

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Ecole Nationale Polytechnique
Département de Génie Mécanique
Laboratoire de Génie Mécanique
et de Développement

LGMD

Thèse

de Doctorat en Sciences

Spécialité : Génie Mécanique

Présenté par :
ATMANI Youcef
Magister en Ingénierie des Systèmes Mécaniques

Intitulé

**Reconnaissance de défauts dans les machines tournantes
par apprentissage machine : cas des roulements**

Devant le Jury Composé de :

<i>Président :</i>	<i>Mr. SMAILI Arezki</i>	<i>Professeur, ENP</i>
<i>Directeur de Thèse :</i>	<i>Mr. RECHAK Saïd</i>	<i>Professeur, ENP</i>
<i>Examineurs :</i>	<i>Mr. BELOUHRANI Mohamed El Amine</i>	<i>Professeur, ENST Alger</i>
	<i>Mr. BELKACEMI Yacine</i>	<i>MCA, ENP</i>
	<i>Mr. BENKEDJOUH Tarek</i>	<i>MCA, EMP</i>
	<i>Mr. BEKHOUCHA Ferhat</i>	<i>MCA, ESTA</i>
<i>Invités :</i>	<i>Mr. MESLOUB Ammar</i>	<i>MCA, EMP</i>
	<i>Mr. BOUAZIZ Mohamed</i>	<i>Professeur, ENP</i>

ENP 2020

Laboratoire de Génie Mécanique et de Développement, Ecole Nationale Polytechnique (ENP)
10, Avenue Pasteur, Hassen Badi, BP. 182, 16200 El Harrach, Alger, Algérie
www.enp.edu.dz

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Ecole Nationale Polytechnique
Département de Génie Mécanique
Laboratoire de Génie Mécanique
et de Développement

LGMD

Thèse

de Doctorat en Sciences

Spécialité : Génie Mécanique

Présenté par :
ATMANI Youcef
Magister en Ingénierie des Systèmes Mécaniques

Intitulé

**Reconnaissance de défauts dans les machines tournantes
par apprentissage machine : cas des roulements**

Devant le Jury Composé de :

<i>Président :</i>	<i>Mr. SMAILI Arezki</i>	<i>Professeur, ENP</i>
<i>Directeur de Thèse :</i>	<i>Mr. RECHAK Saïd</i>	<i>Professeur, ENP</i>
<i>Examineurs :</i>	<i>Mr. BELOUHRANI Mohamed El Amine</i>	<i>Professeur, ENST Alger</i>
	<i>Mr. BELKACEMI Yacine</i>	<i>MCA, ENP</i>
	<i>Mr. BENKEDJOUH Tarek</i>	<i>MCA, EMP</i>
	<i>Mr. BEKHOUCHA Ferhat</i>	<i>MCA, ESTA</i>
<i>Invités :</i>	<i>Mr. MESLOUB Ammar</i>	<i>MCA, EMP</i>
	<i>Mr. BOUAZIZ Mohamed</i>	<i>Professeur, ENP</i>

ENP 2020

Laboratoire de Génie Mécanique et de Développement, Ecole Nationale Polytechnique (ENP)
10, Avenue Pasteur, Hassen Badi, BP. 182, 16200 El Harrach, Alger, Algérie

www.enp.edu.dz

Dédicaces

A mes parents

A Kouki, Medi, Abdou et Chichou

A toute ma famille

A tous mes amis...



REMERCIEMENTS




*Tout d'abord, je tiens à exprimer toute ma gratitude et ma sympathie à Mr **RECHAK Said**, Professeur à l'Ecole Nationale Polytechnique d'ALGER, qu'il trouve ici l'expression de ma profonde reconnaissance, pour sa compétence, ses encouragements et sa disponibilité.*

*Pour Mr **MESLOUB Ammar**, Maître de conférences à l'Ecole Militaire Polytechnique, je lui exprime mes sincères remerciements pour son aide précieuse dans ce travail, ainsi que pour toutes ses qualités scientifiques et humaines.*

*Je tiens à remercier aussi Mr **SMAILI Arezki**, Professeur à l'Ecole Nationale Polytechnique, pour m'avoir fait l'honneur d'accepter la présidence du jury.*

*Mes remerciements vont aussi à Messieurs, **BELOUCHRANI Mohamed El Amine**, Professeur à l'Ecole Nationale Supérieure de Technologie, **BELKACEMI Yacine**, Maître de conférences à l'Ecole Nationale Polytechnique, **BENKEDJOUH Tarek**, Maître de conférences à l'Ecole Militaire Polytechnique et **BEKHOUCHA Ferhat**, Maître de conférences à L'Ecole Supérieure des Techniques de l'Aéronautique pour avoir accepté d'examiner ce travail.*

*Mes remerciements vont également à Mr **BOUAZIZ Mohamed**, Professeur à l'Ecole Nationale Polytechnique, à tous mes amis et collègues.*



ملخص:

يبحث العمل المقدم في التعرف على أخطاء المُدخَرَجَة استناداً إلى نماذج الخلط الغوسي (GMM) لطرق التعلم الآلي. الميزات المستخدمة هي معاملات التنبؤ الخطي (LPC) من ناحية، ومعاملات سيسترال تردد ميل (MFCC) من ناحية أخرى. خلال هذه الدراسة، تم تحديد معيار واحد فقط لتقييم أداء عملية التصنيف. هذا هو متوسط معدل التصنيف (ACR) الذي تم الحصول عليه من مصفوفة الخلط. في كل اختبار يتم إجراؤه، يتم اعتبار ناقلات الميزات المولدة على طول للتمييز بين أربع حالات خطأ مثل المحامل العادية والمحمل ذات أخطاء السباق الداخلية والخارجية وأخطاء الكرة. تم اختبار العديد من التكوينات من أجل تحديد القيم المثلى لمعاملات الإدخال (حجم إطار التحليل وترتيب النموذج وغيرها). يعتمد التطبيق التجريبي على إشارات الاهتزاز من مركز البيانات الحاملة للجامعة الأمريكية والتي تسمى CWRU (جامعة كيس ويسترن ريزيرف). في ظل ظروف معينة، تظهر النتائج التي تم الحصول عليها في هذه الدراسة أداء تصنيفات عالية، في توافق جيد مع تلك من الأدب.

الكلمات المفتاحية: عيوب المُدخَرَجَة، نماذج المخلوط الغوسي، معاملات سيسترال لتردد ميل، استخلاص الميزات، التشخيص.

Abstract:

The presented work investigates the bearing faults recognition based on the Gaussian Mixing Models (GMM) of machine learning methods. The used features are the Linear Prediction Coefficients (LPC) on the one hand and the Mel Frequency Cepstral Coefficients (MFCC) on the other hand. Throughout this study, only one criterion is defined for the evaluation of the performance of the classification process. This is the Average Classification Rate (ACR) obtained from the confusion matrix. In every test performed, the generated features vectors are considered along to discriminate between four fault conditions as normal bearings, bearings with inner and outer race faults and ball faults. Many configurations were tested in order to determinate the optimal values of input parameters (size of analysis frame, order of the model and others). The experimental application is based on vibration signals from the bearing data center of the American University so called CWRU (Case Western Reserve University). Under certain conditions, the results obtained in this study show high classification performances, in good agreement with those from the literature.

Keywords: Bearing faults, Gaussian Mixture Models, Mel Frequency Cepstral Coefficients, Linear Prediction Coefficients, Feature extraction.

Résumé :

Le travail présenté étudie la reconnaissance des défauts de roulement basé sur une méthode d'apprentissage machine (Machine Learning), en l'occurrence les Modèles de Mélange Gaussien (GMM). Les descripteurs utilisés sont les coefficients de prédiction linéaire (LPC) d'une part, et les coefficients cepstraux des fréquences de Mel (MFCC), d'autre part. Tout au long de cette étude, un seul critère est défini pour l'évaluation des performances du processus de classification. Il s'agit du taux de classification moyen (ACR) obtenu à partir de la matrice de confusion. Dans chaque test effectué, les vecteurs d'attributs générés sont pris en compte pour distinguer quatre conditions de défaut à savoir, des roulements normaux, des roulements avec des défauts de course internes et externes et des défauts de bille. De nombreuses configurations ont été testées afin de déterminer les valeurs optimales des paramètres d'entrée (taille de trame d'analyse, ordre du modèle, etc.). L'application expérimentale est basée sur des signaux de vibration en provenance du centre de données de roulement de l'université américaine CWRU (Case Western Reserve University). Dans certaines conditions, les résultats obtenus dans cette étude affichent des performances de classification élevées, en bonne concordance avec ceux issus de la littérature.

Mots clés : Défauts de roulement, Modèles de mélange gaussiens, Coefficients cepstraux de fréquence de Mel, Coefficients de prédiction linéaire, Extraction d'attributs.

SOMMAIRE

LISTE DES NOTATIONS ET ABREVIATIONS	
LISTE DES FIGURES ET TABLEAUX	
INTRODUCTION GENERALE	17

Chapitre 1 :

GENERALITES SUR LES DEFAILLANCES USUELLES DES ORGANES DE MACHINES TOURNANTES - CAS DES ROULEMENTS

1.1 INTRODUCTION	23
1.2 POLITIQUE ET STRATEGIE DE MAINTENANCE	23
1.2.1 Concepts et méthodes de maintenance	24
1.2.2 Maintenance Préventive Conditionnelle	26
1.3 MAINTENANCE CONDITIONNELLE PAR ANALYSE VIBRATOIRE	29
1.3.1 Approche sélective de l'analyse vibratoire	29
1.3.2 Systèmes de mesure des vibrations	30
1.4 TYPOLOGIE DES PRINCIPAUX DEFAUTS MECANIQUES	32
1.4.1 Défauts de déséquilibre (balourd)	32
1.4.2 Défauts d'alignement	34
1.4.3 Défauts des engrenages	35
1.4.4 Défauts de transmission par courroies	41
1.4.5 Défauts sur les paliers lisses	43
1.4.6 Défauts induits par un desserrage	44
1.4.7 Défaut de cavitation	44
1.5 DEFAUTS DES ROULEMENTS	45
1.5.1 Description et caractéristiques du roulement à billes	45
1.5.2 Typologie de défaillances dans les roulements	46
1.5.3 Principales causes des défauts de roulements	50
1.5.4 Fréquences caractéristiques des défauts de roulement	51
1.5.5 Signatures et images spectrales théoriques de défauts de roulement	53
1.6 CONCLUSION	60

Chapitre 2 :

METHODES DE DIAGNOSTIC DES DEFAUTS DE ROULEMENTS – ETAT DE L'ART

2.1 INTRODUCTION	62
2.2 METHODES D'ANALYSE TEMPORELLES	62
2.2.1 Indicateurs statistiques	62
2.2.2 Moyennage temporel synchrone	63

2.2.3 Modèles paramétriques	64
2.3 METHODES D'ANALYSE FREQUENTIELLES	66
2.3.1 Analyse de Fourier	66
2.3.2 Analyse d'enveloppe	67
2.3.3 Analyse cepstrale	69
2.3.4 Méthodes de haute résolution HRM	70
2.4 METHODES D'ANALYSE TEMPS- FREQUENCE	71
2.4.1 Transformée de Fourier à court terme STFT	72
2.4.2 Distribution de Wigner-Ville WVD	73
2.4.3 Distribution de Choi-Williams	73
2.5 KURTOSIS SPECTRAL ET KURTOGRAMME	74
2.6 METHODES D'ANALYSE TEMPS- ECHELLE	74
2.6.1 Transformée en ondelettes continues	75
2.6.2 Transformée en ondelettes discrètes	75
2.6.3 Paquets d'ondelettes	75
2.7 ANALYSE PAR LES METHODES DE DECOMPOSITION	76
2.7.1 Décomposition en mode empirique EMD	76
2.7.2 Décomposition en mode empirique d'ensemble EEMD	77
2.7.3 Décomposition en modes variationnels VMD	78
2.8 ANALYSE DE LA CYCLOSTATIONNARITE	80
2.9 METHODES BASEES SUR L'APPRENTISSAGE MACHINE (ML)	80
2.9.1 Réseaux de neurones artificiels ANN	81
2.9.2 Séparateurs à vaste marge SVM	84
2.9.3 K plus proches voisins KNN	89
2.9.4 Arbres de décision	90
2.9.5 Classifieur bayésien naïf NB	91
2.9.6 Classifieur neuro floue	92
2.10 METHODES BASEES SUR L'APPRENTISSAGE PROFOND (DL)	92
2.11 SYNTHESE	94
2.12 CONCLUSION	95

Chapitre 3 :

RECONNAISSANCE DES DEFAUTS DE ROULEMENTS PAR LES MODELES DE MELANGE GAUSSIENS GMMs

3.1 INTRODUCTION	98
3.2 CLASSIFICATION BASEE SUR LES MODELES PROBABILISTES	98
3.2.1 Formalisme des notions de classe et de règle de décision	98
3.2.2 Règle de classification optimale au sens probabiliste	99
3.2.3 Modélisation par mélange de lois	100
3.3 MODELE DE MELANGE GAUSSIEN ET ALGORITHME EM	101
3.3.1 Distribution gaussienne	101
3.3.2 Modèle de mélange gaussien	103
3.3.3 Maximum de vraisemblance	105
3.3.4 Algorithme Espérance Maximisation – EM	106
3.4 CLASSIFICATION DES DEFAUTS DE ROULEMENT PAR LES GMM	110

3.4.1	Concept de diagnostic par reconnaissance de défauts	110
3.4.2	Schéma global de la classification par les GMM	111
3.5	MODELES ET ORIGINES DES DONNEES	113
3.5.1	Modèles analytiques de roulement	113
3.5.2	Modèle dynamique du roulement à billes	116
3.5.3	Modèle d'excitation générée par des défauts localisés	118
3.5.4	Modèle de signaux générés par la présence de défauts localisés	119
3.5.5	Prétraitements appliqués aux données	121
3.5.6	Organisation et structuration des données	122
3.6	EXTRACTION DES INDICATEURS DISCRIMINANTS	122
3.6.1	Approche basée sur les LPCs	122
3.6.2	Approche basée sur les MFCCs	126
3.7	CONSTRUCTION DE MODELES	128
3.8	TESTS DE PERFORMANCE	131
3.9	CONCLUSION	132

Chapitre 4 :

EXPERIMENTATIONS ET RESULTATS OBTENUS

4.1	INTRODUCTION	134
4.2	4.2. EXPERIMENTATION	134
4.2.1	Description des moyens expérimentaux	135
4.2.2	Base de signaux de vibration sélectionnés	136
4.2.3	Mise en œuvre de la procédure de classification	137
4.3	RESULTATS DE LA CLASSIFICATION BASEE SUR LES LPCs	142
4.3.1	Erreur estimée et ordre de prédiction	143
4.3.2	Choix du nombre de LPC	145
4.3.3	Effet de la longueur de la trame d'analyse	146
4.3.4	Effet de la longueur de la séquence de test	146
4.3.5	Effet du nombre de composantes du GMM	148
4.4	RESULTATS DE LA CLASSIFICATION BASEE SUR LES MFCCs	149
4.4.1	Effet du nombre de composantes du GMM	150
4.4.2	Effet du nombre de MFCC	151
4.4.3	Effet de la longueur de la trame d'analyse	153
4.4.4	Effet de la longueur de la séquence de test	154
4.5	DISCUSSIONS DES RESULTATS	156
4.5.1	Synthèse des performances de la classification avec les LPCs	156
4.5.2	Synthèse des performances de la classification avec les MFCCs	157
4.5.3	Evaluation des approches utilisées	160
4.6	CONCLUSION	161
	CONCLUSION GENERALE	162
	REFERENCES BIBLIOGRAPHIQUES	165
	ANNEXE	178

LISTE DES FIGURES

Figure 1.1 : Statistiques des méthodes de surveillance employées en maintenance [2].	29
Figure 1.2 : Schéma d'une chaîne d'acquisition multivoies de mesure des vibrations .	31
Figure 1.3 : Balourds du rotor rigide. (a) Statique ; (b) De couple ; (c) Dynamique [22]	33
Figure 1.4 : Evolution du spectre du balourd entre deux états (initial et évolué)	34
Figure 1.5 : Types de désalignements dans un accouplement	34
Figure 1.6 : Spectre vibratoire théorique d'un défaut d'alignement	35
Figure 1.7 : Groupes de transmission par engrenages	36
Figure 1.8 : Défauts généralisés sur différentes dentures	37
Figure 1.9 : Défauts localisés sur différentes dentures	37
Figure 1.10 : Schéma simplifié de l'engrènement	38
Figure 1.11 : Spectre de l'engrenage à l'état normal	38
Figure 1.12 : Spectre du défaut d'engrenage ayant une dent détériorée	39
Figure 1.13 : Spectre du défaut d'engrenage ayant l'ensemble de dentures détériorées	39
Figure 1.14 : Spectre du défaut d'engrenage ayant l'entraxe insuffisant	40
Figure 1.15 : Spectre du défaut d'engrenage ayant un jeu de fond trop grand	40
Figure 1.16 : Spectre de l'engrènement modulé par la fréquence de rotation	41
Figure 1.17 : Schéma d'un défaut localisé de courroie	42
Figure 1.18 : Image vibratoire théorique d'un défaut de transmission par courroie ...	43
Figure 1.19 : Éléments constitutifs du roulement à bille [26]	45
Figure 1.20 : Illustrations de deux cas d'écaillage	46
Figure 1.21 : Grippage de roulement à billes [30].....	47
Figure 1.22 : Eraflures de cage d'un roulement à billes à gorge profonde [30].....	47
Figure 1.23 : Usure de contact sur toute la circonférence de la surface de l'alésage [30]	48
Figure 1.24 : Ruptures dans des éléments de roulement [31]	48
Figure 1.25 : Défaut de corrosion sur les bagues de roulements [30]	49

Figure 1.26 : Les causes les plus fréquentes de défaillance des roulements [30]	51
Figure 1.27 : Caractéristiques géométriques d'un roulement à billes	52
Figure 1.28 : Caractéristiques spectrales des différents stades du défaut de roulement [35]	55
Figure 1.29 : Image vibratoire d'un défaut ponctuel de la bague fixe d'un roulement [3]	56
Figure 1.30 : Image spectrale d'un défaut ponctuel de la bague tournante ou d'un élément roulant dans le cas d'une charge axiale pure [3]	56
Figure 1.31 : Image spectrale d'un défaut ponctuel de la bague tournante ou un élément roulant dans le cas d'une charge radiale pure [3]	57
Figure 1.32 : Signature de défauts de déversement de bagues externe et/ou interne [3]	58
Figure 1.33 : Spectre d'une altération de bague de roulement par électroérosion [3] ...	59
Figure 2.1 : Etapes de l'analyse d'enveloppe basée sur la transformée de Hilbert [39]	68
Figure 2.2 : Structure d'un réseau MLP	82
Figure 2.3 : Représentation graphique d'un neurone	83
Figure 3.1 : Schéma de classification des deux approches proposées	112
Figure 3.2 : Déplacements estimés à partir du modèle analytique	113
Figure 3.3 : Repères de base du modèle de calcul	114
Figure 3.4 : Modèle dynamique simplifié de roulement à 3 degrés de liberté, [144] ...	116
Figure 3.5 : Réponse temporelle d'un roulement endommagé du type SKF1210, à $v=11.6$ Hz. (a) défaut localisé (1mm) sur la BE ; (b) défaut localisé sur la BI [144]	121
Figure 3.6 : Schéma de la procédure de classification avec les LPCs	123
Figure 3.7 : Organigramme schématisant les étapes d'obtention des LPCs	124
Figure 3.8 : Schéma de la procédure de classification avec les MFCCs	126
Figure 3.9 : Organigramme schématisant les étapes d'obtention des MFCCs	127
Figure 3.10 : Fenêtrage appliqué pour minimiser les discontinuités aux limites des trames	127
Figure 3.11 : Extraction des attributs à partir des trames d'analyse et de test	129
Figure 4.1 : Banc d'essai de vibration des roulements du CWRU [152]	135
Figure 4.2 : Signaux relatifs aux quatre classes considérées après prétraitements	138
Figure 4.3 : Valeurs des LPC pour quatre trames représentatives des conditions de défaut	139

Figure 4.4 : Représentation graphique de la base d'apprentissage des LPC (chaque condition de défaut correspondant à 1792 trames de 1024 échantillons) ...	140
Figure 4.5 : Valeurs de MFCC correspondant aux différentes conditions de défauts ..	141
Figure 4.6 : Représentation graphique de la base d'apprentissage des MFCC (chaque condition de défaut correspondant à 896 trames de 2048 échantillons) ...	142
Figure 4.7 : Comparaison des signaux original et estimé du modèle de prédiction	143
Figure 4.8 : Courbes d'erreur normalisée relatives aux défauts de roulement	144
Figure 4.9 : Variation du taux de l'ACR en fonction du nombre de LPCs	145
Figure 4.10 : Variation du taux de l'ACR en fonction de la longueur de la trame d'analyse	146
Figure 4.11 : Variation du taux de l'ACR en fonction de la longueur de la séquence de test	147
Figure 4.12 : Variation du taux de l'ACR en fonction du nombre de composantes du mélange	149
Figure 4.13 : Evolution de l'ACR en fonction du nombre de composantes GMM (2048 pts)	150
Figure 4.14 : Evolution de l'ACR en fonction du nombre de composantes GMM (1024 pts)	151
Figure 4.15 : Evolution de l'ACR en fonction du nombre de MFCCs ($N_{GMMs} = 8$) ...	152
Figure 4.16 : Variation du taux ACR en fonction de la longueur de la trame d'analyse	154
Figure 4.17 : Variation du taux ACR en fonction de la longueur de la séquence de test	154
Figure 4.18 : Nuage de points 2D pour les LPCs relatifs à chaque conditions de défaut	157
Figure 4.19 : Nuage de points 3D pour les LPCs relatifs aux conditions de défaut	157
Figure 4.20 : Nuage de points 2D pour les MFCCs relatifs aux conditions de défaut ..	159
Figure 4.21 : Nuage de points 3D pour les MFCCs relatifs aux conditions de défaut ..	159

LISTE DES TABLEAUX

Tableau 1.1 : Tableau comparatif des méthodes de surveillance en maintenance [1] ..	28
Tableau 3.1 : Modèle de matrice de confusion d'un classifieur	131
Tableau 4.1 : Caractéristiques dimensionnelles et cinématiques des roulements utilisés	136
Tableau 4.2 : Nomenclature de fichiers des défauts sélectionnés (roulement DE) à 48KHz	136
Tableau 4.3 : Partition des données entre l'apprentissage et le test selon la taille de la trame	137
Tableau 4.4 : Influence de l'ordre de prédiction sur l'erreur normalisée	145
Tableau 4.5 : Matrices de confusion : NLPC =14 ; NGMM = 2 ; Trame d'analyse = 1024 pts	147
Tableau 4.6 : Matrice de confusion dans le cas : $N_{MFCC} = 4$	153
Tableau 4.7 : Matrices de confusion : $N_{MFCC} = 4$; $N_{GMM} = 8$; Trame d'analyse=2048 points	155
Tableau 4.8 : Matrice de confusion : $N_{LPC}=14$; $N_{GMM}=2$; longueur de trame =1024 points	156
Tableau 4.9 : Matrice de confusion : $N_{MFCC}=4$; $N_{GMM}=8$; longueur de trame =2048 points	158
Tableau 4.10 : Résumé des résultats des deux approches de classification utilisées	160

LISTE DES ABREVIATIONS

ACR : Average Classification Rate

AE : Auto Encoder

AFSA : Artificial Fish Swarm Algorithm

AIC : Akaike Information Criterion

ANN : Artificial Neural Networks

AR : Auto Regressive

ARMA : Auto Regressive Moving Average

BIC : Bayesian Information Criterion

BPFI : Ball Pass Frequency Inner race

BPFO : Ball Pass Frequency Outer race

BSF : Ball Spin Frequency

CNN : Convolutional Neural Networks

CPU : Central Processing Unit

DBN : Deep Belief Network

DE : Drive End

DL : Deep Learning

DSP : Densité Spectrale de Puissance

DWT : Discrete Wavelet Transform

EDM : Electrical Discharge Machining

EEMD : Ensemble Empirical Mode Decomposition

ELM : Extreme Learning Machine

EM : Expectation Maximization

EMD : Empirical Mode Decomposition

ESPRIT : Estimation of Signal Parameters via Rotational Invariance Techniques

FE : Fan End

FFT : Fast Fourier Transform

FIR : Finite Impulse Response
FTF : Fundamental Train Frequency
GA : Genetic Algorithms
GAN : Generative Adversarial Network
GMM : Gaussian Mixture Models
GPU : Graphics Processing Unit
GRU : Gated Recurrent Unit
HFRT: High Frequency Resonance Technique
HMM : Hidden Markov Models
HRM : High Resolution Methods
IMF : Intrinsic Mode Function
KNN : K Nearest Neighbours
LMD : Local Mean Decomposition
LP : Linear Prediction
LPC : Linear Predictive Coding
LPP : Locality Preservation Projection
LSTM : Long Short Term Memory
MA : Moving Average
MDL : Minimum Description Length
MFCC : Mel Frequency Cepstral Coefficient
MFD : Multi Fractal Dimension
MHM : Machine Health Monitoring
ML : Machine Learning
MUSIC : Multiple Signal Classification
MVC : Minimal Variance Cepstrum
PCA : Principal Component Analysis
PDF : Probability Distribution Function
PNN : Probabilistic Neural Networks
PPCM : Plus Petit Commun Multiple
PSO : Particle Swarms Optimization
RBF : Radial Basis Function
RMS : Root Mean Square
RNN : Recurrent Neural Networks

SK : Spectral Kurtosis

SNR : Signal Noise Rate

STFT : Short Time Frequency Transform

SVM : Support Vector Machine

WKNN : Weighted K Nearest Neighbours

WPD : Wavelet Packet Decomposition

WPT : Wavelet Packet Transform

WVD : Wigner-Ville Distribution

LISTE DES NOTATIONS

F_i : Fréquences de rotation
 F_e : Fréquence d'engrènement
 Z_i : Nombre de dents
 N_i : Vitesses de rotation
 F_{co} : Fréquence de coïncidence
 F_c : Fréquence de courroie
 \emptyset : Angle de contact du roulement
 d_e : Distance euclidienne
 α_i : Angle de la bille avec l'axe de référence
 δ_i : Déplacements partiels des bagues
 σ : Ecart type
 M_k : Moment statistique d'ordre supérieur
 T_q : Période
 A_0 : Amplitude
 A_q^k : Amplitude k^{ème} harmonique
 \emptyset_q^k : Phase initiale k^{ème} harmonique
 α_{bs}^j : Facteur d'amortissement du j^{ème} mode
 N_q : Nombre d'harmoniques
 p : Ordre de prédiction
 $\hat{s}(n)$: Echantillon estimé (prédit)
 $s(n)$: Séquence de signal
 $w(n)$: fenêtre
 a_k : Coefficients de prédiction linéaire LPC
 $e(n)$: Erreur de prédiction linéaire
 E : Erreur de prédiction totale
 $R(i)$: Séquence d'autocorrélation

E_p : Erreur minimale totale pour un ordre p

V_p : Erreur normalisée

α : paramètre de contrôle du filtrage

M : Nombre de filtres triangulaires

$c[q]$: Coefficients MFCC

θ : Paramètre du modèle de mélange gaussien

\mathbf{w} : Vecteur des poids w_i

$\boldsymbol{\mu}$: Vecteur des moyennes μ_i

$\boldsymbol{\Sigma}$: Matrice des covariances Σ_i

$p(\mathbf{x}|\theta)$: Densité du mélange gaussien

$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$: Distribution gaussienne multivariée

C : Nombre de classes

P_{cc} : Probabilité de classification correcte

X_i : $i^{\text{ème}}$ classe

P_{cc}^i : Probabilité de classification correcte de X_i

P_{ij} : Probabilité de classification incorrecte de X_i

d_{ij} : Nombre de réalisations ayant la classe X_j en sortie sachant que la classe X_i était en entrée.

M_i : Nombre total des tests sur la classe X_i

\emptyset_i : Diamètre intérieur du roulement

\emptyset_e : Diamètre extérieur du roulement

\emptyset_b : Diamètre de la bille du roulement

\emptyset_m : Diamètre moyen

INTRODUCTION GENERALE

INTRODUCTION GENERALE

L'analyse vibratoire est la méthode privilégiée de diagnostic de défauts des machines tournantes, dans le cadre de la maintenance conditionnelle. Le développement des ressources allouées à l'analyse des défaillances de ces systèmes, a généré des améliorations dans les techniques et méthodologies utilisées [1]. Les roulements se situent aux premiers plans des domaines d'application de ces techniques. En jouant à la fois les rôles de support, de guidage et de transmission de puissance, les roulements sont des composants essentiels dans les machines tournantes. De ce fait, ils représentent des sièges de contraintes sévères résultant des applications de charges élevées dans les directions, radiale et axiale. Des recherches statistiques ont montré que 30% à 40% des défaillances des machines tournantes sont causées par des dommages aux roulements [2]. Il est donc difficile de garantir les diagnostics sur l'état du roulement. Ceci explique le grand intérêt marqué par la recherche scientifique à cet organe sensible, en s'intéressant sans cesse au développement de nouvelles méthodes de diagnostic, depuis de nombreuses décennies à ce jour.

Dans cette voie, la surveillance des défaillances de roulements et leur prédiction sont mises en œuvre à partir de données de sources multiples. Les causes de ces détériorations peuvent être : une conception incorrecte, une mauvaise installation, la corrosion, une mauvaise lubrification, la fatigue, l'usure ou la déformation plastique. Les techniques traditionnelles utilisées sont basées sur le traitement du signal et issues des domaines, temporel, fréquentiel, temps-fréquence, temps-échelle, entre autres [3]. Avec, l'avenue des méthodes issues de l'intelligence artificielle, le modèle de signature de défaut peut être considéré comme un problème de classification. Les procédures de diagnostic sont automatisées et élaborées avec différentes approches émergentes basées que soit sur l'apprentissage machine (Machine Learning - ML -) ou l'apprentissage profond (Deep Learning – DL -).

Dans la première catégorie, la majorité des techniques développées sont anciennes et la plupart d'entre-elles ont été intégrées au milieu industriel. McFadden et Smith [4], ont établi les règles d'utilisation de la méthode de l'enveloppe, appelée aussi la méthode de résonance hautes fréquences (HFRT). Dron et al. [5], ont montré l'intérêt de la soustraction spectrale dans l'amélioration de la sensibilité des indicateurs scalaires (facteur crête, kurtosis, etc.). Randall et Antoni [6], ont étudié l'efficacité des techniques conventionnelles de traitement du signal par le diagnostic de signaux vibratoires (accélérométriques) provenant de roulements.

Dans [7], Yu a proposé une méthode temps-fréquence à base de STFT, nommée la transformée des transitoires extraites. Selon l'auteur, elle peut efficacement caractériser et extraire les composantes transitoires à partir des signaux des défauts de roulement. Quinde et al. [8], ont proposé une méthodologie basée sur la décomposition moyenne locale (LMD) et la WVD pour obtenir un diagnostic de défaut de roulement plus fiable basé sur des signaux de vibration. Li et al. [9], ont présenté l'association de la WVD avec la méthode de décomposition en mode empirique (EMD), pour le diagnostic des défauts de roulement. Bella et al. [10], ont appliquées une multitude de techniques temps-fréquence pour la détection de défauts de roulements issus de l'industrie.

Dans la deuxième catégorie, on retrouve l'ensemble des approches issues de l'intelligence artificielle regroupées sous le nom de Machine Learning. Purushotham et al. [11], ont présenté une méthode de détection des défauts de palier localisés basée sur la transformée en ondelettes discrètes (DWT) et les modèles de Markov cachés (HMM). Tian et al. [12], ont présenté une méthode basée sur le Kurtosis spectral, pour la détection des défauts et la surveillance de la dégradation des roulements dans les moteurs électriques. Benkedjouh et al. [13], ont étudié une méthode de surveillance de l'état des machines basée sur les descripteurs MFCC et la machine à vecteurs de support (SVM). Pandya et al. [14], ont présenté une étude de diagnostic automatisée des défauts de roulements à billes ayant des défauts localisés. Yu et al. [15], ont proposé une approche simple de diagnostic des défauts de machine basée sur le modèle de mélange gaussien (GMM) et la transformée en ondelette. Yu [16], a proposé une approche basée sur l'extraction des attributs appelés les projections de préservation des localités (LPP) avec les modèles de mélange gaussiens (GMM). Nelwamondo et al. [17], ont utilisé des caractéristiques linéaires et non linéaires extraites des signaux de vibration du domaine temporel d'une machine tournante avec des roulements normaux et défectueux. Des indicateurs à base de dimensions fractales multi-échelles (MFD), de MFCCs et de kurtosis, sont utilisés comme attributs pour classer les défauts de roulement à l'aide de GMM et de HMM.

Récemment, une sous-catégorie de méthodes regroupées sous le nom de Deep Learning DL, sont développées en tant que sous-domaine du Machine Learning ML. À ce jour, il existe une variété d'architectures d'apprentissage en profondeur appliquées à la surveillance de l'état des machines (Machine Health Monitoring - MHM). Ainsi, Guo et al. [18], ont proposé un algorithme CNN adaptatif (ADCNN) ayant pour objectif de modifier dynamiquement le rythme d'apprentissage pour obtenir un meilleur compromis entre vitesse d'entraînement et précision. Zhuang et Wei. [19], ont adopté un algorithme CNN à multi-échelles nommé MS-

DCNN. Dans [20], Mao et al., ont utilisé une approche basée sur l'apprentissage machine extrême (ELM) et un auto encodeur (AE). Enfin, Zhao et Shao [21], ont proposé dans une étude récente, une méthode appelée Deep Gated Recurrent Unit (DGRU), pour le diagnostic des signaux de vibration des roulements de locomotive.

L'objectif principal de ce travail est de proposer une méthode de reconnaissance de défauts de roulements basée sur l'apprentissage machine (Machine Learning). Le choix a été porté sur les modèles de mélange gaussiens GMM, principalement pour deux raisons. La première, la rareté de leur usage dans le diagnostic des défauts mécaniques, malgré leur succès fulgurant dans les domaines de traitement de la voix, de l'astronomie, de la biologie, de l'économie, des sciences de l'ingénieur, du marketing, de la reconnaissance d'images, et beaucoup d'autres. La seconde, est leur simplicité d'application et leur flexibilité d'adaptation à la diversité des données d'analyse, ainsi qu'à la multitude des types et du nombre de descripteurs (variables) utilisés dans les procédures de classification.

La contribution de notre étude est basée sur la discrimination par la classification entre un groupe de classes représentatives des conditions de défaut de roulement, moyennant les mélanges gaussiens, tout au long d'un processus supervisé. Une partie de cette recherche est consacrée à l'optimisation des paramètres d'une double classification basée sur l'extraction d'attributs cepstraux à l'échelle de Mel et d'autres issus de la prédiction linéaire, à partir de signaux accélérométriques de vibrations.

Le premier chapitre présente la problématique du diagnostic des défaillances par une analyse vibratoire, dans un contexte global de maintenance conditionnelle. Ainsi, un aperçu général sur les défauts usuels dans les machines tournantes, leurs origines et leurs signatures vibratoires typiques, a été donné. Dans la suite de l'étude, les travaux se sont focalisés sur l'objectif principal de cette recherche à savoir, le diagnostic des défauts dans les roulements. En conséquence, une description détaillée de leurs modes de défaillances, de leurs causes et de leurs signatures caractéristiques est fournie.

