ces enfants (§ 3.4.2.4).
  - f) Calcul de la fonction d'adaptation (3.22) pour chaque enfant et ensuite mise à jour de la *population* ( $P/2$  parents +  $P/2$  enfants).
- 4) Déterminer le sous-ensemble de paramètres ainsi que le nombre de classes structurant l'ensemble d'apprentissage (et les centres de gravité de ces classes) qui correspondent



à la meilleure fonction d'adaptation de toutes les sous-populations finales.

La dernière étape, décrite dans les parties suivantes, consiste à définir une règle de décision qui permettra de classer ou non toute nouvelle observation recueillie sur le processus.

### 3.5. Détermination d'une méthode de décision

Les deux premières étapes de la *phase d'analyse* ont permis de structurer l'ensemble d'apprentissage en classes distinctes et de pouvoir étiqueter les données de telle sorte que chacune des observations appartient à une des  $M$  classes connues.

Soit  $X_a = (\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  l'ensemble d'apprentissage initial et  $(\Omega_1, \Omega_2, \dots, \Omega_M)$  l'ensemble des  $M$  classes connues. Le problème posé est alors d'affecter une nouvelle observation  $\underline{X}_\mu$  à l'une des  $M$  classes à l'aide d'une règle de décision.

Si on souhaite réaliser une présentation non exhaustive des différentes approches de définition d'une règle de décision, il convient tout de même de distinguer les méthodes statistiques, paramétriques ou non paramétriques. Les méthodes paramétriques supposent la connaissance des lois de probabilité des classes et des observations. En effet, tout vecteur  $\underline{X}$  suit, dans une classe donnée  $\Omega_c$ , une loi de probabilité  $f(\underline{X}|\Omega_c)$ . Par ailleurs, les classes  $\Omega_c$  ( $c = 1, \dots, M$ ) sont caractérisées par leurs probabilités a priori  $\Pr(\Omega_c)$ . Ces probabilités vérifient la condition d'orthogonalité suivante :

$$\sum_{c=1}^M \Pr(\Omega_c) = 1 \quad (3.28)$$

Parmi ces méthodes, l'approche la plus commune est la théorie Bayésienne de la décision [Dubuisson, 90].

En pratique, les données sur les lois de probabilités des classes  $\Pr(\Omega_c)$  et des observations  $f(\underline{X}|\Omega_c)$  sont rarement disponibles dans un problème de diagnostic, ce qui conduit alors à utiliser les méthodes non paramétriques.

#### 3.5.1. Méthodes non paramétriques

Les méthodes précédentes supposent les connaissances totales des lois de probabilités pour les classes et les observations et donc, ne peuvent s'appliquer de façon réaliste que sous

l'hypothèse gaussienne. Toutefois si cette connaissance est incomplète, il est nécessaire de s'orienter vers une estimation de ces lois de probabilité.

Parmi les estimateurs usuels, on distingue l'estimateur des  $k$  – plus proches voisins ( $k - ppv$ ), qui est une approche facile d'utilisation.

### 3.5.1.1. Règle des $k$ – plus proches voisins ( $k - ppv$ )

Soit  $X_a = (\underline{X}_1, \underline{X}_2, \dots, \underline{X}_N)$  l'ensemble d'apprentissage composé de  $N$  vecteurs indépendants, chacun étiqueté à l'une  $M$  classes connues. Le principe de la règle des  $k - ppv$  est d'affecter une nouvelle observation  $\underline{X}_u$  à la classe majoritairement représentée parmi ses  $k$  - plus proches voisins.

En réalité, La manière la plus simple d'élaborer cette règle, est de mesurer la distance entre la nouvelle observation et chacun des vecteurs de l'ensemble d'apprentissage. La nouvelle observation sera alors affectée à la classe majoritairement représentée parmi ses  $k - ppv$ .

La définition de la notion de plus proche voisin est liée au choix d'une distance.

#### 3.5.1.1.1. Choix de la distance dans la règle des $k - ppv$

Différentes distances peuvent être utilisées, définies par la formulation générale :

$$d^2(\underline{X}, \underline{Y}) = (\underline{X} - \underline{Y})^T A (\underline{X} - \underline{Y}) \quad (3.29)$$

avec :

$A$  est une matrice définie positive.

$d(\underline{X}, \underline{Y})$  : distance entre  $\underline{X}$  et  $\underline{Y}$

$\underline{X}$  et  $\underline{Y}$  : vecteur de dimension  $d'$

Plusieurs cas particuliers peuvent être tirés de cette formulation :

- La distance *euclidienne* où  $A$  est une matrice identité :

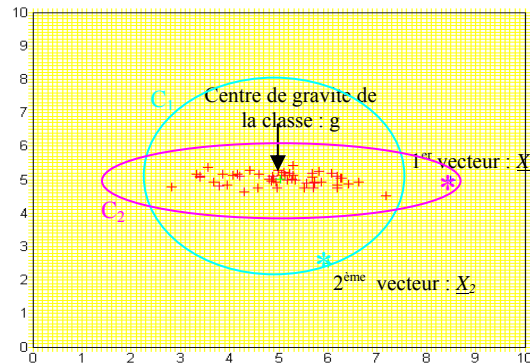
$$d_E^2(\underline{X}, \underline{Y}) = (\underline{X} - \underline{Y})^T \cdot (\underline{X} - \underline{Y}) \quad (3.30)$$

- La distance de *Mahanalobis* où  $A$  est l'inverse de la matrice variance-covariance totale  $\Sigma$  :

$$d_M^2(\underline{X}, \underline{Y}) = (\underline{X} - \underline{Y})^T \Sigma^{-1} (\underline{X} - \underline{Y}) \quad (3.31)$$

L'introduction de la matrice de variance-covariance dans l'expression 3.31 signifie que cette mesure prend en compte la dispersion de l'ensemble d'apprentissage, dans chaque dimension. Pour illustrer cette différence, nous avons simulé une classe (figure 3.12), dans

une représentation en deux dimensions, avec une forte variance sur un axe (axe des « x ») et une faible sur l'autre (axe des « y »).



**Fig. 3. 12.** Visualisation de la classe  $\Omega$  et de son centre de gravité  $g$  ainsi que  $X_1$  et  $X_2$

A partir de deux vecteurs  $\underline{X}_1$  et  $\underline{X}_2$ , la distance aussi bien *euclidienne* que de *Mahanalobis* entre le centre de gravité  $g$  de la classe  $\Omega$  et ces deux vecteurs peut être définie, le but étant de savoir quel est le vecteur le plus apte à être affecté à cette classe.

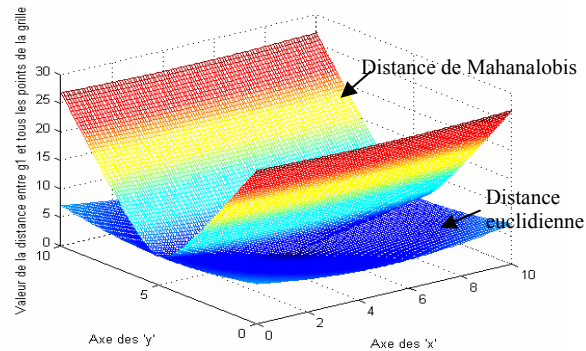
	Distance euclidienne ( $d_E$ )	Distance de Mahanalobis ( $d_M$ )
$d(g, \underline{X}_1)$	3.4607	3.5467
$d(g, \underline{X}_2)$	2.6914	13.3411

**Tableau. 3. 5.** Résultat des mesures de distance

L'affectation d'un vecteur à la classe  $\Omega$  présente deux cas différents. Avec la distance *euclidienne*, le vecteur  $\underline{X}_2$  est affecté à  $\Omega$  bien que la différence entre  $d(g, \underline{X}_1)$  et  $d(g, \underline{X}_2)$  ne soit pas trop importante. Or, la différence est parfaitement établie avec le calcul de la distance de *Mahanalobis*. En effet, ici le vecteur  $\underline{X}_1$  est affecté à  $\Omega$ , ce qui au vu de la figure 3.12 paraît plus judicieux.

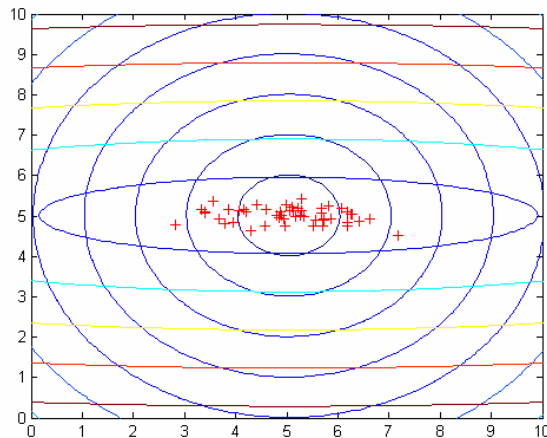
Une méthode plus explicite consiste à visualiser la différence entre les deux distances. Pour cela, une grille est définie dans un espace de représentation à deux dimensions (plan  $Oxy$ ) constituée de 10 000 points comme le montre les « + » sur la figure 3.12. Les axes des abscisses et des ordonnées sont gradués de 0 à 10 en pas de 0.1. Chaque pas représente un point. Calculons la distance *euclidienne*, puis de *Mahanalobis*, entre le centre de gravité  $g$  de la classe  $\Omega$  et tous les points de cette grille. La figure 3.13 représente les distances

*euclidiennes* d'une part et de *Mahanalobis* d'autre part entre le centre de gravité  $g$  de la classe  $\Omega$  et les points de la grille.



**Fig. 3. 13.** Visualisation de la distance euclidienne et de Mahanalobis entre  $g$  et tous les points de la grille

La différence entre les deux distances est bien nette. La distance de *Mahanalobis* tient compte de la variance sur chaque axe de la classe  $\Omega$ , ce que ne réalise pas la distance *euclidienne*. La distance *euclidienne* est plus appropriée lorsque les classes qui ont été définies sont de formes hypersphériques (même variance sur chaque axe). Il est possible de tracer les contours de l'axe 'z' en les projetant sur le plan  $Oxy$ . Ceci est représenté sur la figure 3.14 :



**Fig. 3. 14.** Visualisation des contours de l'axe 'z' projeté sur le plan  $Oxy$

Sur cette figure, les cercles représentent les contours de la distance *euclidienne* entre  $g$  et les points de la grille. On peut noter que la distance est uniforme dans les deux axes 'x' et 'y'.

Les ellipses représentent les contours de la distance de *Mahanalobis* entre  $g$  et les points de la grille. On remarque ici que la distance n'est plus uniforme dans les deux axes 'x' et 'y' du

fait que la distance de *Mahanalobis* fait appel à la matrice variance-covariance qui contient les variances de chaque axe.

De ce fait, la distance de *Mahanalobis* sera retenue par la suite, car elle tient compte de la dispersion des classes.

### 3.5.1.1.2. Choix du nombre de plus proches voisins $k$

Le choix de  $k$  est un choix délicat souvent effectué un peu au hasard. Deux compromis sont à satisfaire cependant : prendre  $k$  grand pour obtenir une estimation « fiable », et être certain que les  $k$  voisins sont proches de la nouvelle observation pour s'assurer de ne pas effectuer un diagnostic erroné. Le choix reste souvent arbitraire,  $k$  est pris relativement faible par rapport à la taille de l'ensemble d'apprentissage  $N$ .  $k$  est toujours ramené à un nombre impair, pour éviter l'ambiguïté d'indécision.  $k$  égal à  $\sqrt{N}$  [Dubuisson, 90] réalise un bon compromis.

### 3.5.1.2. Règle des $k$ – plus proches voisins avec rejet

Comme évoqué précédemment, le cas où la connaissance sur les classes est exhaustive n'est pas usuel en diagnostic. De ce fait, de nouvelles classes peuvent apparaître au cours du temps, et, dans ce cas, l'affectation systématique d'une nouvelle observation  $\underline{X}_u$  à une classe connue peut conduire à une erreur de diagnostic. Il vaut mieux reporter la décision et affecter  $\underline{X}_u$  à une classe dite de rejet. On parlera alors d'utilisation de règle de décision avec rejet [Chow, 57].

En réalité, il existe deux grandes notions de rejet : le rejet en ambiguïté qui concerne une nouvelle observation située entre deux ou plusieurs classes de  $X_a$  et le rejet en distance qui correspond à une nouvelle observation située à une distance éloignée des classes de  $X_a$ .

Ces rejets ont été introduits dans la règle des  $k$  –  $ppv$  et plus généralement, dans les méthodes de diagnostic afin de diminuer les erreurs de classification et, par conséquent, les erreurs de diagnostic.

Globalement, la règle de décision incluant les deux options de rejets sera appliquée pour ( $M + 2$ ) classes :

- $\underline{X}_u \rightarrow \Omega_c$  ( $c = 1, M$ ) :  $\underline{X}_u$  est classé dans  $\Omega_c$
- $\underline{X}_u$  est rejeté en ambiguïté alors :  $\underline{X}_u \rightarrow \Omega_0$  (3.32)
- $\underline{X}_u$  est rejeté en distance alors :  $\underline{X}_u \rightarrow \Omega_d$

avec :

$\Omega_d$  : classe fictive des observations rejetées en distance

$\Omega_0$  : classe fictive des observations rejetées en ambiguïté

### 3.5.1.2.1. Rejet d'ambiguïté

Lorsque deux états de fonctionnement du système sont représentés par deux classes proches dans  $\mathcal{R}^d$  et que l'observation  $\underline{X}_u$  du système à l'instant  $t$  se trouve à la limite de la frontière de décision entre ces deux classes (figure 3.15), il semble intuitivement plus sûr de ne pas prendre de décision sur l'affectation de  $\underline{X}_u$ . Par contre le classifieur indiquera à l'utilisateur qu'il y a ambiguïté sur la décision.

Appliqué à la règle des  $k - ppv$ , le rejet d'ambiguïté se définit ainsi [Hellman, 70], [Tomek, 76] :

- $\underline{X}_u$  est associé à la classe  $\Omega_c$  si  $k'$  au moins parmi ses  $k$  plus proches voisins appartiennent à  $\Omega_c$

$$\begin{cases} \underline{X}_u \rightarrow \Omega_o & \text{si } k_c = \max_{r=1,M} k_r < k' \\ \underline{X}_u \rightarrow \Omega_c & \text{si } k_c = \max_{r=1,M} k_r \geq k' \end{cases} \quad (3.33)$$

avec  $k$  le nombre de plus proches voisins du vecteur  $\underline{X}_u$ ,

$k_c$  le nombre de plus proches voisins de  $\underline{X}_u$  (parmi  $k$ ) dans la classe  $\Omega_c$ .

Les observations rejetées en ambiguïté seront affectées à une nouvelle classe fictive  $\Omega_0$ , appelée classe de rejet en ambiguïté.

Plus  $k'$  est proche de  $k$ , plus le rejet en ambiguïté est important. La règle de rejet d'ambiguïté a été par la suite généralisée par [Devijver, 77] en fixant un nombre minimal de voisin propre à chacune des classes.

D'autre part si  $k' = 0$ , alors le rejet n'est pas appliqué.

### 3.5.1.2.2. Rejet de distance

Seule, la notion de rejet en ambiguïté est insuffisante. Il faut ajouter la possibilité de décision sur des observations résultant de l'apparition de nouvelles classes. Ces nouvelles classes correspondent généralement à des zones de l'espace inexplorées, au sens où aucune observation de  $X_u$  n'y a été rencontrée. Dans ce cas, une deuxième possibilité de rejet a été utilisée, appelée rejet en distance [Dubuisson, 90a].

Ce rejet en distance est basé sur la notion de distance aux classes, qui doit être 'grande'.

L'idée la plus simple consiste à effectuer un seuillage de distance. L'observation  $\underline{X}_u$  est rejetée si la distance à ses  $k - ppv$  est supérieure à un seuil  $T$ .

Le seuil de distance est défini par l'utilisateur. Il est comparé par exemple :

- à la distance moyenne de  $\underline{X}_u$  à ses  $k$  plus proches voisins  $\underline{y}_h$  ( $h=1, k$ ) :

$$\underline{X}_u \rightarrow \Omega_d \text{ si } \frac{1}{k} \sum_{h=1}^k d(\underline{X}_u, \underline{y}_h) > T \quad (3.34)$$

- ou à la distance entre  $\underline{X}_u$  et le centre de gravité de sa classe d'affectation :

$$\underline{X}_u \rightarrow \Omega_d \text{ si } d(\underline{X}_u, \underline{m}_c) > T \quad (3.35)$$

$\underline{m}_c$  étant le centre de gravité de la classe d'affectation de  $\underline{X}_u$

Dans la suite de l'étude, le seuil de distance sera comparé à la distance entre l'observation  $\underline{X}_u$  et le centre de gravité de sa classe d'affectation défini par l'équation 3.35.

Le seuil  $T$  peut dépendre de la classe, donc tenir compte de sa géométrie et être lié au diamètre de la classe où a des considérations statistiques.

Ainsi une première solution peut être d'assimiler le seuil de rejet en distance  $T_c$  de la classe  $\Omega_c$  au diamètre de cette classe. Pour cela, il est nécessaire de calculer les distances entre les vecteurs de celle-ci et son centre de gravité  $\underline{m}_c$  :

$$T_c = \beta \cdot \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_c)] \text{ avec } \beta = 2 \text{ par exemple} \quad (3.36)$$

Avec :

$N_c$ , le nombre de vecteurs de la classe  $\Omega_c$

$\underline{X}_{cv}$ , le  $v^{\text{ième}}$  vecteur de la classe  $\Omega_c$

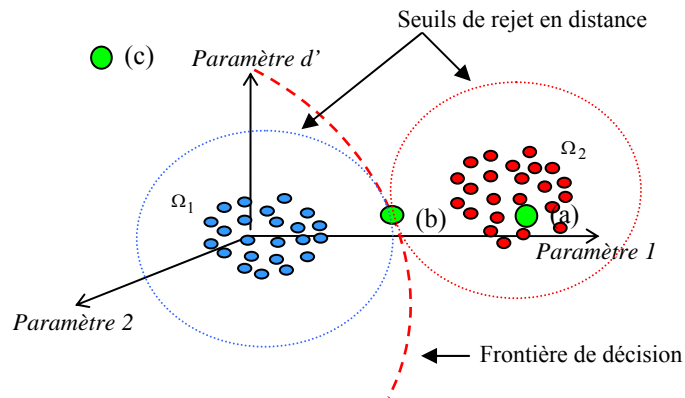
La marge de sécurité (facteur  $\beta$ ) permet de prendre en compte les incertitudes sur la dimension exacte de la classe mais aussi sur le bruit lors de l'acquisition des mesures.

La deuxième solution consiste à définir les caractéristiques statistiques réelles des classes pour ensuite s'en servir lors de la détermination du seuil. Cette solution sera développée dans le chapitre IV.

Finalement, la règle des  $k - ppv$  incluant les deux options de rejet peut s'exprimer par :

$$\begin{cases} \underline{X}_u \rightarrow \Omega_d \text{ si } d(\underline{X}_u, \underline{m}_c) > T_c \\ \underline{X}_u \rightarrow \Omega_o \text{ si } d(\underline{X}_u, \underline{m}_c) \leq T_c \text{ et } k_c = \max_{r=1, M} k_r < k' \\ \underline{X}_u \rightarrow \Omega_c \text{ si } d(\underline{X}_u, \underline{m}_c) \leq T_c \text{ et } k_c = \max_{r=1, M} k_r \geq k' \end{cases} \quad (3.37)$$

Sur la figure 3.15 est montrée un exemple des différentes classifications possibles réalisées par la règle de décision associée aux options de rejet.



**Fig. 3. 15.** Phase de décision par la règle des  $k - ppv$ . Affectation (a), rejet en ambiguïté (b), rejet en distance (c)

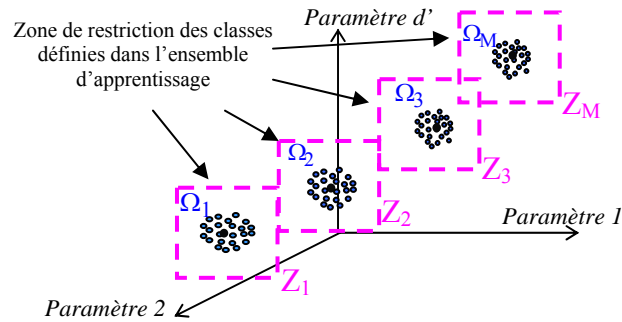
### 3.5.1.3. Réduction des temps de calcul

Un des inconvénients de la méthode des  $k - ppv$  est qu'elle nécessite le calcul des distances de l'observation  $\underline{X}_u$  à tous les éléments de l'ensemble d'apprentissage et le stockage avant classement. Cette contrainte peut être pénalisante en terme de temps de calcul si l'ensemble d'apprentissage initial est de grande taille.

Il est donc préférable de chercher à accélérer cette étape. Dans ce but, de nombreuses études ont été effectuées, d'une part, par la recherche rapide des distances en appliquant des calculs de distance plus courte (distance du max, distance de Manhattan) que la distance euclidienne classique [Kittler, 78], [Richetin, 80], [Friedman, 75] ou d'autre part, par la réduction de la taille de  $X_a$ , ce qui entraîne une diminution des distances à calculer [Wilson, 72].

En effet, il semble intéressant de ne calculer que les distances entre une nouvelle observation  $\underline{X}_u$  et les vecteurs des classes proches physiquement. Pour cela, un maillage de l'espace de représentation peut être effectué. Une grande dimension de l'espace de représentation est susceptible de rendre difficile la construction de ce maillage. Pour cette raison le maillage est effectué seulement dans les régions où les classes de  $X_a$  sont localisées. Ainsi une zone de restriction est définie autour de ces classes (figure 3.16).





**Fig. 3. 16.** Représentation schématique des zones de restriction pour les différentes classes de  $X_a$

Ces zones forment donc des sortes de frontières autour de ces classes. Le choix de la dimension des zones est fait en tenant compte de la variance des classes définies dans  $X_a$ . La variance de chaque classe sur chaque  $d'$  dimension est calculée de la façon suivante :

$$Seuil(j)_c = 2 * \sigma_c^2(j) \tag{3.38}$$

$\begin{matrix} j=1,d' \\ c=1,M \end{matrix}$ 
 $\begin{matrix} j=1,d' \\ c=1,M \end{matrix}$

avec :

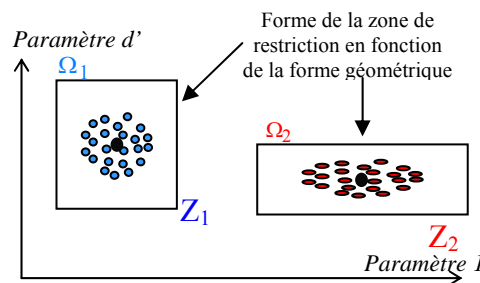
Seuil(j)<sub>c</sub> : frontière sur l'axe 'j' de la classe 'Ω<sub>c</sub>'

σ<sub>c</sub><sup>2</sup>(j) : variance de la classe 'Ω<sub>c</sub>' sur l'axe 'j'

M : nombre de classes de l'ensemble d'apprentissage

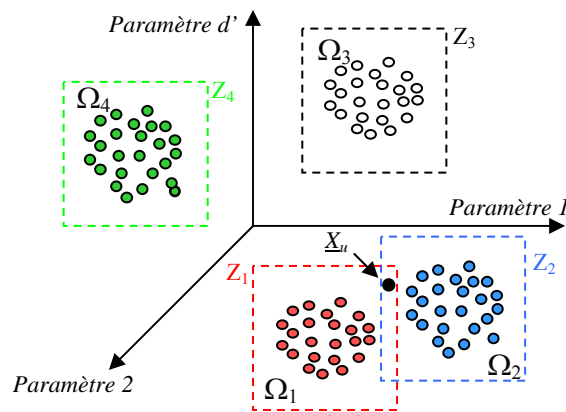
$d'$  : dimension de l'espace de représentation

L'utilisation de la variance, afin de définir les frontières des zones de restriction permet de prendre en compte la dispersion des classes en fonction des différents axes. Sur la figure 3.17, la zone de restriction  $Z_c$  définit une sorte d'enveloppe autour de la classe  $\Omega_c$  ( $c = 1, M$ ).



**Fig. 3. 17.** Forme de la zone de restriction en fonction de la forme géométrique des classes de  $X_a$

Le maillage de l'espace de représentation est une étape préalable à l'opération de discrimination. En effet, grâce à ce maillage des classes, il est possible de rechercher quelles sont les classes proches d'une nouvelle observation. À présent, il est juste nécessaire de vérifier si les coordonnées de cette nouvelle observation sont dans les zones définies autour des classes. Prenons l'exemple de la figure 3.18, où l'observation  $\underline{X}_u$  est située dans les zones  $Z_1$  et  $Z_2$  des classes  $\Omega_1$  et  $\Omega_2$  respectivement. La règle des  $k - ppv$  est appliquée en calculant les distances entre  $\underline{X}_u$  et les vecteurs des classes  $\Omega_1$  et  $\Omega_2$ . Les distances avec les vecteurs des classes  $\Omega_3$  et  $\Omega_4$  ne sont pas calculées. De ce fait, le temps de calcul est quasiment diminué de moitié.



**Fig. 3. 18.** Exemple de localisation d'une observation dans le maillage de l'espace de représentation

Cette approche par maillage pourrait même être utilisée comme une règle de décision à part entière (figure 3.19). Alors les règles liées à l'affectation se définiraient de la façon suivante :

- Si une nouvelle observation  $\underline{X}_u$  est localisée dans la zone  $Z_c$  alors elle appartient à la classe  $\Omega_c$ . Si  $\underline{X}_u$  est localisée dans plusieurs zones, alors elle est rejetée en ambiguïté. D'autre part si  $\underline{X}_u$  n'appartient à aucune des zones alors elle est rejetée en distance.

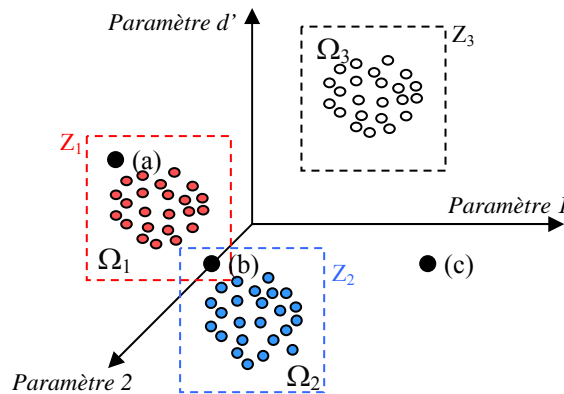
En résumé :

$$\left\{ \begin{array}{l} \text{Si } \underline{X}_u \in \text{à la zone } Z_c \quad \rightarrow \text{ alors } \underline{X}_u \xrightarrow{\text{est affectée}} \Omega_c \quad c = 1, M \\ \text{Si } \underline{X}_u \in \text{à aucune des zones } Z_c \quad \rightarrow \text{ alors } \underline{X}_u \xrightarrow{\text{est rejetée}} \Omega_d \\ \text{Si } \underline{X}_u \in \text{à plusieurs zones } Z_c \quad \rightarrow \text{ alors } \underline{X}_u \xrightarrow{\text{est rejetée}} \Omega_o \end{array} \right. \quad (3.39)$$

$\Omega_o$  classe fictive des observations rejetées en ambiguïté

$\Omega_d$  classe fictive des observations rejetées en distance

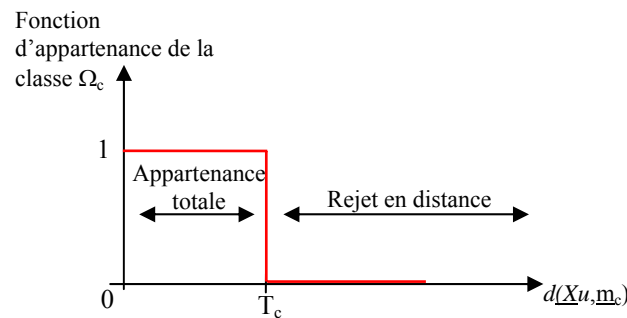
A l'inverse de la règle des  $k - ppv$ , aucun calcul de distance n'aurait besoin d'être réalisé puisque seul les coordonnées de l'observation seraient prises en compte. Le temps de calcul en serait fortement diminué. Mais la classification faite par ce simple maillage pourrait perdre en précision en approximant les frontières de décision.



**Fig. 3. 19.** Phase de décision par maillage de l'espace de représentation. Affectation (a), rejet en ambiguïté (b), rejet en distance (c)

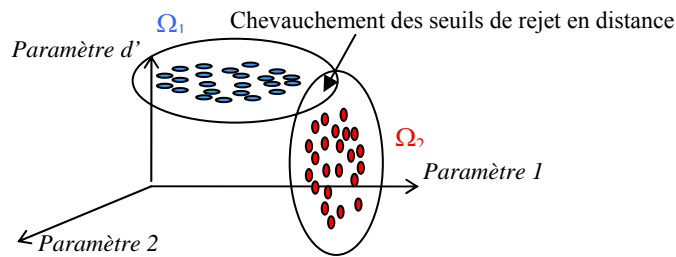
### 3.5.1.4. Classification non exclusive

La règle des  $k - ppv$  classique présentée précédemment affecte donc une nouvelle observation à la classe majoritairement représentée parmi ses plus proches voisins. Ceci peut se résumer de la façon suivante :



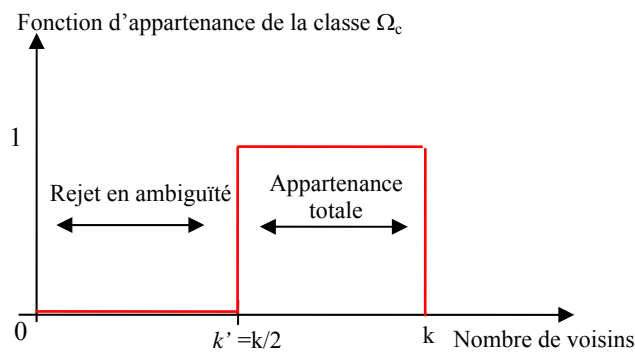
**Fig. 3. 20.** Classification en fonction de la distance par la règle des  $k - ppv$  classique

Il se peut que deux ou plusieurs classes contenues dans  $X_a$  ne soient pas suffisamment éloignées du fait d'un mode de fonctionnement peu différent. Leurs seuils de rejet en distance peuvent donc s'entrecroiser comme le montre la figure 3.21.



**Fig. 3. 21.** *Seuil de rejet en distance de classes peu éloignées*

Avec la règle des  $k - ppv$ , une nouvelle observation  $\underline{X}_u$  située dans la région d'appartenance des deux classes (figure 3.21) sera affectée à la classe majoritairement représentée parmi ses plus proches voisins. Ceci peut se résumer de la façon suivante :



**Fig. 3. 22.** *Classification en fonction du nombre de voisins par la règle des  $k - ppv$  classique*

Le fait d'utiliser une classification exclusive renseigne sur l'appartenance d'une observation à une classe, mais ne donne aucune indication sur sa localisation dans la zone de la classe concernée. En effet, une observation située au centre d'une classe et une observation située en bordure (proche du seuil limite de rejet en distance) de celle-ci n'ont pas la même représentativité, et pourtant, ces deux observations sont affectées de la même manière à cette classe. De même, une observation située en bordure de classes contiguës, ou entre deux classes disjointes, peut signifier un état transitoire entre ces deux classes. Quantifier cette information permet de donner la probabilité d'appartenance à la classe retenue et, de ce fait, fournit un degré de confiance dans la classification réalisée. Cela permet aussi de suivre l'évolution d'observations successives à l'intérieur d'une classe, c'est-à-dire de donner une mesure de l'évolution du système vis-à-vis des états connus.

Pour ces raisons, la mise en place d'une règle de décision non exclusive peut être intéressante. Nous disposerons de plus d'informations sur l'appartenance d'une observation, ce qui permettra d'une part, de donner un diagnostic plus en adéquation avec la réalité d'évolution du processus et d'autre part, d'associer un degré de confiance à la classification.

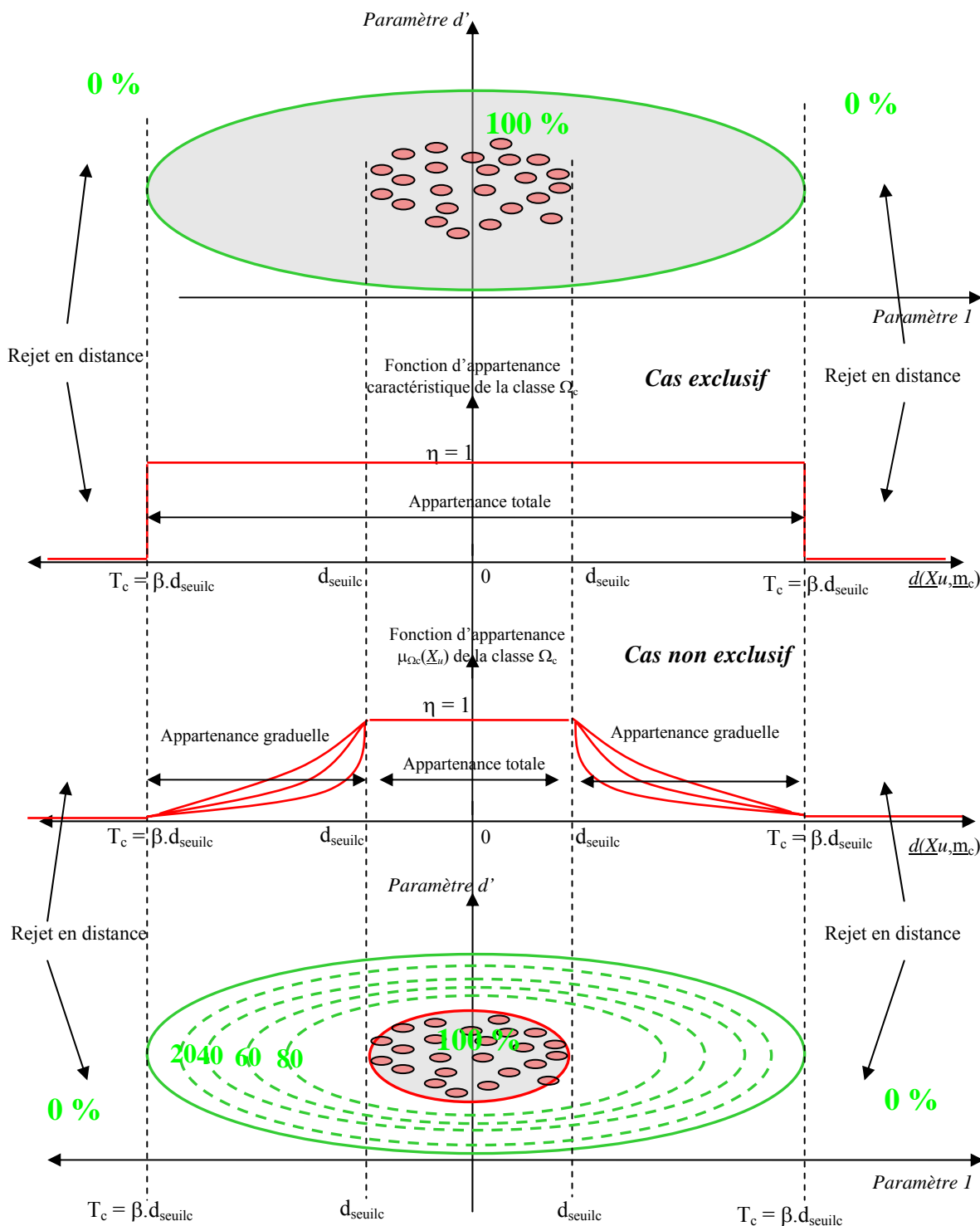
Le passage de l'affectation exclusive à l'affectation non exclusive doit se caractériser par une gradation entre la non appartenance absolue et l'appartenance totale. La variation brutale observée sur la figure 3.20 en  $T_c$  et sur la figure 3.22 en  $k'$  doit être remplacée respectivement par une descente et une montée plus 'douce'. Ainsi, l'appartenance à une classe pourra être modulée et les notions de proximité ou d'éloignement quantifiées.

#### 3.5.1.4.1. Evaluation du degré d'appartenance d'une observation aux classes définies dans $X_u$

La ressemblance d'un objet avec une forme référence étant souvent relative, en général, une fonction d'appartenance caractérisant le degré de compatibilité entre l'observation  $\underline{X}_u$  et la classe  $\Omega_c$  ( $c=1$  à  $M$ ) est introduite. Cette fonction est notée  $\mu_{\Omega}(\underline{X}_u)$ . Par rapport à la théorie des ensembles classiques où la fonction d'appartenance, appelée fonction caractéristique, ne peut prendre que des valeurs 0 ou 1,  $\mu_{\Omega}(\underline{X}_u)$  introduit la notion d'appartenance graduelle puisque c'est une fonction continue à valeurs généralement comprises dans l'intervalle  $[0,1]$ . Elle permet de quantifier le degré de compatibilité entre une observation  $\underline{X}_u$  et une classe  $\Omega_c$  de l'ensemble d'apprentissage de telle sorte que plus  $\underline{X}_u$  ressemble (ne ressemble pas) à la forme référence représentée par la classe  $\Omega_c$ , plus  $\mu_{\Omega_c}(\underline{X}_u)$  tend vers la valeur 1 (0). Cet aspect graduel de la fonction d'appartenance et non plus abrupt, permet de nuancer le degré de ressemblance d'une observation appartenant à une classe. La notion de ressemblance suggère de définir  $\mu_{\Omega_c}$  en termes de mesure de similarité ou de dissimilarité entre  $\underline{X}_u$  et  $\Omega_c$ . Comme en RdF, deux objets sont similaires s'ils sont proches, une distance est généralement choisie comme mesure de dissimilarité, et  $\mu_{\Omega_c}$  peut être définie comme une fonction monotone décroissante de  $d(\underline{X}_u, \underline{m}_c)$ .

Il n'existe pas de méthode générale permettant de construire la fonction d'appartenance associée à un ensemble donné. La forme de cette fonction est étroitement liée au contexte.

Pour notre étude, il est nécessaire de déterminer une fonction qui applique une appartenance totale lorsque la nouvelle observation  $\underline{X}_u$  se situe entre le centre de gravité (cdg)  $\underline{m}_c$  de la classe  $\Omega_c$  et le vecteur le plus éloigné de ce centre mais aussi une décroissance plus ou moins douce lorsque  $\underline{X}_u$  se situe entre le vecteur le plus éloigné et le seuil limite de rejet en distance  $T_c$  (figure 3.23).



**Fig. 3. 23.** Représentation schématique de la fonction d'appartenance dans le cas exclusive et non exclusive

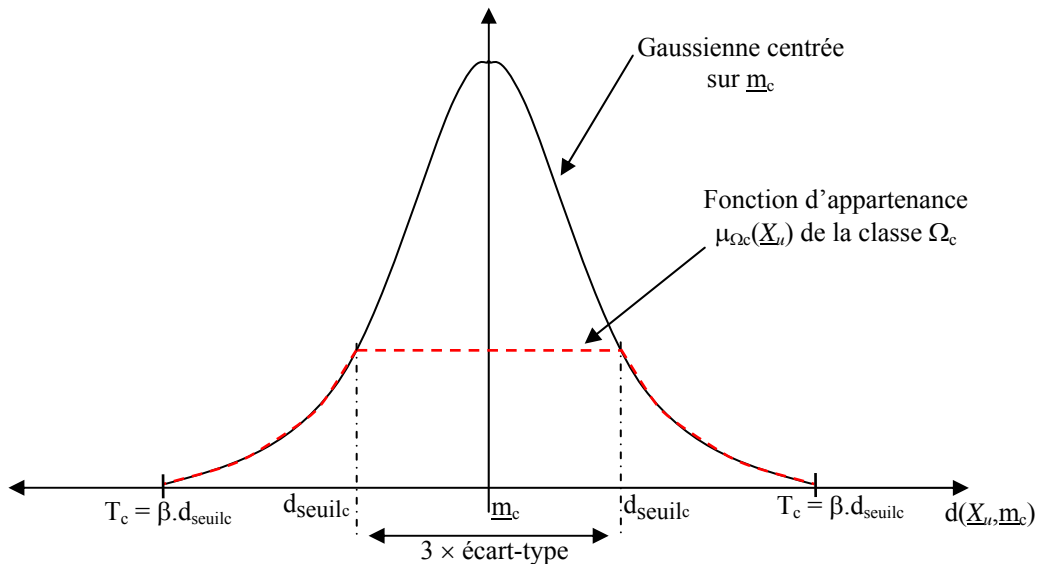
Pour répondre à ces contraintes, plusieurs solutions sont envisageables. Une des solutions possibles, qui n'a pas été utilisée par la suite, se base sur l'utilisation d'une fonction gaussienne centrée sur le centre de gravité de la classe. L'expression de cette fonction pour notre problème prendrait la forme suivante :

$$f_c(\underline{X}_u) = \frac{1}{\sigma_c \cdot \sqrt{2\pi}} \cdot e^{-d\left(\frac{\underline{X}_u, \underline{m}_c}{2\sigma_c}\right)^2} \quad (3.40)$$

Avec  $\sigma_c$ , l'écart type de la classe  $\Omega_c$ .

En supposant que les classes soient caractérisées par une distribution gaussienne (ce qui est généralement le cas), il est possible de fixer  $d_{seuil}$ , correspondant à une appartenance totale, à trois fois l'écart type (ce qui englobe 99% des observations de la classe) et ensuite, de fixer le seuil  $T_c$  pour obtenir une appartenance décroissante jusqu'au rejet en distance. Ceci est représenté sur la figure 3.24. Ainsi, la fonction d'appartenance de la classe  $\Omega_c$  pourrait se définir comme suit :

$$\begin{aligned} \text{Si } d(\underline{X}_u, \underline{m}_c) \leq 3 \times \sigma & \quad \text{alors } \mu_{\Omega_c}(\underline{X}_u) = 1 \\ \text{Si } 3 \times \sigma \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{seuil} & \quad \text{alors } \mu_{\Omega_c}(\underline{X}_u) = f_c(\underline{X}_u) \\ 0 & \quad \text{sinon} \end{aligned} \quad (3.41)$$



**Fig. 3. 24.** Fonction d'appartenance définie à l'aide d'une fonction gaussienne centrée sur  $\underline{m}_c$ .

La solution qui a été retenue par la suite définit une fonction d'appartenance à l'aide d'une fonction  $\pi$  multidimensionnelle. Cette fonction d'appartenance est décrite par la formulation suivante (3.42) :

$$\mu_{\Omega_c}(\underline{X}_u) = \begin{cases} 1 - (1 - \eta) \left( \frac{d(\underline{X}_u, \underline{m}_c)}{d_{seuil}} \right)^g & \text{si } 0 \leq d(\underline{X}_u, \underline{m}_c) \leq d_{seuil} \\ \eta \left( \frac{\beta}{\beta - 1} \right)^g \left( \frac{\beta \cdot d_{seuil} - d(\underline{X}_u, \underline{m}_c)}{\beta \cdot d_{seuil}} \right)^g & \text{si } d_{seuil} \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{seuil} \\ 0 & \text{sinon} \end{cases} \quad (3.42)$$

avec :

$d(\underline{X}_u, \underline{m}_c)$  est la distance entre une observation  $\underline{X}_u$  et le centre de gravité  $\underline{m}_c$  de la classe  $\Omega_c$ .

$d_{seuil_c}$  est le paramètre qui définit la position des points d'inflexion. Il correspond à la distance maximale entre le centre de gravité de la classe  $\Omega_c$  et le point le plus éloigné de cette même classe  $d_{seuil_c} = \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_c)]$

$$\mu_{\Omega_c}(\underline{X}_u) = \eta \text{ si } d(\underline{X}_u, \underline{m}_c) = d_{seuil_c}$$

$\eta$  est le paramètre servant à fixer le plateau de la fonction  $\pi$ .

