

# وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR –ANNABA UNIVERSITY  
UNIVERSITE BADJI MOKHTAR ANNABA



جامعة باجي مختار عنابة

Année : 2017

Faculté des Sciences de l'Ingénieur

Département d'Électronique

## Mémoire

Présenté en vue de l'obtention du diplôme de Master

Domaine : Sciences et Technologie

Filière : Génie Électrique

Spécialité : Automatique industrielle

## Intitulé

CLASSIFICATION DES DEFAUTS DE ROULEMENTS D'UNE  
MACHINE ASYNCHRONE PAR SEPARATEURS A VASTE MARGE

Présenté par :

Mlle. AMARA KORBA Karima

DEVANT Le JURY

Président : **DEBBACHE Nour Eddine** Professeur UBMA

Directeur de Mémoire : **ARBAOUI Fayçal** MCA UBMA

Examineur : **LAKEL Rabeh** Professeur UBMA



## Dédicaces

---

Afin d'être reconnaissante envers ceux qui m'ont appuyée et encouragée à réaliser ce travail de recherche, je dédie ce mémoire.

À la mémoire de mon défunt Père **AMARA KORBA Abdel - Aziz**.

Je prie notre Seigneur **Allah Le Tout-Puissant, Le Clément et Le Miséricordieux**, pour qu'il lui accorde la félicité éternelle en l'accueillant dans sa vaste demeure lui et tous les défunts membres de ma famille.

Tous les mots ne sauraient exprimer ma gratitude envers ma chère Mère, pour ses encouragements réguliers. Elle, qui m'a appris le respect d'autrui, la gentillesse et la persévérance.

À mes frères **Raouf** et **Adel**, à ma sœur **Nora**, Merci à toute ma famille pour leurs soutiens indéfectibles tout au long de mon parcours de Master.

(Qu'**Allah Le Tout-Puissant**, les protège et les garde).

## Remerciements

---

Au terme de ce travail, je remercie Allah le Tout-Puissant

الله سبحانه وتعالى

De m'avoir donné la volonté et le courage qui m'ont permis de réaliser ce travail, puisse-t-il me guider dans le droit chemin. Louange à L'Éternel.

الحمد لله

Ce travail a été réalisé au Laboratoire d'Automatique et Signaux d'Annaba (LASA) du Département de L'électronique, Université Badji Mokhtar.

Mes premiers remerciements vont au **Dr. Arbaoui F. Maître de conférences**, qui m'a encadrée quotidiennement. Merci d'avoir eu confiance en mes capacités en me proposant ce sujet ô combien intéressant, ouvrant beaucoup de perspectives et durant lequel j'ai acquis de nombreuses compétences. Merci pour votre soutien, pour votre disponibilité, pour la pertinence de vos remarques et votre expérience... Ceux-ci ont beaucoup contribué à la réalisation de ce travail.

Je veux remercier les membres du jury :

**Mr. DEBBACHE N**, Professeur de l'université de Badji Mokhtar Annaba.

**Mr. LAKEL R**, Professeur de l'université de Badji Mokhtar Annaba.

Pour avoir pris le temps de soigneusement étudier ce manuscrit et qui m'ont fait l'honneur d'évaluer ce travail.

Mes collègues **Mr. Drici Djalel** et **Mr. Oudjani Brahim**, attachés de recherche à L'URMA-CRTI Annaba, merci pour m'avoir conseillée tout au long de mon travail et surtout d'avoir été disponibles à chacune de mes sollicitations.

Un merci distinct à ceux qui ont rendu le double cursus Ingénieur- Master possible pour moi et mes collègues.

العنوان: تصنيف عيوب المدرجة الكروية للآلة غير المتزامنة باستعمال فاصل الهوامش الكبير

الملخص:

الخلل الناتج عن العيوب مثل تآكل أو تلف في المدرجة الكروية هو أحد الأسباب الرئيسية لسوء الأداء في الآلات غير المتزامنة، الذي يؤدي إلى خسائر مادية معتبرة.

الغرض من هذه المذكرة هو التشخيص الآلي لعيوب المدرجة الكروية على أساس التحليل الاهتزازي.

الشعاع المكون من المعلمات (الخصائص أو المؤشرات)، أو ما يسمى الشعاع الشكلي، تستخرج مؤشرات الإحصائية والطيفية لنتائج أمثل من الإشارات الاهتزازية الأصلية ومن الإشارات الاهتزازية التي تم معالجتها باستخدام التحليل التجريبي المشروط.

أجريت الدراسة باستعمال قاعدة بيانات جامعة: *Case Western Reserve University* ، حيث تم الحصول على الإشارات الاهتزازية في أوضاع وظروف تشغيل مختلفة (في الحالة السليمة وفي حالة وجود العيب، تغيير في السرعة، تغيير في الحمولة وتغيير في شدة العيب).

في مرحلة التصنيف، تم الاحتفاظ بفاصل الهوامش الكبير متعدد الأصناف.

أجريت مقارنة النتائج بين التحليل الزمني مع وبدون استخدام التحليل التجريبي المشروط، وأظهرت النتائج فعالية هذا النهج.

علاوة على ذلك تمت إجراء دراسة مقارنة مع نوعين آخرين من المصنفات. وقد أثبتت الطريقة المنتقاة، أي فاصل الهوامش الكبير متعدد الأصناف، فعالية أفضل من حيث وقت ومعدل تصنيف بالمقارنة مع الشبكة العصبية الاصطناعية والشبكة العصبية ذات المنطق الغامض.

الكلمات الدالة: فاصل الهوامش الكبير، المعلمات الإحصائية، التحليل التجريبي المشروط، إشارة اهتزازية، عيوب المدرجة الكروية.

## Résumé

**Titre : Classification des défauts de roulements d'une machine asynchrone par séparateurs à vaste marge**

Le défaut de roulement est l'une des causes dominantes de dysfonctionnements dans les machines asynchrones, entraînant une perte économique significative.

L'objectif de ce mémoire est le diagnostic automatisé des défauts de roulements basé sur l'analyse vibratoire.

Un vecteur de paramètres (Indicateurs ou caractéristiques), appelé aussi vecteur forme, est extrait de chacun des signaux de vibration bruts et spectres, il se compose d'indicateurs statistiques et spectraux pour une performance optimale dans un premier temps, ensuite les paramètres sont extraits des signaux traités en utilisant la décomposition modale empirique.

L'étude a été menée sur la base de données de la Case Western Reserve University, où les signaux ont été acquis dans différents modes et conditions de fonctionnement (sain et avec défauts, variations de vitesse, variation de charges et variation de la sévérité du défaut).

Dans la phase de classification, le séparateur à vaste marge multi-classes appelé aussi machines à vecteurs de support multi-classes a été retenue.

Une comparaison des résultats dans le domaine temporel avec et sans décomposition modale empirique (DME) a été effectuée, les résultats obtenus montrent l'efficacité de cette approche.

Ensuite une étude comparative a été effectuée sur deux types de classificateurs, RNA (Réseaux de Neurones Artificiels), RNF (Réseaux Neuro-Flous). On a constaté une meilleure efficacité en termes de temps et taux de classification du multi-SVM par rapport aux RNF et RNA.

**Mots clés :** Machines à Vecteurs de Supports (SVM), Paramètres statistiques, décomposition modale empirique (DME), défaut roulement, signal vibratoire.

**Abstract:**

**Title: Classification of bearings defects of an Induction Machine by support vector machine**

The rolling defect is one of the main causes of dysfunctions in induction machines, resulting in a significant economic loss.

The main purpose of this study is an automated diagnosis of bearing defects based on vibratory analysis.

A vector of parameters (Indicators or characteristics), also known as vector form and composed of statistical and spectral indicators for optimal performance, is extracted from the original vibration signals and from the signals processed using empirical modal decomposition.

The study was carried out on the *Case Western Reserve University* database, where the signals were acquired in different modes and operating conditions (healthy and with defaults, speed variations, load variation and variation in defect severity).

In the classification phase, the Multi-Class Support Vector Machine (SVM) was chosen.

A comparison of the results of the temporal domain with and without empirical modal decomposition was performed, the results obtained show the effectiveness of this approach.

As a result, a comparative study was achieved on two other types of classifiers, ANN (Artificial Neural Network), NFN (Neuro-Fuzzy Network). A better efficiency in terms of time and classification rate goes to the multi-class SVM compared to the ANN and NFN.

Key words: Support Vector Machines (SVM), Statistical parameters, empirical modal decomposition (EMD), rolling failure, vibration signal.

## Liste des figures

Figure	Titre de la figure	Page
<b>Chapitre 1</b>		
1.1	Constitution d'un roulement .....	14
1.2	(a) Ecaillage d'une bague intérieure, (b) Criques de corrosion .....	15
1.3	Photos représentant les différents défauts de roulements. (a) Corrosion, (b) : Cratère bille, (c) : Cratère bague intérieure, (d) : Cratère bague extérieure, (e) : mécanismes généraux d'usure en présence de contaminant .....	16
1.4	Spectre théorique d'un défaut .....	16
1.5	Fréquence caractéristique d'un défaut de type écaillage affectant la bague extérieure .....	17
1.6	Fréquences caractéristiques du roulement à billes .....	18
<b>Chapitre 2</b>		
2.1	Phénomène du fretting .....	21
2.2	Signal de tension de sortie de l'accéléromètre proportionnel à la force de vibration du Crystal piézoélectrique .....	22
2.3	Techniques d'extraction des caractéristiques .....	24
2.4	Principe d'estimation de la FI et de AI via le signal analytique $Z_x(t)$ et $x(t)$ .....	31
2.5	Principe de l'EMD .....	33
2.6	Organigramme du Pseudo-code de l'EMD .....	35
2.7	Principe de tamisage pour l'estimation de la 1 <sup>ère</sup> IMF du signal test.....	36
2.8	Signal temporel test .....	37
2.9	Estimation des IMF du signal test.....	37
<b>Chapitre 3</b>		
3.1	Illustration d'une classification binaire linéaire SVM.....	46
3.2	Illustration du choix de la marge avec un hyperplan optimale.....	47
3.3	Schéma de changement des représentations non linéaires vers un nouvel espace.....	49
3.4	Classification SVM binaire avec le noyau linéaire de Kernel .....	50
3.5	Classification SVM binaire avec le noyau polynomial 2 <sup>ème</sup> degré de Kernel .....	50
3.6	Classification SVM binaire avec le noyau gaussien de Kernel .....	50
3.7	Classification un contre tous (SVM One Vs All).....	52
3.8	Schéma des hyperplans linéaires pour la classification un contre tous (One vs all).....	52
3.9	Classification un contre un (SVM One Vs One).....	53
<b>Chapitre 5</b>		
5.1	Le Banc d'essai de la CWRU.....	65
5.2	Signal original cas sain.....	67
5.3	Auto-corrélation du signal original cas sain.....	67
5.4	Variation des paramètres temporels et fréquentielles du signal sans défauts (sain).....	67

5.5	Signal original avec défaut bille.....	68
5.6	Auto-corrélation du signal avec défaut bille.....	68
5.7	Variation des paramètres temporels et fréquentielles du signal avec défaut bille.....	68
5.8	Signal original avec défaut bague interne.....	69
5.9	Auto-corrélation du signal avec défaut bague interne.....	69
5.10	Variation des paramètres temporels et fréquentielles du signal avec défaut bague interne	69
5.11	Signal original avec défaut bague externe.....	70
5.12	Auto-corrélation du signal avec défaut bague externe.....	70
5.13	Variation des paramètres temporels et fréquentielles du signal avec défaut bague externe	70
5.14	Extraction des 3IMFs les plus pertinents cas sain.....	71
5.15	Spectre du signal cas sain.....	71
5.16	Extraction des 3 IMFs les plus pertinentes du signal avec défaut bille.....	72
5.17	Spectre du signal avec défaut bille.....	72
5.18	Extraction des 3 IMFs les plus pertinentes du signal avec défaut Bague interne.....	73
5.19	Spectre du signal avec défaut bague interne.....	73
5.20	Extraction des IMFs les plus pertinentes du signal avec défaut bague externe.....	74
5.21	Spectre du signal avec défaut bague externe.....	74

## Liste des tableaux

Tableau	Titre du tableau	Page
1	Résumé des articles traitant la base de données de la CWRU .....	57,58,59,60
2	Données vibratoires des défauts avec une fréquence d'échantillonnage de 12K HZ .....	66
3	Récapitulatif des résultats .....	75
4	Récapitulatif des résultats avec les matrices de confusions .....	76

## Table des matières

<b>Dédicaces</b> .....	3
<b>Remerciements</b> .....	4
<b>Résumé / Abstract / ملخص</b> .....	5
<b>Liste des figures</b> .....	8
<b>Introduction générale</b> .....	14
<b>Chapitre 1. Défauts Roulements</b> .....	16
<b>1.1. Roulement :</b> .....	16
<b>1.2. Types des défauts roulements :</b> .....	16
1.2.1. Écaillage :.....	17
1.2.2. Grippage :.....	17
1.2.3. Empreintes de corps roulants :.....	17
1.2.4. Usure généralisée :.....	17
1.2.5. Cratères (pitting) et cannelures : .....	17
1.2.6. Coups, fissures, cassures :.....	17
1.2.7. Corrosion de contact : .....	17
1.2.8. Coloration : .....	18
1.2.9. Détérioration des cages :.....	18
<b>1.3. Fréquences caractéristiques et défauts</b> .....	18
1.3.1. Les paramètres caractéristiques d'un roulement : .....	20
<b>Références Bibliographiques</b> .....	21
<b>Chapitre 2. Analyse Vibratoire</b> .....	23
<b>2.1. Introduction</b> .....	23
<b>2.2. Définition d'une vibration :</b> .....	23
2.2.1. Outils d'analyse des vibrations .....	24
2.2.2. Principe de l'accéléromètre : .....	24
<b>2.3. Les Méthodes d'analyse Vibratoire :</b> .....	25
<b>2.4. Extraction des caractéristiques :</b> .....	26
<b>2.5. Sélection et réduction des caractéristiques :</b> .....	27
2.5.1. Caractéristiques ou paramètres temporelles .....	27
2.5.2. Caractéristiques ou paramètres fréquentielles : .....	28
<b>2.6. Analyse vibratoire par la transformée de Hilbert Huang</b> .....	29
2.6.1. La transformée de Hilbert.....	29
2.6.1.1. Introduction.....	29

2.6.1.2 Signal analytique.....	30
2.6.2. Décomposition modale empirique (DME) .....	32
2.6.2.1. Principe de la décomposition .....	33
2.6.2.2. Définition d'une IMF (Intrinsic Mode Function) : .....	34
2.6.2.3. Critère d'arrêt du processus de tamisage .....	39
2.6.2.4. Les différents Types d'EMD : .....	39
2.6.2.5. Bruit et Valeurs aberrantes : .....	40
2.6.2.6. Les nouvelles approches pour la sélection des IMFs : .....	40
<b>Références Bibliographiques</b> .....	42
<b>Chapitre 3. Multi classification avec les séparateurs à vaste marge</b> .....	45
3.1. Les algorithmes d'apprentissage automatique : .....	45
3.1.2. Algorithmes d'apprentissage non supervisé communs.....	45
3.2. Séparateur à vaste marge (SVM) : .....	46
3.3. Hyperplan, marge et support vecteur : .....	46
3.4. Séparateur à vaste marge dans le cas linéaire : .....	47
3.4.1. Maximisation de marge : .....	47
3.4.2. Calcul de la marge : .....	48
3.4.3. Calcul de l'Hyperplan optimal : .....	48
3.5. Séparateur à vaste marge dans le cas non linéaire .....	50
3.5.1. Les Noyaux : .....	51
3.5.2. Calcul de l'hyperplan optimal : .....	52
3.6. Types multi-class des séparateurs à vaste marge.....	52
3.6.1. Multi classification par l'approche un contre tous.....	53
3.6.2. Multi classification par l'approche un contre un .....	55
<b>Chapitre 4. Revue Bibliographique</b> .....	58
<b>Références Bibliographiques</b> .....	62
<b>Chapitre 5. Résultats et discussion</b> .....	66
5.1. Matériels et méthodes : .....	66
5.1.1. Présentation du Banc d'Essai de la CWRU : .....	66
5.2. Discussion des résultats des classifications .....	78
<b>Conclusion Générale</b> .....	80

# **INTRODUCTION GENERALE**

## Introduction générale

Les machines industrielles automatisées sont composées de divers sous-systèmes, tels que le système de commande électrique, unités de commande(API), les actionneurs (vérins, moteurs asynchrones, distributeurs) ces composants sont impliqués dans l'exécution des opérations désirées des machines.

Les composants structurels d'une machine tels que les roulements à billes, sont soumis à l'usure au fil du temps, la surveillance de ces composants est donc nécessaire pour réduire régulièrement le risque de défaillances et de pannes par un diagnostic anticipatif ou précoce des défauts et assurer la disponibilité et l'utilisation efficace et optimale des machines qui sont les facteurs clés de l'économie de l'entreprise manufacturière.

L'objectif principal de ce mémoire est de mettre en œuvre un algorithme de surveillance de l'état des roulements à l'aide des machines à vecteurs de support.

Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent. Il existe en effet un lien direct entre la théorie de l'apprentissage statistique et l'algorithme d'apprentissage du SVM.

La méthode SVM est une technique de classification particulièrement bien adaptée pour la reconnaissance des formes et traite des données très diversifiées ayant plusieurs dimensions telles que les textes, les images, la voix, le diagnostics médicales, la bio-informatique pour la reconnaissance de gènes ...etc.

Il existe des extensions que nous n'avons pas présentées, parmi lesquelles l'utilisation des SVM pour des tâches de régression, c'est-à-dire de prédiction d'une variable continue en fonction d'autres variables, comme c'est le cas par exemple dans de la prédiction de consommation électrique en fonction de la période de l'année, de la température, etc. Le champ d'application des SVM est donc large et représente une méthode de classification intéressante.

Ce mémoire comporte principalement 5 chapitres. Le premier est consacré à la présentation des différents types de défauts des roulements. Le deuxième chapitre, introduit diverses techniques de traitement des signaux vibratoires, notamment la transformé de Hilbert-Huang en mettant l'accent sur la décomposition empirique pour le calcul de fonctions appelées IMFs (Intrinsic Mode Functions). Dans le troisième chapitre, nous détaillons la classification supervisée par les séparateurs à vaste marge. Le 4<sup>ième</sup> chapitre présente les différents travaux publiés sur la base de données de la CWRU.

Le dernier chapitre présente l'application expérimentale du système de diagnostic sur les données du banc d'essai de la CWRU avec trois techniques de classification : les réseaux de neurones(ANN), les réseaux de neurones flous (RNF) et les séparateurs à vaste marge (SVM). Une comparaison de leurs performances est donnée et les résultats obtenus sont interprétés et discutés

# CHAPITRE

# 1

## **DEFAUTS ROULEMENTS**

## Chapitre 1. Défauts Roulements

### 1.1. Roulement :

Le roulement, organe de base pouvant assurer à lui seul plusieurs fonctions parmi lesquelles la liaison mobile en rotation avec un minimum de frottement entre deux éléments, tout en supportant des charges radiales ou axiales plus ou moins importantes (Morel, J. 1992).

De manière générale, les roulements sont constitués ; de deux bagues (externe et interne) l'une fixe et l'autre tournante avec le rotor. Elles sont séparées par un ensemble d'éléments roulants qui peuvent être des billes, des rouleaux ou des aiguilles selon l'application envisagée, une cage séparant à intervalles réguliers et guidant les éléments roulants.

Chaque type de roulement possède des caractéristiques spécifiques en termes de charges radiales et axiales supportées, de vitesse limite ou encore de défaut d'alignement admissible.

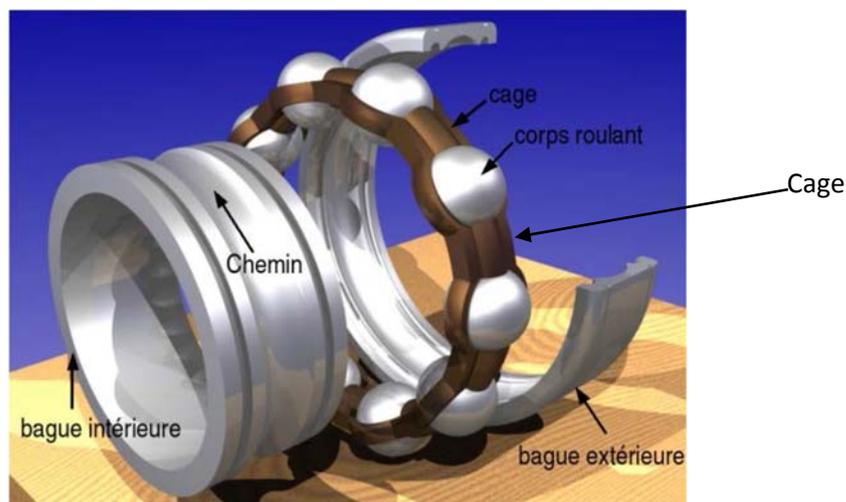


Figure1.1 : Constitution d'un roulement

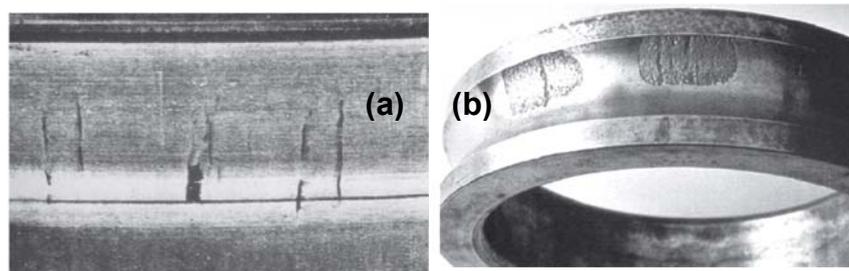
### 1.2. Types des défauts roulements :

Selon leurs applications, et la criticité des machines sur lesquelles ils sont montés, les roulements nécessitent plus ou moins une surveillance et un entretien préventif. Pour cela le suivi des niveaux de vibration et de température, sont les sources d'information les plus utilisées.

Les aspects de ces détériorations se manifestant dans la plupart des cas par un arrachement de matière plus ou moins important, sont décrits comme suit :

### 1.2.1. Écaillage :

Se caractérise par des traces de fissuration et d'arrachement de fragments de matière.



**Figure 1.2 :** (a) Ecaillage d'une bague intérieure, (b) Criques de corrosion

### 1.2.2. Grippage :

En plus de l'arrachement de matière, des traces brunes dues à l'échauffement, des déformations de corps roulants, et des micros fusions peuvent apparaître.

### 1.2.3. Empreintes de corps roulants :

- Par déformation : Des empreintes des corps roulants apparaissent avec un fond brillant, on y retrouve les traces de la rectification mais sans traces d'usure.
- Par abrasion : Dans ce cas les empreintes peuvent ne pas correspondre à l'écartement des corps roulants. L'enlèvement de matière cette fois par usure due à la vibration sans rotation du roulement (faux effet Brinell), ce défaut peut apparaître sur des machines jamais mises en marche (Ilyes, KHELFI. 2014).

### 1.2.4. Usure généralisée :

Apparaissant sur des corps roulants, des pistes et des cages (teinte grise), due à l'intrusion d'une particule abrasive.

### 1.2.5. Cratères (pitting) et cannelures :

Les premiers sont caractérisés par des piqûres il s'agit de trous plus ou moins profonds à bords nets et sont propres aux roulements à billes résulte de la fissuration de la surface par fatigue avec perte de particules, le Pitting est moins à craindre lorsque la viscosité du lubrifiant est bonne quant aux cannelures propres aux roulements à rouleaux, sont caractérisées par une succession de plages étroites parallèles.

### 1.2.6. Coups, fissures, cassures :

Se manifestant avec des traces visibles d'outils, et des ruptures des bagues.