Le second chapitre concerne l'état de l'art de la recherche sur les méthodes existantes dans la thématique de cette étude, afin de positionner les travaux de cette thèse par rapport à ceux déjà réalisés. Il est divisé en deux parties essentielles : la première est liée à une présentation des méthodes dites classiques basées sur les traitements de signaux, la seconde partie est réservée aux approches issues de l'intelligence artificielle.

Le troisième chapitre est relatif à la modélisation mathématique du problème de reconnaissance de défauts de roulement par une approche de classification basée sur une méthode d'apprentissage machine (Machine Learning). Notre choix a été porté sur les GMMs,

et a été largement argumenté. Ainsi, un schéma de classification a été établi, et des codes de calculs ont été élaborés, relatifs aux algorithmes développés dans chacune des étapes de ce dernier.

Dans le dernier chapitre, une validation expérimentale des codes de calcul numérique, développés dans le cadre de la démarche entamée, par l'usage d'une base de données dédiée au développement de nouveaux algorithmes liés à la problématique du diagnostic des défauts de roulements. L'influence des effets des paramètres d'entrée de la classification sur ses performances, basées sur les attributs de types LPCs, ou ceux de types MFCCs, a été largement détaillée et discutée.

Enfin, une conclusion générale est présentée pour rappeler l'essentiel des résultats obtenus ainsi que des perspectives à envisager pour la continuité du présent travail.

Chapitre 1

***GENERALITES SUR LES DEFAILLANCES
USUELLES DES ORGANES DE MACHINES
TOURNANTES - CAS DES ROULEMENTS***

1.1 INTRODUCTION

Dans le monde moderne, où la compétitivité est de plus en plus accrue et la concurrence plus féroce, la disponibilité des machines de production devient parmi les premiers facteurs de survie des systèmes de production. Ainsi, l'établissement d'une politique de maintenance efficace devient une action de premier plan, des systèmes de gestion. C'est pourquoi, la maintenance conditionnelle est reconsidérée depuis peu comme la stratégie la plus efficace, notamment depuis le développement et l'évolution de nombreuses techniques de surveillance et de diagnostic des machines, majoritairement basées sur l'analyse vibratoire.

En effet et dans la réalité, une machine tournante à l'état neuf produit de faibles niveaux vibratoires pendant son fonctionnement. Cependant, au cours de sa durée de vie, elle va subir des dégradations progressives sous les effets du vieillissement de ses composants, du travail des fondations, de l'usure et des déformations des pièces et mécanismes. En conséquence, des changements graduels dans ses propriétés dynamiques apparaissent. Ainsi, les arbres et accouplements subissent des désalignements, les paliers s'usent, les balourds s'affirment sur les rotors, des défauts naissent dans les courroies et les jeux augmentent. Tous ces facteurs vont contribuer à une augmentation de l'énergie vibratoire qui va exciter les modes propres et ajouter des charges dynamiques considérables aux paliers.

Dans ce chapitre, après la mise en évidence de leur criticité, de leurs origines et de leurs images vibratoires, les différentes défaillances mécaniques pouvant altérer les machines tournantes sont énumérées. Un accent particulier est mis sur les défauts de roulement, qui font la problématique principale de cette étude. Leur constitution, leur modélisation, leur fonctionnement ainsi que les techniques et moyens actuels mis en place pour leur surveillance, vont être décrits.

1.2 POLITIQUE ET STRATEGIE DE MAINTENANCE

Conformément aux normes en application, la maintenance est définie au sein d'une entreprise, comme une fonction stratégique et un ensemble d'actions techniques, administratives et de management destinées à maintenir un bien industriel en bon état de marche durant un cycle de vie ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise.

La mise en application de ces actions diffère selon la nature technologique ou la vocation de l'entreprise industrielle. Selon qu'elle opère dans la fabrication de produits, dans le conditionnement, ou autres, chaque entreprise industrielle a ses particularités. Cependant, et dans chaque entité, il existe une fonction «maintenance» qui est régie par des concepts, des

normes, des exigences, etc. Vue l'importance de la politique de maintenance et la stratégie de sa mise en œuvre dans la gestion de l'entreprise, ces fonctions figurent parmi les attributs du comité de direction ou du responsable de maintenance délégué à cet effet.

La réussite d'un plan de maintenance ne peut être garantie qu'avec une véritable volonté de la part de la direction d'engager des moyens et des procédures nécessaires suivants :

- Cibler les machines potentiellement critiques où les avaries entraîneraient des coûts d'arrêt de production élevés, porteraient atteinte à la qualité ou comporteraient des risques pour le personnel et l'environnement ;
- Faire des bilans de pannes et des dégradations connues et possibles à l'aide de l'AMDEC, d'historiques de panne, d'arbres de dépannage ;
- Faire le bilan des dégradations perceptibles par des symptômes, et le recensement des outils adéquats adaptés à leur détection ;
- Choisir judicieusement les personnes affectées à cette surveillance (compétence, pluridisciplinarité, curiosité, travail en équipe, etc. ;
- Effectuer un bilan annuel afin de quantifier les gains et de vérifier la rentabilité économique de la démarche entreprise sous tous les angles, et réfléchir à des améliorations éventuelles.

1.2.1 Concepts et méthodes de maintenance

L'analyse des différentes formes de maintenance se base sur 4 concepts suivants :

- Les événements et les actions associées (échancier, information d'un capteur, mesure d'une usure, apparition d'une défaillance, etc.) ;
- Les méthodes de maintenance qui leur seront respectivement associées (préventive systématique, préventive conditionnelle, corrective) ;
- Les opérations de maintenance (inspection, contrôle, dépannage, réparation) ;
- Les activités connexes (maintenance d'amélioration, rénovation, reconstruction, modernisation, travaux neufs, sécurité, etc.).

La gestion de maintenance prend essentiellement en compte les aspects technique, économique et financier des différentes méthodes utilisables (corrective, préventive systématique et préventive conditionnelle) en vue d'optimiser la disponibilité des matériels. Le choix entre les méthodes de maintenance s'effectue dans le cadre de la politique de la maintenance définie par la direction de l'entreprise. Il dépend des objectifs visés, des orientations et directives de la politique de maintenance, des modes de fonctionnement et des caractéristiques des matériels en exploitation, des conditions d'application de chaque méthode, des coûts de maintenance et des coûts de perte de production.

La maintenance corrective : appelée parfois curative, a pour objet de redonner au matériel des qualités perdues nécessaires à son utilisation. Les défaillances partielles, pannes ou avaries diverses exigeant une maintenance corrective entraînent une indisponibilité immédiate ou à très brève échéance des matériels affectés ou une déperdition en quantité ou en qualité des services rendus.

La maintenance préventive : a pour objectif de réduire la probabilité de défaillance des matériels en cours d'utilisation. Elle est réalisée selon des critères prédéterminés. Elle offre des avantages, parmi lesquels on peut citer :

- L'augmentation de la durée de vie des matériels ;
- La réduction de la probabilité des défaillances en service ;
- La diminution des temps d'arrêt en cas de révision ou de panne ;
- La prévention des interventions coûteuses de maintenance corrective ;
- La programmation de la maintenance corrective dans les meilleures conditions ;
- La diminution du budget de maintenance ;
- Etc.

En mode systématique, elle est effectuée en fonction d'un échancier établi selon le temps ou le nombre d'unités d'usage. Même si le temps est l'unité la plus répandue, d'autres unités peuvent être retenues telles que : la quantité de produits fabriqués, la distance parcourue, le nombre de cycles effectués, etc. La périodicité des interventions est déterminée à partir de la mise en service ou après une révision complète ou partielle. Elle nécessite de connaître le comportement du matériel, les modes de dégradation et le temps moyen de bon fonctionnement entre deux avaries. Elle est préconisée dans les cas d'application suivants :

- Équipements soumis à une législation en vigueur (sécurité réglementée) : appareils de manutention et de levage, extincteurs, réservoirs sous pression, convoyeurs, ascenseurs, monte-charge, etc.
- Équipements dont la panne peut provoquer des accidents graves : moyens de transport en commun des personnes, avions, bateaux, trains, etc.
- Équipement ayant un coût de défaillance élevé : chaînes de production automatisées, processus fonctionnant en continu (industries chimiques ou métallurgiques).

En mode conditionnel, elle permet d'optimiser les opérations de maintenance et surtout de les effectuer au bon moment, au juste coût. Elle permet donc, d'anticiper les pannes à l'aide de techniques préventives conditionnelles. Cette méthode est particulièrement étalée dans le point ci-après.

1.2.2 Maintenance Préventive Conditionnelle

Conditionnée par l'état de la machine sous surveillance, cette méthode indispensable de maintenance est promulguée de nos jours aux premiers rangs. Elle est venue remplacer progressivement la maintenance systématique très coûteuse en temps et en argent. Cette ascension a été gagnée particulièrement grâce aux succès des résultats obtenus, suite aux développements de techniques spécifiques inhérentes.

1.2.2.1 Mise en œuvre

C'est une maintenance selon l'état qui est exécutée à titre préventif en faisant intervenir des informations recueillies en temps réel. Elle est souvent conduite par des mesures pertinentes, périodiques ou continues, sur le matériel en fonctionnement. Selon la nature de la technique utilisée, les paramètres mesurés peuvent porter sur :

- La qualité de l'huile ;
- La température ;
- La pression ;
- La tension et l'intensité, électriques ;
- Les vibrations ;
- Etc.

Dans cette voie, la surveillance d'un parc de machines, peut être exprimé globalement comme étant un suivi de la valeur d'un certain nombre d'indicateurs de l'état de santé, relevés par des capteurs. Une évolution dans le temps de ces indicateurs permet de détecter une anomalie. La planification des interventions repose sur la détermination de seuils critiques de ces indicateurs de dégradation, communément appelés seuils d'alarmes.

Il y a trois étapes primordiales, nécessaires à la mise en œuvre d'une telle maintenance :

- 1) Appliquer une technique de surveillance performante pour la mise en évidence d'un éventuel dysfonctionnement.
- 2) Détecter les avaries par une surveillance en permanence.
- 3) Diagnostiquer la nature et l'ampleur de la dégradation et établir un pronostic conséquent.

Il est à noter que ce type de maintenance est relativement onéreux à mettre en place car il nécessite l'intégration d'instruments de mesures fiables sur le système à surveiller et fait appel à des compétences avérées en termes d'analyses et traitements de données. Cependant, il reflète de la meilleure manière, l'état réel de fonctionnement de l'équipement surveillé, et dans bien

des cas, il peut être rentabilisé rapidement. Les mesures prélevées peuvent être extrapolées pour en déduire la durée de vie résiduelle avant la survenue d'une panne.

1.2.2.2 Techniques de surveillance utilisées en maintenance conditionnelle

Naturellement, le choix de la méthode de surveillance à utiliser dépend de la nature des défauts potentiels à surveiller, des impératifs de la chaîne de production, de sa complexité et d'autres conditions de l'environnement. En d'autres termes, de la politique de maintenance définie dans le plan de management de l'entreprise (le centre de décision). Il revêt une grande importance et doit être le plus approprié. Ainsi, en fonction des contraintes considérées on pourra opter pour un ou plusieurs choix parmi ces types de de surveillance suivants :

La thermographie infrarouge c'est une technique de mesure sans contact de la température de composants de machines. Elle peut être utilisée dans le cas où la présence du défaut s'accompagne d'un échauffement anormal à la surface de l'équipement. Elle permet donc, grâce à des mesures à distance, d'obtenir instantanément une image thermique de la zone inspectée.

L'analyse des huiles Elle consiste à analyser un échantillon d'huile prélevé à partir des machines contenant des fluides de lubrification (réducteurs, motoréducteurs, moteurs thermique, etc.), pour déduire l'état de l'équipement. L'huile étant en contact permanent avec les pièces en mouvement, fournit de nombreuses informations sur l'usure des organes et de la qualité du lubrifiant.

L'analyse acoustique Elle est utilisée dans des applications spécifiques. Etant aussi une mesure sans contact, elle consiste en une détection de tout bruit anormal au moyen de microphones placés soigneusement à une certaine distance de l'équipement.

L'analyse vibratoire C'est la méthode la plus utilisée et elle permet la détection de la majorité des défauts susceptibles d'apparaître dans les machines tournantes en fonctionnement. Elle se base sur l'existence d'une forte corrélation ente le comportement vibratoire et l'état de la machine. Ainsi plus l'état de détérioration de la machine est à un stade avancé plus les niveaux de vibration sont importants.

Il est bien montré dans la littérature et résumé dans le tableau 1.1, que la méthode basée sur la surveillance par analyse vibratoire est l'une des plus efficaces de par sa capacité à détecter, localiser et distinguer les différents types de défauts dès leurs naissance, et avant qu'ils ne deviennent critiques et dangereux [1]. Ces défauts peuvent être distribués ou localisés. Chaque machine a une signature de vibration spécifique et qui est liée à sa conception et son

état. Si ce dernier change, il en serait de même pour la signature vibratoire. Ce changement peut être utilisé pour détecter tous défauts naissants.

Tableau 1.1 : Tableau comparatif des méthodes de surveillance en maintenance [1]

	Principaux avantages	Principales limitations	Champ d'applications privilégié
Analyse vibratoire	<ul style="list-style-type: none"> - Détection de défauts à un stade précoce - Possibilité de réaliser un diagnostic approfondi - Autorise une surveillance continue - Permet de surveiller l'équipement à distance (télémaintenance) 	<ul style="list-style-type: none"> - Spectres parfois difficiles à interpréter - Dans le cas de la surveillance continue, installations relativement coûteuses 	<ul style="list-style-type: none"> - Détection des défauts de tous les organes cinématiques de la machine (balourd, défauts d'alignement, jeux, etc.) et de sa structure
Analyse d'huiles	<ul style="list-style-type: none"> - Détection d'une pollution anormale du lubrifiant, avant que celle-ci n'entraîne une usure ou un échauffement - Possibilité de connaître l'origine de l'anomalie par analyse des particules 	<ul style="list-style-type: none"> - Ne permet pas de localiser précisément le défaut - Nécessite de prendre de nombreuses précautions dans le prélèvement de l'échantillon 	<ul style="list-style-type: none"> - Contrôle des propriétés physico-chimiques du lubrifiant, détection d'un manque de lubrification, analyse des éléments d'usure, analyse de contamination par le process (étanchéité), etc.
Thermographie IR	<ul style="list-style-type: none"> - Permet de réaliser un contrôle rapide de l'installation - Interprétation souvent immédiate des résultats 	<ul style="list-style-type: none"> - Détection de défauts à un stade moins précoce que l'analyse vibratoire - Contrôle limité à ce que « voit » la caméra (échauffements de surface) - Ne permet pas de réaliser un diagnostic approfondi 	<ul style="list-style-type: none"> - Détection de tous les défauts engendrant un échauffement (manque de lubrification en particulier)
Analyse acoustique	<ul style="list-style-type: none"> - Permet de détecter l'apparition de défauts audibles - Autorise une surveillance continue 	<ul style="list-style-type: none"> - Sensibilité au bruit ambiant - Diagnostic souvent difficile à réaliser - Problèmes de répétabilité des mesures 	<ul style="list-style-type: none"> - Détection d'un bruit inhabituel pouvant ensuite être analysé par analyse vibratoire

Par ailleurs, les statistiques de la figure 1.1 confirment bien que l'analyse vibratoire est la méthode la plus connue et la plus largement utilisée, car adaptée à tout type de machines industrielles en fonctionnement. L'état actuel de la recherche aussi bien dans le domaine du diagnostic que celui du pronostic connaît un développement fulgurant depuis quelques

décennies. Cette position le restera encore pour de longues années. C'est pourquoi les travaux de cette thèse s'inscrivent dans la même direction.

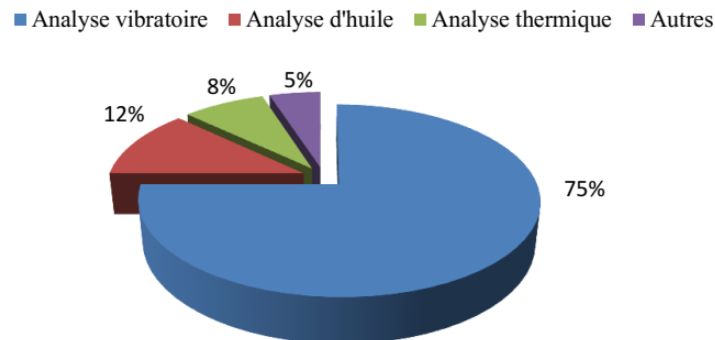


Figure 1.1 : Statistiques des méthodes de surveillance employées en maintenance [2].

1.3 MAINTENANCE CONDITIONNELLE PAR ANALYSE VIBRATOIRE

A l'origine, cette méthode de maintenance était destinée à permettre l'arrêt d'une machine avant qu'elle ne subisse des dégradations importantes. De nos jours, la surveillance vibratoire des machines est devenue le pilier d'une stratégie de maintenance appelée la maintenance préventive conditionnelle (MPC), et s'appuie sur le suivi de l'évolution d'indicateurs objectifs représentatifs de l'état réel d'une machine.

1.3.1 Approche sélective de l'analyse vibratoire

D'un point de vue technologique, les principaux paramètres vibratoires accessibles par des mesures industrielles sont liés aux mouvements relatifs des rotors par rapport à leurs paliers et les mouvements absolus des paliers et carters. Les capteurs utilisés peuvent être de déplacement (Capteur à courant de Foucault), de vitesse (vélocimètre) ou d'accélération (accéléromètre).

La mise en place d'une maintenance conditionnelle efficace basée sur l'analyse vibratoire, exige des choix préalables en termes d'objectifs visés, de stratégie à suivre, de moyens de surveillance adaptés et des méthodes de surveillance et de traitements appropriées. Ces exigences font de l'analyse vibratoire une approche sélective basée sur les éléments suivants :

- a) La sélection des machines tournantes qui feront l'objet de surveillance vibratoire selon leur importance (ou leur criticité) et suivant des critères de sûreté de fonctionnement. Les considérations purement économiques sur les coûts directs et indirects, sont associées à la justification de la sélection.

- b) La sélection des composants et des défaillances potentielles dont la connaissance par retour d'expérience aura permis d'évaluer les risques et leur gravité (AMDEC), et la connaissance des symptômes annonçant la défaillance des composants.
- c) La sélection de la méthode de surveillance la mieux adaptée à la politique de maintenance retenue. Elle peut être :
 - Un suivi périodique (off-line) de l'évolution des paramètres vibratoires, qui peut être accompagné parfois, par un autre mode de surveillance (analyse de lubrifiant, performances, etc.). Cette démarche permet de détecter l'apparition de défauts à évolutions lentes.
 - Un suivi continu (on-line) qui permet de suivre l'évolution de l'ensemble des paramètres. Ce type présente l'avantage de détecter des défauts à évolution rapide et d'assurer la sécurité des installations par déclenchement de la machine à l'approche d'un seuil réputé dangereux.
- d) La sélection des moyens de mesure qui permettront de suivre les paramètres retenus. Elle comporte le choix des types de capteurs et de l'instrumentation de mesure associée (collecteur de données vibratoires, analyseurs multivoies, systèmes d'acquisition multivoies et multifonctions, etc.).
- e) La sélection des moyens de traitement associés aux outils d'acquisition. Le choix est complexe, et dépend principalement des objectifs visés (diagnostic, pronostic) et des compétences disponibles.

1.3.2 Systèmes de mesure des vibrations

L'analyse vibratoire est un concept qui vise à étudier le comportement vibratoire des machines en fonctionnement, en tenant compte de leurs caractéristiques cinématiques, de leur environnement et de la nature des défauts qui peuvent apparaître, en vue d'établir un diagnostic sur leur état de marche. En effet, les vibrations reflètent fidèlement l'ensemble des efforts dynamiques engendrés par les pièces en mouvement.

Les matériels utilisés pour la mesure des vibrations sont diversifiés, et leur choix dépend du mode de surveillance escompté (continu ou périodique), des nombres de paramètres et des voies de mesure. Ils comprennent soient de simples collecteurs de données portatifs, soient des analyseurs monovoies ou multivoies, ou encore des systèmes d'acquisition multivoies et multifonctions complexes.

La qualité de mesure des paramètres de vibration dépend des paramètres liés à l'instrumentation utilisée et du contrôle de l'acquisition. En conséquence, l'analyse vibratoire

exige la maîtrise de la chaîne de mesure et des méthodes de traitement des données acquises. Une chaîne type multivoies d'acquisition et de mesure des vibrations est schématisée dans la figure 1.2. Dans cette figure sont décrites les principales opérations nécessaires à une acquisition de données vibratoires, à savoir :

- Transformation des vibrations mécaniques en signaux électriques grâce à des capteurs adaptés (accéléromètres, vélocimètres ou capteurs de déplacement) ;
- Conditionnement (amplifications de tension, de charge, etc.) des signaux électriques acquis par les capteurs utilisés, pour les rendre exploitables par les appareils d'analyses ;
- Numérisation des signaux analogiques au moyen de convertisseur analogique/numérique (CAN) ;
- Enregistrement des signaux par différents supports matériels tels que les analyseurs de spectres, les collecteurs des données ou les ordinateurs via des cartes d'acquisition associées ;
- Traitements des données acquises, soit en mode continu (on-line) ou en mode périodique (off-line).

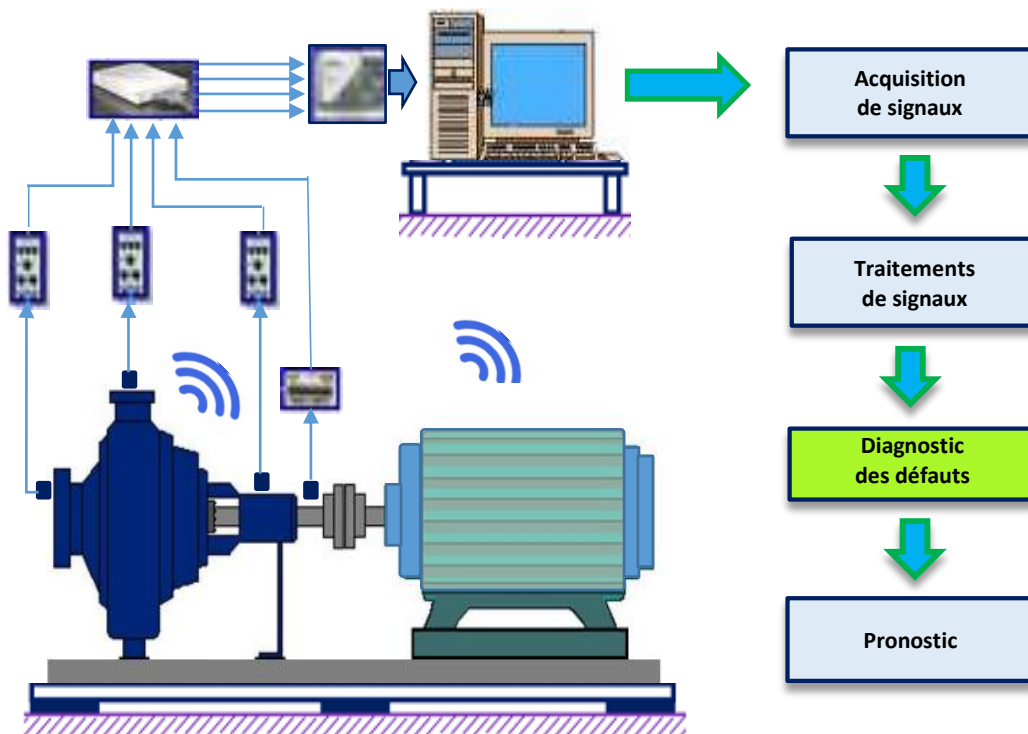


Figure 1.2 : Schéma d'une chaîne d'acquisition multivoies de mesure des vibrations.

L'efficacité de l'analyse vibratoire se base principalement sur deux facteurs. Le premier concerne le bon choix des paramètres d'acquisition (fréquence d'échantillonnage, résolution, fenêtrage, etc.). Le second est lié aux méthodes utilisées en post-traitements.

1.4 TYPOLOGIE DES PRINCIPAUX DEFAUTS MECANIQUES

Les anomalies liées aux machines tournantes en fonctionnement sont de natures très variées et leurs évolutions en absence de surveillance et suivi appropriés, peuvent aboutir à des pannes imprévues aux conséquences dramatiques sur la production. Ainsi, leur détection et diagnostic devient un objectif majeur de toute politique de maintenance menée au sein d'une entreprise.

Vue cette importance, il convient d'énumérer les principaux défauts présents dans les machines tournantes. Les défauts mécaniques usuels seront abordés dans les sous-sections ci-après. Cependant les défauts de roulement seront traités à part dans la section suivante. Quant aux défauts de nature électrique, ils ne font pas l'objet de cette étude.

1.4.1 Défauts de déséquilibre (balourd)

Par définition, un rotor est parfaitement équilibré si son centre de masse appartient à l'axe de rotation et si cet axe est l'axe principal d'inertie. En d'autres termes, sa masse est distribuée symétriquement par rapport à son axe de rotation. Or cette condition théorique est impossible à satisfaire en pratique, car un balourd résiduel existe même pour un système à l'état neuf. C'est pourquoi l'objectif de toute technique d'équilibrage ne consiste pas à l'élimination total de ce déséquilibre, mais seulement de le maintenir en dessous de seuils fixés par les normes en vigueur.

Dans le cas du rotor rigide quelconque, on distingue les déséquilibres usuels dus aux balourd statique, balourd de couple et balourd dynamique qui sont représentés dans la figure 1.3. Dans le premier cas, le centre de masse du rotor ne coïncide pas avec l'axe de rotation. L'axe d'inertie et l'axe de rotation sont parallèles et distants d'une excentricité. En conséquence, durant le mouvement de rotation, il y a apparition d'une force d'inertie centrifuge qui l'entraîne à l'extérieur de sa trajectoire. Dans le second cas, l'axe d'inertie n'est plus colinéaire à l'axe de rotation. Il coupe ce dernier au centre de gravité et les deux axes forment un angle φ . Dans le troisième cas, c'est une combinaison des deux premiers et correspond au cas le plus courant. Cependant, on évoque aussi parfois le balourd quasi statique qui est situé dans un plan différent de celui du centre de gravité, ou le balourd thermique qui génère une modification de la géométrie des parties tournantes en fonction de la température, ou encore le

balourd dit évolutif causé par des frottements de l'arbre avec l'un de ces paliers. Il est à souligner que les balourds, statique, quasi statique et de couple sont des cas particuliers du balourd dynamique [3, 22].

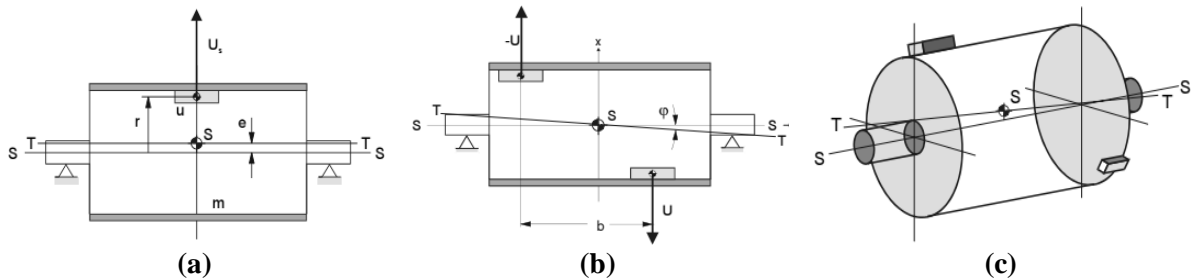


Figure 1.3 : Balourds du rotor rigide. (a) Statique ; (b) De couple ; (c) Dynamique [22].

1.4.1.1 Origines

D'une manière générale, ce déséquilibre provient :

- D'un défaut de fabrication (moulage, usinage, ...) ;
- D'un défaut d'assemblage ou de montage ;
- D'une altération mécanique (usure, encrassement, perte d'ailettes, ...) ;
- D'une déformation thermique (dilatations de matériaux sous l'effet de variations de températures localisées des éléments constitutifs du rotor).

1.4.1.2 Signature spectrale

La présence du balourd engendre un mouvement de vibration dans le plan radial, caractérisé par un pic d'amplitude élevée de fréquence de base égale à la fréquence de rotation F_0 et des harmoniques d'amplitudes plus faibles et de valeurs $2 * F_0$ et $3 * F_0$. La figure 1.4 montre l'évolution du balourd qui se manifeste aussi bien par une augmentation de la raie à F_0 que celles des harmoniques $2 * F_0$, $3 * F_0$, ...

Enfin, il est important de noter des ambiguïtés liées au diagnostic du défaut de balourd. En effet, certains phénomènes sont susceptibles de générer de l'énergie à la fréquence de rotation F_0 sans pour autant qu'ils soient liés à un déséquilibre de masse. En conséquence, les problèmes vibratoires engendrés ne seront donc pas corrigés lors de l'équilibrage de la machine. On peut citer :

- Les jeux de paliers trop importants
- Les chocs périodiques à la rotation (problème de fixation)
- Les défauts rotoriques sur un moteur asynchrone (barres, anneau)
- Les déformations du rotor (effet thermique)
- La résonance de structure.

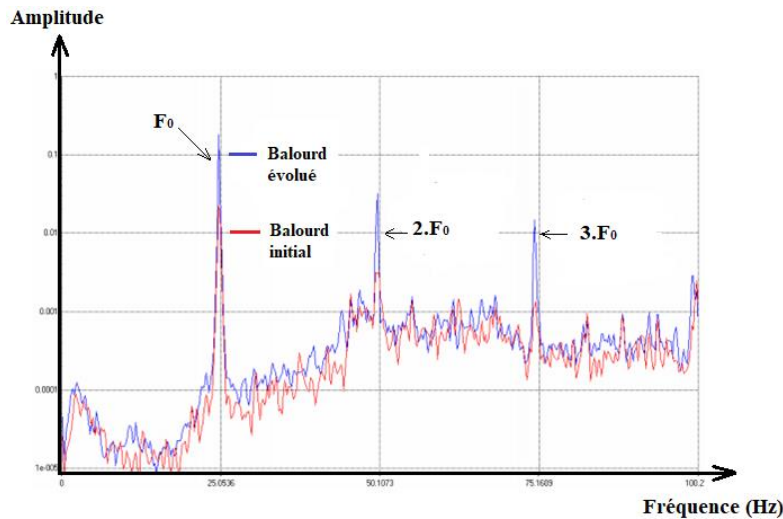


Figure 1.4 : Evolution du spectre du balourd entre deux états (initial et évolué)

1.4.2 Défauts d'alignement

Par définition, c'est l'écart de position relative de l'arbre par rapport à l'axe colinéaire de rotation lorsque la machine tourne dans des conditions normales de fonctionnement. Il peut surgir soit de deux arbres liés par un accouplement, soit de deux paliers soutenant le même axe. Dans tous les cas, il représente l'une des principales causes de réduction de la durée de vie des roulements en particulier.

Il se manifeste au niveau de l'accouplement de deux rotors d'une machine sous forme de défaut de concentricité (décalage radial et axes parallèles), ou sous forme angulaire (axes concourants). La figure 1.5 illustre bien ces deux cas. Il peut se manifester au niveau des deux paliers du même arbre d'une machine si les axes des deux paliers d'un même corps de machine ne sont pas concentriques. Cette anomalie peut être la conséquence d'un défaut de montage d'un palier, mais également d'un mauvais calage des pattes de fixation ou d'une déformation de châssis (par exemple, sous l'effet de contraintes thermiques), qui se traduit par une flexion de l'arbre du rotor [3].

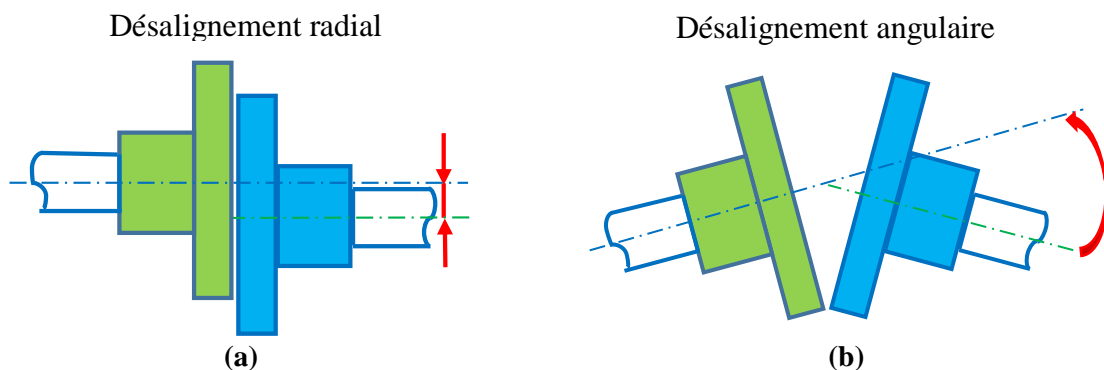


Figure 1.5 : Types de désalignements dans un accouplement

1.4.2.1 Origines

D'une manière générale, ce défaut d'alignement provient :

- De l'usure prématurée des paliers, joints, arbres et accouplements ;
- Des niveaux de vibrations axiale et radiale élevés ;
- De l'élévation de températures dans les corps de palier (températures d'huile élevées) ;
- De la déformation de la structure porteuse ;
- Boulons d'accouplement desserrés ou endommagés ;
- Des fissures d'arbres au niveau ou à proximité des paliers ou des moyeux d'accouplement.

1.4.2.2 Signature spectrale

Qu'il soit du type radial, angulaire ou dû à une flexion d'arbre, le désalignement est un défaut qui va agir sur l'arbre de façon symétrique deux fois à chaque tour. De ce fait, cet ensemble de défauts se manifestera le plus souvent par la présence :

- en direction radiale, de composantes d'ordre 2, rarement 3 ou 4 de la fréquence de rotation dans le cas d'une non concentricité, avec des amplitudes généralement supérieures à celle de la composante d'ordre 1 (figure 1.6).
- en direction axiale, de composantes dont les fréquences correspondent à l'ordre 2 ou 3 de la fréquence de rotation (parfois à l'ordre 1) et dont les amplitudes sont supérieures à celles des composantes radiales correspondantes dans le cas d'un désalignement angulaire ou d'une déformation de l'arbre [3].

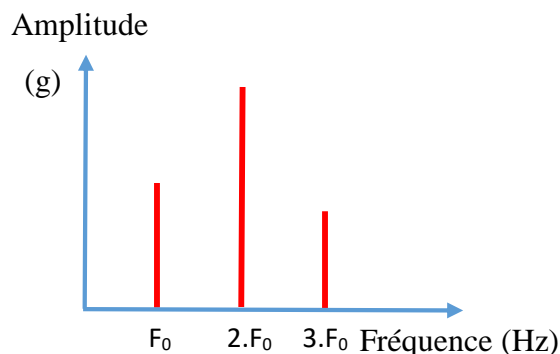


Figure 1.6 : Spectre vibratoire théorique d'un défaut d'alignement

1.4.3 Défauts des engrenages

Un engrenage est l'un parmi les mécanismes utilisés pour transmettre de la puissance entre des éléments de machines en mouvement de rotation. Basiquement, il est composé de deux roues (menante et menée) dentées en contact. Les choix dimensionnels et des profils de dents adaptés permettent la transmission de couples avec ou sans réduction (multiplication) des

vitesses. En fonction de la position relative des deux arbres, on distingue trois groupes d'engrenages (voir figure 1.7) :

- Engrenages parallèles (les axes des arbres sont parallèles) ;
- Engrenages concourants (les prolongements des axes des arbres se coupent) ;
- Engrenages gauches (les axes des arbres occupent une position relative quelconque).