$\beta$  est un paramètre ( $\beta \in \mathfrak{R}^{+*} > 1$ ) qui permet de définir la largeur de bande de la fonction  $\pi$  (les points d'arrêt de la fonction  $\pi$ ) (§ 3.5.1.2.2).

$g$  est un paramètre permettant de faire varier la pente de la fonction.

Mais cette mesure de similarité, liée à la seule distance, ne rend pas forcément optimale la classification. En effet il serait intéressant, pour une nouvelle observation, de prendre en compte non pas seulement la distance avec les différents centre de gravité de chaque classe mais aussi d'avantager les classes qui contiennent le plus grand nombre de plus proches voisins.

Par exemple, si on suppose que les  $k$  plus proches voisins sont tous issus d'une même classe, alors le degré déterminé par la fonction d'appartenance doit être égal au maximum à 1 si  $d(\underline{X}_u, \underline{m}_c) \leq d_{seuil_c}$  et compris entre  $[0,1]$  si  $d_{seuil_c} \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{seuil_c}$ . Les équations 3.42 deviennent alors :

$$\mu_{\Omega_c}(\underline{X}_u) = \begin{cases} \left[ \frac{1}{N_c} - \left( \frac{1}{N_c} - \eta \right) \left( \frac{d(\underline{X}_u, \underline{m}_c)}{d_{seuil_c}} \right)^g \right] \cdot nb\_kppv_c & \text{si } 0 \leq d(\underline{X}_u, \underline{m}_c) \leq d_{seuil_c} \\ \left[ \eta \left( \frac{\beta}{\beta-1} \right)^g \left( \frac{\beta \cdot d_{seuil_c} - d(\underline{X}_u, \underline{m}_c)}{\beta \cdot d_{seuil_c}} \right)^g \right] \cdot nb\_kppv_c & \text{si } d_{seuil_c} \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{seuil_c} \\ 0 & \text{sinon} \end{cases} \quad (3.43)$$

avec :

$N_c$  est le nombre de points contenus dans la classe  $\Omega_c$

$nb\_kppv_c$  est le nombre de  $k$  plus proches voisins issus de la classe  $\Omega_c$

La figure 3.25 représente l'évolution de la fonction d'appartenance  $\mu_{\Omega_c}(\underline{X}_u)$  en fonction de la distance entre  $\underline{X}_u$  et le centre de gravité de cette classe et le nombre de voisins parmi les  $k$



plus proches voisins contenus dans cette classe. Cette figure est obtenue avec  $g = 1, \beta = 2$  et on considère que la classe est formée de 10 échantillons ( $N_c = 10$ ). Si on choisit  $\eta = 1/N_c$ , cela impose de prendre  $k = N_c = 10$ , pour que chaque échantillon de la classe apporte sa contribution dans le pourcentage (degré) d'appartenance. Par exemple, si  $d(\underline{X}_u, \underline{m}_c) \leq d_{seuil_c}$ , alors l'équation 3.43 devient

$$\mu_{\Omega_c}(\underline{X}_u) = \left[ \frac{1}{N_c} \right] \cdot nb\_kppv_c \tag{3.44}$$

De ce fait, le degré d'appartenance d'une nouvelle observation  $\underline{X}_u$  à la classe  $\Omega_c$  est dépendant du nombre de plus proches voisins de cette classe. Si  $nb\_kppv_c = 1$  alors  $\mu_{\Omega_c}(\underline{X}_u) = 0.1$  ou encore si  $nb\_kppv_c = 7$  alors  $\mu_{\Omega_c}(\underline{X}_u) = 0.7$ .

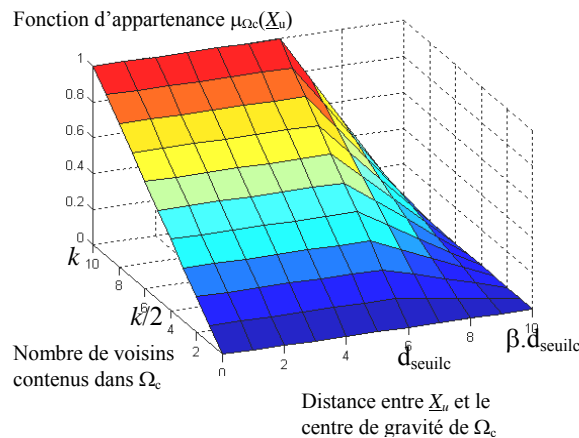


Fig. 3. 25. Fonction d'appartenance de  $\underline{X}_u$  en fonction de la distance et des plus proches voisins

**3.5.1.4.2. Extension de la règle des  $k - ppv$  par association d'une fonction d'appartenance non exclusive**

Le problème de la décision doit être considéré en terme d'appartenance non exclusive aux classes. Les valeurs des fonctions d'appartenance  $\mu_{\Omega_c}(\underline{X}_u)$  relatives aux différentes classes  $\Omega_c$  ( $c=1, M$ ) peuvent être utilisées à ces fins.

La solution intuitive consiste à décider d'associer  $\underline{X}_u$  à la classe  $\Omega_c$  pour laquelle la valeur de  $\mu_{\Omega_c}(\underline{X}_u)$  est maximale. Cette règle, appelée règle du maximum d'appartenance, peut s'écrire :

$$\underline{X}_u \in \Omega_c \quad \text{tel que } \mu_{\Omega_c}(\underline{X}_u) = \max_{r=1, M} \mu_{\Omega_r}(\underline{X}_u) \tag{3.45}$$

Cette règle n'autorise pas la notion de rejet d'ambiguïté. Pour diminuer le risque de mauvaise classification, il convient donc de définir de nouvelles règles de décision non exclusive permettant de contrôler ce type de rejet.

### 3.5.1.4.2.1. Rejet d'ambiguïté

Une règle de décision non exclusive incluant la notion de rejet d'ambiguïté doit être basée sur une analyse non binaire des fonctions d'appartenance. L'idée la plus simple consiste à décider de l'affectation d'une observation  $\underline{X}_u$  aux différentes classes par comparaison des valeurs des fonctions  $\mu_{\Omega_c}(\underline{X}_u)$  ( $c=1,M$ ) à un seuil d'appartenance. Pour que ce rejet puisse être appliqué, il est nécessaire que les  $k$  plus proches voisins soient issus d'au moins deux classes différentes.

Ainsi, une nouvelle observation est rejetée en ambiguïté :

- Si la différence entre les degrés d'appartenance aux classes les plus proches physiquement (i.e. contenant des plus proches voisins) est inférieure ou égale à un seuil d'ambiguïté  $T_{ne}$  (par exemple, pour un seuil  $T_{ne}=0.05$ , si  $\underline{X}_u$  a une appartenance de 0.25 à la classe  $\Omega_c$  et une appartenance de 0.22 à la classe  $\Omega_r$  alors cette observation est rejetée car  $0.25-0.22 < T_{ne}$ ).

Les équations régissant ce rejet sont :

$$\begin{aligned}
 & \text{Si } \mu_{\Omega_c}(\underline{X}_u) \neq 0 \\
 & \text{alors} \\
 & \left\{ \underline{X}_u \in \Omega_o \quad \text{si } \mu_{\Omega_c}(\underline{X}_u) - \mu_{\Omega_r}(\underline{X}_u) \leq T_{ne} \right.
 \end{aligned} \tag{3.46}$$

*Remarque* : Plus  $T_{ne} \rightarrow 0$  (1), moins (plus) les observations sont rejetées en ambiguïté.

Une information supplémentaire est apportée pour l'utilisateur concernant l'incertitude sur l'affectation d'une nouvelle observation. En effet, une règle peut être émise pour alerter sur la classification réalisée. Elle ne s'applique que lorsque les  $k$  plus proches voisins sont issus d'au moins deux classes. Elle indique que les degrés d'appartenance de l'observation aux classes sont inférieurs à l'incertitude, ce qui laisse penser que l'observation se situe dans une zone de l'espace faiblement représentée par les classes de  $X_a$ . La règle émet donc une alerte si :

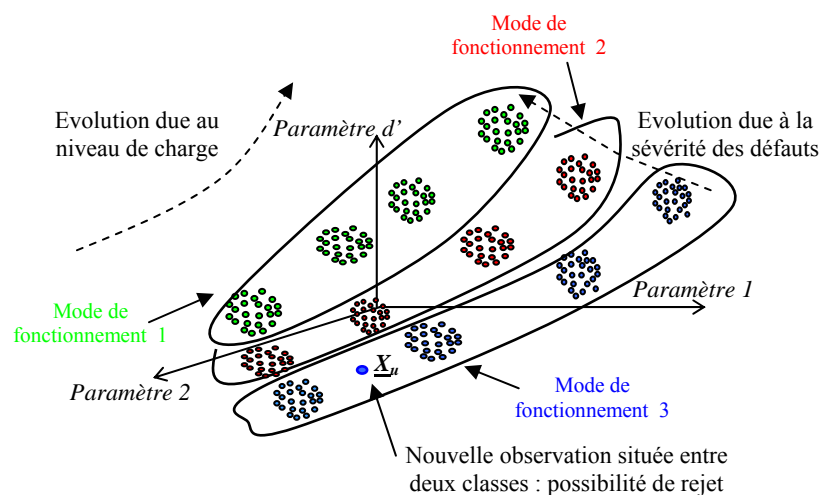
- La somme des degrés d'appartenance est inférieure à l'incertitude sur l'appartenance (par exemple une observation appartient à 5% à la classe  $\Omega_1$  et à 20% à la classe  $\Omega_2$  alors la somme des deux, 25%, est inférieure à l'incertitude, qui est de 75%).



De ce fait, une nouvelle observation localisée entre deux classes peut correspondre à un état intermédiaire de ce mode. En utilisant la règle de décision exclusive ou non exclusive avec les options de rejet, l'observation peut être rejetée. Il est donc nécessaire de prendre en compte l'évolution des différents modes de fonctionnement du système.

### 3.6. Prise en compte de l'évolution des modes de fonctionnement

Comme énoncé précédemment, un des inconvénients de la règle des  $k - ppv$  est qu'une nouvelle observation localisée entre deux classes peut être rejetée en ambiguïté ou en distance alors que celle-ci appartient au même mode de fonctionnement défini par ses classes (figure 3.26). En général, à cause d'un faible retour d'expérience, l'ensemble d'apprentissage initial ne couvre pas exhaustivement tous les modes de fonctionnement du système. Ainsi, dans notre application, il est impossible d'avoir des mesures pour n'importe quel niveau de charge et pour n'importe quel degré de sévérité pour un type de défaut donné.



**Fig. 3. 26.** Représentation de modes de fonctionnement composés de plusieurs classes

Il est donc nécessaire, pour obtenir un diagnostic sans erreur, de prendre en compte l'évolution des modes de fonctionnement par rapport au niveau de charge du moteur mais aussi l'évolution due à la sévérité des défauts (figure 3.26).

### 3.6.1. Approche polynomiale

#### 3.6.1.1. Estimation des états manquants

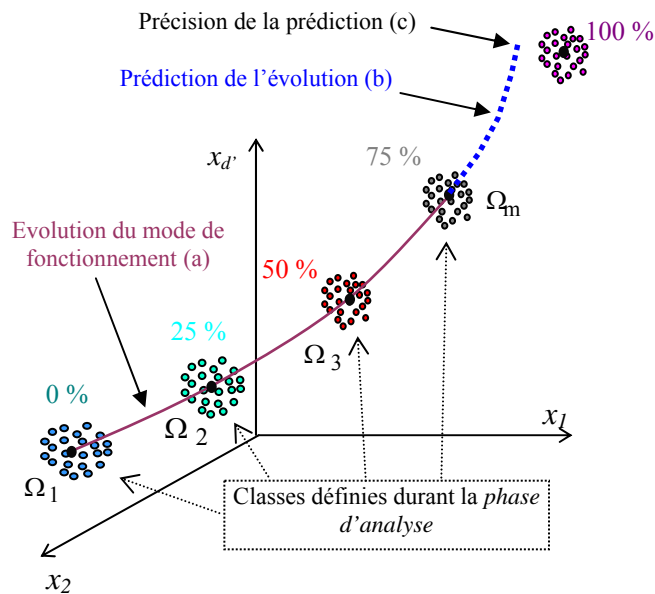
Une interpolation linéaire peut être appliquée entre les différents centre de gravité des L classes constituant un même mode de fonctionnement (par exemple, sur la figure 3.26, L=4). Ceci va permettre d'obtenir des vecteurs références entre les classes de ce mode. Cette interpolation mène à une fonction représentant une '*trajectoire moyenne*' (figure 3.27a). Elle est estimée par une fonction mathématique pour chaque dimension  $d'$  formant l'espace de représentation. Comme cette évolution est lente, la fonction mathématique utilisée est une fonction polynomiale, dont l'expression est la suivante :

$$x_j = a(j) + \sum_{r=1}^{L-1} (b_r(j) z^r) \text{ avec } (j=1, d') \quad (3.51)$$

avec :

- $z$  : évolution du niveau de charge ou de la sévérité des défauts (variable d'évolution)
- $d'$  : dimension de l'espace de représentation après réduction
- $a(j), b(j)$  : coefficients déterminés pour chaque dimension  $j$ . Ces coefficients peuvent être facilement calculés en utilisant la méthode d'identification des moindres carrés.
- $x_j$  :  $j^{\text{ème}}$  composante de  $E'$  (vecteur forme)
- $L$  : nombre de classes constituant un mode de fonctionnement

Dans le cas du niveau de charge, l'évolution des modes de fonctionnement pourrait être décrite par une fonction discrétisée par pas de 1% de charge.



**Fig. 3.27.** Représentation de la trajectoire moyenne d'un mode de fonctionnement

Prenons en exemple le moteur asynchrone. Un ensemble d'apprentissage, comprenant un mode de fonctionnement sain pour différents niveaux de charge (0%, 25%, 50% et 75%), est considéré (figure 3.27). Dans ce cas, il y a 4 classes ( $L=4$ ), donc à partir de (3.51), la fonction de l'évolution de ce mode est définie par :

$$x_j = a(j) + b_1(j) \cdot z + b_2(j) \cdot z^2 + b_3(j) \cdot z^3 \quad (3.52)$$

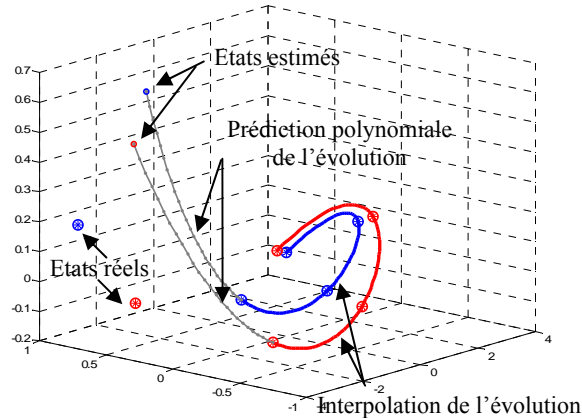
### 3.6.1.2. Vers la prédiction

Pour avoir une complète évolution de ce mode de fonctionnement, il est facile d'extrapoler la fonction (3.52) de 75 % à 100 % de charge (figure 3.27b). De cette façon, un mode de fonctionnement complet est obtenu dans l'ensemble d'apprentissage initial. Cette extrapolation est donc simple à mettre en œuvre et permet d'obtenir des vecteurs références dans une zone de l'espace de représentation encore inconnue. Elle réalise une prédiction de l'évolution du mode de fonctionnement vers des niveaux de charge non étudiés. Reste à savoir quelle est la précision (figure 3.27c) et la confiance de cette prédiction. Cela sera étudié au chapitre suivant.

### 3.6.2. Approche par filtre de Kalman

Dans le paragraphe précédent, l'interpolation des états manquants est réalisée à partir d'une approche polynomiale. Cette méthode peut être étendue à la prédiction. Cependant, elle est mal adaptée lorsque la règle des  $k - ppv$  est employée. La figure 3.28 présente l'interpolation

réalisée entre 4 classes (correspondant aux quatre niveaux de charge 0, 25, 50 et 75% de l'exemple précédent, équation 3.52) et la prédiction vers un horizon inconnu (en l'occurrence 100% de charge).



**Fig. 3. 28.** Exemple d'interpolation et de prédiction polynomiale à partir de 4 classes ( $L=4$ ).

En effet, une nouvelle observation  $\underline{X}_u$  peut être mal affectée si une mauvaise prédiction de l'évolution du mode de fonctionnement est effectuée (figure 3.28). De plus, cette méthode de prédiction présente quelques inconvénients : l'approche polynomiale mène à une relation algébrique et donc, elle ne tient pas compte de la dynamique d'évolution des modes de fonctionnement. C'est pour ces raisons qu'une approche différente est proposée utilisant dans un premier temps l'estimateur récursif de Kalman [Gelb, 74] pour déterminer les paramètres du modèle dynamique décrivant l'évolution d'un mode, et dans un deuxième temps, le prédicteur de Kalman [Brown, 92a] pour prédire cette même évolution vers des zones inconnues de l'espace.

### 3.6.2.1. Estimation des états manquants

Cette approche proposée est complémentaire avec l'approche par RdF. Cependant, il est difficile de transposer le formalisme des équations d'état en reconnaissance des formes. Pour pallier à ce problème, un modèle décrivant, dans l'espace de représentation, l'évolution de chaque composante du vecteur forme, est défini comme suit

$$\begin{cases} x_{k+1} = F_k x_k + G_k u_k \\ y_k = C_k x_k \end{cases} \quad (3.53)$$

$x_k$  : Vecteur de dimension n

$u_k$  : Vecteur de commande ou d'entrée de dimension p

$y_k$  : Vecteur de mesure de dimension  $m$   
 $F_k(n \times n), G_k(n \times p), C_k(m \times n)$  : Matrices  
 $n$  : Ordre du système

$F_k$  et  $G_k$  sont des matrices variant avec la variable d'évolution  $k$  et peuvent être également fonction de  $x_k$  et  $u_k$ . Le problème est qu'aucune information n'est disponible et donc, plusieurs cas peuvent être envisagés.

- 1<sup>er</sup> cas :

Les coefficients des matrices  $F_k$  et  $G_k$ , qu'elles soient pleines ou diagonales, sont considérés comme constants. Dans ce cas, une relation linéaire est obtenue, qui s'avère non suffisante pour décrire l'évolution des modes de fonctionnement. Cette solution a été testée et les résultats ont confirmé que cette hypothèse ne permettait pas une estimation paramétrique correcte. Cette solution ne convient donc pas car elle ne représente pas l'évolution réelle des modes du processus.

- 2<sup>ème</sup> cas :

Il s'agit de la formulation la plus générale. La matrice  $F_k$  est considérée comme pleine et ses coefficients sont non constants. Le fait que la matrice soit pleine indique qu'il y a une contribution de chacune des composantes du vecteur forme sur l'évolution d'une seule composante et un possible couplage entre ces mêmes composantes.

$$F_k = \begin{bmatrix} f_{11k} & f_{12k} & \cdots & f_{1nk} \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ f_{n1k} & \cdots & \cdots & f_{nnk} \end{bmatrix} \quad (3.54)$$

Ceci va entraîner, dans le cas de l'estimation paramétrique par Kalman, la résolution d'un problème de grande dimension ( $n^2+n$ ). Cette solution a pour l'instant été écartée en raison de temps de calcul trop important.

- 3<sup>ème</sup> cas :

La matrice  $F_k$  est considérée comme une matrice diagonale, ce qui suppose qu'il n'y a pas de contribution entre les composantes du vecteur forme.

$$F_k = \begin{bmatrix} f_{11k} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f_{nnk} \end{bmatrix} \quad (3.55)$$



La dimension du problème est fortement diminuée car elle passe d'une dimension  $(n^2+n)$  à une dimension  $(n+n)$ . Cette solution est celle retenue par la suite. Elle a été testée (§ 4.5.2.4, chapitre IV) et permet de bien décrire l'évolution des modes de fonctionnement. Elle présente aussi l'avantage de pouvoir découpler le problème, c'est-à-dire de faire fonctionner en parallèle,  $n$  modèle réalisant chacun l'estimation des paramètres d'évolution d'une composante du vecteur forme.

Cette hypothèse débouche sur la définition d'une représentation d'état modal de l'évolution des composantes, ou les états  $x_k$  sont les  $d$  composantes du vecteur forme. Au final, nous supposons que l'évolution de chaque composante du vecteur forme est représentée par un modèle du premier ordre non linéaire tel que :

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k u_k + w_k \\ y_k &= C x_k + v_k \end{aligned} \quad (3.56)$$

Où

- $k$  est l'instant d'observation ( $k \geq 0$ )
- $x_k$  ( $1 \times 1$ ) est le vecteur d'état du processus pour la valeur  $k$  (vecteur forme en RdF),
- $u_k$  ( $1 \times 1$ ) est le vecteur d'entrée,
- $y_k$  ( $1 \times 1$ ) est la mesure. L'évolution de la variable  $k$  correspond à une variation (augmentation dans notre cas) du niveau de charge. Les mesures qui sont considérées ici, sont l'ensemble des évolutions des composantes du vecteur forme obtenues à l'aide de l'approche par interpolation (*trajectoire moyenne*, § 3.6.1.1).
- Le bruit d'état ( $w_k$ ) et le bruit de mesure ( $v_k$ ) sont des bruits stationnaires non corrélés de valeur moyenne nulle avec pour valeurs :

$$E[w_k w_i^T] = Q_w \quad E[v_k v_i^T] = R \quad , \quad \forall k = i, \text{ et égal à zéro } \forall k \neq i. \quad (3.57)$$

- $F_k$  est la matrice ( $1 \times 1$ ) de transition d'état (matrice d'évolution),  $G_k$  est la matrice ( $1 \times 1$ ) de commande (matrice de poids du niveau de charge) et  $C$  est la matrice ( $1 \times 1$ ) de sortie.

Dans le tableau 3.6, un résumé des différentes variables utilisées, est présenté. La définition de ces variables est faite d'un point de vue système et RdF. Par conséquent, il est possible de connaître la correspondance entre l'approche de Kalman classique et l'approche de Kalman appliquée à la reconnaissance des formes.

	Point de vue système	Point de vue RdF	Dimension
$x_k$	Vecteur d'état	Vecteur forme	$(d' \times 1)$
$u_k$	Vecteur d'entrée	Niveau de charge	$(1 \times 1)$
$y_k$	Mesure	Evolution polynomiale	$(d' \times 1)$
$w_k, v_k$	Bruits	-	$(1 \times 1)$
$F_k$	Matrice de transition d'état	Matrice d'évolution	$(d' \times d')$
$G_k$	Matrice de commande	Matrice poids du niveau de charge	$(d' \times 1)$
$C_k$	Matrice de sortie	Matrice identité	$(d' \times d')$
$L_k$	Matrice d'observation	Matrice identité	$(d' \times d')$
$k$	Variable d'évolution	Variation de niveau de charge	

**Tableau. 3. 6.** Définition des différentes variables utilisées d'un point de vue système et RdF

Comme toutes les composantes du vecteur forme sont disponibles, tous les états sont supposés mesurables ce qui fait que la matrice de sortie  $C$  est une matrice identité.

Ainsi, on peut dire que l'étape d'estimation est utilisée pour déterminer les paramètres du modèle et, par la suite, l'étape de prédiction est utilisée pour réaliser la prédiction de l'évolution du vecteur forme.

Afin de déterminer les valeurs de paramètres du modèle  $\forall k$ , il est nécessaire de définir :

$$\begin{aligned} \lambda_k &= [F_k^{1,1} \quad F_k^{1,2} \quad \dots \quad F_k^{d',d'-1} \quad F_k^{d',d'}]' \\ \gamma_k &= [G_k^1 \quad G_k^2 \quad \dots \quad G_k^{d'-1} \quad G_k^{d'}]' \end{aligned} \tag{3.58}$$

Où  $F_k^{l,c}$ ,  $G_k^l$  sont les valeurs sur la ligne  $l$  et la colonne  $c$  de  $F_k$  and  $G_k$  respectivement.

Un processus de Wiener a été retenu pour décrire l'évolution des deux précédents paramètres  $\lambda$  et  $\gamma$ . Une nouvelle équation est donc définie comme suit

$$\begin{aligned} \lambda_{k+1} &= \lambda_k + \mu_k \\ \gamma_{k+1} &= \gamma_k + \varepsilon_k \end{aligned} \tag{3.59}$$

$\mu_k$  et  $\varepsilon_k$  sont des bruits stationnaires non corrélés de valeur moyenne nulle.

$$E[\mu_k \mu_i^T] = Q_\mu \quad E[\varepsilon_k \varepsilon_i^T] = Q_\varepsilon, \quad \forall k = i, \text{ et égal à zéro } \forall k \neq i. \tag{3.60}$$

Maintenant, nous pouvons définir les nouveaux éléments associés à (3.59)

$$\Psi_k = \begin{bmatrix} \lambda_k \\ \gamma_k \end{bmatrix}, \quad \bar{F} = \bar{G} = I, \quad \omega_k = [\mu_k \quad \varepsilon_k] \tag{3.61}$$

$$\bar{C}_k = \begin{bmatrix} x_k^T & 0 & \dots & 0 & u_k \\ 0 & x_k^T & & \dots & \cdot \\ \dots & \dots & \dots & 0 & \cdot \\ 0 & \dots & 0 & x_k^T & u_k \end{bmatrix} \quad (3.62)$$

Dès lors, une nouvelle représentation de (3.56) est introduite :

$$\begin{aligned} \Psi_{k+1} &= \bar{F} \Psi_k + \bar{G} \omega_k \\ y_k &= \bar{C}_{k-1} \Psi_{k-1} + w_{k-1} + v_k \end{aligned} \quad (3.63)$$

$\omega_k$  est aussi un bruit stationnaire non corrélé de valeur moyenne nulle

$$Q_\omega = E[\omega_k \omega_i^T] = \begin{bmatrix} Q_u & 0 \\ 0 & Q_\varepsilon \end{bmatrix} \quad \forall k = i, \text{ et égal à zéro } \forall k \neq i. \quad (3.64)$$

L'estimateur récursif de Kalman est utilisé pour déterminer l'évolution de  $F_k$  et  $G_k$ , nécessaire à la réalisation de la prédiction. Pour cela, il faut procéder à quatre étapes :

#### 1 – Initialization

$$\hat{\Psi}_{0/0} \quad P_{0/0} \quad (3.65)$$

#### 2 – Estimation a priori

$$\begin{aligned} \hat{\Psi}_{k/k-1} &= \bar{F} \hat{\Psi}_{k-1/k-1} \\ \hat{y}_{k/k-1} &= \bar{C}_{k-1} \hat{\Psi}_{k/k-1} \\ P_{k/k-1} &= \bar{F} P_{k-1/k-1} \bar{F}^T + \bar{G} Q_\omega \bar{G} \end{aligned} \quad (3.66)$$

#### 3 - Innovation

$$\begin{aligned} \tilde{y}_{k/k-1} &= y_k - \hat{y}_{k/k-1} \\ \Gamma_{k/k-1} &= \bar{C}_k P_{k/k-1} \bar{C}_k^T + R + Q_w \end{aligned} \quad (3.67)$$

#### 4 – Estimation a posteriori

$$\begin{aligned} K_k &= P_{k/k-1} \bar{C}_k^T \Gamma_{k/k-1}^{-1} \\ \hat{\Psi}_{k/k} &= \hat{\Psi}_{k/k-1} + K_k \tilde{y}_{k/k-1} \\ P_{k/k} &= (I - K_k \bar{C}_k) P_{k/k-1} \end{aligned} \quad (3.68)$$

*Remarque* :  $\hat{X}$  représente une estimation de la variable.

### 3.6.2.2. Vers la prédiction

Pour réaliser la prédiction, le prédicteur à n pas suivant [Brown, 92a] est utilisé :

$$\hat{x}_{k+n/k} = \hat{F}_k^n \hat{x}_{k/k} + \hat{G}_k u_{k+n-1} + \sum_{j=k}^{k+n-1} \left( \hat{F}_k^{k+n-1-j} \hat{G}_k u_j \right) \quad (3.69)$$

La matrice de covariance est prédite :

$$P_{k+n/k} = \left( \hat{F}_k^n \right) P_{k/k} \left( \hat{F}_k^n \right)^T + \sum_{j=k}^{k+n-1} \left( \left( \hat{F}_k^{k+n-1-j} \right) Q_w \left( \hat{F}_k^{k+n-1-j} \right)^T \right) \quad (3.70)$$

Il est à noter que cette méthode consiste à la détermination de la meilleure estimation des paramètres  $F_k$  et  $G_k$  du système, avant de commencer l'opération de prédiction. La contribution d'une telle technique peut être visualisée en comparant les figures 3.28 et 3.29 mais sera montrée dans le chapitre suivant.

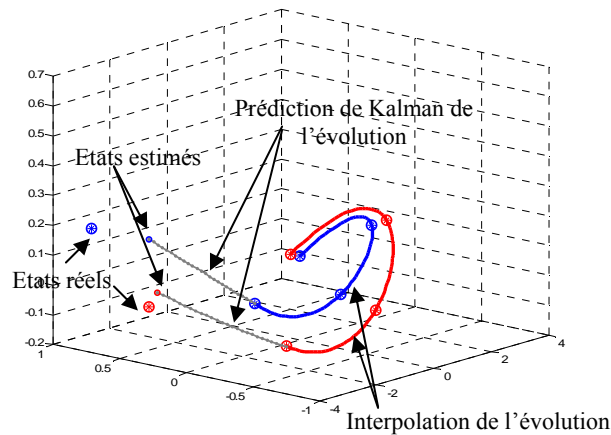


Fig. 3. 29. Exemple d'interpolation et de prédiction par Kalman à partir de 4 classes (L=4).

### 3.6.3. Mise en place du rejet d'affectation

Les observations rejetées par la règle de décision non exclusive sont ensuite analysées à partir de la *trajectoire moyenne* réalisée par l'interpolation linéaire (§ 3.6.1.1). Ainsi, il est possible de déterminer si l'observation rejetée appartient en réalité à un mode de fonctionnement défini dans l'ensemble d'apprentissage mais avec un niveau de charge différent.

Comme pour la règle de décision précédemment exposée, un rejet en distance doit être appliqué pour éviter des erreurs de classification. Ce rejet est représenté par une *trajectoire* dite '*maximale*' (figure 3.30). Il est déterminé de la manière suivante :

A partir d'un mode de fonctionnement composé de L classes, pour chaque classe, le vecteur le plus éloigné, en terme de distance (euclidienne ou Mahalanobis), de leur centre de gravité respectif est recherché. La dispersion maximale de ces classes est, de ce fait, déterminée. Ensuite, une fonction mathématique (fonction polynomiale) est estimée, de la même manière qu'en 3.6.1.1, à l'aide des L points qui ont été calculés. La prédiction de l'évolution de la *trajectoire maximale* est réalisée à partir de l'algorithme de Kalman, développé en 3.6.2.

Ainsi, une *trajectoire maximale* (évolution + prédiction) est obtenue permettant de caractériser la dispersion de l'ensemble du mode de fonctionnement en définissant une enveloppe autour des classes de ce mode. Cette enveloppe est représentée sur la figure 3.30.

Grâce à cela, dans un premier temps, il est possible de déterminer si l'observation rejetée  $\underline{X}_u$  peut correspondre à un niveau de charge non défini, pour le mode de fonctionnement, dans  $X_a$ .

On procède de la façon suivante :

$$\text{Si pour } s = [1\%, 100\%]$$

$$d(\underline{P}_s, \underline{X}_u) \leq \beta \cdot d(\underline{P}_s, \underline{P}_{s \max}) \text{ alors } N_{ch} = s \%$$

Alors

$$\underline{X}_u \text{ appartient au mode de fonctionnement avec un niveau de charge} \quad (3.71)$$

$$\text{ou de sévérité } N_{ch}$$

Sinon

$$\underline{X}_u \text{ est rejetée en distance}$$

avec :

s : représente soit le niveau de charge, soit le degré de sévérité de la défaillance

$\underline{P}_s$  : points de l'évolution du mode de fonctionnement (*trajectoire moyenne*)

$\underline{P}_{s \max}$  : points du seuil d'acceptation (*trajectoire maximale*)

$N_{ch}$  : niveau de charge ou de sévérité de la défaillance de la nouvelle observation

$\beta$  : permet de prendre en compte l'incertitude sur les propriétés des classes

Dans un deuxième temps, si  $\underline{X}_u$  est affectée à un mode de fonctionnement par (3.71) avec un niveau de charge ou de sévérité  $N_{ch}$ , pour obtenir des informations plus précises sur sa localisation dans l'enveloppe, il est nécessaire d'utiliser la règle de décision non exclusive développé en 3.5.1.4.2.

Par conséquent, la fonction d'appartenance  $\mu_{P_s}(\underline{X}_u)$  va être définie, non plus pour une classe, mais pour un point  $\underline{P}_s$  de la *trajectoire moyenne* correspondant au niveau de charge ou de sévérité de l'observation. Cette fonction est alors définie comme suit

$$\mu_{P_s}(\underline{X}_u) = \begin{cases} 1 - (1 - \eta) \left( \frac{d(\underline{X}_u, \underline{P}_s)}{d_{seuil}} \right)^g & \text{si } 0 \leq d(\underline{X}_u, \underline{P}_s) \leq d_{seuil} \\ \eta \left( \frac{\beta}{\beta - 1} \right)^g \left( \frac{\beta \cdot d_{seuil} - d(\underline{X}_u, \underline{P}_s)}{\beta \cdot d_{seuil}} \right)^g & \text{si } d_{seuil} \leq d(\underline{X}_u, \underline{P}_s) \leq \beta \cdot d_{seuil} \\ 0 & \text{sinon} \end{cases} \quad (3.72)$$

avec :

$d(\underline{X}_u, \underline{P}_s)$  est la distance entre une observation  $\underline{X}_u$  et le point  $\underline{P}_s$  de la *trajectoire moyenne*

$d_{seuil}$  est le paramètre qui définit la position des points d'inflexion. Il correspond à la distance entre le  $\underline{P}_s$  de la *trajectoire moyenne* et le  $\underline{P}_{smax}$  de la *trajectoire maximale*

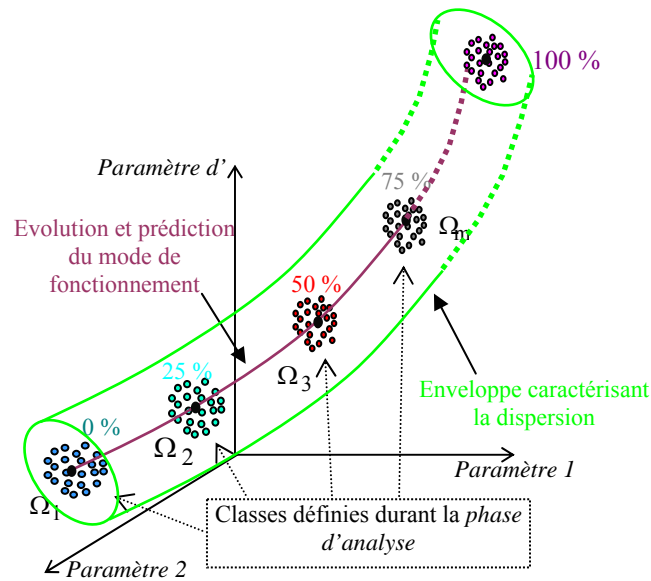
$$d_{seuil} = d(\underline{P}_s, \underline{P}_{smax})$$

$$\mu_{P_s}(\underline{X}_u) = \eta \text{ si } d(\underline{X}_u, \underline{P}_s) = d_{seuil}$$

$\eta$  est le paramètre servant à fixer le plateau de la fonction  $\pi$ .

$\beta$  est un paramètre ( $\beta \in \mathbb{R}^{+*} > 1$ ) qui permet de définir la largeur de bande de la fonction  $\pi$  (les points d'arrêt de la fonction  $\pi$ ) (§ 3.5.1.2.2).

$g$  est un paramètre permettant de faire varier la pente de la fonction.



**Fig. 3.30.** Représentation de l'évolution et de la prédiction du seuil de rejet dans le cas de l'évaluation du niveau de charge

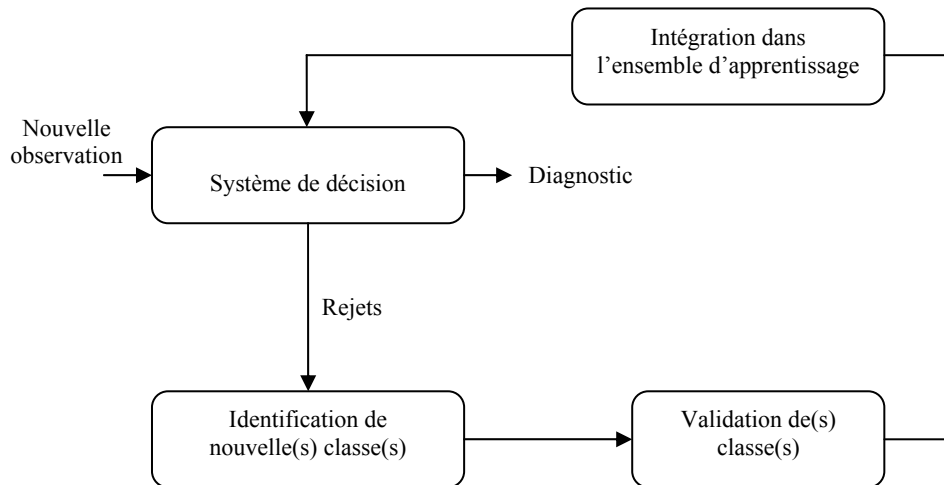
Cette fonction d'appartenance permet donc de localiser la position de l'observation. Le rejet d'ambiguïté défini par (3.46) ne s'applique pas car si la condition (3.71) est respectée alors l'observation est située dans l'enveloppe et par conséquent, elle ne peut appartenir qu'au mode de fonctionnement contenu dans l'enveloppe. Par contre, le rejet en distance défini par (3.49) est lui, tout à fait applicable.

La prise en compte de l'évolution des modes de fonctionnement à l'aide du filtre de Kalman permet de diminuer le nombre d'acquisition nécessaire à la *phase d'apprentissage*, pour constituer l'ensemble d'apprentissage. En effet, à partir de quelques points servant à initialiser le modèle, il est possible de définir l'évolution complète d'un mode de fonctionnement et de déterminer une enveloppe autour de ce mode. De plus, ce modèle permet de prédire l'évolution future de ce mode dans une partie de l'espace encore inconnue. Cette méthode peut être utilisée d'une part, pour estimer les états manquants d'un mode de fonctionnement par rapport au niveau de charge et, d'autre part, pour prédire et évaluer le niveau de sévérité des défaillances.

### ***3.7. Actualisation de la connaissance***

Du fait de la connaissance souvent non exhaustive sur les modes de fonctionnement du système, il est nécessaire d'employer une méthode de décision avec option de rejet (en distance et en ambiguïté) : ainsi, certaines observations sont classées et d'autres rejetées. Ces observations rejetées peuvent définir de nouvelles classes qui correspondent à de nouveaux modes de fonctionnement du système.

Il est donc essentiel pour une méthode diagnostic, de pouvoir identifier l'apparition de nouveaux modes de fonctionnement, puis de pouvoir les intégrer au processus de diagnostic, c'est-à-dire, en un mot, d'être adaptatif.



**Fig. 3. 31.** *Système de RdF adaptatif*

Dans tous les cas, il convient de déterminer à partir de combien d'observations rejetées on peut considérer avoir l'apparition d'une nouvelle classe.

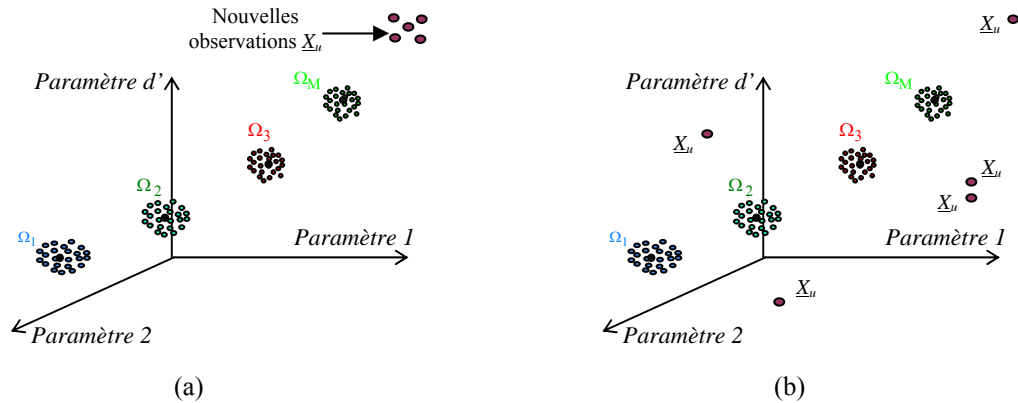
Le choix peut se faire de manière arbitraire tout en éliminant la valeur 1 qui ne représente pas de façon significative l'apparition de nouvelles classes. En effet, un nombre trop petit peut être source d'erreur du fait des incertitudes de mesures (bruit ou problèmes de capteurs de mesures). Mais il ne doit pas être trop grand non plus car si la nouvelle classe correspond à un mode de fonctionnement dégradé ou défectueux, il faut que celui-ci soit détecté le plus rapidement possible pour ne pas endommager le système.

Pour la 'confirmation' d'une nouvelle classe  $\Omega_{new}$ , nous proposons un critère qui tienne compte du nombre d'observations contenues dans cette classe. De ce fait, une nouvelle classe  $\Omega_{new}$  est confirmée si le nombre d'observations  $N_{new}$  est supérieur à un certain seuil :

$$N_{new} \geq N_{min} \quad (3.73)$$

Mais ceci n'est pas la seule condition pour confirmer l'apparition d'une nouvelle classe. En premier lieu, il est nécessaire de disposer d'information sur la localisation de ces observations rejetées. Des observations situées dans une même zone de l'espace de représentation ont une forte probabilité de définir l'apparition d'un nouveau mode de fonctionnement (figure 3.32a). Par contre des observations dispersées dans différentes zones de l'espace de représentation n'ont pas la même signification (figure 3.32b). Il peut s'agir soit, d'erreurs de mesures pour certaines, soit, d'un début d'apparition d'un mode de fonctionnement inconnu mais dans ce cas là, il faut attendre l'arrivée d'autres observations pour conclure à l'occurrence d'une nouvelle classe (c'est-à-dire avoir au moins  $N_{min}$  observations dans une même zone).





**Fig. 3. 32.** Analyse des observations rejetées. Apparition d’une nouvelle classe, observations situées dans une même zone (a), observations dispersées dans différentes zones de l’espace (b)

Ainsi, pour disposer d’information sur la localisation des observations rejetées, deux solutions peuvent être envisageables : la première en tenant compte de la dispersion des observations rejetées, et la deuxième, en utilisant le maillage de l’espace de représentation développé précédemment.

### 3.7.1. Analyse de la compacité des observations rejetées

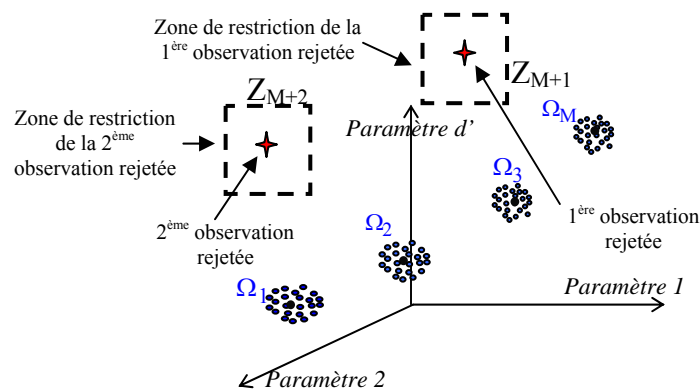
A partir de  $N_{\min}$  observations rejetées, il est possible de déterminer la compacité de ces observations, appelée  $C_r$ , à l’aide de la matrice de variance intra-classe (3.9). Cette compacité est calculée en réalisant une permutation circulaire entre les différentes observations rejetées. Par exemple, dans le cas de 4 observations rejetées  $X_{uP_r}$  ( $P_r=1$  à 4), si on fixe  $N_{\min}=3$ , il y a 4 combinaisons possibles :  $X_{u1} X_{u2} X_{u3}$ ,  $X_{u2} X_{u3} X_{u4}$ ,  $X_{u3} X_{u4} X_{u1}$ ,  $X_{u4} X_{u1} X_{u2}$ . Ainsi pour ces quatre combinaisons, qui représentent potentiellement des nouvelles classes, la compacité est calculée. Ensuite ces compacités sont comparées à la dispersion intra-classe moyenne  $C_m$  (égale à  $\frac{\sum w}{M}$ , M étant le nombre de classes dans  $X_a$ ) des classes définies dans l’ensemble d’apprentissage. Si elles sont du même ordre de grandeur alors on peut dire qu’il y a apparition d’une ou plusieurs nouvelles classes.