### 1.2.7. Corrosion de contact :

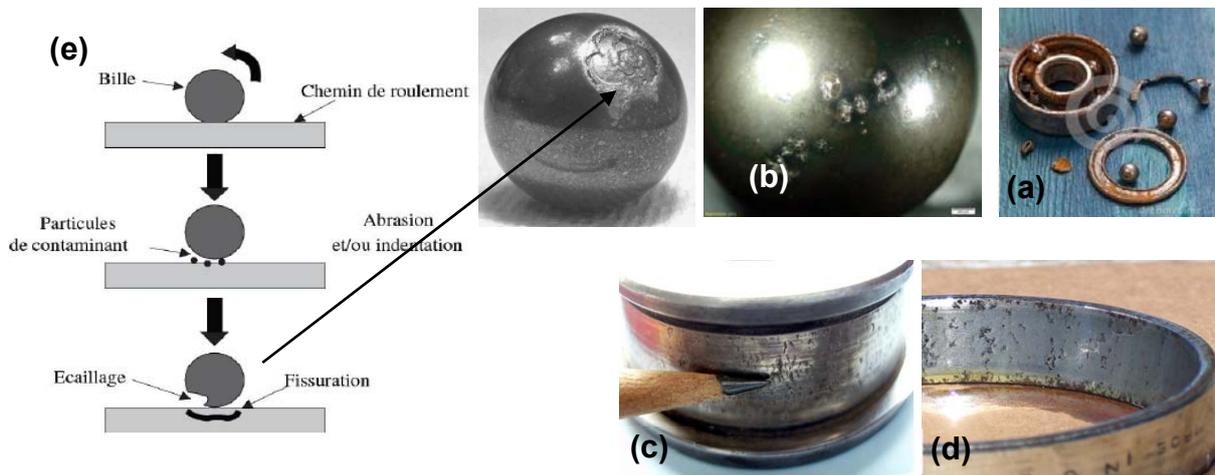
Se manifestant avec des taches d'oxydation sur l'alésage, le diamètre extérieur et les faces d'appui du roulement.

1.2.8. Coloration :

Due à une oxydation localisée ou généralisée elle se manifeste avec une coloration des chemins de roulement et des corps roulants.

1.2.9. Détérioration des cages :

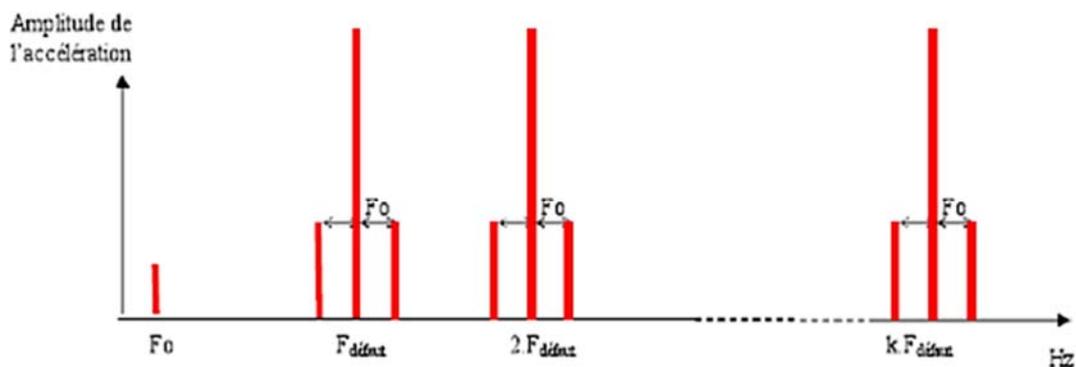
Se manifestant sous plusieurs formes : déformation ; usure ou rupture des cages.



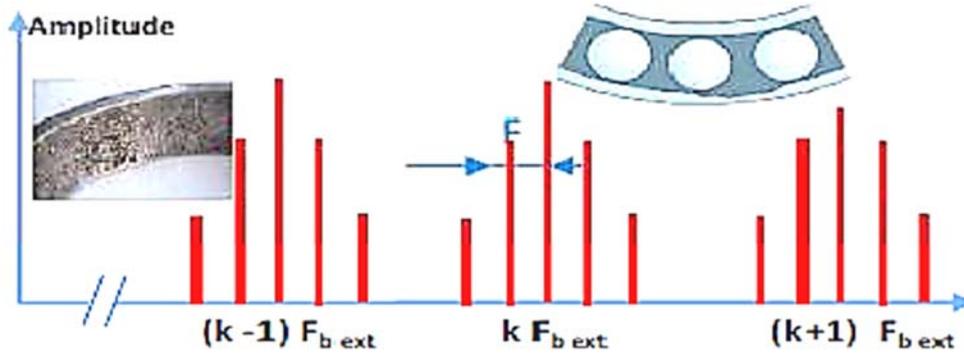
**Figure 1.3 :** photos représentant les différents défauts de roulements. (a) Corrosion, (b) : Cratère bille, (c) : Cratère bague intérieure, (d) : Cratère bague extérieure, (e) : mécanismes généraux d'usure en présence de contaminant.

1.3. Fréquences caractéristiques et défauts

La présence d'un défaut au niveau d'un roulement a un impact sur le spectre de son signal vibratoire tel qu'une augmentation de l'amplitude de la vibration aux fréquences du roulement, mais aussi la présence des vibrations aux harmoniques de ces fréquences ainsi qu'à des fréquences de modulation, le spectre va faire apparaître un peigne de raie à la fréquence du défaut et des bandes latérales à la fréquence de rotation de l'arbre ( $F_0$ ).



**Figure 1.4 :** Spectre théorique d'un défaut (Ishak, 2012)



**Figure 1.5** : Fréquence caractéristique d'un défaut de type écaillage affectant la bague extérieure (Ishak, 2012).

Prenons un exemple où la bague extérieure d'un roulement à billes est endommagée à cause de l'un des mécanismes de défaillance déjà cités. Chaque fois que l'une des billes roule sur le défaut, une force impulsive se produit, et provoque la vibration du roulement. Le roulement répond (sonne) à sa fréquence naturelle (fréquence propre), une réponse qui s'atténue rapidement à cause de l'amortissement.

L'excitation et la réponse se produisent chaque fois qu'une des billes roule sur le défaut, de sorte que la fréquence fondamentale des formes d'onde de réponse sera le taux de passage de la bille sur le défaut.

C'est cette fréquence fondamentale qui nous intéresse dans la détection des défauts de roulement, pas la fréquence sur laquelle le roulement sonne. Elle peut être prédite à partir de la géométrie de palier et la vitesse sous laquelle la bague intérieure et extérieure tourne. Différentes fréquences sont obtenues pour des défauts, les quatre fréquences caractéristiques sont les suivantes (McInerny and Dai, 2003) :

$$\left\{ \begin{array}{l} f_o(Hz) = f_r \left( \frac{N}{2} \right) \left( 1 - \frac{D_b}{D_p} \cos\theta \right) \\ f_i(Hz) = f_r \left( \frac{N}{2} \right) \left( 1 + \frac{D_b}{D_p} \cos\theta \right) \\ f_c(Hz) = f_r \left( \frac{1}{2} \right) \left( 1 - \frac{D_b}{D_p} \cos\theta \right) \\ f_b(Hz) = f_r \left( \frac{D_b}{2D_p} \right) \left( 1 - \frac{D_b^2}{D_p^2} \cos^2\theta \right) \end{array} \right. \quad (1.1)$$

Avec:

- $f_b$  (Ball Frequency [**BF**]):

La fréquence de passage d'un défaut de bille

- $f_i$  (Ball Passing Frequency Inner Race [BPFI]):

La fréquence de passage d'un élément roulant sur un défaut de bague intérieure

- $f_o$  (Ball Passing Frequency Outer Race [BPFO]):

La fréquence de passage d'un élément roulant sur un défaut de bague extérieure

- $f_c$  (Cage Frequency [CF]) :

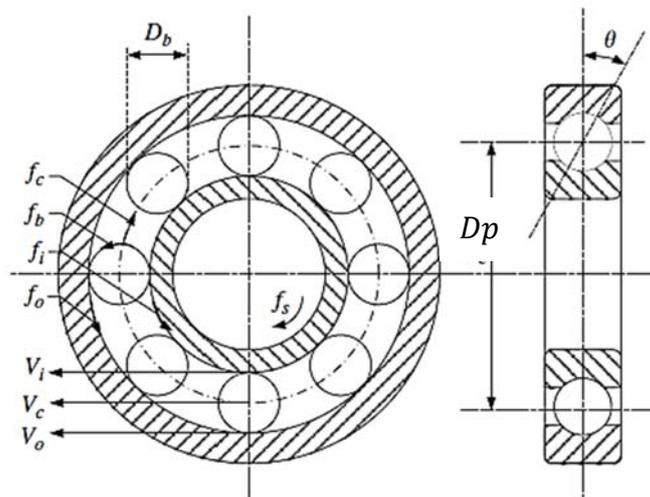
La fréquence de passage d'un défaut de cage

- $f_r$  la fréquence de rotation mécanique

### 1.3.1. Les paramètres caractéristiques d'un roulement :

Sont les suivantes :

- $N$  Le nombre de billes,
- $D_b$  le diamètre de bille (**Figure 1.6**),
- $D_p$  le diamètre moyen du roulement (**Figure 1.6**),
- $\theta$  l'angle de contact (**Figure 1.6**).



**Figure 1.6 :** Fréquences caractéristiques du roulement à billes (Ming, Y,2011)

### **Références Bibliographiques**

Ilyes, KHELFI. (2014) *Diagnostic des machines tournantes par les techniques de l'intelligence artificielle*. Thèse de doctorat. Université de Badji Mokhtar Annaba.

Ishak T. (2012) *Extraction d'indicateurs robustes pour le diagnostic des défauts mécaniques : Comparaison de L'EMD et des ondelettes (WT)*. Thèse de doctorat. Université Ferhat Abbas de Sétif 1.

McInerny S.A. and Dai Y. (2003), Basic vibration signal processing for bearing fault detection, IEEE Transactions On Education, Vol.46, No.1.

Ming, Y., Chen, J., & Dong, G. (2011). Weak fault feature extraction of rolling bearing based on cyclic Wiener filter and envelope spectrum. *Mechanical Systems and Signal Processing*, 25(5), 1773-1785.

Morel, J. (1992). *Vibrations des machines et diagnostic de leur état mécanique*. Collection de la Direction des études et recherches d'électricité de France.

# CHAPITRE

# 2

## ANALYSE VIBRATOIRE

## Chapitre 2. Analyse Vibratoire

### 2.1. Introduction

Les vibrations des machines sont considérées comme la réponse de la fonction de transfert de leurs structures aux différents efforts auxquels elles sont sollicitées, La modification de la vibration d'une machine constitue souvent la première manifestation physique d'une anomalie, cause potentielle de dégradations, voire de pannes par effet de fatigue au sein de la matière ou de fretting entre deux surfaces accolées voir figure (2.1), entraîneront la fissuration puis la rupture et dans d'autres cas ,les vibrations transporteront avec elles une énergie acoustiquement gênante (Tang et al, 2014).

La surveillance des vibrations de ces structures peut permettre donc de détecter l'apparition d'efforts perturbateurs à un stade précoce, la facilité relative de son installation rend l'analyse des signaux vibratoires une des techniques les plus utilisées pour la surveillance et le diagnostic des défauts des machines tournantes, cela dit il existe d'autres techniques de surveillance tel que :

- La thermographie infrarouge
- L'analyse des mesures ultrasonore

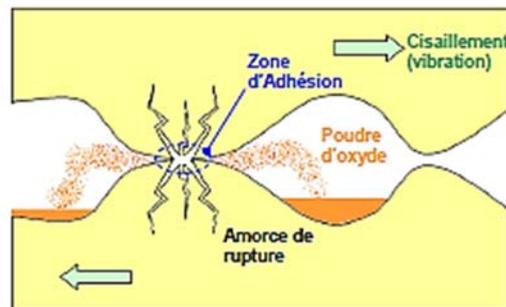


Figure 2.1 : Phénomène de fretting

### 2.2. Définition d'une vibration :

La vibration d'une machine soumise à une force périodique peut être décrite en termes de déplacement, de vitesse ou d'accélération. La vitesse du mouvement vibratoire correspond à la variation de son déplacement pour une unité de temps. L'accélération représente une variation de la vitesse par unité de temps.

Le déplacement exprimé en [ $\mu\text{m}$ ], la vitesse en [ $\text{mm/s}$ ] et l'accélération en [ $\text{m/s}^2$ ].

A ces grandeurs sont associées une période temporelle et une phase. Ainsi, dans le cas d'une machine tournante, une vibration élémentaire peut se mettre sous la forme :

$$x(t) = X \cdot \sin(2\pi \cdot f \cdot t) \quad (2.1)$$

Une propriété fondamentale relie les trois représentations déplacement  $D$ , vitesse  $V$  et accélération  $a$  :

$$D = \frac{V}{2\pi f} = \frac{a}{(2\pi f)^2} \quad (2.2)$$

### 2.2.1. Outils d'analyse des vibrations

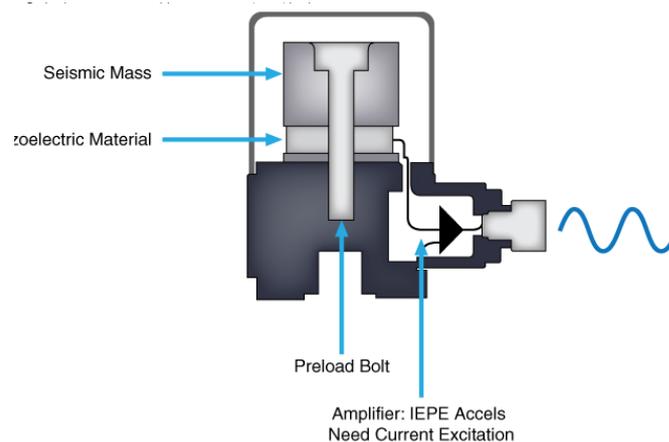
Les outils utilisés pour mesurer les vibrations se sont considérablement améliorés au cours des dernières années. Le capteur de choix pour la plupart des données de vibration sur les machines industrielles est un accéléromètre. Comme son nom l'indique, la sortie est proportionnelle à l'accélération ; Cependant, il est normalement intégré pour afficher en unités de vitesse.

### 2.2.2. Principe de l'accéléromètre :

Un accéléromètre est un capteur qui convertit le mouvement mécanique en un signal électrique.

Sur le plan conceptuel un accéléromètre se comporte comme une masse amortie sur un ressort. Lorsque l'accéléromètre connaît une accélération, la masse est déplacée et le déplacement est ensuite mesuré pour donner l'accélération

Dans ces dispositifs, les techniques piézoélectriques, piézo-résistantes et capacitives sont couramment utilisées pour convertir le mouvement mécanique en un signal électrique. Les accéléromètres piézoélectriques s'appuient sur des piézo-céramiques (par exemple le titanate de zirconate de plomb) ou des monocristaux (par exemple le quartz, la tourmaline)



**Figure 2.2:** Tension de sortie de l'accéléromètre proportionnel à la force de vibration du Crystal piézoélectrique

### **2.3. Les Méthodes d'analyse Vibratoire :**

Les techniques du domaine temps-fréquence sont principalement utilisées pour traiter des signaux vibratoires non stationnaires. On y distingue la transformée de Fourier courte, la distribution de Wigner-Ville, l'analyse en ondelettes notamment la décomposition en paquets d'ondelettes (Wavelet Packet Decomposition : WPD), la décomposition modale empirique et la transformée de Hilbert-Huang. La transformée de Fourier courte segmente le signal en fenêtres temporelles et applique ensuite la transformée de Fourier sur les fenêtres obtenues, cependant, la taille des fenêtres est difficile à choisir.

Concernant la distribution de Wigner-Ville, elle conduit parfois à l'apparition de termes croisés, ce qui peut conduire à des analyses erronées.

La décomposition en paquets d'ondelettes (WPD) consiste à utiliser une paire de filtres passe-haut et passe-bas. Ces filtres permettent de réaliser une division du contenu fréquentiel du signal brut en deux composantes (ou paquets), l'une contient les hautes fréquences et l'autre les basses fréquences. Puis, ces paquets sont successivement injectés dans les filtres pour séparer de nouveau en deux l'espace fréquentiel. Cette opération permet d'atteindre des niveaux de décomposition assez fins qui permettent de suivre l'évolution des fréquences, tout en gardant une bonne résolution pour les phénomènes rapides et lents grâce à l'utilisation d'une ondelette mère et des paramètres d'échelle et de translation. La WPD peut être utilisée pour la détection et le suivi de dégradations, mais son efficacité dépend fortement de la qualité des signaux traités.

L'une des Méthodes la plus efficace pour l'analyse des signaux non stationnaires est la transformée de Hilbert-Huang. Elle s'appuie sur deux techniques : l'EMD et la transformée de Hilbert. L'EMD décompose le signal en fonctions modales intrinsèques (Intrinsic Mode Functions : IMFs) représentant les modes oscillatoires du signal. Ensuite, la transformée de Hilbert est appliquée sur chaque IMF pour extraire les fréquences instantanées et les amplitudes instantanées. La transformée de Hilbert-Huang permet ainsi une représentation tridimensionnelle (amplitude, fréquence, temps) de chaque mode oscillatoire du signal.

## 2.4. Extraction des caractéristiques :

Les paramètres statistiques quantifient les informations sur la forme et l'énergie du signal, elles peuvent refléter l'état de santé mécanique sous différents aspects et peuvent se compenser mutuellement à différents niveaux de défaut.

L'extraction des caractéristiques repose essentiellement sur des techniques issues du traitement du signal. Ces techniques peuvent être regroupées dans deux catégories principales selon que le signal à traiter est stationnaire ou non stationnaire. La figure ci-dessous donne un aperçu des techniques utilisées pour l'extraction des caractéristiques.

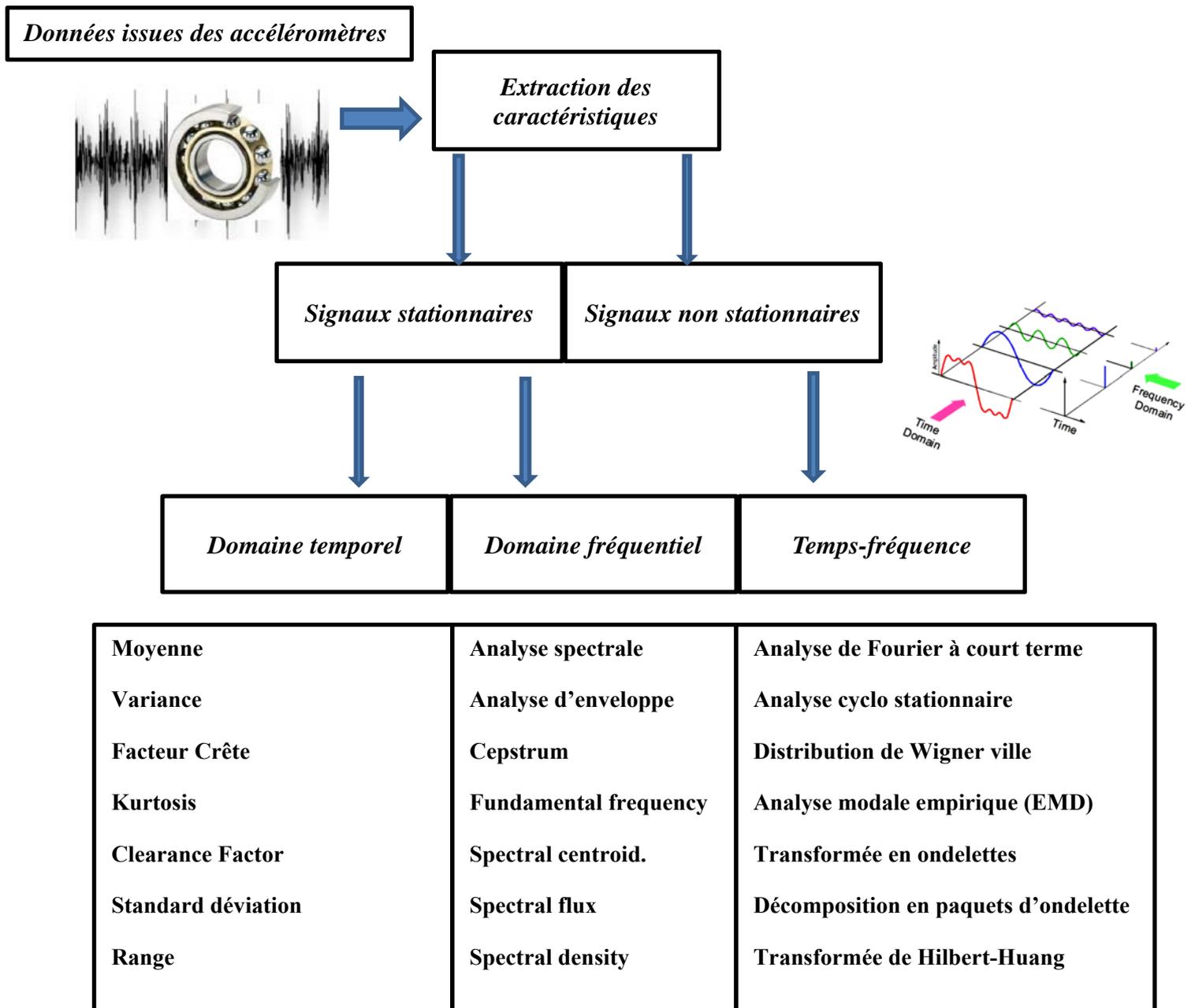


Figure 2.3 : Techniques d'extraction des caractéristiques (W. Yan,2008)

## 2.5. Sélection et réduction des caractéristiques :

Après l'étape d'extraction des paramètres, descripteurs ou indicateurs, l'étape de sélection est cruciale pour trouver quels sont les paramètres les plus importants et les plus fiables pour la détection du défaut. En effet, si les paramètres sont sélectionnés avec peu de pouvoir de discrimination, la conception ultérieure d'un classificateur conduirait à de mauvaises performances.

D'autre part, si les paramètres avec des informations riches sont sélectionnés, la précision de la classification augmente et le temps de calcul est réduit.

Afin d'avoir un système de diagnostic performant, il est nécessaire d'employer un processus de sélection des indicateurs les plus pertinents, permettant d'améliorer les performances de la classification

Les techniques de sélection d'indicateurs sont divisibles en trois catégories, selon la manière dont elles interagissent avec le classificateur (Tang et al, 2014) :

- Les approches avec filtres
- Les approches avec enveloppes (wrapper)
- Les approches intégrées (embedded)

Dans le cas de la classification par le séparateur à vaste marge on peut citer les travaux suivants :

- L'élimination récursive des indicateurs (Recursive Features Elimination (RFESVM)), présentée par Guyon (Guyon & Elisseeff, 2003).
- Dans (Lin et al, 2008), les auteurs proposent une approche SA-SVM basé sur le recuit simulé pour l'optimisation des paramètres des SVMs.
- Dans (Samanta et al. 2003), les auteurs proposent un algorithme génétique AG pour la sélection d'un sous ensemble optimal d'indicateurs, dans le cadre de diagnostic automatisé des machines tournantes. Le sous ensemble sélectionné est utilisé pour entraîner deux classificateurs différents à savoir, les SVMs et les réseaux de neurones artificiels RNA

### 2.5.1. Caractéristiques ou paramètres temporelles

Elles concernent le calcul des paramètres statistiques du signal. Elles sont utilisées en classification, détection et diagnostic de défaillances où elles donnent des résultats acceptables.

Parmi les caractéristiques temporelles, on peut citer la moyenne, le carré de la moyenne quadratique (Root Mean Square : RMS) qui correspond à l'énergie moyenne du signal, la valeur de son pic qui peut renseigner sur la taille d'un défaut, le Kurtosis qui caractérise l'aspect impulsif du signal, le facteur de crête qui mesure les augmentations rapides des vibrations dans

les roulements et le Skewness utilisé pour évaluer l'état de surface d'un composant (Tang et al, 2014).

$$\text{Moyenne : } x_m = \sum_{n=1}^N \frac{x(n)}{N} \quad (2.3)$$

$$\text{Variance : } x_{\text{var}} = \frac{\sum_{n=1}^N (x(n) - x_m)^2}{N} \quad (2.4)$$

$$\text{Kurtosis } x_{\text{kur}} = \frac{\sum_{n=1}^N (x(n) - x_m)^4}{(N - 1)\sigma^4} \quad (2.5)$$

$$\text{Clearance factor : CLF} = \frac{x_{\text{max}}}{\left(\frac{1}{N} \sum_{n=1}^N \sqrt{|x(n)|}\right)^2} \quad (2.6)$$

$$\text{Mean absolute deviation : MAD} = \sum_{n=1}^N \frac{|\sum_{n=1}^N (x(n) - x_m)|}{N} \quad (2.7)$$

$$\text{Range : } x_{\text{range}} = |x_{\text{max}} - x_{\text{min}}| \quad (2.8)$$

$$\text{Écart type (STD): } \sigma = \sqrt{\frac{\sum_{n=1}^N (x(n) - x_m)^2}{N}} \quad (2.9)$$

$$\text{Peak magnitude to RMS ratio: Crest factor} = \frac{x_{\text{max}}}{\text{RMS}} \quad (2.10)$$

Où  $x(n)$  / ( $n= 1, \dots, N$ ) est l'amplitude au point d'échantillonnage  $n$  et  $N$  est le nombre de points d'échantillonnage.  $x_m$  est le moyen de  $x$ ,  $\sigma$  est l'écart type de  $x$ .