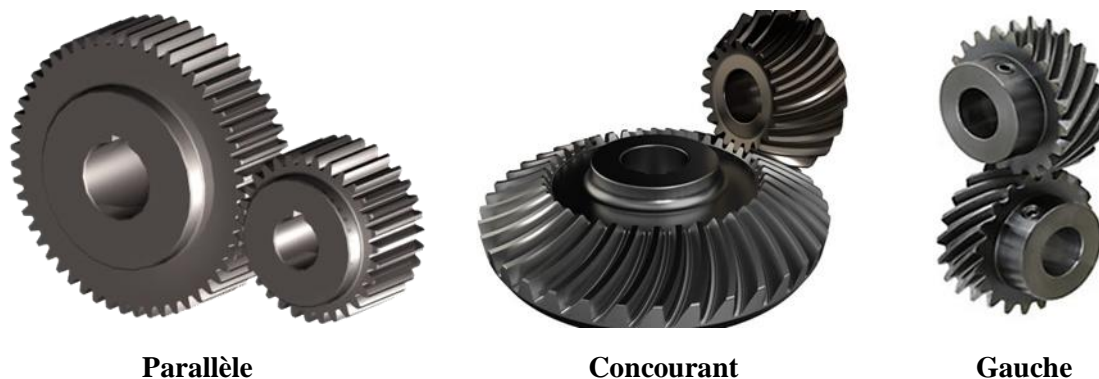


Figure 1.7 : Groupes de transmission par engrenages

Les profils de dentures des engrenages peuvent être droites, hélicoïdales ou à chevrons. Ainsi, et selon le type de denture, les efforts appliqués sont :

- uniquement radiaux sur les engrenages parallèles, à denture droite ou à chevrons ;
- radiaux et axiaux sur les engrenages, à denture hélicoïdale, conique et à roue et vis sans fin.

1.4.3.1 Origines

Les engrenages sont des organes sensibles, et soumis à des charges élevées variées. En conséquence, un grand nombre d'avaries dont les causes sont multiples, surgissent pendant la durée de leur exploitation. Cependant, et hormis les défauts de fabrication et de montage, les défauts peuvent être regroupés selon leur mode de répartition sur la surface de la denture, en deux catégories : les défauts généralisés et localisés [23, 24].

Les **défauts généralisés** sont des défauts répartis sur l'ensemble des dents de l'engrenage. On trouve dans cette catégorie comme le montre la figure 1.8, l'usure abrasive, le Pitting ou piqûres et le grippage généralisé. Le premier défaut est causé par la présence de particules abrasives dans le lubrifiant, ce qui engendre de l'enlèvement de matière lors d'un glissement important entre les deux surfaces frottantes. Quant au deuxième, il se manifeste par la présence de trous plus ou moins profonds sur toutes les dents et en particulier au niveau du cercle primitif. Ces piqûres apparaissent suite à un problème de désalignement ou des déformations de pièces.

Enfin, le troisième est lié à un arrachement de matière au contact entre les surfaces des dents lors de l'engrènement. Sa cause est le film de lubrifiant inadéquat entre les dents.



Figure 1.8 : Défauts généralisés sur différentes dentures [24].

Les **défauts localisés** sont des défauts partiels au niveau de la denture. On y trouve comme le montre la figure 1.9, l'écaillage, la fissuration ou crack et la rupture d'une dent. Le premier défaut se manifeste par des trous moins nombreux que ceux du *pitting*, mais plus profonds et plus étendus. Ils sont produits par fatigue en sous-couche, au point de cisaillement maximal. Ce défaut, évolue très rapidement vers la rupture, sans passer par une phase d'usure. Quant au deuxième, il se forme souvent au pied de la dent qui est l'endroit de concentration de la contrainte maximale de flexion. En plus de la fatigue due à la sollicitation en flexion, ce défaut peut avoir d'autres causes telles que les défauts de fabrication (état des surfaces, criques de rectification) et les criques d'origine thermique. Enfin, le troisième est lié aux chocs dus aux sursurcharges de torsion ou à la fatigue due aux charges cycliques de flexion.



Figure 1.9 : Défauts localisés sur différentes dentures [24, 25].

1.4.3.2 Signature spectrale

Considérons un engrènement composé de deux roues dentées Z_1 et Z_2 tournant avec des vitesses N_1 , N_2 correspondant aux fréquences de rotation F_1 et F_2 (voir figure 1.10). La fréquence d'engrènement F_e (GMF) qui correspond au rythme d'engagement des dents, aura pour expression :

$$F_e = Z_1 F_1 = Z_2 F_2 \quad (1.1)$$

Le rapport des vitesses des deux arbres est lié au nombre de dents de chacun des pignons en contact :

$$Z_1 F_1 = Z_2 F_2 \Rightarrow \frac{N_2}{N_1} = \frac{F_2}{F_1} = \frac{Z_1}{Z_2} \quad (1.2)$$

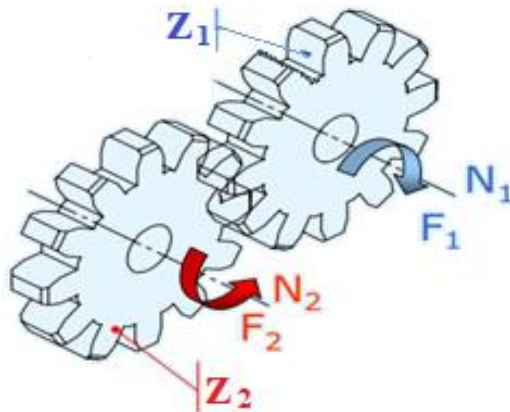


Figure 1.10 : Schéma simplifié de l'engrènement

L'amplitude vibratoire de la raie d'engrènement F_e est très dépendante de la charge de la machine puisque l'engrènement assure la transmission du couple. En l'absence de défaut, on retrouve dans le spectre vibratoire la fréquence d'engrènement F_e et ses harmoniques $2.F_e$, $3.F_e$...

Si la denture est correcte, le spectre illustré dans la figure 1.11, est constitué de composantes dont les fréquences correspondent à la fréquence d'engrènement ou à ses harmoniques [25].

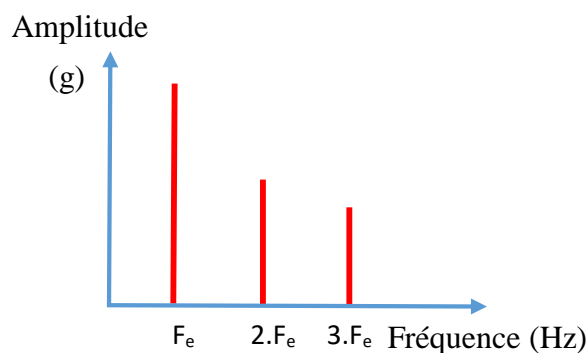


Figure 1.11 : Spectre de l'engrènement à l'état normal

A) Détérioration d'une dent

Si l'une des roues de l'engrènement présente une dent détériorée, il se produit un choc dur, à chaque tour du pignon. Le spectre correspondant montre un peigne de raies dont le pas correspond à la fréquence de rotation du pignon détérioré s'étalant jusqu'aux hautes fréquences. La figure 1.12 montre un modèle de spectre correspondant à ce type de défaut.

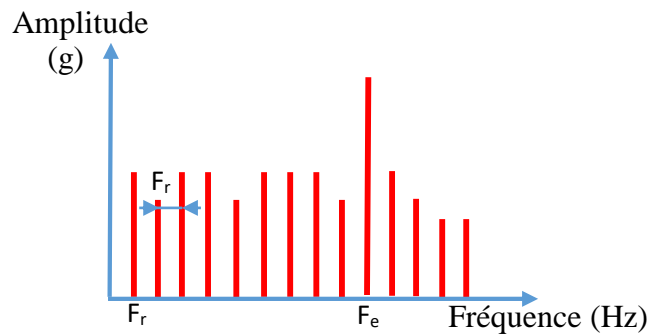


Figure 1.12 : Spectre du défaut d'engrenage ayant une dent détériorée

B) Détérioration de deux dents sur les deux roues

Si les deux roues dentées présentent chacune une dent détériorée, les chocs peuvent être importants lorsque les deux défauts se rencontrent «coïncidence ». La rencontre s'effectue à la fréquence F_{co} , appelée fréquence de coïncidence et telle que :

$$F_{co} = \frac{F_e}{PPCM(Z_1, Z_2)} \quad (1.3)$$

La fréquence de coïncidence F_{co} est inférieure aux fréquences de rotation F_1 et F_2 . Le spectre montre non seulement les deux peignes correspondant aux fréquences de rotation de chaque roue, mais aussi un peigne de raies de pas correspondant à la fréquence de coïncidence F_{co} .

C) Détérioration de l'ensemble de dentures

Lorsque l'ensemble de la denture est usé ou détérioré, les chocs se produisent au passage de chacune des dents. Le spectre illustré dans la figure 1.13, est constitué d'un peigne de raies dont la fréquence correspond à la fréquence d'engrènement, mais cette fois avec une amplitude beaucoup plus élevée. Le spectre peut présenter des raies à la fréquence de rotation correspondant à des équilibrages imparfaits.

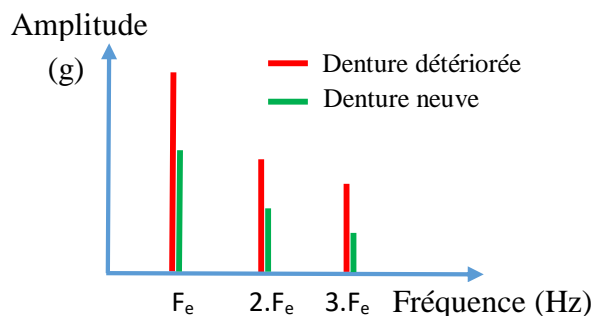


Figure 1.13 : Spectre du défaut d'engrenage ayant l'ensemble de dentures détériorées

D) Entraxe insuffisant (jeu insuffisant au fond de denture)

Si l'entraxe des arbres portant les deux roues est insuffisant, il se produit un engrènement en fond de denture. La dent menante force à l'engagement et au déengagement ce qui provoque une usure des deux côtés de la denture. En analyse spectrale ce phénomène se traduit par l'apparition d'une raie d'amplitude prépondérante à deux fois la fréquence d'engrènement F_e et la disparition partielle, voire totale de l'amplitude à cette fréquence (voir figure 1.14).

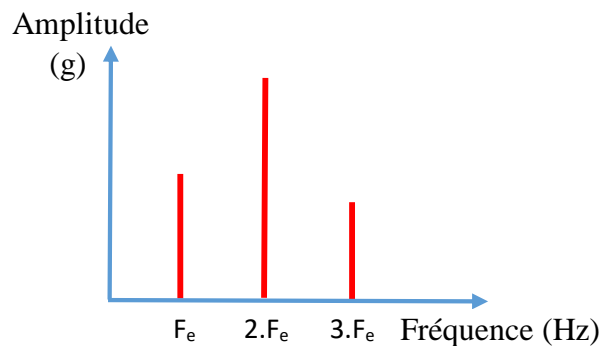


Figure 1.14 : Spectre du défaut d'engrenage ayant l'entraxe insuffisant

E) Jeu de fond de denture trop grand

Un jeu de fond de denture trop grand a pour conséquence un choc dur à chaque passage d'une dent à l'autre ("rattrapage" du jeu). On obtient alors, comme montré à la figure 1.15, un spectre de choc dur, périodique, à la fréquence d'engrènement (présence de nombreuses harmoniques de niveau du même ordre de grandeur).

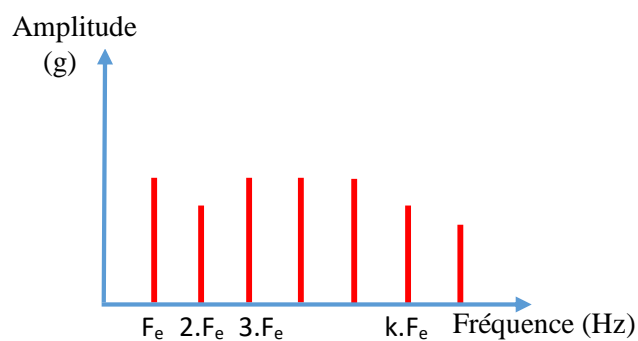


Figure 1.15 : Spectre du défaut d'engrenage ayant un jeu de fond trop grand

F) Défauts induisant une modulation d'amplitude

Si l'arbre ou le pignon présente un défaut d'excentricité, ou de faux rond, il va apparaître une modulation d'amplitude du signal vibratoire par la fréquence de rotation due à la modélisation de l'effort d'engrènement. L'image vibratoire théorique dans la figure 1.16, présente autour de la fréquence d'engrènement (fréquence porteuse) ou de ses harmoniques, des raies latérales dont le pas correspond à la fréquence de rotation de l'arbre qui porte le défaut.

C'est ce qu'on appelle des bandes latérales de modulation d'amplitude. Il faut noter que les amplitudes de ces bandes latérales sont généralement très faibles devant l'amplitude de la fréquence porteuse (entre 2 et 20 fois inférieures) et ne sont visibles que dans des spectres exprimés en décibel.

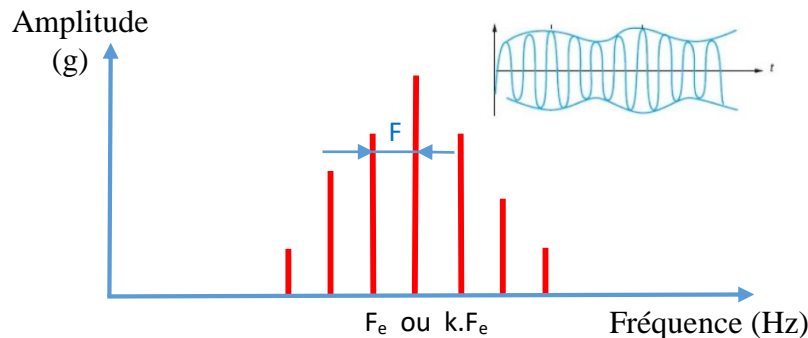


Figure 1.16 : Spectre de l'engrènement modulé par la fréquence de rotation

1.4.4 Défauts de transmission par courroies

Le défaut le plus rencontré sur ce type de transmission se manifeste par une détérioration localisée de la courroie sous forme de partie arrachée, de défaut de jointure, ...etc. En conséquence, cela engendre un effort ou un choc particulier à la fréquence de passage de ce défaut. Néanmoins, d'autres types de défauts liés à ce mode de transmission de puissance existent. On peut citer :

- Courroies détendues,
- Courroies usées,
- Mauvais alignement des poulies,
- Poulie excentrée,
- Courroie et poulie crantées.

1.4.4.1 Origines

Selon la nature du problème de transmission par courroie, on distingue les cas suivants :

A) Courroies usées, détendues ou désaxées

La détérioration localisée d'une courroie comme le montre la figure 1.17, implique généralement un effort ou un choc dont la fréquence de répétition correspond à la fréquence de passage de ce défaut. Elle se situe toujours en dessous des fréquences de rotation des arbres menant et mené. À cause des inévitables glissements poulies/courroies, la fréquence réelle sera toujours inférieure de 2 à 3 % de sa valeur théorique calculée à partir de la relation :

$$F_c = \frac{\pi D_1}{L} F_1 = \frac{\pi D_2}{L} F_2 \quad (1.4)$$

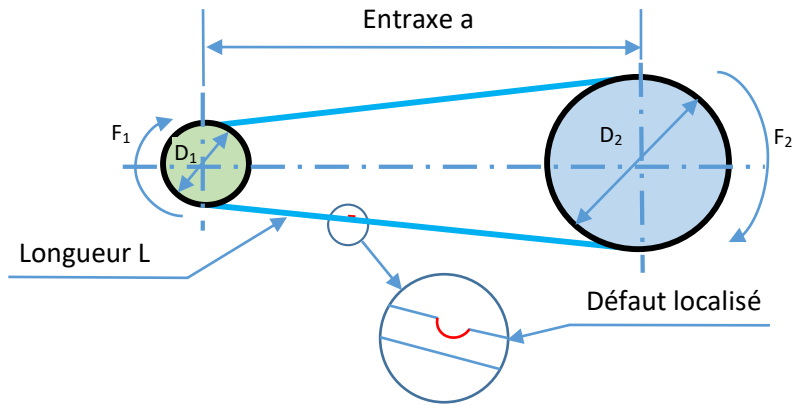


Figure 1.17 : Schéma d'un défaut localisé de courroie

Dans le cas d'une courroie crantée, l'usure ou le désalignement des poulies se manifestent par la présence d'une composante d'amplitude élevée dont la fréquence correspond à la fréquence de synchronisation (produit de la fréquence de passage de courroie par le nombre de crans ou de la fréquence de rotation des roues menante ou menée par le nombre de dents de cette dernière). Pour ce type de transmission, les fréquences rencontrées sont les mêmes que pour les engrenages avec cependant des amplitudes vibratoires moindres du fait de la nature du matériau qui a tendance à amortir les chocs. La fréquence de passage de la courroie s'écrit alors :

$$F_c = \frac{z}{z_c} F_1 = \frac{z}{z_c} F_2 \quad (1.5)$$

B) Mauvais alignement des poulies

Ce désalignement engendre une composante d'amplitude élevée prépondérante en direction axiale à la fréquence de rotation des arbres menant et mené. Le rapport des amplitudes de ces composantes dépend aussi bien de la position du point de mesures que des masses relatives des unités motrice et réceptrice et de la rigidité des châssis supports. Souvent, dans le cas d'un désalignement de poulies, la fréquence de la composante axiale d'amplitude prépondérante sur le moteur correspond à la fréquence de rotation du récepteur. Le diagnostic peut être confirmé par la mesure des déphasages entre composantes axiales de même fréquence.

C) Poulie excentrée

Dans ce cas, le centre de gravité différent du centre de rotation, induit une composante vibratoire d'amplitude élevée à la fois sur les paliers des poulies menante et menée en direction radiale, dont la fréquence correspond à la fréquence de rotation de la poulie excentrée. La direction pour laquelle l'amplitude de cette composante est la plus élevée correspond à la direction définie par la droite joignant les centres des poulies. L'excentricité d'une poulie peut

être confirmée par une mesure de déphasage, le mouvement radial des paliers étant quasiment rectiligne.

1.4.4.2 Signature spectrale

Généralement, le spectre provient d'une mesure radiale prise dans la direction de l'effort de tension de la courroie. Une mesure dans la direction axiale peut, éventuellement, détecter un défaut de désalignement des poulies. L'image vibratoire d'un défaut de courroies (figure 1.18) est constituée d'un peigne de raies dont le pas correspond à la fréquence de passage de courroies avec prépondérance des composantes dont les fréquences sont voisines de fréquences propres de structure. Elle donne un pic d'amplitude importante à la fréquence de passage des courroies F_c , et ses harmoniques $2 \cdot F_c$, $3 \cdot F_c$, ..., $k \cdot F_c$.

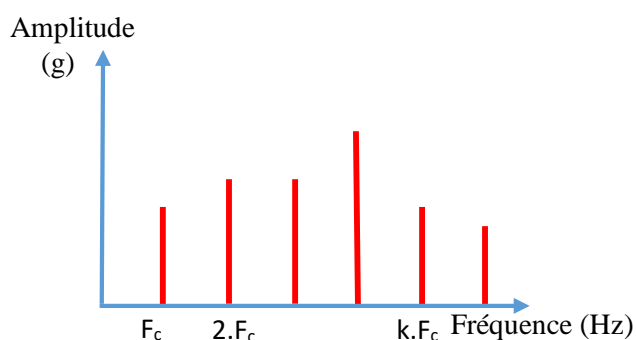


Figure 1.18 : Image vibratoire théorique d'un défaut de transmission par courroie

Lorsque les courroies sont usées ou détériorées, détendues ou désaxées, on observe dans le spectre vibratoire la présence de composantes dont les fréquences correspondent aux premiers ordres de la fréquence de passage de courroie. Souvent, l'amplitude de la composante d'ordre 2 est prépondérante.

1.4.5 Défauts sur les paliers lisses

Les défauts sur les paliers lisses ne provoquent pas de véritables vibrations, comme le cas des roulements. On admet généralement qu'il est difficile de rechercher un défaut sur un palier lisse (enlèvement de matière) à l'aide d'un accéléromètre. Le meilleur moyen pour la surveillance des paliers lisses est l'utilisation de capteurs de déplacement qui permettent de mesurer les déplacements de l'axe de rotation. Ainsi, il est possible d'utiliser l'analyse orbitale. Un problème spécifique aux paliers lisses est l'instabilité du film d'huile à l'intérieur du palier. Des turbulences se produisent dans l'huile et provoquent l'excentricité de l'arbre. Cette excentricité peut aussi être provoquée par un balourd. La fréquence résultante de l'instabilité du film d'huile se situe aux alentours de 0.35 à 0.49 fois la vitesse de rotation.

1.4.6 Défauts induits par un desserrage

Un mauvais serrage de la structure de la machine génère aussi bien des vibrations que du bruit. Le spectre typique mesuré sur une machine présentant ce défaut contient un grand nombre de pic à des fréquences multiples de la fréquence de rotation. Il est également parfois possible de retrouver des pics à l'harmonique égale $1/2 \times$ la fréquence de rotation de l'arbre, et ses multiples. Les causes du jeu sont soit l'usure soit un mauvais montage. Les exemples de jeu, qu'on peut citer sont, l'usure du caoutchouc de l'accouplement, les paliers montés avec des jeux excessifs ou les boulons cassés (plus de jonction mécanique). Une roue à aubes dont les aubes touchent la carcasse nous donne un défaut semblable au jeu, amplifié. L'image vibratoire induite par un desserrage ou par une fissuration d'ancrage ou de bâti est également caractérisée par la présence d'une composante d'amplitude prépondérante dont la fréquence peut également correspondre à l'ordre 1 de la fréquence de rotation. Toutefois, un desserrage ou une fissuration d'ancrage ou de bâti se traduisent par une diminution de raideurs modales avec, pour conséquence dynamique, une diminution significative de fréquence de certains modes propres, fréquences qui peuvent alors coïncider avec un des premiers ordres de la fréquence de rotation dont l'amplitude sera alors prépondérante.

1.4.7 Défaut de cavitation

Il s'agit d'un défaut d'écoulement d'un fluide qui se produit quand la pression, à l'intérieur de la pompe, s'abaisse à un niveau égal à celle de la vapeur saturante du fluide véhiculé. Il y a formation de bulles de gaz qui, en implosant, excitent un grand nombre de modes de structure. La cavitation est causée par un débit d'aspiration insuffisant et se manifeste par un bruit caractéristique identique à celui que provoquerait le passage de graviers dans le corps de la pompe. Ces phénomènes sont souvent intermittents et s'ils persistent, ils peuvent dégrader des éléments internes de la pompe et notamment éroder les aubes des impulseurs. La cavitation induit des vibrations aléatoires à spectre étendu et dont l'énergie se concentre souvent autour de la fréquence de passage des aubes et de ses harmoniques et des fréquences des principaux modes de structure.

La démodulation du signal vibratoire dans les zones de concentration d'énergie, se caractérise par des spectres « plats » démontrant ainsi l'absence de toute périodicité. On pourra distinguer un phénomène de cavitation des autres phénomènes ayant des manifestations vibratoires voisines (usure de bagues de roulement, défaut de lubrification) en prenant une mesure sur le corps de pompe. Pour un défaut lié à un palier, l'énergie vibratoire sera beaucoup

plus élevée sur le palier que sur la volute, alors que pour la cavitation, elle sera généralement supérieure sur la volute.

1.5 DEFAUTS DES ROULEMENTS

Les roulements sont des composants stratégiques des machines tournantes, situés entre les parties fixes et les parties mobiles de la structure. Ils assurent le support ou l'appui, le guidage, la transmission des efforts et la rotation de l'arbre. En conséquence, ils sont le siège de fortes contraintes mécaniques qui entraînent peu à peu une dégradation des matériaux des éléments roulants, et finissent par provoquer l'apparition de défauts ponctuels (des fissures, un écaillage des surfaces, etc.). C'est pourquoi, ils sont considérés comme les organes les plus sensibles des machines tournantes et suscitent donc autant d'intérêt.

1.5.1 Description et caractéristiques du roulement à billes

Parmi les types de roulements existants, notre étude porte sur le roulement à bille, car il est le plus couramment employé dans les machines à puissance moyenne (ordre de quelques KW). Comme montré dans la figure 1.19, ce modèle est composé : d'une bague extérieure (1), d'une bague intérieure (5) avec le chemin de roulement des billes (4), d'une cage (3) et d'une rangée de billes à contact radial (2). Il est conçu de façon à ce que les billes effectuent simultanément une rotation autour de leur propre axe et autour de l'axe des pistes ou chemins de roulement sur les bagues. Le contact entre les billes et les bagues est ponctuel. Les billes sont fabriquées habituellement en acier, ou en céramique pour les roulements hybrides. La cage a pour rôle, le maintient les billes à égale distance les unes par rapport aux autres, tout en les retenant à l'intérieur du roulement. Cependant, elle ne supporte pas directement les efforts appliqués au roulement. On distingue le type de types de cage en fonction de sa fabrication (en tôle emboutie, usinée,

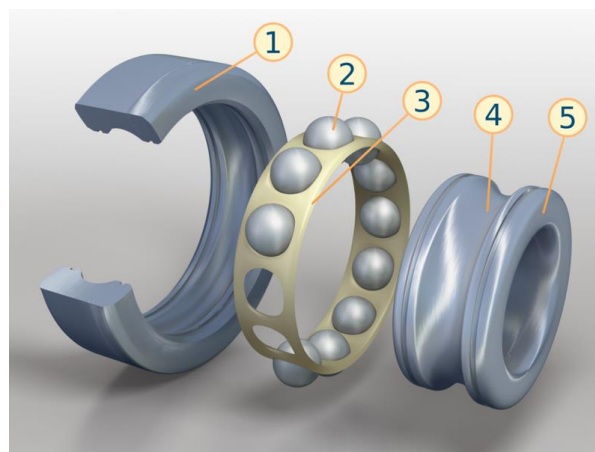


Figure 1.19 : Eléments constitutifs du roulement à bille [26].

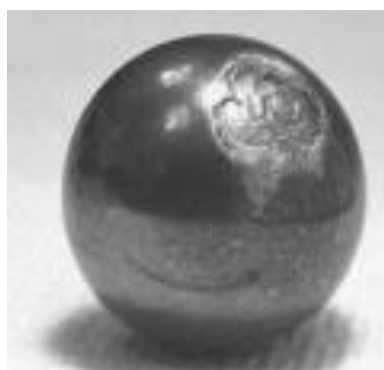
En termes de caractéristiques, les roulements à billes tolèrent à la fois des charges radiales et des charges axiales (figure 1-9). Leur plage de température d'utilisation varie de -20° jusqu'à 120° . Parfois, la rigidité de ces roulements est augmentée par une précharge de contrainte axiale, destinée à forcer tous les éléments roulants à prendre une surface de contact maximale avec les pistes. Quant à la lubrification, la quantité de graisse nécessaire demeure faible. Ainsi, le couple de frottement au démarrage est plus faible comparé à d'autre type de roulements, et n'est que légèrement supérieur au couple de frottement dynamique [27]. Enfin, les roulements à billes sont fabriqués suivant des normes internationales très rigoureuses, et sont interchangeables.

1.5.2 Typologie de défaillances dans les roulements

On peut classer les différents types de défauts et avaries représentatifs de la détérioration des roulements, d'après leurs processus de dégradation. Ils peuvent être localisés ou généralisés (non localisés).

1.5.2.1 Ecaillage

Ce défaut très fréquent est dû à la fatigue des roulements (figure 1.20). C'est un phénomène naturel qui conduit à une défaillance et ce, quel que soit les conditions d'utilisation et de fonctionnement. Il survient sous l'effet de la fatigue due aux contraintes de cisaillement alternées qui sévissent en sous-couche. Durant le fonctionnement, le dépassement de la limite d'élasticité, crée des microdéformations permanentes et des fissures qui se propagent vers la surface et conduisent à l'écaillage. Ce dernier est un processus continu qui s'accélère plus ou moins après l'apparition des premières fissures. Dans la littérature, des études de recherches entières ont été consacrées au suivi de l'évolution de la taille de l'écaillage [28, 29].



(a) Bille



(b) Bague intérieure

Figure 1.20 : Illustrations de deux cas d'écaillage [29]

L'écaillage peut être localisé et prématuré si il résulte d'anomalies caractérisées telles qu'un mauvais montage, une surcharge, un défaut d'alignement, une mauvaise forme de

logement. Ces défauts émettent des signatures vibratoires qui les caractérisent. C'est pourquoi l'analyse vibratoire est un outil puissant qui a pour objectif de surveiller l'état de fonctionnement d'une part et d'autre part de diagnostiquer le défaut et suivre son évolution.

1.5.2.2 Grippage

Il se produit lors de la surchauffe instantanée en cours de rotation, et le roulement ternit ou change de couleur. Puis les bagues de chemin de roulement (figure 1.21), les éléments roulants et la cage se ramollissent, fondent et se déforment en fonction de l'évolution des dommages [30]. Il résulte d'une lubrification insuffisante ou inexistante. Les contacts métal-métal (sous pression, température et des conditions de vitesses moyennes ou élevées), entraînent des échauffements qui facilitent les microsoudures et le transfert de métal.

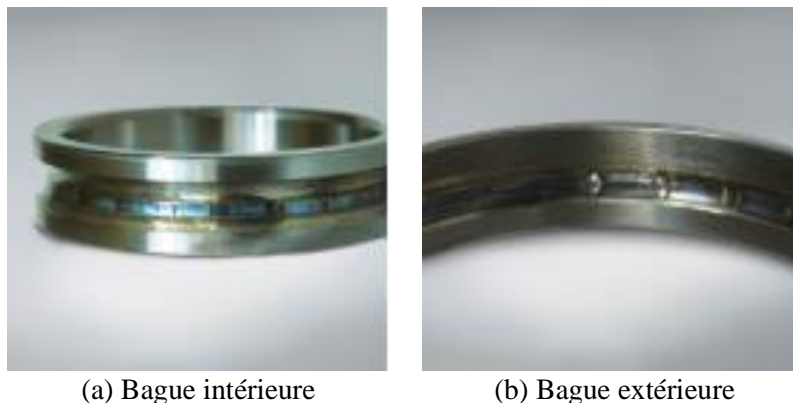


Figure 1.21 : Grippage de roulement à billes [30].

1.5.2.3 Éraflures

Les éraflures consistent en un endommagement de la surface dû à l'accumulation de petits grippages causés par le glissement dans des conditions de lubrification insuffisante ou d'utilisation intense (figure 1.22). Un endommagement linéaire apparaît sur toute la circonférence du chemin de roulement et sur les surfaces de roulement. Un endommagement de forme cycloïdale au niveau des extrémités du roulement ainsi que des éraflures au niveau de la surface de contact de l'épaule peuvent aussi se produire [31].



Figure 1.22 : Éraflures de cage d'un roulement à billes à gorge profonde [30]

1.5.2.4 Usure de contact (fretting)

Il s'agit de l'usure et de l'oxydation due au glissement répétitif entre deux surfaces en acier de composants non tournants. Le frottement se produit à la surface de piquûre et également à la zone de contact entre l'anneau du chemin de roulement et les éléments roulants (figure 1.23). La corrosion par frottement ou de contact est un autre terme utilisé pour décrire une pâte brune (ou les particules usées brun rougeâtre ou noir) formée par la rouille et le lubrifiant [31].



(a) Bague intérieure d'un roulement à billes à gorge profonde

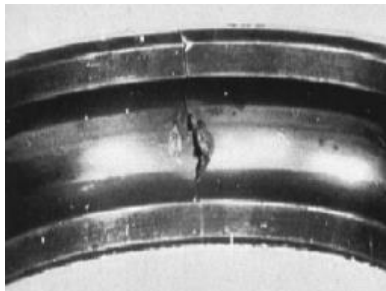


(b) Bague intérieure d'un roulement à billes à contact oblique

Figure 1.23 : Usure de contact sur toute la circonférence de la surface de l'alésage [30]

1.5.2.5 Rupture

La rupture des éléments (Fig. 1.24) est rare, elle est précédée par des craquelures dues à des charges élevées ou à des défauts : portée réduite de la bague extérieure et contraintes pendant l'usinage.



(a) Bague



(b) Cage

Figure 1.24 : Ruptures dans des éléments de roulement [31]

1.5.2.6 Rouille et corrosion

La rouille et la corrosion (figure 1.25), sont le résultat du phénomène d'oxydation de l'acier. Ceci peut provoquer des piquûres à la surface des anneaux et des éléments roulants. Cette avarie surgit suite à une pénétration, d'eau, de fluide ou de gaz corrosif, ou de la condensation d'humidité dans l'air. Les mauvaises conditions d'emballage ou de stockage peuvent y contribuer à l'apparition et la propagation de ce défaut [30].



(a) Corrosion généralisée



(b) Apparition de zones de corrosion

Figure 1.25 : Défaut de corrosion sur les bagues de roulements [30]

1.5.2.7 Autres défauts

D'autres défauts cités brièvement dans la littérature [32], sont énumérés ci-après :

Usure Elle se manifeste sous deux formes. La première est l'endommagement continu, dans des conditions de film fin et de vitesses faibles, quand les interactions entre les rugosités de surface des matériaux peuvent avoir lieu. La seconde est l'usure par un troisième corps (débris d'usure en suspension dans l'huile). Par ailleurs ces corps étrangers (poussières, particules) introduits au montage ou pendant le fonctionnement, peuvent accroître le jeu, créer des cavités, coincer l'élément roulant et par la suite entraîner une rotation des bagues par rapport aux éléments qui les contiennent. L'usure générée par les corps étrangers peut être réduite par filtrage à 10 μm .

Reptation C'est une rotation de la bague intérieure sur l'arbre. Elle est due à un jeu. Il en résulte une usure de la bague et de l'arbre. Elle est révélée par un polissage ou par des stries.

Écaillage superficiel (peeling) C'est un enlèvement superficiel de métal, plus ou moins étendu, sous forme de paillettes très fines. Il est attribué à une épaisseur de lubrifiant trop faible par rapport à la rugosité, ce qui provoque des contacts métal sur métal. Le remède consiste essentiellement à diminuer la rugosité et à augmenter la viscosité du lubrifiant.

Déséquilibre thermique Il est lié au dégagement de chaleur des surfaces en contact. Si cette énergie thermique n'est pas dissipée en continu, une élévation de température est possible et engendre une destruction du lubrifiant ainsi qu'une réduction de la dureté des matériaux en contact.

Piqûres L'érosion électrique, également appelé contournement électrique ou arc électrique, est un courant vagabond qui passe d'une bague de roulement à l'autre par l'intermédiaire des éléments roulants. L'étendue des dommages dépend de la quantité d'énergie et de sa durée. On note alors des dommages par piqûres des éléments roulants, des pistes de roulement et une dégradation rapide du lubrifiant entraînant une défaillance prématurée du roulement.

Vrai Brinelling L'apparition de bosses sur les chemins de roulement qui sont le résultat d'un dépassement de la limite élastique de l'acier. Les causes possibles proviennent de toute surcharge statique ou impact sévère.

Faux Brinelling L'apparition d'une usure elliptique à l'espacement des billes ou des rouleaux due à une vibration externe excessive. La provenance, un petit mouvement relatif entre les éléments roulants et les chemins de roulement dans un roulement non tournant ou des endommagements dus au transport.