De façon plus générale, la détection ou l’apparition d’une nouvelle classe peut se résumer ainsi :

Détection de nouvelles classes
<p><b>1.</b> Calcul de la compacité moyenne <math>C_m</math> des classes définies dans l'ensemble d'apprentissage</p> <p style="padding-left: 20px;"><b>Si</b> le nombre d'observations rejetées <math>&gt; N_{min}</math></p> <p><b>2.</b> Calcul de la compacité des observations rejetées <math>C_r</math></p> <p><b>3.</b> Comparaison de <math>C_m</math> et <math>C_r</math></p> <p style="padding-left: 20px;"><b>Si</b> <math>C_r &lt; 1.5 \times C_m</math></p> <p style="padding-left: 20px;"><b>Alors</b> Apparition d'une ou plusieurs nouvelles classes</p> <p><b>4.</b> Expertise humaine pour étiqueter la ou les nouvelles classes</p> <p><b>5.</b> Intégration de cette ou ces nouvelles classes dans l'ensemble d'apprentissage</p> <p style="text-align: center;"><b>Fin de Si</b></p>

### 3.7.2. Localisation des observations rejetées par maillage de l'espace de représentation

En se basant sur l'approche d'un maillage de l'espace de représentation décrit dans le paragraphe 3.5.1.3, il est envisageable de détecter l'apparition de nouvelles classes. Pour cela, lorsqu'une première observation est rejetée par la règle de décision, une zone de restriction, dont la dimension est déterminée à partir de la dispersion moyenne  $C_m$  des classes de  $X_a$  (figure 3.17), est définie.



**Fig. 3. 33.** Représentation de la 1<sup>ère</sup> et de la 2<sup>ème</sup> observations rejetées ainsi que leur zone de restriction respective

Ensuite, pour la seconde observation rejetée, il faut déterminer si elle appartient à cette zone  $Z_{M+1}$  en regardant ces coordonnées. Si c'est le cas, la zone  $Z_{M+1}$  est alors composée de ces deux observations autrement une nouvelle zone  $Z_{M+2}$ , qui est composée de cette 2<sup>ème</sup> observations est créée (figure 3.33). Par la suite, l'étape est la même pour toutes les

observations rejetées suivantes. Au final, lorsqu'une nouvelle zone est composée d'au moins  $N_{\min}$  observations, on peut considérer qu'il y a apparition d'une nouvelle classe.

Une option a été ajoutée pour éviter un trop grand nombre de zones créées. Une observation peut être rejetée en raison d'une erreur de mesure ou d'une variation inhérente au système. Dans ce cas, une zone sera créée autour de ce point. En supposant que le diagnostic est effectué en ligne et que les acquisitions sont réalisées avec un intervalle de temps fixe, si durant les  $L$  observations rejetées suivantes, aucune n'est assignée à cette zone alors celle-ci est éliminée.  $L$  doit être sélectionné en fonction du nombre de points  $N_{\min}$  mais aussi en fonction du temps entre les acquisitions.

L'algorithme d'analyse des observations rejetées peut être résumé de la façon suivante :

Algorithme d'analyse des observations rejetées
<ol style="list-style-type: none"> <li>1. Détermination d'une zone de restriction ou frontières d'appartenance <math>Z_{M+1}</math> pour la 1<sup>ère</sup> observation rejetée <math>\underline{X}_{u1}</math></li> <li>2. Initialisation: - compteur = 2 ; - ajout = 1 ;</li> <li>3. Pour <math>k = 2</math> à <math>P_r</math>, avec <math>P_r</math>: nombre d'observations rejetées <ul style="list-style-type: none"> <li>• Recherche si la <math>k^{\text{ème}}</math> observation <math>\underline{X}_{uk}</math> appartient aux zones créées oui : <math>\underline{X}_{uk}</math> est associée à la zone correspondante non : création d'une nouvelle zone <math>Z_{M+\text{compteur}}</math> Si durant les <math>L</math> observations suivantes, aucune n'appartient à la zone créée <math>Z_{M+\text{compteur}} \rightarrow</math> Alors cette zone est éliminée compteur = compteur + 1 ;</li> <li>• Pour <math>j = 1</math> à <i>compteur</i>, si <math>Z_{M+j}</math> est composée d'au moins <math>N_{\min}</math> observations, création d'une nouvelle classe <math>\Omega_{M+\text{ajout}}</math> ajout = ajout+1 ;</li> </ul> </li> </ol>

Les nouvelles classes détectées sont insérées, après validation et caractérisation par un expert, dans l'ensemble d'apprentissage initial, ce qui permet d'étoffer la base de connaissance du système. Ainsi, il est nécessaire de refaire une sélection de paramètres pour mieux définir le nouvel ensemble d'apprentissage augmenté et une étape de coalescence pour déterminer une nouvelle partition en classes de cet ensemble. La règle de décision sera aussi adaptée. Il faut aussi redéfinir les matrices de variance intra-classe et inter-classe.

### 3.8. Procédure de diagnostic par RdF

Au final, les différentes phases, la phase d'analyse et la phase d'exploitation, composant un système de reconnaissance des formes, ainsi que l'étape d'actualisation servant à la détection de nouveaux modes de fonctionnement peuvent être représentées de la manière suivante :

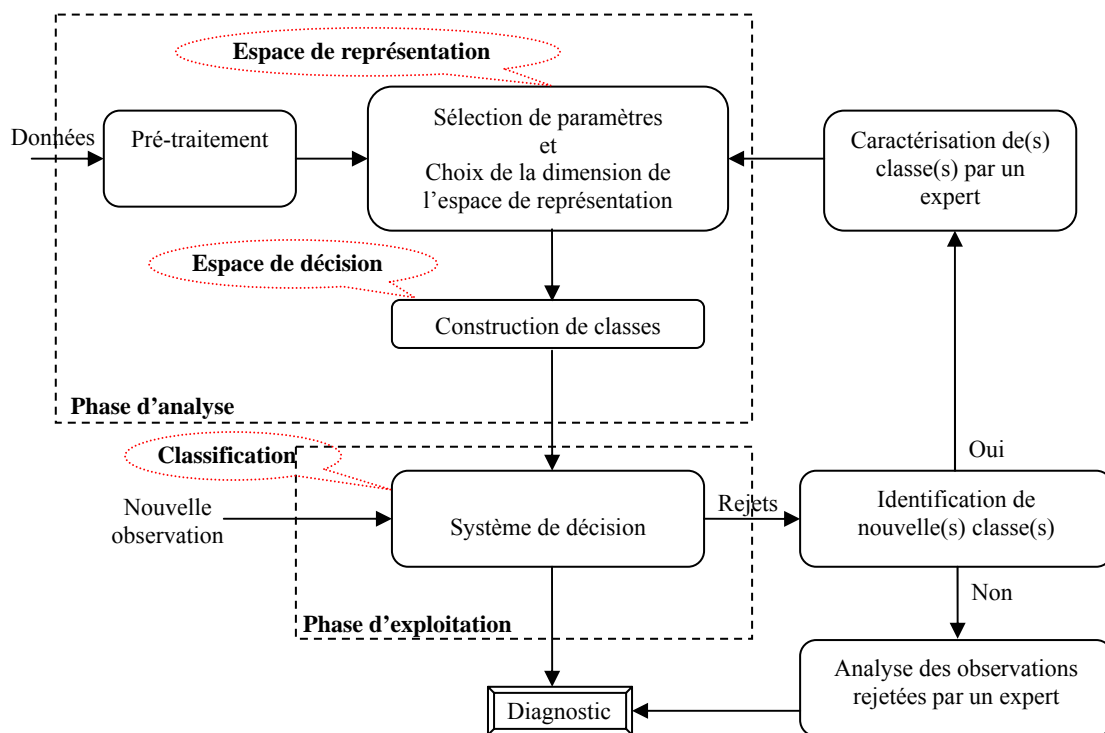


Fig. 3. 34. Schéma de principe de la procédure de diagnostic

### 3.9. Conclusion

Dans ce chapitre nous avons décrit les différentes étapes de l'élaboration d'un système de diagnostic par reconnaissance des formes. Cette approche consiste à confronter les données recueillies sur le processus aux informations sur les modes de fonctionnement disponibles dans la base de connaissance.

Dans la *phase d'analyse*, les méthodes de sélection de paramètres telles que SBS, critère de Fisher ou encore l'approche par Algorithme Génétique (AG), ont été choisies pour déterminer le vecteur forme optimal.

Afin d'obtenir un diagnostic quasi automatique, une méthode basée sur les AGs multi-population a été développée pour réaliser simultanément l'étape de sélection et de partition en classes de l'ensemble d'apprentissage.

L'extraction de paramètres a été réservée à la visualisation de la base de données initiale lorsque celle-ci est composée d'un vecteur forme optimal de dimension supérieure à trois.

Pour notre système de décision, nous avons retenu la règle des *k – plus proches voisins*. Cette règle est simple à élaborer et ne nécessite pas, au départ, la connaissance des lois de probabilités.

Toutefois, du fait de sa simplicité, cette règle présente quelques inconvénients. Pour réduire le temps de calcul, un maillage progressif de l'espace de représentation a été utilisé. Il permet de diminuer le nombre de distances à calculer.

Cette règle réalisant une affectation exclusive, des fonctions d'appartenance ont été associées à chaque classe de  $X_a$  permettant ainsi de prévenir d'éventuel dérive des modes de fonctionnement connus vers des modes inconnus mais aussi, de fournir un indice de fiabilité à la classification réalisée.

Le suivi d'évolution, réalisé à partir d'une approche par interpolation linéaire associée à une approche par filtre de Kalman, sert d'une part, à estimer les états manquants de l'ensemble d'apprentissage initial et, de ce fait, à réduire le nombre d'acquisitions nécessaire à la constitution de cet ensemble. D'autre part, il permet de prédire l'évolution des modes de fonctionnement aussi bien en fonction de la charge qu'en fonction de la sévérité des défauts.

Enfin, pour que ce système soit adaptatif, des approches permettant d'appréhender l'apparition de nouveaux de modes de fonctionnement à partir de l'analyse des observations rejetées ont été mises en places.

Dans le chapitre suivant, la construction du vecteur forme, c'est-à-dire le choix des différents paramètres caractérisant les défaillances du système, est présenté. Ensuite, toutes les différentes étapes qui élaborent le système de diagnostic par RdF vont être appliquées à la détection des défaillances sur un ensemble 'onduleur - machine asynchrone'. Une partie des mesures effectuées sur le procédé vont servir à valider les techniques de sélection, de coalescence et d'analyser l'apport de l'estimation des états manquants par l'approche de Kalman. Enfin, la *phase d'exploitation*, où une nouvelle observation est associée à l'une des classes précédemment définies et constituant l'ensemble d'apprentissage initial, va être traitée. Elle va permettre de caractériser les performances du système décisionnel mis en place et ainsi de valider notre approche de diagnostic.

# CHAPITRE IV

Application au diagnostic de la machine asynchrone



## 4.1. Introduction

Les chapitres précédents ont permis de présenter les principaux défauts pouvant intervenir sur un ensemble « convertisseur – machine asynchrone » ainsi que les différentes méthodes employées pour les diagnostiquer.

Ce chapitre porte sur l'application proprement dite de la méthode de diagnostic à base de reconnaissances des formes, détaillée au chapitre III, à la machine asynchrone détaillée au chapitre I.

Dans une première section, le banc moteur et les modes de fonctionnement sont présentés. L'analyse physique du processus permet de construire le vecteur forme, présenté dans la seconde section.

La troisième section est consacrée à la normalisation des données initiales par rapport à la charge. Cette technique va permettre de diminuer la dispersion causée par la charge. De ce fait, seule la variation due aux différents modes de fonctionnements (sain, défectueux) sera visible. Des exemples de données normalisées et non normalisées seront présentés et comparés pour montrer l'apport de cette normalisation.

Les deux sections suivantes portent sur l'application des méthodes précédentes pour différentes alimentations.

La quatrième section concerne le diagnostic d'une machine alimentée par le réseau. Tout d'abord, le choix du vecteur forme, c'est-à-dire le choix de l'ensemble des signatures possibles (au nombre de  $d$ ) caractéristiques des défaillances du système est présenté. Une fois cette étape réalisée, la recherche des paramètres les plus pertinents, du point de vue de la discrimination des modes de fonctionnement, est effectuée dans le cas *supervisé* et *non supervisé*. Une comparaison des vecteurs forme optimaux, obtenus à partir des différentes méthodes de sélection (Fisher, SBS, AG), est faite, permettant ainsi de choisir celui qui sépare au mieux les classes de l'ensemble d'apprentissage.

Après caractérisation de l'ensemble des classes (ou modes de fonctionnement), soit par expertise humaine, soit par coalescence, les règles de décisions présentées précédemment seront appliquées à la classification des nouvelles observations (en défaut ou non) du système. Les résultats de diagnostic obtenus sont comparés.

Les observations rejetées sont analysées pour mettre en évidence l'apparition d'une ou plusieurs nouvelles classes correspondant à des modes de fonctionnement encore inconnu de



l'ensemble d'apprentissage.

Enfin, le suivi de l'évolution d'une défaillance est effectué en utilisant l'approche polynomiale et l'approche de Kalman. Les résultats de ces deux approches sont comparés pour définir celle qui présente la meilleure précision, c'est-à-dire la plus proche du fonctionnement réel.

Il s'agit dans cette phase d'application de démontrer l'efficacité des méthodes de RdF utilisées pour suivre l'état de la machine asynchrone.

La dernière partie est consacrée au diagnostic de la machine alimentée par un onduleur de tension. La démarche dans cette partie sera la même que dans la partie précédente.

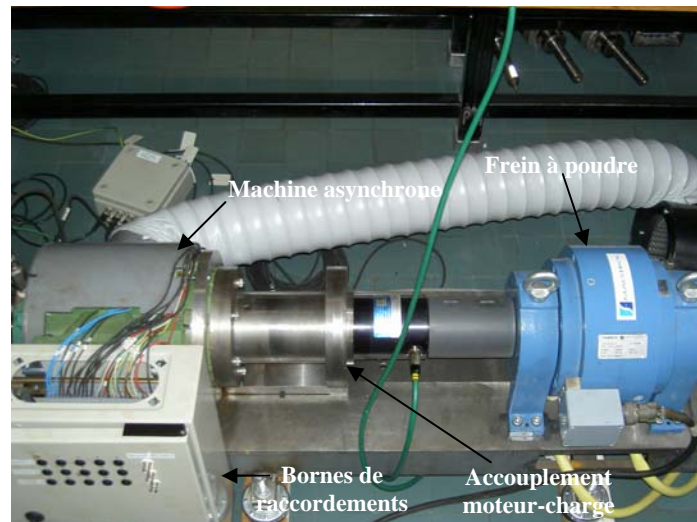
## ***4.2. Description du procédé***

### **4.2.1. Présentation du banc moteur**

Le banc est composé d'un moteur asynchrone triphasé à cage d'écureuil Leroy Somer LS 132S, IP 55, classe F, T°C normalisée = 40°C, régime nominal S1. Ce moteur est caractérisé par :

- La tension nominale entre phases : 400 V
- La fréquence d'alimentation : 50 Hz
- La vitesse nominale : 1440 tr/min
- La puissance utile nominale : 5.5 kW
- Le facteur de puissance :  $\cos \varphi = 0.84$
- Le courant nominal : 11,4 A
- Le nombre de paires de pôles :  $p = 2$
- La résistance stator par phase : 1,315  $\Omega$
- Le nombre d'encoches au rotor  $N_r = 28$
- Le nombre d'encoches au stator  $N_s = 48$

Les enroulements au stator sont couplés en étoile. Le moteur est chargé par un frein à poudre. Son couple maximal (100 Nm) est atteint à la vitesse nominale. Ce frein est dimensionné pour dissiper une puissance maximale de 5 kW. La figure suivante présente le banc moteur.



**Fig. 4. 1.** Moteur 5.5 kW accouplé à sa charge

#### 4.2.2. Acquisitions des données

Les signaux enregistrés par le système d'acquisition sont :

- les trois courants d'alimentation de la machine
- les trois tensions d'alimentation de la machine
- la vitesse de rotation
- les vibrations dans les directions suivantes :
  - direction verticale et perpendiculaire à l'axe du moteur, avec un accéléromètre côté opposé accouplement
  - direction parallèle à l'axe du moteur, côté opposé accouplement
  - direction horizontale et perpendiculaire à l'axe du moteur, côté opposé accouplement

Le système d'acquisition utilisé pour mesurer ces signaux possède :

- Huit entrées différentielles dites voies rapides qui peuvent être échantillonnées jusqu'à 10 MHz sur 14 bits. Ces entrées ont servi à mesurer les tensions ainsi que les courants.
- Huit entrées en mode commun dites voies lentes sur lesquelles la fréquence d'échantillonnage peut atteindre 1 MHz sur 16 bits. Une de ces entrées a servi à mesurer la vitesse de rotation.
- Quatre entrées sont réservées aux accéléromètres mesurant les vibrations. Les entrées "accéléromètres" offrent aussi la possibilité d'obtenir une fréquence d'échantillonnage de 100 kHz. Les accéléromètres piézo-électriques sont

alimentés par des sources de courants constants fournissant entre 1 et 10 mA sous 28 V.

La figure 4.2 représente le dispositif mis en place pour l'acquisition des signaux: ensemble machine charge et acquisition des signaux.

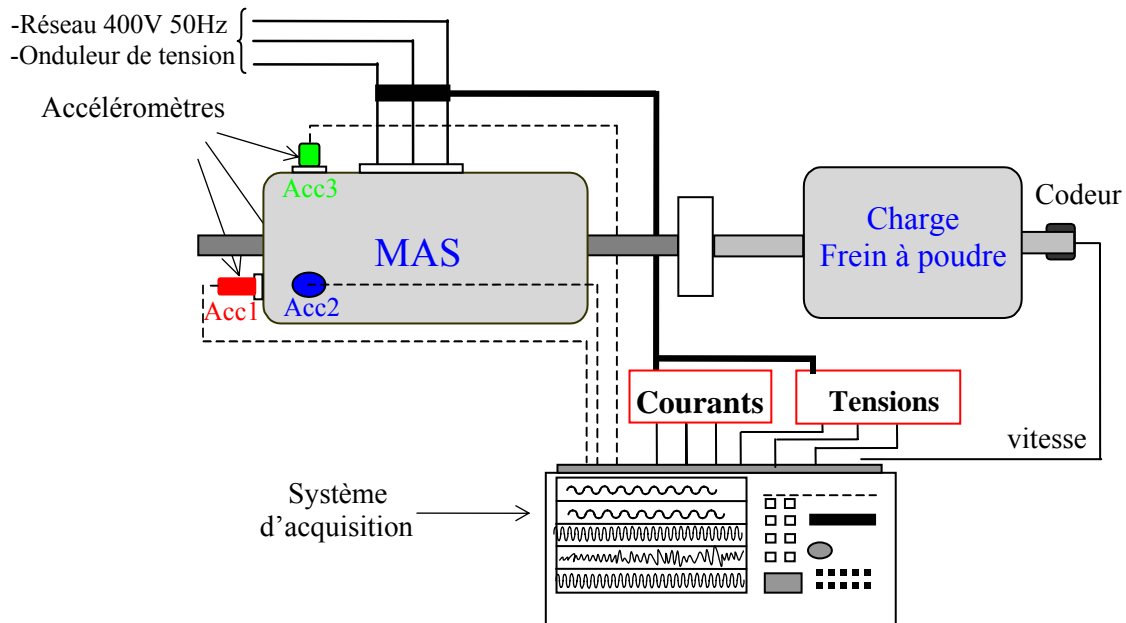


Fig. 4. 2. Système d'acquisition des signaux

Toutes les acquisitions seront réalisées en régime permanent sur une durée de 5 secondes, avec une fréquence d'échantillonnage de 20 kHz. Soit 100.000 points pour chacun des signaux mesurés.

### 4.2.3. Modes de fonctionnement étudiés

Les modes de fonctionnement ayant servi à valider la procédure de diagnostic sont repris dans le tableau 4.1. Chacun d'entre eux est représenté par une classe dans l'espace de décision. Les défauts ont été créés au stator et au rotor de la machine asynchrone, au niveau du filtre d'entrée (vieillesse condensateur) pour le convertisseur, et ceci pour plusieurs niveaux de charge.

	Mode de fonctionnement	Types de défauts	Taux de charge (Pour une puissance nominale : $P_u = 5.5$ kW)	Alimentation
Machine asynchrone	Machine saine	Aucun défaut	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	1 barre cassée au rotor	Défaut rotor (barres cassées)	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	3 barres cassées au rotor			
	4 barres cassées au rotor			
	Usure progressive (vieillessement naturel)	Défaut roulement	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	5% de déséquilibre sur une phase ( $V_b$ )	Déséquilibre d'alimentation au stator	0%, 25%, 50%, 75%, 100%	Réseau
	10% de déséquilibre sur une phase ( $V_b$ )			
	20% de déséquilibre sur une phase ( $V_b$ )			
40% de déséquilibre sur une phase ( $V_b$ )				
Convertisseur	Vieillessement condensateur de filtrage	Défaut convertisseur	0% et 50%	Onduleur

**Tableau. 4. 1.** Modes de fonctionnement servant à constituer l'ensemble d'apprentissage et les ensembles test

Toutes ces mesures ont été réalisées dans le but de valider la mise en place de notre approche diagnostic. Une partie de ces mesures va servir à construire l'ensemble d'apprentissage  $X_a$  ainsi que les ensembles de test. Ces ensembles vont être utilisés pour appréhender les algorithmes de reconnaissance des formes en déterminant la sensibilité des paramètres calculés et la définition du vecteur forme optimal.

Une autre partie des mesures va servir pour la *phase d'exploitation*. De ce fait, ces mesures seront parfaitement disjointes de l'ensemble d'apprentissage et des échantillons de test et serviront ainsi à tester l'efficacité des règles de décision.

L'évolution des défauts, tels que le nombre de barres cassées ou bien encore le pourcentage de déséquilibre d'alimentation, est utilisée pour montrer l'efficacité de l'estimation et la prédiction réalisée par l'approche Kalman.

Au niveau des rotors, les barres ont été cassées en perçant les barres des cages d'écureuil. L'usure obtenue sur les roulements est une usure naturelle d'utilisation (roulements fournis par la société SECCO). Le déséquilibre d'alimentation au stator est obtenu à l'aide d'un alternostat placé sur une des phases du réseau.

Les données obtenues sur la machine alimentée par onduleur vont permettre de montrer l'efficacité des algorithmes de décision. Elles mettront aussi en évidence l'importance de l'utilisation des signatures dites 'temporelles' pour obtenir une bonne séparation des modes de fonctionnement.

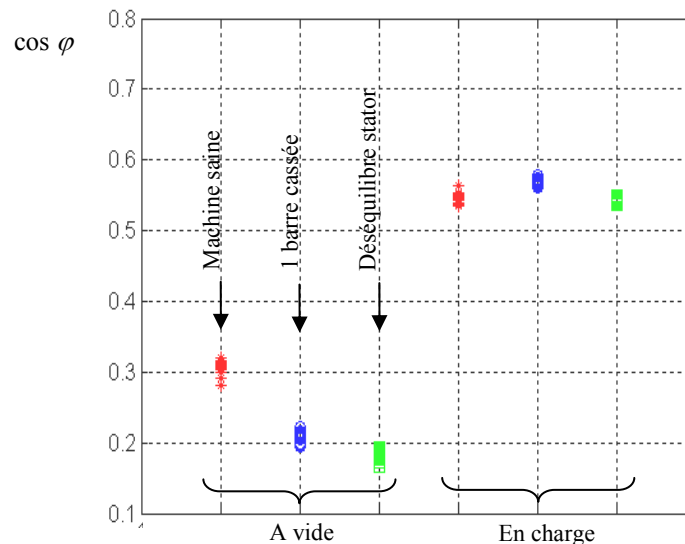
### ***4.3. Choix du vecteur forme***

Il est possible pour l'utilisateur, de choisir les paramètres constituant le vecteur forme en se basant sur les analyses de sensibilité (en fonction des différents défauts) de ces paramètres. L'approche par expertise humaine peut donner de très bons résultats comme c'est le cas pour les études [Casimir, 03a] et [Casimir, 03b].

Notre étude se situe dans le prolongement de l'étude réalisée par [Casimir, 03] au sein du laboratoire CEGELY de Lyon. De ce fait, nous avons choisi comme paramètres représentatifs des défaillances possibles sur la machine, ceux utilisés par cet auteur. Ils sont aux nombres de 35, 27 sont issus d'une analyse fréquentielle et 8 sont des paramètres statistiques. Ces paramètres sont tous décrits dans le chapitre II et résumés dans le tableau 4.3.

A cet ensemble de paramètres, nous ajoutons trois autres paramètres :

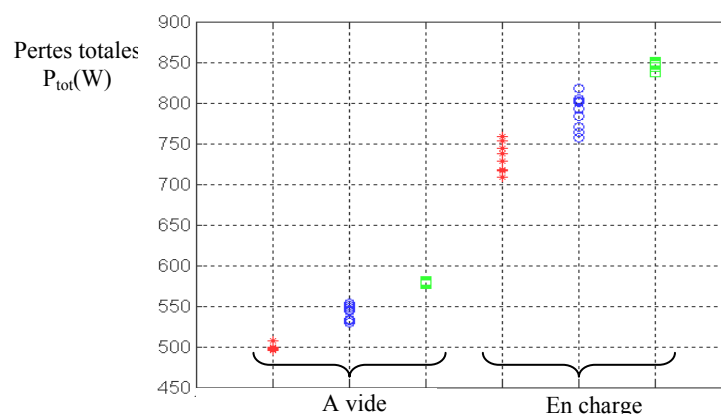
- le facteur de puissance  $\cos\varphi$  qui peut être un bon indicateur pour la détection d'excentricité au rotor [Chang, 94]. Une étude de sensibilité de ce paramètre permet de représenter son évolution en fonction de différents modes de fonctionnement (figure 4.3).



**Fig. 4.3.** Variations du facteur de puissance. Machine saine (\*). Rotor en défaut 1 barre cassée (o). déséquilibre stator de 10 % (□).

Sur l'ensemble des acquisitions, à vide, le facteur de puissance fournit de bons résultats, à savoir une bonne distinction entre le fonctionnement sain, rotor en défaut (1 barre cassée) et aussi pour le déséquilibre d'alimentation. En revanche en charge, cette distinction n'est plus très marquée.

- Les pertes totales de la machine  $P_{\text{tot}}$ , calculées à partir des mesures de courants, de tensions et d'un essai à vide (chapitre II, § 2.2.2.5 et annexe 1).

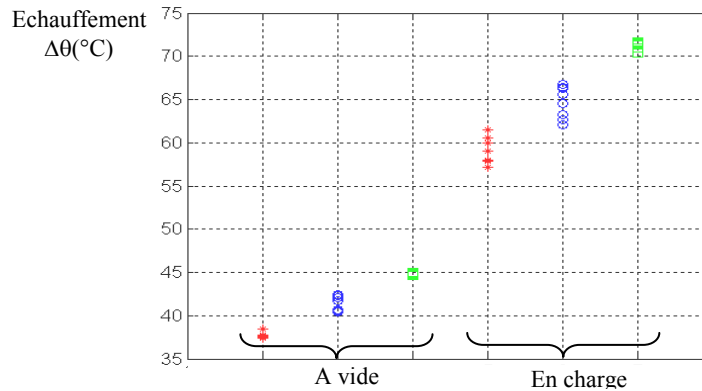


**Fig. 4.4.** Variations des pertes totales. Machine saine (\*). Rotor en défaut 1 barre cassée (o). déséquilibre stator de 10 % (□).

La première observation qui peut être faite, est que le paramètre  $P_{\text{tot}}$  est sensible aux différents états de la machine. Sur l'ensemble des acquisitions, à vide, il permet une bonne discrimination du déséquilibre d'alimentation, qui plus est, de faible amplitude (10%) et du rotor avec 1 barre cassée par rapport au fonctionnement sain. En charge, la constatation est la

même, avec toutefois, une plus forte dispersion pour les fonctionnements sain et rotor en défaut.

- L'échauffement au niveau de l'induit de la machine  $\Delta\theta$ , déterminé par la modélisation thermique de celle-ci par un modèle un corps (chapitre II, § 2.2.2.5 et annexe 1).



**Fig. 4.5.** Variations de l'échauffement. Machine saine (\*). Rotor en défaut 1 barre cassée (o). déséquilibre stator de 10 % (□).

Pour ce paramètre  $\Delta\theta$ , les mêmes constatations que précédemment peuvent être faites. A vide et en charge, malgré le faible niveau de déséquilibre d'alimentation sur une phase (10%), la distinction entre ce défaut et le fonctionnement sain est assez nette. Il en est de même pour le défaut rotor, bien différencié avec le fonctionnement sain et le déséquilibre au stator.

De plus, nous avons vu précédemment que lorsque la machine est alimentée par un onduleur cela entraîne l'apparition d'harmonique qui ont pour pulsation  $h * f_s$ , avec  $h = 5, 7, 11, 13, \dots$  ou plus généralement  $(6 * k \pm 1) f_s$ .

Il semble donc intéressant d'intégrer certaines de ces fréquences dans notre vecteur forme initial et de regarder la sensibilité de celles-ci au travers d'une analyse de leur sensibilité.

### 4.3.1. Analyse de la sensibilité des composantes fréquentielles

Les fréquences engendrées par l'alimentation onduleur et rajoutées dans le vecteur forme initial sont résumées dans le tableau 4.2.

Fréquences	Dénomination
$5f_s$	Harmonique 5 du fondamental
$7f_s$	Harmonique 7 du fondamental
$9f_s$	Harmonique 9 du fondamental
$11f_s$	Harmonique 11 du fondamental
$13f_s$	Harmonique 13 du fondamental
$17f_s$	Harmonique 17 du fondamental
$19f_s$	Harmonique 19 du fondamental
$23f_s$	Harmonique 23 du fondamental
$25f_s$	Harmonique 25 du fondamental
$29f_s$	Harmonique 29 du fondamental

**Tableau. 4. 2.** *Fréquences contenues dans le vecteur forme et induites par l'alimentation onduleur*

Au total, 36 amplitudes de raies sont calculées pour les paramètres fréquentiels. Elles sont décrites dans le tableau 4.3.

Afin d'analyser la sensibilité de ces fréquences, nous avons utilisé trois modes de fonctionnement différents pour une machine alimentée par onduleur :

- Machine saine, à vide et en charge,
- Machine avec 4 barres cassées au rotor, à vide et en charge,
- Machine avec un déséquilibre d'alimentation (10%), à vide et en charge.

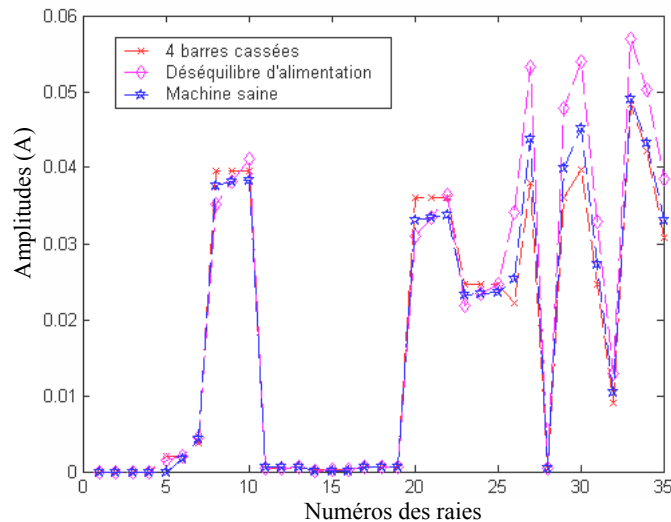
Sur les figures 4.6 et 4.7, les amplitudes des raies correspondent à une moyenne des 15 acquisitions réalisées par mode de fonctionnement. L'analyse fréquentielle (FFT) est réalisée sur la partie réelle de la composante  $\underline{I}_{sd}$  et elle est calculée sur les 100.000 points du signal.

Les fréquences reprises sur les figures 4.6 et 4.7 sont classées dans l'ordre suivant :

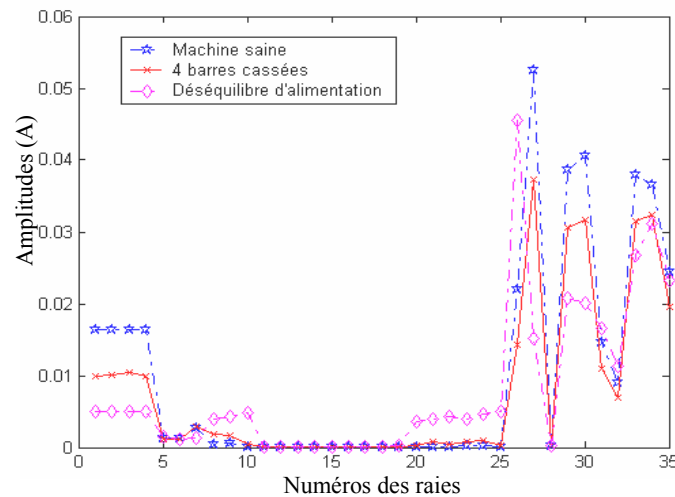
- Les quatre premières valeurs sont les moyennes des amplitudes des raies de modulation autour du fondamental, rangées dans l'ordre croissant d'apparition des fréquences, notées  $\alpha_4$ ,  $\alpha_5$ ,  $\alpha_6$  et  $\alpha_7$  dans le tableau 4.3,
- Les valeurs 5, 6, et 7 correspondent respectivement aux fréquences suivantes :  $f_s - f_r$ ,  $f_s + f_r$  et  $3 * f_s$ , notées  $\alpha_9$ ,  $\alpha_{10}$  et  $\alpha_{11}$  dans le tableau 4.3,



- Les valeurs 8 à 25 correspondent aux fréquences d'encoches rangées dans l'ordre croissant d'apparition des fréquences, notées  $\alpha_{11}$  à  $\alpha_{28}$  dans le tableau 4.3,
- Enfin les valeurs 26 à 35 correspondent aux harmoniques créés par l'alimentation par convertisseur, notées  $\alpha_{29}$  à  $\alpha_{38}$  dans le tableau 4.3.



**Fig. 4.6.** Variation des raies correspondant aux fréquences recherchées. Fonctionnement à vide



**Fig. 4.7.** Variation des raies correspondant aux fréquences recherchées. Fonctionnement en charge

Les amplitudes des raies 1 à 25 varient très peu selon les modes de fonctionnement à vide. Par contre en charge, l'analyse des composantes à  $f_s(1 \pm 2 \cdot n \cdot g)$  (raies 1 à 4) fournit un résultat net du point de vue de l'identification du défaut rotor et du déséquilibre d'alimentation.

Par contre, les fréquences d'encoches (raies 5 à 25) ne permettent pas de différencier efficacement les différents modes en charge.

Avec une alimentation par onduleur, les raies les plus discriminantes sont les raies 26 à 35. Elles correspondent aux harmoniques créées par cette alimentation. En effet, à vide on remarque que les raies 26, 27, 29, 30 et 31 respectivement  $5 \cdot f_s$ ,  $7 \cdot f_s$ ,  $11 \cdot f_s$ ,  $13 \cdot f_s$ ,  $17 \cdot f_s$  permettent de différencier le défaut stator du défaut rotor mais aussi de la machine saine. Ces raies permettent aussi la même différenciation en charge. En revanche les raies 33, 34 et 35 respectivement  $23f_s$ ,  $25f_s$  et  $29f_s$  ne permettent de différencier que le fonctionnement sain par rapport au déséquilibre d'alimentation aussi bien à vide qu'en charge.

Il est donc intéressant de conserver ces raies dans le vecteur forme initial car l'analyse des résultats a mis en évidence la possibilité de différencier les différents types de défaut aussi bien à vide qu'en charge.

De plus, pour ces méthodes de diagnostic à base de RdF, il est primordial de disposer d'un vecteur forme constitué d'un nombre important de paramètres représentatifs du plus grand nombre de modes de fonctionnement possibles (sain et défectueux). Cela permet d'obtenir, par la suite, une meilleure discrimination entre les différentes classes.

Ce grand nombre de paramètres n'est pas pénalisant en termes de temps de calcul. En effet, une étape de sélection de paramètres est réalisée, permettant ainsi de choisir seulement les paramètres représentatifs des modes de fonctionnement contenus dans l'ensemble d'apprentissage initial.

$\alpha_i$	Dénomination		$\alpha_i$	Dénomination	
1	$m_p$	Puissance active moyenne	17	$N_r f_r - 2f_s - 2g f_s$	Amplitudes autour de $N_r f_r \pm 2f_s$
2	$m_q$	Puissance réactive moyenne	18	$N_r f_r - 2f_s$	
3	$\zeta$	Dispersion des points de la caractéristique $Q'(P')$	19	$N_r f_r - 2f_s + 2g f_s$	
4	$f_s(1-4g)$	Amplitudes autour de $f_s$	20	$N_r f_r + 2f_s - 2g f_s$	
5	$f_s(1-2g)$		21	$N_r f_r + 2f_s$	
6	$f_s(1+2g)$		22	$N_r f_r + 2f_s + 2g f_s$	
7	$f_s(1+4g)$		23	$N_r f_r - 3f_s - 2g f_s$	Amplitudes autour de $N_r f_r \pm 3f_s$
8	$f_s - f_r$		24	$N_r f_r - 3f_s$	
9	$f_s + f_r$		25	$N_r f_r - 3f_s + 2g f_s$	
10	$3f_s$		Harmonique 3 du fondamental	26	
11	$N_r f_r - f_s(1+2g)$	Amplitudes autour de $N_r f_r \pm f_s$	27	$N_r f_r + 3f_s$	
12	$N_r f_r - f_s$		28	$N_r f_r + 3f_s + 2g f_s$	
13	$N_r f_r - f_s(1-2g)$		29	$5f_s$	Harmonique 5 du fondamental
14	$N_r f_r + f_s(1-2g)$		30	$7f_s$	Harmonique 7 du fondamental
15	$N_r f_r + f_s$		31	$9f_s$	Harmonique 9 du fondamental
16	$N_r f_r + f_s(1+2g)$		32	$11f_s$	Harmonique 11 du fondamental

$\alpha_i$	Dénomination		$\alpha_i$	Dénomination	
33	$13f_s$	Harmonique 13 du fondamental	46	$\cos \varphi$	Déphasage entre la tension et le courant d'une phase
34	$17f_s$	Harmonique 17 du fondamental	47	$P_{tot}$	Pertes totales de la machine
35	$19f_s$	Harmonique 19 du fondamental			
36	$23f_s$	Harmonique 23 du fondamental	48	$\Delta\theta$	Echauffement au niveau de l'induit de la machine
37	$25f_s$	Harmonique 25 du fondamental			
38	$29f_s$	Harmonique 29 du fondamental			
39	$P_{1d}$	Puissance moyenne de la composante directe des courants			
40	$\delta_\alpha$	Déformation de la caractéristique $I_{s\beta}(I_{s\alpha})$ dans l'axe $\alpha$			
41	$\delta_\beta$	Déformation de la caractéristique $I_{s\beta}(I_{s\alpha})$ dans l'axe $\beta$			
42	$\sigma_\alpha$	Ecart – type de la composante $I_{s\alpha}$ des courants			
43	$\sigma_\beta$	Ecart – type de la composante $I_{s\beta}$ des courants			
44	$\sigma_s$	Ecart – type du module de $I_s = I_{s\alpha} + jI_{s\beta}$			
45	$Z_d$	Impédance directe			

**Tableau. 4. 3.** Panorama des paramètres calculés constituant le vecteur forme initial

Au final, la signature de la machine (vecteur forme) est définie par un vecteur de 48 paramètres.

Comme le montre le tableau 4.3, la connaissance du glissement, et par conséquent, de la vitesse de rotation, est indispensable pour le calcul des paramètres fréquentiels formant le vecteur forme. Ainsi, lors d'acquisitions sur le système, il est nécessaire de disposer de capteur de vitesse (dynamo tachymétrique ou codeur). Si ce n'est pas le cas (pour des raisons de coût par exemple), il est nécessaire de pouvoir déterminer cette vitesse de façon automatique par le biais d'analyse temporelle et/ou fréquentielle du signal mesuré (en l'occurrence le courant stator).

De plus, cette procédure rentre dans le cadre de la réalisation d'un système de diagnostic quasi automatique.

### 4.3.2. Recherche automatique du glissement

Le glissement de la machine asynchrone a pour expression :

$$g = \frac{n_s - n}{n_s} \quad (4.1)$$

Avec :

$n_s$  : vitesse de synchronisme en tr/min

$n$  : vitesse de rotation de la machine en tr/min

ou bien encore :

$$g = 1 - \frac{p^* f_r}{f_s} \quad (4.2)$$

Avec :

$f_s$  : fréquence du fondamental

$f_r$  : fréquence de rotation du rotor

$p$  : nombre de paires de pôles

Ainsi, pour déterminer le glissement et donc la vitesse de rotation, il est nécessaire de rechercher automatiquement les fréquences  $f_s$  et  $f_r$ .

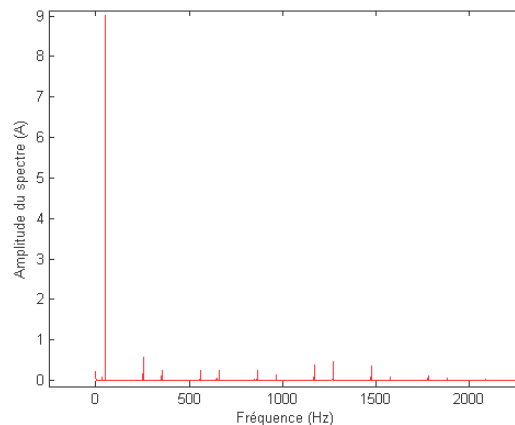
#### 4.3.2.1. Détermination de la fréquence fondamentale $f_s$

Deux approches sont possibles pour déterminer cette fréquence, soit une analyse temporelle du signal, soit une analyse fréquentielle.

Pour valider notre approche, la machine est alimentée par onduleur avec un déséquilibre d'alimentation sur une phase (10%) en charge. Ce choix d'alimentation onduleur correspond à un cas très défavorable (signal perturbé). Le signal étudié est le courant de phase  $i_a$ .

#### 4.3.2.1.1. Analyse fréquentielle

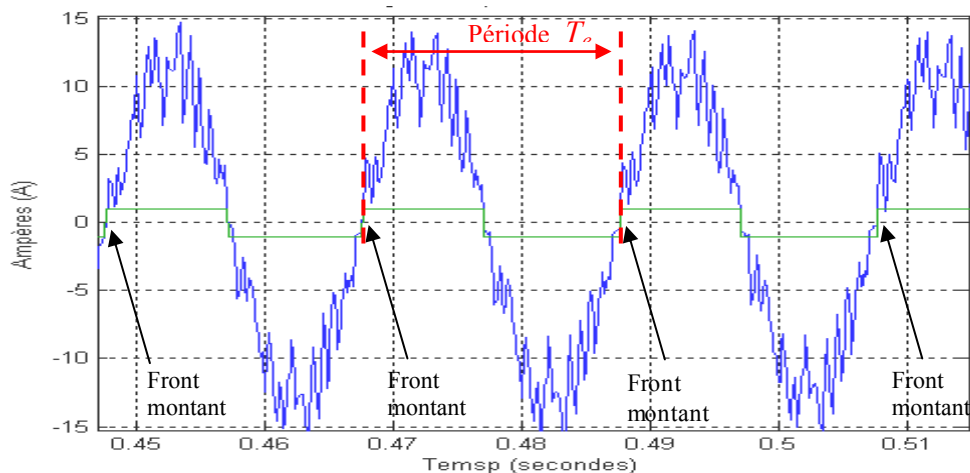
A partir d'une FFT réalisé sur le signal  $i_a$  (figure 4.8), l'amplitude la plus importante, qui correspond à la fréquence fondamentale du signal, est recherchée.