### 2.5.2. Caractéristiques ou paramètres fréquentielles :

L'analyse fréquentielle consiste à visualiser l'énergie du signal à différentes fréquences. En connaissant les conditions critiques de fonctionnement du composant (vitesse, charge, géométrie des éléments constitutifs, ...etc.), on peut extraire des caractéristiques permettant d'isoler les défauts. Ces caractéristiques sont aussi appelées signatures fréquentielles.

La fréquence moyenne :

$$\text{Mean Frequency : } f_{\text{mean}} = \frac{\sum_{n=1}^N I_n \cdot f_n}{\sum_{n=1}^N I_n} \quad (2.11)$$

Où  $f_n$  et  $I_n$  sont les fréquences et les amplitudes au point  $n$  du spectre du signal

$$\text{Mean power frequency: MNF} = C_s = \frac{\sum_{n=1}^N P_n \cdot f_n}{\sum_{n=1}^N P_n} \quad (2.12)$$

Où  $P_n$  est la puissance spectrale du signal au point  $n$  ( Oskoei & Hu, 2008).

$$\text{Densité spectrale de puissance DSP} = \frac{\sum_{n=1}^N P_n}{N} \quad (2.13)$$

$$\text{Spectral Entropy En}_s = -\sum_{n=1}^N P_n \log P_n \quad (2.14)$$

$$\text{Short time Energy} = \sum_{n=1}^N [x[m] W[n - m]]^2 \quad (2.15)$$

## 2.6. Analyse vibratoire par la transformée de Hilbert Huang

### 2.6.1. La transformée de Hilbert

#### 2.6.1.1. Introduction

Tout signal sinusoïdal réel  $A \cos(\omega t + \varphi)$  peut être transformée en un signal complexe  $Ae^{j(\omega t + \varphi)}$  en générant une composante en quadrature de phase :  $A \sin(\omega t + \varphi)$ , servant de partie imaginaire :

$$Ae^{j(\omega t + \varphi)} = A \cos(\omega t + \varphi) + j A \sin(\omega t + \varphi) \quad (2.16).$$

Pour les signaux plus compliqués qui sont exprimables par une somme de plusieurs sinusoïdes, nous pouvons utiliser un filtre qui déphase chaque composante d'un quart de période, cette méthode est appelée la transformée de Hilbert.

Soit  $H\{x(t)\}$  la sortie de l'instant  $t$  du filtre quadrature de Hilbert appliqué au signal  $x(t)$ . Idéalement ce filtre a une magnitude de 1 pour toutes les fréquences et introduit un décalage de  $-\frac{\pi}{2}$  pour chaque fréquence positive et  $+\frac{\pi}{2}$  pour chaque fréquence négative.

La transformée de Hilbert est définie par :

$$H\{x(t)\} = y(t) = \frac{1}{\pi} VP \int_{-\infty}^{+\infty} \frac{x(t')}{t-t'} dt' \quad (2.17)$$

VP est la valeur principale de Cauchy définie par :

$$VP \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt' = \lim_{\epsilon \rightarrow 0} [\int_{-\infty}^{t-\epsilon} \frac{x(t')}{t-t'} dt' + \int_{t+\epsilon}^{\infty} \frac{x(t')}{t-t'} dt'] \quad (2.18)$$

$Y(t)$  : est le conjugué de  $x(t)$ .

$x(t)$  peut être représentée comme étant l'inverse de la transformée de Hilbert de  $y(t)$ , tel que :

$$x(t) = \{H^{-1}(y(t))\} = -\frac{1}{\pi} VP \int_{-\infty}^{\infty} \frac{y(t')}{t-t'} dt' \quad (2.19)$$

D'après la relation (2.20) nous pouvons observer que la fonction  $y(t)$  est déterminée par la convolution de  $x(t)$  avec la réponse impulsionnelle du filtre de Hilbert égal à  $\frac{1}{\pi t}$  :

$$y(t) = x(t) * \frac{1}{\pi t} \quad (2.20)$$

Cette relation permet de calculer la densité spectrale de  $y(t)$  qui est :

$$Y(j\omega) = TF\{y(t)\} = TF\{x(t)\} \cdot TF\left\{\frac{1}{\pi t}\right\} \quad (2.21)$$

$$\text{Nous avons : } Y(j\omega) = X(j\omega) \cdot TF\left\{\frac{1}{\pi t}\right\} \quad (2.22)$$

Le gain complexe du filtre de Hilbert est :

$$TF\left\{\frac{1}{\pi t}\right\} = -j\text{sgn}(\omega) \quad (2.23)$$

$$\text{Ainsi : } Y(j\omega) = X(j\omega) \cdot [-j\text{sgn}(\omega)] \quad (2.24)$$

$$Y(j\omega) = \begin{cases} -jX(j\omega), & \omega > 0 \\ jX(j\omega), & \omega < 0 \end{cases} \quad (2.25)$$

Comme résultat ; nous pouvons dire que la fonction de la densité spectrale du signal conjugué de  $x(t)$  est obtenu par le changement de phase de la densité spectrale de  $X(j\omega)$  par  $\mp \frac{\pi}{2}$ .

$$\text{Le résultat est : } H\{x(t)\} = y(t) = TF^{-1}\{Y(j\omega)\} \quad (2.26)$$

La transformée inverse de Hilbert est définie dans la relation (2.17), nous donne :

$$x(t) = H^{-1}\{y(t)\} = -H\{y(t)\} \quad (2.27)$$

En prenant en compte la relation 2.25, il en résulte que :

$$x(t) = \begin{cases} TF^{-1}\{jY(j\omega)\}, & \omega > 0 \\ TF^{-1}\{-jY(j\omega)\}, & \omega < 0 \end{cases} \quad (2.28)$$

Afin d'introduire la méthode de Hilbert permettant l'estimation de la FI et de la AI, il est nécessaire d'associer à un signal réel  $x(t)$  un signal analytique.

### 2.6.1.2 Signal analytique

Soit un signal  $x(t)$  tel que :

$$x(t) = a(t) \cos \theta(t) \quad (2.29)$$

Le signal analytique associé à  $x(t)$  est :

$$Z_x(t) = S[x(t)] \quad (2.30)$$

$S[.]$  : opération pour avoir le signal analytique.

Il est alors nécessaire d'imposer au signal  $Z_x(t)$  les deux conditions ci-dessous.

- Premièrement, comme il s'agit d'une extension au cas général d'une approche qui doit s'appliquer au cas sinusoïdal, on impose la condition :

$$S[\cos(\omega t)] = \exp(j\omega t), \forall \omega > 0 \quad (2.31)$$

- Deuxièmement, l'opération doit être linéaire : si l'AI de  $x(t)$  est  $a(t)$  alors celle  $\alpha \cdot x(t)$  doit être  $\alpha \cdot a(t)$  Par extension, il suffit d'imposer que l'application  $S[.]$  soit un filtrage linéaire.

Un signal analytique  $Z_x(t)$  est défini par :

$$Z_x(t) = x(t) + iy(t) = x(t) + jH\{x(t)\} = a(t)e^{j\varphi(t)} \quad (2.32)$$

$$\text{Et : } Z_x(j\omega) = \text{TF}\{Z_x(t)\} = X(j\omega) + jY(j\omega) \quad (2.33)$$

On note que la partie réelle de  $Z_x(t)$  lui-même et que la partie imaginaire est obtenue à partir d'un filtre linéaire de Hilbert noté  $H$  en quadrature avec  $x(t)$ . La fonction  $Z_x(t)$  ainsi déterminée à la propriété de filtrer toutes les fréquences négatives de  $x(t)$ .

Les paramètres  $a(t)$  et  $\theta(t)$  du signal  $x(t)$  peuvent alors être déterminés de manière claire comme le module et l'argument  $Z_x(t)$  pour qui la transformée de Fourier est nulle pour les fréquences négatives. Il est à mettre en opposition avec un signal sinusoïdale réel (ex :  $\cos \omega t$ ) dont le spectre est constitué de  $\omega$  et  $-\omega$ .

En se référant à la relation (2.23) on a :

$$\begin{aligned} Z_x(j\omega) &= X(j\omega) + j[-j\text{sgn}\omega]X(j\omega) = \\ X(j\omega)[1 + \text{sgn}\omega] &= 2X(j\omega)U(\omega) \end{aligned} \quad (2.34)$$

Où  $U(\omega)$  désigne l'échelon unité de Heaviside.  $Z_x(t)$  est complexe avec un encombrement spectrale deux fois plus petit que celui du signal réel  $x(t)$ . La relation 2.24 revient à supprimer les fréquences négatives du spectre original ce qui ne diminue en rien l'information, puisque pour un signal  $X(t) \in \mathbb{R}$ , nous avons la relation suivante :

$$\begin{cases} X(j\omega) = \frac{1}{2}[Z(j\omega) + Z^*(-j\omega)] \\ Y(j\omega) = \frac{1}{2}[Z(j\omega) - Z^*(-j\omega)] \end{cases} \quad (2.35)$$

$$Z^*(j\omega) = X(j\omega) - jY(j\omega)$$

Ainsi par construction, les notions d'amplitude et de phase d'un signal réel  $x(t)$  non-stationnaire sont associées à un unique signal complexe  $Z_x(t)$  dont le module et l'argument définissent respectivement les grandeurs :

**Amplitude instantanées AI**

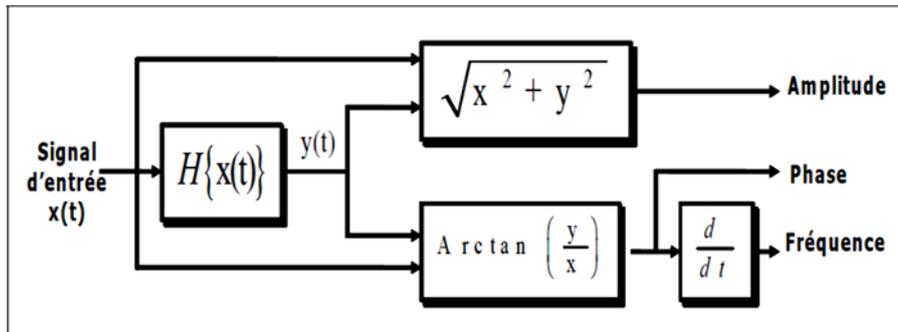
$$AI = |Z_x(t)| = \sqrt{x^2(t) + y^2(t)} \quad (2.36)$$

**Et phase instantanée**

$$\theta(t) = \arg z_x(t) = \arctan\left(\frac{y(t)}{x(t)}\right) \quad (2.37)$$

La fréquence instantanée FI de  $z(t)$ , et donc de  $x(t)$  ; n'est autre que la dérivée de la phase instantanée :

$$F(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \quad (2.38)$$



**Figure 2.4** : Principe d'estimation de la FI et de AI via le signal analytique  $Z_x(t)$  et  $x(t)$

La transformée de Hilbert présente une limitation forte dans la mesure où elle n'est applicable qu'aux signaux qui peuvent se représenter par une fréquence unique à chaque instant. En effet, si le signal contient deux composantes de cette forme, la définition du couple amplitude / fréquence instantanée à partir du signal analytique ne fournit qu'un seul couple amplitude / fréquence instantanées pour les deux composantes et l'interprétation de ce couple est alors loin des fréquences et amplitude instantanées des deux composantes. Pour éviter ce problème, une nouvelle méthode appelée DME pour « Décomposition Modale Empirique » a été développée afin de séparer les composantes d'un signal.

**2.6.2. Décomposition modale empirique (DME)**

La décomposition modale empirique est une méthode algorithmique de décomposition spectrale adaptative (entièrement pilotée par les données), qui est définie par un processus appelé tamisage (*sifting*) permettant de décomposer un signal en fonctions de base : au lieu d'analyser le signal dans une base fixe comme avec Fourier. Ces fonctions de base, appelées (Intrinsic Mode Functions), sont des signaux de type *AM - FM* de moyenne nulle.

La décomposition en IMFs est locale, c'est-à-dire qu'on considère les signaux à l'échelle de leurs oscillations locales (la méthode opère à l'échelle d'une oscillation). Elle est itérative, séquentielle et entièrement pilotée par les données. En plus de sa capacité à décrire ponctuellement et de manière instantanée les fréquences d'un signal, problème non résolu par

l'analyse de Fourier, la *DME* est bien adaptée à l'étude des signaux non-stationnaires et/ou génères par des systèmes non-linéaires.

Cependant la *DME* est simplement définie par un algorithme et elle ne possède pas de cadre théorique pour le moment. Dans la plupart des cas, les études menées sur la *DME* se font avec des simulations numériques extensives et dans des conditions contrôlées. Malgré le manque actuel de formalisme mathématique pour décrire la *DME*, il a été démontré que ce manque ne remet pas en cause les réelles capacités d'analyse des signaux par la *DME* ou par ses dérivées telle que la THH. En effet, cette technique a été appliquée avec succès à des données réelles dans des domaines tels que l'océanographie et l'étude des phénomènes climatiques, la sismologie, le contrôle non destructif, l'acoustique sous- marine, ou la biologie.

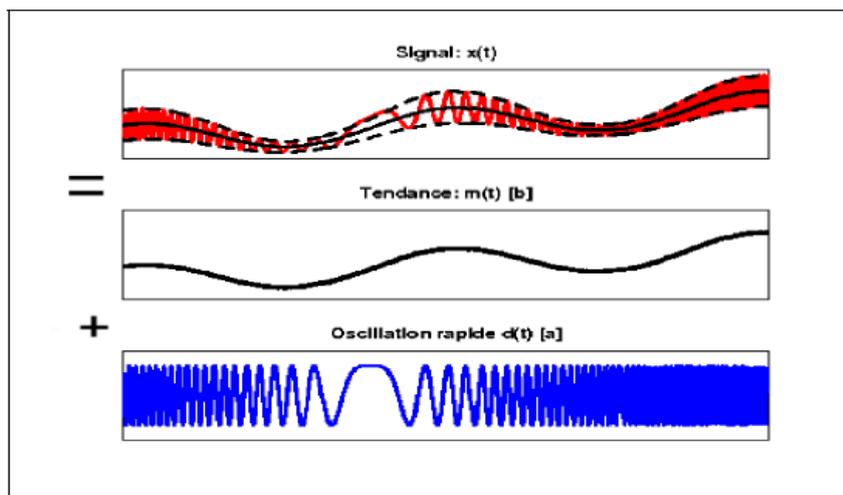
### 2.6.2.1. Principe de la décomposition

On veut décomposer le signal en plusieurs fonctions, pour lesquelles une définition de la fréquence instantanée sera possible. On introduit ainsi la notion de *fonction de mode intrinsèque* ou *intrinsic mode function* (IMF), classe de fonctions vérifiant deux conditions de symétrie, nécessairement vérifiées par les sinusoïdes, mais plus faibles :

- Le nombre de zéros et le nombre d'extrema différent d'au plus un.
- La moyenne *locale* est nulle.

Pour illustrer la décomposition, nous considérons l'exemple simple d'un signal  $x(t)$  formé de deux contributions : une oscillation rapide et une autre lente :

$$\underbrace{x(t)}_{\text{signal}} = \underbrace{d(t)}_{\text{oscillation}(a)} + \underbrace{m(t)}_{\text{tendance}(b)} \quad (2.39)$$



**Figure 2.5** : Principe de l'EMD :  
« Oscillation rapide » (a) + « oscillation lente » (tendance) (b)

Ainsi, la reconstruction du signal  $x(t)$  est réalisée en sommant les deux courbes point par point (on parle de décomposition géométrique) (Fig. 2.5). Notons que, par construction la tendance  $b(t)$  du signal contient localement des oscillations de plus basse fréquence que celle de l'oscillation rapide. Ce résultat peut être généralisé à toutes les contributions comme suit : chacune contient localement des oscillations de plus basse fréquence que celle extraite précédemment.

Conceptuellement, la mise en place de la méthode permettant de décrire cette décomposition est assez simple :

Il suffit de considérer le signal à l'échelle de ses oscillations locales, de soustraire l'oscillation la plus rapide du signal et de répéter le processus sur le résidu de manière à écrire le signal  $x(t)$  comme une combinaison finie d'oscillations.

Le signal  $x(t)$  peut alors s'écrire :

$$x(t) = \sum_{j=1}^N IMF_j(t) + r(t), \quad N \in \mathbb{N}^* \quad (2.40)$$

Où  $IMF_j(t)$  est la  $j^{\text{e}}$  oscillation,  $r(t)$  le résidu de la décomposition et  $N$  le nombre d'IMF (ou modes).

#### 2.6.2.2. Définition d'une IMF (Intrinsic Mode Function) :

Une IMF est une fonction telle que :

$$IMF: \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ t \rightarrow IMF_j(t) \end{cases} \quad (2.41)$$

qui doit être :

- a. De moyenne nulle,
- b. Dont les nombres d'extrema et de passages à zéro diffèrent au plus d'un (en d'autres termes, cela signifie qu'entre un minimum et un maximum successif, une IMF passe par zéro),
- c. Qui suit une loi de modulation en amplitude et en fréquence (comportement oscillant) naturellement de type mono-composante.

#### 2.6.2 .1.1. Processus de tamisage : principe algorithmique

Le *sifting process*, que l'on pourrait traduire par *processus de tamisage*, correspond à l'extraction d'une IMF, à partir d'un signal donné. La fonction de mode intrinsèque, ou IMF, ainsi obtenue peut-être vue comme les détails du signal, de variation rapide. Pour cela, on extrait de façon itérative ces détails, jusqu'à conserver une IMF :

1. Calculer par interpolation  $E_{min}$  (resp.  $E_{max}$ ), enveloppe des minima (resp.  $E_{min}$ ) de  $x(t)$ .
2. Calculer la moyenne locale  $m(t) = \frac{E_{min}(t) + E_{max}(t)}{2}$  (2.42)
2. Si  $m \equiv 0$ ,  $x$  est déjà une IMF.

3. Sinon, itérer le processus sur les détails  $c(t) = x(t) - m(t)$ .

On peut observer toutes les étapes du *sifting process* dans la figure 2.5 : en rouge les deux enveloppes d'interpolation, en noir la moyenne locale, et enfin la dernière courbe qui est la différence de la courbe bleue précédente avec sa moyenne locale. En noir la nouvelle moyenne locale.

Il est important de mentionner que le tamisage est répété plusieurs fois ( $i$ ) jusqu'à ce que  $h$  vérifie les conditions nécessaires d'une IMF.

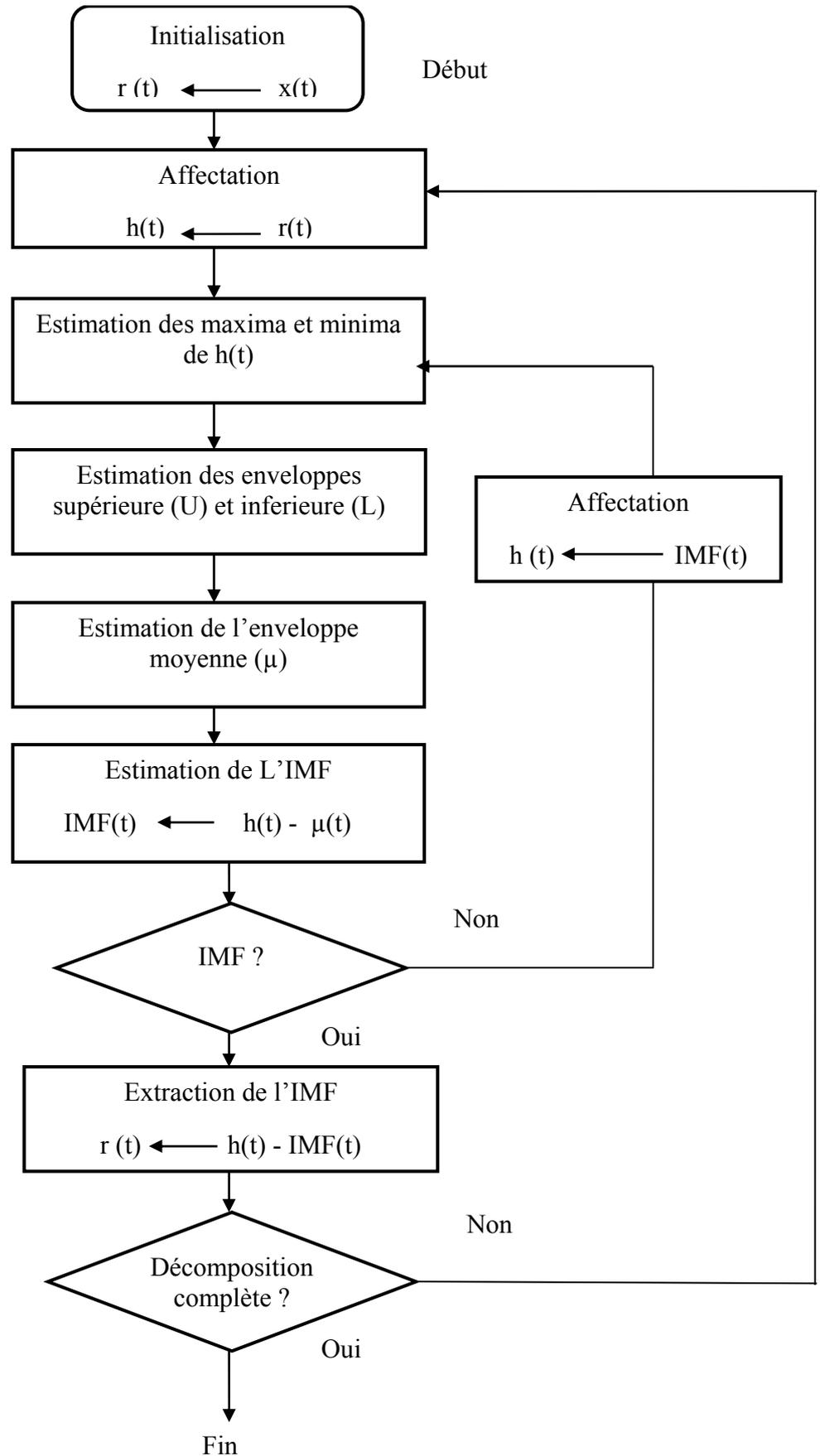


Figure 2.6 : Organigramme du Pseudo-code de l'EMD.