1.5.3 Principales causes des défauts de roulements

Principalement, peu importe la nature des avaries, elles peuvent être structurées en deux catégories. Celles dites naturelles dues à la fatigue des roulements et les avaries dues à un mauvais montage ou une mauvaise utilisation du roulement. Néanmoins, le diagramme de la figure 1.26 illustre bien de manière générale la provenance et les origines des principaux défauts rencontrés dans les roulements. Il s'en suit que la majorité des défaillances prématurées des roulements (54%), sont causées par une lubrification inadéquate [30]. En principe, les éléments roulants des roulements sont antifriction, et conçus pour avoir une fine pellicule d'huile entre les éléments roulants et les surfaces du chemin de roulement. Lorsque ce film se dégrade ou devient trop mince, les éléments roulants entrent en contact avec les surfaces du chemin de roulement et une usure se développe.

Les causes de lubrification inadéquate sont nombreuses :

- Quantité insuffisante ou excessive de lubrifiant (graisse) ;
- Utilisation d'un lubrifiant aux caractéristiques inappropriées ou mélange de lubrifiants ;
- Humidité ou contamination par les particules dures de l'environnement d'exploitation. La contamination peut dégrader les surfaces d'appui ou dégrader le film d'huile, ce qui entraînera également une usure ;
- Température ou vitesse de fonctionnement du roulement, excessives. Plus un roulement fonctionne à des vitesses élevées, plus sa température devient élevée. Les roulements et les lubrifiants ont des limites de température et de vitesse.

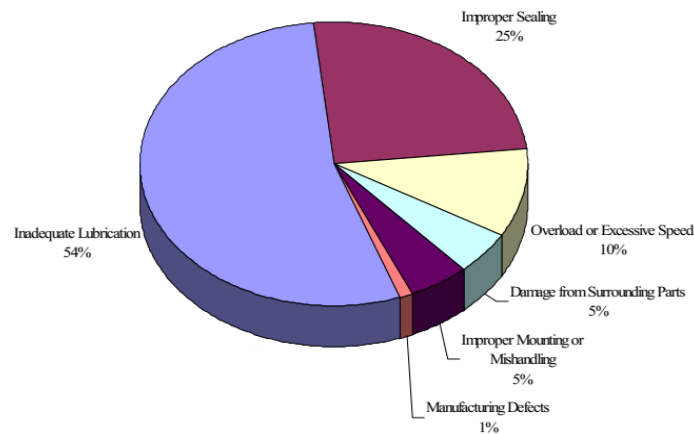


Figure 1.26 : Les causes les plus fréquentes de défaillance des roulements [30].

L'étude de la défaillance d'un roulement implique généralement l'examen de son application. Les chemins ou pistes de roulement, ont tendance à laisser les meilleurs indices sur ce qui a pu causer la défaillance des roulements. Ces derniers devront d'abord être démontés pour accéder et visualiser les chemins de roulement des bagues.

Les autres facteurs d'influence ont les contributions suivantes :

- Étanchéité incorrecte (25%) ;
- Surcharges ou vitesses excessives : vibrations (10%) ;
- Organes environnants : désalignement des accouplements, fléchissement des arbres (5%) ;
- Défauts de montage : mauvaise précontrainte, erreurs de cote, erreur de lignage (5%)
- Défauts de fabrication : matériaux non homogènes, tolérance sur les côtes (1%).

1.5.4 Fréquences caractéristiques des défauts de roulement

Théoriquement, cinq fréquences caractéristiques liées aux défauts dans les éléments roulants d'un roulement fixé sur un arbre tournant, peuvent être répertoriées (figure 1.27). Il s'agit de la fréquence de rotation de l'arbre F_r , de la fréquence de bague intérieure, de la fréquence des billes BPFI, de la fréquence fondamentale de la cage FTF, de la fréquence des billes BSF et de la fréquence de bague extérieure BPFO. Les résonances à haute fréquence entre le roulement et la réponse du capteur, sont excitées lorsque les éléments roulants (billes) heurtent un défaut local sur l'un de ces composants. Ces fréquences caractéristiques des défauts peuvent être calculées à l'aide des équations suivantes [6, 33, 34] :

- 1) La fréquence de passage d'un élément roulant sur un défaut de bague intérieure, supposée montée sur l'arbre tournant dénotée BPFI est donnée par l'équation suivante :

$$BPFI = \frac{nF_r}{2} \left(1 + \frac{d}{D} \cos\phi \right) \quad (1.6)$$

Avec,

n est le nombre d'éléments roulants (billes), D est le diamètre moyen, d est le diamètre des éléments roulants (billes), ϕ est l'angle de contact et F_r est la fréquence de rotation de l'arbre ($F_r = \frac{N_r}{60}$) où N_r est la vitesse de rotation de l'arbre.

2) La fréquence de passage d'un défaut de bille sur la bague externe ou sur la bague interne dénotée BSF , est donnée par l'équation suivante :

$$BSF = \frac{DF_r}{2d} \left(1 - \left(\frac{d}{D} \cos\phi \right)^2 \right) \quad (1.7)$$

3) La fréquence de passage d'un élément roulant sur un défaut de bague extérieure notée $BPFO$ est donnée par l'équation suivante :

$$BPFO = \frac{nF_r}{2} \left(1 - \frac{d}{D} \cos\phi \right) \quad (1.8)$$

4) La fréquence de passage d'un défaut de cage dénotée FTF est donnée par l'équation suivante :

$$FTF = \frac{F_r}{2} \left(1 - \frac{d}{D} \cos\phi \right) \quad (1.9)$$

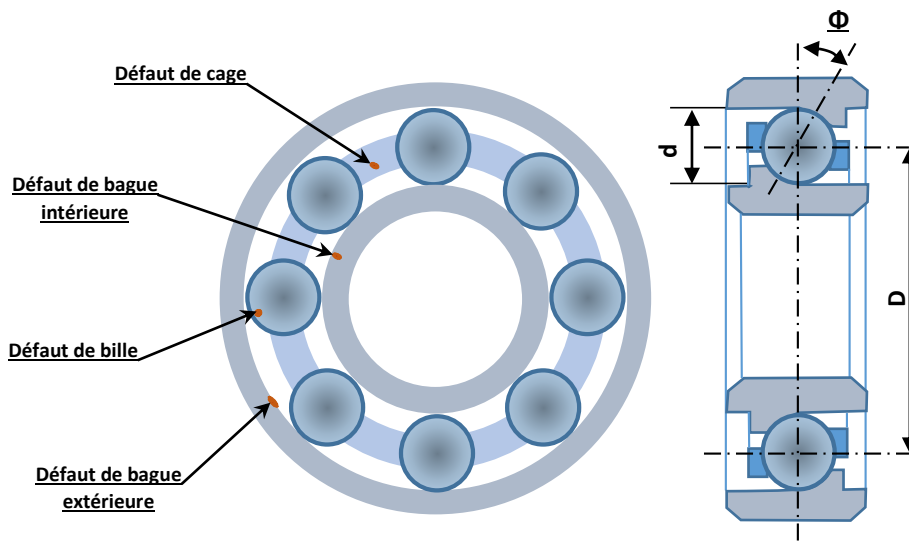


Figure 1.27 : Caractéristiques géométriques d'un roulement à billes

Notons que la fréquence de passage de bille BSF est la fréquence avec laquelle le défaut heurte la même bague (intérieure ou extérieure). En général il existe deux chocs par période de base. En conséquence, les harmoniques paires de BSF sont souvent dominantes, en particulier dans les spectres d'enveloppe. Le calcul de ces fréquences caractéristiques nécessite une connaissance précise des caractéristiques dimensionnelles du roulement.

1.5.5 Signatures et images spectrales théoriques de défauts de roulement

En analyse vibratoire, et parmi les techniques simples de surveillance à partir d'indicateurs scalaires dédiés, l'analyse fréquentielle a l'avantage de déterminer la fréquence de répétition de chocs et, de ce fait, d'identifier la nature du défaut de roulement ou d'une toute autre cause (dégradation d'un engrenage, turbulence aérodynamique, défaut d'origine électromagnétique ...) [3]. Dans le cas d'un défaut de roulement, cette technique est capable :

- d'identifier le processus de dégradation et le stade de dégradation dans lequel il se trouve,
- de localiser le roulement en cause et ses éléments constitutifs en voie d'altération,
- de statuer s'il s'agit d'un écaillage, d'un manque de graisse ou d'une usure des pistes,
- de pouvoir estimer de manière objective les risques encourus à court terme.

1.5.5.1 Relation de la signature en fonction de la progression de la dégradation

Le processus d'écaillage par fatigue est le seul mode de vieillissement normal d'un roulement et c'est ce mode de dégradation qui est pris en compte dans les relations permettant le calcul des durées de vie statistiques. Il existe cependant bien d'autres processus de dégradation de roulement (usure, grippage, réduction de jeux, défaut de serrage du roulement dans son palier ou sur l'arbre, blocage axial du palier guide...), qui présentent des manifestations vibratoires pouvant être très différentes de celles précédemment décrites.

D'une manière générale, le processus normal de dégradation d'un roulement présente quatre stades de dégradation, bien distincts, auxquels sont associées des typologies vibratoires bien différenciées [35]. Le spectre du roulement avec le défaut pendant ses différents stades d'évolution (voir figure 1.28), contient les harmoniques de la fréquence de rotation, des fréquences de défauts, de fréquence de résonance (naturelle) et des hautes fréquences dans le domaine des ultrasons.

Stade 1

Les indications les plus précoces de dégradations, dues à la fissuration en sous-couche d'une bague par fatigue sous l'effet de la pression d'Hertz, apparaissent dans la bande ultrasonique entre 250 et 350 kHz et, plus tard, avec l'accroissement du nombre de fissures, dans la bande [20-60 kHz]. A ce stade, il faudra des capteurs spécifiques pour la détection du défaut. L'inspection physique du roulement à ce stade ne peut identifier le défaut.

Stade 2

Les fissures migrent progressivement vers la surface et le passage des éléments roulants sur ces fissures commence à exciter les modes propres de déformation de bagues de roulement,

dont les premières fréquences propres se situent généralement dans la bande [1500-3500 Hz]. Dans le domaine fréquentiel, on observe alors dans cette bande la présence d'un dôme d'amplitude élevée. Le spectre de modulation du signal filtré dans cette bande permet d'identifier la cadence de répétition des chocs et, de ce fait, de localiser le défaut. À ce stade, l'inspection visuelle peut déceler la présence des défauts.

Stade 3

Certaines fissures qui se propageaient finissent par se rejoindre et forment un morceau de métal qui s'arrache. Cette dégradation devient donc visible à l'œil nu. L'apparition du premier écaillage va immédiatement se manifester par la présence d'un peigne de raies parfaitement identifiable en basses et moyennes fréquences avec un pas correspondant à la fréquence de défaut caractéristique de l'élément altéré (bague interne, bague externe, éléments roulants). L'élargissement de l'étendue de ces zones écaillées se traduit dans le domaine spectral par une forte augmentation de l'amplitude des premières composantes du peigne de raies. Si le roulement est lubrifié à la graisse, les copeaux métalliques restent prisonniers et sont laminés par le passage des éléments roulants (qui ont une dureté très supérieure à celle des bagues) en créant de très nombreuses empreintes. Ces empreintes vont à leur tour engendrer de nombreux chocs de très courte durée, exciter les fréquences propres de bagues et augmenter fortement l'amplitude efficace du signal dans la bande [2 k-20 kHz]. Si le roulement est lubrifié à l'huile, les copeaux sont rapidement évacués et le nombre d'empreintes sera beaucoup plus faible, générant de ce fait dans cette même bande fréquentielle une amplitude efficace considérablement plus faible [3].

Stade 4

À ce stade, l'écaillage est généralisé sur la quasi-totalité des surfaces de roulage suivi d'une forte augmentation des jeux internes du roulement. C'est la dernière étape avant une défaillance catastrophique du roulement. Dans le domaine fréquentiel, elle est associée à de nombreuses fréquences fondamentales et harmoniques modulées indiquant que les défauts sont répartis autour des chemins de roulement. Les grands jeux internes permettent à l'arbre de vibrer plus librement et les défauts induits de déséquilibre et de désalignement viennent s'associer aux fréquences de dégradation accrues. Ainsi, les peignes de raies relatives aux différentes fréquences fondamentales de roulement, sont remplacés par un bruit de fond aléatoire à des fréquences plus élevées. L'importance des chocs conduit rapidement à une rupture de la cage ou à un grippage suite à la mise en travers des éléments roulants.

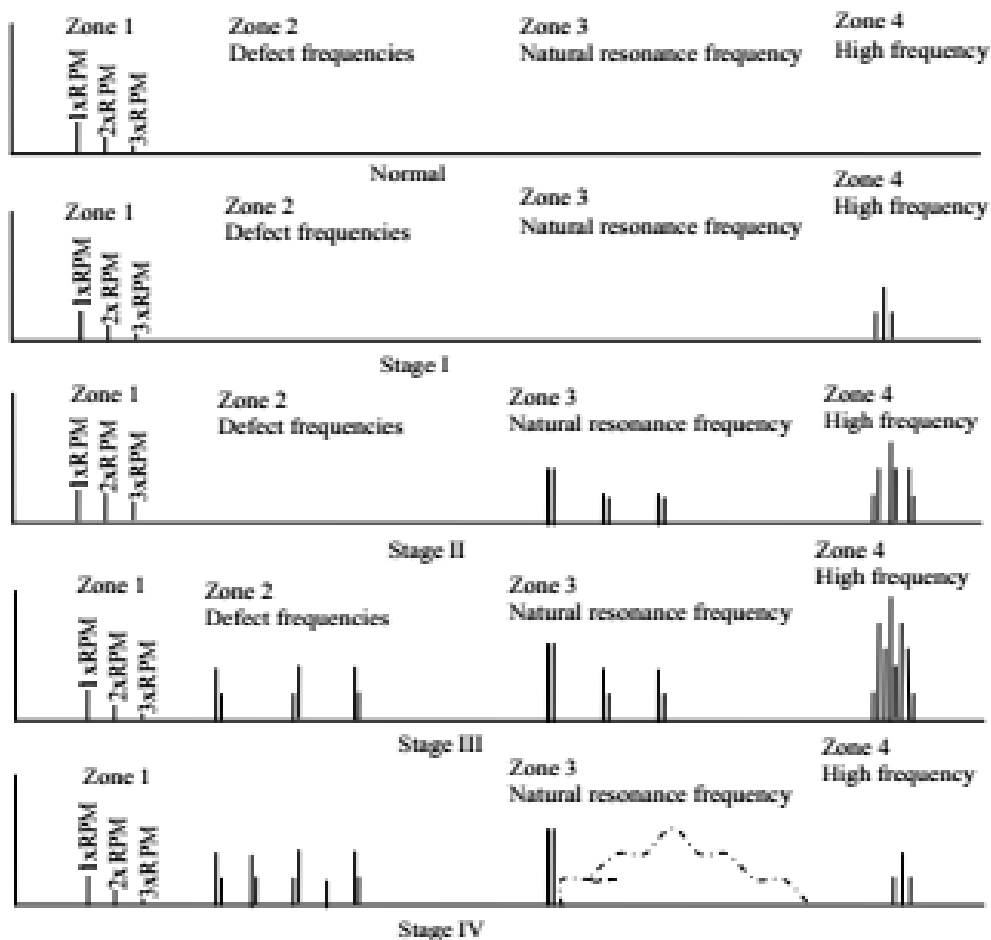


Figure 1.28 : Caractéristiques spectrales des différents stades du défaut de roulement [35]

Pour assurer une bonne fiabilité d'exploitation, il semble indispensable de remplacer le roulement au cours du stade 3, bien avant la disparition des peignes de raies dont les pas correspondent aux fréquences de défauts du roulement.

1.5.5.2 Signature d'un défaut ponctuel peu étendu

L'image vibratoire de ce type de défaut est un peigne de raies dont le pas F_d correspond à la fréquence de défaut associé à l'élément constitutif du roulement qui est affecté par ce défaut (bague externe, bague interne, élément roulant, cage). La forme spectrale du peigne de raies dépend de la position du défaut dans la zone de charge et de sa mobilité éventuelle dans cette dernière.

1) Affectant la bague fixe

Si le défaut affecte la bague fixe du roulement, généralement la bague externe, le défaut se manifestera, dans le domaine spectral, par un peigne de raies de pas égal à la fréquence BPFO (indiquée f_{be} dans la figure 1.29), et dont l'amplitude dépendra essentiellement des réponses

impulsionnelles du roulement et de son palier, l'amplitude des forces d'excitation étant constante.

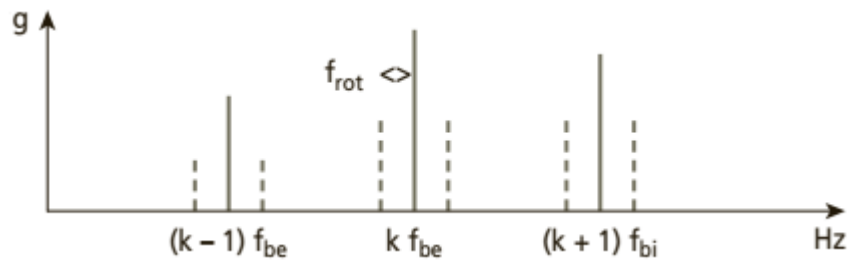


Figure 1.29 : Image vibratoire d'un défaut ponctuel de la bague fixe d'un roulement [3]

En présence d'une charge dynamique importante due, par exemple, à un déséquilibre conséquent, à chaque composante de ce peigne peut être associée une paire de bandes latérales espacées de la fréquence de rotation.

2) Affectant la bague tournante ou un élément roulant

Charge axiale pure

Dans ce cas, la charge est constante quelle que soit la position du défaut. L'amplitude de la force excitatrice est indépendante de la position du défaut, mais la position du défaut par rapport à celle du capteur n'est plus constante et se retrouve à l'identique à l'issue de chaque rotation de l'arbre. Le transfert vibratoire devient alors une fonction périodique du temps.

Il en résulte une modulation de l'amplitude et de la fréquence des chocs, qui se manifeste autour de chaque composante du peigne par la présence d'une à deux paires de bandes latérales (figure 1.30) dont l'espacement correspond, selon la localisation du défaut, à la fréquence de rotation de l'arbre ou de la cage :

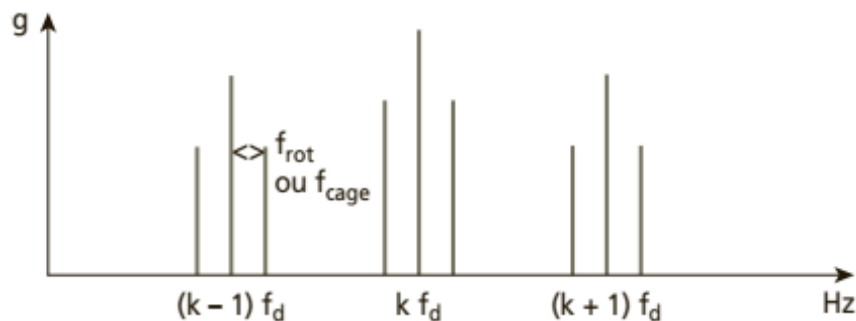


Figure 1.30 : Image spectrale d'un défaut ponctuel de la bague tournante ou d'un élément roulant dans le cas d'une charge axiale pure [3]

Charge radiale pure

Dans ce cas, l'étendue de la zone de charge dépend de la valeur du jeu interne radial fonctionnel. La charge statique est supportée par un nombre restreint d'éléments roulants. En dehors de la zone de charge, les éléments roulants ne sont pas en contact avec le défaut, et de ce fait, il n'y a aucun choc.

À chaque composante du peigne est associé un grand nombre de paires de bandes latérales dont l'espacement correspond, selon la localisation du défaut, à la fréquence de rotation de l'arbre ou de la cage (figure 1.31). Plus la zone de charge du roulement sera étroite ou plus son jeu radial interne fonctionnel sera élevé, plus le nombre de bandes latérales d'amplitude significative associées à chaque composante du peigne sera important et vice versa.

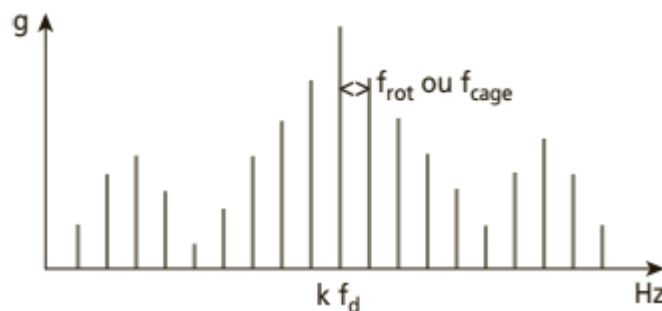


Figure 1.31 : Image spectrale d'un défaut ponctuel de la bague tournante ou un élément roulant dans le cas d'une charge radiale pure [3]

La fréquence du défaut f_d est égale à la BPFI ou BPFO, et la fréquence de modulation correspond à la fréquence de rotation de l'arbre ou de la cage selon que le défaut affecte la bague tournante ou un élément roulant.

1.5.5.3 Déversement de bague

Pour assurer une durée de vie optimale, les bagues d'un roulement doivent être parfaitement alignées, c'est-à-dire que l'angle formé par les axes des bagues intérieure et extérieure doit être nul. En pratique, il existe toujours un certain défaut d'alignement entre l'arbre et le logement. L'effet de ce défaut d'alignement sur la durée de vie dépend de l'importance du déversement exprimé en degrés, du type et de la géométrie interne du roulement et de la charge appliquée. Le défaut d'alignement provoque une distribution de contraintes non uniforme le long de la ligne de contact rouleau/chemin de roulement, voire des contraintes aux extrémités des rouleaux qui réduisent considérablement la durée de vie d'un roulement de manière plus ou moins importante selon son type :

- les roulements dits rigides (roulements rigides à billes, roulements à rouleaux cylindriques), ne supportent guère des déversements d'un angle supérieur à 5 minutes sans induire des contraintes importantes.
- les roulements à alignement automatique (roulements à rotule sur billes ou sur rouleaux, butées à rotule sur rouleaux) sont capables de compenser les défauts d'alignement. Les valeurs maximales admissibles sont données pour chaque type de roulement par les constructeurs.

Un déversement de bague se manifeste dans le domaine vibratoire par la présence d'une raie d'amplitude importante dont la fréquence correspond à la fréquence de défaut de la bague déversée, l'amplitude de cette composante étant largement prépondérante face à ses premiers harmoniques avant que ce défaut n'induisse, par fatigue, des écaillages (figure 1.32).

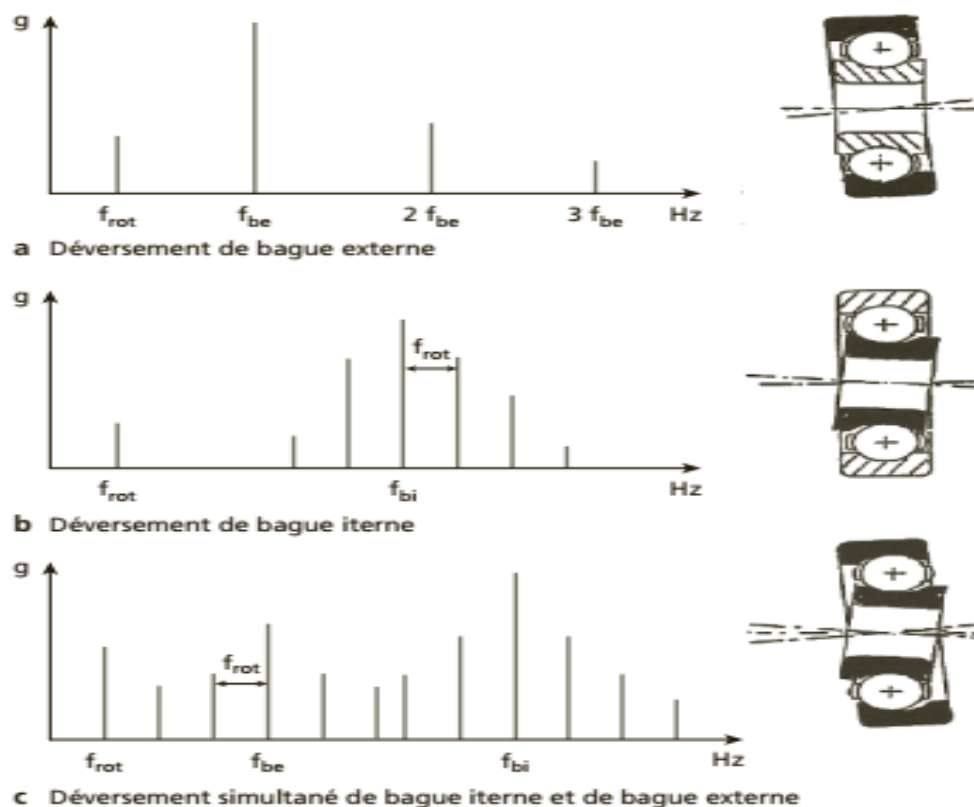


Figure 1.32 : Signature de défauts de déversement de bagues externe et/ou interne [3].

L'amplitude de cette composante dépend de l'importance du déversement, du type de roulement, c'est-à-dire de sa tolérance à accepter ce déversement, et de la proximité ou non d'une résonance de palier. Elle est généralement plus élevée en direction axiale qu'en directions radiales. Dans le cas d'un déversement de la bague externe (figures 1.32 a), l'amplitude de cette composante n'est pas modulée, alors qu'elle l'est fortement dans le cas de déversement de la bague interne (figure 1.32 b). Le déversement simultané des deux bagues présente évidemment

une image vibratoire conjuguant les images de ces deux types de déversement (figure 1.32 c).

Un déversement de bague se manifeste par une vibration quasi sinusoïdale dont la fréquence correspond à la fréquence de défaut de la bague déversée et dont le spectre présente un nombre restreint d'harmoniques de très faibles amplitudes face à celle de la composante fondamentale.

1.5.5.4 Altération due au passage de courants de fuite

L'apparition de cannelures affectant les pistes d'un roulement induite par le passage de courants de fuite se traduit par la présence d'un peigne de raies dont le pas correspond le plus souvent à la fréquence de défaut de bague externe, même si souvent les cannelures affectent à la fois les bagues externe et interne. L'amplitude du peigne de raies dans la bande fréquentielle [1500 - 5000 Hz] est très élevée alors qu'elle est très faible dans la bande [0-1500 Hz] (figure 1.33).

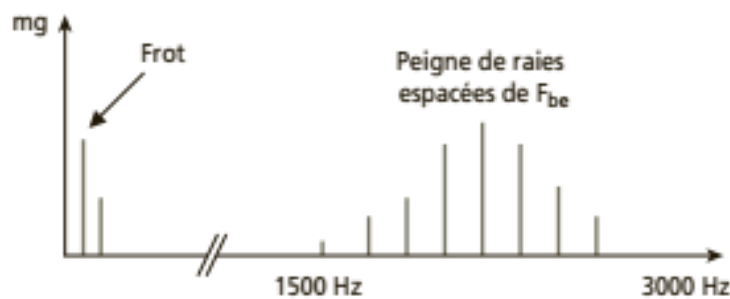


Figure 1.33 : Spectre d'une altération de bague de roulement par électroérosion [3]

1.5.5.5 Jeux de paliers

Les typologies vibratoires liées à des jeux de palier se caractérisent selon l'importance du défaut par :

- l'excitation des résonances hautes fréquences des différents éléments constitutifs du roulement et palier par chocs. La fréquence de répétition correspond à la fréquence de rotation et à la fréquence de cage dans le cas d'un jeu interne conséquent entre éléments roulants et cage (usure ou rupture de cage). Cependant, si la bague tournante ou les éléments tournants présentent des altérations, nous avons vu précédemment que l'augmentation du jeu fonctionnel interne se manifestait par une diminution de l'étendue de la zone de charge avec, pour conséquence, une forte augmentation du nombre de bandes latérales associées à la fréquence de défaut de l'élément altéré. Globalement, ces

excitations de résonances se traduisent par une forte élévation de la valeur efficace et crête de l'accélération hautes fréquences.

- un peigne de raies basses et moyennes fréquences plus ou moins étendu de pas correspondant à la fréquence de rotation, dû à la troncature par la fonction de transfert de la force d'excitation sinusoïdale induite par le déséquilibre résiduel (ou non) de la ligne d'arbres.

1.5.5.6 Rupture de la cage

L'évolution rapide de ce type de défaut en rend la surveillance difficile. Dans la plupart des cas, un roulement dont la cage est cassée va très vite se détériorer. Les éléments roulants (billes) n'étant plus maintenus, ils vont se regrouper et créer un fléchissement de l'arbre, entraînant ainsi le blocage immédiat du rotor.

Le peigne de raies basses fréquences associé à des chocs induits par la rotation de la cage passe généralement inaperçu du fait qu'il est le plus souvent noyé dans le bruit de fond. Par contre, l'existence de ces chocs pourra plus facilement être détectée dans les spectres « enveloppe » issus de la démodulation des réponses des modes de déformation du roulement. La rupture de la cage étant souvent la conséquence d'efforts dynamiques axiaux trop importants (mauvais serrage du roulement, mauvais alignement des courroies de transmission ...), la meilleure façon de prévenir ce défaut est de surveiller les vibrations des paliers en direction axiale.

1.6 CONCLUSION

La maintenance des installations industrielles se dirige de nos jours principalement vers une maintenance préventive conditionnelle. Nous nous intéressons particulièrement au cas des machines tournantes dont il est indispensable de suivre l'état des organes assurant les transmissions de puissance les plus sujets à l'usure : les engrenages, les courroies, ... et particulièrement les roulements.

À travers ce chapitre, une liste non exhaustive des principales défaillances rencontrées dans les machines tournantes, a été établie avec un accent particulier mis sur la nature de ces défaillances, leurs origines et leurs signatures vibratoires caractéristiques. La modélisation de roulement nous renseigne sur les considérations statiques qui seront utilisés pour comprendre ultérieurement, le modèle dynamique simplifié permettant de déterminer l'évolution de la réponse dynamique au cours du temps, due aux excitations causées par le passage de billes sur la piste de roulement en présence de défauts.

Chapitre 2

METHODES DE DIAGNOSTIC DES DEFAUTS DE ROULEMENTS – ETAT DE L'ART

2.1 INTRODUCTION

La détection de défauts nécessite d'une part une prise de mesure du signal vibratoire et d'autre part une exploitation du signal recueilli. La prise de mesure est une étape essentielle dans la procédure d'analyse vibratoire car elle conditionne l'efficacité des méthodes de traitement mises en œuvre. Le choix du ou des point(s) de mesure, le mode de fixation des capteurs, les paramètres de l'appareil de mesure (gamme de fréquence, résolution spectrale) sont des paramètres à prendre en considération.

Dans ce chapitre, on passe en revue, les méthodes de traitement des signaux vibratoires en vue du diagnostic des roulements. Sont présentées en premier, les méthodes dites classiques que l'on retrouve dans la plupart des analyseurs de signaux commercialisés (analyse temporelle, l'analyse de Fourier, l'analyse d'enveloppe, cepstrale, temps fréquence, etc.). Ensuite, les méthodes dites avancées enchaînent avec celles basées sur les statistiques d'ordres supérieurs. En dernier, on termine avec les méthodes issues de l'intelligence artificielle, dont l'une d'entre-elles, a fait l'objet des travaux de cette étude.

2.2 METHODES D'ANALYSE TEMPORELLES

Ces méthodes utilisent des indicateurs scalaires qui permettent de suivre l'évolution d'une grandeur décrivant de la puissance ou de l'amplitude crête du signal. Sa valeur peut ne pas avoir de signification intrinsèque, c'est son évolution dans le temps qui est significative du défaut.

2.2.1 Indicateurs statistiques

L'indicateur statistique résume l'information contenue dans le signal de vibration extrait à partir du capteur (accéléromètre) et permet de générer un repère. Le suivi de l'évolution du repère donne une information plus ou moins pertinente, sur l'état de fonctionnement de l'équipement sous surveillance. Cependant, l'utilisation de ces indicateurs globaux dans le cadre du diagnostic, nécessite d'être accompagnée par des investigations plus poussées.

De nombreux indicateurs existent dans la littérature, plus ou moins adaptés au dépistage de certaines défaillances plus que d'autres. Dans les applications aux défauts de roulements, les indicateurs les plus utilisés sont décrits ci-après [36 - 38].

La valeur efficace (RMS) nommée aussi moyenne quadratique d'un signal, correspond à la racine carré du moment d'ordre deux et donnée par :

$$x_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N [x(n)]^2} \quad (2.1)$$

Où,

N est le nombre de points du signal numérique.

Grâce à sa simplicité, cet indicateur compte parmi les premiers utilisés en industrie. Une variation excessive de sa valeur, signifie un changement de l’état de fonctionnement et donc une présence de défaillance. Cependant, et particulièrement dans le cas des défauts de roulements où la variation du signal due à l’apparition du défaut est masquée par d’autres composantes d’amplitudes plus élevées, il donne généralement une alarme assez tardive.

L’**asymétrie (Skewness)** correspond à une mesure de l’asymétrie de la distribution d’une variable aléatoire réelle. Ce coefficient de dissymétrie qui est aussi le moment statistique centré d’ordre 3 (M_3), est donné par :

$$x_{SKEW} = \frac{1}{(N-1)\sigma^3} \sum_{n=1}^N [x(n) - \bar{x}]^3 \quad (2.2)$$

Où,

σ est l’écart type.

Le **Kurtosis** particulièrement adapté au dépistage des défauts de roulements, le *kurtosis* est une grandeur statistique permettant de caractériser le degré d’aplatissement d’une distribution. Dérivé du moment statistique d’ordre quatre, il est défini comme le rapport de la valeur moyenne du signal élevée à la puissance 4 sur le carré de son énergie. Dans le cas d’un roulement sans défaut, la distribution des amplitudes du signal recueilli est gaussienne, avec une valeur du kurtosis proche de 3, [39]. Ce paramètre est très utilisé dans la détection précoce d’un défaut de roulement [40, 41]. Parfois, et pour améliorer la qualité d’analyse fournie par le kurtosis, des méthodes de filtrage du signal vibratoire peuvent être mises en place [5, 42]. L’expression du kurtosis est donnée par :

$$x_{Ku} = \frac{\frac{1}{N} \sum_{n=1}^N [x(n) - \bar{x}]^4}{\left[\frac{1}{N} \sum_{n=1}^N [x(n) - \bar{x}]^2 \right]^2} = \frac{M_4}{M_2^2} \quad (2.3)$$

Il est à noter que des indicateurs basés sur des moments statistiques d’ordres supérieurs (M_k avec $k \geq 5$), ont été utilisés. D’autres encore, tels que le facteur de crête, le facteur d’impulsion, le facteur de clairance ou le facteur de forme qui sont basés sur des combinaisons d’indicateurs classiques, ont été également utilisés dans la littérature [43 - 46].

2.2.2 Moyennage temporel synchrone

C’est une technique de traitement du signal qui peut être utilisée pour extraire des signaux répétitifs du bruit additif. Elle nécessite soit une connaissance précise de la fréquence de répétition du signal souhaité, s’il est périodique, soit un second signal synchrone avec le premier mais exempt de bruit. En utilisant soit la fréquence de répétition soit le signal synchrone, des blocs successifs du signal bruyant peuvent être échantillonnés et moyennés ensemble. Lorsque des moyennes suffisantes sont prises, on constate que le bruit dans la moyenne d’ensemble a tendance à s’annuler, laissant une meilleure estimation du signal souhaité [6, 47, 48]. Le moyennage peut être modélisé par la convolution du signal bruité avec un train fini d’impulsions

dans lequel le temps entre les impulsions est égal à la période du signal souhaité. Par définition, il est donné par l'expression :

$$a(t) = \frac{1}{N} \sum_{n=0}^{N-1} y(t + nT) \quad (2.4)$$

Où,

$$a(t) = c(t) * y(t) ;$$

$c(t) = \frac{1}{N} \sum_{n=0}^{N-1} \delta(t + nT)$ Fonction représentative de N impulsions idéales d'amplitudes $\frac{1}{N}$ espacées à des intervalles de T.