**Fig. 4.8.** Spectre du signal  $i_a$

#### 4.3.2.1.2. Analyse temporelle

A partir du signal mesuré  $i_a$ , la période  $T_e$  est déterminée et ainsi sa fréquence est en déduite. Pour se faire, une détection du passage à zéro du signal est opérée.

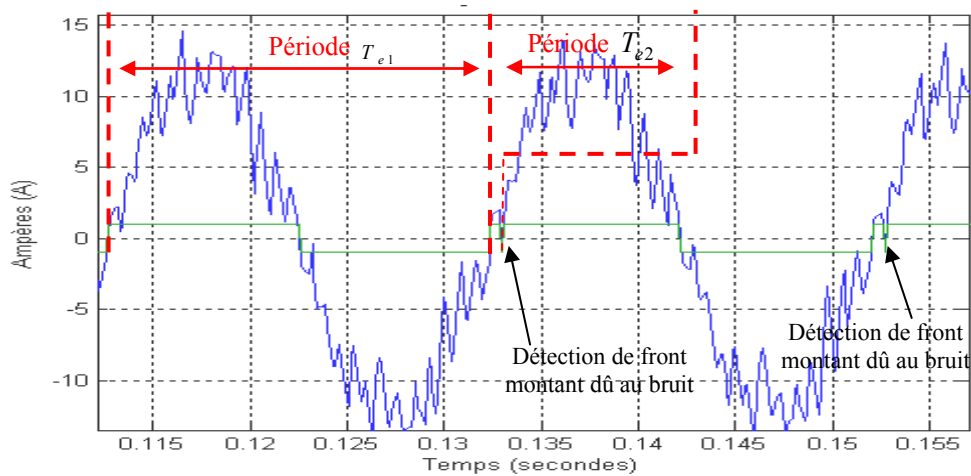


**Fig. 4.9.** Signal temporel du courant  $i_a$

Chaque front montant successif (figure 4.9) est détecté, ce qui nous permet d'obtenir la

période  $T_e$  et donc la fréquence  $f_s = \frac{1}{T_e}$ .

Mais si le signal est fortement bruité, notamment par la présence d'un onduleur, la détection du passage à zéro peut être faussée sur certaines périodes comme le montre la figure 4.10.



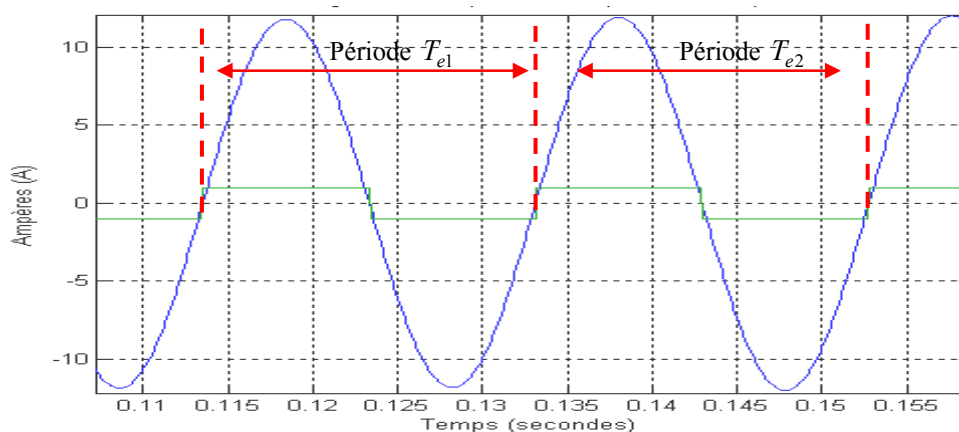
**Fig. 4.10.** Signal temporel du courant  $i_a$  bruité

Sur la figure 4.10, des fronts montants sont détectés par erreur, dû au bruit du signal, ce qui par la suite va fausser le résultat du calcul de fréquence. La période  $T_{e1}$  est correcte mais la période  $T_{e2}$ , du fait de la présence d'un front montant dû au bruit, est erronée.

Pour éviter ces erreurs de calcul, le filtrage du signal autour de la composante fondamentale est nécessaire. Pour cela, il convient de réaliser une FFT pour déterminer la fréquence fondamentale et ensuite filtrer, à l'aide d'un filtre passe bande autour de cette fréquence.

Par conséquent, l'analyse temporelle lorsque la machine est alimentée par onduleur, n'a que peu d'intérêt. Il faut, en effet, déterminer au préalable la fréquence fondamentale par FFT (qui est le but final). Elle peut, cependant, améliorer la précision en utilisant des techniques de moyennage. Par contre cette analyse temporelle reste tout à fait valide avec une alimentation classique par le réseau.

La figure 4.11 représente le signal mesuré  $i_a$  filtré avec ce filtre passe-bande.



**Fig. 4.11.** Signal temporel du courant  $i_a$  filtré par un filtre passe bande

Il devient alors très facile de déterminer la valeur exacte de la période et ainsi en déduire la fréquence fondamentale  $f_s$  du signal avec une grande précision.

#### 4.3.2.2. Calcul de la fréquence de rotation du rotor $f_r$

Connaissant  $f_s$ , il reste à déterminer la fréquence de rotation  $f_r$  qui a pour expression :

$$f_r = \frac{(1-g) * f_s}{p} \quad (4.3)$$

Par conséquent, pour déterminer  $f_r$ , il faut donc connaître au préalable la valeur du glissement  $g$ .

Pour résoudre ce problème, nous avons choisi de rechercher la valeur de la fréquence caractéristique  $f_s + f_r$  sur le spectre de  $i_a$ .

Pour retrouver la valeur de cette fréquence caractéristique, il est nécessaire de connaître une valeur approximative de  $f_r$  et, de ce fait, pouvoir définir une plage de variation autour de cette fréquence. Nous allons prendre une valeur de  $g$  maximale et minimale.

$$\begin{aligned} g_{\min} &= 0 \\ g_{\max} &= 0.1 \end{aligned} \quad (4.4)$$

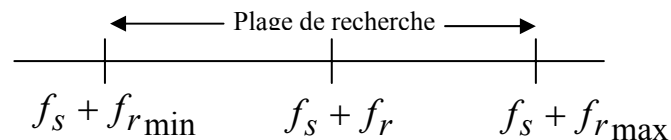
Ainsi nous pouvons calculer une valeur maximale et minimale de  $f_r$ .

$$f_{r_{\max}} = \frac{(1-g_{\min}) * f_s}{p} = \frac{f_s}{p} \quad (4.5)$$

$$f_{r_{\min}} = \frac{(1-g_{\max}) * f_s}{p} = 0.9 * \frac{f_s}{p} \quad (4.6)$$

avec  $p = 2$  paires de pôles dans notre application

Avec ces valeurs déterminées précédemment, nous disposons d'une plage de fréquence pour déterminer la fréquence caractéristique  $f_s + f_r$  sur le spectre du signal mesuré  $i_a$ .



Dans le cas du fonctionnement étudié ci-dessus, après avoir déterminé la valeur de  $f_s = 50.877 \text{ Hz}$ , les valeurs min et max de  $f_r$  sont calculées à partir des expressions 4.5 et 4.6 :



$$fr_{\max} = \frac{50.877}{2} = 25.438 \text{ Hz}$$

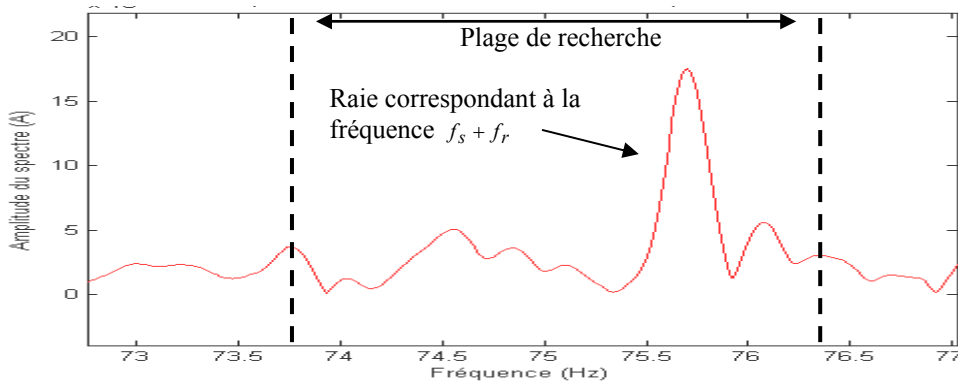
$$fr_{\min} = 0.9 * \frac{50.877}{2} = 22.894 \text{ Hz}$$

Ensuite, l'intervalle de recherche pour déterminer la fréquence caractéristique  $f_s + f_r$  est défini

$$fs + fr_{\min} = 50.877 + 22.894 = 73.771 \text{ Hz} \quad (4.7)$$

$$fs + fr_{\max} = 50.877 + 25.438 = 76.315 \text{ Hz}$$

La raie qui a la plus grande amplitude dans cette plage de fréquence, est alors recherchée sur le spectre :



**Fig. 4. 12.** Spectre de  $i_a$  autour de la fréquence caractéristique  $f_s+f_r$

La valeur exacte de la fréquence caractéristique trouvée automatiquement sur le spectre est :

$$f_s + f_r = 75.7027 \text{ Hz}$$

On peut alors en déduire la valeur de  $f_r$  :

$$fr = 75.7027 - fs = 75.7027 - 50.877$$

$$fr = 24.8257 \text{ Hz}$$

#### 4.3.2.3. Calcul du glissement $g$

$f_r$  et  $f_s$  étant connues, le glissement peut être déduit d'après l'équation 4.2:

$$g = 1 - \frac{p * f_r}{f_s} \quad (4.8)$$

$$A.N : g = 0.024 = 2.4\%$$

Pour vérifier l'exactitude du résultat trouvé, il suffit de calculer la vitesse de rotation du moteur à partir de la valeur du glissement obtenue, et ensuite de la comparer à la valeur de la vitesse donnée par un capteur disposé sur le système lors de l'acquisition.

#### 4.3.2.4. Calcul de la vitesse de rotation du moteur 'n'

D'après l'expression générale suivante, on peut déterminer  $n$  :

$$g = \frac{n_s - n}{n_s} \quad (4.9)$$

avec  $n_s$  : vitesse de synchronisme du moteur

La vitesse de synchronisme du moteur  $n_s$  est calculée à partir de l'expression suivante :

$$n_s = \frac{f_s}{p} \quad (4.10)$$

$$A.N : n_s = \frac{50.877}{2} = 25.4385 \text{ tr/s}$$

$$n_s = 1526.3 \text{ tr/min}$$

Au final, la vitesse de rotation du rotor est calculée :

$$n = (1 - g) * n_s \quad (4.11)$$

$$A.N : n = (1 - 0.024) * 1526.3$$

$$n = 1489.66 \text{ tr/min}$$

La vitesse réelle mesurée par la dynamo tachymétrique est 1493.9 tr/min, ce qui donne un taux d'erreur avec l'approche présentée, de 0.29%.

Grâce à ce calcul du glissement, l'ensemble des paramètres formant le vecteur forme initial peut être déterminé.

Cependant, quelques-uns de ces paramètres risquent d'être non pertinents ou redondants selon l'ensemble d'apprentissage à définir. La dimension de l'espace est donc réduite grâce à la sélection de paramètres. Cette sélection de caractères rend la recherche de l'espace de représentation plus rapide et plus globale.

Des algorithmes de sélection ont été présentés au cours du chapitre précédent. Dans le cas du *mode supervisé*, deux d'entre eux seront utilisés pour définir la signature caractérisant l'ensemble d'apprentissage. Il s'agit de la méthode dite "Sequential Backward Selection" (SBS) et de la méthode à base d'algorithme génétique notée "AG". Le rapport de Fisher

(équation 3.16, chapitre III) sera utilisé pour déceler le pouvoir discriminant de chaque paramètre.

Les algorithmes de sélection sont chargés de rechercher les paramètres les plus pertinents conformément à la dimension de l'espace souhaitée. Cette dimension peut être fixée, soit par l'utilisateur, soit de manière automatique comme cela est présenté par la suite.

Néanmoins, le vecteur forme doit permettre de distinguer clairement les différentes classes. La sélection sera validée par une procédure de test.

Les paramètres de départ ne sont pas homogènes. La sélection des paramètres risque d'être biaisée si l'ensemble d'apprentissage n'est pas homogénéisé. Pour cela, les données de l'ensemble d'apprentissage doivent être de valeurs moyennes nulles et d'écart type égal à 1. Nous n'utiliserons que des données 'centrées – réduites' par la suite.

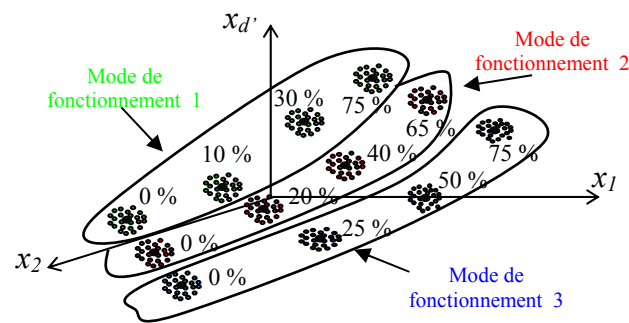
Il semble intéressant aussi d'étudier l'influence du niveau de charge sur l'évolution des paramètres. En effet, si des paramètres comme la puissance active ou réactive ou bien encore le facteur de puissance, permettent de détecter l'apparition de défaillances, ils sont aussi sensibles aux variations de la charge. De ce fait, il pourrait devenir difficile de différencier ces deux aspects.

C'est pour ces raisons que dans le paragraphe suivant, la normalisation des données par rapport à la charge est développée. Cette normalisation va permettre de s'affranchir du niveau de charge et ainsi, seule la variation due à des défaillances restera présente.

#### ***4.4. Normalisation des données***

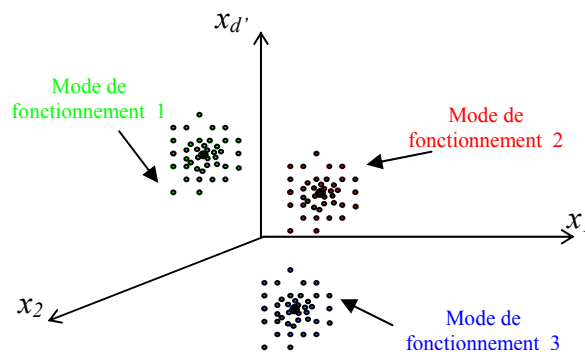
Pour rendre plus simple et plus efficace la phase de décision, la normalisation des données initiales est réalisée. En effet, avant cette normalisation, un mode de fonctionnement présente une grande dispersion de ses vecteurs, due à la variation de charge. Ce mode de fonctionnement est représenté par plusieurs classes dans  $X_a$  [Ondel, 05 a], [Casimir, 03]. Ces différentes classes correspondent à plusieurs niveaux de charge de la machine. En normalisant, c'est-à-dire en s'affranchissant de ce niveau de charge, un mode de fonctionnement est représenté par une seule classe regroupant toutes les variations de charge.

Schématisons simplement cette approche :



**Fig. 4. 13.** Représentation schématique de modes de fonctionnement dépendant du niveau de charge

Sur la figure 4.13 trois différents modes de fonctionnement sont représentés. Chaque mode est composé de quatre classes correspondant à 4 niveaux de charge. Le but est d’obtenir la configuration suivante (figure 4.14), en s’affranchissant du niveau de charge :



**Fig. 4. 14.** Représentation schématique de modes de fonctionnement indépendant du niveau de charge

Pour cela, les données sont normalisées à partir de la valeur de l’amplitude fondamentale de chaque signal mesuré. Ainsi, les signaux suivants sont obtenus, et vont être utilisés pour le calcul des paramètres (signatures des défauts) :

- Trois courants normalisés

$$Ia_n = \frac{Ia}{Ia_1} \quad ; \quad Ib_n = \frac{Ib}{Ib_1} \quad ; \quad Ic_n = \frac{Ic}{Ic_1} \tag{4.12}$$

où l’indice 1 représente la valeur de 1<sup>er</sup> harmonique.

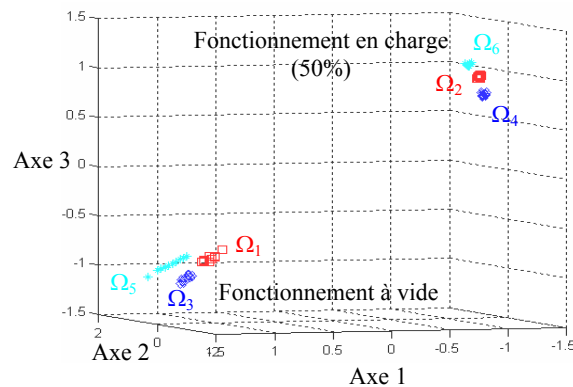
- Trois tensions normalisées

$$Va_n = \frac{Va}{Va_1} \quad ; \quad Vb_n = \frac{Vb}{Vb_1} \quad ; \quad Vc_n = \frac{Vc}{Vc_1} \tag{4.13}$$

où l’indice 1 représente la valeur de 1<sup>er</sup> harmonique.

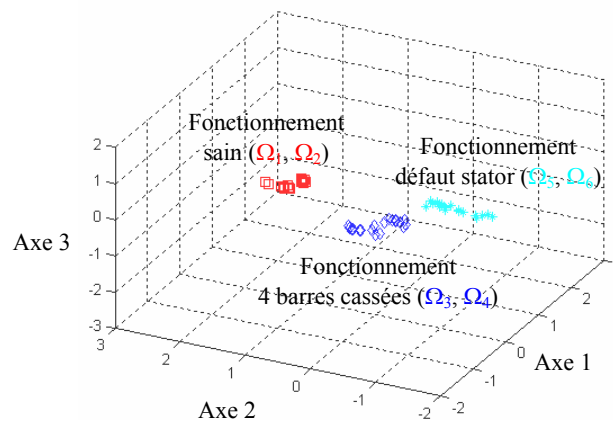
Cette normalisation peut permettre une amélioration du diagnostic. Tous les niveaux de charge (de 0 à 100%) pour chaque mode de fonctionnement ne peuvent être caractérisés, ce qui sous-entend que l'ensemble d'apprentissage  $X_a$  ne peut pas être exhaustif. Or, quand une nouvelle mesure est réalisée sur le système pour un niveau de charge non défini pour ce mode, alors cette mesure peut être rejetée par la règle de décision. Avec la normalisation, quelque soit le niveau de charge, ce mode de fonctionnement est seulement représenté par une classe, la nouvelle mesure pourra être affectée et non pas rejetée.

Les figures 4.15 et 4.16 présentent des données issues de la machine asynchrone dans le cas normalisées et non normalisées pour un fonctionnement à vide et à 50% de charge.



**Fig. 4.15.** Données non normalisées. Machine saine «  $\square$  », 4 barres cassées «  $\blacklozenge$  », déséquilibre stator (10%) «  $*$  »

Sur la figure 4.15, la dispersion la plus importante est due à l'évolution du niveau de charge (de 0 à 50%) et non à la variation des états (sain ou défectueux) de la machine.



**Fig. 4.16.** Données normalisées. Machine saine «  $\square$  », 4 barres cassées «  $\blacklozenge$  », déséquilibre stator (10%) «  $*$  »

Sur la figure 4.16, l'effet de la normalisation des données est nettement visible. La dispersion due à la variation du niveau de charge n'existe quasiment plus. En revanche la dispersion entre les différents états de la machine est restée intacte.

A présent, grâce à cette normalisation, la seule évolution possible entre les classes est due aux différents types de défauts possibles et à l'évolution de leur gravité (sévérité).

Ainsi, l'approche présentée dans le chapitre III (§. 3.6) peut être utilisée, non plus pour suivre et prédire l'évolution des différents fonctionnements en fonction du niveau de charge, mais pour suivre et faire un pronostic sur l'évolution de la sévérité des défaillances. Ceci est présenté dans le paragraphe suivant.

Il est important de noter, que pour la suite de l'étude, les données utilisées sont de valeurs moyennes nulles et d'écart type égal à 1 (homogénéisation des paramètres du vecteur forme), c'est-à-dire des données 'centrées – réduites', et sont normalisées par rapport à la charge.

#### 4.4.1. Constitution des classes

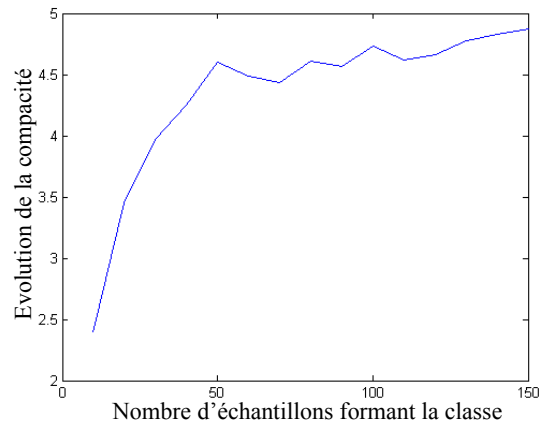
Lors de nos expérimentations sur la machine asynchrone, pour chaque mode de fonctionnement, 15 acquisitions ont été réalisées pour différents niveaux de charge. 10 de ces acquisitions sont utilisées pour constituer la classe caractérisant le mode correspondant, et les 5 restantes sont utilisées pour constituer l'ensemble test servant à valider le diagnostic.

Pour définir précisément les caractéristiques statistiques de chaque classe, il est nécessaire de prendre un nombre plus important d'échantillons. Or pour des raisons de temps, de capacité mémoire mais aussi à cause du faible retour d'expérience mis à disposition sur ces systèmes, il est difficile d'obtenir un grand nombre de données.

Ainsi, pour caractériser correctement nos classes, nous avons réalisé, pour une machine saine (sans défauts) à vide, 150 acquisitions (environ 2h30 pour réaliser les 150 mesures). On supposera, toujours pour des raisons de temps, que la dispersion (compacité des classes) est la même pour tous les modes de fonctionnement étudiés ainsi que pour les différents niveaux de charge.

A partir de ces acquisitions, nous avons étudié l'évolution de la dispersion des classes : il s'agit de déterminer le nombre minimal d'acquisition à partir duquel cette évolution devient stationnaire. La compacité est calculée à l'aide de la matrice de dispersion intra-classe (équation 3.9). Pour se ramener à un scalaire, le critère de la trace est utilisé (équation 3.12).

Au départ, la dispersion totale des 150 échantillons est calculée. Ensuite, à chaque itération, 10 échantillons sont enlevés aléatoirement puis cette même dispersion est recalculée et ainsi de suite jusqu'à ce que la classe soit formée par 10 échantillons. Sur la figure 4.17 est représentée l'évolution de la compacité en fonction du nombre d'échantillons formant la classe.



**Fig. 4. 17.** Evolution de la dispersion intra-classe en fonction de nombre d'échantillons

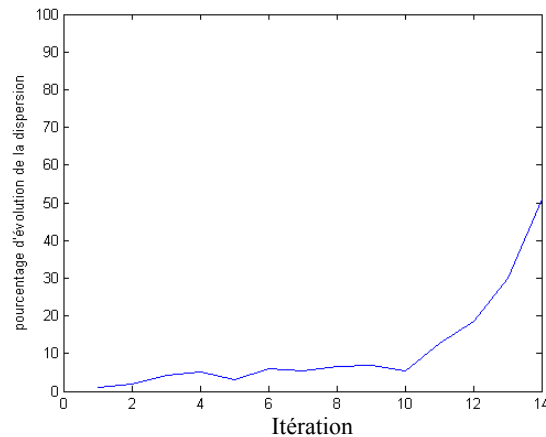
La première constatation faite à partir de la figure 4.17 est que pour un nombre d'échantillons compris entre 50 et 150, la compacité est quasi constante. En dessous de 50 échantillons par classe, cette compacité diminue fortement. Par conséquent, il semble que 50 échantillons suffisent pour caractériser correctement les propriétés de la classe.

Pour vérifier cela, le tableau suivant présente l'évolution de la dispersion entre une classe définie par 150 échantillons et toutes les autres classes définies avec un nombre décroissant d'échantillons.

Nombre d'échantillons	Evolution de la dispersion (en %)	itération
150 → 140	1	1
150 → 130	2	2
150 → 120	4.3	3
150 → 110	5.2	4
150 → 100	3	5
150 → 90	6	6
150 → 80	5.46	7
150 → 70	6.5	8
150 → 60	6.88	9
<b>150 → 50</b>	<b>5.5</b>	<b>10</b>
150 → 40	12.81	11
150 → 30	18.62	12
150 → 20	30	13
150 → 10	50.94	14

**Tableau. 4. 4.** Comparaison de la compacité des classes

Jusqu'à 50 échantillons, la dispersion est quasi stationnaire, la différence maximale étant de 7% alors qu'en dessous de 50, cette différence est croissante et atteint 51% (pour 10 échantillons). Visualisons les données du tableau 4.4 sur la figure suivante :



**Fig. 4.18.** *Différence entre les compacités des classes*

On constate, sur la figure 4.18, qu'à partir de la 10<sup>ème</sup> itération, qui correspond à la différence de dispersion entre la classe composée de 150 échantillons et la classe composée de 50 échantillons (cf. tableau 4.4), une nette rupture correspondant à une très grande variation de compacité.

Au final, pour caractériser correctement une classe, il est nécessaire que celle-ci soit au moins constituée de 50 échantillons.

Par la suite, comme nos classes ne sont constituées que de 10 échantillons, il est nécessaire de tenir compte de cette erreur faite sur la dispersion. Elle sera prise en compte dans le seuil de rejet en distance.

#### **4.4.2. Définition du seuil de rejet en distance pour des données normalisées et non normalisées**

Le seuil de rejet en distance permet la prise en compte des incertitudes liées aux classes. Dans notre cas, il est fixé à partir de la dispersion calculée pour une classe formée de 50 échantillons. Pour cela, la distance maximale est recherchée, c'est-à-dire la distance entre le centre de gravité de cette classe et l'échantillon le plus éloigné de ce centre. La métrique utilisée est la distance de Mahalanobis, pour tenir compte des variances sur chaque dimension.

$$T_{50} = \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_{50})] \quad (4.14)$$

Avec :

$T_{50}$  : distance maximale de la classe composée de 50 échantillons

$N_c$  : nombre d'échantillons formant la classe (ici  $N_c = 50$ )

$\underline{m}_{50}$  : centre de gravité de la classe



$\underline{X}_{cv}$  :  $v^{\text{ième}}$  vecteur de la classe  $\Omega_c$

Concernant la règle de décision des  $k$ -ppv exclusive (§ 3.5.1.2, chapitre III), cette distance est appliquée comme seuil de rejet en distance pour nos classes définies avec 10 échantillons, c'est-à-dire lorsque les données non normalisées par rapport à la charge sont utilisées. Concernant les données normalisées, le seuil de rejet en distance utilisé pour chaque classe est celui défini au chapitre III par l'équation 3.36

$$T = \beta \cdot \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_c)], \text{ avec } \beta = 2 \quad (4.15)$$

Avec :  $N_c$ , le nombre de vecteurs de la classe  $\Omega_c$ ,

$\underline{X}_{cv}$ , le  $v^{\text{ième}}$  vecteur de la classe  $\Omega_c$ ,

$\underline{m}_c$ , centre de gravité de la classe  $\Omega_c$ .

Le seuil de rejet en distance n'est pas le même pour des données normalisées et non normalisées (par rapport à la charge). En effet, les classes définies pour des données non normalisées ont été caractérisées pour un seul niveau de charge (en l'occurrence à vide) et il a été supposé que la dispersion était la même pour les autres niveaux de charge. Par contre, les classes définies par des données normalisées regroupent l'ensemble des niveaux de charge. Par conséquent, la dispersion de ces classes pour ces deux approches n'est pas la même. La dispersion des classes normalisées n'a pas pu être définies pour des raisons de temps. Il aurait été nécessaire d'effectuer un certain nombre d'acquisitions (150 par exemple) pour chaque niveau de charge, ensuite normaliser ces données par rapport à la charge et définir la dispersion de l'ensemble de ces acquisitions.

Pour la règle de décision des  $k$ -ppv non exclusive (§ 3.5.1.4, chapitre III), le seuil de rejet en distance est fixé par les équations 3.43 et 3.49 du chapitre III. En fonction des données, seul le paramètre  $d_{\text{seuil } c}$  (qui définit la position des points d'inflexion de la fonction d'appartenance) est modifié. Pour des données non normalisées, ce paramètre  $d_{\text{seuil } c}$  sera égal à  $T_{50}$  alors que pour des données normalisées, il sera égal à  $\max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_c)]$ . En résumé,

- pour des données non normalisées

$$d_{\text{seuil } c} = T_{50} = \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_{50})] \quad (4.16)$$

- pour des données normalisées

$$d_{\text{seuil } c} = \max_{v=1, N_c} [d(\underline{X}_{cv}, \underline{m}_c)] \quad (4.17)$$

## 4.5. Diagnostic par RdF d'une machine alimentée par le réseau

La section suivante présente la sélection des paramètres en fonction de la dimension désirée et de la méthode utilisée. Afin d'illustrer les résultats, les algorithmes de sélection sont appliqués, dans un premier temps, aux données recueillies sur les modes de fonctionnement en alimentation réseau et définis dans le tableau suivant (tableau 4.1) :

Classe	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_1$	Machine saine	0 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_2$	1 barre cassée	0 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_3$	Déséquilibre d'alimentation au stator (10%)	0 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_4$	Usure des roulements	0 %	10
		50 %	10
		75 %	10
		100 %	10

**Tableau. 4. 5.** Composition de l'ensemble d'apprentissage  $X_{al}$  dans le cas d'une machine alimentée par le réseau

L'ensemble d'apprentissage  $X_{al}$  est donc constitué de ces quatre modes de fonctionnement représentés par quatre classes dans l'espace de représentation. Chaque classe est composée de 10 échantillons par niveau de charge, ce qui fait au final, que les classes sont composées de 40 échantillons.

Les ensembles test utilisés pour définir, par la suite, le meilleur vecteur forme sont les suivants :

Ensembles test	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$X_{t1}$	Machine saine	0 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{t2}$	1 barre cassée	0 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{t3}$	Déséquilibre d'alimentation au stator (10%)	0 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{t4}$	Usure des roulements	0 %	5
		50 %	5
		75 %	5
		100 %	5

**Tableau. 4. 6.** Présentation des différents ensembles test dans le cas d'une machine alimentée par le réseau

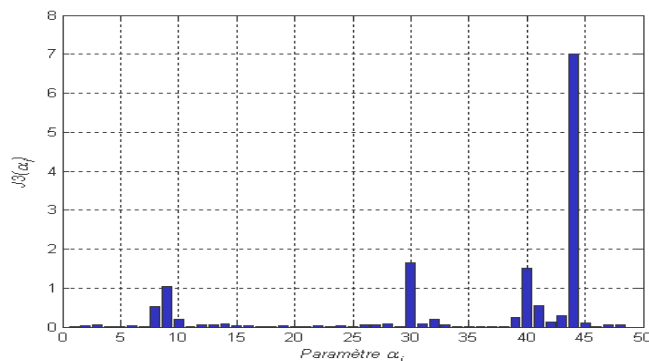
Chaque ensemble test est composé de 5 échantillons par niveau de charge, ce qui fait au final que ces ensembles sont constitués de 20 échantillons.

#### 4.5.1. Réduction de l'espace de représentation (sélection des $d'$ paramètres)

##### 4.5.1.1. Dans le cas du mode *supervisé*

Avant d'utiliser les méthodes de sélection, il est intéressant de déterminer le pouvoir de discrimination de chaque paramètre du vecteur forme par rapport à l'ensemble d'apprentissage  $X_{al}$ . Pour cela le rapport de Fisher (équation 3.16, chapitre III) est calculé pour chaque paramètre. Un paramètre  $\alpha$  aura un pouvoir de séparation d'autant meilleur que le rapport de Fisher qui lui est associé est grand.

Afin d'illustrer le pouvoir de chacun, l'histogramme suivant montre la variation du critère pour les 48 paramètres calculés.



**Fig. 4. 19.** Variation du rapport de Fisher pour chaque paramètre pour des données normalisées

Le critère de Fisher montre qu'il y a 6 paramètres  $[\alpha_8, \alpha_9, \alpha_{30}, \alpha_{40}, \alpha_{41}, \alpha_{44}]$  qui, pris séparément, ont un pouvoir discriminant. La figure 4.19 montre clairement que les paramètres  $\alpha_{44}$ ,  $\alpha_{30}$  et  $\alpha_{40}$  ont les pouvoirs discriminants les plus élevés. Mais ce résultat ne signifie pas pour autant qu'il faille limiter la taille du vecteur forme à trois. En effet, celui – ci doit être défini sur la base d'un apport collectif des paramètres discriminants. Quelques paramètres comme  $\alpha_{10}$ ,  $\alpha_{32}$ ,  $\alpha_{39}$  ou encore  $\alpha_{43}$  présentent des pouvoirs de discrimination mais bien plus faible que ceux apportés par les 6 paramètres décrits précédemment.

Donc d'après le résultat de la figure 4.19, la dimension du vecteur forme et donc de l'espace de représentation, peut aller de 2 à 6. Au vue des résultats de la figure 4.19, les méthodes de sélection SBS ou AG devraient sélectionner tout ou partie des paramètres  $[\alpha_8, \alpha_9, \alpha_{30}, \alpha_{40}, \alpha_{41}, \alpha_{44}]$  pour obtenir la meilleure séparation des classes possible.

#### 4.5.1.1.1. Sélection par la méthode SBS

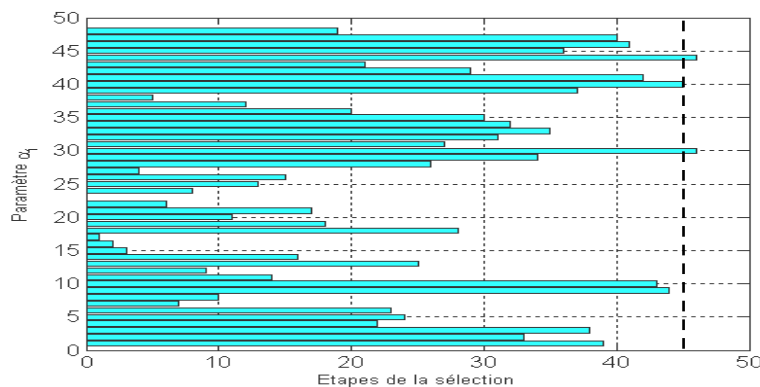
L'algorithme SBS utilise le calcul d'un critère sur des ensembles de paramètres.

Le critère utilisé est la trace des valeurs propres de  $\sum_w^{-1} \cdot \sum_B$ . Il représente un indicateur du rapport de la dispersion inter-classes à la dispersion intra-classe.

$$J = \text{trace}\left(\sum_w^{-1} \cdot \sum_B\right) \quad (4.18)$$

L'espace de départ est l'ensemble des paramètres calculés. A chaque étape de la procédure, le sous – espace qui maximise le critère est pris en considération. Il ne s'agit pas d'éliminer le paramètre le moins pertinent, mais de conserver le sous – ensemble qui caractérise le mieux l'espace de représentation en accord avec la dimension désirée. Pour cela, à chaque étape k de la sélection, le sous – espace  $E_k$ , pris en considération pour l'étape suivante, est celui pour lequel le critère est optimal (équation 3.20, chapitre III).

La figure 4.20 illustre la réalisation de la sélection de sous – ensembles des paramètres les plus pertinents à chaque étape de la sélection.



**Fig. 4. 20.** Procédure de sélection des paramètres par la méthode SBS

A chaque étape  $k$  de la sélection, un vecteur forme est choisi, constitué de  $d-k$  paramètres. Pour définir quel est le meilleur vecteur forme, et de ce fait, choisir la dimension de l'espace de représentation, on analyse celui qui donnera les meilleures performances au sens du nombre d'échantillons test mal classés.

Le tableau 4.7 donne le nombre d'échantillons mal classés pour des configurations de vecteur forme possibles obtenues d'après les résultats de la figure 4.20. Cette analyse a été réalisée à l'aide de la règle de décision des '5 - ppv' exclusive. Le seuil de rejet en distance utilisé est celui défini par l'équation 4.15. D'autre part, le nombre minimal de voisin  $k'$  que doit avoir un échantillon test dans la classe la mieux représentée est fixé à  $k/2$ . Concrètement, étant donné que  $k$  a été pris égal à 5, le nombre minimal de voisin est 3. Le choix de prendre  $k = 5$  est arbitraire. Une étude de sensibilité sera réalisée par la suite pour déterminer précisément le choix optimal de  $k$ .

Paramètres du vecteur forme	Nombre de mal classés
$[\alpha_{30}, \alpha_{44}]$	14
$[\alpha_{30}, \alpha_{40}, \alpha_{44}]$	0
$[\alpha_9, \alpha_{30}, \alpha_{40}, \alpha_{44}]$	0
$[\alpha_9, \alpha_{10}, \alpha_{30}, \alpha_{40}, \alpha_{44}]$	0
$[\alpha_9, \alpha_{10}, \alpha_{30}, \alpha_{40}, \alpha_{41}, \alpha_{44}]$	0

**Tableau. 4. 7.** Résultats de test pour définir le vecteur forme à l'aide de l'algorithme SBS pour des données normalisées

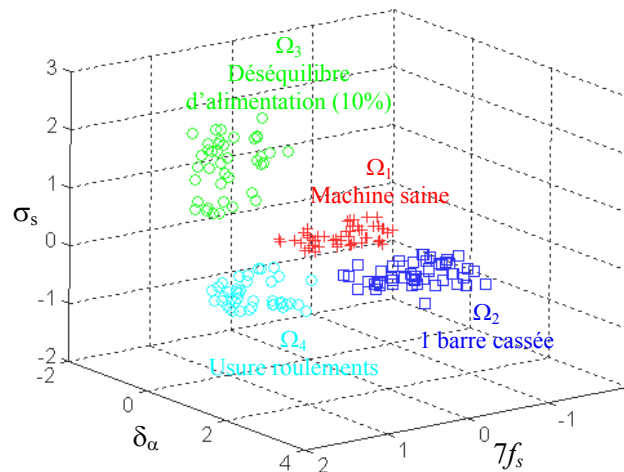
Par conséquent, la procédure de sélection semble optimale lorsque  $d' = 3$  (c'est-à-dire à la 45<sup>ème</sup> étape). Le sous - ensemble de paramètres constituant le vecteur forme optimal, noté  $V_{Fnorm(SBS)}$ , est composé de 3 paramètres :

$$V_{Fnorm(SBS)} = [\alpha_{30}, \alpha_{40}, \alpha_{44}] \quad (4.19)$$

Les paramètres de ce vecteur correspondent respectivement à :

- L'harmonique 7 du fondamental  $7f_s$ ,
- La déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\alpha$ , notée  $\delta_\alpha$ ,
- L'écart type du module de  $I_s = I_{s\alpha} + jI_{s\beta}$ , noté  $\sigma_s$ .

Les résultats obtenus avec la méthode SBS corroborent bien les résultats de l'analyse du rapport de Fisher. En effet, l'analyse de ce rapport avait permis de détecter trois paramètres avec un fort pouvoir de discrimination par rapport à l'ensemble d'apprentissage  $X_{al}$  utilisé. Ces trois paramètres,  $[\alpha_{30}, \alpha_{40}, \alpha_{44}]$ , suffisent donc pour obtenir une très bonne séparation des classes comme le montre la figure 4.21.



**Fig. 4.21.** Position des classes de  $X_{al}$  en dimension 3. Les paramètres sont sélectionnés avec la méthode SBS.

Sur cette figure, on peut voir que les trois paramètres permettent une bonne séparation des différents modes de fonctionnement.

De plus, on remarque que l'utilisation de la normalisation par rapport à la charge a permis d'éliminer la variation de celle-ci. Seule la dispersion due aux différents modes de fonctionnement est visible.

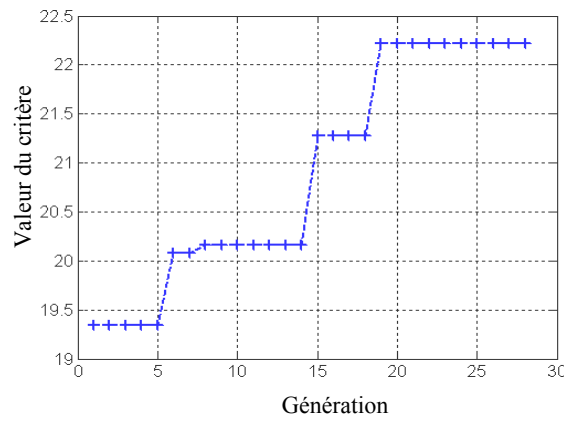
#### 4.5.1.1.2. Sélection par la méthode AG

Contrairement à la méthode précédente, l'algorithme AG utilise le calcul d'un critère sur des *individus* codés pour différentes valeurs de  $d'$  (dimension de l'espace). Ainsi, il permet de réaliser l'optimisation simultanée des meilleurs paramètres ainsi que la dimension optimale de l'espace de représentation.

Le critère utilisé (appelé aussi fonction d'adaptation) est le même que celui de la méthode SBS (équation 3.23).

La taille de la *population* initiale est prise égale à  $P = 500$  *individus*. Ces *individus* sont constitués de manière aléatoire (choix des paramètres). L'arrêt de l'algorithme est réalisé lorsque la meilleure fonction d'adaptation n'a pas évolué de la génération  $t$  à la génération  $t+10$ .

La figure 4.22 présente l'évolution du critère durant le fonctionnement de l'algorithme. A chaque génération la meilleure fonction d'adaptation est retenue en ne tenant pas compte du nombre de paramètres sélectionnés. Ainsi, deux générations successives peuvent présenter comme solution optimale, deux valeurs différentes pour  $d'$ .



**Fig. 4.22.** Evolution du critère durant le fonctionnement de l’algorithme AG

On constate que l’algorithme converge assez rapidement vers une solution stable (28 itérations). La solution finale donne directement le vecteur forme optimal ainsi que sa dimension

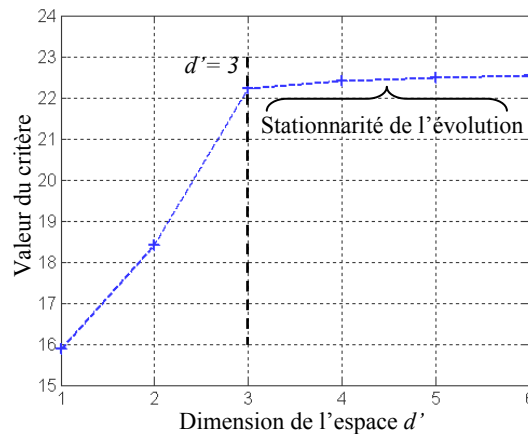
$$V_{Fnorm(AG)} = [\alpha_{30}, \alpha_{40}, \alpha_{44}] \tag{4.20}$$

Les paramètres de ce vecteur  $V_{Fnorm(AG)}$  correspondent respectivement à :

- harmonique 7 du fondamental :  $7f_s$
- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l’axe  $\alpha$  :  $\delta_\alpha$
- écart-type du module  $I_s = I_{s\alpha} + jI_{s\beta}$  :  $\sigma_s$

Il s’agit de la même solution que celle donnée par l’algorithme SBS. Ceci peut s’expliquer par le fait que, comme le montre le rapport de Fisher, ces trois paramètres sont les seuls à avoir un fort pouvoir de discrimination. Cela expliquerait aussi la rapide convergence vers la solution optimale.