Pour illustrer cette méthode considérons un exemple :

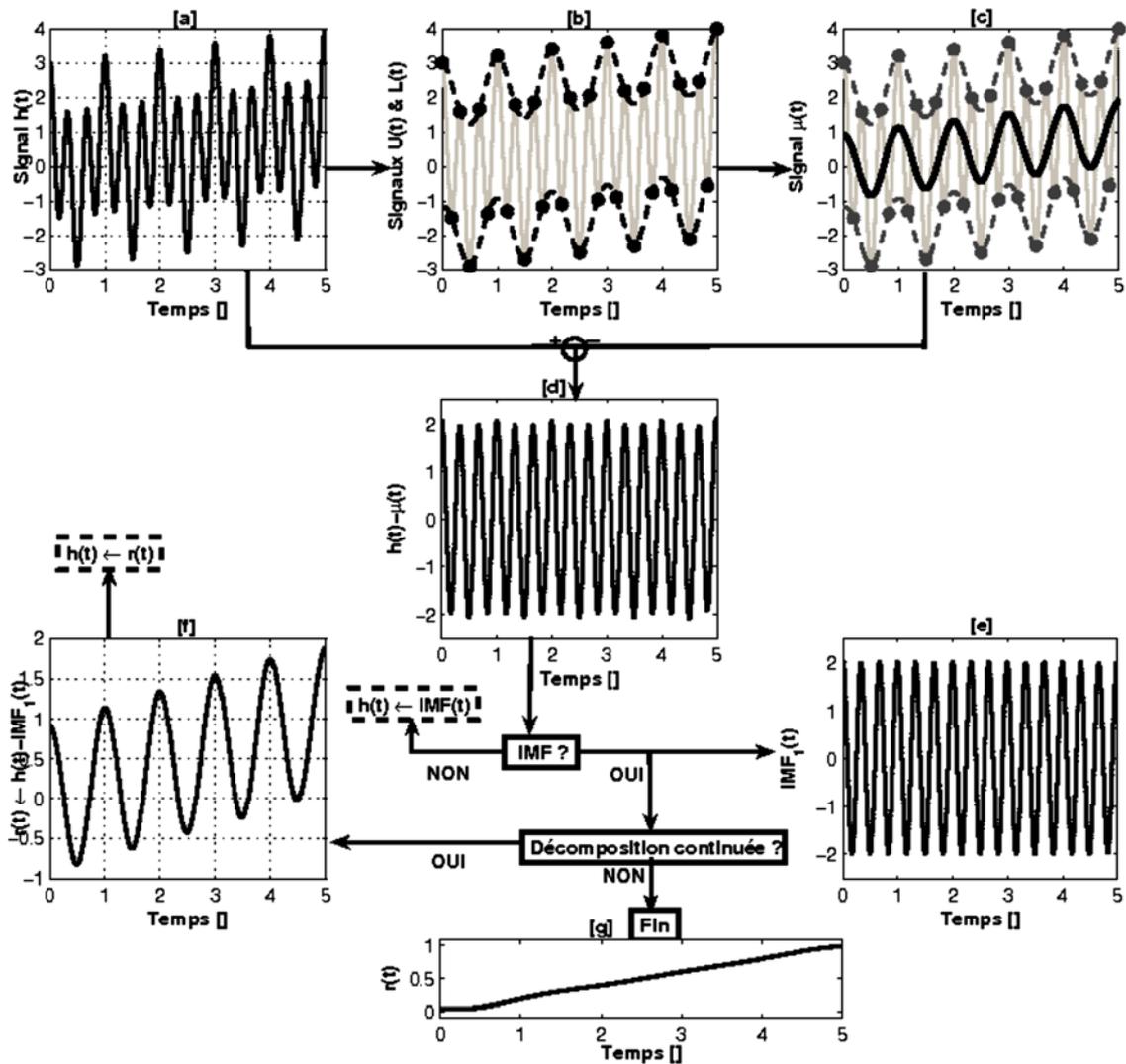


Figure 2.7: Principe de tamisage pour l'estimation de la 1<sup>ère</sup> IMF du signal test

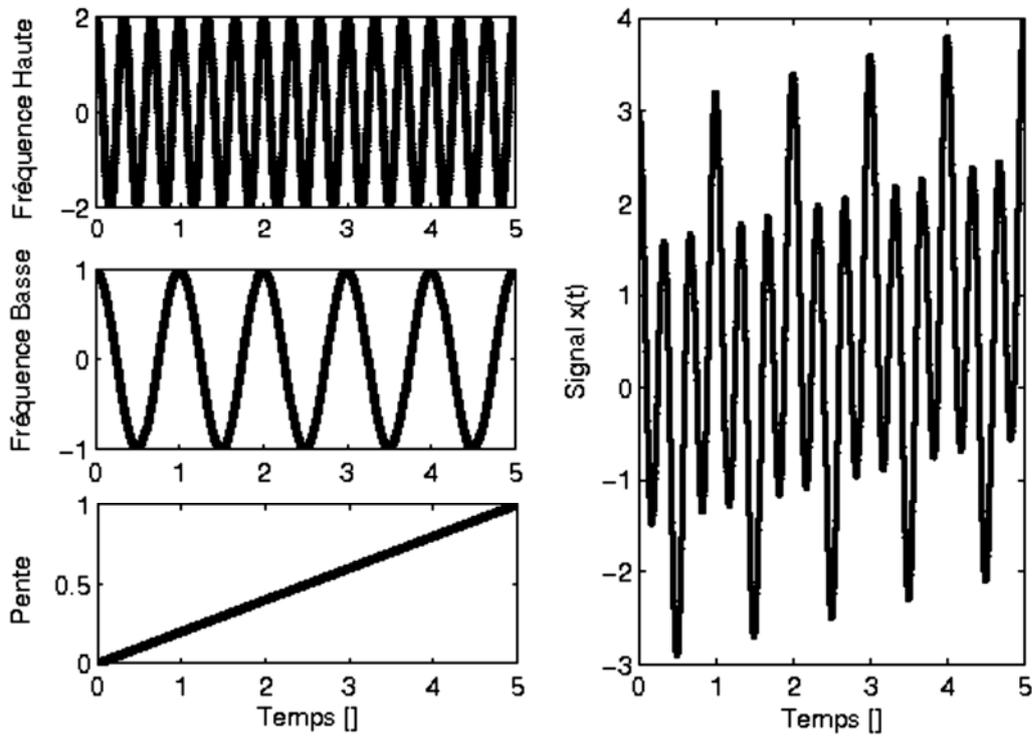


Figure 2.8: Signal temporel test

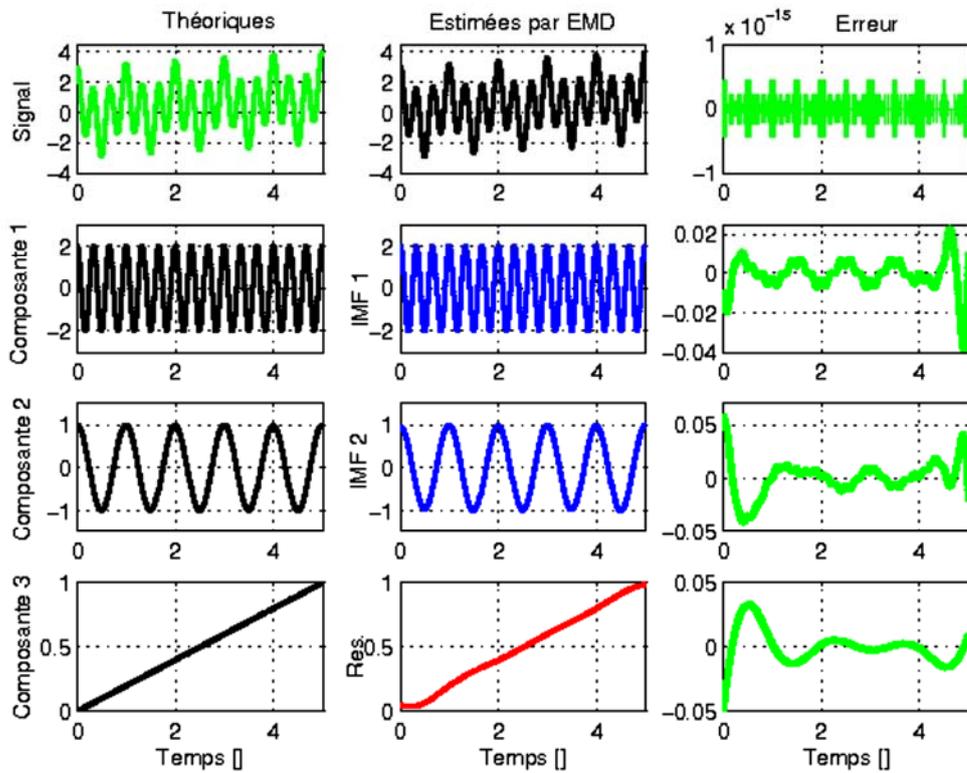


Figure 2.9: Estimation des IMF du signal test.

### 2.6.2.3. Critère d'arrêt du processus de tamisage

Comme indiqué précédemment, l'EMD est composée de deux boucles : l'une permet d'extraire l'IMF (qui détermine le niveau de profondeur de la décomposition) et l'autre conditionne la fonction IMF  $j(t)$  de manière à valider la définition d'une IMF.

La boucle de tamisage qui permet d'extraire des IMF correctement conditionnées est pilotée par un critère note  $SD(i)$  de manière générale, les définitions du critère d'arrêt du tamisage reposent toutes sur le même principe :  $h_{j,i}(t)$  est de moyenne nulle et symétrique.

Ainsi dans ; les auteurs proposent un critère d'arrêt  $SD(i)$  reposant sur la déviation standard et est défini par :

$$SD(i) = \sum_{t=0}^T \frac{|h_{i-1}(t) - h_i(t)|^2}{[h_{i-1}(t)]^2} \quad (2.43)$$

### 2.6.2.4. Les différents Types d'EMD :

La conception d'un système de diagnostic performant n'est pas une tâche facile à réaliser, puisque la majorité des systèmes réels sont de nature non-linéaire et non-stationnaire souvent corrélées et bruitées.

Les performances du système de diagnostic se basent sur la robustesse de la méthode face à la présence de bruit, perturbation et erreur de modélisation. La manipulation du prétraitement des données peut améliorer les performances du diagnostic. de nouvelles méthodes offrent la possibilité de réduire le bruit et d'extraire seulement l'information pertinente.

#### **L'EEMD :**

Est une méthode d'analyse de données bruitées, Noise Assisted Data Analysis (NADA).

L'EEMD consiste à "tamiser" un bruit blanc ajouté au signal.

Le bruit blanc est calculé en moyenne avec un nombre suffisant d'essais, la seule partie persistante qui survit au processus de calcul de la moyenne est le signal, qui est ensuite traité comme la vraie et plus significative réponse physique. En outre, L'EEMD représente une amélioration substantielle par rapport à l'EMD d'origine.

#### **CEEMDAN :**

Complete Ensemble EMD with Adaptive Noise est une variation de l'algorithme EEMD qui fournit une reconstruction exacte du signal d'origine et une meilleure séparation spectrale des IMFs avec moins de bruits et plus d'information physiques.

2.6.2.5. Bruit et Valeurs aberrantes :

En pratique un signal que l'on cherche à mesurer est toujours accompagné de signaux perturbateurs, que l'on peut appeler des bruits :

- Le bruit électronique issu des composants électroniques peut être considéré comme blanc et peu énergétique. On appelle bruit blanc un signal aléatoire dont l'énergie spectrale est la même pour toutes les fréquences.
- Le bruit de quantification résultant de la conversion des données analogiques en données numériques.

Le rapport signal sur bruit est défini par :

$$\text{SNR}=10 \log\left(\frac{\sum_{i=1}^n y^2(i)}{\sum_{i=1}^n e^2(i)}\right) \quad (2.44)$$

$$\text{FNR}=10 \log_{10} \frac{\hat{\sigma}_f^2}{\hat{\sigma}_v^2} \quad (2.45)$$

Les signaux peuvent également contenir des données erronées ou fausses. La manipulation et la détection des valeurs aberrantes est nécessaire à toute modélisation car leurs présences détériorent l'élaboration d'un bon modèle.

- Parmi les techniques de détection des valeurs aberrantes dans les bases de données on cite la distance de Mahalanobis et l'analyse de résidu de la régression PLS avec les réseaux de neurones dans les travaux publiés par Hugo and Sbarbaro en 2011.

2.6.2.6. Les nouvelles approches pour la sélection des IMFs :

L'EMD est utilisé comme outil de prétraitement qui est particulièrement adaptée pour traiter des signaux non stationnaires.

Plusieurs IMFS sont obtenues dont on choisit les plus significatives pour extraire l'enveloppe qui contient l'information requise.

- Sélection des IMFs par la densité spectrale de Hilbert, les IMFs retenues sont celles qui maximisent la densité spectrale de Hilbert autour des fréquences caractéristiques des défauts. Pour chaque fréquence caractéristique, on regarde parmi toutes les IMFs celle qui maximise la densité spectrale de Hilbert sur cette fréquence et on la garde. On obtient ainsi une seule IMF pour une fréquence donnée. Cela permet d'avoir au minimum une IMF (la même pour toutes les fréquences caractéristiques) et au maximum trois IMFs différentes (une pour chaque fréquence). La valeur de chaque IMF retenue est prise comme valeur de l'indicateur de santé, ce qui conduit au total à trois indicateurs (un par fréquence caractéristique) (Medjaher, K. 2014).
- L'EMD est une procédure numérique sujette aux erreurs pouvant persister dans la décomposition en IMFs en tant que supplément additionné.

Dans les travaux publiés par De Souza en 2014, un nouvel algorithme est proposé pour sélectionner les IMFs pertinents, basé sur l'approche énergie- bruit par comparaison avec la technique traditionnelle qui utilise la corrélation.

- Dans un d'autres travaux publiés par Md. Burhan Uddin en 2016 un algorithme de sélection adaptatif des IMFs qui exploite le Caractéristiques du seuillage basé sur l'énergie et du potentiel empirique de SVM pour la classification des IMFs sans bruit : les IMFs dominants avec bruit sélectionnés de manière adaptative sont désactivés à travers le filtre SGF et Les IMFs sans bruit sont ensuite reconstruits. Une conclusion ultérieure est tirée ; Les résultats obtenus par la comparaison du SNR (Signal-to-Noise Ratio), PRD et Max Error

## Références Bibliographiques

- A. Boashash, (2003): Time-Frequency Signal Analysis and Processing: A Comprehensive Reference. Elsevier Ltd., Oxford, UK.
- B. Boashash, (1992): Time-Frequency Signal Analysis-Methods and Applications. Longman-Cheshire, Melbourne and John Wiley Halsted Press, New York.
- De Souza, D. B., Chanussot, J., & Favre, A. C. (2014, May). On selecting relevant intrinsic mode functions in empirical mode decomposition: An energy-based approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on* (pp. 325-329). IEEE.
- Garces, H., & Sbarbaro, D. (2011). Outliers detection in environmental monitoring databases. *Engineering Applications of Artificial Intelligence*, 24(2), 341-349.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.
- Jean-Christophe Cexus, (2005) Analyse des signaux non-stationnaires par transformation de Huang, Operateur de Teager-Kaiser, et Transformation de Huang-Teager (THT). Acoustics. Université Rennes 1. French. <tel-00012141>.
- Lin, S. W., Lee, Z. J., Chen, S. C., & Tseng, T. Y. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied soft computing*, 8(4), 1505-1512.
- M. G. Rilling, (2007). Décompositions Modales Empiriques Contribution à la théorie, l'algorithme et l'analyse de performances. Thèse de doctorat. Université de Lyon.
- management (PHM)-a survey. Technical report, Air Force Research Laboratory.
- Medjaher, K. (2014). *Contribution au pronostic de défaillances guidé par des données* (Doctoral dissertation, Université de Franche-Comté).
- Morel, J. (1992). Vibrations des machines et diagnostic de leur état mécanique. *Collection de la Direction des Etudes et recherches d'Electricité de France*.
- N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung et H.H. Liu, (1998): The empirical mode decomposition and the Hilbert spectrum for Nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, 454 :903–995.

Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE transactions on biomedical engineering*, 55(8), 1956-1965.

Samanta, B., Al-Balushi, K. R., & Al-Araimi, S. A. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence*, 16(7), 657-665.

Tang, D., Wei, F., Qin, B., Liu, T., & Zhou, M. (2014, August). Coooolll: A deep learning system for twitter sentiment classification. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)* (pp. 208-212).

Uddin, M. B., Uddin, J., Sultana, R., & Islam, S. (2016, December). A new machine learning approach to select adaptive IMFs of EMD. In *Electrical, Computer & Telecommunication Engineering (ICECTE), International Conference on* (pp. 1-4). IEEE.

W. Yan, H. Qiu, and N. Iyer, (2008). Feature extraction for bearing prognostics and health

# CHAPITRE

# 3

MULTICLASSIFICATION  
AVEC LES SEPARATEURS  
A VASTE MARGE

## **Chapitre 3. Multi classification avec les séparateurs à vaste marge**

### **3.1. Les algorithmes d'apprentissage automatique :**

Les algorithmes d'apprentissage automatique ont pour objet d'extraire de la connaissance à partir de données. Les algorithmes d'apprentissage automatique dit supervisés construisent un modèle de prédiction en utilisant un ensemble d'objets préalablement étiquetés, dit ensemble d'apprentissage.

À partir de l'ensemble d'apprentissage, ou base d'exemples, l'algorithme estime les paramètres du modèle de prédiction les plus performants possible, c'est-à-dire qui produisent le moins d'erreurs en prédiction.

À partir de ce modèle qui a été construit avec les paramètres optimaux, il sera possible de prédire la classe ou la catégorie d'un nouvel objet qui lui est soumis.

Le but de l'algorithme est donc de généraliser pour des entrées inconnues ce qu'il a pu « apprendre » grâce aux données de la base d'exemples, ceci de façon « raisonnable » en se basant sur des calculs de distance et de similarité. Un de ces algorithmes d'apprentissage automatique supervisé est le séparateur à vaste marge (SVM) ou machines à vecteurs de support.

#### 3.1.1. Algorithmes d'apprentissage supervisé communs

Techniques d'apprentissage supervisé (étiquetés) pour construire des modèles prédictifs à partir de données d'entrée et de réponses connues :

- Arbres de décision boostés et agrégés
- Machines à vecteur de support
- Classificateur bayésien naïf
- Classification des k plus proches voisins
- Analyse discriminante
- Réseaux de neurones

#### 3.1.2. Algorithmes d'apprentissage non supervisé communs

Techniques d'apprentissage non supervisé (non étiquetés) pour trouver des motifs cachés ou des structures intrinsèques au cœur de données :

- k-means et groupement hiérarchique
- Mélanges gaussiens (utilisant la maximisation de l'espérance conditionnelle)
- Les chaînes de Markov cachées
- Les cartes auto-organisatrices (SOM)
- Les réseaux de neurones par couche compétitive

- Les réseaux de neurones d'apprentissage par quantification de vecteur (LVQ) [Mathworks].

### 3.2. Séparateur à vaste marge (SVM) :

Les séparateurs à vastes marges reposent sur deux idées clés : la notion de marge maximale et la notion de fonction noyau.

L'idée des hyperplans à marge maximale a été explorée dès 1963 par Vladimir Vapnik et A. Lerner, et en 1973 par Richard Duda et Peter Hart dans leur livre *Pattern Classification*.

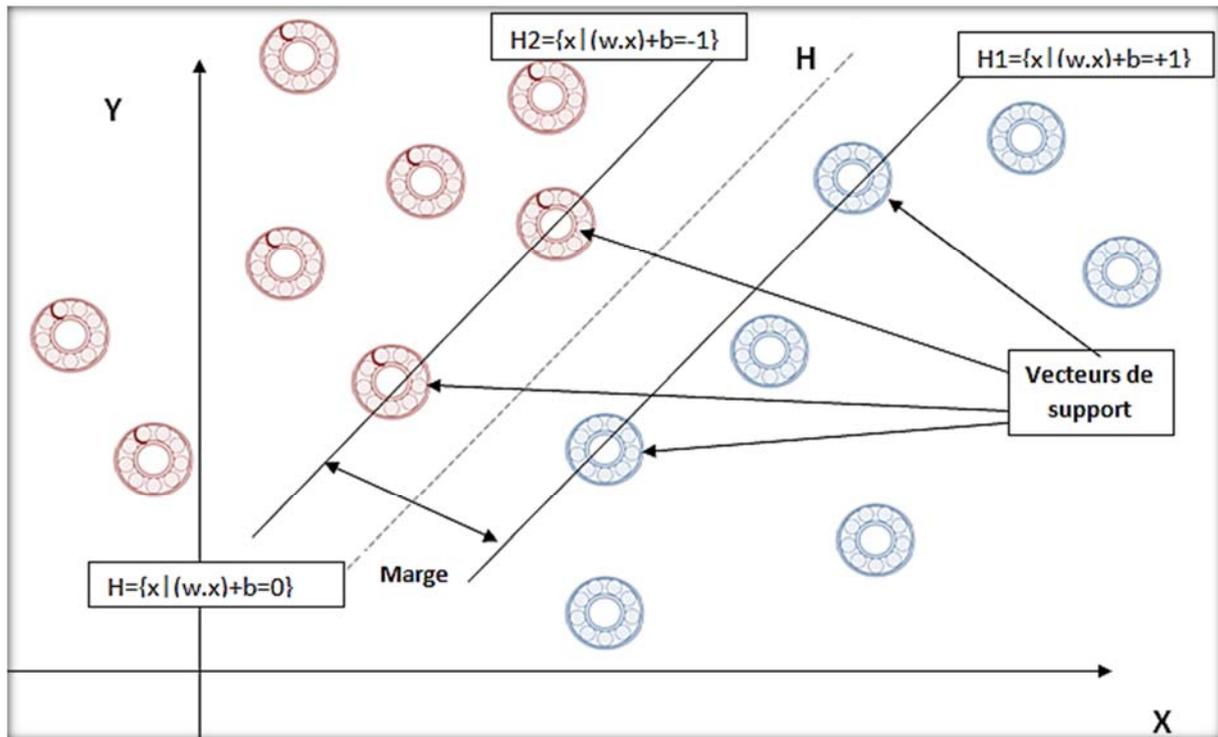
Un SVM appartient à un ensemble de techniques d'apprentissage comme un perceptron, il trouve un séparateur linéaire entre les points de données de deux classes différentes,

En général, il peut y avoir plusieurs séparateurs possibles entre les classes (en supposant le problème linéairement séparable) et qu'un perceptron n'a pas de préférence parmi celles-ci, dans les SVMs, cependant, nous faisons un choix particulier parmi tous les séparateurs possibles : nous voulons celui avec la "marge" maximale.

### **3.3. Hyperplan, marge et support vecteur :**

Le but du SVM-binaire est de trouver un classificateur qui va séparer les données et maximiser la distance entre deux classes. Dans l'exemple qui suit le classificateur est un classificateur linéaire appelé hyperplan qui sépare les deux ensembles de points.

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de support.



**Figure 3.1 :** Illustration d'une classification binaire linéaire SVM

Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe « au milieu » des points des deux classes d'exemples. Intuitivement, cela revient à chercher l'hyperplan le « plus sûr ». En effet, supposons qu'un exemple n'ait pas été décrit parfaitement, une petite variation ne modifiera pas sa classification si sa distance à l'hyperplan est grande. Formellement, cela revient à chercher un hyperplan dont la distance minimale aux exemples d'apprentissage est maximale.

On appelle cette distance « marge » entre l'hyperplan et les exemples. L'hyperplan séparateur optimal est celui qui maximise la marge. Comme on cherche à maximiser cette marge, on parlera de séparateurs à vaste marge

### 3.4. Séparateur à vaste marge dans le cas linéaire :

#### 3.4.1. Maximisation de marge :

La maximisation de la marge est une méthode de régularisation qui réduit la complexité du classifieur. Ce processus produit un ensemble réduit de prototypes faisant partie de l'ensemble d'apprentissage qu'on appelle communément vecteurs de support. Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classifieur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie (b) nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie (a) qu'avec une plus petite marge, l'exemple se voit mal classé.

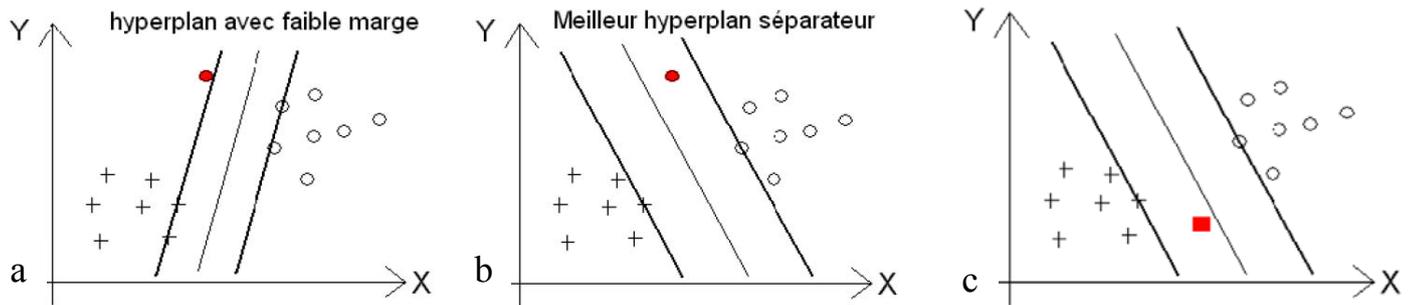


Figure 3.2 : Illustration du **choix** de la marge avec un hyperplan optimal

En général, la classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma (c), le nouvel élément sera classé dans la catégorie des « + ».

### 3.4.2. Calcul de la marge :

Rappelons qu'en deux dimensions la distance entre un point  $(x_0, y_0)$  et une droite

$Ax + By + c = 0$ , est donnée par la relation suivante :

$$\frac{|Ax_0 + By_0 + c|}{\sqrt{A^2 + B^2}} \quad (3.1)$$

De façon similaire, dans un modèle linéaire (voir figure ( 3.1)), la distance entre un point et le hyperplan séparateur qui a pour équation  $wx + b = 0$  est donnée par :

$$d(x) = \frac{|wx + b|}{\|w\|} \quad (3.2)$$

### 3.4.3. Calcul de l'Hyperplan optimal :

Donc l'hyperplan optimal est celui pour lequel la distance aux points les plus proches (Marge) est maximale. L'équation générale de l'hyperplan est  $H : wx + b = 0$  où :

- $w$  est le vecteur orthogonal à l'hyperplan.
- $x$  un vecteur de l'hyperplan.
- $b/\|w\|$  est la distance entre l'hyperplan et l'origine.

Doit satisfaire les conditions suivantes pour une classification binaire :

$$\begin{cases} Si y_i = +1 ; w x_i + b \geq 1 \\ Si y_i = -1 ; w x_i + b \leq -1 \\ Pour tous \{x_i, y_i\} y_i ; y_i(w x_i + b) \geq 1 \end{cases} \quad (3.3)$$

Soient  $f(x) = w \cdot x + b$  tel que  $x_1$  et  $x_2$  deux points de classes différentes :

$$(f(x_1) = +1 \text{ et } f(x_2) = -1).$$

$$(w \cdot x_1) + b = +1 \text{ et } (w \cdot x_2) + b = -1 \text{ donc } (w \cdot (x_1 - x_2)) = 2$$

D'où :  $(w \cdot (x_1 - x_2) / \|w\|) = 2/\|w\|$ .