$y(t) = x(t) + e(t)$ est le signal bruité ;

$x(t)$ est un signal périodique de période T ;

$e(t)$ est la composante du bruit additif.

Ce processus est équivalent dans le domaine fréquentiel à la multiplication de la transformée de Fourier du signal bruyant par un filtre en peigne, ne laissant passer que les composantes correspondantes aux fréquences fondamentales et harmoniques du signal souhaité.

Cette approche basée sur l'hypothèse du signal parfaitement stationnaire, a été appliquée dans de nombreuses études de suivi vibratoire des organes de machines tournantes. Dans le cas particulier des roulements, nous pouvons citer les travaux des références [49, 50].

2.2.3 Modèles paramétriques

Ce sont des approches de l'analyse temporelle qui appliquent des modèles de série chronologique au signal vibratoire considéré comme processus aléatoire. Dans ces modèles, l'évolution du signal vibratoire est interprétée par les valeurs passées dans une approche autorégressive du prédicteur. Les coefficients du filtre autorégressif sont ajustés de telle sorte à minimiser l'erreur entre le signal d'entrée et le signal filtré. Le modèle linéaire paramétrique ARMA, tiré d'un modèle plus générale dans l'identification des séries temporelles appelé modèle de Box et Jenkins ou un modèle non-linéaire de Prony [51], est un cas général.

Un processus stationnaire (signal temporel) x_t , suit un modèle ARMA d'ordre « p, q » et noté ARMA (p, q), s'il vérifie la relation suivante [52] :

$$\begin{cases} x_t - \sum_{k=1}^p a_k x_{t-k} = \varepsilon_t + \sum_{j=1}^q b_j \varepsilon_{t-j} \\ a_p \neq 0, b_q \neq 0 \\ \{\varepsilon_t\} \sim BB(0, \sigma_\varepsilon^2) \end{cases} \quad (2.5)$$

Avec,

a_1, a_2, \dots, a_p et b_1, b_2, \dots, b_q sont les paramètres du modèle ;

ε_t : un bruit blanc de moyenne nulle et de variance σ_ε^2 .

Un processus suit un modèle autorégressif AR d'ordre p , et noté AR (p), où l'on écrit une observation au temps t comme une combinaison linéaire des observations passées plus un certain bruit blanc. Il vérifie la relation suivante :

$$\begin{cases} x_t = \sum_{k=1}^p a_k x_{t-k} + \varepsilon_t \\ a_p \neq 0, \\ \{\varepsilon_t\} \sim BB(0, \sigma_\varepsilon^2) \end{cases} \quad (2.6)$$

Un processus est dit à moyenne mobile MA d'ordre q et noté MA (q), si celui-ci peut s'écrire sous la forme suivante :

$$\begin{cases} x_t = \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \\ b_q \neq 0 \\ \{\varepsilon_t\} \sim BB(0, \sigma_\varepsilon^2) \end{cases} \quad (2.7)$$

Les modèles, AR et MA représentent des cas particuliers du modèle ARMA avec $q=0$ et $p=0$. Ils ont été introduits pour la première fois en 1927 par respectivement, George Udny Yule et Eugen Slutsky. Depuis, ils ont démontré dans de nombreux domaines, qu'ils peuvent être parfaits dans certain cas, mais il peut y avoir un besoin d'estimer un grand nombre de paramètres pour ajuster le modèle. Si l'on dispose de peu d'observations, ces estimations auront tendance à être peu précises. De plus, si un modèle contenant p paramètres convient pour la situation, il n'est pas bon de tenter d'ajuster un modèle qui en contiendra plus que p . Cependant, Herman Wold [53], a montré que les modèles ARMA qui sont une combinaison d'une partie autorégressive et une autre à moyenne mobile, pouvaient être utilisés pour modéliser n'importe quelle série stationnaire pour autant que les ordres p et q soient bien choisis. Box et Jenkins [54], ont travaillé dans le développement de méthodologie d'estimation du modèle d'une série chronologique.

Bien que le modèle ARMA possède un avantage, dans la plupart des cas, on préfère recourir à la modélisation autorégressive notée AR(p). Si l'ordre p est connu, les paramètres du modèle (coefficients a_k) sont estimés par la méthode du maximum de vraisemblance ou la méthode des moindres carrés et leurs variantes. Ainsi, de nombreux travaux dans le diagnostic des défauts dans les organes de machines tournantes, sont cités dans la littérature. Mechefske et Mathew ont utilisé l'estimation paramétrique pour détecter et diagnostiquer des défauts dans les roulements à basse vitesse de rotation (< 100 tr/min) [55, 56]. Les modèles paramétriques ayant des ordres variant entre 20 et 60, sont utilisés pour générer des spectres de fréquence, une détection et un diagnostic réussis des défauts, obtenus à partir de longueurs de signal considérablement plus courtes que celles utilisées dans des procédures de surveillance conventionnelles. Ils ont démontré que la modélisation autorégressive fournit une approche alternative pour le diagnostic des défauts dans les roulements, à partir de quantités limitées de

données de vibration. Elle pourrait donc être idéale pour classer les défauts dans les machines à vitesse lente ou variable. Baillie et Mathew ont comparé les performances et la fiabilité de trois techniques de modélisation autorégressive dans le cas du diagnostic de roulements, pour différentes longueurs de signal [57].

2.3 METHODES D'ANALYSE FREQUENTIELLES

2.3.1 Analyse de Fourier

C'est l'une des méthodes les plus anciennes et les plus utilisées dans le domaine industriel, grâce à sa facilité de mise en œuvre et d'interprétation. Elle est devenue l'outil fondamental pour le traitement des signaux vibratoires, qui permet de connaître le contenu spectral du signal temporel et de localiser les fréquences caractéristiques de défauts des composants mécaniques dans une machine tournante [58]. La transformée de Fourier est définie par l'équation suivante :

$$X(f) = \int_{-\infty}^{+\infty} x(t). e^{-j2\pi ft} dt \quad (2.8)$$

Où,

$X(f)$ est la transformée de Fourier ;

t et f les variables respectives de temps et de fréquence.

En pratique, on utilise la transformée de Fourier discrète rapide FFT, sur des signaux numérisés. Elle est donnée par l'expression suivante :

$$X(k\Delta f) = \frac{1}{N} \sum_{n=0}^{N-1} x(nT_e). e^{-j2\pi k \frac{n}{N}} \quad (2.9)$$

Où,

$X(k\Delta f)$ est la transformée de Fourier discrète ;

T_e période d'échantillonnage du signal temporel ;

N nombre d'échantillons prélevés ;

Δf résolution en fréquence ;

k numéro de ligne fréquentielle.

L'analyse du spectre permet d'identifier les pics révélateurs de présence de défauts dans les composants de la machine tournante (roulement, engrenage, courroie, etc.). En effet, chacun de ces éléments tournants est caractérisé par une ou plusieurs fréquences cinématiques connues au préalable. Ces fréquences fondamentales dépendent entre autres, de la géométrie du composant et de sa vitesse de rotation. Cependant les fréquences de défauts calculées ne correspondent pas souvent aux fréquences qui apparaissent dans le spectre de vibration. Les raisons peuvent être attribuées, soit à une augmentation de charges sur le roulement qui va causer un changement de l'angle de contact, ou à une variation de la vitesse de rotation [59].

Par ailleurs, le contenu fréquentiel d'un signal à énergie finie peut être représenté par la densité spectrale d'énergie. Cette grandeur est le carré du spectre d'amplitude défini via la transformée de Fourier discrète, dans le cas de signaux discrets. Dans le cas de signaux bruités et donc aléatoires, on définit la densité spectrale de puissance (DSP), comme la transformée de Fourier de l'autocorrélation du signal. Parmi ses applications, on cite l'étude [60], destinée à identifier l'état du roulement en utilisant des indicateurs simples basés sur la DSP et les moments statistiques centraux du signal de vibration temporel, dans un processus de classification par les ANN et SVM. Il a été démontré qu'en utilisant ces indicateurs, l'état du roulement peut être classé avec une grande précision.

Il est à noter que l'utilisation quantitative directe des spectres obtenus par la transformée de Fourier comme indicateurs, n'est pas pratique en diagnostic des défauts dans les machines tournantes. Par contre, l'usage d'indicateurs extraits à partir des spectres présente un grand intérêt. En effet, les indicateurs spectraux sont sensibles aussi bien aux évolutions de la forme d'un signal qu'à celles de son énergie, et de ce fait, sont insensibles aux effets de masque à condition que les résolutions d'analyse choisies pour les élaborer soient en adéquation avec les fréquences de répétition des phénomènes recherchés. En plus d'être d'indicateurs de défauts, ils offrent la facilité de leur comparaison graphique par rapport à un état de référence, ce qui favorise l'interprétation de leurs évolutions. Parmi ces nombreuses caractéristiques statistiques fréquentielles, on cite le centre fréquentiel (Frequency centre) qui permet d'indiquer la concentration des fréquences dans le spectre, la valeur efficace fréquentielle (RMS frequency), la moyenne fréquentielle (Mean frequency) et l'écart type fréquentiel (Standard deviation frequency), entre autres [61 - 63].

2.3.2 Analyse d'enveloppe

L'historique de cette méthode d'après la revue de Howard [64], remonte au début des années 1970, où Darlow et al [65], l'avaient appliquée au diagnostic des défauts de roulements dans un réducteur d'hélicoptère, au profit du « Air Mobility Research and Development Laboratory ». Elle était à l'origine appelée la technique de résonance à haute fréquence HFRT. Depuis, elle a connue différentes appellations, notamment la démodulation d'amplitude, l'analyse de résonance démodulée, l'analyse d'enveloppe à bande étroite ou simplement l'analyse d'enveloppe, qui semble maintenant être la désignation la plus populaire.

Cette méthode est particulièrement adaptée au diagnostic des défauts de roulements. Elle permet la détection des chocs périodiques à partir des résonances de structure. La méthode de l'enveloppe utilise la modulation de l'amplitude de la fréquence de résonance du roulement,

par la fréquence du défaut. En effet, les chocs dus aux passages des éléments roulants sur les défauts de roulement excitent les hautes fréquences des structures, ainsi un phénomène de modulation se produit entre la fréquence de défaut (basse fréquence) et la résonance de structure (haute fréquence). Après la démodulation du signal, les raies d’impulsions dues au défaut apparaissent clairement dans le spectre d’enveloppe. D’une façon générale, les différentes étapes de la méthode sont élucidées dans la figure 2.1.

En passant en revue les applications antérieures de la méthode, McFadden et Smith [4] ont relevé notamment les difficultés liées aux estimations des fréquences de résonance (à cause de la grande variété des montages roulement-palier-structure), de la largeur de bande optimale au filtrage, de la nature des mesures (déplacement ou accélération) et de la position du défaut sur l’organe du roulement. Ils ont particulièrement mis l’accent sur la nécessité d’utiliser les fréquences de résonance mesurées et une largeur de bande du filtre qui doit être au moins deux fois et de préférence plus de quatre fois la fréquence de défaut caractéristique maximale attendue. Prashad et al. [66], ont mené des investigations en vue d’évaluer ses performances, à travers des comparaisons des fréquences de défauts déterminés théoriquement avec celles évaluées expérimentalement. Les fréquences mesurées et leurs niveaux d’énergie sont utilisés à la fois pour la surveillance et la gravité des défauts existants, dans différentes conditions de fonctionnement et à différentes vitesses.

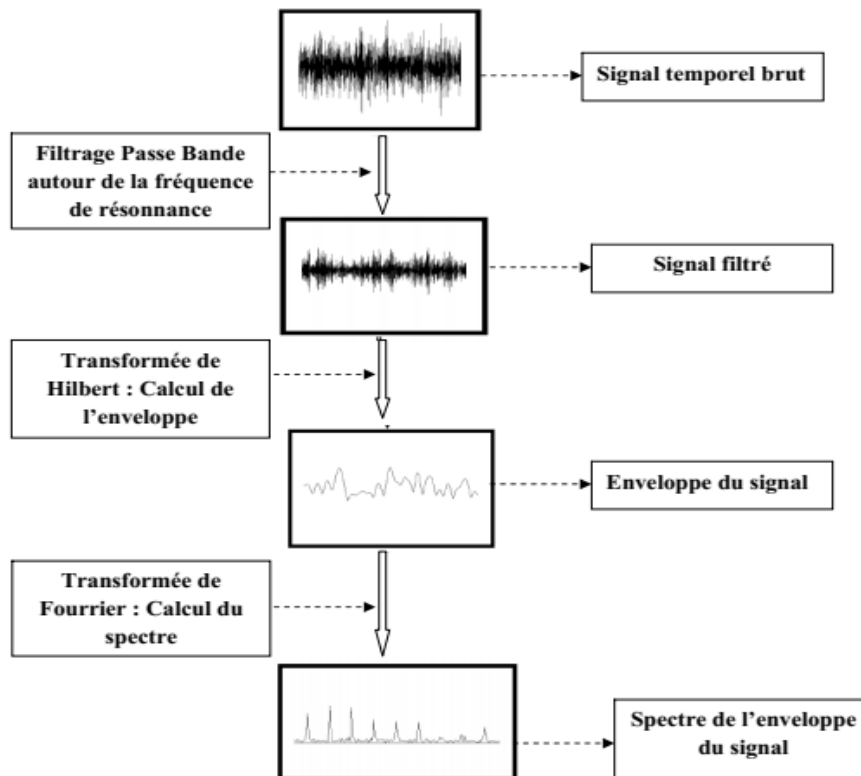


Figure 2.1 : Etapes de l’analyse d’enveloppe basée sur la transformée de Hilbert [39].

Tse et al. [67], ont comparés les résultats obtenus par la méthode d’enveloppe avec ceux des ondelettes, appliquées sur les roulements d’une motopompe. Certains défauts ont été révélés avec efficacité par les deux méthodes. Étant donné que l’analyse par les ondelettes peut détecter à la fois des signaux périodiques et non périodiques, elle offre donc un certain avantage dans la détection facile des défauts qui posent des difficultés à la méthode d’enveloppe basée sur l’analyse du spectre obtenu par FFT, exigeant donc en termes de périodicité. Sheen [68] a développé un algorithme d’estimation de l’enveloppe, basé sur les modes de résonance du système mécanique. La décroissance exponentielle de la fréquence du signal d’enveloppe pourrait en outre être estimée comme un indice pour quantifier le diagnostic de défaut de roulement. Le lecteur intéressé trouvera également dans certains travaux de Randall et al. [6, 69], de nombreux cas d’application de cette méthode, aux défauts de roulements.

2.3.3 Analyse cepstrale

Cette méthode constitue un outil de diagnostic destiné à distinguer des défauts qui donnent des images spectrales complexes dues à plusieurs modulations d’amplitude coexistantes. Elle est donc plus appréciée dans la surveillance des engrenages, où la fréquence d’engrènement est souvent modulée par les fréquences de rotation des roues menantes et menées. Ainsi, Le cepstre permet de séparer et d’identifier, sur une seule image, toutes les fréquences de modulation (fréquences de rotation des arbres d’entrée, intermédiaire, et de sortie dans un réducteur). Mathématiquement, le cepstre d’un signal temporel $x(t)$ est la transformée de Fourier inverse du logarithme décimal de sa transformée de Fourier directe :

$$C(q) = TF^{-1}\{\log_{10}[X(f)]\} \quad (2.10)$$

Dans une autre définition [70], il est donné par :

$$C(q) = TF^{-1}\{\log_{10}|X(f)|^2\} \quad (2.11)$$

Où,

$$X(f) = TF[x(t)]$$

$C(q)$ désigne le cepstre, anagramme du mot « spectre » ;

TF désigne la transformée de Fourier ;

TF^{-1} désigne la transformée de Fourier inverse ;

La variable q du cepstre a la dimension d’un temps et désigne la quéfrence, anagramme du mot fréquence. Elle représente les périodes d’oscillations des réponses impulsionnelles de la structure, et les périodes de répétition ou de modulation des forces d’excitation. Les pics du cepstre ainsi obtenus, s’appellent les « rhamoniques » et leurs positions sur l’axe des « quéfrences » indiquent les périodes

Bien que l’application de l’analyse cepstrale à la surveillance des engrenages soit très répandue dans la littérature, elle demeure relativement restreinte dans le cas des roulements. Dans ce contexte, Park et al [71], ont étudiés la détection des défauts précoces, dans les roulements à billes des roues automobiles. Ils ont introduit un cepstre de variance minimale (MVC) basé sur un spectre de puissance logarithmique, pour l'observation du signal d'impulsion périodique dans des environnements bruyants. Ils ont montré son efficacité à détecter les défauts naissants dans les roulements sous étude aux conditions de fonctionnement. Borghesani et al. [72], ont proposés une technique de pré-blanchiment (pre-whitening) dite de cepstrum, basée sur l'analyse cepstrale. Celle-ci permet d'éliminer les composantes déterministes, dans diverses applications industrielles caractérisées par des sources de vibrations hétérogènes, avant de procéder à l'analyse de l'enveloppe. Une comparaison des résultats obtenus avec ceux des techniques traditionnelles a montré les potentiels de la méthode en termes de simplicité et d'efficacité. Zhang et al. [73], ont combiné le pré-blanchiment du signal basé sur l'édition du cepstre et la résonance stochastique (SR), souvent utilisée dans la détection de signaux faibles, pour améliorer la détection des défauts de roulement. Le kurtosis du spectre d'enveloppe est utilisé comme indicateur des défauts de roulement. La méthode proposée a été validée après comparaison de ses résultats avec d'autres considérés de référence.

2.3.4 Méthodes de haute résolution HRM

A la base, ces méthodes avancées ont été développées particulièrement dans le domaine des télécommunications, afin de détecter les raies spectrales infiniment proches. Elles sont largement adoptées dans le diagnostic des machines à induction, et dans la quasi-totalité des études, elles exploitent les signaux de courants statoriques. Toutefois, les travaux de recherche effectués avec ces méthodes HRM, ne mettent pas en évidence les paramètres de précision, le niveau de robustesse de chaque approche et le temps de calcul écoulé. Ces techniques ont une résolution supérieure à la résolution de Fourier donnée par :

$$\Delta f = \frac{\text{Fréquence d'échantillonnage}}{\text{Taille du signal}}$$

Parmi les méthodes HRM, on cite, celles, de Prony, de Pisarenko, MUSIC, des vecteurs propres EV, de norme minimale MN et ESPRIT. La performance de ces méthodes indiquées, a été largement étudiée dans la littérature, et basée sur les critères de, détection de défauts, vitesse de calcul, précision et degré de dispersion d'estimation fréquentielle, pour différents niveaux du rapport signal sur bruit SNR et pour des valeurs fixes d'amplitude de défauts. Les méthodes Root-MUSIC et ESPRIT deviennent plus précises lorsque le rapport SNR augmente, mais

exigeantes en termes des temps de calcul relativement élevés [74]. Ces deux estimateurs sont issus de la théorie des sous-espaces (la décomposition de la matrice d'autocovariance), l'un exploite le sous-espace bruit comme MUSIC ou Root-MUSIC et l'autre exploite le sous-espace signal comme ESPRIT. Il existe encore d'autres méthodes dans la littérature, basées sur la minimisation de fonctions objectives données, telles que MODE et PLEDGE, entre autres.

L'usage des méthodes HRM dans le diagnostic des défauts de roulement, basé sur les signaux vibratoires est relativement rare. Agrawal et al. [75], ont combiné la transformée de Hilbert avec deux méthodes de haute résolution. Ainsi, deux variantes ont été explorées. Dans la première variante, elle est associée à la méthode ESPRIT, et dans la seconde, elle est associée à une autre dite Root-MUSIC. Pour l'extraction des fréquences de défaut de roulement du moteur à induction à partir du signal vibratoire, une fenêtre glissante est introduite pour lutter contre les effets de variation, de la charge et de la fréquence d'alimentation. Une série de résultats montre que les deux variantes, Hilbert & ESPRIT et Hilbert & Root-MUSIC, sont capables d'extraire les fréquences des défauts de roulement avec précision, avec un léger désavantage à la première, en termes de temps de calcul. Garcia-Perez et al. [76], ont proposé une méthode de surveillance des défauts simples et multiples, spécifiques aux moteurs à induction. Elle combine une banque de filtres à réponse impulsionnelle finie avec un algorithme d'analyse spectrale de haute résolution MUSIC, pour une identification précise du défaut lié à la fréquence. En général, les résultats obtenus montrent une grande concordance avec les valeurs des fréquences de défaut uniques ou combinés, prédéterminées analytiquement.

2.4 METHODES D'ANALYSE TEMPS-FREQUENCE

D'une manière générale, l'analyse de Fourier donne de bons résultats dans les cas où les signaux sont stationnaires (vibrations dues, aux balourds, au désalignement, etc.). Cependant, elle devient souvent inefficace dans les cas où les signaux sont non stationnaires composés de chocs multiples et des signaux transitoires dont la durée dépend de la nature du choc (roulement, engrenage, moteur à combustion internes, desserrage de paliers, etc.). Pour remédier à ces limitations, les représentations temps-fréquence ont été développées dans le but de décrire conjointement le contenu des signaux en temps et en fréquence, et mettre ainsi en évidence, certaines caractéristiques non-stationnaires. Elles sont classiquement issues de la classe de Cohen ou du groupe affine (classe affine et transformée en ondelettes) [77, 78]. Elles offrent l'avantage d'une représentation tridimensionnelle (amplitude-temps-fréquence). Ainsi, on peut superviser, le suivi, la détection et l'évolution des défauts qui génèrent une faible puissance

vibratoire. Le choix du type de représentation dépend souvent d'une alternative entre propriétés théoriques et lisibilité de la représentation.

2.4.1 Transformée de Fourier à court terme STFT

Afin d'échapper à l'inconvénient du caractère global de la transformée de Fourier, la transformée de Fourier à court terme consiste à opérer à la fois, seulement sur une portion du signal supposée localement stationnaire, grâce à une fenêtre glissante. Ainsi, on calcule la transformée de Fourier des segments limités par la fenêtre dans son glissement le long de toutes ses positions temporelles [79].

La transformée de Fourier à court terme STFT est donnée par :

$$F_x(t, f) = \int_{-\infty}^{\infty} x(t + \tau) w(\tau) e^{-j2\pi f\tau} d\tau \quad (2.12)$$

Avec,

$x(t)$ est le signal complexe quelconque ;

$w(\tau)$ est la fenêtre à valeur réelle

Le spectrogramme du signal $x(t)$ est estimé par le carré du module de la STFT, soit l'expression :

$$S_x(t, f) = |F_x(t, f)|^2 \quad (2.13)$$

Cependant, le spectrogramme a une certaine limitation dans la résolution de temps-fréquence due à la segmentation de signal. Il peut seulement être appliqué aux signaux lentement variables.

Un certain nombre de travaux basés sur la STFT, en tant que méthode d'analyse temps-fréquence, ont été introduits. Parmi ces applications, son usage dans l'identification des composantes fréquentielles caractéristiques dans le cas des vitesses variables, dans [80]. L'effet de modulation des signaux de vibration des roulements rend difficile une telle identification. Ainsi, Liu et al. ont proposé un procédé basé sur la STFT et un algorithme d'estimation de fréquence instantanée, moyennant un filtrage passe bande par Kurtogramme. Ensuite, la fréquence caractéristique de défaut instantanée est extraite de la représentation temps-fréquence d'enveloppe en adoptant conjointement la transformée de Hilbert et la STFT au signal filtré. Arun Kumar et al. [81], ont proposé un diagnostic du défaut de roulement d'une machine, en utilisant l'algorithme STFT et un arbre de décision basé sur l'algorithme nommé « J48 ». La STFT peut être utilisée afin d'identifier les défauts dans le roulement à partir du signal de vibration acquis. L'analyse du signal de vibration est réalisée dans Lab VIEW. Une sélection des indicateurs permet de réaliser un apprentissage automatique qui aboutit à une meilleure optimisation des paramètres du classifieur indiqué. Les résultats obtenus ont atteint une précision de classification de 94,5%.

2.4.2 Distribution de Wigner-Ville WVD

C’est l’outil probablement le plus largement utilisé dans la classe de Cohen pour l’analyse temps-fréquence. Cette distribution permet de décrire comment le contenu spectral d’un signal évolue dans le temps. Elle possède une bonne résolution temps-fréquence et un nombre important de propriétés théoriques (marginales en temps et en fréquence, conservation des supports temporel et fréquentiel, ...) [82]. Elle est bilinéaire (signal multiplié par lui-même), et définie par :

$$w_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \quad (2.14)$$

Avec,

$x(t)$ est le signal complexe quelconque ;

$w_x(t, f)$ est la valeur de la distribution qui représente la mesure de la corrélation du signal avec sa version décalée en fréquence et considérée selon l’axe du temps inverse.

Comparée à la STFT, la distribution de Wigner-Ville n’est pas basée sur la segmentation du signal. Elle ne souffre donc pas de la limitation de résolution de temps-fréquence du spectrogramme. Néanmoins, sa structure quadratique conduit à la présence de termes d’interférence, ne permettent pas une interprétation simple de la distribution d’énergie, à cause de la lisibilité réduite dans le cas de signaux à grand nombre de composantes.

Singru et al. [83], ont utilisé la cartographie de Poincaré modifiée et la WVD pour détecter la défaillance des roulements. Ils ont conclu que l’association des deux méthodes peut permettre la prédiction de l’emplacement et de la gravité du défaut de roulement. Zhou et al. [84], ont présenté la WVD basée sur la densité spectrale cyclique (CSD). Ils ont démontré que son application à une longue série temporelle, peut rendre la distribution temps-fréquence moins sensible au bruit et restreindre les termes croisés dans le domaine temps-fréquence. La validation de l’approche, est effectuée aux moyens de simulations et essais expérimentaux.

2.4.3 Distribution de Choi-Williams

Pour remédier à l’inconvénient principal de la distribution de Wigner Ville, dû à la présence de termes croisés, qui peut engendrer des interprétations trompeuses dans le cas des applications sur des signaux à composantes multiples, un certain nombre de versions de cette représentation ont été élaborées. Parmi celles-ci, on cite : la pseudo Wigner Ville, la Pseudo Wigner Ville lissée, les représentations, de Rihaczek, de Cohen, de Choi-Williams, etc.

La distribution de Choi-Williams a été proposée pour la première fois par Hyung-Il Choi et William J. Williams en 1989. Pour la suppression des termes croisés, elle adopte un noyau exponentiel [85]. Ce dernier est donné par la fonction :

$$f(t, \tau) = \frac{\sqrt{\pi\sigma}}{|\tau|} e^{-\pi^2 \sigma t^2 / \tau^2} \quad (2.15)$$

La distribution de Choï-Williams est donnée par :

$$P_x(t, \tau) = \iint \frac{\sqrt{\pi\sigma}}{|\tau|} e^{-\frac{\pi^2 \sigma (t-\mu)^2}{\tau^2}} x\left(\mu + \frac{\tau}{2}\right) x^*\left(\mu - \frac{\tau}{2}\right) e^{-j2\pi f \tau} d\mu d\tau \quad (2.16)$$

Wei et al. [86], ont combiné l’EMD avec la distribution de Choï-Williams, pour le diagnostic des défauts de roulement. En premier, les signaux originaux ont été décomposés en IMFs de différentes échelles. Ensuite, l’analyse temps-fréquence Choï-Williams est appliquée à chaque IMF. Après, une superposition linéaire de l’ensemble permet de retrouver le signal de vibration du roulement de la distribution de Choï-Williams. Les auteurs assurent que cette méthode peut supprimer efficacement le repliement de fréquence et les interférences causées par les termes croisés, ainsi que l’extraction précise des fréquences des éléments roulants du roulement. Gu et al. [87], ont développés une analyse de signal non stationnaire, basée sur la décomposition en valeurs singulières et l’analyse temps-fréquence Choï-Williams, pour la surveillance on-line de l’état transitoire d’une machine. Grâce aux essais expérimentaux sur le processus d’usinage et la surveillance de l’état des roulements, l’efficacité de la technique proposée est démontrée.

2.5 KURTOSIS SPECTRAL ET KURTOGRAMME

La notion de Kurtosis spectral est introduite pour la première fois comme étant un Kurtosis dans le domaine fréquentiel pour détecter et caractériser les évènements transitoires dans un signal. Antoni [88] a proposé une définition et une formalisation compréhensible du Kurtosis spectral pour les processus non stationnaires. L’étude présente plusieurs exemples d’application du Kurtosis spectral pour la détection et le diagnostic des défauts. A partir de cette définition, plusieurs auteurs ont utilisé le Kurtosis spectral pour la détection des défauts, ces derniers étant considérés comme des réalisations non-stationnaires.

Le Kurtogram est défini initialement pour généraliser le Kurtosis spectral pour la détection des composantes non stationnaires sans connaissance des paramètres du signal (Blind Detection). Parmi ses avantages, la possibilité d’identification des fréquences caractéristiques de défauts.

2.6 ANALYSE TEMPS- ECHELLE

La transformée en ondelettes est, l’un des outils mathématiques les plus utilisés dans l’analyse de signaux non-stationnaires et transitoires. Les ondelettes continues sont plus adaptées à l’analyse temps-fréquence et à la visualisation, tandis que les ondelettes discrètes sont utiles en décomposition, compression et pour l’extraction des indicateurs.

2.6.1 Transformée en ondelettes continues

La transformée en ondelette est une transformation linéaire qui décompose un signal en fonctions élémentaires localisées en temps et appelées ondelettes, obtenues par translation et dilatation de l’ondelette mère. Une infinité d’ondelettes existent, mais rien n’empêche de créer sa propre ondelette [89, 90]. Le choix de l’ondelette peut être délicat mais très important car il conditionne la qualité des résultats obtenus. Une transformée en ondelettes continues est définie dans le domaine temporel et dans le domaine fréquentiel. Il existe un grand nombre de fonctions réelles et complexe qui peuvent être utilisées comme ondelettes mères. Le choix de l’ondelette mère optimale dépend de l’application envisagée.

Un avantage principal de la transformation par ondelettes est sa capacité de produire une très grande résolution fréquentielle à de basses fréquences et une très grande résolution temporelle à de hautes fréquences pour des signaux longue durée avec de basses fréquences et de courte durée avec hautes fréquences. Un autre avantage de la transformée en ondelettes est sa capacité de réduire le bruit des signaux.

2.6.2 Transformée en ondelettes discrètes

Dans le cas de la transformée en ondelettes continues, les paramètres de dilatation et d’échelle, prennent une infinité de valeurs. Il est possible de limiter le nombre de coefficients sans perdre d’information sur le signal de départ. On parle alors de transformée en ondelettes discrètes (DWT).

La transformation en ondelettes discrète est réalisée en utilisant des ondelettes orthogonales. L’ondelette orthogonale la plus simple est celle de HAAR. Le développement notable des ondelettes orthogonales est, attribué à la famille d’ondelettes proposées par Daubchies en 1992. Celles-ci ne peuvent pas être représentées dans une forme mathématique car elles résultent des fonctions d’échelles qui correspondent à des réponses des filtres d’impulsions. Un grand nombre d’applications au diagnostic des roulements, existe dans la littérature [91].

2.6.3 Paquets d’ondelettes

La décomposition en paquets d’ondelettes (WPD) est similaire à la transformation en ondelettes discrètes (DWT). Toutes les deux sont classées comme méthodes d’analyse multi-résolution. Cependant, la différence entre elles est que la WPD décompose simultanément les versions des détails et les approximations, tandis que la DWT décompose uniquement les

approximations. A ajouter aussi que la WPD a la même largeur de bande fréquentielle dans chaque résolution, propriété qu’on ne trouve pas en transformée en ondelettes discrète. Enfin, il est à signaler que le choix de niveau de décomposition en paquet d’ondelettes constitue un paramètre important dans la phase d’extraction d’indicateurs.

Parmi les nombreux travaux de la littérature basés sur les ondelettes, on cite ceux de Deriche [92], qui a montré que l’analyse des ondelettes est une approche efficace pour surveiller l’état du roulement. Dans son étude, il présente spécifiquement, la transformée en ondelettes discrète et la transformée en paquets d’ondelettes discrètes, ainsi que le choix de l’ondelette mère et les différentes caractéristiques qui peuvent être extraites de ces transformées. Son constat est que les informations contenues dans les emplacements des pics dans le spectre peuvent être utilisées efficacement dans la détection d’un défaut de roulement à billes. Pour le diagnostic de la localisation des défauts et de leur taille, le RMS extrait des nœuds terminaux à partir d’un arbre d’ondelettes se révèle être une caractéristique robuste. Le choix de l’ondelette mère n’est pas aussi critique et peut être : db4, db5, db6, sym4, sym5 ou sym6. Nikolaou et antoniadis [93], ont proposé une démarche basée sur l’analyse des signaux de vibration de roulements avec défauts localisés, en utilisant la transformée de paquets en ondelettes (WPT) comme outil systématique. Les composantes portant les informations de diagnostic importantes et issues de la décomposition temps-fréquence des signaux de vibration, sont sélectionnées pour l’exploitation des concepts physiques sous-jacents du mécanisme de modulation, présents dans la réponse aux vibrations des roulements défectueux. La méthode a été évaluée à l’aide de signaux simulés et expérimentaux.

2.7 ANALYSE PAR LES METHODES DE DECOMPOSITION

2.7.1 Décomposition en mode empirique EMD

La décomposition en mode empirique, EMD (Empirical Mode Decomposition) a été proposée par Huang [94], pour l’analyse des signaux non stationnaires. Contrairement à la représentation temps-fréquence ou aux ondelettes, la base de décomposition de l’EMD est intrinsèque au signal. L’extraction des composantes oscillantes appelées modes empiriques est non-linéaire, mais leur recombinaison est linéaire.

Comme l’EMD n’a pas de formulation analytique, elle est définie par un algorithme et par un processus appelé tamisage, permettant de décomposer le signal en modes empiriques ou IMFs (Intrinsic mode functions), selon la procédure suivante :

1. Initialisation : $r = f, k = 1$

2. Calcul de l’enveloppe moyenne e de r (moyenne de l’enveloppe des minima et de l’enveloppe des maxima de r).
3. Extraction de fonctions intermédiaires $p_i = r - e$ et posons $r = e$.
4. Tant que p_i n’est pas un IMF répéter
 - Calcul de l’enveloppe moyenne e de p_i
 - $p_{i+1} = p_i - e_i ; i = i + 1$
5. $d_k = p_i , r = r - d_k$
6. Si r n’est pas monotone, retour à l’étape 2 et posons $k = k + 1$, sinon la décomposition est terminée.

Cette décomposition est locale, itérative, séquentielle et entièrement pilotée par les données [31]. L’EMD a souvent été utilisée comme un filtre passe bande, pour déterminer avec précision les IMFs qui contiennent les fréquences de résonance avant d’appliquer l’analyse d’enveloppe pour la détection des défauts multiples de roulements.

L’usage de l’EMD est très répandu dans la littérature, mais on se restreint de décrire deux méthodes dans ce contexte. Tavakkoli et Teshnehlab [95] ont proposé une méthode de diagnostic de défaut de roulement à billes basée sur l’ondelette et la décomposition en mode empirique (EMD). Après la décomposition des signaux vibratoires en IMF et composantes d’ondelettes, l’extraction de la moyenne d’énergie d’entropie de différentes bandes de fréquences a été opérée pour identifier les défauts de roulements à billes. Zhang et al. [96], se sont intéressés à la détection précoce des défauts et l’extraction des caractéristiques de dégradation pour les roulements. L’algorithme EMD révisé, est utilisé pour débruiter le signal porteur de défaut avant le traitement par la déconvolution de kurtosis corrélée maximale (MCKD). L’analyse de la dégradation, est assurée par l’usage de la valeur de l’aplatissement corrélé (CK) comme indicateur qui reflète l’état de santé du roulement.