La figure 4.23 montre l’évolution du critère en fonction des meilleures solutions pour chaque dimension.



**Fig. 4.23.** Evolution du critère en fonction de la meilleure solution de chaque dimension  $d'$

Cette figure montre que l'évolution du critère devient stationnaire lorsque la dimension devient supérieure à 3. Cela permet de valider le fait que la dimension  $d' = 3$  est bien la dimension optimale pour définir l'espace de représentation.

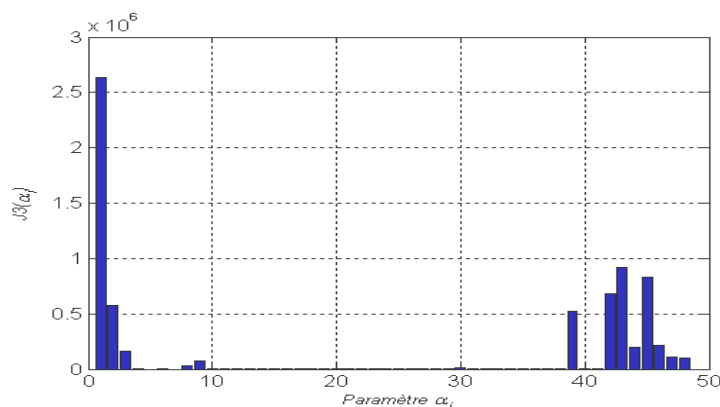
Comme les deux méthodes de sélection utilisées donnent le même résultat au niveau du choix du vecteur forme optimal, nous utiliserons par la suite une seule méthode pour définir le vecteur forme. L'algorithme AG sera retenu : cette méthode à l'avantage de déterminer simultanément la combinaison de paramètres la plus adaptée et la taille du vecteur forme optimal.

#### 4.5.1.1.3. Comparaison avec des données non normalisées

A titre de comparaison, il semble intéressant de regarder le choix du vecteur optimal, ainsi que les performances au sens du nombre d'échantillons test mal classés pour des données non normalisées (dépendantes du niveau de charge). Pour cela, le même ensemble d'apprentissage  $X_{al}$  et les mêmes ensembles de test sont utilisés. La seule différence est qu'un mode de fonctionnement n'est plus représenté par une classe mais par 4 classes, correspondant aux 4 niveaux de charge (0, 50, 75 et 100%). De ce fait,  $X_{al}$  est constitué de 16 classes, toutes composées de 10 échantillons répartis de la façon suivante :

- Machine saine :  $\Omega_1, \Omega_3, \Omega_4, \Omega_5$  respectivement pour 0, 50, 75 et 100% de charge,
- 1 barre cassée :  $\Omega_6, \Omega_8, \Omega_9, \Omega_{10}$  respectivement pour 0, 50, 75 et 100% de charge,
- Déséquilibre d'alimentation (10%) :  $\Omega_{11}, \Omega_{13}, \Omega_{14}, \Omega_{15}$  respectivement pour 0, 50, 75 et 100% de charge,
- Usure roulements :  $\Omega_{16}, \Omega_{18}, \Omega_{19}, \Omega_{20}$  respectivement pour 0, 50, 75 et 100% de charge.

Tout d'abord, le rapport de Fisher est utilisé pour déterminer le pouvoir discriminant de chaque paramètre par rapport à cet ensemble  $X_{al}$ .



**Fig. 4. 24.** Variation du rapport de Fisher pour chaque paramètre pour des données non normalisées



Sur cette figure, on remarque que 6 paramètres  $[\alpha_1, \alpha_2, \alpha_{39}, \alpha_{42}, \alpha_{43}, \alpha_{45}]$  ont un fort pouvoir de séparation des classes. D'autres paramètres  $[\alpha_3, \alpha_{44}, \alpha_{46}, \alpha_{47}, \alpha_{48}]$  présentent aussi un pouvoir de discrimination mais à une échelle moindre.

D'après le résultat de la figure 4.24, la dimension du vecteur forme peut aller de 2 à 6.

En utilisant la méthode de sélection AG, celle-ci donne comme dimension optimale  $d' = 5$ .

Le vecteur forme utilisé, dans le cas de données non normalisées et pour l'ensemble d'apprentissage  $X_{al}$  défini dans le tableau 4.5, est le vecteur de dimension 5 suivant :

$$V_{F(AG)} = [\alpha_1, \alpha_2, \alpha_{43}, \alpha_{44}, \alpha_{45}] \quad (4.21)$$

Les paramètres de ce vecteur  $V_{F(AG)}$  correspondent respectivement à :

- puissance active moyenne :  $m_p$
- puissance réactive moyenne :  $m_q$
- écart-type de la composante  $I_{s\beta}$  des courants :  $\sigma_\beta$
- écart-type du module  $I_s = I_{s\alpha} + jI_{s\beta}$  :  $\sigma_s$
- impédance directe :  $Z_d$

Vérifions la classification obtenue à l'aide de la règle de décision exclusive des '5 - ppv' pour différentes dimensions du vecteur forme.

Paramètres du vecteur forme	Nombre de mal classés
$[\alpha_1, \alpha_{43}]$	11
$[\alpha_1, \alpha_2, \alpha_{43}]$	11
$[\alpha_1, \alpha_2, \alpha_{43}, \alpha_{45}]$	3
$[\alpha_1, \alpha_2, \alpha_{43}, \alpha_{44}, \alpha_{45}]$	1
$[\alpha_1, \alpha_2, \alpha_3, \alpha_{43}, \alpha_{44}, \alpha_{45}]$	3

**Tableau. 4. 8.** Résultats de test pour définir le vecteur forme à l'aide de l'algorithme AG pour des données non normalisées

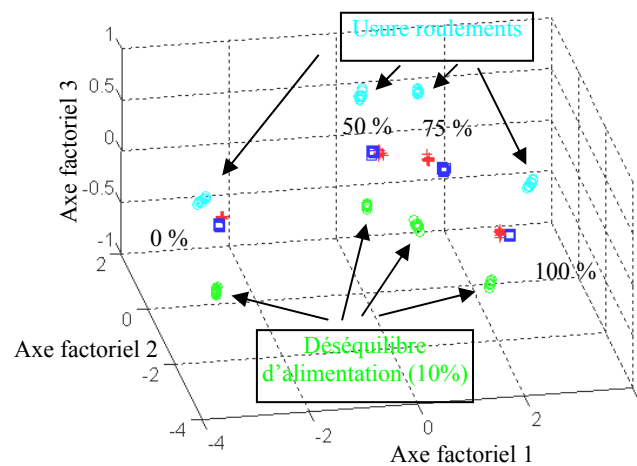
Le nombre de mal classés minimal est obtenu pour un vecteur forme composé de 5 paramètres, c'est-à-dire pour  $d' = 5$ . La dimension trouvée par l'AG semble donc être la meilleure solution.

Si une comparaison est faite avec les résultats obtenus pour des données normalisées, le taux d'échantillons test bien classés est quasiment identique. En revanche, ce taux n'est pas obtenu pour la même dimension de l'espace de représentation. En effet, pour des données normalisées, la dimension optimale est seulement de 3 alors que pour des données non normalisées celle-ci est de 5. Par conséquent, les temps de calcul avec une dimension plus

grande vont être augmentés et de plus, il va être nécessaire de réaliser une analyse factorielle pour visualiser les données.

L'utilisation de la normalisation des données par rapport à la charge est donc primordiale pour optimiser et augmenter les performances de la méthode de diagnostic.

La figure 4.25 montre la position des classes de  $X_{al}$  dans le sous – espace de dimension 3 obtenu par ACP.



**Fig. 4. 25.** Position des classes dans le sous – espace de dimension 3. Machine saine «+», 1 barre cassée «□», déséquilibre d'alimentation (10%) «o», usure roulements «o».

#### 4.5.1.1.4. Conclusion

Au final, les deux méthodes de sélection utilisées donnent quasiment les mêmes résultats au niveau des paramètres sélectionnés. En ce qui concerne le choix de la dimension optimale, elles donnent le même résultat, c'est-à-dire  $d' = 3$ . Ce résultat paraît logique au vue de l'analyse réalisée à l'aide du rapport de Fisher. En effet, on a pu constater, grâce à cette analyse, que sur les 48 paramètres initiaux, 3 étaient nettement plus discriminants que les autres.

Au niveau de la classification, les résultats sont d'autant plus satisfaisants que les défaillances utilisées dans  $X_a$  sont de faible amplitude (1 barre cassée et faible déséquilibre d'alimentation).

La comparaison de ces deux techniques de sélection aura montré l'efficacité de la méthode à base d'algorithme génétique (AG). Par conséquent, pour la suite de notre étude, nous utiliserons cette méthode pour définir le vecteur forme optimal en fonction de l'ensemble d'apprentissage utilisé dans le cas d'un mode *supervisé* (partitionnement en classes de l'ensemble d'apprentissage supposé connu).

Cette partie a montré aussi la nécessité d'effectuer une normalisation des données pour éliminer l'effet de la variation du niveau de charge.

#### 4.5.1.2. Dans le cas du mode *non supervisé*

Les méthodes appliquées précédemment ne peuvent être utilisées que dans le cas *supervisé*, c'est-à-dire en connaissant le partitionnement de l'ensemble  $X_{al}$  en classes. Lorsque cette connaissance n'est pas disponible, il est nécessaire de réaliser une étape de coalescence. Cette étape va permettre de structurer automatiquement l'ensemble d'apprentissage en classes. Ensuite, une expertise humaine est nécessaire pour valider cette décomposition et identifier à quels modes de fonctionnement correspondent ces classes.

Pour optimiser cette recherche, une méthode a été développée au paragraphe 3.4 du chapitre III utilisant les algorithmes génétiques multi-population AG<sub>m</sub>. Cette approche combine simultanément l'étape de coalescence et la sélection des paramètres pertinents.

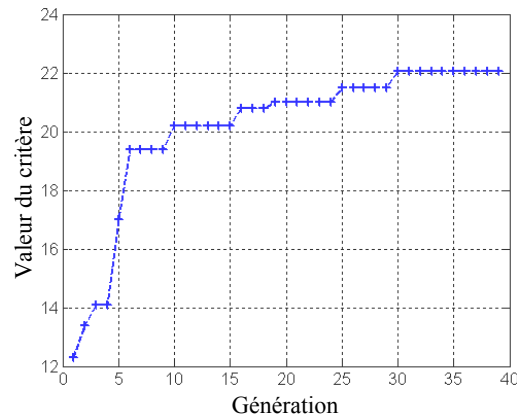
La fonction d'adaptation utilisée est définie par l'équation 3.23. Il s'agit d'un critère qui cherche à maximiser la séparation des classes et à minimiser la dispersion à l'intérieur de ces classes.

La taille de la *population* initiale est prise égale à  $P = 500$  individus. Dans chaque individu est codé la partie 'sélection des paramètres' et la partie 'centre des classes'. Ce codage est défini dans le paragraphe 3.4.2.2 du chapitre III. Le nombre initial de classes  $M_k$  dans l'individu  $k$  est choisi de manière aléatoire dans l'intervalle  $[2, M_{kmax}]$  avec  $M_{kmax}$  égal à  $\sqrt{N}$ , d'où  $M_{kmax} = 12$ . Les opérations de croisement et de mutation sont décrites dans le paragraphe 3.4.2.3. Le fonctionnement de l'algorithme complet est décrit et détaillé dans le chapitre III au paragraphe 3.4.

Cet algorithme a la particularité de diviser en différentes 'niches' les  $P$  individus. Le nombre de niches principales peut aller de 2 à  $M_{kmax}$ . Ces niches principales correspondent au nombre de classes codées dans les individus. Ensuite chaque niche principale peut être divisée en sous-niches en fonction du nombre de paramètres codés dans les individus (tableau 3.4, chapitre III). Au final, dans chacune de ces sous-niches, est calculée la fonction d'adaptation et sont appliquées les différentes opérations (croisement, mutation, opération de recherche locale, ...).

L'arrêt de l'algorithme s'effectue lorsque, de la génération  $t$  à la génération  $t+10$ , la meilleure fonction d'adaptation de toutes les sous-niches ne présente pas d'évolution.

La figure 4.26 présente l'évolution de la meilleure fonction d'adaptation en fonction des générations successives. Ainsi, d'une génération à l'autre, la meilleure fonction peut venir de sous-niches différentes.



**Fig. 4. 26.** Evolution de la meilleure fonction d'adaptation durant le fonctionnement de l'algorithme  $AG_m$

On constate que l'algorithme converge vers une solution optimale en 30 générations. L'arrêt de l'algorithme étant une stationnarité dans l'évolution de la valeur du critère pendant 10 générations, il faut attendre la 40<sup>ème</sup> génération (30 + 10) peut que l' $AG_m$  s'arrête définitivement.

La solution finale donne le vecteur forme optimal suivant. La dimension de ce vecteur donne la dimension de l'espace de représentation optimale.

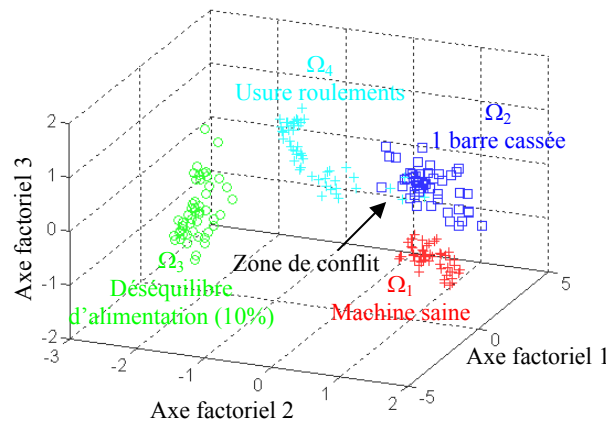
$$V_{Fnorm(AGm)} = [\alpha_9, \alpha_{30}, \alpha_{40}, \alpha_{44}] \quad (4.22)$$

Les paramètres de ce vecteur correspondent respectivement à :

- amplitude de la fréquence  $f_s + f_r$
- harmonique 7 du fondamental :  $7f_s$
- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\alpha$  :  $\delta_\alpha$
- écart-type du module  $I_s = I_{s\alpha} + jI_{s\beta}$  :  $\sigma_s$

L'algorithme  $AG_m$  considère que la dimension optimale est  $d'=4$ . Cette sélection semble correcte au regard du pouvoir discriminant de ces paramètres (figure 4.19). Trois des quatre paramètres sélectionnés sont identiques aux vecteurs forme optimaux obtenus avec la méthode SBS et AG (équations 4.19 et 4.20) dans le mode *supervisé*. De plus, si on regarde l'évolution du critère (ou fonction d'adaptation) en fonction de la dimension sur la figure 4.23, on constate que la valeur de ce critère est meilleure pour  $d'=4$  et que  $d'=3$ . Cependant, cette faible évolution n'apporte pas d'amélioration supplémentaire au niveau de la classification comme le montre le tableau 4.7. La raison de ce choix est due au fait que dans cet algorithme, aucune étude sur la stationnarité de l'évolution du critère par rapport à la dimension n'est réalisée. Seule la meilleure fonction est choisie, avec ses  $d'$  paramètres sélectionnés, sans analyser la solution avec des dimensions inférieures.

La figure 4.27 représente la position des classes de  $X_{al}$  dans le sous – espace de dimension 3 obtenu par ACP.



**Fig. 4. 27.** Position des classes déterminées par l'AG<sub>m</sub> dans le sous – espace de dimension 3. Machine saine «+», 1 barre cassée «□», déséquilibre d'alimentation (10%) «o», usure roulements «+».

Pour le choix du nombre de classes et la répartition des vecteurs dans ces classes, la solution optimale a été obtenue pour M = 4 classes avec la répartition suivante :

	$\Omega_1$	$\Omega_2$	$\Omega_3$	$\Omega_4$
$\Omega_1$	40	0	0	0
$\Omega_2$	0	39	0	<b>1</b>
$\Omega_3$	0	0	40	0
$\Omega_4$	0	<b>4</b>	0	36

**Tableau. 4. 9.** Résultats de la répartition des vecteurs dans les M classes déterminées par la méthode AG<sub>m</sub>

L'algorithme retrouve bien les quatre classes contenues dans l'ensemble d'apprentissage  $X_{al}$ . Connaissant l'étiquetage de chaque vecteur, il est possible d'en déduire le taux de vecteurs bien classés  $T_B$ :

$$TB = \frac{155}{160} \times 100 = 96.8\% \quad (4.23)$$

Ceci montre la bonne répartition effectuée par l'algorithme. Seuls 5 vecteurs sont mal répartis. Il s'agit de vecteurs se situant dans une zone ambiguë localisée entre deux classes (zone de conflit sur la figure 4.27).

Au final, l'algorithme  $AG_m$  a déterminé une sélection de paramètres assez pertinente même s'il apparaît que la dimension de l'espace de représentation n'est pas optimale. Pour la dimension déterminée ici, la classification obtenue avec ce vecteur forme est tout de même de 100% d'échantillons bien classés (tableau 4.7). Concernant la structuration en classes, les quatre classes sont bien identifiées avec un taux de vecteurs bien classés important, traduisant la bonne répartition de l'ensemble des vecteurs.

## 4.5.2. Décision par la règle des $k - ppv$

Les premières étapes de la *phase d'analyse*, qui sont le choix du vecteur forme, la réduction de sa dimension et la détermination de l'espace de décision, viennent d'être traitées dans le paragraphe précédent. A noter que l'espace de décision, c'est-à-dire la structuration en classes de l'ensemble d'apprentissage dans le cas du mode *supervisé*, est supposé connu, alors que dans le cas *non supervisé*, il est déterminé par la méthode à base d'algorithme génétique multi-population (§ 4.5.1.2).

Pour établir la règle des  $k - ppv$ , deux paramètres sont à définir : la métrique et le choix de  $k$ . Le choix de la distance a été étudié dans le chapitre III (§ 3.5.1.1.1). Par conséquent, la distance de Mahalanobis est utilisée comme métrique car elle tient compte de la variance des classes de  $X_a$ .

### 4.5.2.1. Etude de sensibilité sur $k$ (nombre de plus proches voisins)

Concernant le choix de  $k$ , comme il a été dit auparavant (§ 3.5.1.1.2, chapitre III), son choix est souvent arbitraire. En pratique,  $k$  égal à  $\sqrt{N}$  [Dubuisson, 90] réalise souvent un bon compromis. Mais il est possible de faire une étude sur  $k$ . Ainsi, pour évaluer la valeur de  $k$  permettant d'obtenir les meilleurs résultats en terme de classification, l'ensemble d'apprentissage  $X_{al}$  et les ensembles test définis dans les tableaux 4.5 et 4.6 respectivement, sont utilisés.

Ces ensembles sont en quelque sorte variables car nous utilisons la méthode du 'leave 5 out'. La procédure est la suivante :

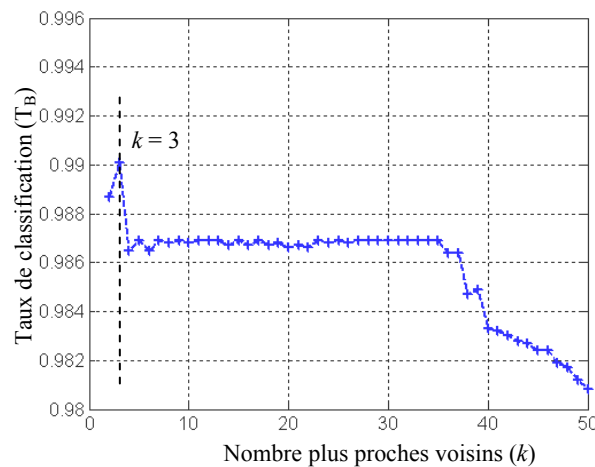
Le vecteur forme optimal est fixé (équation 4.20) pour  $X_{al}$  choisi. Pour chaque mode de fonctionnement et pour chaque niveau de charge, 10 échantillons sont tirés aléatoirement et vont dans l'ensemble d'apprentissage et les 5 restants vont dans l'ensemble de test. A chaque itération (tirage aléatoire), l'espace de décision (les classes) est redéfini en fonction des

nouvelles données d'apprentissage. De cette manière, les ensembles d'apprentissage et de test sont toujours disjoints.

Après chaque sélection des éléments de l'ensemble d'apprentissage et des ensembles de test, le nombre d'échantillons bien classés est calculé. Cela permet d'avoir une estimation des taux moyens de bonne classification :

$$T_B = \frac{\overline{N}_B}{N_t} \quad (4.24)$$

Avec :  $\overline{N}_B$  la moyenne du nombre de bien classés sur les cent itérations et  $N_t$  le nombre d'échantillons de test. La figure 4.28 montre l'évolution de ce taux en fonction des différentes valeurs de  $k$ .

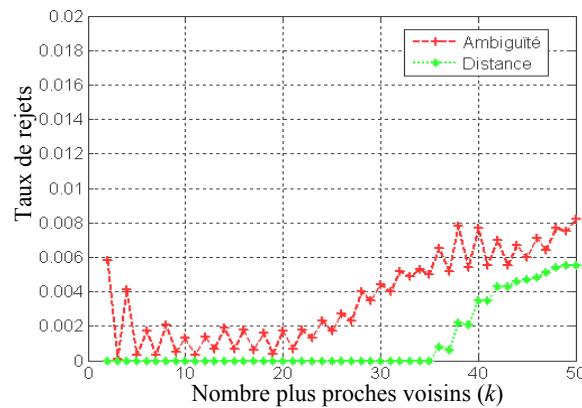


**Fig. 4. 28.** Taux de classification en fonction du nombre de voisins pour un ensemble d'apprentissage normalisé

Sur cette figure, on remarque que la meilleure classification est obtenue pour un  $k=3$  avec un taux égal à 99%. Pour un  $k$  variant de 3 à 35, ce taux est quasiment constant (98.7%). A partir de  $k=35$ , le taux diminue assez nettement.

Cette étude permet de définir le nombre de voisins  $k$  que nous utiliserons dans la règle de décision des  $k - ppv$  lors de la phase d'exploitation dans le cas de données normalisées. Ce nombre est donc pris égal à 3,  $k = 3$  pour la suite de notre étude.

Pour valider ce choix, la figure 4.29 présente le taux d'échantillons test rejetés en différenciant le rejet de distance du rejet d'ambiguïté.



**Fig. 4.29.** Taux de rejets en fonction du nombre de voisins

Les résultats de la figure 4.29 confirment bien le choix pris auparavant. En effet le taux de rejet en ambiguïté est minimal pour  $k = 3$ . De même en ce qui concerne le taux de rejet en distance.

Il est important de noter que quelque soit le nombre  $k$  de voisins, le taux de classification est élevé (varie de 98 à 99%). Ceci permet de valider, d'une part, le choix des paramètres formant le vecteur forme optimal. Ces paramètres présentent un fort pouvoir de séparation entre les différentes classes de  $X_{al}$ , rendant la classification plus efficace.

D'autre part ce fort taux permet de montrer la robustesse de notre approche et notamment de la règle de décision.

#### 4.5.2.2. Phase d'exploitation : classification pour les nouvelles observations $\underline{X}_u$

Il s'agit maintenant de valider les méthodes de décision mises en place. Comme précisé au début de ce chapitre, nous utilisons les mesures ne faisant pas partie de l'ensemble d'apprentissage et des ensembles de test. Ainsi, le diagnostic final de la règle des  $k - ppv$  est établi pour les nouvelles observations suivantes :

Observations	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$C_1 : \underline{X}_{u1} \text{ à } \underline{X}_{u5}$	Machine saine	25 %	5
$C_2 : \underline{X}_{u6} \text{ à } \underline{X}_{u10}$	1 barre cassée	25 %	5
$C_3 : \underline{X}_{u11} \text{ à } \underline{X}_{u15}$	3 barres cassées	0 %	5
$C_4 : \underline{X}_{u16} \text{ à } \underline{X}_{u20}$	Déséquilibre d'alimentation stator (5%)	50 %	5
$C_5 : \underline{X}_{u21} \text{ à } \underline{X}_{u25}$	Déséquilibre d'alimentation stator (10%)	25 %	5
$C_6 : \underline{X}_{u26} \text{ à } \underline{X}_{u30}$	Usure des roulements	25 %	5

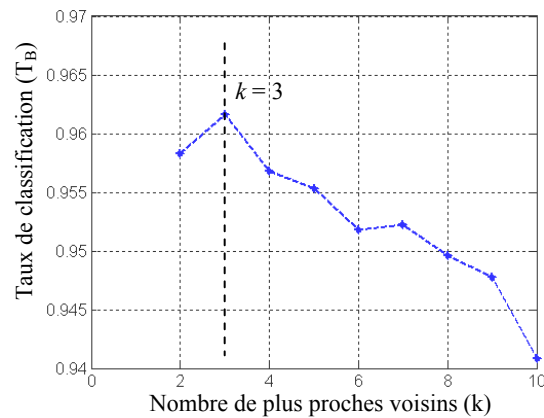
**Tableau. 4.10.** Observations utilisées pour la validation de la règle de décision



Cinq acquisitions sont prises en compte pour chacun de ces modes de fonctionnement. Dans un premier temps, nous allons appliquer la règle des  $k - ppv$  exclusive avec rejet pour des données non normalisées puis normalisées et ensuite comparer les résultats de classification obtenus dans les deux cas.

#### 4.5.2.2.1. Classification par la règle des $k - ppv$ exclusive pour un ensemble d'apprentissage non normalisé

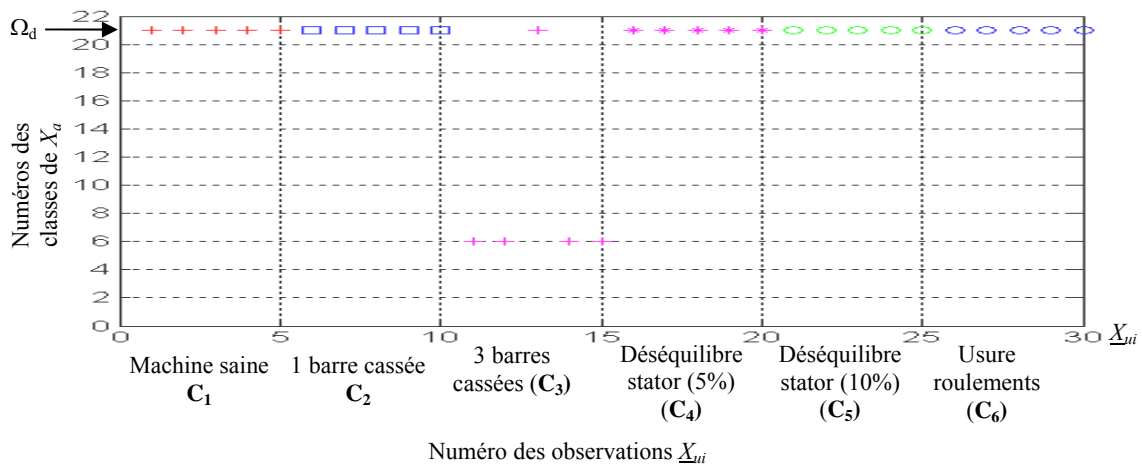
Les données de l'ensemble d'apprentissage non normalisé sont définies par le vecteur forme  $V_{F(AG)}$  (équation 4.21). La règle des  $k - ppv$  exclusive est utilisée. Une étude pour définir le choix de  $k$  peut être faite. La procédure ('leave 5 out') est la même que dans le paragraphe 5.2. La figure 4.30 présente le taux de classification obtenu pour différentes valeurs de  $k$ .



**Fig. 4.30.** Taux de classification en fonction du nombre de voisins pour un ensemble d'apprentissage non normalisé

Le meilleur taux de classification est obtenu par  $k = 3$  avec 96%. Ainsi, pour la phase d'exploitation, nous utiliserons comme nombre de plus proches voisins dans la règle des  $k - ppv$ ,  $k = 3$ . Le nombre minimal de voisin  $k'$  concernant le rejet en ambiguïté est fixé à  $k/2$ . Par conséquent, comme  $k = 3$ , le nombre minimal de voisin est 2. Le seuil de rejet en distance ( $T_{50}$ ) est défini par l'équation 4.14.

La figure 4.31 donne les affectations des nouvelles observations aux classes de l'ensemble d'apprentissage non normalisé.



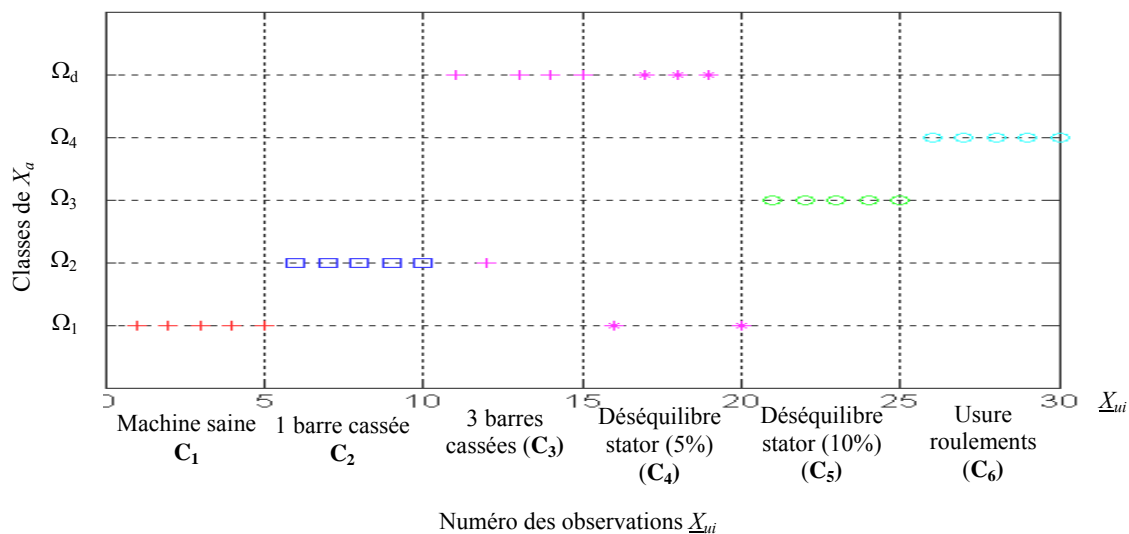
**Fig. 4. 31.** Répartition des nouvelles observations dans le cas de données non normalisées

Les observations ‘déséquilibre d’alimentation 5%’,  $C_4$ , sont rejetées en distance, ce qui est logique car ce mode de fonctionnement n’est pas connu dans l’ensemble d’apprentissage initial. L’affectation d’une partie des observations de  $C_3$  (3 barres cassées) à la classe  $\Omega_6$  (défaut 1 barre cassée à 0% de charge) est un échec du diagnostic. Cela confirme la difficulté d’identifier des modes de fonctionnement pour des faibles niveaux de charge. En revanche, concernant les autres mesures, c’est-à-dire ‘machine saine’, ‘1 barre cassée’, ‘déséquilibre d’alimentation 10%’ et ‘usure roulements’,  $C_1$ ,  $C_2$ ,  $C_5$  et  $C_6$ , celles-ci sont toutes rejetées en distance, ce qui peut être assimilé à une mauvaise classification et donc à une erreur dans le diagnostic. En effet, ces mesures correspondent à des modes de fonctionnement connus et contenus dans l’ensemble d’apprentissage. Le fait qu’elles soient rejetées est simplement dû à la différence du niveau de charge de la machine. Les nouvelles observations sont obtenues pour un niveau de charge de 25%, or ce niveau de charge n’est pas inclus dans  $X_{al}$ . Il est donc nécessaire de définir, initialement dans l’ensemble d’apprentissage, les modes de fonctionnement pour l’ensemble des niveaux de charge, ce qui ne paraît pas réalisable (problème de stockage des données et de temps de calcul). Pour pallier à ce problème, la normalisation des données par rapport à la charge peut être une solution.

#### 4.5.2.2.2. Classification par la règle des $k - ppv$ exclusive pour un ensemble d’apprentissage normalisé

Les données de l’ensemble d’apprentissage  $X_{al}$  (tableau 4.5) sont définies par le vecteur forme  $V_{Fnorm(AG)}$  (équation 4.20). La règle des  $k - ppv$  exclusive est utilisée pour  $k = 3$ . Le seuil de rejet en distance est défini par l’équation 4.15 et le nombre minimal de voisin  $k'$  concernant le rejet en ambiguïté est fixé à  $k/2$ . Par conséquent, comme  $k = 3$ , le nombre minimal de voisin est 2.

La figure 4.32 donne les affectations des nouvelles observations aux classes de l'ensemble d'apprentissage normalisé.



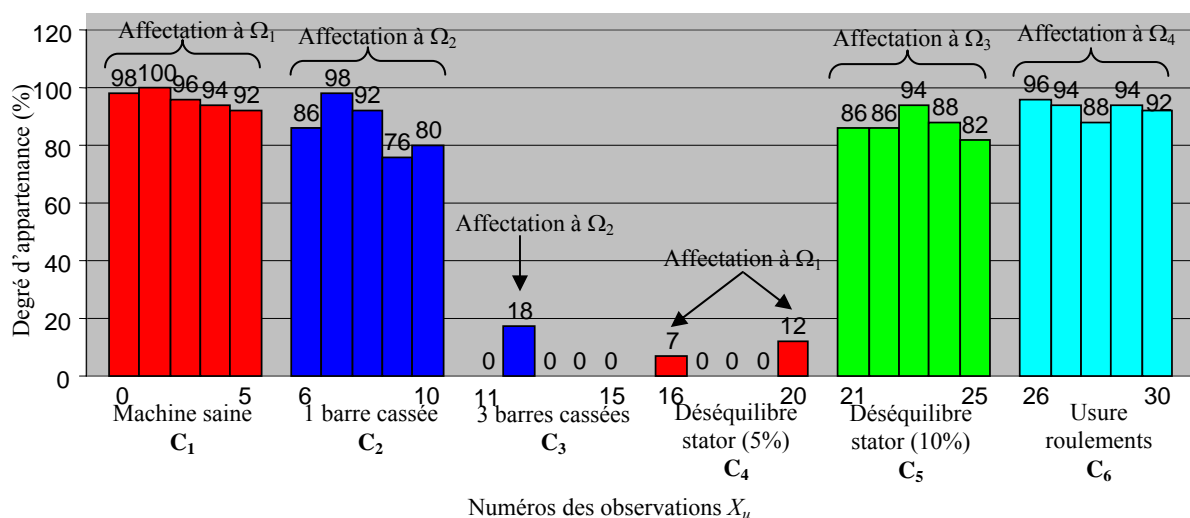
**Fig. 4.32.** Répartition des nouvelles observations réalisées par la règle des  $k - ppv$  exclusive dans le cas de données normalisées

Les observations '3 barres cassées',  $C_3$ , sont bien rejetées en distance, ce qui montre l'intérêt d'associer la notion de rejet à la règle de décision dans le cas où les modes de fonctionnement des observations ne sont pas contenus dans l'ensemble d'apprentissage. Seule l'observation  $\underline{X}_{u12}$  (3 barres cassées à vide) est affectée à la classe  $\Omega_2$  (défaut 1 barre cassée), ce qui constitue une erreur de classification. Les observations  $\underline{X}_{u17}$ ,  $\underline{X}_{u18}$  et  $\underline{X}_{u19}$  de  $C_4$  (déséquilibre d'alimentation 5%), sont logiquement rejetées en distance. Par contre  $\underline{X}_{u16}$  et  $\underline{X}_{u20}$  sont affectées à la classe  $\Omega_1$  (machine saine). Cela s'explique par le fait que ces deux observations sont issues d'un fonctionnement avec un déséquilibre d'alimentation au stator de 5%, proche du fonctionnement sain même s'il s'agit d'un début de défaillance. Concernant les autres mesures, 'machine saine', '1 barre cassée', 'déséquilibre d'alimentation 10%' et 'usure roulements',  $C_1$ ,  $C_2$ ,  $C_5$  et  $C_6$ , elles ne sont plus rejetées en distance, mais affectées respectivement aux classes  $\Omega_1$ ,  $\Omega_2$ ,  $\Omega_3$ ,  $\Omega_4$  de  $X_{a1}$ . Du fait de la normalisation, les classes de  $X_{a1}$  regroupent l'ensemble des niveaux de charge pour un même mode de fonctionnement. Cela a pour effet de diminuer le nombre de mesures nécessaires au départ et d'éviter les erreurs de diagnostic comme celles vues précédemment (§ 4.5.2.2.1, chapitre IV).

Par conséquent, pour la suite de notre étude, nous utiliserons uniquement les données normalisées par rapport à la charge.



(3 barres cassées 0%) sont bien rejetées car ce mode de fonctionnement n'est pas connu de l'ensemble d'apprentissage. Seule  $X_{u12}$  (3 barres cassées 0%) est affectée à la classe  $\Omega_2$  (1 barre cassée), ce qui constitue une erreur d'affectation. De même  $X_{u16}$  et  $X_{u20}$  (déséquilibre d'alimentation 5%) sont affectées à  $\Omega_1$  (machine saine). Par contre  $X_{u17}$ ,  $X_{u18}$  et  $X_{u19}$  (déséquilibre d'alimentation 5%) sont bien rejetées en distance. Toutes les autres observations sont bien affectées à la classe correspondant à leur mode de fonctionnement. En revanche maintenant, le degré d'appartenance n'est plus, soit de 100 % ou de 0% mais évolue entre ces deux valeurs. La figure 4.33 présente le degré d'appartenance de chaque nouvelle observation  $X_{u1}$  à  $X_{u30}$  aux classes de  $X_{a1}$ . Le degré présenté sur cette figure est le degré d'appartenance maximal de l'observation à l'une des classes.



**Fig. 4.33.** Répartition des nouvelles observations réalisées par la règle des  $k$ -ppv non exclusive dans le cas de données normalisées

Cette règle non exclusive apporte des informations intéressantes par rapport à la règle exclusive sur les observations mal classées ( $X_{u12}$ ,  $X_{u16}$  et  $X_{u20}$ ). En effet concernant  $X_{u12}$ , son degré d'appartenance à la classe  $\Omega_2$  est assez faible (18%). Cette indication suppose que l'observation se situe loin de centre de gravité de la classe et que, par conséquent, l'indice de fiabilité donné à cette affectation est faible. De plus, cela peut indiquer une certaine dérive d'un mode fonctionnement, en l'occurrence ici, une dérive du mode défaillant '1 barre cassée' vers une aggravation de la sévérité du défaut (3 barres cassées). Concernant les observations  $X_{u16}$  et  $X_{u20}$ , le constat est le même. Leur degré d'appartenance respectif à la classe  $\Omega_1$  est de 7 et 12%. Par conséquent, cette affectation doit être prise avec beaucoup de précaution car l'indice de fiabilité est très faible. Il peut s'agir ici de l'évolution du fonctionnement sain vers les prémisses d'un défaut, en l'occurrence un faible déséquilibre d'alimentation au stator (5%). Concernant toutes les autres observations, leur affectation est réalisée avec un fort degré

d'appartenance (entre 80 et 100%), ce qui permet d'assurer une forte fiabilité dans le diagnostic réalisé.

L'intérêt d'associer un degré d'appartenance à chaque affectation réalisée donne une information sur la localisation des observations et ainsi définit un indice de fiabilité et prévient d'une éventuelle dérive d'un mode de fonctionnement. Par conséquent, pour la suite de notre étude, la règle des  $k - ppv$  non exclusive sera appliquée dans la phase *d'exploitation*, pour réaliser l'étape de classification.

#### 4.5.2.3. Actualisation de la connaissance

Les observations rejetées par la règle de décision peuvent représenter des modes inconnus de fonctionnement. Pour déterminer si ces observations caractérisent l'apparition d'une nouvelle classe, c'est-à-dire d'un nouveau mode de fonctionnement, la méthode décrite dans le paragraphe 3.7.1 du chapitre III est utilisée. Ainsi, une permutation circulaire est réalisée sur les observations rejetées en considérant  $N_{\min}$  observations pour chaque sous-ensemble. On fixe ici le nombre minimal d'observations à partir duquel l'apparition d'une nouvelle classe  $\Omega_{\text{new}}$  est considérée,  $N_{\min} = 3$ . Dans notre cas, ce nombre est choisi en fonction de nombre d'observations rejetées, qui en l'occurrence est assez faible (7 observations rejetées). Lors d'une application réelle (en ligne par exemple), il serait intéressant de faire une étude de sensibilité pour définir le choix de  $N_{\min}$ . Le tableau 4.11 présente la recherche d'une ou plusieurs nouvelles classes par rapport aux permutations circulaires réalisées sur les différentes observations rejetées.

Permutation circulaire des observations rejetées	Test d'apparition d'une nouvelle classe	Nouvelle classe possible	Nouvelle classe définitive
$S_1 = [X_{u11}, X_{u13}, X_{u14}]$	$C_{r1} < 1.5 \times C_m \rightarrow \text{oui}$	$\Omega_{\text{new1}}$	$\Omega_{\text{new1}}$
$S_2 = [X_{u13}, X_{u14}, X_{u15}]$	$C_{r2} < 1.5 \times C_m \rightarrow \text{oui}$	$\Omega_{\text{new2}}$	
$S_3 = [X_{u14}, X_{u15}, X_{u17}]$	$C_{r3} > 1.5 \times C_m \rightarrow \text{non}$	-	-
$S_4 = [X_{u15}, X_{u17}, X_{u18}]$	$C_{r4} > 1.5 \times C_m \rightarrow \text{non}$	-	-
$S_5 = [X_{u17}, X_{u18}, X_{u19}]$	$C_{r5} < 1.5 \times C_m \rightarrow \text{oui}$	$\Omega_{\text{new3}}$	$\Omega_{\text{new2}}$
$S_6 = [X_{u18}, X_{u19}, X_{u11}]$	$C_{r6} > 1.5 \times C_m \rightarrow \text{non}$	-	-
$S_7 = [X_{u19}, X_{u11}, X_{u13}]$	$C_{r7} > 1.5 \times C_m \rightarrow \text{non}$	-	-

**Tableau. 4. 11.** Détection de l'apparition d'une nouvelle classe

La compacité des sous-ensembles  $S_1$  à  $S_7$ , notée  $C_{r1}$  à  $C_{r7}$ , est calculée pour ensuite être comparée à la compacité moyenne des classes de  $X_{ai}$  notée  $C_m$ . Dans le tableau 4.11, 3 sous-ensembles,  $S_1$ ,  $S_2$  et  $S_5$  sont susceptibles de représenter l'apparition de nouveau mode de

fonctionnement. Il reste une condition à vérifier : si  $N_{\min}-1$  (ici égal à 2) observations sont identiques (c'est-à-dire si elles correspondent à un même mode de fonctionnement) dans ces sous-ensembles alors ceux-ci sont regroupés et ne forment alors qu'une seule classe. C'est le cas ici des observations  $\underline{X}_{u13}$  et  $\underline{X}_{u14}$  (3 barres cassées à vide) qui se retrouvent dans les sous-ensembles  $S_1$  et  $S_2$ . Au final, l'algorithme détecte l'apparition de deux nouvelles classes regroupant les observations rejetées de la manière suivante :

- $\Omega_{\text{new1}} = [\underline{X}_{u11}, \underline{X}_{u13}, \underline{X}_{u14}, \underline{X}_{u15},]$
- $\Omega_{\text{new2}} = [\underline{X}_{u17}, \underline{X}_{u18}, \underline{X}_{u19}]$

L'intervention d'un expert est nécessaire pour définir les modes de fonctionnement associés à ces nouvelles classes et, par la suite, intégrer ces classes à l'ensemble d'apprentissage. Dans notre cas, les classes  $\Omega_{\text{new1}}$  et  $\Omega_{\text{new2}}$  correspondent respectivement à l'apparition des modes '3 barres cassées' et 'déséquilibre stator (5%)', modes qui ne sont pas inclus dans l'ensemble d'apprentissage  $X_{a1}$ .