On peut donc en déduire que maximiser la marge revient à minimiser son inverse  $\|w\|/2$  sous la contrainte  $y_i(w \cdot x + b) \geq +1$ .

$$\min \frac{1}{2} \|w\| = \min \frac{1}{2} w^T w \quad (3.4)$$

On a 
$$\begin{cases} \min \frac{1}{2} w^T w \\ \text{sous la contrainte} \\ y_i(w \cdot x + b) \geq +1 \end{cases} \quad (3.5)$$

Ce genre de problème d'optimisation peut être résolu en associant un multiplicateur de Lagrange  $\alpha_i \geq 0$ .

La règle dit que pour des contraintes de la forme  $C_i \geq 0$  pour former le lagrangien il faut multiplier le multiplicateur de Lagrange par la contrainte et le soustraire à de la fonction objective (fonction à minimiser), pour l'égalité des contraintes les multiplicateurs de Lagrange ne sont pas contraints. Cela donne Le Lagrangien suivant (Borges, C. J.1998) :

$$L(w, s, \alpha) = \frac{1}{2} w^t w - \sum_{i=1}^N \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^N \alpha_i \quad (3.6)$$

Pour trouver la solution du problème primal, il faut minimiser le lagrangien par rapport à  $w$  et le maximiser par rapport aux  $\alpha_i$ . Pour cette solution, les dérivées partielles du lagrangien s'annulent, selon les conditions de Kuhn-Tucker, on obtient :

$$\frac{\partial L(w,b,\alpha)}{\partial w} = 0, \frac{\partial L(w,b,\alpha)}{\partial b} = 0 \text{ et } \alpha_i \geq 0 \quad (3.7)$$

Cela revient à trouver les  $\alpha_i$  et  $w$  qui annulent ses dérivées partielles:

On trouve :

$$w = \sum_{i=1}^N \alpha_i y_i x_i \text{ et } \sum_{i=1}^N \alpha_i y_i = 0 \quad (3.8)$$

Et en les réinjectant dans le Lagrangien on obtient le Lagrangien dual qui doit être maximisé

$$L(w, b, \alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j \quad (3.9)$$

Si les  $\alpha_i^*$  sont solutions de ce problème c.à.d. les multiplicateurs optimaux correspondant aux points les plus proches et sont non-nuls. Alors les points  $x_i$  sont appelés **vecteurs de supports**.

Et on a le  $w^*$  optimal :  $w^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (3.10)$

En le remplaçant dans l'équation de l'hyperplan séparateur on obtient l'hyperplan optimal

$$h(x) = \sum_{i=1}^n \alpha_i^* y_i x_i \cdot x + b \quad (3.11)$$

La fonction de décision associée est donc :

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i^* y_i x_i \cdot x + b) \text{ (Hasan, M., \& Boris, F. 2006)} \quad (3.12)$$

3.5. Séparateur à vaste marge dans le cas non linéaire

Pour surmonter les inconvénients des cas non linéairement séparables, l'idée des SVM est d'utiliser des fonctions noyaux et de projeter les données de l'espace d'entrée (non linéairement séparables) dans un espace de plus grande dimension, de façon à ce que les données deviennent linéairement séparables. On va donc avoir un changement de dimension.

Plus la dimension du nouvel espace est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les classes est élevée, ceci est illustré dans la figure (3.3).

En pratique, il revient à l'utilisateur du SVM d'effectuer des tests pour déterminer **la fonction noyau** qui convient le mieux pour son application.

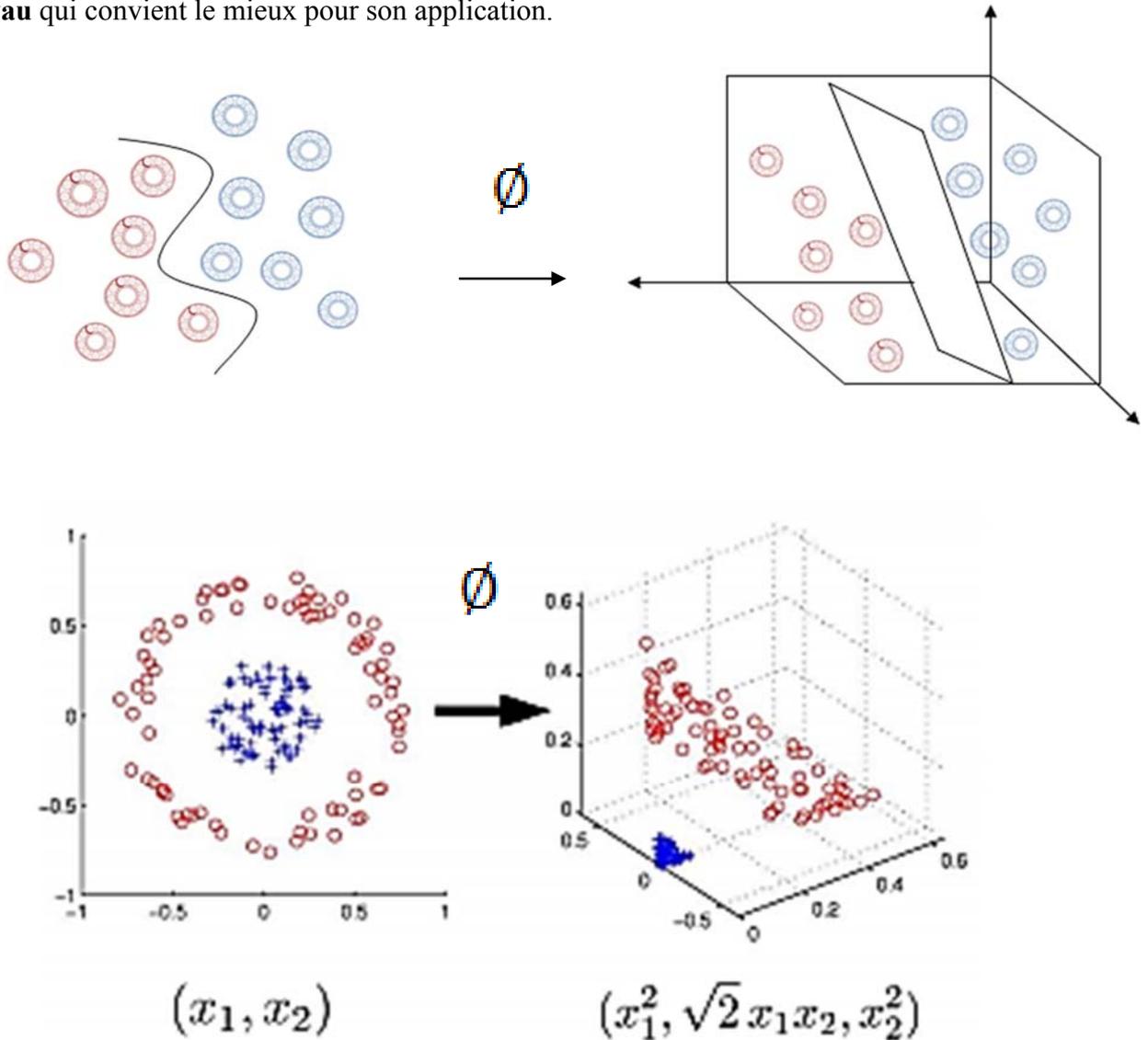


Figure 3.3 : Schéma de changement des représentations non linéaires vers un nouvel espace

3.5.1. Les Noyaux :

Le noyau permet de projeter les données dans un espace transformé *espace-φ* de grande dimension dans lequel s'opère plus facilement la séparation des classes.

Pour qu'il existe un *espace-φ* et une fonction  $\phi$  correspondant à un noyau  $K$ , il suffit que  $K$  soit un "noyau de Mercer",  $K$  doit vérifier les propriétés suivantes :

$K$  est une fonction continue, symétrique défini-positive qui associe à tout couple d'observations  $(x, x')$  une mesure de leur « influence réciproque » calculée à travers leur corrélation (les noyaux projectifs) ou leur distance (les noyaux radiaux)

On peut citer les exemples de noyaux : polynomiale, gaussien, sigmoïde et Laplacien.

**Les différentes fonctions du Noyau (Kernel) :** Les différentes fonctions du noyau Kernel sont les suivantes :

- Noyau linéaire :  $K(x, x') = x \cdot x'$  (3.13)

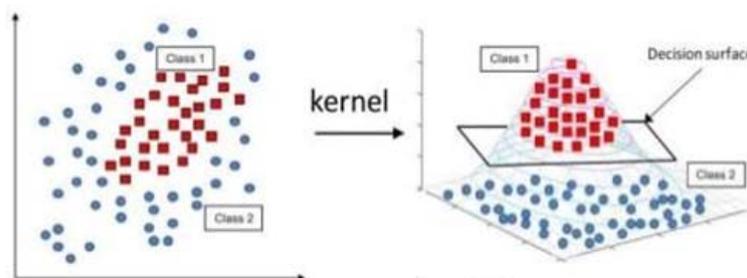


Figure 3.4 : Classification SVM binaire avec le noyau linéaire de Kernel.

- Noyau polynomial de degré  $d$  :  $K(x, x') = (a * \langle x, x' \rangle + b)^d$  .....(3.14)

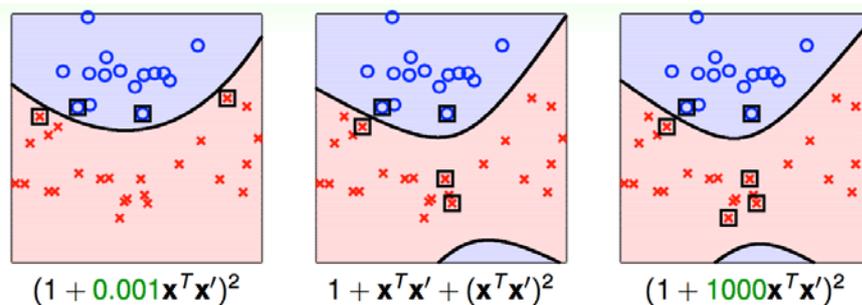


Figure 3.5 : Classification SVM binaire avec le noyau poly 2<sup>ième</sup> degré de Kernel.

- Noyau Gaussien (Radial Basis Function, RBF) :  $K(x, x') = e^{-\frac{\|x-x'\|^2}{2\sigma^2}}$  (3.15)

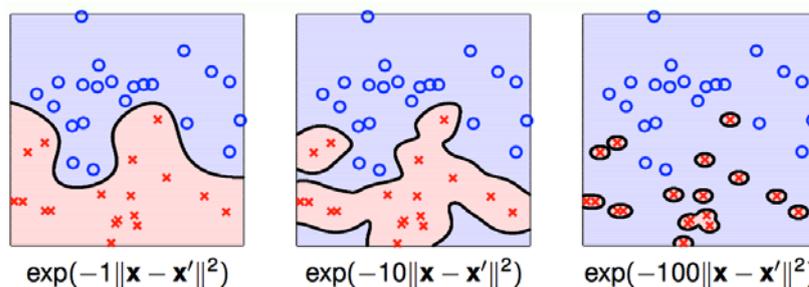


Figure 3.6 : Classification SVM binaire avec le noyau Gaussien de Kernel

### 3.5.2. Calcul de l'hyperplan optimal :

Le problème de l'application SVM aux cas non-linéairement séparables, se présente comme suit :

Étant donné un ensemble d'apprentissage de  $d$  observations dont les classes sont  $\{(x_i, c_i), i = 1, \dots, N\}$ , avec  $x_i = (x_{i1}, x_{i2}, \dots, x_{iM})$  une observation de dimension  $M$  et  $c_i \in \{+1, -1\}$ , non linéairement séparable, on cherche une transformation  $x \rightarrow \phi(x)$  telle que les nouvelles données  $\{(\phi(x_i), c_i), i = 1, \dots, N\}$  soient linéairement séparables dans  $\phi$ .

Généralement, l'image de  $x \rightarrow \phi(x)$  est de dimension  $N'$  supérieure à  $N$ , la dimension de l'espace d'origine.

La formulation mathématique de l'hyperplan optimal pour le SVM non linéaire est similaire à la formulation dans le cas linéaire sauf qu'une observation  $x_i$  est remplacée par la fonction non linéaire  $\phi(x_i)$ , et le problème dual est donné par :

$$\begin{cases} \max_{\alpha} \left\{ \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j c_i c_j \langle \phi_i(x_i) \phi_j(x_j) \rangle \right\} \\ \forall i \alpha_i \geq 0 \quad i = 1 \dots \dots \dots N \\ \sum_{i,j=1}^N \alpha_i c_i = 0 \end{cases} \quad (3.16)$$

L'objection immédiate concerne le produit scalaire  $\phi(x_i) \cdot \phi(x_j)$  qui devient rapidement impossible à calculer quand la dimension de  $\phi(x)$  augmente., il existe des fonctions noyau définies positives comme par exemple la fonction **noyau Kernel**,  $k(x_i, x_j)$  faciles à calculer et qui correspond à un produit scalaire dans un espace de grande dimension.

Alors On a  $\phi(x_i) \cdot \phi(x_j) = k(x_i, x_j)$ , où  $k$  est appelée fonction **noyau Kernel** (Ziani, R & al).

$$\begin{cases} \max_{\alpha} \left\{ \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j c_i c_j K(x_i, x_j) \right\} \\ \forall i \alpha_i \geq 0 \\ \sum_{i,j=1}^N \alpha_i c_i = 0 \end{cases} \quad (3.17)$$

La solution du problème dual donne les valeurs du multiplicateur de Lagrange  $\alpha^*$  et du vecteur  $w^*$  qui détermine l'hyperplan séparateur optimal tel que :

$$h(x) = (w^* \cdot x) + b^* = \sum_{i=1}^N \alpha_i^* c_i \langle \phi_i(x_i) \phi_j(x_j) \rangle + b^* \quad (3.18)$$

La classe d'une nouvelle observation  $x_k$  est déterminée par la règle suivante (Burges, C. J.1998). :

$$classe(x_k) = \text{signe}(\sum_{i=1}^d \alpha_i^* c_i \langle \phi_i(x_i) \phi_j(x_j) \rangle + b^*). \quad (3.19)$$

### 3.6. Types multi-class des séparateurs à vaste marge

On discute dans la suite deux méthodes utilisées dans les vecteurs de support traitant des problèmes multi classes.

- un contre tous.

- un contre un.

Dans la méthode "un contre tous", un problème de classification de  $n$  classes est converti en  $n$  classificateurs binaires et pour le classificateur 'i' binaire, la classe  $i$  est séparé des autres classes.

L'approche "un contre un" transforme le problème de  $n$  classes en  $n(n-1)/2$  classificateurs binaire qui couvrent tous les paires de classe (kotta, A, 2015).

### 3.6.1. Multi classification par l'approche un contre tous.

Pour un problème de  $N$  classes, le classificateur multiclassés1vsR établit d'abord  $N$  classificateurs bi-classe, chacun d'eux est responsable de chaque classe respectivement.

Pour le  $i$ ème classificateur bi-classe, toutes les données d'entraînement ont besoin d'être impliqué ; le résultat positif est le point de données dans la classe  $N$ , et le résultat négatif est le point de données dans les autres classes  $N-1$ .

Chacun des  $N$  classificateurs bi-classes, est formé pour trouver la fonction de décision, et alors toutes les fonctions de décision sont combinées pour obtenir la fonction finale de décision pour le problème de classification multi-classes figure (3.7).

Considérons un problème de  $n$  classes, pour l'approche "un contre tous", on détermine  $n$  fonctions de décision discrètes qui séparent une classe des autres classes avec une marge maximale :

$$D_{i(x)} = w_i^T \phi(x) + b_i \quad (3.21)$$

Si on a pour un vecteur d'entrée  $x$  :

$$D_{i(x)} > 0 \quad (3.22)$$

Les vecteurs de support doivent satisfaire :

$$y_i D_{i(x)} = 1$$

Le  $i$ ème SVM est formé avec tous les exemples dans la  $i$ ème classe avec des étiquettes positives et tous les autres exemples avec des étiquettes négatives, ainsi, l'hyperplan  $D_i(x) = 0$  forme l'hyperplan de séparation optimal avec  $D_i(x) \geq 1$  pour la  $i$ ème classe et ceux qui appartiennent aux classes restantes satisfont  $D_i(x) \leq -1$  (Abe, S. 2005).

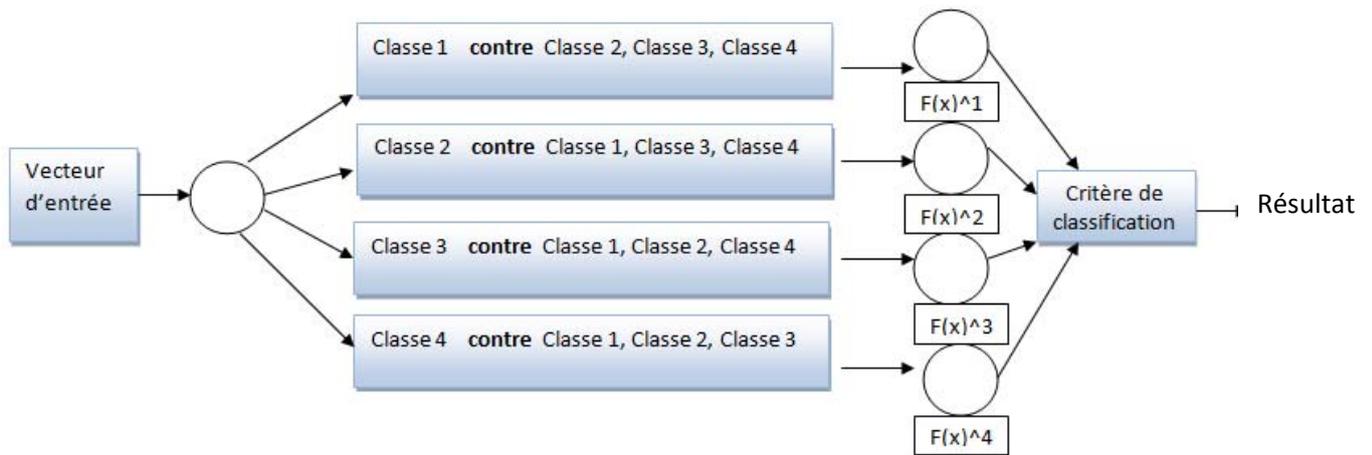


Figure 3.7 : Classification SVM un contre tous (One Vs All)

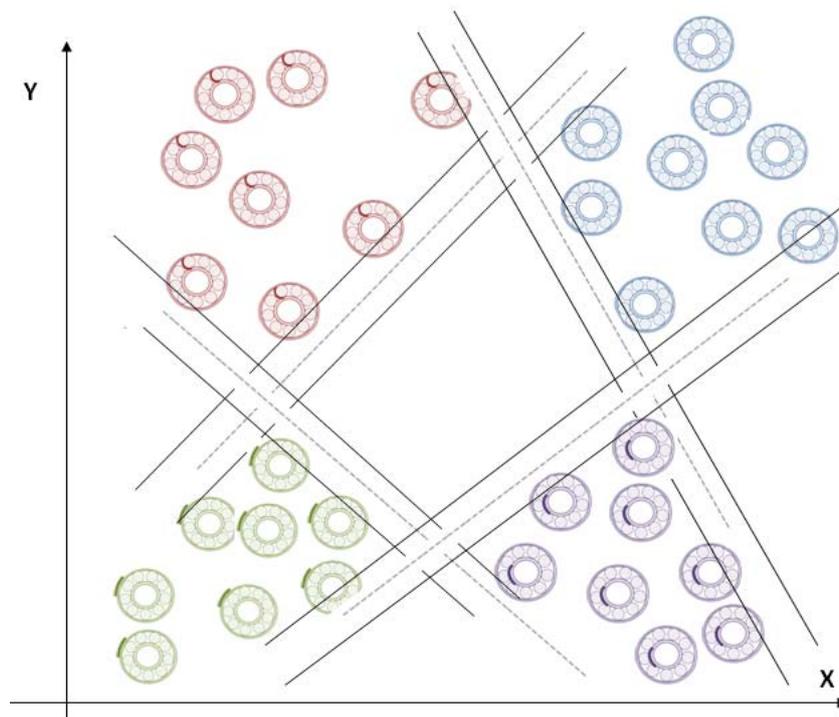


Figure 3.8 : Schéma des hyperplans linéaires pour la classification un contre tous (One vs all) /n=4

3.6.2. Multi classification par l'approche un contre un

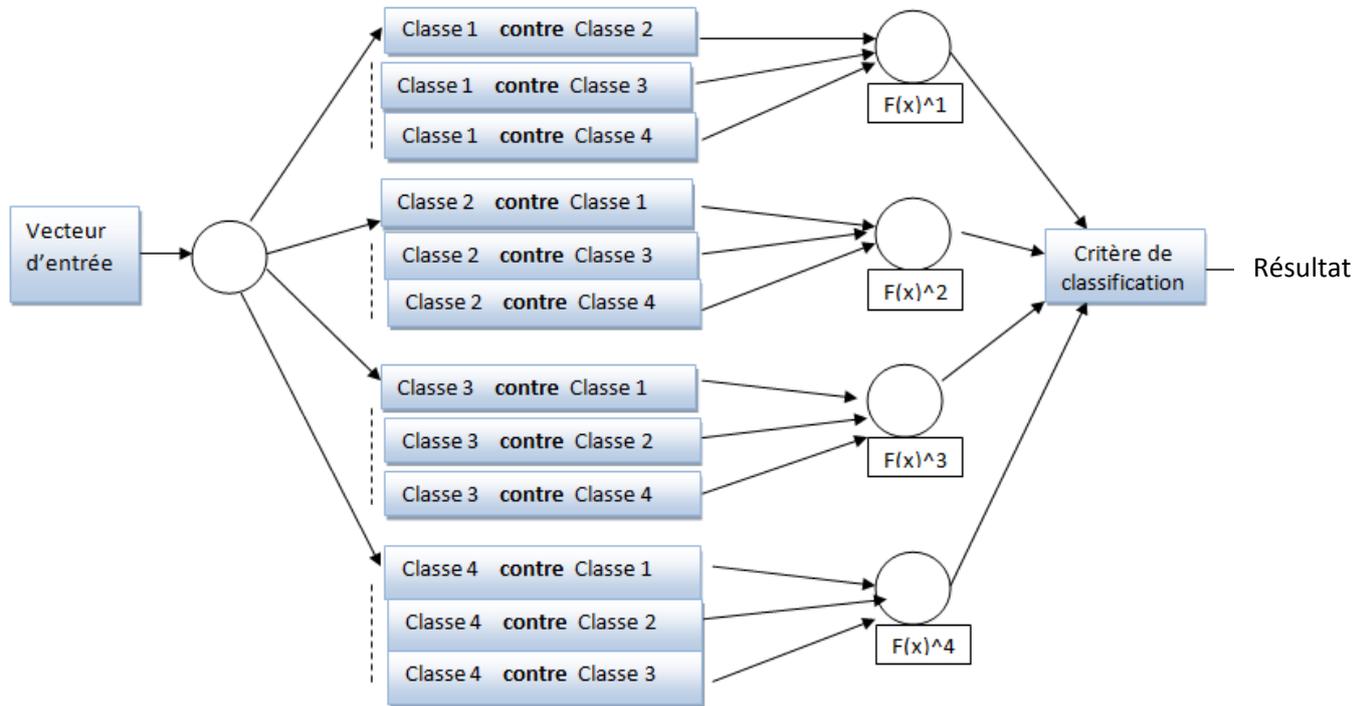


Figure 3.9 : Classification SVM un contre un (One Vs One)

Dans l'approche "un contre un", les fonctions de décision sont déterminées à partir des combinaisons des deux classes. Pour déterminer une fonction de décision pour une classe paire, nous utilisons les données d'apprentissage pour les deux classes correspondant. Ainsi, dans chaque session d'apprentissage, le nombre de données d'apprentissages est considérablement réduit par rapport à "un contre tous", qui utilisent tous les données d'apprentissage.

Mais le nombre de fonctions de décision est  $n(n-1)/2$  comparativement avec  $n$  de "un contre tous", où  $n$  est le nombre de classes.

Soit la marge à maximiser pour la classe  $i$  contre la classe  $j$  est :

$$D_{ij}(x) = w_{ij}^* \phi(x) - b_{ij}^* \quad (3.23)$$

Avec  $D_{ij} = -D_{ji}$ . (Widodo, 2007).

## **Références Bibliographiques**

Abe, S. (2005). *Support vector machines for pattern classification* (Vol. 2). London: Springer.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167.

Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines.

Hasan, M., & Boris, F. (2006). Svm: Machines à vecteurs de support ou séparateurs a vastes marges. Rapport technique, Versailles St Quentin, France. Cité, 64.

kotta, A. Condition Monitoring: Using Computational intelligence methods. 2015

Maïga, A. (2010). Détection Et Correction Automatique Des Défauts De Conception Au Moyen De L'apprentissage Automatique Pour L'amélioration De La Qualité Des Systèmes.

Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing*, 21(6), 2560-2574.

Ziani, R., Felkaoui, A., & Zegadi, R. (2007). Performances de la classification par les Séparateurs à Vaste Marge (SVM): application au diagnostic vibratoire automatisé. In *4th International Conference on Computer Integrated Manufacturing CIP* (pp. 03-04).

# CHAPITRE

# 4

## REVUE BIBLIORAPHIQUE

## Chapitre 4. Revue Bibliographique

**Tableau 4.1** : Résumé des articles traitant la base de données de la CWRU

Année	Pays	Référence	Nbre de classe	Nombre échant apprentissage /Largeur fenêtrage (bins)	Nbre échan test	Prétraitement	Classificateur géométrique statistique neural ou neuro-floue	Nombre de paramètres & Méthode d'extraction et de sélection	Taux de bonne classification %
2004	USA	[31]				The wavelet transform	ANFIS	6features /inputs &one output For training the targets are coded as: 1 (normal), 0 (inner race fault) and -1 (ball fault).	
2007	Canada	[27]	04			Time Synchronous Averaging TSA as a signal processing tool	SVM <i>one-against-ones</i>	Independent Component Analysis (ICA)	<i>DE bearing</i> Ball 88.9% IR 94.4% OUR 97.62%
2007	Chine	[28]	04	120			SVMs	Combination CoD and InD /3 features Combination FDF and TDF/3 features	99.579 100
2007	chine	[32]	4/10	30/A,B,C,D	30		6 ANFISs + GAs	EMD	A /100for the 06 classifier
2008	Inde	[29]	04			Training algorithm Levenberg-Marquardt, LM Resilient propagation, RP scaled conjugate gradient, SCG Powell-Beale conjugate gradient, PBCG	FFNN	10/8/2 time domain statistical selected features from: Pv,RMS,SD,Kv,Crf ,Clf,lmf,Shf,Wnl,Nnl And the sensitivity of the output nodes is studied per number of feature selected	100 MSE=0
2009	Chine	[1]	4	30/120 bins	30 120 bins		ANN SVM	32 8 selected: RMS,StD,ShF,CLF,Kurtosis Normalized Energy 5 /3 /14	85%-91% 97%-98%
2010	chine	[30]	04	2048pts 1.118 2.354			BP IL NCL NCCE	11 time-domain statistical features	BP 99.78 IL 99.84 NCL 99.81 NCCE 99.97
2011	Chine	[2]	4	6000 bins		Envelope demodulation technique based on: 1- wavelet packet transform (multi-band filter, 2- energy operator is applied to extract the fault feature of vibration signal	KNN	2: The relative spectrum entropy $rHr$ the gravity frequency $F_c$	Analysis
2011	Malaysia	[9]						DWT-RMS	Analysis
2011	Chine	[11]		20/6000pts	20		KNNC NBC	(MFDs) based on morphological cover (MC)	100 ~99

Année	Pays	Référence	Nomb de classe	Nombre d'échant apprentissage /Largeur fenêtrage (bins)	Nbre échan test	Prétraitement	LS-SVM Classificateur géométrique statistique neural ou neuro-floue	Nombre de paramètres & Méthode d'extraction et de sélection	Taux de bonne classification %
2012	Taiwan	[3]	4	2048 bins 10/20/30/40/50	200 test		SVM One <i>Versus</i> one	TDFDSFs PE MSE multiscale permutation entropy (MPE) is more robust on the variation of the training size. 1, 2, 3, 4, 5, 6, 8, 10, 12, 16, 20	~91% ~74% ~98% ~99%
2012	Livre	[10]	4/10	40	40	Based on morphological cover (MC)	KNN+MCS	EMD + fractal feature parameter extraction MFDs	100
2012	chine	[33]							
2013		[7]				comparison of parameters with 1. Robustness index. 2. Detectivity index. 3. sensitivity index			Analysis
2013	Chine	[12]		Case1 :426 pts Case2 :1706 pt			LSSVM	EMD-Principal Component Analysis (PCA),	96.92 95.53
2013		[17]		16/ 4096 pts	16	particle swarm optimization (PSO) wavelet packet	SVM	particle swarm optimization (PSO) algorithm based on the combining of the DWT preprocessor and the Kurtosis for feature extraction.	100
2013	Chine	[18]	10	15	35		SVM	1.EMD and Intrinsic Mode Function (IMF) envelope with sample entropy (SampEn) 2.EMD-IMF-SampEn	98 90
2014		[4]	4			There are 287 features in seven domains (time, real Cepstrum, minimum phase reconstruction, discrete cosine transform, discrete Fourier transform, envelope analysis signal and the Hilbert transform).	RBF_NN Bees Algorithm	First 6 features were selected by Total Distance  2 :Lognormal1 (DCT) and gev3 (minimum phase)	100%
2014	chine	[14]	6/9	500/4000pts	209	EEMD	ELM DE-ELM EELM EO-ELM	1.IMPE  2.EEMD + permutation entropy of each IMF	Max 98.85  Max 98.10

Année	Pays	Référence	Nomb de classe	Nombre d'échant apprentissage /Largeur fenêtrage (bins)	Nbre échan test	Prétraitement	Classificateur géométrique statistique neural ou neuro-floue	Nombre de paramètres & Méthode d'extraction et de sélection	Taux de bonne classification %
2014	Chine	[15]	10	5/10/20			1NN Off line modeling& on line monitoring	PCA LPP SLPP MMC LDA MFA TR-LDA1 TR-LDA2	90 90 92 93 93 93 95 96
2014	Iran	[21]	06	6/500pts 12 30 60 120 240 300 576	54 104 270 540 1080 2160 2700 5184		MSVM	FFT Multi -class Support Vector Machine 99.91 / time :0.043s Artificial Neural Network 96.46 /time 13.787s	84.48 85.63 99.26 99.63 99.91 99.95 99.88 100
2014	Chine	[26]	04	30	30	Calculate the RMSE	two-layer SVRMs	SVR with the selected parameters and optimization. Comparisons with other SVM and ANN methods.	1rst layer 98-100 2nd layer 98-100
2014	Tunisie	[34]	04	20 78 58	12 50 38		SVM-OAA classifier	Sum of logarithmic amplitudes of the bispectrum F1, Sum of logarithmic amplitudes of diagonal elements in the bi-spectrum F2, First-order spectral moment of amplitudes of diagonal elements in the bi-spectrum F3, Normalized bi-spectral entropy P1 ,Normalized bi-spectral squared entropy P2 ,Bi-spectrum phase entropy Pe	95 /n- selected.f 100/selected f
2015	chine	[13]	10	10/20/30/40				trace ratio criterion-based kernel discriminant analysis (TR-KDA-BIGA)	
2015	Australie	[24]				Method1: <i>envelope analysis of the raw signal</i> Method2: <i>cepstrum prewhitening</i> Method3: <i>benchmark method</i> All methods used the squared envelope spectrum		Ball 118, 119, 120DE, not diagnosable with any of the applied methods (only N1 or N2 diagnoses).	Analysis

2016	Chine	[5]	11	10	40		GRA algorithm	multifractal theory and gray relation theory	92~100
2016	Chine	[6]						CK with the RSGWPT	Analysis
2016	Chine	[8]	04			<i>particle swarm optimization (PSO) algorithm</i>	<i>Back-propagation neural network (BP)</i>	<i>particle swarm optimization (PSO) algorithm</i>	89~100
2016		[19]	11	10	40		AGRA	improved generalized fractal box-counting dimension algorithm with: 1. gray relation algorithm 2. adaptive gray relation algorithm	GRA 90-100 AGRA +2~+5
2016	Korè	[20]	04	144 96 72 57 48	432 288 216 171 144	1.Rectification 2.Segmentation (cycle length slices) 3.Stacking (multiple cycles)	KNN	Construction of vibration images with microtexture analysis of the grayscale vibration acceleration fault images	90-94 97 96-99 96-99 99-100
2016	chine	[22]	04				SVM	<i>IMASFD</i> : EMD- IMF components, the energy of IMF components is selected to construct the fault feature vectors. Then the adaptive controlling pheromone, adaptive controlling stochastic selection threshold and dynamic evaporation rate are <i>The improved ACO algorithm is used to optimize the parameters of SVM model in order to obtain the optimal values of parameter</i>	94.93%,96.04%, 99.74% and100.00%.
2016	Algérie	[23]				Fast Fourier transform(FFT) Ceptrum analysis (CA) Envelope Analysis (EA) Wavelet transform (WT)		1.The fault can be detected(FD) FD& with certain amount of confidence FD& it's nature can be detected FD& its nature can be detected	Analysis
2016		[25]		Load0≥5sampls Load1 Load2 Load3			Euclid distance discriminance	1. FFT are processed with two-dimensional principal component analysis (2DPCA) to reduce the dimensions 2. minimum distance applied toclassify	PCA 2DPCA 75~100
2017	chine	[16]	10			Anti-Noise and Domain Adaptation	WDCNN	1. used wide kernels, and then use successive small 3*1 kernels to acquire better feature . 2.W denoting wide kernels in first layer and D denoting the deep structure	FFT-SVM/68 FFT-MLP/82 FFT-DNN/82 WDCNN/99



Travaux effectués avec le classificateur à vaste Marges (SVM)

## Références Bibliographiques

- [1] Sui, W. T., & Zhang, D. (2009, July). Rolling element bearings fault classification based on SVM and feature evaluation. In *Machine Learning and Cybernetics, 2009 International Conference on* (Vol. 1, pp. 450-453). IEEE.
- [2] Wei, Z., Gao, J., Zhong, X., Jiang, Z., & Ma, B. (2011). Incipient fault diagnosis of rolling element bearing based on wavelet packet transform and energy operator. *WSEAS TRANSACTIONS on SYSTEMS*, 10(3), 81-90.
- [3] Wu, S. D., Wu, P. H., Wu, C. W., Ding, J. J., & Wang, C. C. (2012). Bearing fault diagnosis based on multiscale permutation entropy and support vector machine. *Entropy*, 14(8), 1343-1356.
- [4] Attaran, B., & Ghanbarzadeh, A. (2014). Bearing fault detection based on maximum likelihood estimation and optimized ann using the bees algorithm. *Journal of Applied and Computational Mechanics*, 1(1), 35-43.
- [5] Li, J., Cao, Y., Ying, Y., & Li, S. (2016). A Rolling Element Bearing Fault Diagnosis Approach Based on Multifractal Theory and Gray Relation Theory. *PloS one*, 11(12), e0167587.
- [6] Chen, X., Feng, F., & Zhang, B. (2016). Weak Fault Feature Extraction of Rolling Bearings Based on an Improved Kurtogram. *Sensors*, 16(9), 1482.
- [7] Shakya, P., Darpe, A. K., & Kulkarni, M. S. (2013). Vibration-based fault diagnosis in rolling element bearings: ranking of various time, frequency and time-frequency domain data-based damage identification parameters. *International journal of condition monitoring*, 3(2), 53-62.
- [8] Li, S., & Zhou, D. (2016). Study on a New Fault Diagnosis Method Based on Combining Intelligent Technologies. *International Journal of Multimedia and Ubiquitous Engineering*, 11(6), 61-72.
- [9] Chebil, J., Hrairi, M., & Abushikhah, N. (2011). Signal analysis of vibration measurements for condition monitoring of bearings. *Australian Journal of Basic and Applied Sciences*, 5(1), 70-78.
- [10] Huang, J. (2012). *Bearing Fault Diagnosis Using Information Fusion and Intelligent Algorithms*. INTECH Open Access Publisher.
- [11] Zhang, P. L., Li, B., Mi, S. S., Zhang, Y. T., & Liu, D. S. (2012). Bearing fault detection using multi-scale fractal dimensions based on morphological covers. *Shock and Vibration*, 19(6), 1373-1383.
- [12] Zhang, Y., Qin, Y., Xing, Z., Jia, L., & Cheng, X. (2013). Safety region estimation and state identification of rolling bearing based on statistical feature extraction. *Shock and Vibration*, 20(5), 833-846.

- [13] Yang, W. A., Xiao, M., Zhou, W., Guo, Y., Liao, W., & Shen, G. (2016). Trace Ratio Criterion-Based Kernel Discriminant Analysis for Fault Diagnosis of Rolling Element Bearings Using Binary Immune Genetic Algorithm. *Shock and Vibration*, 2016.
- [14] Zhou, J., Xiao, J., Xiao, H., Zhang, W., Zhu, W., & Li, C. (2014). Multifault diagnosis for rolling element bearings based on intrinsic mode permutation entropy and ensemble optimal extreme learning machine. *Advances in Mechanical Engineering*.
- [15] Zhao, M., Jin, X., Zhang, Z., & Li, B. (2014). Fault diagnosis of rolling element bearings via discriminative subspace learning: visualization and classification. *Expert Systems with Applications*, 41(7), 3391-3401.
- [16] Zhang, W., Peng, G., Li, C., Chen, Y., & Zhang, Z. (2017). A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. *Sensors*, 17(2), 425.
- [17] Thelaidjia, T., & Chenikher, S. (2013, December). A New approach of preprocessing with SVM optimization based on PSO for bearing fault diagnosis. In *Hybrid Intelligent Systems (HIS), 2013 13th International Conference on* (pp. 319-324). IEEE.
- [18] Zhu, K., Song, X., & Xue, D. (2013). Fault diagnosis of rolling bearings based on IMF envelope sample entropy and support vector machine. *JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE*, 10(16), 5189-5198.
- [19] Cao, Y., Ying, Y., Li, J., Li, S., & Guo, J. (2016). Study on rolling bearing fault diagnosis approach based on improved generalized fractal box-counting dimension and adaptive gray relation algorithm. *Advances in Mechanical Engineering*, 8(10), 1687814016675583.
- [20] Khan, S. A., & Kim, J. M. (2016). Automated Bearing Fault Diagnosis Using 2D Analysis of Vibration Acceleration Signals under Variable Speed Conditions. *Shock and Vibration*, 2016.
- [21] ESTILAF1a, H. A., & Fatemi, S. J. R. (2014). Bearing fault diagnosis of electrical machine base on vibration signal using multi-class support vector machine. *Indian J. Sci. Res*, 2(1), 46-53.
- [22] Deng, W., Li, X., & Zhao, H. (2016). Study on A Fault Diagnosis Method of Rolling Element Bearing Based on Improved ACO and SVM Model. *International Journal of Future Generation Communication and Networking*, 9(3), 167-180.
- [23] Boudiaf, A., Moussaoui, A., Dahane, A., & Atoui, I. (2016). A Comparative Study of Various Methods of Bearing Faults Diagnosis Using the Case Western Reserve University Data. *Journal of Failure Analysis and Prevention*, 16(2), 271-284.
- [24] Smith, W. A., & Randall, R. B. (2015). Rolling element bearing diagnostics using the Case Western Reserve University data : A benchmark study. *Mechanical Systems and Signal Processing*, 64, 100-131.

- [25] Li, W., Qiu, M., Zhu, Z., Wu, B., & Zhou, G. (2016). Bearing fault diagnosis based on spectrum images of vibration signals. *Measurement Science and Technology*, 27(3), 035005.
- [26] Shen, C., Wang, D., Liu, Y., Kong, F., & Tse, P. W. (2014). Recognition of rolling bearing fault patterns and sizes based on two-layer support vector regression machines. *Smart Structures and Systems*, 13(3), 453-471.
- [27] Christian, K., Mureithi, N., Lakis, A., & Thomas, M. (2007). On the use of time synchronous averaging, independent component analysis and support vector machines for bearing fault diagnosis. In *First international conference on industrial risk engineering*.
- [28] Yang, J., Zhang, Y., & Zhu, Y. (2007). Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. *Mechanical Systems and Signal Processing*, 21(5), 2012-2024.
- [29] Sreejith, B., Verma, A. K., & Srividya, A. (2008, December). Fault diagnosis of rolling element bearing using time-domain features and neural networks. In *Industrial and Information Systems, 2008. ICIIS 2008. IEEE Region 10 and the Third International Conference on* (pp. 1-6). IEEE.
- [30] Wang, Q., Zhang, Y., & Zhu, Y. (2010). Neural network compact ensemble and its applications. *Chinese Journal of Mechanical Engineering*, (2), 209.
- [31] Lou, X., & Loparo, K. A. (2004). Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mechanical systems and signal processing*, 18(5), 1077-1095.
- [32] Lei, Yaguo, et al. "Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs." *Mechanical systems and signal processing* 21.5 (2007): 2280-2294.
- [33] Zhang, P. L., Li, B., Mi, S. S., Zhang, Y. T., & Liu, D. S. (2012). Bearing fault detection using multi-scale fractal dimensions based on morphological covers. *Shock and Vibration*, 19(6), 1373-1383.
- [34] Saidi, L., Ali, J. B., & Fnaiech, F. (2015). Application of higher order spectral features and support vector machines for bearing faults classification. *ISA transactions*, 54, 193-206.

# CHAPITRE

# 5

## RESULTATS ET DISCUSSION

## Chapitre 5. Résultats et discussion

### 5.1. Matériels et méthodes :

#### 5.1.1. Présentation du Banc d'Essai de la CWRU :

Comme représenté sur la figure 5.1, le banc d'essai se compose principalement d'un moteur (à gauche), d'un accouplement « transducteur/encodeur » (au centre), d'un dynamomètre (à droite) et d'un module d'acquisition des données.

Les signaux ont été enregistrés à l'aide des accéléromètres, qui ont été fixés à la cage du moteur par aimant. Pour différents couples résistants (donc différentes vitesses de rotations).

Les couples sont respectivement 0, 1, 2, 3 Hp dont des vitesses respectives sont 1797, 1772, 1750, 1730 tr/min.

Les capteurs ont été placés à la position 12 heures verticale des deux côtés accouplement et ventilateur de la cage du moteur en plus de deux autres capteurs à 6 et 3 heures.

La fréquence d'échantillonnage est  $f_e = 12$  KHz et le nombre de points total est de  $N_p = 243938$  pts.

Les données ont été enregistrées pour :

- Les roulements sans défaut.
- Les défauts uniques du cotés accouplement (12.000 et 48.000 points/second).
- Les défauts uniques du cotés ventilateur (12.000 points/second.)

Dans ce travail, les roulements que nous avons considérés, sont le roulement du coté accouplement pour l'apprentissage et les tests. Les fréquences caractéristiques sont :

- Fréquence de la bague intérieure :  $f_{bi} = 5.415 \times f_r$
- Fréquence de la bague extérieure :  $f_{be} = 3.584 \times f_r$
- Fréquence de la cage :  $f_c = 0.3982 \times f_r$
- Fréquence de la bille :  $f_b = 4.7135 \times f_r$

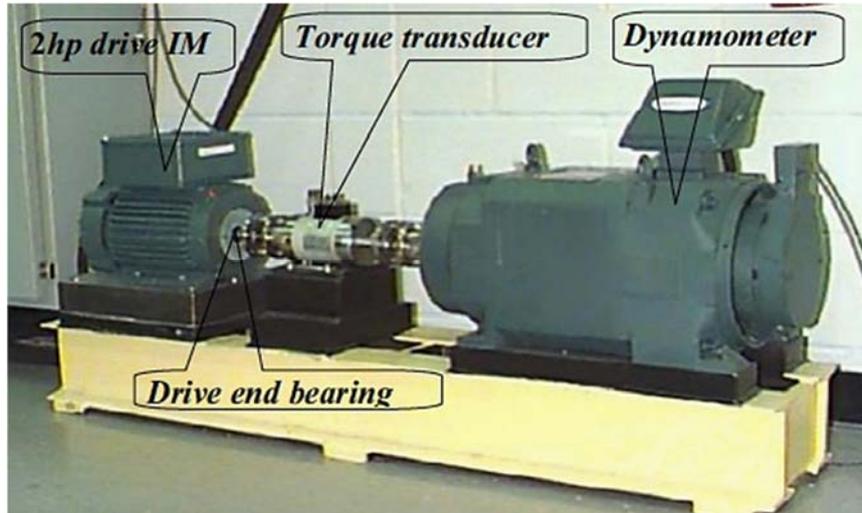
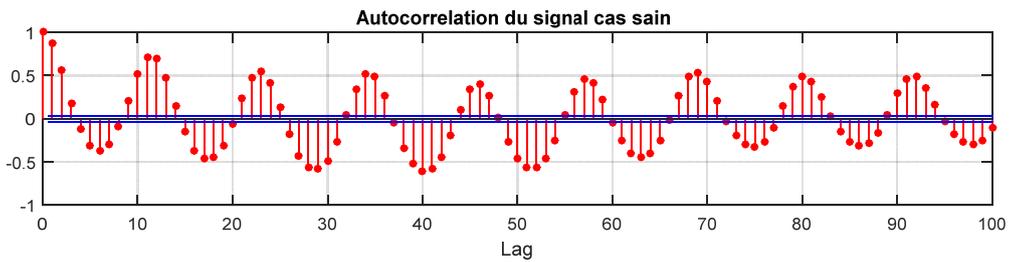
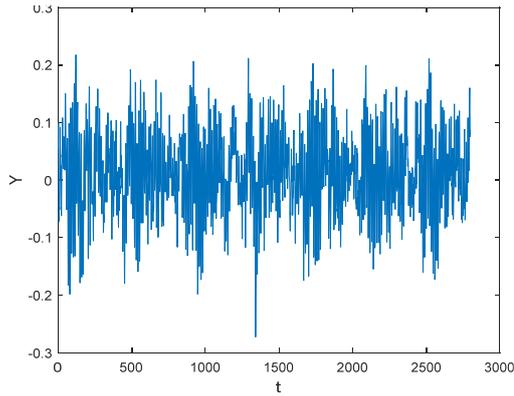


Figure 5.1 : Le Banc d'essai de la CWRU

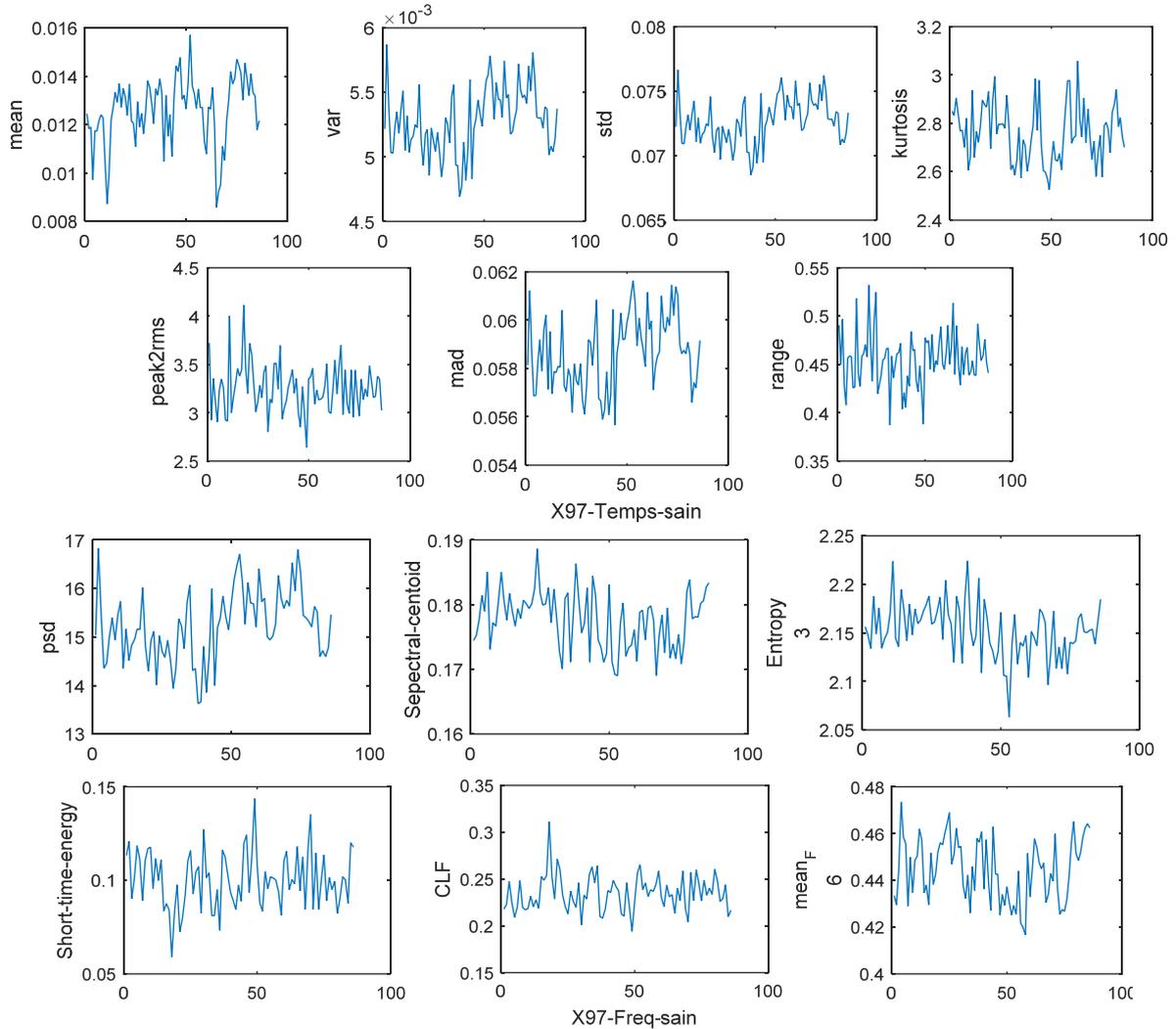
Tableau 2 : données vibratoires des défauts avec une fréquence d'échantillonnage de 12K HZ / CWRU