2.7.2 Décomposition en mode empirique d’ensemble EEMD

Il est bien connu que l’EMD souffre d’un problème de mélange de modes, de sorte que les IMF individuels ont tendance à être mélangés ensemble [97]. L’EEMD est proposé pour pallier au problème du mélange de modes en insérant du bruit gaussien dans le signal d’origine avant de passer à l’EMD, et en itérant le processus de moyenne hors des IMF pour éliminer les modes mixtes [98]. Etant donné un signal $x(t)$, le principe de l’EEMD est décrit comme suit :

1. Générer N réalisations $n_i(t)$, $1 \leq i \leq N$ de bruit blanc gaussien de variance σ^2
2. Calculer N jeux d’IMFs $d_k^{(i)}(t)$ $1 \leq i \leq N$ à partir des N signaux $x(t) + n_i(t)$
3. Les IMFs EEMD sont alors les moyennes de l’ensemble des IMFs précédents

$$d_k^{EEMD}(t) = \frac{1}{N} \sum_{i=1}^N d_k^i + r \quad (2.17)$$

Depuis la présentation de l’EEMD par Wu et Huang [99] en 2009, de nombreux travaux sont apparus, parmi lesquels on cite l’étude de Lee et al. [100], qui ont proposé la détection et le diagnostic de défauts pour les systèmes de roulements en utilisant l'extraction d'indicateurs, basée sur l’EEMD utilisée conjointement avec l'optimisation de l'essaim de particules (PSO), l'analyse en composantes principales (ACP) et Isomap. Des signaux simulant les différents types de défauts sont générés et utilisés pour la mise au point de cette méthode. Dans [101], une méthode combinant la décomposition en mode empirique d'ensemble (EEMD) et l’algorithme nommé Random Forest (RF) est proposée. En premier, le signal original est décomposé par EEMD, puis les IMF efficaces sont sélectionnés. Ensuite, leur énergie d’entropie est calculée en tant que attribut. Enfin, la classification est effectuée par RF. Dans une seconde étape, le même processus a été répété avec la méthode des ondelettes. Selon les auteurs, la comparaison des résultats a montré que la méthode EEMD est plus précise que la méthode par ondelettes.

2.7.3 Décomposition en modes variationnels VMD

La décomposition en mode variationnel (VMD) est une méthode récente de décomposition du signal basée sur le filtrage de Wiener, la transformation de Hilbert unidimensionnelle et la démodulation hétérodyne [102]. Elle définit la composante de mode comme des signaux modulés en amplitude et en fréquence comme suit :

$$u_k(t) = A_k(t) \cos(\phi_k(t)) \quad (2.18)$$

Où,

$\phi_k(t)$ est une fonction non décroissante ;

$\phi_k'(t) \geq 0 = \omega_k(t)$ est la fréquence instantanée.

L’enveloppe est non négative, $A_k(t) \geq 0$.

La décomposition en mode variationnel VMD est adaptative et qui peut réaliser une analyse de signal non linéaire et non stationnaire. Elle résout les problèmes de l’aliasing modal, d'effets de bord et de sensibilité au bruit, causés par la décomposition EMD. L'essence de la méthode VMD est une pluralité de jeux de filtres Wiener adaptatifs. En termes de séparation modale, VMD détermine le centre de fréquence et la bande passante de chaque composant en recherchant de manière itérative la solution optimale du modèle variationnel, réalisant de manière adaptative la division du domaine de fréquence du signal et la séparation effective du VMD de chaque composant.

Pour un signal d'entrée f , le but ultime de la méthode de décomposition VMD est de décomposer ce dernier en un certain nombre de sous-signaux u_k , qui non seulement

reproduisent l'entrée, mais assurent également la parcimonie. Les étapes nécessaires pour construire un modèle variationnel sont :

1. Obtention du signal analytique de chaque fonction modale u_k par transformée de Hilbert
2. Correction de la fréquence centrale estimée respective $\{\omega_k\}$ de façon exponentielle
3. Effectuer un lissage gaussien sur le signal démodulé pour obtenir la bande passante de chaque segment. En supposant que le signal est décomposé en k modes par VMD, le modèle variationnel contraint est construit comme :

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 \right\} \quad (2.19)$$

$$\text{Sujet à } \sum_k u_k = f \quad (2.20)$$

Où,

$u_k = \{u_1, u_2, \dots, u_k\}$ sont les composantes modales ;

$\omega_k = \{\omega_1, \omega_2, \dots, \omega_k\}$ sont les fréquences centrales.

Afin d'obtenir la solution optimale du modèle variationnel, le facteur de pénalité α est introduit, et la fonction de Lagrange augmentée est construite :

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 + \|f(t) - \sum_k u_k(t)\| + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \quad (2.21)$$

Par conséquent, les multiplicateurs de Lagrange, les fonctions modales et leur fréquence centrale correspondante sont mis à jour de manière itérative en utilisant la méthode du multiplicateur de direction alternative (ADMM) pour obtenir le point de selle dans l'expression. L'algorithme spécifique de VMD classique est donné comme suit :

Etape 1 : Initialiser $\{\hat{u}_k^1\}, \{\omega_k^1\}, \lambda^1, n \leftarrow 0$

Etape 2 : Mise à jour de u_k, ω_k et λ selon les formules suivantes :

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i^{n+1}(\omega) + (\hat{\lambda}(\omega)/2)}{1 + 2\alpha(\omega - \omega_k)^2}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (2.22)$$

$$\lambda^{n+1} = \lambda^n + \tau(f - \sum_k u_k^{n+1})$$

Etape 3 : Répéter le processus itératif de (2) jusqu'à ce que la fonction converge, laquelle satisfait la condition de $\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < e$, avec e est la précision donnée.

L'usage de la méthode de décomposition VMD dans la littérature pour le diagnostic des défauts de roulement, est varié. Mohanty et al. [103], ont procédé à l'analyse du signal de vibration du roulement à billes en utilisant la décomposition en mode variationnel (VMD). Après l'extraction des fonctions de mode intrinsèque en premier temps, la Transformée de

Fourier rapide, a été appliquée pour statuer sur l'état du roulement (défectueux ou sain). Dans ce travail, un accent a été mis particulièrement sur les performances de l'algorithme VMD par rapport à l'EMD, en termes de ses qualités dans la détection de signatures de vibrations et sonores proches, et en temps de calcul réduit. Une autre étude [104], a proposé une approche destinée à la détermination du nombre de mode pour la méthode VMD en tirant parti des avantages du concept de similitudes entre la somme des fonctions de mode variationnel (VMF) et les signaux d'entrée. La validation de la méthode a été obtenue grâce à son application à des données de roulements et d'engrenages.

2.8 ANALYSE DE LA CYCLOSTATIONNARITE

Un signal est dit cyclostationnaire si ses propriétés statistiques d'ordre 1 (moyenne, variance) ou d'ordre 2 (fonction de corrélation) sont périodiques. D'une manière générale, l'analyse de la cyclo-stationnarité permet de révéler la présence de la modulation dans le signal vibratoire, ce qui permet de trouver :

- la fréquence modulante (fréquence cyclique) qui permet de trouver la signature de défaut.
- la fréquence de la porteuse qui dépend de la structure ou en d'autres termes la fonction de transfert entre la source des vibrations et le capteur. Cela facilite la détection des défauts de structure.

Antoni [105,106], a largement exploité cette propriété pour la détection de défauts de roulement, où les vibrations des éléments roulants sont considérés cyclostationnaires aléatoires. Il considère qu'elle est si symptomatique lorsqu'un défaut naissant se développe, qu'elle peut être exploitée à des fins de diagnostic. Parmi les outils spectraux cycliques, la cohérence cyclique (spectre de puissance cyclique normalisé en puissance) est la quantité spectrale cyclique qui offre la meilleure capacité de détection du défaut. Plus précisément, il est démontré que son optimalité peut non seulement prouver la présence d'un défaut dans des niveaux élevés de bruit de fond, mais peut également renvoyer une mesure relative de sa gravité.

2.9 METHODES BASEES SUR L'APPRENTISSAGE MACHINE (ML)

En pratique, les méthodes appliquées dans le diagnostic automatisée des défauts sont nombreuses, et ayant des typologies diverses et variées. Néanmoins, on distingue principalement deux groupes : le premier dit « systèmes expert » et basés sur les connaissances à priori (élaboré au préalable par des experts), le deuxième est basé sur les connaissances issues des observations. Les méthodes basées sur l'apprentissage machine (ML), font partie du deuxième groupe et font l'objet d'un usage très répandu dans le domaine de recherche en

diagnostic des défauts de machines tournantes. Elles ont prouvées leur efficacité face à la difficulté liée à la génération de modèles physiques représentatifs du fonctionnement d’un système et de ses différents composants, rencontrées par celles décrites dans le premier groupe.

Les méthodes basées sur l’apprentissage machine (ML) exploitent les informations délivrées par les signaux vibratoires des capteurs installés sur les organes de machines, pour construire une image représentative « instance, observation » et l’assigner à son mode de fonctionnement, par les techniques de classification adéquates. Ainsi, un schéma de classification général est réalisé globalement sur quatre étapes suivantes :

- Acquisition des signaux délivrés par les capteurs
- Extraction des indicateurs (attributs) appropriés
- Transformation et de sélection d’indicateurs (robustesse et pertinence)
- Construction des règles de décision

La dernière étape relative à l’établissement des règles de décision, peut être réalisée via l’utilisation des outils de classification offerte par l’intelligence artificielle et la fouille de données « *Data Mining* ». Parmi les outils de classification les plus populaires on cite à titre indicatif, les méthodes décrites ci-après.

2.9.1 Réseaux de neurones artificiels ANN

Ces modèles de réseaux ont été développés en s’inspirant des réseaux biologiques du cerveau humain. Ce sont des assemblages de neurones formels associés en couches fonctionnant en parallèle. L’information donnée au réseau va donc se propager de la couche d’entrée à la couche de sortie, en passant une ou plusieurs couches intermédiaires [107, 108]. Ils ont trouvé une très large application dans l’analyse des données.

2.9.1.1 Perceptron multicouche MLP

Il est l’un des réseaux de neurones les plus utilisés pour la classification de données. Il se compose d’une couche d’entrée, une ou plusieurs couches cachées et d’une couche de sortie. Le choix du nombre de neurones dans la couche cachée se fait en fonction du nombre de neurones dans les couches d’entrée et de sortie, pour améliorer les possibilités de généralisation et éviter les problèmes de sur-ajustement. Le calcul des sorties des neurones de la couche cachée, se fait en tenant compte des différentes entrées, de la fonction d’activation (qui est le plus souvent une fonction identité, sigmoïde, tangente hyperbolique ou une fonction linéaire par morceaux), du biais et du poids du neurone.

Dans le cas d'un réseau typique de MLP montré dans la figure 2.2, et composé de trois couches respectives : entrée, cachée et sortie, le vecteur d'entrée $\chi = [x_1, x_2, \dots, x_N]^T$, est transformé en un vecteur intermédiaire de variables cachées H par utilisation de fonction d'activation f_h . La sortie h_j du $j^{\text{ème}}$ neurone de la couche cachée comme indiqué sur la figure 2.3, est obtenue comme suit :

$$h_j = f_h\left(\sum_{i=1}^N w_{i,j}x_i + b_j\right) \quad (2.23)$$

Le vecteur de sortie désirée $Y = [y_1, y_2, \dots, y_M]^T$ du réseau est obtenu à partir du vecteur des variables intermédiaires H par une fonction d'activation f_o de la couche de sortie. Ainsi, la sortie du neurone k peut être exprimée comme suit :

$$y_k = f_o\left(\sum_{m=1}^M w_{m,k}h_m + b_k\right) \quad (2.24)$$

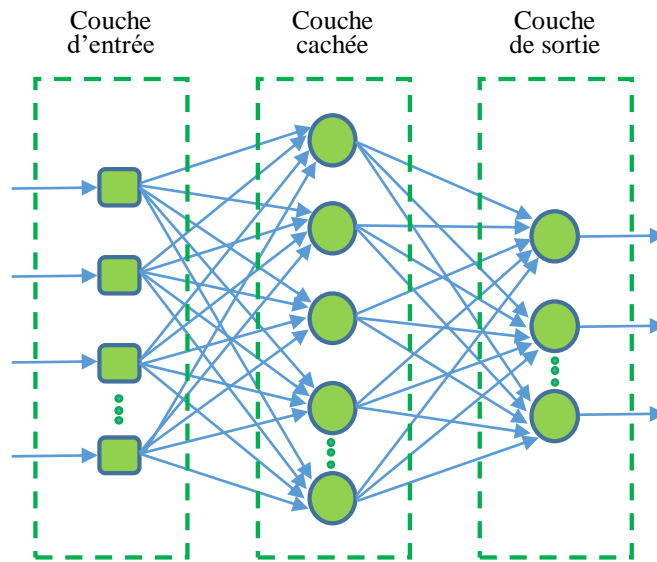


Figure 2.2 : Structure d'un réseau MLP

Les x_i représentent les vecteurs d'entrées. Les $w_{i,j}$ sont les poids synaptiques du neurone j . Le b_j est le biais qui permet d'ajouter de la flexibilité au réseau grâce à son ajustement (avec les poids) lors de l'apprentissage. Le noyau : intègre toutes les entrées et le biais et calcul la sortie du neurone selon une fonction d'activation qui est souvent non linéaire pour donner une plus grande flexibilité d'apprentissage.

Les fonctions d'activation f_h et f_o peuvent être sous différentes formes (sigmoïde, tangente hyperbolique, linéaire, etc.). On utilise aussi des fonctions «seuil» (en anglais «hard limit»). Comme son nom l'indique, la fonction seuil applique un seuil sur son entrée.

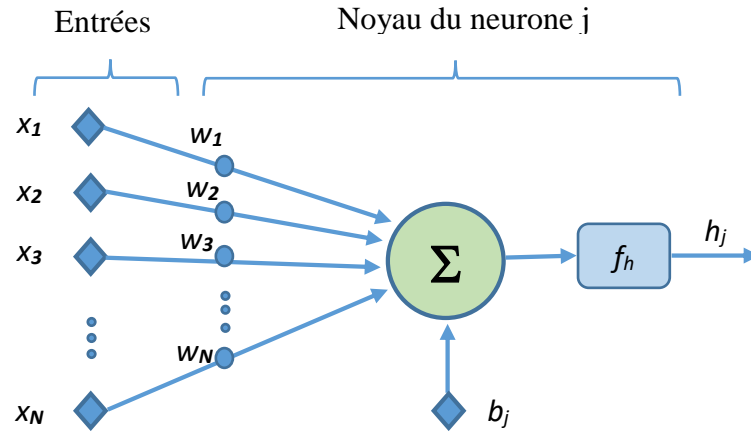


Figure 2.3 : Représentation graphique d'un neurone.

2.9.1.2 Réseaux à fonctions de base radiale RBF

La structure d'un réseau RBF est similaire à celle d'un MLP. C'est un réseau de type propagation directe (feedforward) avec une couche d'entrée, une couche cachée « une couche RBF » et une couche de sortie. Chaque couche est complètement connectée à sa suivante. Chaque neurone de la couche cachée contient une gaussienne centrée sur un point de l'espace d'entrée. Sa forme générale est donnée par :

$$h(x) = e\left(\frac{-|x-c|^2}{2\sigma^2}\right) \quad (2.25)$$

Les paramètres de cette fonction sont le centre C et le rayon noté par σ . La sortie du neurone caché donne une mesure de la distance entre le vecteur d'entrée x et le centroïde C du cluster de données. La sortie du réseau est une combinaison linéaire des sorties des neurones de la couche cachée, pondérées par les poids de leurs connexions respectives. Le paramètre σ , représentant le rayon de l'hypersphère, est généralement déterminé à l'aide d'un processus itératif sélectionnant une largeur optimale sur la base des ensembles de données complets.

2.9.1.3 Réseaux de neurones probabilistes (PNN)

La structure d'un PNN est similaire à celle d'un RBF, les deux utilisent des fonctions radiales plutôt que des fonctions d'activation sigmoïdes pour construire une fonction de décision locale centrée sur un sous-ensemble de l'espace d'entrée. La fonction de décision globale est la somme de toutes les fonctions locales. De cette façon, le problème des minima locaux est résolu. L'inconvénient majeur de l'utilisation des PNN est le coût de calcul pour la taille potentiellement grande de la couche cachée qui pouvait être égale à la taille du vecteur d'entrée.

Le PNN peut être un classificateur bayésien, approximant la fonction de densité de probabilité (PDF) d'une classe à l'aide de fenêtres de Parzen. L'estimateur généralisé utilisé pour approcher des fonctions de densités (PDF) dans le cas des réseaux PNN est :

$$f_A(x) = \frac{1}{(2\pi)^2 \sigma^p N_A} \sum_{i=1}^{N_A} e^{-\frac{|x-C_i|^2}{2\sigma^2}} \quad (2.26)$$

2.9.1.4 Cas d'applications au diagnostic des défauts de roulement

Les réseaux de neurones sont parmi les méthodes de classification des défauts de roulement les plus utilisés. De nombreux travaux existent dans la littérature. On peut citer [109] qui ont présenté une étude comparative pour les performances de trois types de réseaux de neurones artificiels (ANN), à savoir, le perceptron multicouche (MLP), le réseau de fonction de base radiale (RBF) et le réseau de neurones probabilistes (PNN), dans la détection de défauts de roulement. Les algorithmes génétiques ont été utilisés dans l'optimisation des paramètres des réseaux utilisés. Deux classes sont obtenues à la sortie (état normal et état défaillant). Dans [110], les auteurs ont présenté dans une autre étude similaire, les performances des trois types de réseaux de neurones (MLP, RBF et PNN) basés sur la transformée en ondelettes discrètes (DWT), dans la discrimination de quatre conditions de défauts de roulement. Ces classes représentatives de sortie se composent respectivement du défaut de bague intérieure, du défaut de bille, du défaut de bague extérieure et de l'état normal. Enfin, dans [111], et pour développer un système expert, les auteurs ont utilisé un réseau de neurones basé sur des descripteurs temporels et fréquentiels issues d'un signal vibratoire mesuré sur un roulement endommagé, capable de remonter à la source du défaut.

2.9.2 Séparateurs à vaste marge SVM

Ces algorithmes sont une classe de techniques d'apprentissage introduite par Vladimir Vapnik au début des années 90. A l'origine, les SVMs ont été développés pour la classification binaire et la régression. Leur principe est basé sur la séparation des classes à l'aide d'une frontière, de sorte que la distance entre les différents groupes de données et la frontière qui les sépare soit maximale [112]. Cette distance est aussi appelée « marge », d'où la qualification des SVMs de « séparateurs à vaste marge ».

2.9.2.1 SVMs binaires

On parle de classification binaire, si les données d'entraînement viennent uniquement de deux classes différentes (+1 ou -1). L'idée des SVMs est de rechercher un hyperplan (droite dans le cas de deux dimensions) qui sépare le mieux ces deux classes.

A) SVM à marge dure

On donne un ensemble d'apprentissage $\{(x_i, y_i)\}_{i=1, \dots, n}$ où $x_i \in \mathcal{X}$ ($\mathcal{X} = \mathbb{R}^m$) et $y_i \in \{-1, +1\}$. Dans un problème de classement à deux classes, le but est de construire une

fonction $f: \chi \rightarrow \mathbb{R}$ qui permet de prédire si un nouvel exemple $x \in \chi$ appartient à la classe -1 ou à la classe $+1$. La fonction f recherchée doit minimiser les erreurs de classification sur les deux ensembles : d'entraînement et de test.

L'hyperplan séparateur est représenté par l'équation suivante :

$$H(x) = w^T x + b \quad (2.27)$$

Où,

w est un vecteur de m dimensions et b est un terme. La fonction de décision, pour un exemple x , peut être exprimée comme suit :

$$\begin{cases} \text{Classe} = 1 & \text{Si } H(x) > 0 \\ \text{Classe} = -1 & \text{Si } H(x) < 0 \end{cases} \quad (2.28)$$

Puisque les deux classes sont linéairement séparables, il n'existe aucun exemple qui se situe sur l'hyperplan, c.-à-d. qui satisfait $H(x) = 0$. Il convient alors d'utiliser la fonction de décisions suivante :

$$\begin{cases} \text{Classe} = 1 & \text{Si } H(x) > 1 \\ \text{Classe} = -1 & \text{Si } H(x) < -1 \end{cases} \quad (2.29)$$

Les inégalités précédentes sont équivalentes à l'équation :

$$y_i(w^T x_i + b) \geq 1, i = 1, \dots, n \quad (2.30)$$

La marge est définie comme étant la distance entre l'hyperplan séparateur des deux classes $w^T x + b = 0$, et l'exemple le plus proche. La région comprise entre les deux hyperplans $w^T x + b = -1$ et $w^T x + b = +1$ est nommée région de généralisation de la machine d'apprentissage. La maximisation de cette dernière est le but de la phase d'entraînement. Elle consiste à rechercher un « hyperplan de séparation optimale » qui maximise la région de généralisation (marge). Ceci peut être obtenu en résolvant le problème d'optimisation (appelé problème primal) suivant :

$$\begin{cases} \text{Minimiser } \frac{1}{2} \|w\|^2 \\ \text{Sous contraintes} \\ y_i(w^T x_i + b) \geq 1, \forall i = 1, \dots, n \end{cases} \quad (2.31)$$

Ce problème de programmation quadratique avec contraintes linéaires, peut être convertit en un problème dual équivalent sans contraintes qui introduit les multiplicateurs de Lagrange, de la façon suivante :

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^n \alpha_i \{y_i(w^T x_i + b) - 1\} \quad (2.32)$$

Où,

les α_i sont les multiplicateurs de Lagrange, non négatifs ;

L est la fonction objective à optimiser

Le problème dual à maximiser devient :

$$\left\{ \begin{array}{l} \text{Maximiser } L(\alpha) = \sum_{i=1}^n \alpha_i - \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{Sous contraintes} \\ \sum_{i=1}^n \alpha_i y_i = 0 \\ \alpha_i \geq 0 \end{array} \right. \quad (2.33)$$

Dans le cas où la classification est linéairement séparable, la fonction de décision est donnée par :

$$H(x) = \sum_S \alpha_i y_i x^T x_i + b \quad (2.34)$$

Où,

S représente l'ensemble des vecteurs supports ;

b peut être calculé à partir de n'importe quel vecteur support par l'équation :

$$b = y_i - w^T x_i \quad (2.35)$$

Il est possible (précision) de prendre la moyenne de b pour tous les vecteurs supports :

$$b = \frac{1}{|S|} \sum_{i \in S} y_i - w^T x_i \quad (2.36)$$

Donc, pour chaque nouvel exemple , la fonction de décision H peut être calculée. La décision prise peut être comme suit :

$$\left\{ \begin{array}{l} x \in \text{Classe } + 1 \text{ si } H(x) > 0 \\ x \in \text{Classe } - 1 \text{ si } H(x) < 0 \\ x \text{ inclassifiable si } H(x) = 0 \end{array} \right. \quad (2.37)$$

B) SVM à marge souple

Dans le cas des données qui ne sont pas linéairement séparables, ou contiennent du bruit, les contraintes précédentes (2.30) ne peuvent être vérifiées, et il y a nécessité de les relaxer, en admettant une certaine erreur de classification des données. On introduit sur les contraintes des variables ξ_i dites de relaxation. On obtient :

$$y_i(w^T x_i + b) \geq 1 - \xi_i, i = 1, \dots, n \quad (2.38)$$

Le problème primal devient :

$$\left\{ \begin{array}{l} \text{Minimiser } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{Sous contraintes} \\ y_i(w^T x_i + b) \geq 1 - \xi_i, \forall i = 1, \dots, n \\ \xi_i \geq 0 \end{array} \right. \quad (2.39)$$

Où,

C est un paramètre positif fixe qui représente une balance entre les deux termes de la fonction objective, à savoir la maximisation de la marge et la minimisation de l'erreur de classification.

En introduisant les multiplicateurs de Lagrange α_i et β_i , on obtient :

$$L(w, b, \alpha, \xi, \beta) = \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i y_i (w^T x_i + b) - 1 + \xi_i - \sum_{i=1}^n \beta_i \xi_i \quad (2.40)$$

Le problème dual s'obtient de la manière suivante :

$$\left\{ \begin{array}{l} \text{Maximiser } L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{Sous contraintes} \\ \sum_{i=1}^n \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C \end{array} \right. \quad (2.41)$$

La fonction de décision est alors calculée de la même manière que dans le cas des SVMs à marge dure mais uniquement à base des vecteurs supports non bornés.

C) Utilisation des noyaux

Dans la plupart des problèmes réels, il est difficile de séparer n'importe quel jeu de données par un simple hyperplan. Alors, la solution consiste à projeter les points d'apprentissage dans un espace de dimension plus élevée, grâce à une fonction non-linéaire Φ qu'on appelle fonction noyau choisie a priori. L'application de la maximisation de marge permet dans le cadre du diagnostic, de garantir une meilleure robustesse face au bruit, et donc une capacité de généralisation. Le nouvel espace ainsi obtenu est appelé espace des caractéristiques (features space).

Dans ce nouvel espace de caractéristiques, la fonction objective à optimiser devient :

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \langle \Phi(x_i), \Phi(x_j) \rangle \quad (2.42)$$

Où,

$\langle \Phi(x_i), \Phi(x_j) \rangle$ est le produit scalaire des deux images des vecteurs x_i et x_j dans le nouvel espace et dont le résultat est un scalaire.

Dans le calcul de l'optimum de la fonction (2.42), au lieu de calculer $\Phi(x_i), \Phi(x_j)$ et leur produit scalaire, on calcule plutôt une fonction $K(x_i, x_j)$ qui peut être vue comme une matrice $G[n, n]$ dite de Gram et représente à la fois les deux transformations et leur produit scalaire. Cette fonction permet de surmonter le problème de détermination de la transformation Φ et ainsi d'apprendre des relations non linéaires par des machines linéaires. Pour que celle-ci soit un « Noyau » (Kernel), il faut qu'elle soit semi-définie positive (respect des conditions de Mercer). Il existe des noyaux très utilisés et considérés comme standards, parmi lesquels on cite, le Noyau linéaire, le Noyau polynomial, le Noyau RBF, etc.

Une fois le noyau choisi, la fonction objective 2.42 peut être calculée comme suit :

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (2.43)$$

Et la fonction de décision devient :

$$H(x) = \sum_S \alpha_i y_i K(x_i, x) + b \quad (2.44)$$

Où,

S représente l'ensemble des vecteurs supports.

2.9.2.2 SVMs multiclasse

Dans le monde réel, les problèmes sont dans la plupart des cas de nature multiclasse. Dans de telles situations, on ne cherche pas à affecter un nouvel exemple à l'une de deux classes mais à l'une parmi plusieurs. Autrement dit, la décision n'est plus binaire et un seul hyperplan ne suffit plus.

Les SVMs multiclasse, réduisent le problème de multitude des classes à une composition de plusieurs hyperplans biclasses permettant de tracer les frontières de décision entre les différentes classes. Ces méthodes décomposent l'ensemble d'exemples en plusieurs sous-ensembles, représentant chacun un problème de classification binaire. Ainsi, pour chaque problème un hyperplan de séparation est déterminé par la méthode un SVM binaire. On construit lors de la classification une hiérarchie des hyperplans binaires, pour décider de la classe d'un nouvel exemple. Dans la littérature, plusieurs méthodes de décomposition ont été élaborées (une contre une, une contre reste, etc.).

2.9.2.3 Cas d'applications au diagnostic des défauts de roulement

Dans l'étude très élaborée [113], les auteurs ont utilisé les machines à vecteurs de support multiclasse (MSVM) pour la tâche de classification des défauts de roulement. L'extraction des indicateurs est basée sur les ondelettes. Plusieurs stratégies différentes de sélection de fonctionnalités et de réduction de la dimensionnalité, ont été employées. Pour valider la méthode, quatre conditions de défauts ont été testés au banc d'essai, et la comparaison des résultats obtenus a permis de sélectionner la meilleure approche en termes de précision et en temps de calcul. Dans une autre étude [114], la classification des défauts de roulement par SVM basée sur les paquets d'ondelettes a été proposée. Deux approches ont été utilisées dans l'optimisation des paramètres du classificateur lors de la construction des modèles, l'optimisation des essaims de particules (PSO-SVM) et l'algorithme artificiel des essaims de poissons (AFS-SVM). Ces approches ont amélioré la précision de la classification des défaillances. Enfin, dans [115], les travaux proposés sont destinés au diagnostic de défaut de roulement à bille par les SVM. L'étude a été menée suivant deux approches, l'une basée sur l'optimisation de la sélection des indicateurs grâce aux algorithmes génétiques (GA), et l'autre est basée sur la validation croisée. L'extraction des indicateurs issus du domaine temporel a été utilisée. La comparaison des résultats obtenus par les deux approches, révèle l'obtention d'une précision relativement meilleure grâce à la validation croisée.

2.9.3 K plus proches voisins KNN

En apprentissage automatique (ML), la méthode des k plus proches voisins k-NN ou KNN (k-Nearest Neighbors), est une méthode d'apprentissage supervisé. Cet algorithme est une méthode non paramétrique utilisée aussi bien dans les problèmes de classification que dans ceux de régression. Il consiste à classer l'entrée dans la catégorie à laquelle appartiennent les k plus proches voisins dans l'espace des caractéristiques identifiées par apprentissage.

Pour effectuer une prédiction, la méthode KNN a besoin du jeu de données en entier, du nombre k de plus proches voisins (à fixer au préalable par l'utilisateur) et d'une fonction de calcul de similarité (distance entre deux observations). Il existe plusieurs fonctions de calcul de distance à savoir, la distance euclidienne, la distance de Manhattan, la distance de Minkowski, distance de Jaccard, la distance de Hamming, etc. Le choix du modèle de fonction dépend de la nature et des types de données à traiter. La distance euclidienne d_e reste la plus utilisée et donnée par l'expression :

$$d_e(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.45)$$

L'algorithme KNN pour le problème de classification, passe par les étapes suivantes :

1. Charger les données de la base d'apprentissage
2. Initialiser k au nombre choisi
3. Pour chaque donnée dans la base d'apprentissage :
 - a) Calculez la distance entre la nouvelle donnée (inconnue) et la donnée courante.
 - b) Mémoriser cette distance et la donnée associée dans une liste de couple (distance, donnée).
4. Trier dans ordre croissant cette collection ordonnée contenant distances et indices
5. Sélectionner les k premières entrées de la collection de données triées
6. Obtenir les étiquettes des k entrées sélectionnées (retourner le mode ou valeur la plus fréquente).

Le choix de la valeur de k à utiliser pour effectuer une prédiction avec KNN, varie en fonction du jeu de données. Le choix d'un nombre k petit, induit un problème de sous apprentissage (underfitting). Toutefois, le choix d'un nombre très grand risque d'engendrer du sur-apprentissage (overfitting) et par conséquent, un modèle qui se généralise mal sur des observations qu'il n'a pas encore vu. En règle générale, sa valeur est choisie après des tests effectués par l'utilisateur.

Parmi ses applications au diagnostic des défauts de roulement, on cite dans la littérature les travaux de Sharma et a. [116], qui ont proposé une méthodologie basée sur le classificateur

k plus proches voisins pondéré (WKNN). La méthode proposée utilise de simples indicateurs temporels pour classer trois conditions de défaut, à partir du signal de vibration. Plusieurs variantes ont été testées pour optimiser la précision de la classification et la complexité des calculs, sans et avec analyse en composantes principales (PCA). Les résultats expérimentaux ont atteint une précision de 100% lors de l'utilisation de la métrique basée sur la distance de Mahalanobis. Une autre étude [117], basée sur la comparaison de plusieurs classificateurs utilisés dans le diagnostic des roulements, a été présentée. Conjointement à la méthode KNN, les réseaux de neurones ANN, l'arbre de décision, et les machines à vecteurs de support ont été utilisés. Les résultats ont montrés la supériorité des SVM, selon les auteurs.

2.9.4 Arbres de décision

Ces classifieurs se basent dans leur décision sur une suite arborescente de tests (séquences) où chacun est associé à un attribut. Ce type de structure en arborescence les rend lisibles par l'utilisateur. Autrement dit, un arbre de décision est une représentation graphique d'un ensemble de règles, conduisant à une procédure de classification basée sur un parcours passant à travers l'arbre. Ce processus entame sa progression à la racine de l'arbre et y progresse ensuite, selon les réponses aux tests qui étiquettent les nœuds internes. Chaque nœud est associé à un attribut, jusqu'à aboutissement à la feuille finale, associé à la classe correspondante.

Il existe de nombreux algorithmes d'apprentissage par arbre de décision. Parmi lesquels, deux sont principalement les plus utilisés, CART (Breiman en 1984) et C4.5 (Quinlan 1993). Le schéma de l'algorithme générique d'apprentissage, peut être décrit comme suit :

<p>entrée : échantillons S</p> <p>début</p> <p>Initialiser l'arbre courant à l'arbre vide ; la racine est le nœud courant</p> <p>répéter</p> <p> Décider si le nœud courant est terminal</p> <p> si le nœud est terminal alors</p> <p> Lui affecter une classe</p> <p> sinon</p> <p> Sélectionner un test et créer autant de nouveaux nœuds fils qu'il y a de réponses possibles au test</p> <p> finsi</p> <p> Passer au nœud suivant non exploré s'il en existe</p> <p>Jusqu'à obtenir un arbre de décision</p> <p>fin</p>

L'utilisation des arbres de décision dans le diagnostic des défauts de roulements des machines tournantes, est répandue dans la littérature. Boumahdi et al. [118], ont présenté une méthodologie pour l'extraction de règles expertes dans l'identification des défauts des roulements dans les machines tournantes en utilisant les arbres de décision. Un montage expérimental a servi à l'obtention de signaux vibratoires, à partir desquels une extraction des indicateurs temporels et fréquentiels, a été opérée. L'arbre de décision et un modèle graphique de roulement, sont ensuite construits en appliquant l'algorithme J48 sur l'ensemble de données, et ainsi des règles expertes sont établies. L'étude [119], est basée sur l'arbre de décision comme choix de l'outil de diagnostic de roulement. L'extraction d'indicateurs temporel a été opérée sur des signaux en provenance de capteur triaxial. La réduction de dimensionnalité est réalisée à l'aide de l'analyse en composantes principales PCA. Les performances de classification ont été commentées grâce à l'algorithme d'apprentissage C4.5 dans les deux cas, avec et sans réduction de dimensionnalité. Un accent a été mis particulièrement sur la sensibilité de l'arbre de décision aux données bruitées. Ainsi, et selon cette étude, la présence de bruit minime soit-il génère une dégradation des performances.

2.9.5 Classifieur bayésien naïf NB

Ce classifieur est un algorithme d'apprentissage supervisé qui permet de classifier un ensemble d'observations selon des règles déterminées par l'algorithme lui-même. Son idée de base, classer une observation selon les valeurs des attributs qui la décrivent. Le Classifieur Bayésien Naïf (NB) calcule les probabilités d'appartenance de l'observation à chacune des classes et l'assigne à celle qui obtient la plus haute valeur. La classe de l'observation ainsi que sa probabilité sont estimées en utilisant le théorème de Bayes avec l'hypothèse simplificatrice, dite naïve, d'indépendance entre toutes les paires de variables.