#### 4.5.2.4. Estimation de l'évolution des modes de fonctionnements

Comme on a pu le constater, le fait de travailler sur des données normalisées améliore le diagnostic. Les données ne présentant plus de dispersion relative à la charge, la seule évolution possible est celle due à la sévérité d'un défaut. Nous allons appliquer le suivi de cette évolution et la prédiction vers des zones inconnues de l'espace en utilisant les approches (polynomiale et Kalman) décrites au paragraphe 3.6 du chapitre III. Pour cela, l'ensemble d'apprentissage suivant  $X_{a2}$  est utilisé :

Classe	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_1$	Machine saine	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_2$	Déséquilibre d'alimentation au stator (5%)	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_3$	Déséquilibre d'alimentation au stator (10%)	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_4$	Déséquilibre d'alimentation au stator (20%)	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10

**Tableau. 4. 12.** Ensemble d'apprentissage  $X_{a2}$  pour le suivi d'évolution

Cet ensemble est composé de quatre classes, correspondant à deux modes de fonctionnement (sain et défaillance d'alimentation au stator). Il regroupe le fonctionnement d'une machine saine, l'apparition d'une défaillance et l'évolution de celle-ci (degré de sévérité). 10 acquisitions ont été réalisées par niveau de charge, et chaque classe est caractérisée par 5 niveaux de charge, ce qui fait qu'au final, une classe est constituée de 50 échantillons.

Les ensembles test utilisés pour déterminer le meilleur vecteur forme sont les suivants :

<i>Ensembles test</i>	<i>Mode de fonctionnement</i>	<i>Niveau de charge</i>	<i>Nombre d'échantillons</i>
$X_{11}$	Machine saine	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{12}$	Déséquilibre d'alimentation au stator (5%)	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{13}$	Déséquilibre d'alimentation au stator (10%)	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5
$X_{14}$	Déséquilibre d'alimentation au stator (20%)	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5

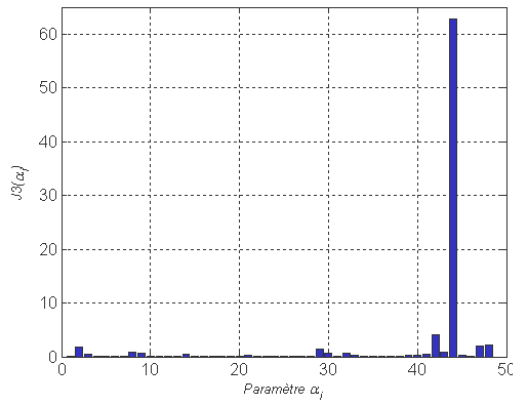
**Tableau. 4. 13.** *Ensembles test pour le suivi d'évolution*

Chaque ensemble test  $X_i$  est composé 25 échantillons, cinq par niveau de charge.

Comme nous utilisons un ensemble d'apprentissage différent de celui de paragraphe 4.5 précédent (tableau 4.5), il est nécessaire de déterminer le pouvoir discriminant de chaque paramètre et ensuite de sélectionner ceux qui permettent la meilleure séparation des classes.

En appliquant le rapport de Fisher sur chaque paramètre, la figure 4.34 présente le pouvoir de séparation des 48 paramètres formant le vecteur forme initial.





**Fig. 4. 34.** Variation du rapport de Fisher pour chaque paramètre pour l'ensemble d'apprentissage défini dans le tableau 4.12.

Ce critère montre que le paramètre  $\alpha_{44}$  a un très fort pouvoir de discrimination par rapport aux autres. Il serait logique de le retrouver, par la suite, dans la sélection faite par l'algorithme AG. Cinq autres paramètres  $[\alpha_2, \alpha_{29}, \alpha_{42}, \alpha_{47}, \alpha_{48}]$  présentent aussi un pouvoir de séparation mais nettement moins important que  $\alpha_{44}$ . On peut noter que le paramètre  $\alpha_{47}$  et  $\alpha_{48}$ , respectivement les pertes totales de la machine  $P_{\text{tot}}$  et l'échauffement au niveau de l'induit  $\Delta\theta$ , permettent de discriminer les différentes classes de  $X_a$ , c'est-à-dire un défaut d'alimentation au stator. Cela confirme bien l'analyse faite au paragraphe 4.3 de ce chapitre et justifie le fait d'avoir intégré ces paramètres dans le vecteur forme initial.

À la vue des résultats de la figure 4.34, la dimension optimale devrait se situer entre 1 et 6.

L'algorithme AG donne, en se basant sur un critère de séparabilité/compacité (équation 3.23), le vecteur forme optimal suivant :

$$V_{F2norm(AG)} = [\alpha_{44}, \alpha_{48}] \quad (4.28)$$

Les paramètres de ce vecteur  $V_{F2norm(AG)}$  correspondent respectivement à :

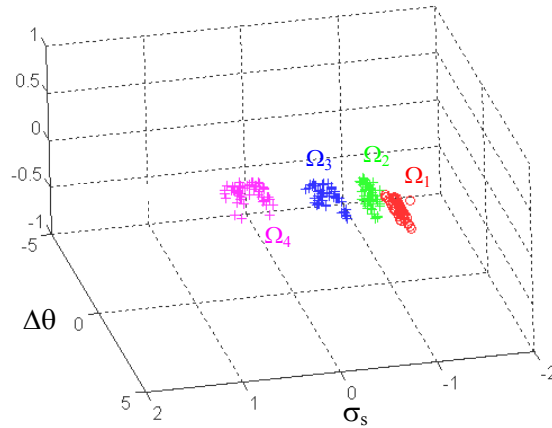
- écart-type du module  $I_s = I_{s\alpha} + jI_{s\beta} : \sigma_s$
- échauffement au niveau de l'induit  $:\Delta\theta$

Le choix des paramètres formant ce vecteur paraît cohérent à la vue de l'analyse faite à l'aide du rapport de Fisher. Le paramètre le plus discriminant ( $\alpha_{44}$ ) est bien sélectionné et un des paramètres avec un plus faible pouvoir de séparation est choisi aussi (en l'occurrence  $\alpha_{48}$ , l'échauffement  $\Delta\theta$ ).

Pour vérifier si le choix est pertinent, il est possible de regarder la classification obtenue par les  $k - ppv$  exclusive (avec  $k = 3$ ), pour ce vecteur forme, sur les ensembles test. Ainsi, le taux de bien classés  $T_B$  est

$$TB = \frac{\text{Nombre de bien classés}}{\text{Nombre de points test}} = \frac{125}{125} \times 100 = 100\% \quad (4.29)$$

La figure 4.35 représente les classes de l'ensemble d'apprentissage  $X_{a2}$  défini par le vecteur forme  $V_{F2norm(AG)}$ .



**Fig. 4.35.** Visualisation de l'ensemble d'apprentissage du tableau 4.12 par le vecteur forme  $V_{F2(AG)}$ . Machine saine «o», déséquilibre d'alimentation (5%) «+», déséquilibre d'alimentation (10%) «+», déséquilibre d'alimentation (20%) «+».

A titre de comparaison, la méthode SBS, pour la même dimension ( $d' = 2$ ) donne le vecteur forme suivant :

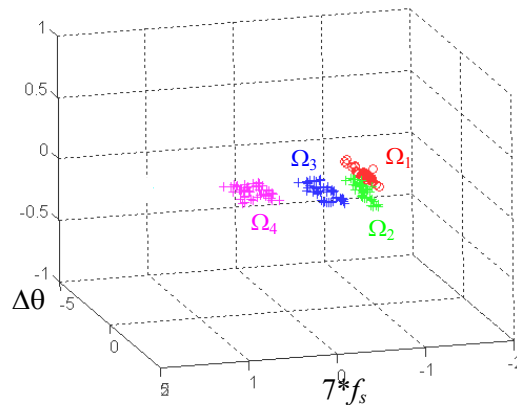
$$V_{F2norm(SBS)} = [\alpha_{30}, \alpha_{44}] \quad (4.30)$$

Le taux de classification obtenu avec ce vecteur forme est :

$$TB = \frac{120}{125} \times 100 = 96\% \quad (4.31)$$

Par conséquent, on peut considérer que le vecteur fourni par l'AG est meilleur que celui donné par la méthode SBS. Cela valide le choix d'utiliser seulement l'AG pour sélectionner les paramètres les plus discriminants.

La figure 4.36 représente l'ensemble d'apprentissage  $X_{a2}$  défini par le vecteur forme  $V_{F2norm(SBS)}$ .



**Fig. 4.36.** Visualisation de l'ensemble d'apprentissage du tableau 4.12 par le vecteur forme  $V_{F2(SBS)}$ . Machine saine «o», déséquilibre d'alimentation (5%) «+», déséquilibre d'alimentation (10%) «+», déséquilibre d'alimentation (20%) «+».

De façon générale, sur la figure 4.36, les classes présentent une plus grande dispersion et sont moins bien séparées surtout pour les classes  $\Omega_1$  et  $\Omega_2$ , ce qui explique le taux de classification plus faible.

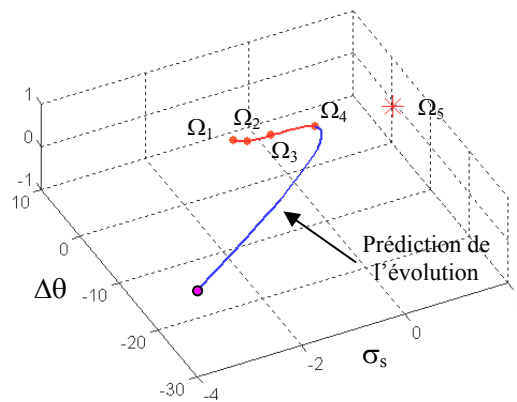
Pour la suite de ce chapitre, le vecteur forme  $V_{F2norm(AG)}$  est utilisé. Nous allons appliquer l'approche polynomiale et l'approche par filtre de kalman (§ 3.6, chapitre III) pour estimer et prédire l'évolution de la sévérité de la défaillance. Ainsi, les quatre classes de  $X_{a2}$ , c'est-à-dire 'machine saine', 'déséquilibre d'alimentation 5%', 'déséquilibre d'alimentation 10%' et 'déséquilibre d'alimentation 20%' vont servir à initialiser le système en réalisant une interpolation entre ces classes. Ensuite la prédiction est faite pour déterminer l'évolution de ce défaut vers des zones inconnues de l'espace. Par conséquent la classe  $\Omega_5$ , 'déséquilibre d'alimentation 40%' décrite dans le tableau 4.14, servira de test pour connaître la précision et la confiance que l'on peut accorder à la prédiction réalisée.

Classe test	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_5$	Déséquilibre d'alimentation au stator (40%)	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10

**Tableau. 4.14.** Classe test servant à déterminer la précision de la prédiction faite par les différentes approches

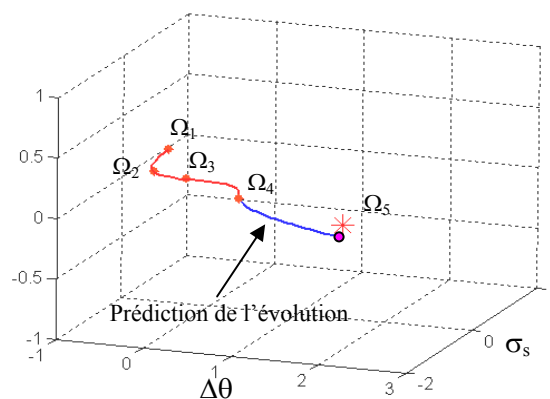
Les figures 4.37 et 4.38 présentent l'interpolation ainsi que la prédiction réalisées respectivement par l'approche polynomiale et l'approche de Kalman dans l'espace de représentation à 2 dimensions défini par le vecteur  $V_{F2norm(AG)}$ . En rouge est représenté l'interpolation entre les centres de gravité des classes de  $X_{a2}$  et en bleu la prédiction effectuée.

La prédiction, dans les deux cas, commence à partir de la classe  $\Omega_4$  et est réalisée jusqu'à la classe  $\Omega_5$ . En d'autres termes, la prédiction est faite pour déterminer l'évolution de la défaillance 'déséquilibre d'alimentation' de 20 à 40 % de déséquilibre.



**Fig. 4. 37.** *Interpolation et prédiction par approche polynomiale*

La figure 4.37 montre que le point d'arrivée « • » de la prédiction polynomiale est très éloignée de la classe test  $\Omega_5$  représentant le mode 'déséquilibre d'alimentation 40%'. Ceci est dû au fait que la fonction polynomiale n'a pas d'effet mémoire, c'est-à-dire qu'elle ne prend pas en considération l'influence des variations antérieures de l'état.



**Fig. 4. 38.** *Interpolation et prédiction par approche de Kalman*

Cette figure permet de voir que le point d'arrivée « • » de la prédiction réalisée par Kalman est proche de la classe test  $\Omega_5$ . La prédiction de l'évolution de la classe  $\Omega_4$  à la classe  $\Omega_5$  est donc plus pertinente que dans le cas précédant. En effet, l'approche par Kalman prend en compte l'évolution des états antérieurs pour réaliser la prédiction à l'instant présent. De plus, il est important de noter que cette prédiction est réalisée vers un horizon assez lointain.

Pour définir précisément les performances de prédiction, c'est-à-dire la précision obtenue, la distance Euclidienne séparant la prédiction à 40% de déséquilibre (« • » sur la figure 4.37

et 4.38) avec la position effective du centre de gravité de la classe test  $\Omega_5$  est calculée. Les résultats sont résumés dans le tableau 4.15.

	Approche	
	Polynomiale	Kalman
Distance Euclidienne	22	0.53

**Tableau. 4. 15.** *Précision de la prédiction dans les deux cas (polynomiale et Kalman)*

Ce tableau permet de constater l'augmentation de la précision offerte par l'approche de Kalman. Par conséquent, cette approche donne une prédiction de l'évolution proche du fonctionnement réel du système. Ceci va permettre d'établir un diagnostic plus précis.

Le tableau 4.16 montre la contribution de l'approche de Kalman comparée à l'approche polynomiale. Les résultats sont présentés en pourcent.

	Polynomiale → Kalman
Augmentation de la précision	+ 97.6 %

**Tableau. 4. 16.** *Comparaison de la prédiction obtenue dans les deux cas (polynomiale et Kalman)*

L'utilisation de l'approche de Kalman pour prédire l'évolution des modes de fonctionnement est intéressante car elle permet d'étoffer l'ensemble d'apprentissage initial, et, de ce fait, améliore le diagnostic. En effet, prenons comme exemple la classification des observations suivantes :

Observations	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$C_1 : \underline{X}_{u1}$ à $\underline{X}_{u25}$	Déséquilibre d'alimentation stator (40%)	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5

**Tableau. 4. 17.** *Observations utilisées pour valider l'apport de la prédiction*

#### 4.5.2.4.1. Classification sans l'utilisation de l'interpolation et de la prédiction

Seul l'ensemble d'apprentissage  $X_{a2}$ , définit dans le tableau 4.12, est utilisé. La classification est réalisée par la règle de décision non exclusive. La fonction d'appartenance ainsi que le rejet d'ambiguïté et d'appartenance sont définis par les équations 4.25, 4.26 et

4.27 respectivement. La fonction pour chaque classe  $\Omega_c$  est définie pour  $g=1$ ,  $\beta=2$ ,  $k_c=50$  et  $\eta=0.02$  ( $1/N_c$ ).

Le tableau ci-dessous présente les résultats de la classification obtenue :

	Classe de $X_{a2}$				
	$\Omega_1$	$\Omega_2$	$\Omega_3$	$\Omega_4$	$\Omega_d$
Observations	0	0	0	0	25 $\underline{X}_{u1}$ à $\underline{X}_{u25}$

**Tableau. 4. 18.** Classification des observations sans interpolation ni prédiction

Le résultat de la classification est logique. En effet, le mode ‘déséquilibre d’alimentation 40%’ n’étant pas contenu dans  $X_{a2}$ , il est normal que ces observations soient rejetées en distance. Par contre, cela représente une erreur de diagnostic, car ce résultat laisse supposer que les observations correspondent à un mode de fonctionnement différent de celui de  $X_{a2}$ . Or il s’agit bien du même mode mais avec un degré de sévérité différent.

**4.5.2.4.2. Classification avec l’utilisation de l’interpolation et de la prédiction**

Pour utiliser l’interpolation et la prédiction dans le but de réaliser une classification des observations, l’approche présentée dans le paragraphe 3.6.3 du chapitre III est mise en place. Ainsi, dans un premier temps, on recherche si l’observation appartient au mode de fonctionnement défini par la *trajectoire moyenne* (interpolation+prédiction). Pour cela, l’opération suivante est effectuée pour chaque observation :

$$\begin{aligned}
 & \text{Si pour } s = [1\%, 40\%] \\
 & d(\underline{P}_s, \underline{X}_u) \leq \beta \cdot d(\underline{P}_s, \underline{P}_{s \max}) \text{ alors } N_{ch} = s \% \\
 & \text{Alors} \\
 & \underline{X}_u \text{ appartient au mode de fonctionnement avec un niveau de sévérité } N_{ch}
 \end{aligned}
 \tag{4.32}$$

*Sinon*

$\underline{X}_u$  est rejetée en distance

avec :

$\underline{P}_s$  : points de l’évolution du mode de fonctionnement (*trajectoire moyenne*) ( $s = 1\%, \dots, 40\%$ )

$\underline{P}_{s \max}$  : points du seuil d’acceptation (*trajectoire maximale*) ( $s = 1\%, \dots, 40\%$ )

$N_{ch}$  : niveau de sévérité de l’observation

$\beta$  : permet de prendre en compte l’incertitude sur les propriétés des classes

En appliquant l’équation 4.32 sur chaque observation, les résultats suivants sont obtenus :

	Points de la <i>trajectoire moyenne</i> les plus proches de l'observation (en %)
Observations : $X_{u1}$ à $X_{u25}$	37 % à 40 %

**Tableau. 4. 19.** *Classification des observations avec interpolation et prédiction*

Ce tableau montre que toutes les observations se situent dans l'enveloppe déterminée à partir de la *trajectoire maximale* (§ 3.6.3, chapitre III). Par conséquent, elles sont considérées comme appartenant aux modes de fonctionnement 'déséquilibre d'alimentation'. Leur localisation est proche de la prédiction réalisée vers un niveau de sévérité de 40%, ce qui correspond bien à leur degré de sévérité.

Dans un deuxième temps, il est possible de localiser leur positionnement dans l'enveloppe. Pour cela, il est nécessaire d'appliquer la fonction d'appartenance  $\mu_{P_s}(X_u)$ , qui est définie non plus pour une classe, mais pour un point  $P_s$  de la *trajectoire moyenne*. Cette fonction est définie par l'équation 3.72 du chapitre III.

#### 4.5.2.4.3. Conclusion

La prédiction réalisée par l'approche de Kalman permet d'estimer l'évolution du mode de fonctionnement 'déséquilibre d'alimentation' vers des degrés de sévérité non définis dans l'ensemble d'apprentissage. Elle permet d'étoffer la base de connaissance, et, de ce fait, améliore le diagnostic.

De façon générale, cette prédiction rend possible l'évaluation de l'évolution d'un mode de fonctionnement vers une défaillance et ainsi permet d'anticiper sur une éventuelle avarie du système.

## 4.6. Diagnostic par *RdF* d'une machine asynchrone alimentée par onduleur de tension

Lorsque la machine asynchrone est alimentée par l'intermédiaire d'un convertisseur, les courants sont perturbés par les harmoniques multiples de la fréquence de commutation. Les formes d'ondes des signaux (figure 2.6, chapitre II) ainsi que leur contenu fréquentiel rendent plus difficile la détection des modes défaillants. Cette étude a pour but de justifier la possibilité d'appliquer des méthodes de reconnaissance de formes aussi bien au diagnostic de la machine qu'au diagnostic de l'alimentation, en l'occurrence, le convertisseur de puissance.

### 4.6.1. Détection des défauts sur la machine asynchrone

La démarche pour le diagnostic est la même que celle faite précédemment dans la partie 4.5, dans le cas d'une machine alimentée par le réseau. Dans un premier temps, pour illustrer les résultats, les modes de fonctionnement utilisés pour composer l'ensemble d'apprentissage et les ensembles test en alimentation onduleur sont les suivants :

Classe	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_1$	Machine saine	0 %	10
		25 %	10
		50 %	10
		100 %	10
$\Omega_2$	1 barre cassée	0 %	10
		25 %	10
		50 %	10
		100 %	10
$\Omega_3$	3 barres cassées	0 %	10
		25 %	10
		50 %	10
		100 %	10
$\Omega_4$	Usure des roulements	0 %	10
		25 %	10
		50 %	10
		100 %	10

**Tableau. 4. 20.** Ensemble d'apprentissage  $X_{a3}$ . Machine alimentée par onduleur

L'ensemble d'apprentissage  $X_{a3}$  est donc constitué de ces quatre modes de fonctionnement représentés par quatre classes dans l'espace de représentation. Chaque classe est composée de 10 échantillons par niveau de charge, ce qui fait au final, que les classes sont composées de 40 échantillons. Cet ensemble est bien distinct de l'ensemble d'apprentissage utilisé précédemment sans onduleur. Il en est de même pour les ensembles test.

Les ensembles test utilisés pour définir, par la suite, le meilleur vecteur forme sont :

Ensembles test	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$X_{t1}$	Machine saine	0 %	5
		25 %	5
		50 %	5
		100 %	5
$X_{t2}$	1 barre cassée	0 %	5
		25 %	5
		50 %	5
		100 %	5
$X_{t3}$	3 barres cassées	0 %	5
		25 %	5
		50 %	5
		100 %	5
$X_{t4}$	Usure des roulements	0 %	5
		25 %	5
		50 %	5
		100 %	5

**Tableau. 4. 21.** Ensembles test dans le cas d'une machine alimentée par onduleur

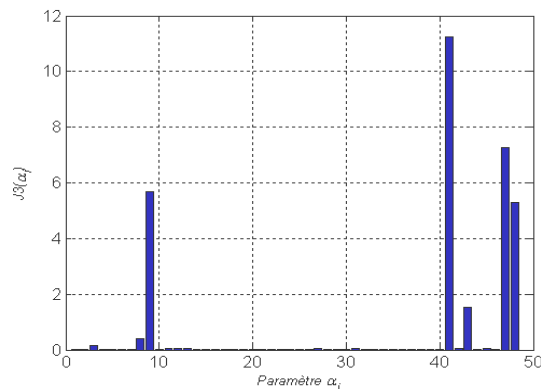


Chaque ensemble test est composé de 5 échantillons par niveau de charge, ce qui fait au final que chaque ensemble est constitué de 20 échantillons.

#### 4.6.1.1. Réduction de l'espace de représentation

En appliquant le rapport de Fisher (équation 3.16, chapitre III), on cherche à déterminer les paramètres qui ont le plus fort pouvoir de discrimination par rapport aux classes de  $X_{a3}$ . Ainsi cela permet de vérifier si l'algorithme de sélection AG réalise le bon choix.

La figure 4.39 présente la variation du critère pour les 48 paramètres calculés du vecteur forme initial.



**Fig. 4.39.** Variation du rapport de Fisher pour chaque paramètre dans le cas d'une alimentation par onduleur

Le critère de Fisher montre que 5 paramètres ont les pouvoirs de discrimination les plus élevés, en l'occurrence  $\alpha_{41}$ ,  $\alpha_{47}$ ,  $\alpha_{48}$ ,  $\alpha_9$  et  $\alpha_{43}$ .

Donc d'après ce résultat, la dimension optimale du vecteur forme se situe entre 2 et 5.

L'application de l'algorithme AG nous donne comme vecteur forme optimal :

$$V_{F3norm(AG)} = [\alpha_{41}, \alpha_{47}, \alpha_{43}] \quad (4.33)$$

Les paramètres de ce vecteur  $V_{F3norm(AG)}$  correspondent respectivement à :

- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\beta$  :  $\delta_\beta$
- pertes totales de la machine :  $P_{tot}$
- écart-type de la composante  $I_{s\beta}$  des courants :  $\sigma_\beta$

Ce choix confirme les résultats obtenus avec le critère de Fisher même si la sélection du paramètre  $\alpha_{43}$  peut être discutable devant le paramètre  $\alpha_9$  ou encore  $\alpha_{48}$ .

Pour valider ce choix, la règle de décision des  $k - ppv$  exclusive est appliquée sur les ensembles test définis dans le tableau 4.21. Le seuil de rejet en distance utilisé est celui défini par l'équation 4.15. D'autre part, le nombre minimal de voisin  $k'$  que doit avoir un échantillon

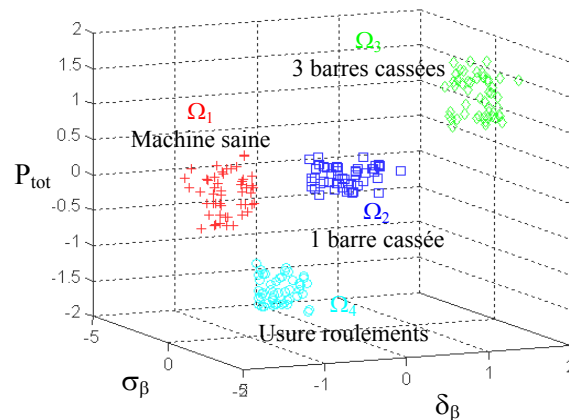
test dans la classe la mieux représentée est fixé à  $k/2$ . Concrètement, étant donné que  $k$  est pris égal à 3, le nombre minimal de voisin est 2.

La classification obtenue pour les différentes dimensions déterminées par l'algorithme AG est :

Paramètres du vecteur forme	Nombre de mal classés
$[\alpha_{41}]$	10
$[\alpha_{41}, \alpha_{43}]$	8
$[\alpha_{41}, \alpha_{43}, \alpha_{47}]$	2
$[\alpha_9, \alpha_{41}, \alpha_{43}, \alpha_{47}]$	2
$[\alpha_9, \alpha_{41}, \alpha_{43}, \alpha_{47}, \alpha_{48}]$	5

**Tableau. 4. 22.** Nombre d'échantillons mal classés en fonction de la dimension de l'espace et des paramètres formant le vecteur forme  $V_{F3norm(AG)}$

Le tableau 4.22 confirme que le vecteur forme défini par l'équation 4.33 est bien la solution optimale. Il permet d'obtenir la meilleure classification. La figure 4.40 présente la position des classes de  $X_{a3}$  (tableau 4.20) dans l'espace de représentation formé par le vecteur forme  $V_{F3norm(AG)}$ .



**Fig. 4. 40.** Position des classes dans l'espace formé par  $V_{F3norm(AG)}$ . Machine saine «+», 1 barre cassée «□», 3 barres cassées «◇», usure roulements «○».

Sur cette figure, on constate que les trois paramètres du vecteur  $V_{F3norm(AG)}$  permettent une très bonne séparation des classes de  $X_{a3}$  malgré l'utilisation d'un onduleur de tension pour alimenter la machine. La normalisation des données par rapport à la charge est toujours efficace puisque les quatre niveaux de charge d'un mode de fonctionnement (0, 25, 50 et 100%) sont regroupés et ne créent pas de dispersion.

### 4.6.1.2. Décision par la règle des $k - ppv$ non exclusive

La fonction d'appartenance associée à chaque classe  $\Omega_c$  ( $c=1$  à  $M$ ), le rejet en ambiguïté et le rejet en distance sont définis par les équations 4.25, 4.26 et 4.27 respectivement.

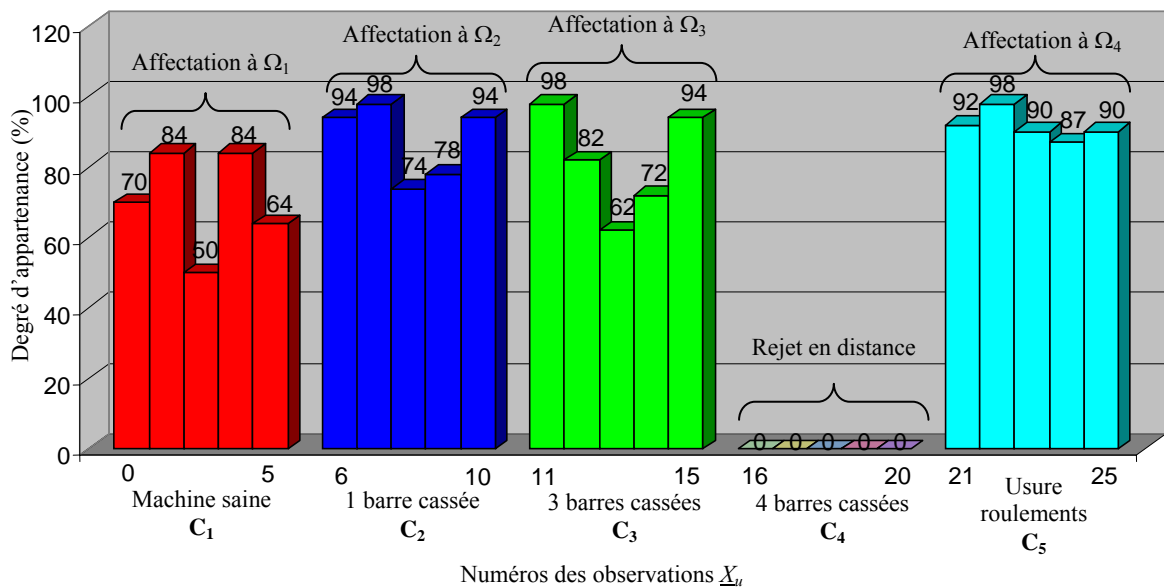
Le diagnostic final de la règle des  $k - ppv$  non exclusive est établi pour les nouvelles observations suivantes :

Observations	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$C_1 : \underline{X}_{u1}$ à $\underline{X}_{u5}$	Machine saine	75 %	5
$C_2 : \underline{X}_{u6}$ à $\underline{X}_{u10}$	1 barre cassée	75 %	5
$C_3 : \underline{X}_{u11}$ à $\underline{X}_{u15}$	3 barres cassées	75 %	5
$C_4 : \underline{X}_{u16}$ à $\underline{X}_{u20}$	4 barres cassées	0 %	5
$C_6 : \underline{X}_{u21}$ à $\underline{X}_{u25}$	Usure des roulements	75 %	5

**Tableau. 4. 23.** Observations utilisées pour la validation de la règle des  $k - ppv$  non exclusive dans le cas d'une alimentation onduleur

Ces mesures ne sont pas utilisées pour constituer l'ensemble d'apprentissage et les ensembles de test. De cette manière, l'ensemble d'apprentissage, les ensembles test et les observations sont disjoints.

La figure 4.41 présente le degré d'appartenance de chaque nouvelle observation  $\underline{X}_{u1}$  à  $\underline{X}_{u25}$  aux classes de  $X_{a3}$ . Le degré présenté sur cette figure est le degré d'appartenance maximal de l'observation à l'une des classes.



**Fig. 4. 41.** Répartition des nouvelles observations réalisées par la règle des  $k - ppv$  non exclusive dans le cas d'une alimentation onduleur

Le taux de classification obtenu est excellent ( $T_B = 100\%$ ). Toutes les observations sont affectées à leur mode de fonctionnement contenu dans l'ensemble d'apprentissage. Le mode '4 barres cassées' n'étant pas présent dans  $X_{a3}$ , les observations  $\underline{X}_{u16}$  à  $\underline{X}_{u20}$  sont rejetées en distance. Ces observations devront être par la suite, traitées pour détecter l'apparition éventuelle d'une nouvelle classe. Le degré d'appartenance pour chaque observation varie entre 70 et 98%. Cela donne une bonne confiance à la classification réalisée. Seule l'observation  $\underline{X}_{u3}$  a un degré plus faible que les autres (50%). Cela est certainement dû à une imprécision de mesures (appareillages, bruits,...). Par contre, si pour les observations suivantes, cela venait à se confirmer, alors il se pourrait que le mode de fonctionnement concerné dérive vers un nouveau mode. Ce résultat montre d'une part, que les paramètres permettent de bien différencier les modes de fonctionnement de la machine et, d'autre part, que la sélection du sous-ensemble de paramètres faite par l'algorithme AG est pertinente.

#### 4.6.1.3. Suivi d'évolution par approche de Kalman

Le suivi de l'évolution et la prédiction vers des zones inconnues de l'espace va être appliqué en utilisant l'approche de Kalman décrite au paragraphe 3.6 du chapitre III. Dans la partie 4.5 précédente, le suivi a été réalisé pour l'apparition et l'évolution de la sévérité d'un défaut d'alimentation (déséquilibre). La prédiction de l'évolution de cette défaillance vers des degrés de sévérité plus importants a été validée. Notre démarche consiste, maintenant, à voir si cette prédiction est généralisable à d'autres types de défaillances (en l'occurrence l'évolution du nombre de barres cassées au rotor) et quel est le nombre de points minimum (ces points sont en fait les centres de gravité des classes de l'ensemble d'apprentissage) nécessaire pour initialiser le système et obtenir une bonne précision sur la prédiction. Pour cela, l'ensemble d'apprentissage suivant  $X_{a4}$  est utilisé :

Classe	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_1$	Machine saine	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_2$	1 barre cassée	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10
$\Omega_3$	3 barres cassées	0 %	10
		25 %	10
		50 %	10
		75 %	10
		100 %	10

**Tableau. 4. 24.** Ensemble d'apprentissage  $X_{a4}$  pour le suivi d'évolution dans le cas d'une alimentation onduleur

Cet ensemble est constitué de trois classes, chacune formée de 50 échantillons. De ce fait, le système est initialisé en réalisant l'interpolation entre les trois centres de gravité de ces classes. Trois points apparaissent comme le nombre minimum de points nécessaire à l'initialisation. En effet, le résultat d'une interpolation réalisée entre deux points, ne peut être qu'une droite, ce qui ne représente certainement pas l'évolution réelle. Reste à connaître la précision obtenue avec ces trois points.

La prédiction détermine l'évolution de ce défaut vers des zones inconnues de l'espace. Par conséquent la classe  $\Omega_4$ , '4 barres cassées' décrite dans le tableau 4.25, servira de test pour connaître la précision que l'on peut accorder à la prédiction réalisée en utilisant seulement les trois points pour l'étape d'initialisation.

Classe test	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_4$	4 barres cassées	0 %	5
		25 %	5
		50 %	5
		75 %	5
		100 %	5

**Tableau. 4. 25.** Classe test servant à déterminer la précision de la prédiction

L'algorithme AG donne, en se basant sur le critère de séparabilité/compacité (équation 3.23), le vecteur forme optimal suivant :

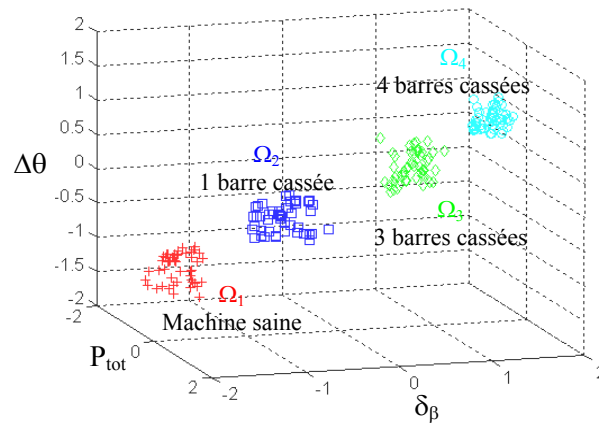
$$V_{F4norm(AG)} = [\alpha_{41}, \alpha_{47}, \alpha_{48}] \quad (4.34)$$

Le choix des paramètres formant ce vecteur paraît cohérent. En effet, l'ensemble d'apprentissage  $X_{a4}$  n'ayant quasiment pas changé par rapport au précédent (§ 4.6.1, tableau 4.20), le mode de fonctionnement 'usure roulements' a été remplacé par le mode '4 barres cassées', il semble normal de retrouver quasiment les mêmes paramètres sélectionnés. Seul le paramètre  $\alpha_{43}$  a été remplacé par le paramètre  $\alpha_{48}$ .

Les paramètres de ce vecteur  $V_{F4norm(AG)}$  correspondent respectivement à :

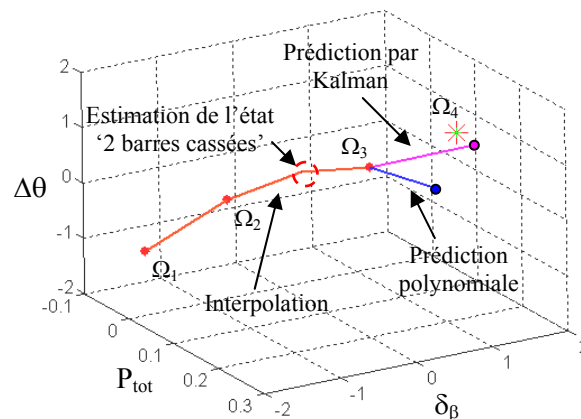
- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\beta$  :  $\delta_\beta$
- pertes totales de la machine :  $P_{tot}$
- échauffement au niveau de l'induit :  $\Delta\theta$

La figure 4.42 présente la position des classes de  $X_{a4}$  et de la classe test  $\Omega_4$  dans l'espace de représentation formé par le vecteur forme  $V_{F4norm(AG)}$ .



**Fig. 4. 42.** Position des classes dans l'espace formé par  $V_{F4norm(AG)}$ . Machine saine «+», 1 barre cassée «□», 3 barres cassées «◇», 4 barres cassées «○».

La figure 4.43 présente l'interpolation et la prédiction réalisées respectivement par l'approche polynomiale et l'approche de Kalman dans l'espace de représentation à 3 dimensions défini par le vecteur  $V_{F4norm(AG)}$ . En rouge est représentée l'interpolation entre les centres de gravité des classes de  $X_{a4}$ , en bleu la prédiction polynomiale et en magenta la prédiction réalisée par Kalman. La prédiction, dans les deux cas, commence à partir de la classe  $\Omega_3$  et est réalisée jusqu'à la classe  $\Omega_4$ . En d'autres termes, la prédiction est faite pour déterminer l'évolution de la défaillance 'barres cassées' passant de trois à quatre barres cassées.



**Fig. 4. 43.** Interpolation et prédiction obtenues par l'approche polynomiale « - » et de Kalman « - » dans l'espace défini par le vecteur  $V_{F4norm(AG)}$

Sur la figure 4.43, la prédiction réalisée par l'approche de Kalman apparaît encore de meilleure qualité que celle obtenue par l'approche polynomiale. Le point « ● » représentant la prédiction '4 barres cassées' est proche physiquement du fonctionnement réel « + ».

Pour définir la précision de la prédiction, la distance Euclidienne séparant la prédiction pour quatre barres cassées avec la position effective du centre de gravité de la classe test  $\Omega_4$  est calculée. Les résultats sont résumés dans le tableau 4.26.

	Approche	
	Polynomiale	Kalman
Distance Euclidienne	0.9	0.1

**Tableau. 4. 26.** *Précision de la prédiction dans les deux cas (polynomiale et Kalman)*

Ce tableau permet de constater l'augmentation de la précision offerte par l'approche de Kalman. Le tableau 4.27 montre la contribution de l'approche de Kalman comparée à l'approche polynomiale. Les résultats sont présentés en pourcent.

	Polynomiale $\rightarrow$ Kalman
Augmentation de la précision	+ 88.9 %

**Tableau. 4. 27.** *Comparaison de la prédiction obtenue dans les deux cas (polynomiale et Kalman)*

La prédiction réalisée par l'approche de Kalman permet d'estimer l'état manquant (2 barres cassées) et de prédire l'évolution de cette défaillance vers des degrés de sévérité plus importants. D'après le tableau 4.26, les trois points utilisés pour initialiser le système (réalisation de l'interpolation) suffisent pour obtenir une bonne précision. L'estimation du mode '2 barres cassées' n'a pas pu être validé car nous ne disposons pas du matériel nécessaire (rotor contenant 2 barres cassées). L'horizon de la prédiction, au-delà de 4 barres cassées, n'a pas pu être étudié pour les mêmes raisons. Il aurait été intéressant de connaître la précision obtenue en s'éloignant vers un horizon plus lointain et ainsi déterminer jusqu'à quel nombre de barres cassées la prédiction est encore acceptable.

#### 4.6.2. Détection des défauts sur l'onduleur de tension

Il a été montré que les méthodes de reconnaissance de formes peuvent être appliquées avec succès pour le diagnostic des machines asynchrones alimentées par un convertisseur de puissance. Comme il a été précisé au chapitre I, de plus en plus d'associations convertisseur – machine sont utilisées dans le domaine industriel notamment pour la variation de vitesse. Il devient donc nécessaire de réaliser un diagnostic aussi bien sur la machine que sur son alimentation (en général onduleur de tension).

Nous allons ainsi nous intéresser à la défaillance qui présente le taux de pannes le plus élevé, en l'occurrence, la défaillance des condensateurs électrolytiques.

En raison de problèmes survenus sur la commande des interrupteurs de l'onduleur de tension, nous n'avons pu effectuer qu'un nombre restreint de mesures à tension réduite. Ainsi les mesures ont été réalisées sur une machine saine, alimentée par un onduleur sans défaut et alimentée par l'onduleur présentant un vieillissement des condensateurs de filtrage d'entrée. Ce vieillissement se caractérise par l'augmentation de la résistance équivalente série ESR (§ 1.5, chapitre I). Pour réaliser cette défaillance, ne disposant que de condensateurs neufs, nous avons supprimé une partie des condensateurs constituant le filtre d'entrée LC. Cette démarche simule une augmentation de l'ESR et a pour conséquence l'augmentation de l'ondulation de la tension d'entrée redressée.

#### 4.6.2.1. Présentation des caractéristiques de l'onduleur de tension

L'onduleur de tension utilisé est un onduleur pouvant débiter jusqu'à 45 kW (figure 4.44).

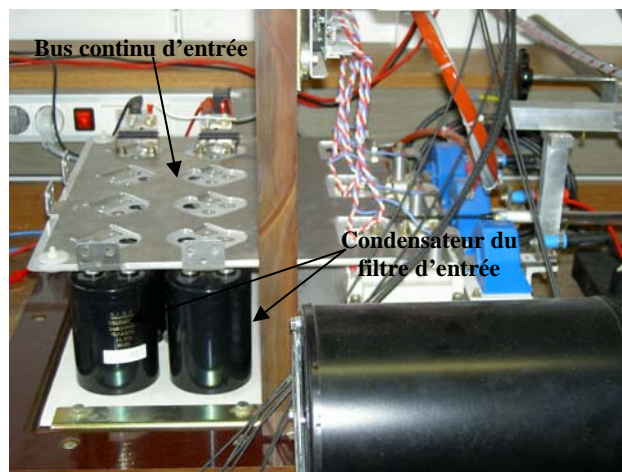
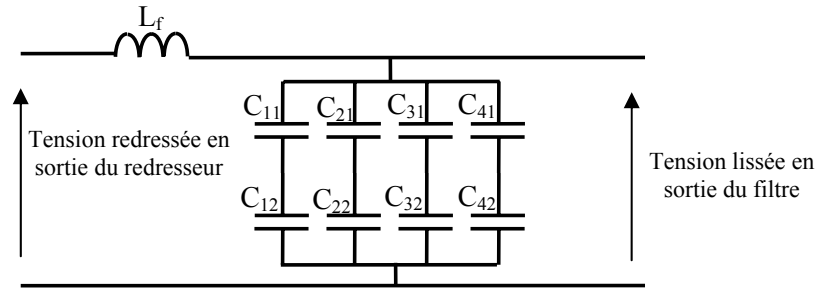


Fig. 4. 44. Onduleur de tension utilisé

La commande réalisée est une commande vectorielle. La partie qui nous intéresse est le filtrage de la tension d'entrée. Cette tension est obtenue à partir d'un redresseur tout thyristor commandé en pleine onde. Ce redresseur est alimenté par le réseau domestique triphasé. Le filtre d'entrée de l'onduleur est constitué d'une inductance de lissage de 100 mH et de huit condensateurs montés en série-parallèle. Les caractéristiques de ces condensateurs sont : 3300  $\mu$ F, 400 V, 10,1A, 85°C.