Fault Diameter (inches)	Motor Speed (rpm)	Inner Race	Ball	Outer Race Position Relative to Load Zone (Load Zone Centered at 6:00)		
				Centered @6:00	Orthogonal @3:00	Opposite @12:00
0.007"	1797	I07_0	B07_0	O07@6_0	O07@3_0	O07@12_0
	1772	I07_1	B07_1	O07@6_1	O07@3_1	O07@12_1
	1750	I07_2	B07_2	O07@6_2	O07@3_2	O07@12_2
	1730	I07_3	B07_3	O07@6_3	O07@3_3	O07@12_3
0.014"	1797	I14_0	B14_0	O14@6_0		
	1772	I14_1	B14_1	O14@6_1		
	1750	I14_2	B14_2	O14@6_2		
	1730	I14_3	B14_3	O14@6_3		
0.021"	1797	I21_0	B21_0	O21@6_0	O21@3_0	O21@12_0
	1772	I21_1	B21_1	O21@6_1	O21@3_1	O21@12_1
	1750	I21_2	B21_2	O21@6_2	O21@3_2	O21@12_2
	1730	I21_3	B21_3	O21@6_3	O21@3_3	O21@12_3
0.028"	1797	I28_0	B28_0			
	1772	I28_1	B28_1			

**Figure 5.2 : Signal original cas sain**



**Figure 5.3 : Auto corrélation du signal original cas sain**



**Figure 5.4 : Variation des paramètres temporels et fréquentiels du signal sans défauts (sain)**

Figure 5.5 : Signal original avec défaut bille

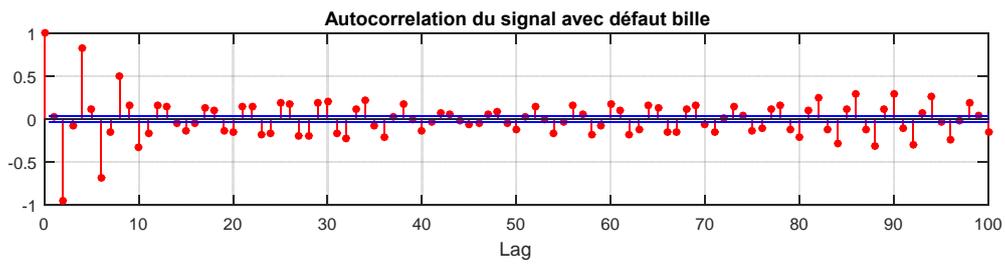
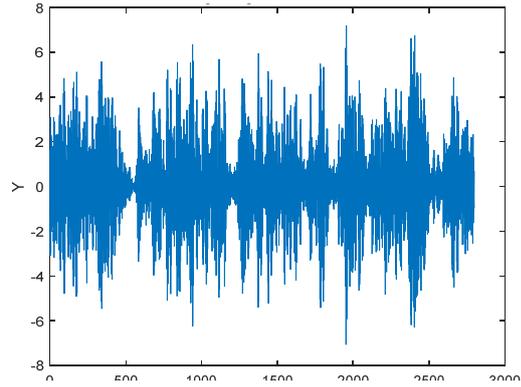


Figure 5.6 : Auto corrélation du signal ave défaut bille

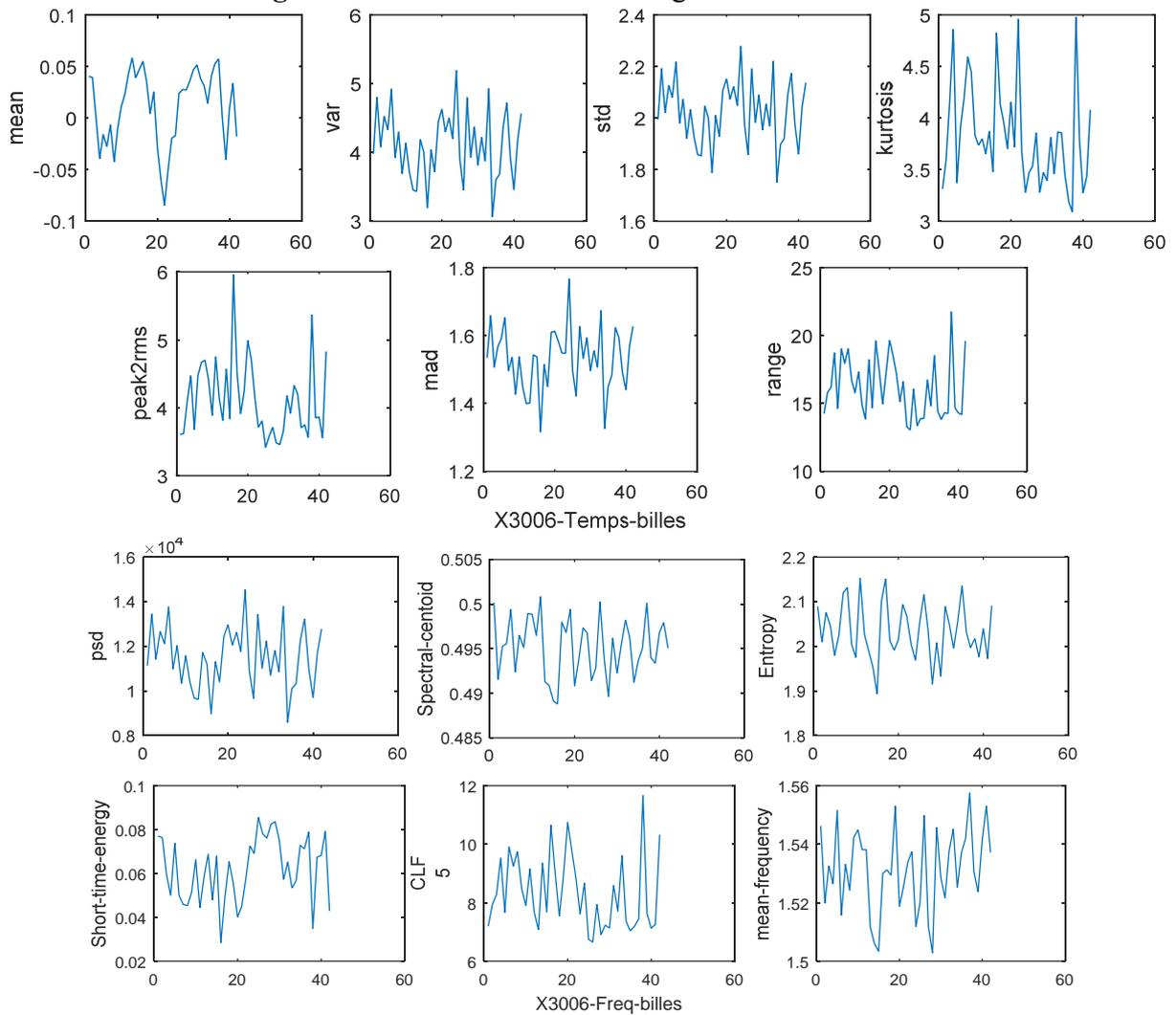
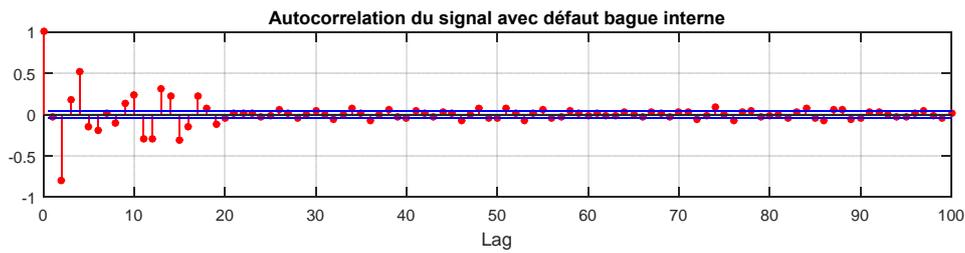
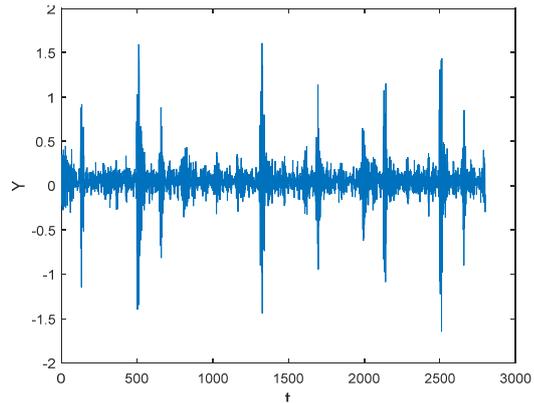
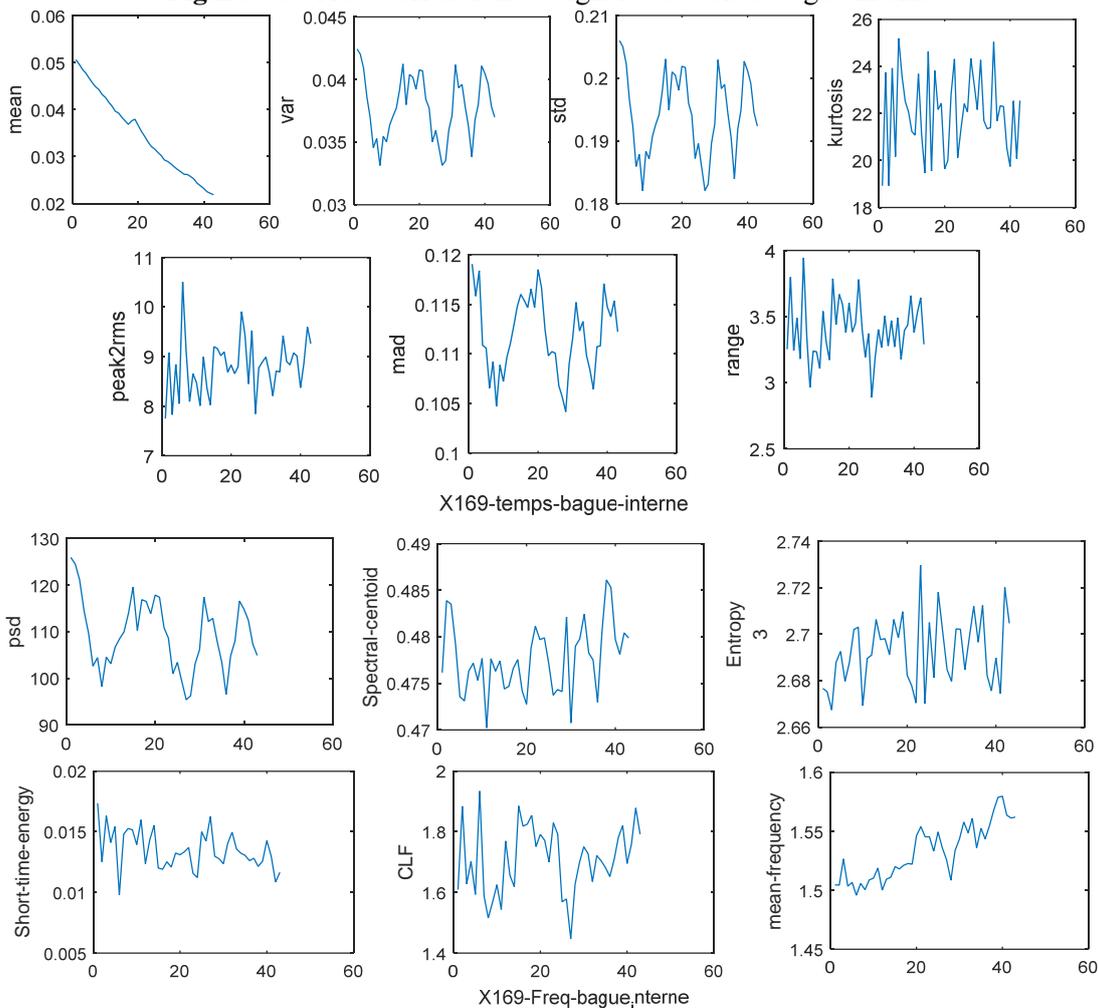


Figure 5.7 : Variation des paramètres temporelles et fréquentielles du signal avec défaut bille

**Figure 5.8** : Signal original avec défaut bague interne



**Figure 5.9** : Auto corrélation du signal avec défaut bague interne



**Figure 5.10** : Variation des paramètres temporels et fréquentiels du signal avec défaut bague interne.

Figure 5.11: Signal original avec défaut bague externe

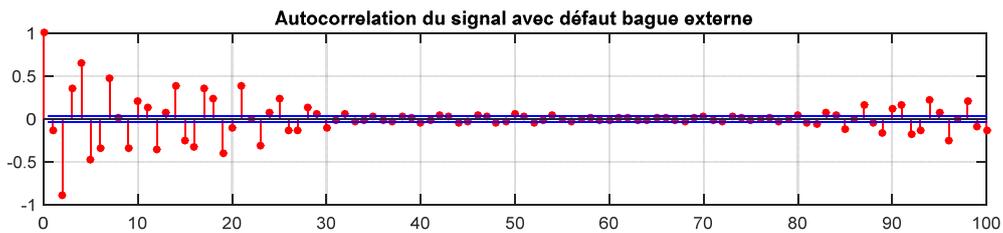
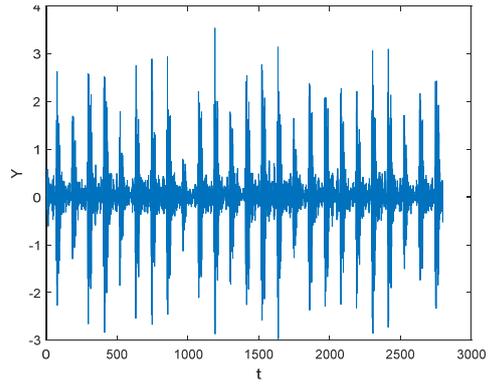


Figure 5.12 : Auto corrélation du signal avec défaut bague externe.

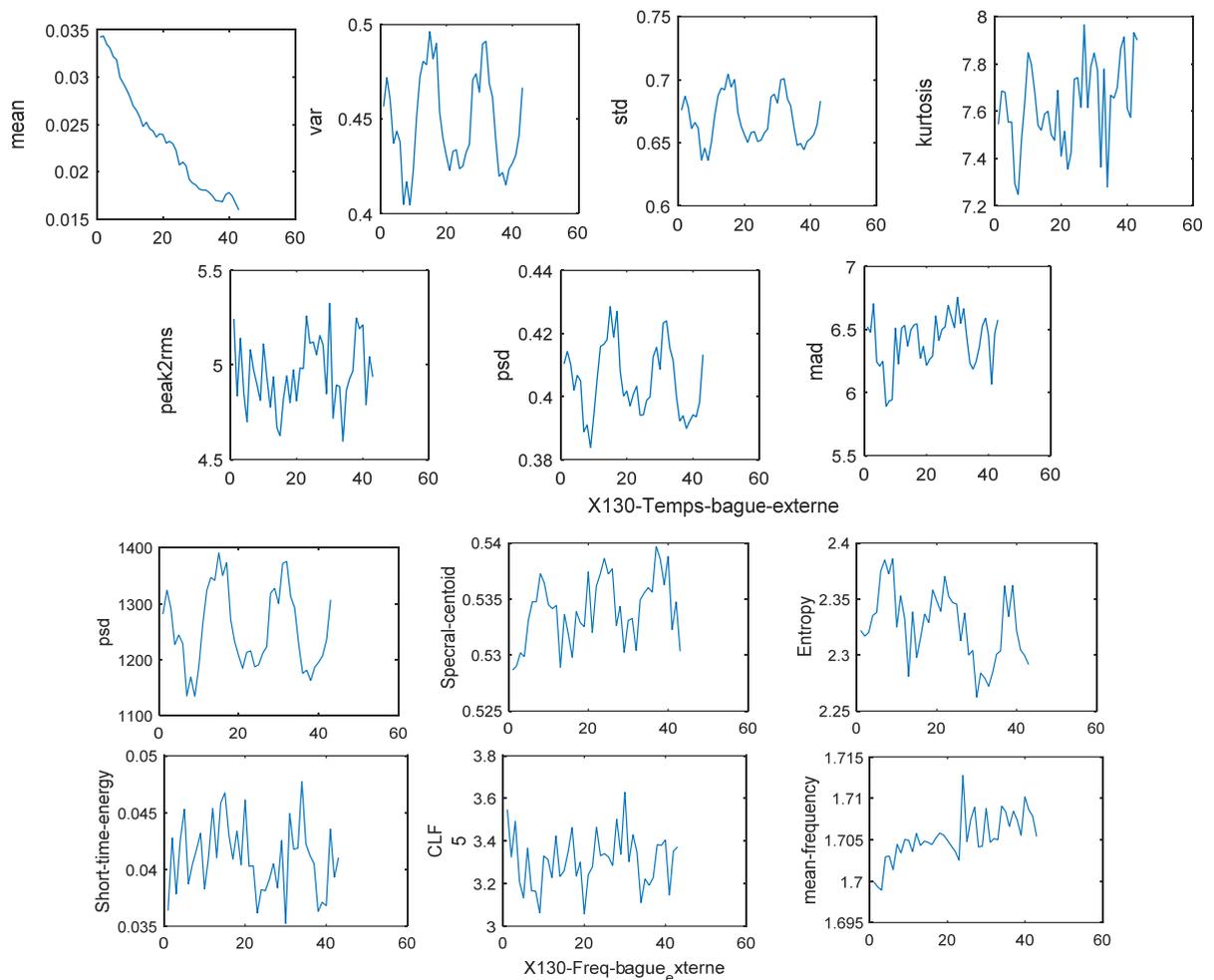


Figure 5.13 : Variation des paramètres temporels et fréquentielles du signal avec défaut bague externe