Dans l'étude [120], les auteurs ont utilisé les algorithmes, Naïf de Bayes et réseau de Bayes pour le diagnostic des défauts de roulement, de cavitation et de turbine d'une pompe centrifuge monobloc. L'extraction des attributs à partir des signaux de vibration dans les bonnes et les mauvaises conditions des composants de la pompe, est basée sur les ondelettes discrètes. Les performances de classification de différentes familles d'ondelettes discrètes ont été testées et comparées lors de cette étude. Kumar et al. [121], ont présenté une étude destinée à l'évaluation de l'état des roulements, basée sur deux classifieurs bayésiens, Naïf de Bayes et réseau de Bayes. Des descripteurs statistiques extraits à partir des données de vibration acquises, ont été utilisés. Une réduction de dimensionnalité est réalisée par arbre de décision basé sur l'algorithme J48. Les modèles construits ont été testés avec la méthode de validation croisée.

2.9.6 Classifieur neuro floue

Les classifieurs à base de réseaux neuro-flous sont nés de combinaison entre les réseaux de neurones et la logique floue. Cette association est conçue de manière à tirer avantage des deux techniques. Leur principale propriété est la capacité de traiter dans un même outil des connaissances numériques et symboliques d’un système. Généralement, ce sont des réseaux neuronaux dont l’élément de base est un opérateur flou souvent appelé neurone flou. Celui-ci possède une architecture similaire à celle du neurone formel, à la différence que ces paramètres sont décrits en totalité ou en partie par la logique floue.

Dans l’étude [122], Lou et Loparo ont exploré le diagnostic des défauts localisés dans les roulements à billes à l’aide de la transformée en ondelettes et de la classification neuro-floue. Les signaux de vibration représentatifs des conditions de défaut de roulements, ont été acquis à partir d’un système expérimental motorisé. La transformée en ondelettes a été utilisée pour le traitement et la génération des attributs (indicateurs). Un système d’inférence adaptative neuro-floue (ANFIS) a été formé et utilisé comme classificateur de diagnostic. Les résultats sont comparés avec ceux obtenus à partir d’autres méthodes, et démontrent la fiabilité de la méthode proposée pour différents cas de charge. Dans un autre travail [123], une approche de diagnostic automatisé des défauts a été développée par Tiwari et al., à l’aide de l’entropie de permutation multi-échelles (MPE) et d’un classifieur neuro-floue adaptatif (ANFC). Pour tenir compte de la non-linéarité dynamique et de l’effet de couplage entre les pièces mécaniques, des MPE sur 16 échelles sont calculés. Les résultats expérimentaux ont montré que cette combinaison de l’ANFC avec MPE peut atteindre une précision de 100%. Selon ses auteurs, cette technique est fiable et peut être utilisée dans système de diagnostic des pannes en ligne.

2.10 METHODES BASEES SUR L’APPRENTISSAGE PROFOND (DL)

Ces approches issues de l’intelligence artificielle, sont apparues depuis 2006, avec l’avènement des méga-données. Il a été prouvé que les méthodes du Machine Learning ML deviennent parfois inadaptées pour cette augmentation de taille. L’apprentissage profond (DL) est venu donc pour surmonter ce problème. Mieux encore, plus la taille des données est grande, plus les méthodes d’apprentissage profond seront efficaces. Ainsi, Ce segment de méthodes d’apprentissage profond devient un axe de recherche très prometteur, redéfinissant les performances dans un large éventail de domaines tels que la segmentation d’images, la reconnaissance d’objets, la traduction automatique et la reconnaissance vocale. Dérivé du réseau de neurones artificiels (ANN) et composé de plusieurs couches de traitement non linéaire, le DL tente d’apprendre des représentations de données hiérarchiques.

À ce jour, il existe une variété d'architectures d'apprentissage en profondeur qui peuvent être appliquées à la surveillance de l'état des machines (Machine Health Monitoring - MHM) en général et au diagnostic des défauts de roulement par analyse des signaux vibratoires en particulier. Le réseau de neurones convolutifs (CNN ou ConvNet) est l'un des algorithmes les plus populaires pour l'apprentissage en profondeur. C'est un type d'apprentissage automatique dans lequel un modèle apprend à effectuer des tâches de classification directement à partir d'images, de vidéos, de textes ou de sons. Dans le domaine de l'identification de défaut dans les roulements, l'apprentissage et par la suite la classification, sont basés sur des images vibratoires obtenues à partir de signaux accélérométriques représentatifs des conditions de défauts. Le réseau neuronal convolutif (CNN) est l'une des méthodes les plus utilisées pour identifier les défauts de roulement [124, 125].

Les auto-encodeurs sont des algorithmes d'apprentissage non supervisé à base de réseaux de neurones artificiels, qui permettent de construire une nouvelle représentation, plus compacte et avec moins de descripteurs, d'un jeu de données, ce qui permet de réduire sa dimensionnalité. Son architecture est composée d'un encodeur constitué par un ensemble de couches de neurones, qui traitent les données afin de construire de nouvelles représentations dites « encodées », et d'un décodeur constitué par d'autres couches de neurones, qui reçoivent ces représentations et les traitent afin d'essayer de reconstruire les données de départ. L'entraînement consiste à modifier les paramètres de l'auto-encodeur afin de réduire l'erreur de reconstruction (différence entre les données reconstruites et les données initiales) mesurée sur les différents exemples du jeu de données. L'usage des méthodes de DL basées sur les auto-encodeurs (AE) dans le diagnostic des défauts de roulements, figure également en bonne position [126, 127].

Les réseaux récurrents (ou RNN pour Recurrent Neural Networks) sont des réseaux de neurones dans lesquels l'information peut se propager dans les deux sens. Ils possèdent des connexions récurrentes au sens où elles conservent des informations en mémoire sur des états passés. Pour cette raison, ils sont particulièrement adaptés au traitement des séquences temporelles, car les données forment une suite et ne sont pas indépendantes les unes des autres. Cependant, dans leurs formes classiques simples ou Vanilla, les RNNs ne sont capables de mémoriser que le passé dit proche. Ce n'est qu'avec l'avènement des réseaux LSTMs (Long Short Term Memory) que leurs domaines d'application ont été révolutionnés [128].

Parmi les autres méthodes de DL appliquées au diagnostic des roulements, on cite encore les réseaux profonds de croyance (Deep Belief Network - DBN), qui sont des modèles graphiques génératifs composés de plusieurs couches de variables latentes (unités cachées) avec

des connexions entre les couches [129]. Enfin, le réseau antagoniste génératif (Generative Adversarial Network - GAN) est une classe d'algorithmes d'apprentissage non-supervisé. Ces algorithmes permettent de générer des images avec un fort degré de réalisme. Il est formé de deux réseaux qui sont placés en compétition. Le premier est le générateur, tandis que son adversaire est le discriminateur [130].

2.11 SYNTHÈSE

En passant en revue l'ensemble des techniques employées dans le diagnostic des défauts de roulement, des constats quant à leur employabilité sont faits. Les indicateurs statistiques permettent une première appréciation de l'état vibratoire, mais ne peuvent ni identifier le type ni la position du défaut. L'analyse spectrale à base de la transformée de Fourier, et ses variantes, sont venues pour remédier au problème et ont donné des résultats plus encourageants en termes d'existence de défauts. D'ailleurs, l'analyse d'enveloppe est l'outil le plus approprié dans cette catégorie pour l'analyse des signaux de roulements. Cependant, on s'est aperçu vite que le problème d'hypothèse de stationnarité des signaux n'est pas compatible avec la présence de composantes transitoires issues de différentes sources dans les signaux traités. Les signatures de défauts recherchés sont caractérisées par de faibles fréquences, et sont généralement modulés par les fréquences de résonance de la structure. Ce problème a été résolu par le développement des techniques d'analyse en temps fréquence, mais ces dernières souffrent du problème de résolution généré par la taille de la fenêtre employée... Les méthodes temps échelle sont bien adaptées aux signaux non stationnaires, mais restent tributaires du choix des types d'ondelettes adaptées à chaque application, etc. Soit en résumé, pour faire un diagnostic de qualité il faut faire usage simultané d'une multitude de techniques plus au moins avancées, ce qui signifie posséder des compétences illimitées dans de nombreux domaines...

Les méthodes de diagnostic automatisé, issues de l'intelligence artificielle, viennent pour ses problèmes et exigences des méthodes à base de traitement de signal. Ainsi, ces méthodes dites du Machine Learning explorent l'espace des données, en mode supervisé ou non, pour établir des règles de décision qui permettent des affectations de classes représentatives de conditions de défauts. Cependant, de toute cette panoplie de méthodes existantes dans ce domaine, il n'y a pas de méthode idéale adaptée pour toutes les problématiques posées. L'utilisateur est contraint d'expérimenter un certain nombre de techniques, avant de se prononcer sur la plus adaptée à sa problématique. Pire encore, l'augmentation des tailles des données et l'arrivée des méga données, certaines méthodes du Machine Learning deviennent inadaptées à cette augmentation de taille. C'est pourquoi, des méthodes dites du Deep Learning

ont été développées, pour accompagner ces augmentations de tailles des données. La précision des résultats obtenus, n'a jamais été aussi élevée auparavant avec le ML, et parfois, ne nécessite aucun prétraitement de données au préalable. Cependant, il faut particulièrement souligner que ce modèle d'apprentissage ne peut être efficace que si les données sont de très grande taille d'une part, et les moyens de calcul très puissants (multitude de CPU, GPU). Ce qui représente un handicap à leur expansion aux temps actuels.

En conséquence, face à notre problématique de diagnostic de défauts de roulement, ou plus concrètement la reconnaissance de conditions de défauts par processus de classification, les modèles de mélange gaussiens GMMs qui ont prouvé leur efficacité dans de nombreux domaines, ont été adoptés comme méthode de classification dans cette thèse. Ajouté à leur simplicité, flexibilité et autres adaptations à des données variées, représentent autant d'arguments pour notre choix du classificateur, et plus encore : en mode supervisé.

2.12 CONCLUSION

L'ensemble des techniques utilisées dans le diagnostic des défauts de roulement, a été passé en revue dans ce chapitre. Il est à conclure que toutes les méthodes conventionnelles basées sur le traitement de signal offrent des possibilités d'exploration relativement encourageantes dans certaines conditions, et restent limitées dans d'autres considérations. Leur bon usage nécessite forcément un savoir-faire conséquent de l'utilisateur et de hautes compétences avérées, pour faire un diagnostic précis. Les méthodes basées sur le diagnostic automatisé, viennent pour palier à cet inconvénient, par la construction de règles de décision à partir des échantillons de données de la base d'apprentissage. Ainsi, un diagnostic par reconnaissance automatique de condition de défauts, offre des possibilités réelles de vulgarisation au sein des structures de maintenance. Ce qui explique l'engouement de la recherche actuelle dans cette direction.

Le diagnostic à base de méthodes d'apprentissage machine nécessite le passage par plusieurs étapes décrites au cours de ce chapitre, et discutées. La première étape dans ces procédures est l'extraction de l'information complète à partir des signaux acquis sur les machines. Cette information nécessite un traitement afin de la rendre exploitable par les techniques de classification et d'intelligence artificielle. Dans cette voie, plusieurs techniques de sélection et de transformation de l'information peuvent être utilisées. Une fois le vecteur d'entrée construit, l'étape suivante est la construction des règles de décision. Ceci peut être fait par le biais de nombreux classifieurs. Certains ont été abordés, avec une description de leurs

principes, chose nécessaire pour la compréhension de leurs résultats et le choix qui s'adapte le mieux à notre problématique.

Aussi, nous avons conclu à travers la revue opérée dans la littérature, que malgré la multitude d'approches abordées, aucune n'a pu démontrer de façon catégorique sa supériorité absolue par rapport aux autres. L'absence de règle ou norme qui guide le choix du classifieur approprié à chaque problématique, oblige l'utilisateur à expérimenter un certain nombre d'algorithmes avant de choisir le plus adapté à son application. Ce qui nous conforte d'avantage dans le choix du GMM pour la classification des défauts de roulement, déjà trié sur la base d'autres critères tranchants, élucidés précédemment.

Chapitre 3

***RECONNAISSANCE DES DEFAUTS DE
ROULEMENTS PAR LES MODELES DE
MELANGE GAUSSIENS-GMM***

3.1 INTRODUCTION

La classification est définie comme étant une méthode d'analyse des données qui a pour objectif de regrouper un ensemble d'observations sous formes de classes homogènes. Chaque année, les besoins en analyse de données et particulièrement la classification, augmentent d'avantage car de plus en plus de domaines scientifiques en font appel à des fins descriptives ou décisionnelles. Le diagnostic des défauts est l'un de ces domaines qui n'a cessé de développer des approches complexes basées sur l'apprentissage machine pour automatiser les procédures de reconnaissance et améliorer les performances.

Ce chapitre, sera consacré au diagnostic des défauts de roulements par l'usage de la méthode des modèles de mélange de gaussiennes (GMM). Il présente en premier lieu le problème de la classification et sa modélisation probabiliste. Une attention particulière est portée à l'approche de classification supervisée dans le cadre des démarches de diagnostic envisagées.

Pour ce faire, le schéma directeur du processus de la classification va s'exécuter en plusieurs étapes. La première, vise à extraire des descripteurs ou indicateurs spectraux à partir des différentes sources de données disponibles. La seconde s'attèle la création des modèles à partir des données relatives à la base d'apprentissage, moyennant l'application de l'algorithme d'espérance-Maximisation. La troisième procède à la génération de règles de décision, réalisée par le biais du classifieur. La dernière étape est consacrée à l'évaluation du processus de classification, par l'élaboration des tests de performance nécessaires.

3.2 CLASSIFICATION BASEE SUR LES MODELES PROBABILISTES

Historiquement, les premières méthodes de la classification étaient de nature, algorithmique, heuristique ou géométrique et reposaient principalement sur la dissimilitude entre les objets à classer. Plus récente, l'approche statistique se base sur des modèles probabilistes qui formalisent la notion de classe. Cette approche permet en outre d'interpréter de façon statistique la classification obtenue. La modélisation la plus classique est celle du modèle de mélange fini qui peut être paramétrique ou non.

3.2.1 Formalisme des notions de classe et de règle de décision

Le but de la classification étant de répartir un ensemble d'objets en classes homogènes. Alors, il est plus commode de définir d'abord en premier les notions de classe et partition (ou groupe) et les règles de décision associées aux éléments que l'on cherche à classer.

Si nous considérons \aleph comme ensemble des objets à classer, une classe sera un sous-ensemble de ce dernier. Les deux approches différentes qui permettent de décrire une classe, sont de nature, générative ou discriminative. Dans la première, elle est décrite par les propriétés caractéristiques de ses objets alors que dans la seconde, par sa frontière avec ses voisins [131]. La structure de la « partition » couramment utilisée se définit de manière suivante :

$P = \{\Omega_1, \dots, \Omega_k\}$ est défini comme une partition de l'ensemble \aleph en k classes Ω , si et seulement si :

- 1) $\Omega_i \neq \emptyset$ pour $i = 1, \dots, k$;
- 2) $\bigcup_{i=1}^k \Omega_i = \aleph$;
- 3) $\Omega_i \cap \Omega_j = \emptyset$ pour tout $i \neq j$.

Soit \mathbf{X} un vecteur décrit par N variables quantitatives, qui modélise un objet de l'ensemble \aleph des objets à classer. Le but dans la classification est d'associer ce vecteur \mathbf{X} à une des k classes. Soit Z la variable auxiliaire à valeurs dans $\{1, \dots, k\}$ et telle que $Z = i$ si \mathbf{X} appartient à la $i^{\text{ème}}$ classe.

Ainsi, le problème de la classification revient à établir une règle de décision δ qui associe au vecteur $\mathbf{X} \in \mathbf{R}^N$ un vecteur $Z \in \{1, \dots, k\}$:

$$\begin{aligned} \delta : \mathbf{R}^N &\rightarrow \{1, \dots, k\} \\ \mathbf{X} &\mapsto Z \end{aligned}$$

Cette règle de décision est construite à partir des données d'apprentissage. Dans le cas supervisé, les données utilisées sont dites « complètes » car elles contiennent à la fois les valeurs prises par les N variables explicatives et leur appartenance aux k classes. Les données complètes sont donc l'ensemble des couples observations - labels. Par contre, dans le cas non supervisé, les données utilisées ne sont pas « complètes » car elles ne contiennent que les valeurs prises par les N variables explicatives.

3.2.2 Règle de classification optimale au sens probabiliste

On suppose que les observations x_1, \dots, x_n de l'ensemble \aleph des observations à classer sont des réalisations d'un vecteur aléatoire \mathbf{X} à valeurs dans \mathbf{R}^N . On suppose en outre que les valeurs z_1, \dots, z_n , décrivant l'origine des observations x_1, \dots, x_n , sont des réalisations de la variable aléatoire Z à valeurs dans $\{1, \dots, k\}$. Ainsi, le fait de dire que x est une réalisation de la variable aléatoire \mathbf{X} conditionnellement au fait que $Z = i$ revient à dire que l'observation x appartient à la $i^{\text{ème}}$ classe Ω_i . On introduit également le vecteur aléatoire $S \in \{0,1\}^k$ tel que si $Z = i$ alors $S = (0, \dots, 0, 1, 0, \dots, 0)$ où le $i^{\text{ème}}$ terme vaut 1, et ce pour $i = 1, \dots, k$.

Ce cadre probabiliste permet de construire la règle de décision optimale δ^* , dite également règle de Bayes, qui minimise le risque conditionnel $R(\delta|x)$ pour chaque observation x . En associant un coût nul à une bonne affectation et un coût de 1 à une mauvaise affectation, le risque conditionnel s'écrit :

$$R(\delta|x) = 1 - P(Z = \delta(x)|X = x)$$

La règle δ^* consiste donc à affecter l'observation x à la classe la plus probable a posteriori :

$$\delta^*(x) = \arg \max_{i=1,\dots,k} P(Z = i|X = x) \quad (3.1)$$

Cette règle porte également le nom de MAP pour *maximum a posteriori*

3.2.3 Modélisation par mélange de lois

Les données rencontrées dans les applications modernes, sont multidimensionnelles et de nature complexe. D'une manière générale, ces données ne peuvent être modélisées par une loi classique. Ainsi, les modèles de mélanges représentent des moyens qui peuvent apporter la solution, en se basant sur des lois simples.

3.2.3.1 Modèle de mélange

Il suppose que chaque groupe est caractérisé par une distribution de probabilité. Ainsi, cette approche de modélisation est souple et permet de prendre en compte un grand nombre de situations. Dans un modèle de mélange, on considère que les données x_1, \dots, x_n constituent un échantillon de n réalisations indépendantes du vecteur aléatoire \mathbf{X} à valeur dans \mathbf{R}^N dont la fonction de densité peut s'écrire de la façon suivante :

$$p(x) = \sum_{i=1}^k w_i p_i(x) \quad (3.2)$$

Où,

k est le nombre de classes (connu dans le cas de la classification supervisée appelée « analyse discriminante », mais inconnu dans le cas de la classification non supervisée appelée « classification automatique »), p_i est la densité de la distribution de \mathbf{X} conditionnellement à $Z = i$ (de la $i^{\text{ème}}$ composante du mélange) et les w_i sont les proportions du mélange ($w_i \in [0, 1]$ et $\sum_{i=1}^k w_i = 1$). A noter que, l'identifiabilité d'un modèle de mélange est conditionnée par la numérotation des classes dans le cadre de l'analyse discriminante, vu que les labels des observations sont connus, et arbitraire dans le cas de la classification automatique, puisque les labels sont inconnus.

3.2.3.2 Modèle de mélange paramétrique et règle de Bayes

De plus, on suppose généralement que les densités p_i des classes appartiennent à une famille paramétrée : $p_i(\cdot) = p(\cdot, \theta_i)$. Le modèle de mélange s'écrit alors :

$$p(x) = \sum_{i=1}^k w_i p(x, \theta_i) \quad (3.3)$$

Où,

$\theta = \{w_1, \dots, w_k, \theta_1, \dots, \theta_k\}$ est l'ensemble des paramètres du modèle.

D'une manière générale, dans le cas du modèle de mélange paramétrique (gaussien ou non gaussien), la formule de Bayes permet d'écrire :

$$P(Z = i | X = x, \theta) = \frac{w_i p(x, \theta_i)}{p(x)} \quad (3.4)$$

Où,

$p(x) = \sum_{i=1}^k w_i p(x, \theta_i)$ est une quantité commune à chacune des classes.

La règle de décision de Bayes peut alors être écrite de la façon suivante :

$$\delta^*(x) = \arg \max_{i=1, \dots, k} \{w_i p(x, \theta_i)\} \quad (3.5)$$

Une telle formulation, indique clairement que, dans le cadre du modèle de mélange paramétrique, le problème de la classification se résume à l'estimation des paramètres du modèle.

3.3 MODELE DE MELANGE GAUSSIEN ET ALGORITHME EM

Dans cette partie, les distributions de mélanges gaussiens sont développées et discutées en considérant le problème de la recherche de groupes dans un ensemble de points de données. Ensuite, une technique générale utilisée pour trouver des estimateurs du maximum de vraisemblance dans des modèles à variables latentes, est introduite sous forme d'un algorithme nommé Espérance-Maximisation (EM).

3.3.1 Distribution Gaussienne

La distribution gaussienne, également connue sous le nom de distribution normale, est un modèle largement utilisé pour la distribution de variables continues. Elle a de nombreuses propriétés analytiques importantes qui lui permettent d'intervenir dans de nombreux contextes différents. Elle peut être motivée aussi, à partir d'une variété de perspectives différentes.

Dans le cas d'une seule variable x , la distribution gaussienne peut s'écrire sous la forme :

$$\mathcal{N}(x | \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\} \quad (3.6)$$

Où,

μ est la moyenne et σ^2 est la variance.

Pour un vecteur \mathbf{x} de dimension D , la distribution gaussienne multivariée prend la forme

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} \quad (3.7)$$

Où,

$\boldsymbol{\mu}$ est un vecteur moyen de dimension D , $\boldsymbol{\Sigma}$ est une matrice de covariance $D \times D$, et $|\boldsymbol{\Sigma}|$ désigne le déterminant de $\boldsymbol{\Sigma}$.

En considérant la forme géométrique de la distribution gaussienne, la dépendance fonctionnelle de la Gaussienne de \mathbf{x} , passe par la forme quadratique, qui apparaît dans l'exposant :

$$\Delta^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \quad (3.8)$$

La quantité Δ est appelée distance de Mahalanobis de $\boldsymbol{\mu}$ à \mathbf{x} et se réduit à la distance euclidienne dans le cas où, $\boldsymbol{\Sigma}$ est la matrice d'identité. La distribution gaussienne sera constante sur les surfaces de l'espace \mathbf{x} pour lesquelles cette forme quadratique est constante.

Par ailleurs, il est à noter une propriété importante de la distribution gaussienne multivariée. Si deux ensembles de variables sont conjointement gaussiens, la distribution conditionnelle de l'un conditionné de l'autre, est à nouveau gaussienne. De même, la distribution marginale de l'un ou l'autre ensemble est également gaussienne.

Enfin, l'estimation des paramètres de la distribution, est obtenue via la maximisation de la vraisemblance. Étant donné un ensemble de données $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}^T$ dans lequel les observations $\{\mathbf{x}_n\}$ sont supposées être tirées indépendamment d'une distribution gaussienne multivariée. La fonction de vraisemblance logarithmique est donnée par

$$\ln p(\mathbf{X}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{ND}{2} \ln(2\pi) - \frac{N}{2} \ln|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{n=1}^N (\mathbf{x}_n - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \boldsymbol{\mu}) \quad (3.9)$$

Par simple réarrangement, on voit que la fonction de vraisemblance ne dépend de l'ensemble de données que par les deux grandeurs $\sum_{n=1}^N \mathbf{x}_n$ et $\sum_{n=1}^N \mathbf{x}_n \mathbf{x}_n^T$, connues comme les statistiques suffisantes pour la distribution gaussienne. La dérivée de la vraisemblance logarithmique par rapport à $\boldsymbol{\mu}$ est donnée par :

$$\frac{\partial}{\partial \boldsymbol{\mu}} \ln p(\mathbf{X}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^N \boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \boldsymbol{\mu}) \quad (3.10)$$

En fixant cette dérivée à zéro, nous obtenons la solution de l'estimation du maximum de vraisemblance de la moyenne considérée comme moyenne de l'ensemble observé de points de données. Elle est donnée par :

$$\boldsymbol{\mu}_{ML} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (3.11)$$

La maximisation de (3.9) par rapport à $\mathbf{\Sigma}$ est un peu plus complexe, pour cause des contraintes de symétrie et de définition positive. Le résultat qui implique $\boldsymbol{\mu}_{ML}$ car conséquence d'une maximisation conjointe par rapport à $\boldsymbol{\mu}$ et $\mathbf{\Sigma}$, prend la forme :

$$\mathbf{\Sigma}_{ML} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \boldsymbol{\mu}_{ML})(\mathbf{x}_n - \boldsymbol{\mu}_{ML})^T \quad (3.12)$$

A notez que la solution (3.11) pour $\boldsymbol{\mu}_{ML}$ ne dépend pas de $\mathbf{\Sigma}_{ML}$, et donc on peut d'abord évaluer $\boldsymbol{\mu}_{ML}$ et ensuite l'utiliser pour évaluer $\mathbf{\Sigma}_{ML}$. L'évaluation des espérances des solutions du maximum de vraisemblance, donne :

$$\mathbb{E}[\boldsymbol{\mu}_{ML}] = \boldsymbol{\mu} \quad (3.13)$$

Et,

$$\mathbb{E}[\mathbf{\Sigma}_{ML}] = \frac{N-1}{N} \mathbf{\Sigma} \quad (3.14)$$

On constate que l'espérance de l'estimation du maximum de vraisemblance pour la moyenne est égale à la vraie moyenne. Cependant, l'estimation du maximum de vraisemblance pour la covariance a une espérance qui est inférieure à la valeur réelle et, par conséquent, biaisée. La correction de ce biais est obtenue via un estimateur différent $\tilde{\mathbf{\Sigma}}$ donné par :

$$\tilde{\mathbf{\Sigma}} = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}_n - \boldsymbol{\mu}_{ML})(\mathbf{x}_n - \boldsymbol{\mu}_{ML})^T \quad (3.15)$$

3.3.2 Modèle de mélange gaussien

Dans la section précédente, nous avons motivé le modèle de mélange gaussien comme une simple superposition linéaire de composantes gaussiennes, visant à fournir une classe de modèles de densité plus riche que le seul gaussien. Nous passons maintenant à une formulation de mélanges gaussiens en termes de variables latentes discrètes. Cela nous donnera un aperçu plus approfondi de cette importante distribution et servira également à motiver l'algorithme de maximisation des espérances.

$$p(\mathbf{x}) = \sum_{k=1}^M w_k \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_k, \mathbf{\Sigma}_k) \quad (3.16)$$

Introduisons une variable aléatoire binaire \mathbf{z} à K dimensions ayant une représentation (1 sur K) dans laquelle un élément particulier z_k est égal à 1 et tous les autres éléments sont égaux à 0. Les valeurs de z_k satisfont donc $z_k \in \{0, 1\}$ et $\sum_k z_k = 1$. Ainsi, on peut voir qu'il y a K états possibles pour le vecteur \mathbf{z} où l'élément est différent de zéro. Nous définirons la distribution conjointe $p(\mathbf{x}, \mathbf{z})$ en termes de distribution marginale $p(\mathbf{z})$ et de distribution conditionnelle $p(\mathbf{x} | \mathbf{z})$, tel que :

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{z}) p(\mathbf{x} | \mathbf{z})$$

La distribution marginale sur \mathbf{z} est spécifiée en termes de coefficients de mélange w_k , tels que :

$$p(z_k = 1) = w_k$$

Où, les paramètres $\{w_k\}$ doivent satisfaire :

$$0 \leq w_k \leq 1 \quad (3.17)$$

Avec,

$$\sum_{k=1}^K w_k = 1 \quad (3.18)$$

Comme \mathbf{z} utilise une représentation 1 sur K , nous pouvons également écrire cette distribution sous la forme :

$$p(\mathbf{z}) = \prod_{k=1}^K w_k^{z_k} \quad (3.19)$$

D'une manière similaire, la distribution conditionnelle de \mathbf{x} pour une valeur particulière de \mathbf{z} est une gaussienne :

$$p(\mathbf{x}|z_k = 1) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

Qui peut être écrite sous la forme :

$$p(\mathbf{x}|\mathbf{z}) = \prod_{k=1}^K \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k} \quad (3.20)$$

La distribution conjointe est donnée par $p(\mathbf{z}) p(\mathbf{x}|\mathbf{z})$, et la distribution marginale de \mathbf{x} est alors obtenue en additionnant la distribution conjointe sur tous les états possibles de \mathbf{z} pour donner :

$$p(\mathbf{x}) = \sum_{\mathbf{z}} p(\mathbf{z}) p(\mathbf{x}|\mathbf{z}) = \sum_{k=1}^K w_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (3.21)$$

Ainsi la distribution marginale de \mathbf{x} est un mélange gaussien de la forme (3.16). Dans le cas de plusieurs observations x_1, \dots, x_N , et comme nous avons représenté la distribution marginale sous la forme $p(\mathbf{x}) = \sum_{\mathbf{z}} p(\mathbf{x}, \mathbf{z})$, il s'ensuit que pour chaque point de données observées \mathbf{x}_n , il existe une variable latente \mathbf{z}_n correspondante.

Cette écriture a abouti à une formulation équivalente du mélange gaussien impliquant une variable latente explicite. Elle permet de travailler avec la distribution conjointe $p(\mathbf{x}, \mathbf{z})$ au lieu de la distribution marginale $p(\mathbf{x})$, ce qui conduit à des simplifications significatives, notamment via l'introduction de l'algorithme de maximisation des espérances (EM).

Une autre grandeur qui joue un rôle important est la probabilité conditionnelle de \mathbf{z} étant donné \mathbf{x} . Nous utiliserons $\gamma(z_k)$ pour désigner $p(z_k = 1|\mathbf{x})$, dont la valeur peut être trouvée en utilisant le théorème de Bayes :

$$\gamma(z_k) \equiv p(z_k = 1|\mathbf{x}) = \frac{p(z_k=1)p(\mathbf{x}|z_k=1)}{\sum_{j=1}^K p(z_j=1)p(\mathbf{x}|z_j=1)} = \frac{w_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K w_j \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \quad (3.22)$$

Nous considérerons w_k comme la probabilité à priori de $z_k = 1$, et la quantité $\gamma(z_k)$ comme la probabilité postérieure correspondante une fois que nous avons observé \mathbf{x} . D'ailleurs, $\gamma(z_k)$ peut également être considéré comme la responsabilité que prend la composante k pour «expliquer» l'observation \mathbf{x} .

3.3.3 Maximum de vraisemblance

Soit un ensemble d'observations $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, et on souhaite modéliser ces données en utilisant un mélange de gaussiens. Il peut être représenté sous la forme d'une matrice \mathbf{X} de dimension $N \times D$ dans laquelle la $n^{\text{ème}}$ ligne est donnée par \mathbf{x}_n^T . De même, les variables latentes correspondantes seront désignées par une matrice \mathbf{Z} de dimension $N \times K$ avec \mathbf{z}_n^T lignes. Si nous supposons que les points de données sont tirés indépendamment de la distribution, alors nous pouvons exprimer le modèle de mélange gaussien pour cet i.i.d. ensemble de données. D'après (3.16), le log de la fonction de vraisemblance est donné par :

$$\ln p(\mathbf{X}|\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^N \ln\{\sum_{k=1}^K w_k \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\} \quad (3.23)$$

Avant d'aborder la façon de maximiser cette fonction, il convient de souligner qu'il existe un problème important associé au cadre du maximum de vraisemblance appliqué aux modèles de mélange gaussien, en raison de la présence de singularités. Pour simplifier, considérons un mélange gaussien dont les composantes ont des matrices de covariance données par $\boldsymbol{\Sigma}_k = \sigma_k^2 \mathbf{I}$, où \mathbf{I} est la matrice unitaire, bien que les conclusions soient valables pour les matrices de covariance générales [132]. Supposons que l'une des composantes du modèle de mélange, soit la $j^{\text{ème}}$ composante qui a sa moyenne $\boldsymbol{\mu}_j$ exactement égale à l'un des points de données de sorte que $\boldsymbol{\mu}_j = \mathbf{x}_n$ pour une valeur de n . Ce point de données contribuera ensuite à un terme dans la fonction de vraisemblance du formulaire :

$$\mathcal{N}(\mathbf{x}_n|\mathbf{x}_n, \sigma_j^2 \mathbf{I}) = \frac{1}{(2\pi)^{1/2}} \frac{1}{\sigma_j} \quad (3.24)$$

Pour toute limite $\sigma_j \rightarrow 0$, ce terme tend vers l'infini et donc la fonction log de vraisemblance ira également à l'infini. Ainsi, la maximisation de la fonction de vraisemblance logarithmique n'est pas un problème bien posé parce que de telles singularités seront toujours

présentes et se produiront chaque fois qu'une des composantes gaussiennes se rétrécit sur un point de données spécifique.

Il est à noter que ce problème ne s'est pas posé dans le cas d'une seule distribution gaussienne. La différence, est que si une seule gaussienne s'effondre sur un point de données, elle contribuera à des facteurs multiplicatifs à la fonction de vraisemblance provenant des autres points de données et ces facteurs iront à zéro de manière exponentielle rapidement, donnant une probabilité globale qui va plutôt à zéro, que l'infini. Cependant, une fois que nous avons (au moins) deux composantes dans le mélange, l'une des composantes peut avoir une variance finie et donc attribuer une probabilité finie à tous les points de données tandis que l'autre composante peut rétrécir sur un point de données spécifique et ainsi contribuer à une valeur additive toujours croissante de la vraisemblance logarithmique. Ces singularités fournissent un autre exemple du sur-ajustement sévère qui peut se produire dans une approche du maximum de vraisemblance.

La maximisation de la fonction log de vraisemblance pour un modèle de mélange gaussien s'avère être un problème plus complexe que pour le cas d'une seule gaussienne. La difficulté vient de la présence de la sommation sur k qui apparaît à l'intérieur du logarithme dans (3.23), de sorte que la fonction logarithme n'agit plus directement sur la Gaussienne. Si nous fixons les dérivées de la vraisemblance logarithmique à zéro, nous n'obtiendrons plus une solution de forme fermée.

Une approche consiste à appliquer des techniques d'optimisation basées sur le gradient [133]. Bien que celles-ci soient réalisables et qu'elles peuvent jouer un rôle important dans certains cas de densité de mélange, il existe une autre approche alternative connue sous le nom d'algorithme « EM » qui a une très large applicabilité.

3.3.4 Algorithme Espérance Maximisation - EM

Une méthode élégante et puissante pour trouver des solutions de maximum de vraisemblance pour les modèles avec des variables latentes est appelée l'algorithme de maximisation des espérances, ou algorithme EM [134]. C'est un algorithme itératif à deux étapes par itération : l'étape d'espérance (E) et l'étape de maximisation (M). La mise à jour des paramètres du mélange gaussien à partir d'une étape E suivie d'une étape M garanti l'augmentation de la fonction du logarithme de la vraisemblance (la vraisemblance dont les paramètres du mélange \mathbf{w} , $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$, sont générés par l'ensemble des observations donné).