La figure 4.45 représente schématiquement le filtre d'entrée de l'onduleur.





**Fig. 4. 45.** Filtre d'entrée de l'onduleur

L'ESR de chaque condensateur a été caractérisée à l'aide d'un impédancemètre à une fréquence de 10 KHz. Les valeurs sont présentées dans le tableau suivant :

	C <sub>11</sub>	C <sub>12</sub>	C <sub>21</sub>	C <sub>22</sub>	C <sub>31</sub>	C <sub>32</sub>	C <sub>41</sub>	C <sub>42</sub>
Valeur de l'ESR (mΩ)	26,06	30,7	28,6	31,2	28,5	29,4	31	27,4

**Tableau. 4. 28.** Valeur de l'ESR pour chaque condensateur du filtre d'entrée

A partir de ces valeurs, il est possible de définir la valeur de l'ESR totale,  $ESR_{tot}$ , formée par l'ensemble de ces condensateurs. Pour cela,  $ESR_{1eq}$  ( $ESR_{11}+ESR_{12}$ ),  $ESR_{2eq}$  ( $ESR_{21}+ESR_{22}$ ),  $ESR_{3eq}$  ( $ESR_{31}+ESR_{32}$ ) et  $ESR_{4eq}$  ( $ESR_{41}+ESR_{42}$ ) sont calculées. Au final l'ESR totale est égal à :

$$ESR_{tot}(sain) = \frac{\left( \frac{ESR_{1eq} \times ESR_{2eq}}{ESR_{1eq} + ESR_{2eq}} \right) \times \left( \frac{ESR_{3eq} \times ESR_{4eq}}{ESR_{3eq} + ESR_{4eq}} \right)}{\left( \frac{ESR_{1eq} \times ESR_{2eq}}{ESR_{1eq} + ESR_{2eq}} \right) + \left( \frac{ESR_{3eq} \times ESR_{4eq}}{ESR_{3eq} + ESR_{4eq}} \right)} = 14,66 m\Omega \quad (4.35)$$

Le vieillissement du condensateur est caractérisé en partie par l'augmentation de son ESR. En général, le constructeur préconise un changement des condensateurs lorsque la valeur de l'ESR a doublé par rapport à la valeur initiale. En pratique, pour accélérer ce vieillissement, il est nécessaire de placer le condensateur dans une enceinte climatique à des températures élevées mais cette opération est longue avant d'obtenir un vieillissement représentatif. Par conséquent, pour simuler ce défaut, nous avons supprimé six condensateurs sur les huit présents (seul reste C<sub>11</sub> et C<sub>12</sub>). De ce fait, l'ESR totale est augmentée d'un facteur 4.

$$ESR_{tot}(vieilli) = 56,76 m\Omega \quad (4.36)$$

Le choix d'un facteur d'augmentation élevé (facteur 4) a été fait dans un premier temps, pour déterminer s'il était possible de diagnostiquer un vieillissement des capacités à partir des

paramètres décrits dans le tableau 4.3. De plus, comme l'onduleur est dimensionné pour débiter jusqu'à 45 kW, ses composants sont surdimensionnés par rapport à la puissance de la machine (5 kW). Par conséquent, pour être sûr de visualiser une défaillance, nous avons opté pour la simulation d'un fort vieillissement.

Pour la suite, la sélection des paramètres discriminants vis-à-vis des modes de fonctionnement contenus dans l'ensemble d'apprentissage est réalisée pour des données normalisées. Ensuite la règle des  $k - ppv$  non exclusive est appliquée pour réaliser le diagnostic de l'ensemble convertisseur – machine. Les modes de fonctionnement, ainsi que les observations utilisées afin d'illustrer ces résultats, sont définis dans les tableaux 4.29 et 4.30 respectivement.

Classe	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$\Omega_1$	Machine saine	0 %	10
		50 %	10
$\Omega_2$	Vieillissement condensateur de filtrage	0 %	10
		50 %	10

**Tableau. 4. 29.** Ensemble d'apprentissage utilisé pour caractériser un défaut convertisseur

Ainsi, l'ensemble d'apprentissage  $X_{a5}$  est constitué de deux modes de fonctionnement représentés par deux classes (regroupant les deux niveaux de charge 0 et 50 %) dans l'espace de représentation. Chaque classe est composée de 20 échantillons.

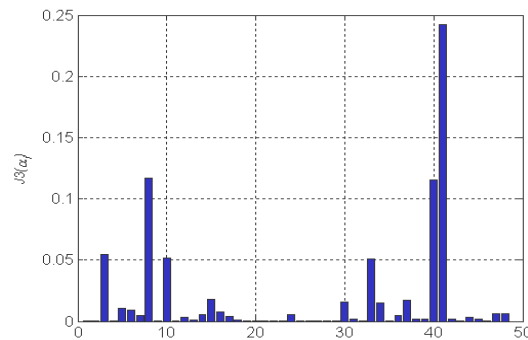
Chaque ensemble d'observations est composé de 10 échantillons, 5 par niveau de charge.

Observations	Mode de fonctionnement	Niveau de charge	Nombre d'échantillons
$C_1 : X_{u1}$ à $X_{u10}$	Machine saine	0 %	5
		50 %	5
$C_2 : X_{u11}$ à $X_{u20}$	Vieillissement condensateur de filtrage	0 %	5
		50 %	5

**Tableau. 4. 30.** Observations utilisées pour valider la classification lors d'un défaut convertisseur

#### 4.6.2.2. Sélection des paramètres

Tout d'abord, avant d'utiliser l'algorithme AG pour la sélection, le rapport de Fisher est utilisé pour déterminer le pouvoir discriminant de chaque paramètre par rapport à l'ensemble d'apprentissage  $X_{a5}$ . La figure 4.46 présente ces résultats.



**Fig. 4. 46.** Pouvoir discriminant de chaque paramètre par rapport à l'ensemble d'apprentissage du tableau 4.20

Sur cette figure, on constate que six paramètres  $[\alpha_3, \alpha_8, \alpha_{10}, \alpha_{33}, \alpha_{40}, \alpha_{41}]$  ont un pouvoir de séparation des modes de fonctionnement contenus dans  $X_{a5}$  avec un paramètre  $\alpha_{41}$ , largement plus discriminant.

On peut en déduire que la dimension optimale du vecteur forme peut aller de 2 à 6.

L'algorithme de sélection AG converge assez rapidement (25 itérations) et donne comme solution optimale, le vecteur forme suivant :

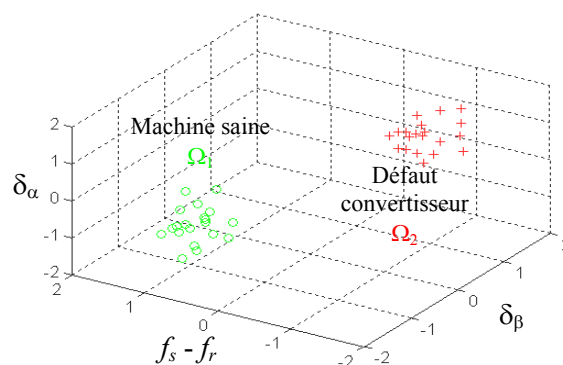
$$V_{F5norm(AG)} = [\alpha_{41}, \alpha_8, \alpha_{40}] \tag{4.37}$$

Ce résultat paraît conforme aux attentes avec la sélection de trois des six paramètres présentant un fort pouvoir de discrimination et une dimension optimale comprise entre 2 et 6, en l'occurrence  $d'=3$ .

Les paramètres de ce vecteur  $V_{F5norm(AG)}$  correspondent respectivement à :

- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\beta$  :  $\delta_\beta$
- amplitudes de la fréquence  $f_s - f_r$
- déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  dans l'axe  $\alpha$  :  $\delta_\alpha$

La figure 4.47 représente l'ensemble d'apprentissage  $X_{a5}$  du tableau 4.29 dans l'espace défini par le vecteur forme  $V_{F5norm(AG)}$ .



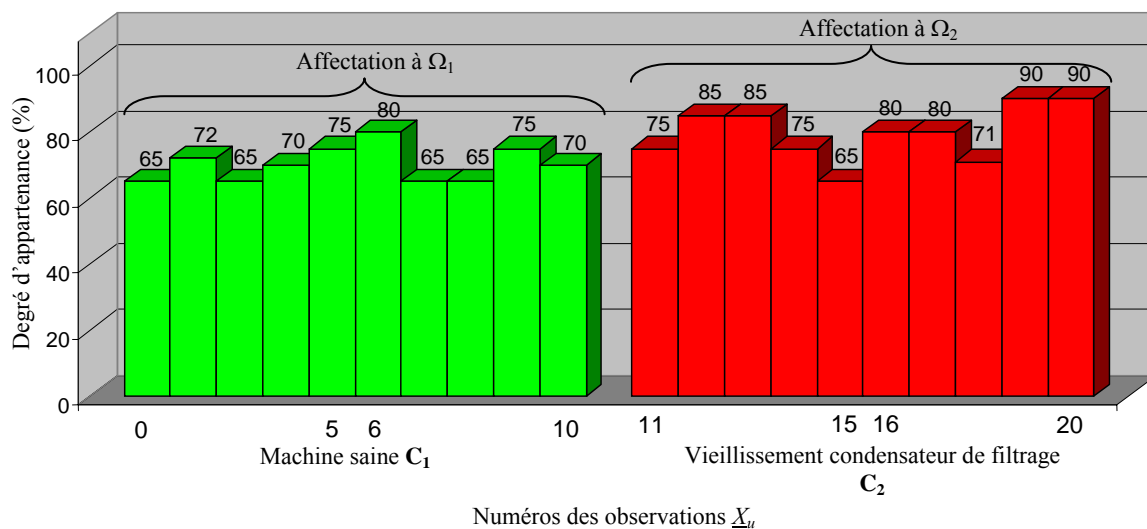
**Fig. 4. 47.** Position des classes dans l'espace de représentation formé par le vecteur  $V_{F5norm(AG)}$ . Machine saine «o», défaut convertisseur«+».

Cette figure montre que les paramètres sélectionnés permettent de différencier très nettement le fonctionnement sain de l'ensemble 'onduleur – machine asynchrone' avec le fonctionnement présentant une défaillance de l'onduleur (vieillessement des condensateurs de filtrage). La classification en sera d'autant plus facile.

#### 4.6.2.3. Classification par la règle des $k - ppv$ non exclusive

Le résultat de la classification des observations du tableau 4.30 est présenté sur la figure 4.48.

La fonction d'appartenance de chaque classe est définie par l'équation 4.25. Les paramètres réglables ont pour valeurs :  $g=1$ ,  $\beta=2$ ,  $k_c=20$  et  $\eta=1/N_c$  avec  $N_c$ , le nombre d'échantillons dans la classe  $\Omega_c$ , ici  $N_c=20$ , d'où  $\eta=0.05$ . Le rejet d'ambiguïté et le rejet d'appartenance sont définis par les équations 4.26 et 4.27 respectivement.



**Fig. 4. 48.** Répartition des nouvelles observations réalisées par la règle des  $k - ppv$  non exclusive

La constatation, faite précédemment, se confirme. En effet, le taux de classification est maximal, aucune observation n'est mal classée. Le fait d'avoir des paramètres qui permettent une bonne séparation des modes de fonctionnement facilite la décision d'affectation. Le degré d'appartenance de chaque observation à sa classe d'affectation est bon, variant entre 65 et 90%. Ce degré donne une bonne fiabilité à la classification réalisée, aucune ambiguïté n'est possible.

Malgré le nombre restreint de mesures réalisées en fonctionnement défaillant du convertisseur, il a été possible de montrer que certains paramètres contenus dans le vecteur

forme peuvent aussi servir à détecter des défauts sur le convertisseur de puissance. Pour pouvoir distinguer d'autres types de défauts sur l'onduleur, il sera sans doute nécessaire d'ajouter de nouveaux paramètres. Le choix de ceux-ci pourra être effectué en réalisant une analyse sur les conséquences possibles que peuvent engendrer certains défauts. Cela fait partie des perspectives de cette étude.

## 4.7. Conclusion

Les résultats ont mis en évidence les performances d'une approche par Reconnaissance de Formes pour diagnostiquer les différents états de la machine mais aussi du convertisseur. Les performances de cette méthode dépendent fortement du choix de la signature (autrement dit du vecteur forme).

L'algorithme AG, dans cette étape de sélection des paramètres les plus pertinents, a donné les meilleurs résultats. En effet, le diagnostic des ensembles test était meilleur dans certains cas, comme la discrimination des différentes classes représentant l'apparition et l'évolution d'un déséquilibre d'alimentation (5%, 10%, 20% et 40% de déséquilibre) au stator (§ 4.5.2.4) et aussi bon dans d'autres cas comme la discrimination de différentes défaillances (barre cassée, usure de roulements et déséquilibre d'alimentation de 10%) par rapport au fonctionnement sain (§ 4.5.1.1).

Dans le cas où aucune connaissance sur le partitionnement des données en classes n'est disponible (mode non supervisé), l'algorithme développé (§ 4.5.1.2) permet de déterminer simultanément le choix des paramètres les plus discriminants et la structuration en classes de l'ensemble d'apprentissage.

Nous avons testé les algorithmes de décision ( $k$ -ppv) exclusive et non exclusive. Ces derniers permettent d'associer un degré d'appartenance de l'observation à une ou plusieurs classes. Ainsi, il est possible d'avoir une information sur la localisation de cette observation par rapport aux classes mais aussi de définir un indice de fiabilité accordé à l'affectation réalisée. De plus, pour toutes les affectations des observations réalisées dans ce chapitre, la réduction du temps de calcul a été appliquée en utilisant le maillage de l'espace de représentation (§ 3.5.1.3, chapitre III). Le gain de temps obtenu par cette approche, dans chaque étape de classification (§ 4.5.2.2, § 4.6.1.2 et § 4.6.2.3), est de l'ordre de 60 à 80%.

Ces algorithmes de décision sont appliqués avec les rejets d'ambiguïté et de distance (ou d'appartenance). Cette notion de rejet est très importante dans la mesure où elle permet

d'enrichir la base de connaissance et de construire de nouvelles classes. Ceci nécessite une analyse des observations rejetées, soit de manière automatique pour la détection de nouveaux modes de fonctionnements (§ 4.5.2.3), soit par l'intervention d'un expert.

La base connaissance peut aussi être enrichie en faisant l'estimation des états manquants. Pour cela, l'estimateur de Kalman a été appliqué au suivi d'évolution de la sévérité des défauts. Dans un premier cas, cet estimateur est utilisé pour prédire l'évolution du fonctionnement vers des zones inconnues de l'espace (§ 4.5.2.4). La fonction est donc la prédiction. Dans un deuxième temps, l'estimateur est utilisé pour déterminer un état manquant dans l'évolution du mode de fonctionnement (§ 4.6.1.3). Sa fonction est donc l'estimation. Dans les deux cas, les résultats de l'estimation et de la prédiction présentent une bonne précision (par rapport à une simple interpolation ou extrapolation). Cette approche pourrait aussi permettre de diminuer le nombre d'acquisitions nécessaires pour construire la base de données initiale.

Le système de diagnostic devient alors évolutif et prédictif.



## Conclusion générale

Etant donné les fréquences d'utilisation des machines asynchrones dans diverses applications, une détection rapide et précoce de leurs défaillances est nécessaire. Leur association avec un convertisseur de puissance type onduleur de tension étant de plus en plus utilisée, il devient nécessaire de prendre en compte les interactions entre ces deux éléments. Le travail présenté dans ce mémoire expose l'apport des méthodes de Reconnaissance des Formes au diagnostic des défauts de cet ensemble.

Pour aborder l'étude, nous avons établi un état de l'art des défauts pouvant intervenir aussi bien dans la machine asynchrone que dans l'onduleur de tension et des méthodes classiques permettant de les identifier. Ce travail préliminaire a mis en évidence deux catégories d'approches en vue du diagnostic spécifique de la machine asynchrone :

- le premier type d'approches consiste à utiliser des modèles analytiques du processus machine asynchrone. Ceux-ci fournissent une estimation en ligne des grandeurs non mesurables, inhérentes à la machine et variant en fonction des états de cette dernière.
- La seconde catégorie d'approches consiste à analyser les signaux mesurés sur la machine en cours de fonctionnement. Les analyses sont réalisées, pour la plupart, dans le domaine fréquentiel et en régime permanent. Les informations issues de ces analyses permettent souvent une identification efficace des défaillances présentes au stator, au rotor et sur les paliers (roulements).

Des différents types d'approches (décrites dans le chapitre I et II) sont extraits un ensemble de paramètres représentatifs des défaillances du processus. Ces paramètres, appelés aussi signature caractéristique du système, forment le vecteur forme initial. Ceux-ci servent de point de départ à la mise en place d'un éventuel système de décision.

La deuxième partie (chapitre III) du travail de la thèse consiste à introduire la procédure de mise en œuvre d'un système de décision basé sur la Reconnaissance des Formes. Cette méthodologie est générique : elle peut s'appliquer à tout système complexe (être humain, réseau téléphonique, véhicule automobile...). La mise en place d'un tel système de diagnostic nécessite l'existence d'une base de données initiale contenant les mesures réalisées sur le processus. De cette base sont extraits un ensemble d'apprentissage et des ensembles de test, totalement disjoints. Différentes méthodes sont utilisées pour effectuer le choix des



paramètres pertinents par rapport aux défauts contenus dans l'ensemble d'apprentissage. Le critère de Fisher indique le pouvoir de discrimination de chaque paramètre par rapport à ces modes de fonctionnement. La méthode "Sequential Backward Selection" ou bien encore la méthode AG utilisent des indicateurs (critères) permettant de juger de la qualité des sous – ensembles de paramètres retenus. Ceux-ci forment le vecteur forme optimal, appelé aussi signature caractéristique du processus à surveiller. Toutes ces méthodes supposent l'espace de décision (structure en classes de l'ensemble d'apprentissage) connu. Dans le cas contraire une étape de coalescence doit être réalisée simultanément avec la sélection des paramètres pour obtenir la structuration en classes ainsi que les paramètres pertinents. Le vecteur forme optimal permet de construire l'espace de représentation.

L'ajout d'une règle de décision permet de classer automatiquement une mesure ou observation, c'est à dire de décider automatiquement du mode de fonctionnement que représente cette mesure. L'algorithme des  $k$  – plus proches voisins ( $k$  – PPV) consiste à associer une mesure à la classe majoritairement représentée parmi les mesures voisines les plus proches (au sens d'une distance qui peut être euclidienne ou de Mahalanobis par exemple). L'association à cette règle, d'une fonction d'appartenance définie pour chaque classe, permet de déterminer un indice de fiabilité (confiance) à l'affectation réalisée.

De plus, deux types de rejet, en ambiguïté et en distance, permettent d'affiner le diagnostic en réduisant le risque de mauvaise classification et en donnant la possibilité de rendre adaptatif le système de décision. De ce fait, des méthodes d'actualisation de la connaissance sont appliquées sur les observations rejetées, pour déterminer l'apparition d'une ou plusieurs nouvelles classes inconnues de la base de données initiales. Elles sont ensuite intégrées dans l'ensemble d'apprentissage après identification par un expert.

Le bon fonctionnement d'une approche diagnostic par des méthodes de reconnaissance des formes nécessite une bonne connaissance du système, ce qui se traduit par une base de données importante. Or, pour de multiples raisons, notamment le faible retour d'expérience, il est difficile de disposer de mesures correspondant à chaque mode de fonctionnement du processus. L'estimation de ces états manquants et la prédiction de l'évolution de certains modes peuvent ainsi être réalisées en appliquant l'approche de Kalman. Cette approche permet de prédire l'évolution de la sévérité d'une défaillance vers un horizon plus ou moins lointain et, de ce fait, permet d'anticiper sur une éventuelle panne destructrice pour le processus.

La dernière partie de la thèse (chapitre IV) consiste à appliquer cette méthode générique de diagnostic par RdF à la machine asynchrone. L'objectif est de proposer un système

automatisé de diagnostic de la machine asynchrone. Au niveau de l'application, sont considérés certains défauts créés au stator et au rotor ainsi que sur l'onduleur de tension. Les défauts étudiés sont les cassures de barres au rotor, le déséquilibre d'alimentation et l'usure (vieillesse naturelle) des roulements pour la machine, et un vieillissement des condensateurs de filtrage d'entrée concernant l'onduleur. La machine est alimentée à vide et à plusieurs niveaux de charge pour une alimentation prélevée soit sur le réseau, soit par le biais d'un onduleur de tension. Les mesures accessibles en direct sur la machine sont les courants et les tensions électriques.

Deux bibliothèques de mesure ont été constituées :

- l'une dans le cas sans onduleur
- l'autre avec onduleur

Dans un premier temps, une normalisation des données est réalisée. Elle permet de s'affranchir de la dispersion due à la variation de la charge. Ainsi, la seule dispersion existante représente les différents modes de fonctionnement. Cette normalisation facilite le diagnostic en optimisant le nombre de classes présentes dans l'espace de représentation.

Cet espace est défini par le vecteur forme. Le nombre élevé de paramètres dans ce vecteur conduit à une dimension trop importante et éventuellement à une mauvaise définition des observations. De ce fait les algorithmes de sélection sont utilisés pour rechercher les paramètres les plus pertinents et pour réduire le temps de calcul. Trois méthodes de sélection de paramètres sont utilisées afin de choisir le vecteur forme optimal dans le cas supervisé :

- la première utilise le calcul du critère de Fisher pour chaque paramètre,
- la seconde méthode est la procédure appelée "Sequential Backward Selection".
- la troisième est la méthode à base d'algorithme génétique

Si la structure en classes de l'ensemble d'apprentissage n'est pas connue (mode non supervisé), une méthode basée sur les algorithmes génétiques multi-population, combinant simultanément la coalescence et la sélection des paramètres est développée.

Pour pouvoir détecter la présence d'un défaut sur la machine ou sur l'onduleur, nous avons appliqué une règle de décision : la règle des  $k - ppv$ . Cette règle de décision exclusive donne de très bons résultats de diagnostic pour les modes de fonctionnement considérés. Toutefois, elle ne permet pas de localiser précisément l'observation. Une amélioration est apportée en ajoutant une fonction d'appartenance à chaque classe. Ainsi, un indice de fiabilité est défini à chaque affectation réalisée, et fournit une indication sur le positionnement à sa classe d'affectation.

Les options de rejet dans les règles de décision jouent un rôle prépondérant dans la détection des nouveaux modes de fonctionnement. Les observations rejetées peuvent alors être intégrées dans l'ensemble d'apprentissage afin d'enrichir la base de données du système de diagnostic.

Enfin, l'approche de Kalman donne de très bons résultats sur la prédiction de l'évolution de la sévérité de certains défauts. En associant à cette approche, les options de rejets appropriées, l'ensemble peut servir de règle de décision.

En conclusion, ce travail a permis de monter l'efficacité de l'application des méthodes de reconnaissance des formes au diagnostic des machines asynchrones et à l'apparition de défaillances sur l'onduleur. Au final, une approche globale du diagnostic d'un système électrique a été réalisée.

Les travaux réalisés dans le cadre de cette thèse pourraient donner lieu à des études supplémentaires sur plusieurs points :

- Les développements de notre méthode:
  - Calculer d'autres paramètres spécifiques aux défaillances de l'onduleur, pour améliorer la pertinence du vecteur forme. Etudier la répercussion des défauts onduleur et les interactions possibles sur le fonctionnement de la machine.
  - Observer la précision de la prédiction obtenue par l'approche de Kalman vers des horizons de l'espace de représentation plus lointain. Valider cette prédiction pour d'autres défaillances du système.
  - Etendre l'approche diagnostic proposée à l'identification des défaillances sur d'autres types de machines électriques notamment la machine synchrone.
- Les nouvelles voix d'études :
  - Prédiction de la durée de vie en introduisant une prise en compte de l'histoire (chaîne de Markov).
  - Extension du diagnostic au régime transitoire (variation de vitesse, ...).
  - Passer du diagnostic à la commande en mode dégradé.

---

## BIBLIOGRAPHIE

---



## ***Contribution Personnelle***

### ***Revues internationales***

O. Ondel, E. Boutleux, G. Clerc, " Diagnostic évolutif par reconnaissance des formes: application aux machines asynchrones " *Revue International de Génie Electrique (RIGE)*, Vol.9, n°2-3, pp.299-331, Mars/Juin 2006.

O. Ondel, E. Boutleux, G. Clerc, " Un système adaptatif de diagnostic et de suivi d'évolution: application aux machines asynchrones " *Revue International de Génie Electrique (RIGE)*, Parution courant 2006.

O. Ondel, E. Boutleux, G. Clerc, " A method to detect broken bars in induction machine using pattern recognition techniques ", *IEEE Transactions on Industrial Application Society (IEEE IAS)*, Vol. 42, n°4, pp.916-923, July/August 2006.

O. Ondel, E. Boutleux, G. Clerc, E. Blanco, " FDI based on Pattern Recognition using Kalman Prediction: application to an induction machine ", *Soumise à IFAC EAAI (Engineering Applications of Artificial Intelligence)*, Elsevier.

O. Ondel, E. Boutleux, G. Clerc, " Diagnosis system of a set "inverter – induction machine – load" by pattern recognition ", *International Society for Computers and Their Applications (ISCA)*, Parution en 2006.

### ***Conférences Internationales (avec comités de lecture et actes)***

O. Ondel, E. Boutleux, P. Venet, "A decision system for electrolytic capacitors diagnosis ", *Proc. IEEE 35<sup>th</sup> Annual Power Electronics Specialists Conference (PESC)*, Vol.6, pp.4360-4364, June 2004.

O. Ondel, A. Yazidi, E. Boutleux, G. Clerc, H. Henao, R. casimir, G.-A. Capolino, "Comparative Study of Two Diagnosis Methods for Induction Machine", Proc. *IEEE International Conference on Industrial Technology (ICIT)*, Vol.1, pp.159-165, Tunisia, December 8-10, 2004.

O. Ondel, E. Boutleux, G. Clerc, " Adaptive diagnosis by pattern recognition: application on an induction machine ", Proc. *IEEE SDEMPED 2005 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 367 - 373, 7-9 September 2005 Vienna, Autriche

E. Blanco, O. Ondel, A. Llor, " Using linear Interpolation and Kalman Prediction in Pattern Recognition: application to an induction machine", Proc. *IEEE SDEMPED 2005 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 95 - 100, 7-9 September 2005 Vienna, Autriche

O. Ondel, E. Boutleux, G. Clerc, " Feature Selection by Evolutionary Computing: application on Diagnosis by Pattern Recognition", *International Society for Computers and their Applications (ISCA)*, 18<sup>th</sup> *International Conference on Computer Applications in Industry and Engineering (CAINE 2005)*, Vol.1, pp.219-225, November 9-11, 2005, Hawaii, USA

O. Ondel, E. Boutleux, G. Clerc, "Diagnosis system of a set “converter-induction machine-load” by pattern recognition", *International Society for Computers and their Applications (ISCA)*, 18<sup>th</sup> *International Conference on Computer Applications in Industry and Engineering (CAINE 2005)*, Vol.1, pp.350-357, November 9-11, 2005, Hawaii, USA

O. Ondel, E. Boutleux, G. Clerc, " Application of pattern recognition method to the diagnosis in induction machine " Proc. *International Conference on Electrical Machines (ICEM)*, 2 - 5 September 2006, Chania, Grèce.

O. Ondel, E. Boutleux, G. Clerc, " Use of data standardization to improve inverter – induction machine fault detection " *32<sup>nd</sup> Annual conference of the IEEE Industrial Electronics Society (IEEE IECON'06)*, 7 - 10 Novembre 2006, Paris, France.

***Conférence Nationale (avec comités de lecture et actes)***

O. Ondel, "Un système adaptatif de diagnostic et suivi d'évolution : application aux machines asynchrones" *Conférence de Jeunes Chercheurs en Génie Electrique (JCGE'05)*, Montpellier, 5 et 6 juin 2005 pp 30 – 35.

O. Ondel, E. Boutleux, G. Clerc, "Application de la reconnaissance des formes au diagnostic des machines asynchrones" *Société de l'Electricité, de l'Electronique et des Technologies de l'information et de la communication (SEE)*, Paris, 23 et 24 Mars 2006.





## **Chapitre I et Introduction générale**

- [Allianz, 88] "Présentation des avaries. Machines électriques tournantes" cahier des préventions. CP2, Allianz, 1988
- [Baudesson, 98] P. Baudesson, F. Richardeau, T. Meynard, "Caractérisation du convertisseur multicellulaire série en présence de défauts de commutation" *Congrès Electronique de Puissance du Futur*, pp. 129-134, 16-18 décembre 1998
- [Beguenane, 97] R. Beguenane et M. E. H. Benbouzid, "Détection de défauts dans les machines asynchrones en utilisant la signature spectrale du courant statorique" *Journée Vibrations et bruit acoustique des machines électriques*, LESIR, ENS-Cachan, Avril 1997
- [Besson, 90] R. Besson, "Technologie des composants électroniques" *Radio*, 6<sup>ème</sup> édition. Tome 1, pp. 107-114 et 184-212, 1990
- [Bigret, 94] R. Bigret et J. L. Féron, "*Diagnostic – maintenance – disponibilité des machines tournantes*" Editions MASSON. Paris. Décembre 1994
- [Bonal, 97] J. Bonal, "Entraînements électriques à vitesse variable", édition Lavoisier, Prométhée Groupe Schneider, Vol.1, pp.144-152, 1997.
- [Bonal, 98] J. Bonal, G. Séguier, "Entraînements électriques à vitesse variable", édition Lavoisier, Prométhée Groupe Schneider, Vol.2, pp.373-377 et 388-395, 1998.
- [Bonnett, 92] A. H. Bonnett, G. C. Soukup, "Cause and analysis of stator and rotor failures in three-phase squirrel cage induction motors" *IEEE Transactions on Industry Applications*, Vol.28 no.4, pp 921-937, Juillet-Août 1992
- [Bonnett, 93] A. H. Bonnett, "Cause and analysis of Anti-Friction Bearing Failures in A.C Induction Motors" *IEEE Transactions on Industry Application*, pp 14 - 23, Sept/Oct 1993
- [Cabanas, 96] M. F. Cabanas, M. G. Melero, J.G. Aleixandre et J. Solares, "Shaft misalignment diagnosis of induction motors using current spectral analysis : A theoretical approach" *Proc. International Conference on Electrical*

- Machines (ICEM)*, pp 384-390, 10 - 12 Septembre 1996, Vigo, Espagne.
- [Calis, 99] H. Calis et P. J. Unsworth, "Fault diagnosis in induction motors by current signal analysis" *IEEE SDEMPED 1999 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp.237-241, Septembre 1999 Gijon, Espagne
- [Devanneaux, 2001] V. Devanneaux, H. Kabbaj, B. Dagues et J. Faucher, "An accurate model of squirrel cage induction machines under static, dynamic or mixed eccentricity" *Proc. IEEE SDEMPED 2001 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp.121-126, 1-3 Septembre 2001 Grado, Italie
- [Dubuisson, 90] B. Dubuisson, *Diagnostic et reconnaissance des formes*. Traité des nouvelles technologies. Série diagnostic et Maintenance. Hermès, Paris, 1990
- [Elkasabgy, 92] N. M. Elkasabgy, A. R. Eastham, G. E. Dawson, "Detection of broken bars in the cage rotor on an induction machine" *IEEE Transactions on Industry Applications*, Vol.28, no.1, Part 1, pp. 165-171, Jan.-Fév. 1992
- [Hamdi, 94] E. S. Hamdi, "Design of small electrical machines" *Wiley*, 1994
- [Henao, 99] H. Henao, G.-A. Capolino, M. Garcia-Melero et M. F. Cabanas, "A new model of the induction motor rotor cage for diagnostics" *IEEE SDEMPED 1999 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp383-388, Septembre 1999 Gijon, Espagne
- [Joksimovic, 00] G. M. Joksimovic, J. Penman, "The detection of inter-turn short circuits in the stator windings of operating motors" *IEEE Transactions on Industrial Electronics*, Vol.47, no.5, pp. 1078-1084, Octobre 2000
- [Kerszenbaum, 84] L. Kerszenbaum et C. F. Landy, "The existence of large interbar currents in the three phase squirrel cage motors with rotor-bar and end-ring faults" *IEEE Transactions on Power Apparatus Systems*, Vol. PAS-103, n°7, pp.1854-1862, July 1984.
- [Muller, 94] G. H. Muller et C. F. Landy, "A novel method to detect rotor bars in squirrel

- cage induction motors when interbar currents are present" *IEEE Power Engineering Society Meeting*, Vol. 3, number of pages: 4, July 2003.
- [Naït-Saïd, 00] M. S. Naït-Saïd, M. E. H. Benbouzid et A. Benchaid, "Detection of broken bars induction motors using an extended Kalman filter for rotor resistance sensorless estimation" *IEEE Transactions on energy conversion*, Vol.15, no.1, pp. 66-70, Mars 2000
- [Nandi, 99] S. Nandi, H. A. Toliyat, "Condition monitoring and fault diagnosis of electrical machines – a review" *IEEE Industry Applications Conference. Thirty-Fourth IAS Annual Meeting*, Vol.1, pp.197-204, Phoenix, USA, 1999
- [O'Donnell, 85] P. O'Donnell, "Report of large reliability survey of industrial and commercial installations, Part 1", *IEEE Transaction on Industry Applications*, Vol. IA 21, N° 4, july/august 1985 pp 853 - 864
- [Pauling, 51] L. Pauling, "Chimie générale" *Dunod. Paris* :1951, pp. 587
- [Raison, 2000] B. Raison, "Détection et localisation de défaillances sur un entraînement électrique" *Thèse de doctorat de l'Institut National Polytechnique de Grenoble*, Septembre 2000
- [Rifa, 88] Rifa, "Capacitors. Data sheets" pp.80-85, 1988
- [Salles, 97] G. Salles, " Surveillance et diagnostic des défauts de la charge d'un entraînement par machine asynchrone" *Thèse de doctorat de l'Université Claude Bernard Lyon1*, Janvier 1997
- [Saunders, 95] L. A. Saunders, K. S. Crawford, R. A. Epperly et G. Murphy, "Applying adjustable frequency controllers to high-starting – torque loads" *IEEE Transactions on Industry Applications*, Vol.31, no.1, pp.68-76, Janvier-Février 1995
- [Schoen, 97] R. Schoen et T. Habetler, "A new method of current-based condition monitoring in induction machines operating under arbitrary load conditions" *Electrical Machines and Power Systems*, Vol.25, n°2, pp.141-152, 1997.

- [Sic Safco, 93], Sic Safco, "Condensateurs électrolytiques aluminium. Données constructeurs" pp. 12-29, 1993
- [Sobczyk, 98] T. J. Sobczyk, A. Izvorski, "Recognition of rotor eccentricity of induction motor based on the fourier spectra of phase currents" *Proc. ICEM'98*, pp. 408 – 413 , Vol 1, september 2 - 4 1998 Istanbul Turkey
- [Thomson, 99] W. T. Thomson, "A review of on-line condition monitoring techniques for three-phase squirrel-cage induction motors: Past, present and future" *Proc. IEEE SDEMPED 1999 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp.3-18, Septembre 1999 Gijon, Espagne
- [Thorsen, 95] O. V. Thorsen, M. Dalva, "A survey of fault on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries" *IEEE Transactions on Industry Applications*, Vol.31, no.5, pp.1186-1196, September 1995
- [Venet, 93] P. Venet, "Surveillance d'alimentation a découpage. Application à la maintenance prédictive" *Thèse de doctorat de l'Université Claude Bernard Lyon1*, Décembre 1993
- [Walliser, 94] R. F. Walliser et C. F. Landy, "Determination of interbar current effect in the detection of broken bars in squirrel cage induction motors" *IEEE Transactions on Energy conversion*, Vol. 9, Issue 1, pp.152-158, March 1994
- [Yahoui, 95a] H. Yahoui, J. Seetohul, G. Grellet, A. Jammal, "Detection of broken bar or end-ring fault in asynchronous machines by spectrum analysis of the observed electromagnetic torque through supply cable" *Revue Européenne de Diagnostic et Sécurité de Fonctionnement*, Vol.5, no.4, 1995
- [Zadeh, 65] L. A. Zadeh, "Fussy sets", *Information and Control* 8, pp. 338-353, 1965

**Chapitre II**

- [Alty, 91] J. L. Alty, L. I. Mikulich, "Industrial applications of artificial intelligence" *2<sup>nd</sup> ed.*, Kluwer Academic Publisher, North Holland, 1991.
- [Andria, 94] G. Andria, M. Savino, A. Trotta, "Application of Wigner – Ville distribution to measurement on transient signals" *IEEE Transactions on Instrumentation and Measurement*, Vol. 45, Issue 4, pp. 187-193, April 1994.
- [Ah-Jaco, 97] A. Ah-Jaco, "Modélisation des moteurs asynchrones triphasés en régime transitoire avec saturation et harmoniques d'espace. Application au diagnostic", *Thèse de doctorat, Université Claude Bernard – Lyon1*, Juillet 1997.
- [Ahrens, 93] S. Ahrens, "Système expert d'aide au diagnostic et à la maintenance de grands entraînement électrique" *Thèse de doctorat – Institut National Polytechnique de Grenoble*, 1993.
- [Arkan, 01] M. Arkan, D. K. Perovic, and P. Unsworth, "Online stator fault diagnosis in induction motors", *Proc. Inst. Elect. Eng., Electr. Power Appl.*, Vol. 148, pp. 537-546, November 2001.
- [Assaf, 01] T. Assaf, H. Henao, G. A. Capolino, "Detection of voltage source dissymmetry using the measurement of symmetrical components in working induction motors", *IEEE SDEMPED'01 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives* pp. 441 – 447, 1 – 3 September 2001, Grado Italy.
- [Barros, 01] D. Barros, M. Castilla, and J. C. Montano, "Wavelet and neural structure: A new tool for diagnostic of power system disturbances" *IEEE Trans. Ind. Appl.*, Vol. 37, no. 1, pp. 184-190, Jan./Feb. 2001.
- [Beguenane, 97] R. Beguenane, M. E. H. Benbouzid, "Détection de défauts dans les machines asynchrones en utilisant la signature spectrale du courant statorique" *Journée vibrations et bruit acoustique des machines électriques*, LESIR, ENS-Cachan, Avril 1997.

- [Benbouzid, 99] M. E. H. Benbouzid, M. Vieira, C. Theys, "Induction motor's faults detection and localization using stator current advanced signal processing techniques" *IEEE Transaction on Power Electronics*, Vol. 14, N° 1, pp 14 – 22, January 1999.
- [Bigret, 94] R. Bigret & J. L. Féron, *Diagnostic – maintenance – disponibilité des machines tournantes*. Editions MASSON. Paris. Décembre 1994.
- [Boudaoud, 97] N. Boudaoud, "Conception d'un système de diagnostic adaptatif en ligne pour la surveillance des systèmes évolutifs" *Thèse de doctorat Université de technologie de Compiègne*, 1997.
- [Boumégoura, 99] T. Boumégoura, H. Yahoui, G. Clerc, G. Grellet, "Oservation des paramètres du moteur asynchrone à cage d'écureuil avec un observateur non linéaire" *Colloque EF'99* - pp 375 – 379 Lille 30&31 mars 1999.
- [Bourbao, 96] E. Bourbao, "Adapter votre analyse temps – fréquence à chaque signal" *Mesures N°685*, pp. 79, Mai 1996.
- [Boutleux, 96] E. Boutleux, "Diagnostic et suivi d'évolution de l'état d'un système, par reconnaissance des formes floues. Application au modèle du réseau téléphonique français" *Thèse de doctorat Université de technologie de Compiègne*, 1996.
- [Cabanas, 98] H. F. Cabanas, M. Garcia-Melero, "Experimental applications of axial leakage flux to the detection of rotor asymmetries, mechanical anomalies and interturn short-circuits in working induction motors" *Proc ICEM'98*, Vol 1/3, September 2 -4, 1998, Istanbul Turkey.
- [Calis, 99] H. Calis, P. J. Unsworth, "Fault diagnosis in induction motors by motor current signal analysis" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 237 – 241, September 1 –3, Gijon, Spain.
- [Cameron, 86] J. R. Cameron, W. T. Thomson, "Vibration and current monitoring for detecting airgap eccentricity in large induction motors" *IEE Proceedings B. Electric Power Applications*, Vol. 133, pp.155-163, May 1986.

- [Casimir, 03] R. Casimir, "Diagnostic des défauts des machines asynchrones par reconnaissance des formes" *Thèse de l'Ecole Centrale de Lyon, CEGELY, 2003*.
- [Casimir, 03a] R. Casimir, E. Boutleux, G. Clerc, "Fault diagnosis in an induction motor by pattern recognition methods" *SDEMPED2003\_Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp.294-299, Atlanta, GA, USA, 24-26 August 2003.
- [Casimir, 03b] R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, "Broken bars detection in an induction motor by pattern recognition", *Proc. IEEE Power Tech'03 International Conference*, Vol.2, pp 282 - 288, 23-26 June 2003, Bologna Italy.
- [Chang, 94] C. H. Chang, "A fast method for determining electrical and mechanical qualities of induction motors using a PC-aided detector" *IEEE transactions on Instrumentation and Measurement*, Vol. 43, Issue 4, pp. 624-629, Aug. 1994.
- [Chow, 95] T. W. S. Chow, G. Fei, "Three phase induction machines asymmetrical faults identification using bispectrum" *IEEE Transactions on Energy Conversion*, Vol. 10, Issue 4, pp. 688-693 , Dec. 1995.
- [Cruz, 00] S. M. A. Cruz and A. J. M Cardoso, "Rotor cage fault diagnosis in three-phase induction motors by Extended Park's Vector approach" *Electrical Machines and Power Systems*, Vol.28, no 4, pp. 289-299, 2000.
- [Deleroi, 82] W. Deleroi, "Squirrel cage motor with broken bar in the rotor – physical phenomena and their experimental assesment" *Proc. International Conference on Electrical Machines (ICEM'82)*, Budapest, 1982, pp. 767-771.
- [Diallo, 05] D. Diallo, M. E. H. Benbouzid, D. Hamad, X. Pierre, "Fault detection and diagnosis in induction machine drive: A pattern recognition approach based on Concordia stator mean current vector" *IEEE Transactions on Energy Conversion*, Vol. 20, Issue 3, pp. 512-519, Sept. 2005.
- [Eykhoff, 74] P. Eykhoff, "System Identification – Parameter and State Estimation" *Wiley-Interscience* – 1974.