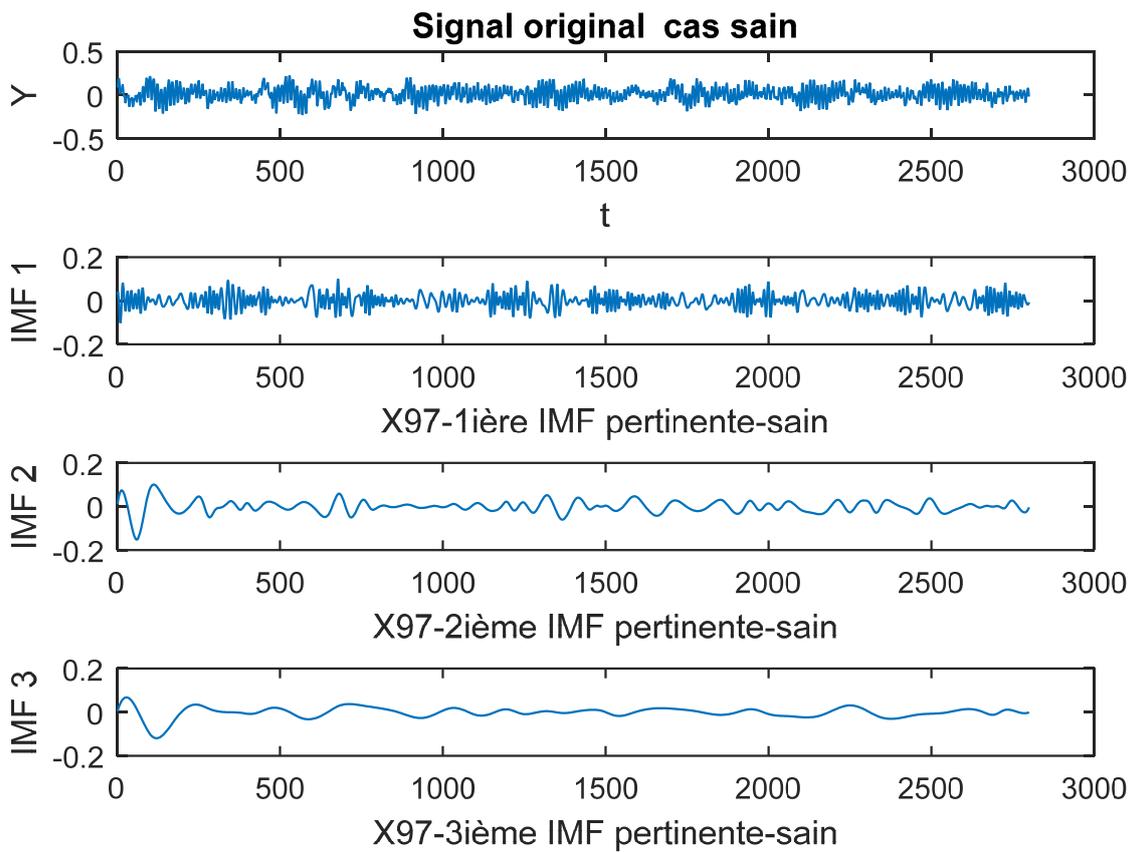


Figure 5.14: Extraction des 3IMFs les plus pertinentes cas sain

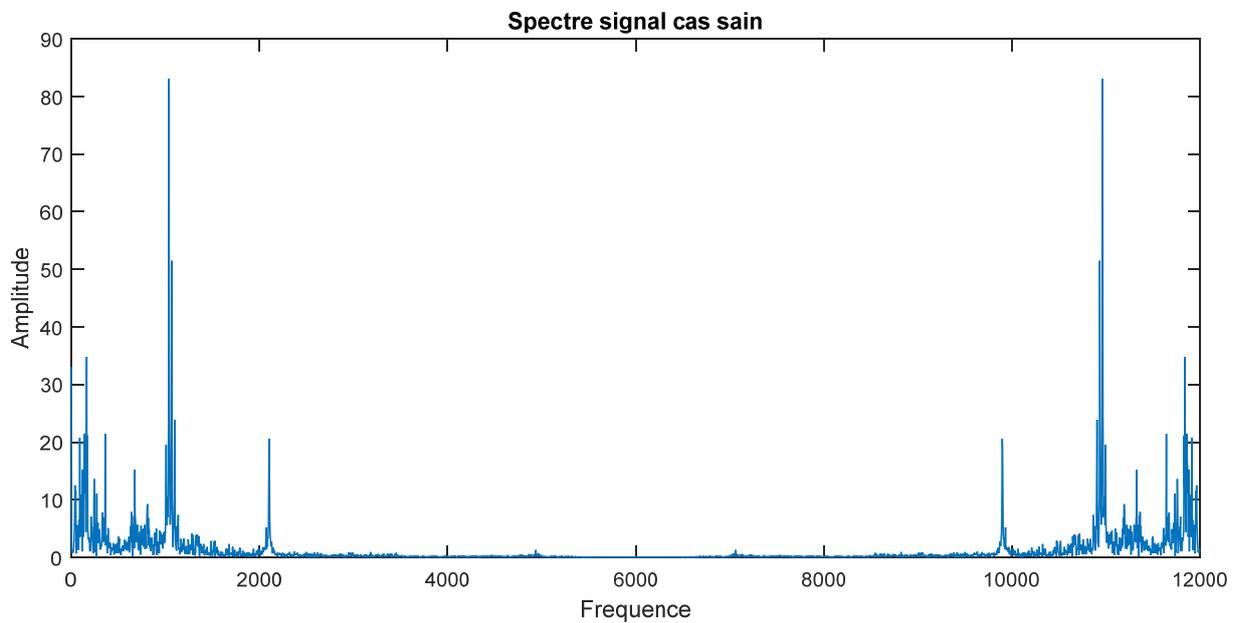
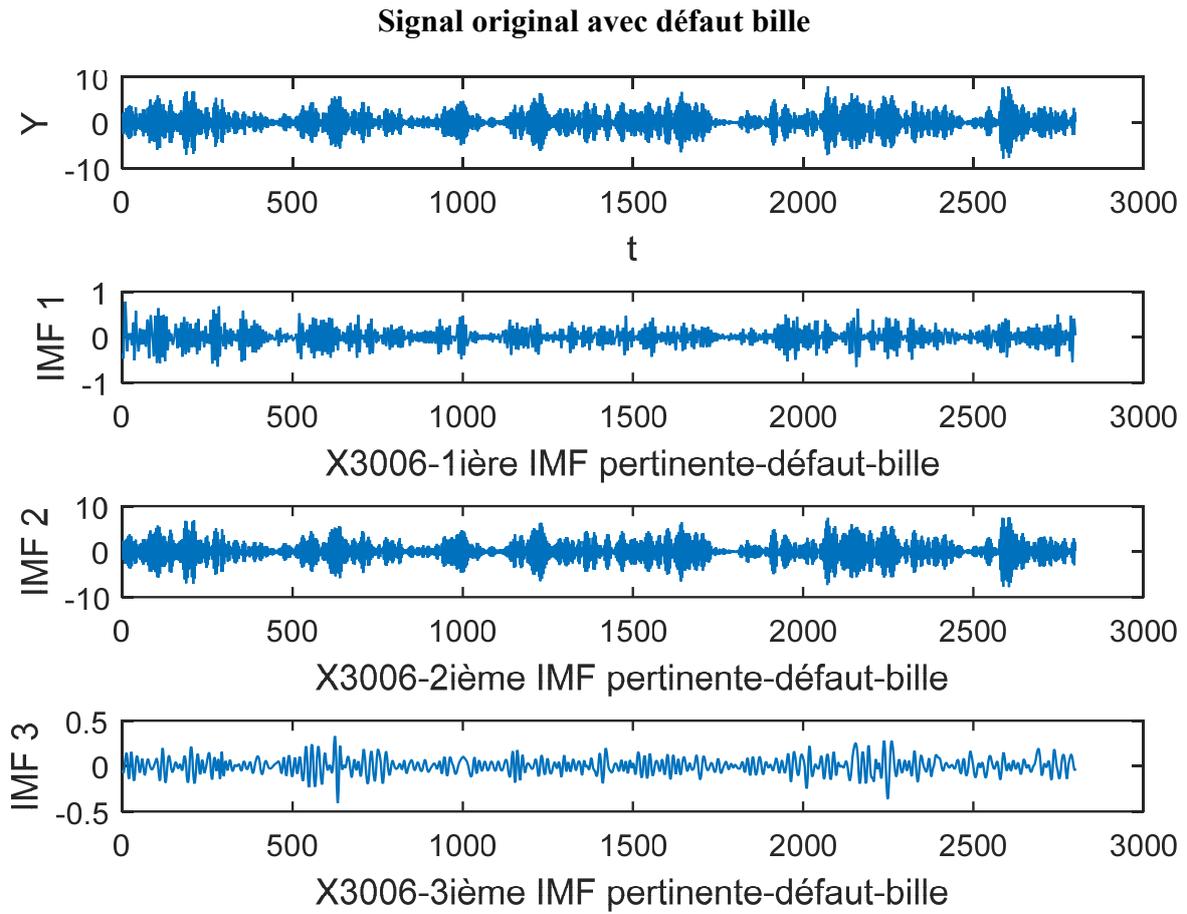
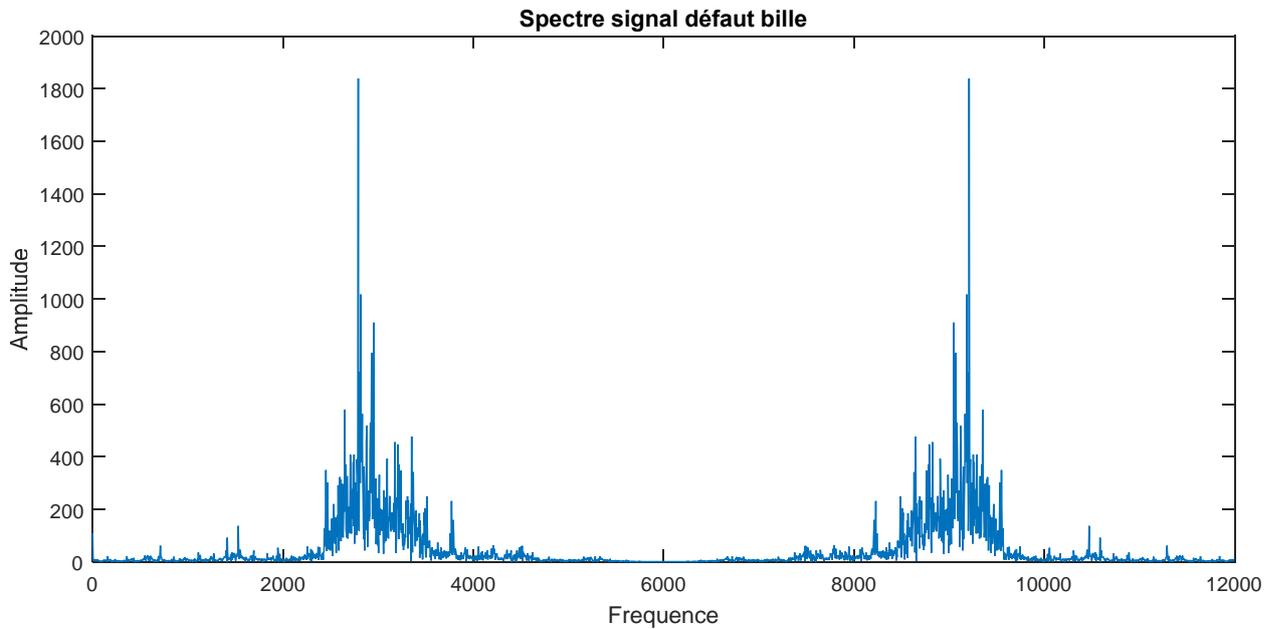


Figure 5.15 : Spectre du signal cas sain



**Figure 5.16** : Extraction des 3 IMFs les plus pertinentes du signal avec défaut bille



**Figure 5.17** : Spectre du signal avec défaut bille

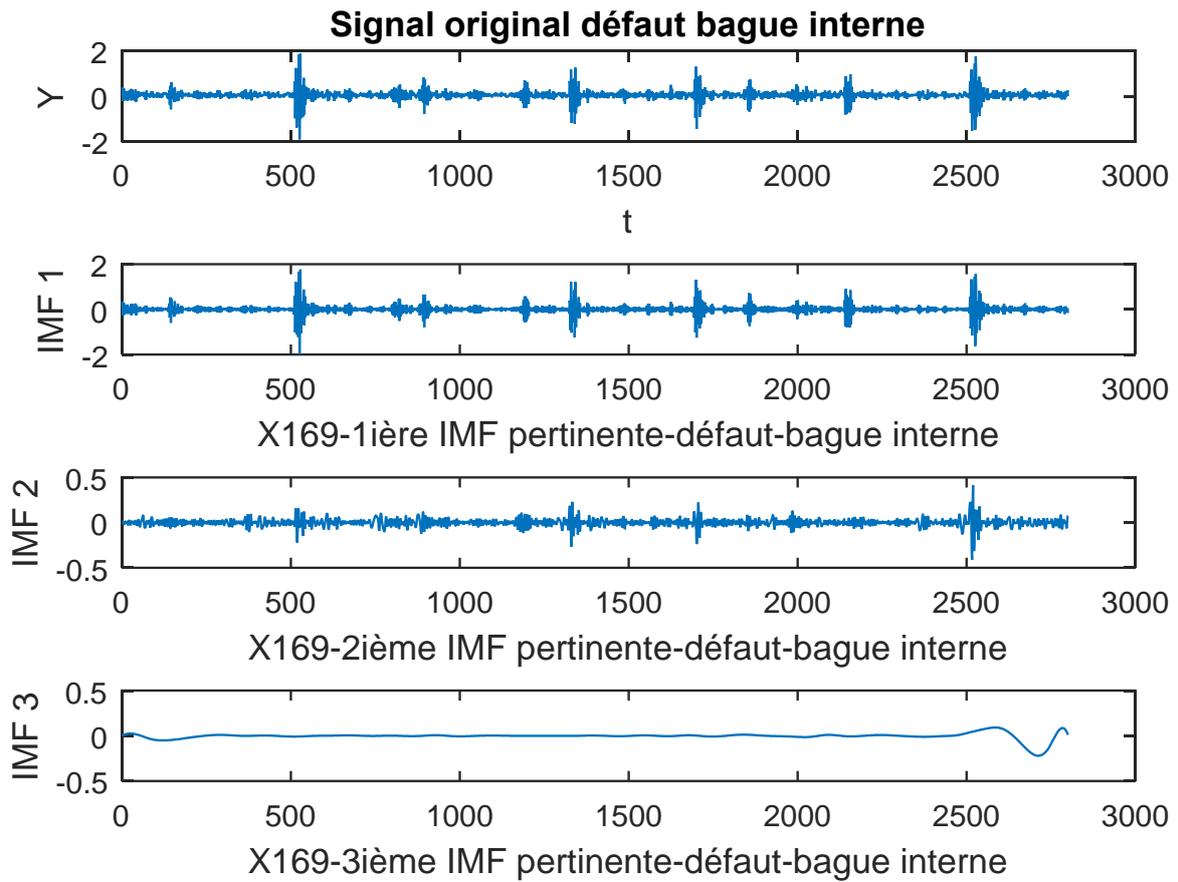


Figure 5.18 : Extraction des 3 IMfs les plus pertinentes du signal avec défaut Bague interne

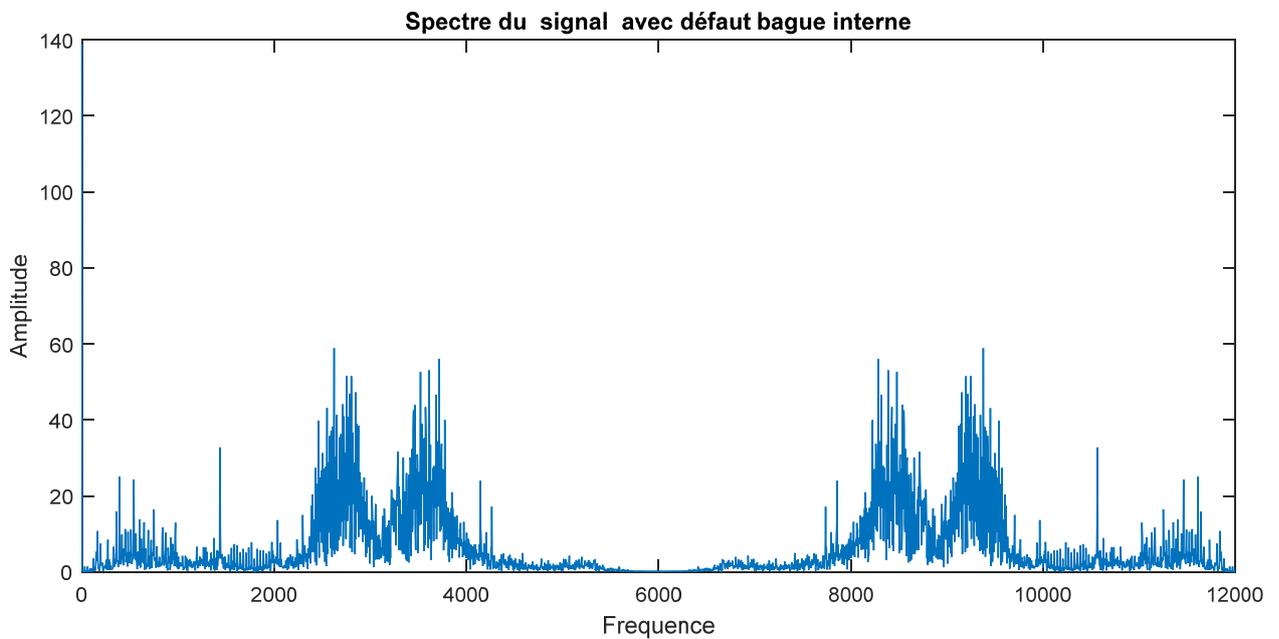
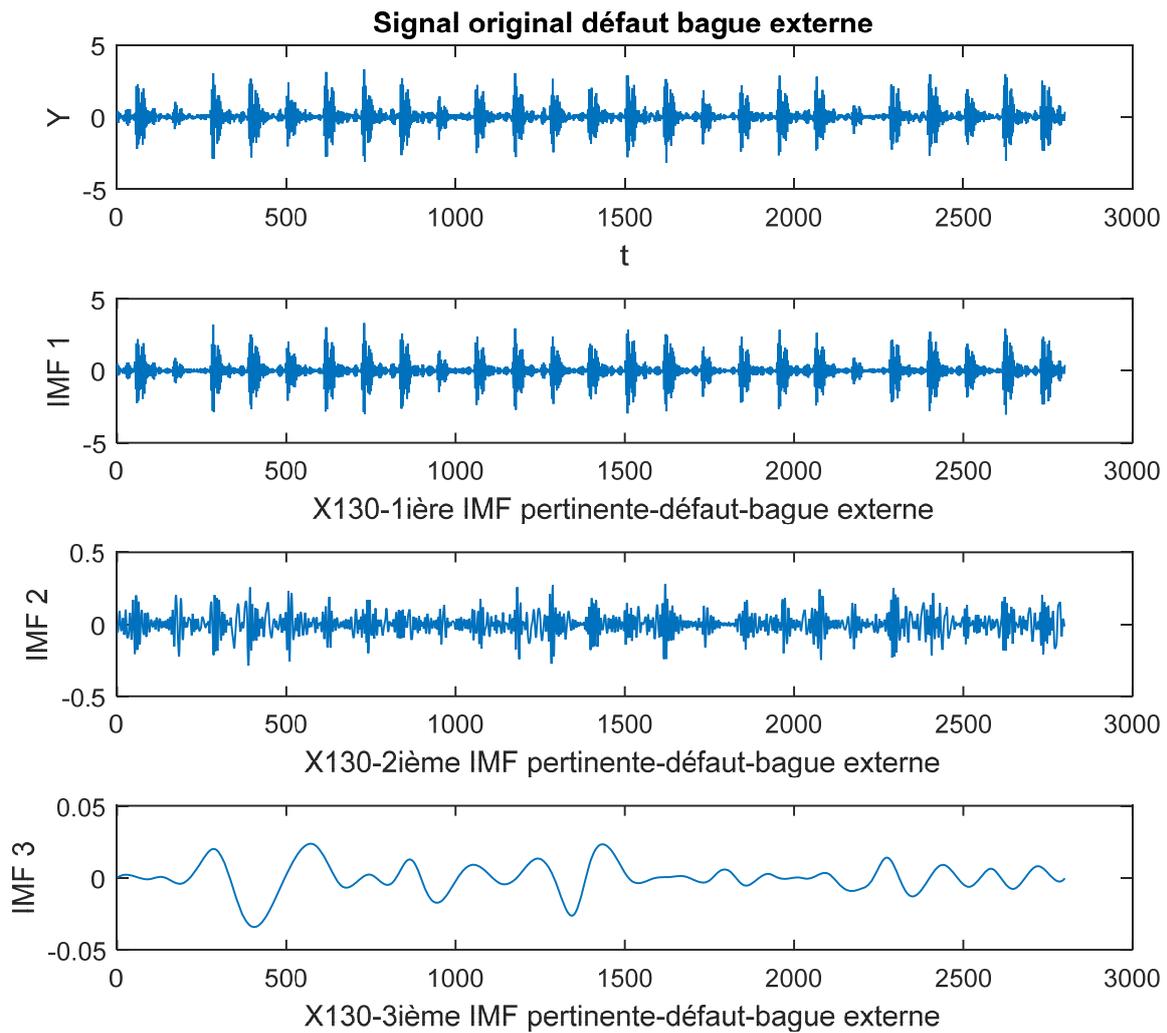
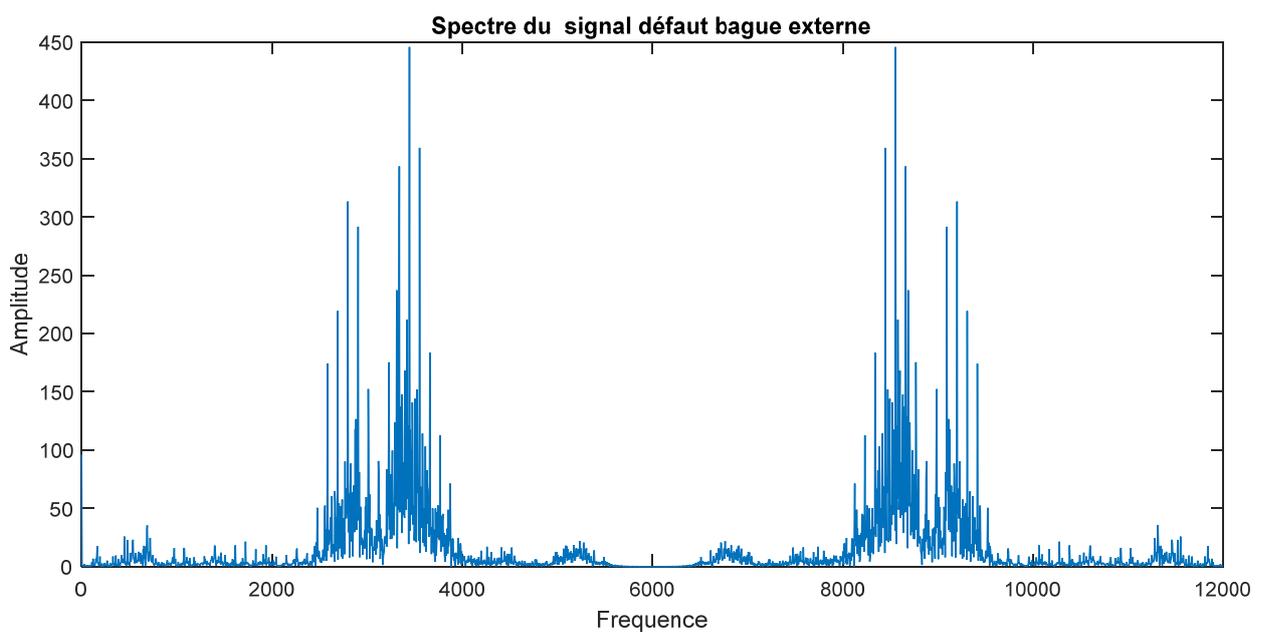


Figure 5.19 : Spectre du signal avec défaut bague interne



**Figure 5.20** : Extraction des IMFs les plus pertinentes du signal avec défaut bague externe



**Figure 5.21** : Spectre du signal avec défaut bague externe

Tableau3 : Récapitulatif des résultats

	NA Nombre d'attributs des Vecteurs	Signal d'origine SVM			Neuronal net_ work	Neuro Floue RNF	EMD / une IMF SVM			Neuronal net work	Neuro Floue RNF	EMD/trois IMFs SVM			Neuronal Net Work	Neuro Floue RNF
		Taux	temps Tot_script	temps classificat			Taux-t	Taux	Taux			Temps- tot Script	Temps classifc	Taux-t		
Fenêtrage du signal 28000	09	100/24	10	0.000	25-0.02	100	/	/	/			/	/	/		
	10	/	/	/			/	/	/			96.88/16	/	/	90/16	96.87
	11	/	/	/			100/4	118.3	0.016	25-0.015	100	93.75/4	144	0.030	75/4	81.25
	12	/	/	/			/	/	/			93.75/4	137	0.023	68.75/4	81.25
	13	/	/	/			/	/	/			93.75/4	141	0.028	75/4	75
Fenêtrage du signal 2800	09	92/24	/	/			/	/	/			/	/	/		
	10	100/12	5	0.077	25-0.02	95.83						/	/	/		
	11	/	/	/			100/4	3.882	0.047	25-0.014	100	/	/	/		
	13	/	/	/			/	/	/			/	/	/		
Fenêtrage du signal 280	09	/	/	/			/	/	/			/	/	/		
	10	87.50/24	9.38	0.016			/	/	/			/	/	/		
	11	/	/	/			50	/	/			/	/	/		
	13	/	/	/			87.5/4	1.83	0.000	56.25-0.013	62.5	/	/	/		

$$C\% = \text{Taux de bonne classification} = \frac{\text{Nombre d'échantillons bien classifiés}}{\text{Nombre d'échantillons mal classifiés}} \times 100$$

4, 12, 16,24 → nombre d'échantillons utilisé pour le test

Exemple : Nombre d'échantillons X fenêtrage = 28000 X 24 = 672000 échantillons utilisés à une fréquence de 12KHz

0.02, 0.016, 0.047 → temps de classification : Self Time - le temps CPU passé à exécuter la fonction de classification à l'exclusion des autres fonctions qu'elle appelle.

10, 5,118.3, 3.882,1.83, 144, 137,141 → Temps total - le temps CPU passé à exécuter tous le script

**Tableau 4** : Récapitulatif des résultats avec les matrices de confusions et le multi-classes SVM

	NA Nombre d'attributs Des Vecteurs	Taux de Bonne classification	Signal d'origine Matrice de confusion Avec SVM	NA Nombre d'attributs Des Vecteurs	Taux de Bonne classification	EMD / une IMF Matrice de confusion Avec SVM
Fenêtrage du signal 28000	09	100/24	$\begin{bmatrix} 24 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 \\ 0 & 0 & 24 & 0 \\ 0 & 0 & 0 & 24 \end{bmatrix}$	11	100/4	$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$
Fenêtrage du signal 2800	10	100/12	$\begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 12 \end{bmatrix}$	11	100/4	$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$
Fenêtrage du signal 280	10	87.50/24	$\begin{bmatrix} 22 & 0 & 0 & 2 \\ 0 & 20 & 3 & 1 \\ 0 & 3 & 18 & 3 \\ 0 & 0 & 0 & 24 \end{bmatrix}$	13	87.50/4	$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

## 5.2. Discussion des résultats des classifications

Nous avons appliqué la phase d'apprentissage et la phase de test à la base de données de la CWRU préalablement détaillé. Le tableau (3) montre les taux de bonne classification ainsi que les différents temps de calcul, le tableau (4) montre les matrices de confusions ainsi que les taux de bonne classification obtenu.

Le taux de classification obtenu sans utiliser la DME, c'est-à-dire, les attributs statistiques sont calculés directement sur les signaux vibratoires, est de 100%. Lorsque les attributs statistiques sont calculés sur les 3 IMFs les plus pertinentes obtenues par l'application de la DME ensuite par la classification des IMFs ayant le facteur le plus élevé de corrélation avec le signal original, le taux de classification atteint 96.88%. En sélectionnant une IMF la plus pertinente, ce taux de classification augmente pour atteindre 100%.

Le meilleur résultat obtenu avec le classificateur neuronal est de 96.87 avec 3IMFs et nombre d'échantillons deux fois plus que les résultats précédents.

Cependant les taux de classifications sont proches pour les deux méthodes SVM et neuro-floue avec une légère avance pour le SVM.

Ces résultats montrent bien l'intérêt de la procédure de sélection des IMFs et des paramètres statistiques ainsi que la nécessité de combiner avec la DME pour la rapidité de la convergence, l'efficacité du séparateur à vaste marge par rapport aux autres méthodes pour la classification des défauts roulements

## CONCLUSION GENERALE

## Conclusion Générale

En reconnaissance de formes, plusieurs méthodes ont été introduites pour le diagnostic mécanique des pannes, parmi lesquelles et les plus couramment utilisées sont les réseaux neuronaux artificiels, le séparateur à vaste marge et les Réseaux de neurones flous.

La formation des réseaux de neurones artificiels nécessite toujours un grand nombre d'échantillons, alors que les séparateurs vastes marges qui sont basés sur la théorie de l'apprentissage statistique sont spécifiques pour un nombre d'échantillons réduit et ont une meilleure généralisation que les réseaux de neurones artificiels et c'est l'utilisation de noyaux qui permet cette généralisation. Cependant la précision d'une machine vectorielle de support est hautement déterminée par la sélection des paramètres optimaux.

La plus grande différence entre le SVM et le RNA est dans la méthode d'optimisation des paramètres du modèle, le RNA utilise la retro propagation du gradient qui ne garantit pas l'obtention d'une solution optimale par contre l'implémentation du SVM trouvera les paramètres optimaux.

Cependant, ils peuvent être complémentaire avec des combinaisons SVM et RNA. Par exemple, on peut utiliser plusieurs couches RNA et avoir la dernière classification via SVM à la couche de sortie.

L'algorithme développé dans cette étude peut être étendu à plusieurs types de défauts (composants, actionneurs, capteurs, ...etc.) sur les machines jusqu'à la conception d'un système d'avertissement adapté dans lequel l'apparition des défauts est rapportée à l'utilisateur sur l'écran de commande de la machine pour un diagnostic en temps réel, ou bien sur une manette portative pour un diagnostic à la demande.

Une autre idée est d'utiliser les informations existantes sur les machines et les exploiter avec cet algorithme pour en tirer un diagnostic fiable, sans pour autant déployer un matériels supplémentaire (capteurs et interfaces de calculs).