- [Farreny, 89] H. Farreny, "*Les systèmes experts – Principes et exemples*" Cépadues, 1989.
- [Filbert, 90] D. Filbert, "Technical diagnosis of electrical drive systems" *Measurement*, Vol. 8, N°3, Jul.-Sep. 1990.
- [Filippetti, 88] F. Filippetti, G. Franceschini, M. Martelli, C. Tassoni, "An approach to knowledge representation about induction machine diagnosis in expert systems" *Proc. Int. Conf. Electrical Machines*, Vol. 3, pp. 289-296, Pisa, Italie, Sep. 1988.
- [Filippetti, 93] F. Filippetti, "Neural networks approach to electric machine on-line diagnostics" *Proc. 5<sup>th</sup> Eur. Conf. Power Electron. Applicat.*, pp 213-218, 1993.
- [Filippetti, 95] F. Filippetti, G. Franceschini, C. Tassoni, "Neural networks aided on-line diagnostics of induction motor rotor faults" *IEEE Transactions on Industry Applications*, Vol. 31, Issue 4, pp. 892-899, July/August 1995.
- [Filippetti, 00] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "Recent developments of induction motors drives fault diagnosis using AI techniques" *IEEE Transactions on Industrial Electronics*, Vol. 47, no. 5, pp. 994-1003, Oct. 2000.
- [Flandrin, 93] P. Flandrin, *Temps – Fréquence. Traité des nouvelles technologies*. Edition Hermès, Paris, 1993.
- [Franck, 92] P. M. Franck, "Principles of model-based fault detection", *Proc. IFAC/IFIP/IMACS/ International Symposium on Artificial Intelligence in Real-Time Control*, Delft, NL, 1992.
- [Frelicot, 92] C. Frelicot, "Un système adaptatif de diagnostic prédictif par reconnaissance des formes floue" *Thèse de doctorat Université de technologie de Compiègne*, 1992
- [Haji, 01] M. Haji, H. A. Toliyat, "Pattern recognition – A technique for induction machine rotor broken bar detection" *IEEE Transactions on Energy Conversion*, Vol.16, Issue 4, pp. 312-317, December 2001.
- [Heno, 99a] H. Heno, M. Garcia-Melero, "Méthodes de détection de défauts dans les machines à induction par utilisation de la mesure du flux axial" *Colloque*

- EF'99 - Lille, 30 et 31 mars 1999 pp 106 – 111.*
- [Isermann, 84] R. Isermann, "Process fault detection based on modeling and estimation methods", *Automatica*, 30, pp. 387-404, 1984.
- [Jianguo, 92] J. Jianguo, Z. Zhiping, S. Pengshang, W. Xiangheng, "Time frequency spectrum (TSF) of line current during starting process – A tool for diagnosing failure of induction motor", *IEEE ICEM'92*, Manchester, United Kingdom, 15-17 September 1992.
- [Kalman, 61] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory", *Transactions of ASME Journal of Basic Engineering*, Vol.83D, pp. 35-45, March 1961.
- [Khelif, 94] M. Khelif, "Contribution à l'étude et la prédiction des défauts de vieillissement par fatigue thermique des composants d'électronique de puissance", *Thèse de doctorant, Ecole Centrale de Lyon, CEGELY*, 1994.
- [Kral, 99] C. Kral, F. Pirker, "Rotor eccentricity detection of induction machines by torque estimation" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 283 - 287, September 1 –3, Gijon, Spain.
- [Kryter, 89] R. C. Kryter, H. D. Haynes, "Condition Monitoring of Machinery Using Motor Current Signature Analysis" *Sound and Vibration Magazine*, September 1989, pp.14-21.
- [Leith, 88] D. Leith, N. D. Deans, L. I. D. Stewart, "Condition monitoring of electrical machine using real time expert systems" *Proc. Int. Conf. Electrical Machines*, Vol. 3, pp. 297-302, Pisa, Italie, Sept. 1988.
- [Ljung, 99] L. Ljung, "System identification: Theory for the user", 2<sup>nd</sup> ed., PTR Prentice Hall Saddle River, N.J., 1999.
- [Moreau, 99] S. Moreau, J. C. Trigeassou, G. Champenois, "Diagnosis of electrical machines : a procedure for electrical fault detection and localization" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 225 - 229, September 1 –3, Gijon, Spain.

- [Naït-Saïd, 00] M. S. Naït-Saïd, M. E. H. Benbouzid et A. Benchaib, "Detection of broken bars in induction motors using an extended kalman filter for rotor resistance sensorless estimation" *IEEE Transactions on Energy Conversion, Vol.15, no 1, pp.66-70*, Mars 2000.
- [Nandi, 99] S. Nandi, H. A. Toliyat, "Condition monitoring and fault diagnosis of electrical machines – a review" *IEEE Industry Applications Conference. Thirty-Fourth IAS Annual Meeting, Vol.1, pp.197-204*, Phoenix, USA, 1999.
- [Ondel, 05] O. Ondel, E. Boutleux, G. Clerc, "Adaptive Diagnosis by Pattern Recognition: Application on an Induction Machine" *SDEMPED2005\_Symposium on Diagnostics for Electric Machines, Power Electronics and Drives, pp.367-373*, Vienna, Austria, 7-9 September 2005.
- [Pau, 74] L. F. Pau, "Diagnosis of equipment failure by pattern recognition" *IEEE Transactions on Reliability, Vol.1, pp. 750-756*, 1974.
- [Peltier, 93] M. A. Peltier, B. Dubuisson, "A human operator monitoring process based on a fuzzy approach" *In Tooldiag'93, International conference on Fault diagnosis*, 1993.
- [Penman, 94] J. Penman, C. M. Yin, "Feasibility of using unsupervised learning artificial neural networks for the condition monitoring of electrical machines" *IEE Proceedings - Electric Power Applications, Volume 141, Issue 6, p. 317-322*, November 1994.
- [Schaeffer, 99] E. Schaeffer, "Diagnostic des machines asynchrones : modèles et outils paramétriques dédiés à la simulation et à la détection de défauts" *Thèse de doctorat - Ecole Centrale de Nantes 1999*.
- [Schaeffer, 99a] E. Schaeffer, E. Le Carpentier, E. H. Zaim, L. Loron, "Diagnostic des entraînements électriques : détection de courts-circuits statoriques dans la machines asynchrone par identification paramétrique" *17<sup>ième</sup> colloque GRETSI sur le traitement du signal et des images, Vol 4, pp 1037 - 1040*, Vannes, France, 13 - 17 septembre 1999.
- [Schoen, 94] R. Schoen & T. Habetler, "A new method of current-based condition monitoring in induction machines operating under arbitrary load conditions"

- Proc. *International Conference on Electrical Machines (ICEM)*, September 1994, Paris, France.
- [Schoen, 95a] R. R. Schoen, T. G. Habetler, F. Kamran, R. G. Bartheld, "Motor bearing damage detection using stator current monitoring" *IEEE Transaction on Industry Applications*, Vol. 31, N° 6, November/December 1995 pp 1274 – 1279.
- [Schoen, 95b] R. R. Schoen, T. G. Habetler, "Effect of time-varying loads on rotor fault detection in induction machines" *IEEE Transaction on Industry applications*, Vol. 31, Issue 4, pp.900-906, July-Aug. 1995.
- [Schoen, 95c] R. R. Schoen, B. K. Lin, T. G. Habetler, "An unsupervised, on-line induction motor fault detection using stator current monitoring" *IEEE Transaction on Industry Applications*, Vol. 31, N° 6, november/december 1995 pp 1280 – 1286.
- [Sidahmed, 97] M. Sidahmed, "Panorama des méthodes d'analyse et de diagnostic des machines tournantes" *Revue française de mécanique RFM*, n°4, pp.291-298, 1997.
- [Stein, 88] J. Stein, R. D. Endicott, "Noninvasive detection of broken rotor bars in operating induction motors" *IEEE Transactions on Energy Conversion*, Vol. 3, Issue 4, december 1988.
- [Tallam, 01] R. M. Tallam, T. G. Habetler., and R. G. Harley, "Continual on-line training of neural networks with applications in electric machines fault diagnosis" *Proc. IEEE 32<sup>nd</sup> Power Electronics Specialists Conf.*, 2001, pp. 2224-2228.
- [Tavner, 87] P. J. Tavner, J. Penman, "Condition Monitoring of Electrical Machines" *Research Studies Press*, Letchworth, Hertfordshire, pp.286, 1987.
- [Thollon, 93] F. Thollon, G. Grellet, A. Jammal, "Asynchronous motor cage fault detection through electromagnetic torque measurement" *Proc ETEP* Vol. 3, N° 3, September/October 1993.
- [Thomson, 01] W. T. Thomson, M. Fenger, "Current signature analysis to detect induction motor faults" *IEEE Industry Applications Magazine*, Juillet-Août 2001.

- [Thomson, 03] W. T. Thomson, R. J. Gilmore, "Motor current signature analysis to detect faults in induction motor drives – fundamentals, data interpretation, and industrial case histories" *Proceedings of the thirty-second turbomachinery symposium*, pp. 145-156, 2003.
- [Vas, 92] P. Vas, "Electrical machines and drives, a space vector theory approach" Oxford University Press, New York, USA 1992.
- [Vil, 90] D. Vilain, F. Madar, "Surveillance et diagnostic en ligne d'un four à ciment à l'aide d'un système expert temps réel" Actes de la convention IA'91, Paris 1990.
- [Yahoui, 95] H. Yahoui, "Contribution à la surveillance et au diagnostic des machines électriques. Application aux machines triphasées asynchrones à cage." *Thèse de doctorat - Université Claude Bernard - Lyon1 1995*.
- [Yahoui, 96] H. Yahoui, J. Seetohul, G. Grellet, "Detection of broken bar or ring in asynchronous machines by the spectrum analysis of the observed electromagnetic torque" *Electricmacs'96*, Saint Nazaire (France), septembre 1996.
- [Zwingelstein, 95] G. Zwingelstein, "Diagnostic des défaillances. Théorie et pratique pour les systèmes industriels" *Traité des nouvelles technologies - série Diagnostic et Maintenance 1995. Editions HERMES*.

**Chapitre III**

- [Anderson, 79] A. C. Anderson, K. S. Fu, "Design and development of a linear binary tree for leukocytes", *Purdue University, West Lafayette, IN*, Tech. Rep. TR-EE-79-31, 1979.
- [Bartolucci, 76] L. A. Bartolucci, P. H. Swain, C. Wu, "Selective radiant temperature mapping using a layered classifier", *IEEE Trans. Geosci. Electron.*, Vol. GE-14, pp. 101-106, 1976
- [Bezdek, 74] J.C. Bezdek, "Cluster validity with fussy sets", *Journal of Cybernetics*, Vol.3, n°3, pp. 58-72, 1974.
- [Boudaoud, 97] N. Boudaoud, "Conception d'un système de diagnostic adaptatif en ligne pour la surveillance des systèmes évolutifs" *Thèse de doctorat Université de Technologie de Compiègne*, 1997.
- [Boutleux, 96] E. Boutleux, "Diagnostic et suivi d'évolution de l'état d'un système, par reconnaissance des formes floues. Application au modèle du réseau téléphonique français" *Thèse de doctorat Université de Technologie de Compiègne*, 1996
- [Breiman, 84] L. Breiman, J. H. freidman, R. A. Olshen, C. J. stone, "Classification and regression Trees", *Belmont, CA:Wadsworth Int.*, 1984.
- [Brill, 92] F. Z. Brill, D. E. Brown and W. N. Martin, "Fast genetic selection of features for neural network classifiers", *IEEE Trans. Neural Networks*, Vol. 3, n°2, pp. 324-328, 1992.
- [Brown, 92] D. E. Brown, C. L. Huntley, "A practical application of simulated annealing to clustering", *Pattern Recognition*, Vol. 25, n°4, pp.401-412, 1992.
- [Brown, 92a] R. G. Brown, P. Y. C. Hwang, "Introduction to random signals and applied Kalman filtering", *J. Wiley & Sons Edition*, 1992.
- [Casimir, 03] R. Casimir, "Diagnostic des défauts des machines asynchrones par reconnaissance des formes" *Thèse de l'Ecole Centrale de Lyon, CEGELY*, 2003.
- [Chen, 03]. X. Chen, "An improved branch and bound algorithm for feature selection". *Pattern Recognition Letters*, Vol. 24, pp.1925 – 1933, 2003.

- [Chow, 57] C. K. Chow, "An optimum character recognition system using decision functions", *Institute of Radio Engineers (IRE) Trans. On Electronic Computers*, Vol.EC-6, n°4, pp. 247-254, December 1957.
- [Davis, 91] L. Davis, "*Handbook of Genetic Algorithms*". New York: Van Nostrand Reinhold, 1991.
- [Devijver, 77] P. A. Devijver, "Reconnaissance des formes par la méthode des plus proches voisins", *Thèse de 3<sup>ème</sup> cycle, Université Paris 6, juin 1977*
- [Diday, 76] E. Diday, J. C. Simon, "*Cluster Analysis*", dans *Digital Pattern Recognition*, (K. S. FU edition), pp 47 – 94, Springer – Verlag, Berlin, 1976
- [Didelet, 92] E. Didelet, "Les arbres de neurones avec rejet d'ambiguïté. Application au diagnostic pour le pilotage en temps réel du réseau téléphonique français" *Thèse de doctorat Université de Technologie de Compiègne, 1992*
- [Dubes, 79] R.C. Dubes et A.K. Jain, "Validities studies in clustering methodologies." *Pattern Recognition*, N°11, pp. 235-254, 1979.
- [Dubuisson, 90] B. Dubuisson, *Diagnostic et Reconnaissance des Formes*. Traité des nouvelles technologies. Série diagnostic et Maintenance. Hermès, Paris, 1990
- [Dubuisson, 90a] B. Dubuisson, "Decision with reject options", *Eusipco*, Barcelone, 1990.
- [Duda, 73] R. O. Duda, P. E. Hart, "*Pattern Classification and Scene Analysis*", John Wiley and Sons, New York, 1973.
- [Eshelman, 93] L. J. Eshelman, J. D. Schaffer, "Real-coded Genetic Algorithms and Interval-Schemata", *Foundations of Genetic Algorithms*, 2, pp. 187-202, 1993.
- [Fisher, 36] R. A. Fisher, "The use of multiple measurements in taxonomic problems". *Annals of Eugenics*, v. 7, pp 179 – 188, 1936.
- [Friedman, 75] J. H. Friedman, J. L. Bentley, R. A. Finkel, "An algorithm for finding nearest neighbors", *IEEE Trans. On Computer*, Vol.24, pp. 1000-1006, July 1975.
- [Fu, 74] K. S. Fu, "*Syntactic Methods in Pattern Recognition*" Academic Press, Inc,

- New York, 1974
- [Gelb, 74] A. Gelb, "Applied optimal estimation", *MIT Press*, 1974.
- [Goldberg, 94] D. E. Goldberg, "*Algorithmes génétiques. Exploration, optimisation et apprentissage automatique*", Editions : Addison-Wesley France, Juin 1994.
- [Gu, 83] Y. X. Gu, Q. R. Wang, C. Y. Suen, "Application of multilayer decision tree in computer recognition of Chinese characters", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-13, pp. 163-174, 1991.
- [Hall, 99] L. O. Hall, I. B. Ozyurt, J. C. Bezdek, "Clustering with a genetically optimized approach", *IEEE Trans. Evolutionary Computation*, Vol. 3, n°2, pp. 103-112, 1999.
- [Hellman, 70] M. E. Hellman, "The nearest neighbor classification rule with a reject option", *IEEE. Trans on System, Man and Cybernetics*, pp 179 – 185, Vol 6, no 3, July 1970.
- [Holland, 75] J. H. Holland, "Adaptation in natural and artificial systems", *University of Michigan Press*, Ann Arbor, MI, Internal report, 1975.
- [Jain, 88] A.K. Jain et R.C. Dubes, "*Algorithms for Clustering Data*" Prentice Hall, 1988.
- [Kittler, 78] J. Kittler, "A method for determining k nearest neighbors", *Kibernetes*, Vol. 7, pp. 313-315, 1978.
- [Krishna, 99] K. Krishna and M. N. Murty, "Genetic k-means algorithm", *IEEE Trans. On Systems, Man and Cybernetics*, Part B: Cybernetics, Vol. 29, Issue 3, pp.433-439, June 1999.
- [Kudo, 00] M. Kudo, J. Sklansky, "Comparison of algorithms to select features for pattern classifiers". *Pattern Recognition* Vol. 33, Issue 1, pp 25-41 January 2000
- [Kulkarni, 76] A. V. Kulkarni, L. N. Kanal, "An optimization approach to hierarchical classifier design", *Proc. 3<sup>rd</sup> Int. Joint Conf. Pattern Recognition*, San Diego, CA, pp. 459-466, 1976.



- [Landeweerd, 83] G. Landeweerd, T. Timmers, E. Gresema, M. Bins, M. Halic, "Binary tree versus single level tree classification of white blood cells", *Pattern Recognition*, Vol. 16, pp. 571-577, 1983.
- [Li, 86] X. Li, R. C. Dubes, "Tree classifier design with a permutation statistic", *Pattern Recognition*, Vol. 19, n°3, pp. 229-235, 1986.
- [Mahfoud, 92] S. M. Mahfoud, "Crowding and preselection revisited", *Parallel Problem Solving from Nature*, 2, North-Holland, 1992.
- [Michalewicz, 94] Z. Michalewicz, "*Genetic Algorithms + Data Structures = Evolution Programs*", Berlin: Springer-Verlag, seconde édition, 1994.
- [Ouaghram, 99] R. B. Ouaghram, "Contribution au diagnostic des machines tournantes complexes. Application à un laminoir", *Thèse de doctorat Université de Technologie de Compiègne 1999*
- [Pal, 95] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy *c*-means model", *IEEE Trans. Fuzzy Systems*, Vol.3, Issue 3, pp. 370-379, August 1995.
- [Payne, 77] H. J. Payne, W. S. Meisel, "An algorithm for constructing optimal binary decision trees", *IEEE Trans. Comput.*, Vol. C-26, pp. 905-916, 1977.
- [Poublan, 93] H. Poublan, "Reconnaissance automatique de signaux à évolution continue" *Thèse de doctorat, Institut National Polytechnique de Toulouse*, 1993.
- [Quinlan, 86] J. R. Quinlan, "Induction of decision trees", *Machine Learning*, Vol. 1, pp. 81-106, 1986.
- [Raymer, 00] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Khun and A. K. Jain, "Dimensionality reduction using genetic algorithms", *IEEE Trans. Evolutionary Computation*, Vol. 4, n°2, pp. 164-171, 2000.
- [Richetin, 80] M. Richetin, G. Rives, M. Naranjo, "Algorithme rapide pour la détermination des k plus proches voisins", *R.A.I.R.O. Informatique*, Vol. 14, n°4, pp. 369-378, 1980.
- [Safavian, 91] S. R. Safavian, D. Landgrede, "A survey of decision tree classifier methodology", *IEEE Trans. Systems, Man, cyber.*, Vol. SMC-21, pp. 660-674, 1991.
- [Selim, 91] S. Z. Selim, K. Alsultan, "A simulated annealing algorithm for the clustering problem", *Pattern Recognition*, Vol. 24, n°10, pp. 1003-1008, 1991.

- [Somol, 04] P. Somol, P. Pudil and J. Kittler, "Fast branch and bound algorithms for optimal feature selection". *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol.26, Issue 7, pp.900-912, July 2004.
- [Swain, 77] P. Swain, H. Hauska, "The decision tree classifier design and potential", *IEEE Trans. Geosci. Electron.*, Vol. GE-15, pp. 142-147, 1977.
- [Tomek, 76] I. Tomek, "A generalisation of the kNN rule", *IEEE Trans on System, Man and Cybernetics*, Vol. 6, no 2, pp 121 – 126, february 1976.
- [Wilson, 72] D. L. Wilson, "Asymptotic properties of nearest neighbor rules using edited data", *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-2, n°3, pp. 408-421, July 1972.
- [Wu, 75] C. Wu, D. Landgrede, P. Swain, "The decision tree approach to classification", Rep. RE-EE 75-17, *School of Elec. Eng., Purdue University*, West Lafayette, IN, 1975.
- [Yang, 98] J. H. Yang and V. Honavar, "Feature subset selection using a genetic algorithm", *IEEE Intelligent Systems*, Vol. 13, n°2, pp. 44-49, 1998.
- [Zieba, 95] S. Zieba, "Une méthode de suivi d'un système évolutif. Application au diagnostic de la qualité d'usinage" *Thèse de doctorat Université de Technologie de Compiègne 1995*

**Chapitre IV**

- [Casimir, 03] R. Casimir, "Diagnostic des défauts des machines asynchrones par reconnaissance des formes" *Thèse de l'Ecole Centrale de Lyon, CEGELY, 2003.*
- [Casimir, 03 a] R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, "Broken bars detection in an induction motor by pattern recognition", *Proc. IEEE Power Tech'03 International Conference*, pp 282 - 288, 23-26 June 2003, Bologna Italy
- [Casimir, 03 b] R. Casimir, E. Boutleux, G. Clerc, " Fault diagnosis in an induction motor by pattern recognition methods", *IEEE SDEMPED 2003 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 294 - 299, 24-26 August 2003 Atlanta, GA, USA
- [Chang, 94] C. H. Chang, "A fast method for determining electrical and mechanical qualities of induction motors using a PC-aided detector" *IEEE Transactions on Instrumentation and Measurement*, Vol. 43, Issue 4, pp. 624-629, Aug. 1994.
- [Dubuisson, 90] B. Dubuisson, *Diagnostic et Reconnaissance des Formes*. Traité des nouvelles technologies. Série diagnostic et Maintenance. Hermès, Paris, 1990
- [Ondel, 05a] O. Ondel, "Un système adaptatif de diagnostic et suivi d'évolution : application aux machines asynchrones" *Conférence de Jeunes Chercheurs en Génie Electrique (JCGE'05)*, pp 30 – 35, Montpellier, 5 et 6 juin 2005.

# ANNEXE



## Annexe 1 : Calcul des pertes totales et de l'échauffement de la machine asynchrone

Les pertes présentes dans la machine provoquent un échauffement des matériaux et en particulier des isolants.

Il existe une température limite supportable pour chaque type d'isolant. On sépare ceux-ci en classes normalisées et repérées par une lettre (tableau A1.1).

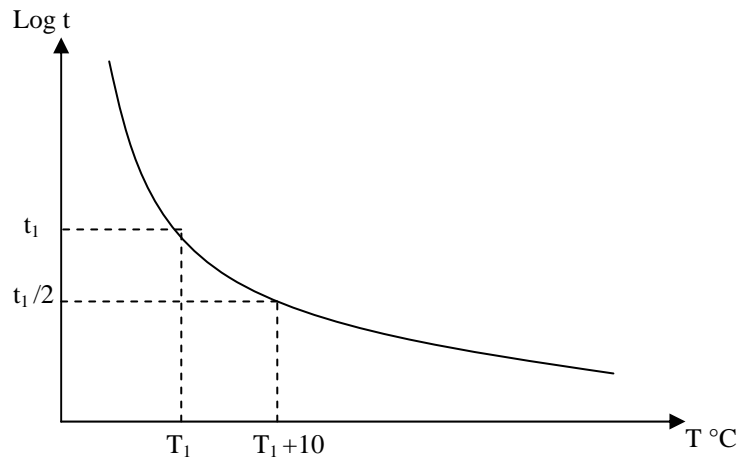
A ces températures limite  $\theta_l$  correspond un échauffement moyen noté  $\theta_{moy}$  défini comme l'augmentation moyenne de la température au-dessus de la température ambiante (généralement normalisée à 40°C). Pour une machine, cet échauffement correspond à celui de ses enroulements, calculé à partir de la mesure de la résistance de ceux-ci. On peut constater que  $\theta_{moy} + 40 < \theta_l$ . L'écart croissant avec  $\theta_l$  tient compte d'éventuels points chauds des enroulements que la méthode ne peut déceler.

Classes	A	B	F	H	C
$\theta_l$ °C	105	130	155	<180	≤180
$\theta_{moy}$ °C	60	80	105	<120	≤120

**Tableau. A1. 1.** *Température limite normalisée et échauffement moyen pour les différentes classes de la machine*

La tenue dans le temps et donc la durée de vie des isolants (égale à celle de la machine), est fortement conditionnée par la tension électrique et par les contraintes vibratoires mécaniques qui dégradent sa texture mais aussi par sa température de fonctionnement. Un échauffement élevé vieillit rapidement les isolants. Le vieillissement correspond à une réaction chimique d'oxydation qui dégrade l'isolant à une vitesse liée à sa température selon la **loi d'Arrhenius**. La durée de vie  $t$  de l'isolant, caractérisée par un seuil de dégradation normalisé, est donnée en fonction de la température absolue à laquelle ce seuil apparaît :

$$\text{Log } t = \frac{a}{T} + b \text{ où } a \text{ et } b \text{ sont des constantes du matériau} \quad (\text{A1.1})$$



**Fig. A1. 1.** Durée de vie de l'isolant en fonction de la température

Pour les matériaux usuels dans la gamme des températures normalisées, à un suréchauffement de 10°C correspond une réduction de 50% de la durée de vie.

### **Modélisation thermique de la machine**

Une machine réelle est un corps complexe avec solides et fluides, non homogène, non isotrope, avec des sources de chaleurs distribuées non uniformément. Les échanges de chaleur à l'intérieur sont principalement de type conductif dans les solides, rayonnants et convectifs dans les fluides internes et externes entourant ces solides. Si l'on n'est intéressé que par une solution approchée rapide et des valeurs limites moyennes des températures des parties principales, il est alors possible de ne considérer qu'un très petit nombre  $N$  de blocs représentant ces différentes parties, de forme simple considérés comme homogènes et isotropes, avec un seul mode de transfert global et une uniformité de la source de chaleur si elle existe, et de les modéliser séparément. Chaque bloc  $i$  répondra à une équation (A1.2) de la forme :

$$P_i = C_{ti} \cdot \frac{d\Delta\theta_i}{dt} + G_{ia} \cdot \Delta\theta_i + \sum_{k=1}^N G_{ik} \cdot (\Delta\theta_i - \Delta\theta_k) \quad (\text{A1.2})$$

- $P_i$  : puissance dissipée par le bloc  $i$  en Watts
- $C_{ti}$  : capacité thermique du bloc  $i$  en  $J/^\circ C$
- $G_{ia}$  : conductance thermique entre le bloc  $i$  et le milieu ambiant en  $J/^\circ C$
- $G_{ik}$  : conductance thermique entre le bloc  $i$  et le bloc  $k$  en  $J/^\circ C$
- $\Delta\theta_i$  : échauffement du bloc  $i$
- $\Delta\theta_k$  : échauffement du bloc  $k$

Cette solution suppose de prendre en compte les conductances globales d'échange avec l'extérieur et chaque autre bloc  $k$ .

La solution générale en supposant les conductances constantes est de la forme (A1.3) :

$$\Delta\theta_i(t) = \sum_1^N \Delta\theta_{im} \cdot \left(1 - e^{-\frac{t}{\tau_{ii}}}\right) \quad (\text{A1.3})$$

- $\Delta\theta_i(t)$  : échauffement du bloc  $i$  en fonction de temps  
 $\Delta\theta_{im}$  : échauffement maximal du bloc  $i$   
 $\tau_{ii}$  : constante de temps thermique du bloc  $i$

Pour les actionneurs tels que les machines électriques, compte tenu de leur faible taille généralement, il est possible de simplifier à l'extrême et de ne considérer que deux cas de figure.

*Si la machine est de construction classique*, avec circuit magnétique d'induit, il est possible d'admettre que les constantes de temps thermiques moyennes de l'induit et de l'inducteur sont du même ordre de grandeur et ainsi que, la machine peut être schématisée par un seul corps (1 seul bloc).

*Si la machine est à induit sans circuit magnétique*, alors les constantes de temps thermiques moyennes d'induit et d'inducteur sont très différentes et la discrétisation ne peut se réduire qu'à deux blocs (modèle à deux corps), l'induit d'une part et l'inducteur et la carcasse d'autre part. Ces machines étant la plupart du temps à courant continu à excitation par aimant permanent, seules les pertes Joule d'induit sont considérées comme conséquentes.

Dans le cadre de notre étude, nous nous placerons dans le cas d'une machine modélisable par un seul corps, ce qui paraît justifiable du fait de la constitution du rotor utilisé (rotor à cage d'écureuil).

### **Modélisation 1 corps**

En première approximation, on considère que la partie active (fer et cuivre) constitue un seul corps homogène et isotrope de propriétés thermiques moyennes et que les sources de chaleur se résument à la puissance perdue  $P_{tot}$  diffusée dans la masse du corps. Alors l'échauffement  $\Delta\theta$  uniforme de ce corps, par rapport à la température ambiante, s'obtient en écrivant que l'énergie perdue pendant le temps  $dt$  sert d'une part à échauffer le corps (quantité de chaleur  $Q_1$ ) et d'autre part à échanger la quantité de chaleur  $Q_2$  avec le milieu ambiant, considéré comme source infinie et donc à température constante.

$$P_{tot} \cdot dt = Q_1 + Q_2 \quad (\text{A1.4})$$

La température du corps s'obtient en ajoutant l'échauffement à la température ambiante.

Si  $M$  est la masse du corps et  $c$  sa chaleur massique, alors  $Q_1$  s'écrit :



$$Q_1 = M \cdot c \cdot d\Delta\theta \quad (\text{A1.5})$$

M.c :  $C_t$  capacité thermique du corps exprimée en J.K<sup>-1</sup>

La quantité de chaleur  $Q_2$  s'échange selon trois processus très différents :

- la conduction, c'est-à-dire par diffusion dans les solides sans déplacement de matière,
- La convection, c'est-à-dire par diffusion au sein d'un fluide avec déplacement de matière,
- Le rayonnement, par émission de radiations au sein d'un gaz.

Les lois de ces trois modes d'échange thermique sont différentes. Mais pour des échauffements inférieurs à 250°C, ces trois lois, en première approximation, peuvent s'identifier les unes aux autres et prennent la forme :

$$Q_i = S \cdot \alpha_i \cdot \Delta\theta \cdot dt \quad (\text{A1.6})$$

où S est la surface normale au flux de chaleur et d'échauffement  $\Delta\theta$ . Le coefficient  $\alpha_i$  est caractéristique du mode de transmission. La quantité de chaleur  $Q_2$  peut alors s'obtenir en sommant les trois coefficients de transmission pour s'écrire :

$$Q_2 = S \cdot \alpha_t \cdot \Delta\theta \cdot dt \quad (\text{A1.7})$$

$\alpha_t = \Sigma \alpha_i$  coefficient total de transmission

$$S \cdot \alpha_t = G_t = \frac{1}{R_t}$$

$G_t$  : conductance thermique

$R_t$  : résistance thermique

Au final, on peut en déduire que la puissance perdue  $P_{tot}$ , équivalente aux pertes de la machine, est de la forme

$$P_{tot} \cdot dt = M \cdot c \cdot d\Delta\theta + S \cdot \alpha_t \cdot \Delta\theta \cdot dt \quad (\text{A1.8})$$

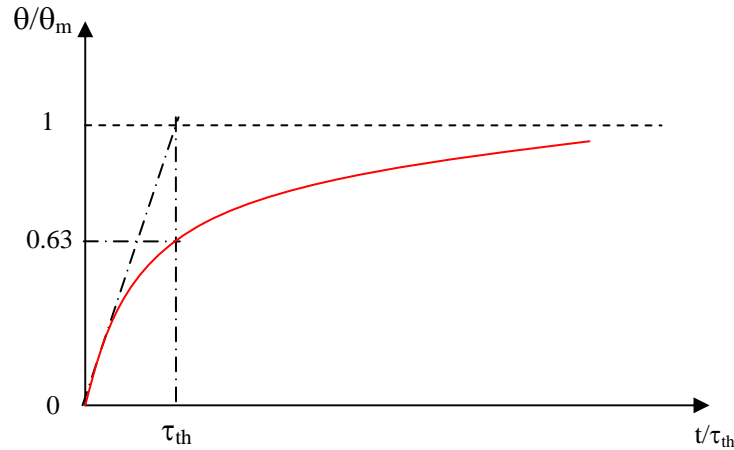
$$P_{tot} = C_t \frac{d\Delta\theta}{dt} + \frac{\Delta\theta}{R_t} \quad (\text{A1.9})$$

Ainsi, l'échauffement peut se déduire de (A1.9)

$$\Delta\theta(t) = \Delta\theta_m \cdot (1 - e^{-\frac{t}{\tau_{th}}}) \quad (\text{A1.10})$$

$\Delta\theta_m = R_t \cdot P_a$  Avec  $\Delta\theta_m$ , échauffement maximal permanent qui ne devra jamais dépasser le  $\theta_1$  limite des isolants

$\tau_{th} = R_t \cdot C_t$  constante de temps thermique



**Fig. A1. 2.** Evolution de l'échauffement de la machine asynchrone

Cependant, la valeur de la résistance d'induit est liée à la variation de température du rotor. L'expression d'une résistance en fonction de la température est donnée par :

$$R = R_a \cdot (1 + \alpha \cdot \Delta\theta), \text{ avec } R_a : \text{valeur de } R \text{ à } T_0 = 40^\circ\text{C} \quad (\text{A1.11})$$

$\alpha$  : coefficient de température de la résistance

De ce fait, la puissance dissipée  $P_{tot}$  va aussi évoluer en fonction de la variation de température de la manière suivante

$$P = P_{tot} \cdot (1 + \alpha \cdot \Delta\theta) \quad (\text{A1.12})$$

Ainsi l'expression (A1.9) va devenir

$$P_{tot} = C_t \cdot \frac{d\Delta\theta}{dt} + \left( \frac{1}{R_t} - P_{tot} \cdot \alpha \right) \cdot \Delta\theta \quad (\text{A1.13})$$

Si on pose  $R_t' = \frac{R_t}{1 - \alpha \cdot \Delta\theta_m}$ , alors l'expression de l'échauffement tenant compte de la

variation de la résistance d'induit en fonction de la température est

$$\Delta\theta(t) = \Delta\theta'_m \cdot (1 - e^{-\frac{t}{\tau_{th}'}}) \quad (\text{A1.14})$$

$\Delta\theta'_m = R_t' \cdot P_{tot} = \frac{\Delta\theta_m}{1 - \alpha \cdot \Delta\theta_m}$  : valeur finale de l'échauffement,

$\tau_{th}' = C_t \cdot R_t' = \frac{\tau_{th}}{1 - \alpha \cdot \Delta\theta_m}$  : constante de temps thermique

Ainsi pour connaître l'échauffement présent au niveau de la machine, il va falloir calculer la puissance dissipée par celle-ci, c'est-à-dire ses pertes.

### Calcul des pertes de la machine asynchrone

A partir de signaux mesurés (courants et tensions), et de certaines données constructeurs telles que la résistance d'enroulement statorique  $R_s$ , la résistance thermique  $R_{th}$  et la capacité thermique  $C_{th}$ , il est possible de déterminer les pertes de puissance auxquelles la machine est soumise (joules, mécanique,...). Avant tout, nous allons présenter les principales puissances mises en jeu sur la machine asynchrone :

- La puissance absorbée :  $P_a = U_{eff} \cdot I_{eff} \sqrt{3} \cdot \cos \varphi$
- Les pertes par effet joule dans le stator :  $P_{js} = 3 \cdot R \cdot I_{eff}^2$
- Les pertes magnétiques dans le stator :  $P_{fs}$

Les pertes magnétiques dans les tôles de son armature (appelées aussi pertes dans le fer) sont pratiquement indépendantes de la charge, en revanche, elles sont liées à la valeur efficace  $U_{eff}$  et à la fréquence d'alimentation  $f_s$ .

Pour un réseau donné, ces pertes sont considérées comme constantes, elles sont soit données, soit calculées par un essai à vide.

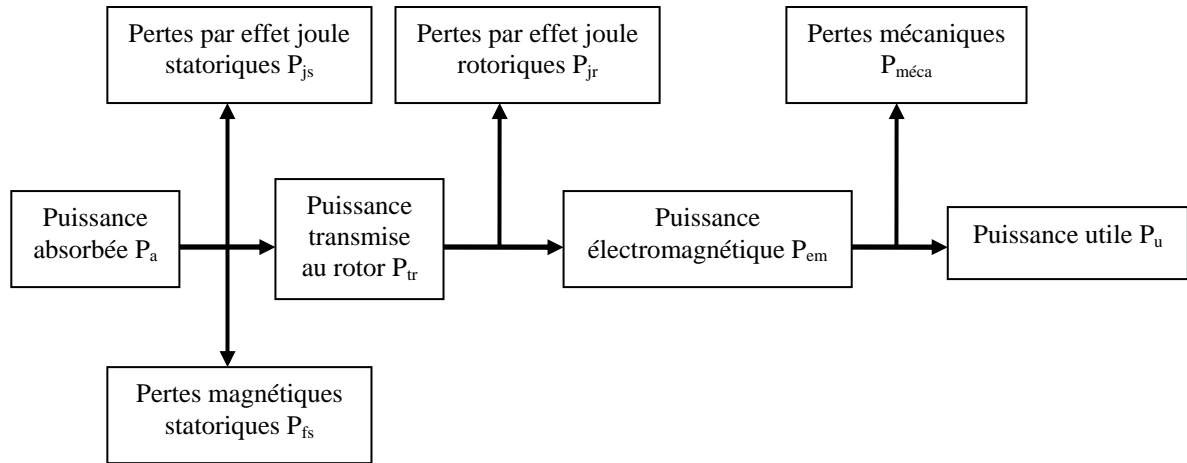
- La puissance transmise au rotor :  $P_{tr} = P_a - P_{js} - P_{fs}$
- Les pertes par effet joule dans le rotor :  $P_{jr} = g \cdot P_{tr}$

Le rotor est également le siège de pertes d'un autre type, les pertes magnétiques. Cependant ces pertes seront toujours négligées devant les autres, la fréquence des courants rotoriques étant très faible.

- La puissance électromagnétique :  $P_{em} = P_{tr} - P_{jr}$
- Les pertes mécaniques :  $P_{méca}$

Ces pertes sont considérées comme constantes, elles sont données ou calculées par un essai à vide.

- La puissance utile :  $P_u = P_a - \sum pertes$



**Fig. A1.3.** Arbre des puissances mises en jeu sur la machine asynchrone

Si on suppose que les pertes mécaniques et les pertes magnétiques statoriques ne sont pas connues, il faut les déterminer à partir d'un essai à vide.

### Pertes à vide

À vide, le couple utile est nul. On calcule la puissance absorbée à vide  $P_{ao}$  et ensuite les pertes par effet joule dans le stator  $P_{jso}$ .

La puissance utile est nulle par définition (couple utile  $C_u = 0$ ) d'où  $P_u = 0$ , alors

$$P_{ao} = P_{jso} + P_{fs} + P_{meca} \quad (\text{A1.15})$$

La puissance absorbée se mesure aisément à partir des signaux tensions et courants mesurés. Connaissant la résistance des enroulements du stator, les pertes par effet joule sont facilement calculables. La différence entre la puissance absorbée et la puissance perdue par effet joule, donne les deux puissances restantes, c'est-à-dire les pertes magnétiques au niveau du stator et les pertes mécaniques.

Plusieurs cas sont possibles :

- Soit les pertes magnétiques dans le stator sont données, on calcule alors la puissance absorbée, les pertes par effet joule dans le stator et on en déduit les pertes mécaniques.
- Soit les pertes mécaniques sont données, on calcule la puissance absorbée, les pertes par effet joule dans le stator et on en déduit les pertes magnétiques dans le stator.

- Soit, le cas le plus fréquent, on suppose que les pertes mécaniques et les pertes magnétiques dans le stator sont égales, on calcule la puissance absorbée, les pertes par effet joule dans le stator et on en déduit les pertes mécaniques ainsi que les pertes magnétiques dans le stator.

Pour notre étude, on se placera dans le dernier cas, où l'on suppose que  $P_{méca} = P_{fs}$ .

On notera ainsi  $P_{cste} = P_{méca} + P_{fs}$ .

Donc pour déterminer ces pertes, dites constantes, il suffit de résoudre l'équation suivante :

$$P_{cste} = P_{méca} + P_{fs} = P_{ao} - P_{jso} \quad (A1.16)$$

### **Pertes en charge**

D'après l'arbre des puissances (figure A1.3), il reste à déterminer les pertes par effet joule statoriques  $P_{js}$  et les pertes par effet joule rotoriques  $P_{jr}$  pour connaître la somme totale des pertes. Ces pertes se calculent à partir d'un essai en charge.

Tout d'abord, on calcule  $P_{js}$  :

$$P_{js} = 3 \cdot R_s \cdot I_{eff}^2 \quad (A1.17)$$

Puis la puissance absorbée

$$Pa = U_{eff} \cdot I_{eff} \sqrt{3} \cdot \cos \varphi \quad (A1.18)$$

De la on peut en déduire la puissance transmise au rotor :

$$P_{tr} = P_a - P_{js} - \left( \frac{P_{cste}}{2} \right) \quad (A1.19)$$

Pour finir on en déduit les pertes joules rotoriques :

$$P_{jr} = g \cdot P_{tr} \quad (A1.20)$$

La somme totale des pertes vaut alors :

$$P_{tot} = P_{js} + P_{jr} + P_{cste} \quad (A1.21)$$

### **Estimation de l'échauffement et de la température**

A partir de l'expression (A1.21), connaissant maintenant la puissance dissipée, il est possible d'utiliser l'expression suivante pour estimer l'échauffement au niveau de l'induit. L'équation suivante ne tient pas compte de la variation de la résistance en fonction de la température :

$$\Delta\theta(t) = R_t \cdot P_{tot} \cdot (1 - e^{-\frac{t}{R_t \cdot C_t}}) \quad (\text{A1.22})$$

Pour tenir compte de cette variation de résistance d'induit, la formulation utilisée sera :

$$\Delta\theta(t) = R_t' \cdot P_{tot} \cdot (1 - e^{-\frac{t}{R_t' \cdot C_t}}) \quad (\text{A1.23})$$

Ces deux formulations donnent l'échauffement présent au niveau de l'induit. Pour connaître la valeur de la température  $\theta$ , il faut additionner à cet échauffement la valeur de la température ambiante  $T_{amb}$  (généralement normalisée à  $T_{amb} = 40 \text{ °C}$ ), ainsi on obtient :

$$\theta = \Delta\theta + T_{amb} \quad (\text{A1.24})$$

Les deux signatures extraites à partir du signal température sont, les pertes totales  $P_{tot}$  présentes dans la machine, ainsi que l'échauffement  $\Delta\theta$  que celles-ci engendrent.







**Titre :**

**Diagnostic par reconnaissance des formes : application à un ensemble ‘convertisseur – machine asynchrone’.**

**Mots clés :**

Diagnostic - prédiction - machines asynchrones - convertisseur de puissance - reconnaissance des formes - règles de décision - suivi d'évolution.

**Résumé :**

Les progrès réalisés en matière d'électronique de puissance, de circuits de commande ont contribué à l'utilisation grandissante des machines asynchrones dans les systèmes d'entraînements électriques. Le recours aux machines asynchrones est surtout lié à leur robustesse, leur puissance massique et à leur coût de fabrication. L'apparition dans les années 1980 des variateurs permettant de faire varier la fréquence de rotation dans une large gamme a largement favorisé son développement. En effet, il entre dans la conception de nombreux procédés industriels associant des convertisseurs statiques et des machines électriques.

La maintenance et la surveillance de ces deux systèmes permettent de rentabiliser les installations. Il est donc important de développer des outils de diagnostic pour détecter de manière précoce les défauts pouvant apparaître aussi bien sur le convertisseur que sur la machine.

Notre approche est basée sur l'utilisation des méthodes de reconnaissance des formes. Un vecteur de paramètres, appelé vecteur forme, est extrait de chacune des mesures effectuées sur la machine. Les règles de décisions utilisées permettent de classer les observations, décrites par le vecteur forme, par rapport aux différents modes de fonctionnement connus avec ou sans défaut.

Des défauts ont été créés au rotor et au stator de la machine asynchrone, alimentée soit à partir du réseau, soit par le biais d'un onduleur de tension.

La procédure de décision, basée sur la règle des  $k$  - plus proches voisins, associée à une fonction d'appartenance, permet de détecter l'évolution des modes de fonctionnements ainsi que les défauts avérés. Par la suite, le suivi d'évolution de ces modes est réalisé par une approche de type Kalman : un estimateur récursif de Kalman est utilisé pour déterminer les paramètres du modèle dynamique rendant compte de l'évolution d'un mode et un prédicteur de Kalman pour prévoir une évolution vers de nouvelles zones de l'espace. Ces algorithmes ont montré l'efficacité de l'application de la reconnaissance des formes au diagnostic.

**Title:**

**Diagnosis by Pattern Recognition: application on a set inverter – induction machine.**

**Keywords:**

Diagnosis - prediction - induction machines - inverter - pattern recognition - decision rules - evolution tracking.

---

**Direction de Recherche :**

Monsieur Guy CLERC, Professeur des Universités  
Centre de Génie Electrique de Lyon (CEGELY) – UMR CNRS 5005  
Université Claude Bernard Lyon 1, bât. Omega  
43 Boulevard du 11 Novembre 1918  
69622 Villeurbanne (France)

<http://cegely.cnrs.fr>