Commençons par noter les conditions qui doivent être satisfaites au maximum de la fonction de vraisemblance. En mettant égales à zéro dans (3.23), les dérivées par rapport aux moyennes $\boldsymbol{\mu}_k$ des composantes gaussiennes, de $\ln p(\mathbf{X}|\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, on obtient

$$0 = -\sum_{n=1}^N \frac{w_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_j w_j \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \boldsymbol{\Sigma}_k (\mathbf{x}_n - \boldsymbol{\mu}_k) = -\sum_{n=1}^N \gamma(z_{nk}) \boldsymbol{\Sigma}_k (\mathbf{x}_n - \boldsymbol{\mu}_k) \quad (3.25)$$

Où,

la distribution gaussienne utilisée est de la forme :

$$\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}$$

A noter que les probabilités postérieures, ou responsabilités, données par (3.21) apparaissent bien au membre droit. En multipliant par $\boldsymbol{\Sigma}_k^{-1}$ (sous hypothèse de non singularité) et en réarrangeant, on obtient :

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) \mathbf{x}_n \quad (3.26)$$

Où,

$$N_k = \sum_{n=1}^N \gamma(z_{nk}) \quad (3.27)$$

On peut interpréter N_k comme le nombre effectif de points attribués au cluster k . A noter aussi qu'à partir de la forme de cette solution, la moyenne $\boldsymbol{\mu}_k$ pour la $k^{\text{ème}}$ composante gaussienne est obtenue en prenant une moyenne pondérée de tous les points de l'ensemble de données, dans laquelle le facteur de pondération pour le point de données \mathbf{x}_n est donné par la probabilité postérieure $\gamma(z_{nk})$ dont la composante k est responsable de la génération de \mathbf{x}_n .

Si nous fixons la dérivée de $\ln p(\mathbf{X} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ par rapport à $\boldsymbol{\Sigma}_k$ à zéro, et suivons un raisonnement similaire, en utilisant le résultat de la solution du maximum de vraisemblance pour la matrice de covariance d'une seule gaussienne, on obtient :

$$\boldsymbol{\Sigma}_k = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) (\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^T \quad (3.28)$$

Avec,

chaque point de données pondéré par la probabilité postérieure correspondante et le dénominateur donné par le nombre effectif de points associés à la composante correspondante.

Enfin, nous maximisons $\ln p(\mathbf{X} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ par rapport aux coefficients de mélange w_k . Il faut ici tenir compte de la contrainte (3.18), qui oblige les coefficients de mélange à s'additionner à un. Ceci peut être réalisé en utilisant un multiplicateur de Lagrange et en maximisant la quantité suivante :

$$\ln p(\mathbf{X} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) + \lambda (\sum_{k=1}^K w_k - 1) \quad (3.29)$$

Qui donne :

$$0 = \sum_{n=1}^N \frac{\mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_j w_j \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} + \lambda \quad (3.30)$$

Où,

à nouveau on constate l'apparition des responsabilités. Si nous multiplions maintenant les deux côtés par w_k et somme sur k en utilisant la contrainte (3.18), nous trouvons $\lambda = -N$. En utilisant ceci pour éliminer λ et en réorganisant nous obtenons :

$$w_k = \frac{N_k}{N} \quad (3.31)$$

de sorte que le coefficient de mélange pour le $k^{\text{ème}}$ composant est donné par la responsabilité moyenne que prend ce composant pour expliquer les points de données.

Les étapes de l'algorithme EM sont énumérées ci-dessous.

1. Initialisez les moyennes μ_k , les covariances Σ_k et les coefficients de mélange w_k , et évaluez la valeur initiale du logarithme de la vraisemblance. L'algorithme k-means est couramment utilisé pour regrouper les observations données, ces groupes (clusters) sont ensuite fournis comme point de départ pour l'algorithme EM
2. Etape E : Évaluer les responsabilités (probabilités à postériori) à l'aide des valeurs de paramètres actuelles.

$$\gamma(z_{nk}) = \frac{w_k \mathcal{N}(\mathbf{x}_n | \mu_k, \Sigma_k)}{\sum_{j=1}^K w_j \mathcal{N}(\mathbf{x}_n | \mu_j, \Sigma_j)}$$

3. Etape M : Réestimer les paramètres en utilisant les responsabilités actuelles

$$\begin{aligned} \mu_k^{new} &= \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) \mathbf{x}_n \\ \Sigma_k^{new} &= \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) (\mathbf{x}_n - \mu_k^{new}) (\mathbf{x}_n - \mu_k^{new})^T \\ w_k^{new} &= \frac{N_k}{N} \end{aligned}$$

Où,

$$N_k = \sum_{n=1}^N \gamma(z_{nk})$$

4. Évaluer le logarithme de la vraisemblance

$$\ln p(\mathbf{X} | \mu, \Sigma, w) = \sum_{n=1}^N \ln \left(\sum_{k=1}^K w_k \mathcal{N}(\mathbf{x}_n | \mu_k, \Sigma_k) \right)$$

Vérifiez la convergence des paramètres ou du logarithme de la vraisemblance. Si le critère de convergence n'est pas satisfait, revenez à l'étape 2.

On considère maintenant l'application de cette forme à variable latente de l'EM au cas spécifique d'un modèle de mélange de gaussiennes pour les données complètes. L'objectif attendu est de maximiser la fonction de vraisemblance logarithmique (3.23), calculée en

utilisant l'ensemble de données observé \mathbf{X} , sachant que cela était plus difficile que dans le cas d'une distribution gaussienne unique en raison de la présence de la sommation sur k qui se produit à l'intérieur du logarithme.

Supposons donc qu'en plus de l'ensemble de données observées \mathbf{X} , nous avons également reçu les valeurs des variables discrètes correspondantes \mathbf{Z} . Alors, on considère le problème de la maximisation de la probabilité pour l'ensemble de données complètes $\{\mathbf{X}, \mathbf{Z}\}$. De (3.19) et (3.20), cette fonction de vraisemblance prend la forme :

$$p(\mathbf{X}, \mathbf{Z} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{n=1}^N \prod_{k=1}^K w_k^{z_{nk}} \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_{nk}} \quad (3.32)$$

Où,

z_{nk} désigne la $k^{\text{ème}}$ composante de \mathbf{z}_n .

En prenant le logarithme, on obtient

$$\ln p(\mathbf{X}, \mathbf{Z} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^N \sum_{k=1}^K z_{nk} \{\ln w_k + \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\} \quad (3.33)$$

La comparaison avec la fonction de vraisemblance logarithmique (3.22) pour les données incomplètes montre que la sommation sur k et le logarithme ont été échangés. Le logarithme agit désormais directement sur la distribution gaussienne, elle-même membre de la famille exponentielle. Sans surprise, cela conduit à une solution beaucoup plus simple au problème du maximum de vraisemblance. Considérons d'abord la maximisation par rapport aux moyennes et covariances. Parce que \mathbf{z}_n est un vecteur de dimension K avec tous les éléments égaux à 0, à l'exception d'un seul élément ayant la valeur 1, la fonction logarithme de la vraisemblance pour les données complètes, est simplement une somme de K contributions indépendantes, une pour chaque composante du mélange. Ainsi, la maximisation par rapport à une moyenne ou une covariance est exactement la même que pour une seule gaussienne, sauf qu'elle ne concerne que le sous-ensemble de points de données qui sont «attribués» à cette composante. Pour la maximisation par rapport aux coefficients de mélange, on note que ceux-ci sont couplés pour différentes valeurs de k grâce à la contrainte de sommation (3.18). Encore une fois, cela peut être appliqué en utilisant un multiplicateur de Lagrange comme précédemment, et conduit au résultat :

$$w_k = \frac{1}{N} \sum_{n=1}^N z_{nk} \quad (3.34)$$

de sorte que les coefficients de mélange sont égaux aux fractions de points de données attribuées aux composants correspondants.

Ainsi, on voit que la fonction logarithmique de la vraisemblance relative aux données complètes peut être maximisée trivialement sous forme fermée. En pratique, on considère en ce qui concerne la distribution postérieure des variables latentes, l'espérance de la fonction logarithmique de la vraisemblance relative aux données complètes. En utilisant (3.19) et (3.20) avec le théorème de Bayes, nous voyons que cette distribution postérieure prend la forme

$$p(\mathbf{Z}|\mathbf{X}, \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto \prod_{n=1}^N \prod_{k=1}^K [w_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)]^{z_{nk}} \quad (3.35)$$

et donc factorise sur n de sorte que sous la distribution postérieure les $\{z_n\}$ sont indépendants. La valeur attendue de la variable indicatrice z_{nk} sous cette distribution postérieure est alors donnée par

$$\mathbb{E}[z_{nk}] = \frac{\sum_{z_{nk}} z_{nk} [w_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)]^{z_{nk}}}{\sum_{z_{nj}} [w_j \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)]^{z_{nj}}} = \frac{w_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K w_j \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} = \gamma(z_{nk}) \quad (3.36)$$

qui est la responsabilité de la composante k pour le point de données \mathbf{x}_n . La valeur attendue de la fonction logarithmique de la vraisemblance relative aux données complètes est donc donnée par

$$\mathbb{E}_Z[\ln p(\mathbf{X}, \mathbf{Z} | \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma})] = \sum_{n=1}^N \sum_{k=1}^K \gamma(z_{nk}) \{\ln w_k + \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\} \quad (3.37)$$

On peut maintenant procéder comme suit. On choisit d'abord des valeurs initiales pour les paramètres $\boldsymbol{\mu}^{old}$, $\boldsymbol{\Sigma}^{old}$ et \mathbf{w}^{old} , et les utilisons pour évaluer les responsabilités (étape E). Nous gardons ensuite les responsabilités fixes et maximisons (3.37) par rapport à $\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}_k$ et w_k (l'étape M). Cela conduit à des solutions sous forme fermée pour $\boldsymbol{\mu}^{new}$, $\boldsymbol{\Sigma}^{new}$ et \mathbf{w}^{new} données par (3.26), (3.28) et (3.31) comme précédemment. Il s'agit précisément de l'algorithme EM pour les mélanges gaussiens tel que dérivé précédemment.

3.4 CLASSIFICATION DES DÉFAUTS DE ROULEMENT PAR LES GMM

3.4.1 Concept de diagnostic par reconnaissance de défauts

Dans les mélanges gaussiens, nous considérons que les variables latentes discrètes peuvent être interprétées comme définissant des affectations de points de données à des composantes spécifiques du mélange. Cela se traduit dans notre problématique par l'affectation automatisée des vecteurs d'attributs (LPC et MFCC) aux différents groupes ou classes représentatives de conditions de défauts, définies au préalable. Ce concept qui peut être intégré dans des systèmes de mesure industriels (en mode online ou en offline), offre d'une manière systématique des possibilités d'une reconnaissance des défauts de roulements dans un cadre de diagnostic automatisé.

3.4.2 Schéma global de la classification par les GMM

Le problème de diagnostic de défaut de roulement peut être modélisé comme un problème de classification, où l'on veut obtenir le meilleur degré de séparabilité entre les classes représentatives des conditions de défaut. Afin d'appliquer ce concept pour construire un système de diagnostic de défaut de roulement capable d'identifier la condition de défaut, une méthodologie basée sur les GMMs a été développée. Ce classifieur a fait ses preuves dans de nombreux problèmes de classification, comme les radars [135], la reconnaissance vocale [136, 137] et la bio-informatique. Basé sur l'algorithme bien connu d'Espérance-Maximisation EM (Expectation-Maximization) pour l'estimation des paramètres, il peut approximer n'importe quelle fonction de densité de probabilité (PDF) par un mélange gaussien fini [138]. Il peut également être complètement représenté avec trois paramètres : les vecteurs de moyenne, les matrices de covariance et les poids du mélange.

Le mélange gaussien construit des modèles pour tous les types de défauts possibles et l'état normal. Le diagnostic du défaut de roulement est obtenu en calculant la probabilité maximale du vecteur des attributs, compte tenu des paramètres du modèle de défaut précédemment construit. En d'autres termes, le GMM détermine l'état du roulement, grâce à l'utilisation de modèles de référence créés suite à l'extraction des indicateurs (ou descripteurs) pertinents à partir de la base d'apprentissage.

Dans cette étude, l'espace des indicateurs (attributs) obtenus est utilisé dans sa totalité son intégrité est gardée intacte. Les procédures habituelles de transformation et réduction des espaces de données, sont volontairement omises car le problème de leur pertinence ne se pose pas dans ce cas de figure (notre problématique consiste à utiliser seulement un seul type d'indicateurs dans chaque processus de classification réalisé). A l'aide de ce schéma de classification, plusieurs configurations seront testées, afin de trouver les paramètres optimaux qui maximisent l'efficacité et la précision de cette approche. Le nombre des itérations est fixé à 100, afin d'éviter que l'algorithme EM, ne soit piégé dans l'un des nombreux maxima locaux de la fonction de vraisemblance.

En conséquence, l'architecture de la plateforme proposée, comme le montre la figure 3.1, comprend quatre étapes principales. La première étape consiste à créer deux ensembles de données distincts, l'un est destiné à la base d'apprentissage (entraînement), et l'autre est destiné la base des tests. Ces données sont en totalité, des signaux de vibration de roulements, ayant subits des prétraitements au préalable. La deuxième étape consiste à l'extraction des indicateurs sélectionnés pour cette étude (LPC et MFCC). Elle est bien clarifiée dans la section

correspondante, pour montrer la pertinence des attributs proposés. La troisième est la construction des modèles, par l'estimation des paramètres du GMM liés à chaque classe et l'élaboration des décisions liées à la discrimination des classes relatives aux conditions de défauts. Enfin, la dernière étape est réservée à l'évaluation des performances (matrice de confusion et probabilité de classification correcte). Le critère de performance choisi est le taux de classification moyen ACR (Average Classification Rate) le plus élevé, prélevé de la matrice de confusion.

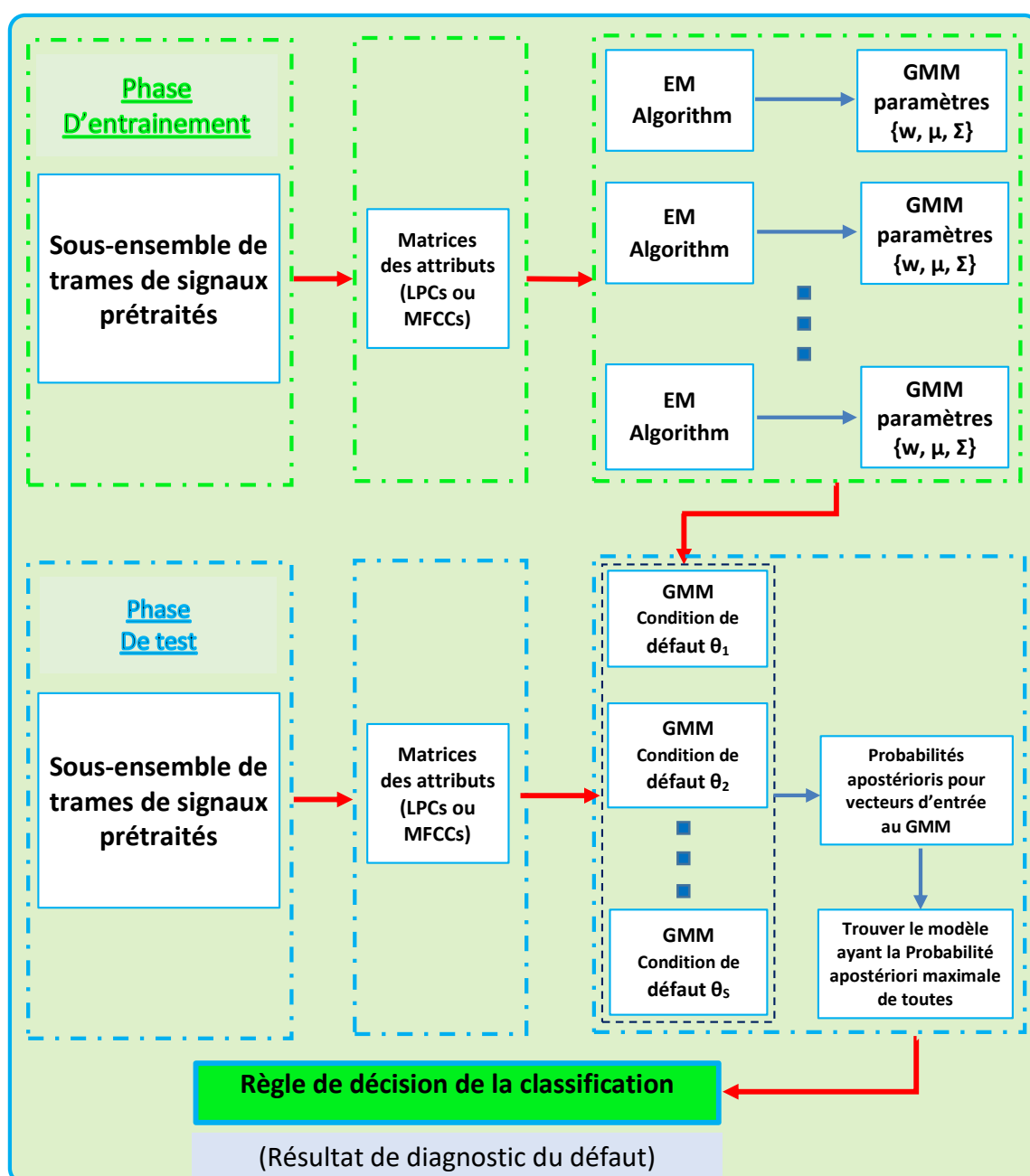


Figure 3.1 : Schéma de classification des deux approches proposées

3.5 MODELES ET ORIGINES DES DONNEES

Les données utilisées dans la classification (diagnostic), sont des signaux vibratoires représentatifs de différentes conditions de défauts de roulement. Dans la pratique, ils sont souvent mesurés à l'aide de capteurs accélérométriques placés au niveau des paliers. Pour mieux comprendre les difficultés liées à leur exploitation et les interprétations y afférentes, il est préférable de préciser les origines des composantes stationnaires et transitoires de ces données, à travers une modélisation des signaux relatifs aux défauts dans les roulements.

3.5.1 Modèles analytiques de roulement

D'une manière globale, trois modèles analytiques sont cités en littérature : le modèle de Palmgren, le modèle de De Mul et le modèle de Bourdon. Le premier sera développé ci-après. Quant aux deux autres, le lecteur peut se référer aux références sous mentionnées pour les procédures développées du calcul relatif à ces modèles.

3.5.1.1 Modèle de Palmgren

Ce modèle de calcul se base sur les hypothèses et considérations suivantes :

- Les déformations structurales des bagues et bâti négligés ;
- Les petits déplacements ;
- Les forces d'inertie négligées ;
- Pas de rotation de la bague intérieure par rapport à la bague extérieure.

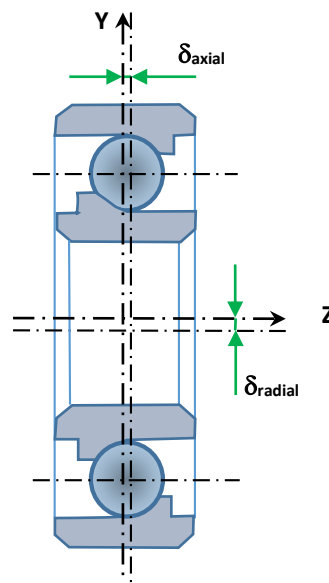


Figure 3.2 : Déplacements estimés à partir du modèle analytique.

Une synthèse des travaux collectés à partir de la littérature [139, 140] montre qu'avec l'estimation des déflexions radiales et axiales (voir la figure 3.2), on obtient des modèles analytiques du comportement, d'un point de vue relationnel, force/déplacement [141].

Le roulement est soumis à un déplacement tel que :

$$\vec{\delta} = \delta_{radial}\vec{Y} + \delta_{axial}\vec{Z} \quad (3.38)$$

Où,

δ_{radial} et δ_{axial} sont les déplacements radiaux et axiaux d'une bague par rapport à l'autre.

La figure 3.3 montre le déplacement local au niveau de la bille, qui s'exprime par :

$$\delta_i = \delta_{axial}\sin(\varnothing) + \delta_{radial}\cos(\varnothing)\cos(\alpha_i) \quad (3.39)$$

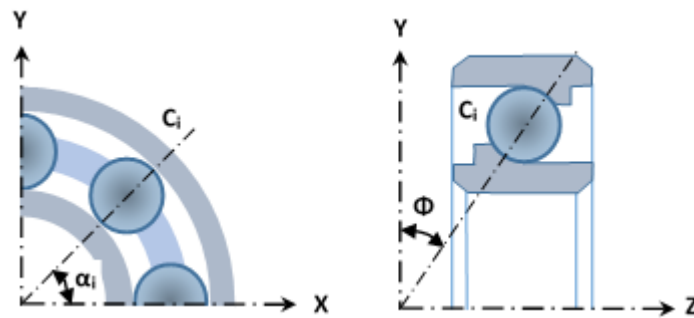


Figure 3.3 : Repères de base du modèle de calcul

En considérant deux solides en contact (bille et bague), il existe une relation entre l'effort et la déflexion occasionnée d'après la théorie d'Hertz :

$$F_i = K\delta_i^{3/2} \quad (3.40)$$

Où,

K : constante liée à la géométrie du roulement ;

F : effort appliqué à la $i^{\text{ème}}$ bille.

Cette relation permet d'étudier l'équilibre global de la bague intérieure du roulement. En sommant sur l'ensemble des billes, les efforts sur le roulement s'écrivent :

$$\begin{cases} F_{axial} = \sum_{i=1}^Z F_i \sin(\varnothing) \\ F_{radial} = \sum_{i=1}^Z F_i \cos(\varnothing)\cos(\alpha_i) \end{cases} \quad (3.41)$$

Où,

Z = nombre de billes

\varnothing = angle de contact

α_i = angle de la bille avec l'axe de référence.

Une force de compression (F_i) sollicite les bagues extérieure et intérieure de chacune des billes ayant rigidité radiale du contact exprimée par les relations suivantes :

$$\begin{cases} F_i = K_{b1} \delta_1^{3/2} \\ F_i = K_{b2} \delta_2^{3/2} \end{cases} \quad (3.42)$$

Avec,

δ_1 et δ_2 : déplacements partiels de chacune des bagues.

L'écrasement total sur la bille (i) peut s'exprimer par la somme des déplacements, tel que :

$$\delta_i = \delta_1 + \delta_2$$

De ce fait, δ_i peut s'écrire :

$$\delta_i = \left(\frac{1}{K_{b1}^{2/3}} + \frac{1}{K_{b2}^{2/3}} \right) F_i^{2/3} \quad (3.43)$$

Ainsi, on déduit que :

$$F_i = \left(\frac{1}{\left(\frac{1}{K_{b1}^{2/3}} + \frac{1}{K_{b2}^{2/3}} \right)} \delta_i \right)^{3/2} \quad (3.44)$$

La rigidité axiale de chaque bille s'écrit :

$$K_b = \frac{K_{b1} K_{b2}}{\left(K_{b1}^{2/3} + K_{b2}^{2/3} \right)^{3/2}} \quad (3.45)$$

La variation de la rigidité et surtout le désalignement des arbres sont les causes principales de l'apparition des vibrations des roulements à billes.

3.5.1.2 Autres modèles

Modèle de De Mul se base essentiellement sur les hypothèses suivantes [142] :

- le frottement interne est négligé ;
- le matériau est supposé élastiquement linéaire ;
- les bagues sont supposées rigides, à l'exception de la déformation de contact locale.

Il permet de déterminer la distribution des charges et des déplacements des éléments roulants du roulement. Les déplacements considérés, ont cinq degrés de liberté. Le modèle prend en compte des effets des forces centrifuges dues aux grandes vitesses. La matrice de rigidité du roulement est calculée analytiquement et utilisée en interne dans la mise en œuvre du calcul itératif de l'équilibre du roulement.

Modèle de Bourdon tient compte d'un éventuel angle de basculement, entre les bagues, intérieure et extérieure [143]. Deux nœuds sont arbitrairement choisis sur les deux bagues de part et d'autre de chaque bille. La connaissance des déplacements de ces nœuds pour chaque

bille, permet de résoudre le problème statique du roulement en utilisant aussi la démarche de De Mul. D'une manière similaire à l'approche précédente, le modèle donne la matrice de rigidité qui peut être exploitée dans des modèles externes par éléments finis de systèmes mécaniques complexes (calculs de rotors), dans le but de prédire le comportement mécanique et les distributions de charge et de déformation.

3.5.2 Modèle dynamique du roulement à bille

Le modèle de roulement pris en considération pour décrire le comportement dynamique dans le cas de présence de défauts localisés de surface et de vitesse de rotation constante. Ce modèle simplifié à 3 degrés de liberté (figure. 3.4), tient compte des deux bagues, externe et interne, de la bille et des films d'huile [144]. Les effets, du socle et du rotor, et l'effet gyroscopique sont négligés.

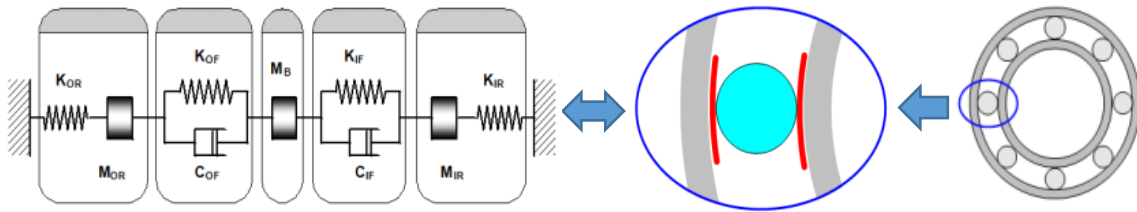


Figure 3.4 : Modèle dynamique simplifié de roulement à 3 degrés de liberté, [144].

Les différents paramètres issus de formulations analytiques et régissant le comportement de ce modèle ont été développés pour les éléments roulants ci-après.

3.5.2.1 Les bagues

Considérées comme des systèmes masse-ressort et ayant un amortissement structural négligeable, les bagues, interne et externe ont leurs rigidités respectives K_{IR} et K_{OR} données par les expressions suivantes :

$$K_{IR} = M_{IR} \cdot \omega_{IR}^2 \quad (3.46)$$

$$K_{OR} = M_{OR} \cdot \omega_{OR}^2 \quad (3.47)$$

Avec, les fréquences naturelles ω_n , sont données par [145] :

$$\omega_n = \frac{n[n^2-1]}{\sqrt{1+n^2}} \sqrt{\frac{EI}{\mu R^4}} \quad (3.48)$$

Où,

E est le module de Young [N/m²] ;

I est le moment d'inertie de la section transversale de la bague [m⁴] ;

μ est la masse par unité de longueur [Kg/m] ;

R est le rayon moyen de la bague [m] ;

n est l'ordre du mode de flexion.

Pour $n = 2$,

$$\omega_n = 2.68 \sqrt{\frac{EI}{\mu R^4}}$$

3.5.2.2 Les Billes

D'une manière générale, il est indiqué dans la littérature que les rigidités des billes sont nettement supérieures à celles des bagues (choix de matériaux spéciaux pour les billes, etc.). A titre d'exemple, elle est de l'ordre de $8.3 \cdot 10^9$ N/m pour un roulement de type SKF 6206. D'où, les billes sont considérées souvent comme infiniment rigides. Ainsi, la modélisation de la bille est considérée uniquement comme un élément de masse.

3.5.2.3 Le film d'huile

Sous un régime de lubrification Elasto-Hydro-Dynamique (EHD), le film d'huile peut se modéliser comme rigidité et amortissement, moyennant l'utilisation des expressions suivantes [146] :

$$K = \frac{4}{w_r \lambda_k^2} \left[\frac{\varepsilon_0}{(1-\varepsilon_0^2)^2} \sin^2 \phi_0 + \frac{3\pi\varepsilon_0^2}{4(1-\varepsilon_0^2)^{5/2}} \sin \phi_0 \cos \phi_0 + \frac{2\varepsilon_0(1+\varepsilon_0^2)}{(1-\varepsilon_0^2)^3} \cos^2 \phi_0 \right] \quad (3.49)$$

$$C = \frac{4}{w_r \lambda_k^2} \left[\frac{\pi}{2(1-\varepsilon_0^2)^{3/2}} \sin^2 \phi_0 + \frac{4\varepsilon_0}{4(1-\varepsilon_0^2)^2} \sin \phi_0 \cos \phi_0 + \frac{\pi(1+2\varepsilon_0^2)}{2(1-\varepsilon_0^2)^{5/2}} \cos^2 \phi_0 \right] \quad (3.50)$$

Où,

$$\frac{4}{w_r \lambda_k^2} = \frac{(1-\varepsilon_0^2)^2}{\varepsilon_0 [16\varepsilon_0^2 + \pi^2(1-\varepsilon_0^2)]^{1/2}} \cos^2 \phi_0 \quad (3.51)$$

$$\tan \phi_0 = \frac{\pi(1-\varepsilon_0^2)^{1/2}}{4\varepsilon_0} \quad (3.52)$$

$$\varepsilon_0 = 1 - \frac{h}{c} \quad (3.53)$$

Avec,

ε_0 est un coefficient d'excentricité ;

c est le jeu diamétral du roulement ;

h est l'épaisseur du film d'huile, calculée d'après la théorie EHD.

3.5.2.4 Equation du mouvement

Une fois les paramètres du modèle dynamique (masse-ressort) du roulement seul connus, le système d'équations différentielles peut s'écrire sous la forme :

$$[M].\{\ddot{x}\} + [C].\{\dot{x}\} + [K].\{x\} = \{F\} \quad (3.54)$$

Avec,

$$[M] = \begin{bmatrix} M_{OR} & 0 & 0 \\ 0 & M_B & 0 \\ 0 & 0 & M_{IR} \end{bmatrix} \quad (3.55)$$

$$[C] = \begin{bmatrix} C_{OF} & -C_{OF} & 0 \\ -C_{OF} & C_{OF} + C_{IF} & -C_{IF} \\ 0 & -C_{IF} & C_{IF} \end{bmatrix} \quad (3.56)$$

$$[K] = \begin{bmatrix} K_{OR} + K_{OF} & -K_{OF} & 0 \\ -K_{OF} & K_{OF} + K_{IF} & -K_{IF} \\ 0 & -K_{IF} & K_{IR} + K_{IF} \end{bmatrix} \quad (3.57)$$

Où,

$[M]$, $[C]$ et $[K]$ représentent respectivement les matrices de masses, d'amortissement et de rigidité du roulement ;

$$\{F\} = \begin{Bmatrix} F_{OR} \\ F_B \\ F_{IR} \end{Bmatrix}$$

$\{F\}$ est le vecteur des forces d'excitation.

$$\{x\} = \begin{Bmatrix} x_{OR} \\ x_B \\ x_{IR} \end{Bmatrix}$$

$\{x\}$ est le vecteur des déplacements.

3.5.3 Modèle d'excitation générée par des défauts localisés

Un signal de vibration du roulement est la réponse dynamique de ce dernier mesurée par un capteur (accéléromètre), aux forces d'excitation produites par la présence de défauts. Ces forces peuvent provenir de différentes sources à savoir : le défaut du chemin de roulement intérieur, le défaut du chemin de roulement extérieur et le défaut de l'élément roulant (la bille dans notre cas).

Un défaut de roulement génère également une force impulsive donnée par [33] :

$$F = \sum_{k=1}^{\infty} A_3^k \delta \left(t - \frac{k}{f_c} \right) \approx f_s \sum_{m=1}^{N_2} A_3^m \cos(2\pi m f_c t + \theta_2^m) \quad (3.58)$$

Où,

A_3^m est l'amplitude appropriée de la $m^{\text{ième}}$ harmonique déterminée par la taille du défaut, f_c est la fréquence caractéristique de l'élément porteur défectueux, θ_2^m est la phase initiale de la $m^{\text{ième}}$ harmonique et N_2 est le nombre d'harmoniques dans le train d'impulsions.

Pour prendre en compte le glissement des éléments roulants, une incertitude aléatoire peut être ajoutée au temps d'apparition de chaque impulsion de telle sorte que

$$F = \sum_{k=1}^{\infty} A_3^k \delta \left(t - \frac{k}{f_c} + \delta_k \right) \quad (3.59)$$

Où,

δ_k est généralement de quelques pour cent de $\frac{1}{f_c}$

En supposant une position fixe pour le capteur de vibration, l'équation (3.59) doit être modifiée pour tenir compte des variations du trajet de transmission pour les défauts de rotation (défauts de bille et bague intérieure). Lorsque le défaut tourne, la distance du défaut à un capteur fixe change périodiquement. De plus, dans le cas d'une répartition de charge non uniforme, lorsqu'un défaut rotatif frappe les chemins de roulement, il se fera sentir différemment dans les différentes zones de charge du roulement. Ceux-ci peuvent être traduits en force d'entrée par modulation avec la série d'impulsions cosinus appropriée donnée par

$$\sum_{k=1}^{N_3} A_4^k \cos(2\pi k f_q t + \theta_3^k) \quad (3.60)$$

Où,

A_4^k et θ_3^k sont l'amplitude initiale et la phase de la $k^{\text{ième}}$ harmonique, respectivement,

N_3 est le nombre d'harmoniques dans le train d'impulsions,

f_q est la fréquence du train d'impulsions.

Un défaut de roulement augmente le frottement en rotation du rotor et, par conséquent, chaque défaut génère des spectres de vibration avec des composantes de fréquence uniques, qui sont une fonction linéaire de la vitesse de fonctionnement. Les deux fréquences de piste sont également des fonctions linéaires du nombre de billes [147].

Les systèmes de surveillance de l'état des roulements du moteur sont mis en œuvre en analysant le signal de vibration de tous les défauts de roulement. Le signal de vibration est produit par l'impulsion d'impact générée à chaque passage de la bille sur un défaut dans les pistes de roulement.

3.5.4 Modèle de signaux générés par la présence de défauts localisés

Les types de données nécessaires à la classification, sont des signaux de vibration, comme le modèle simulé de la figure 3.5. Ainsi, un signal (t) , émis par un roulement est le signal résultant mesuré par un capteur (accéléromètre), des effets des forces d'excitation générées par la présence de défauts et de l'environnement de la machine [134, 148].

Pour un modèle de défaut localisé sur un des éléments du roulement, on a :

$$x(t) = x_f(t) + x_q(t) + x_{bs}(t) + x_s(t) + n(t) \quad (3.61)$$

Où,

$x_f(t)$ est la série d'impulsions de base produite par l'impact du défaut de manière répétitive

avec une période $T = \frac{1}{f_f}$;

f_f est la fréquence caractéristique du défaut ;

$x_q(t)$ est l'effet de modulation provoqué par une distribution de charge non uniforme des roulements et la variation cyclique du trajet de transmission entre le site d'impact du défaut et le transducteur (capteur) ;

$x_{bs}(t)$ est la vibration induite par le roulement. C'est l'effet de modulation causé en partie par la géométrie du roulement, mais plus significativement par la position du capteur de vibration par rapport à l'emplacement du défaut.

$x_s(t)$ est la vibration induite par les machines. C'est un signal harmonique amorti indésirable, structuré et prévisible qui est également considéré comme un bruit.

$n(t)$ est un bruit blanc gaussien de variance σ^2 .

1) Impulsions dues à une défaillance

$$x_f(t) = d_0 \sum_{j=1}^{\infty} \delta(t - jT_f) \approx A_0 \sum_{l=1}^{N_f} \cos(2\pi l f_f t + \phi_f^l) \quad (3.62)$$

Où,

$A_0 = d_0 f_f$ est l'amplitude ;

ϕ_f^l est la phase initiale de la $l^{\text{ème}}$ harmonique ;

N_f est le nombre des harmoniques induites par les séries d'impulsions ;

f_f est la fréquence caractéristique de défaut (f_{BPFI} , f_{BPFO} , f_{BSF})

2) Chargement non uniforme

$$x_q(t) = \sum_{k=1}^{N_q} A_q^k \cos(2\pi k f_q t + \phi_q^k) \quad (3.63)$$

Où,

Laquelle est une fonction périodique de période T_q ;

$f_q = \frac{1}{T_q}$;

A_q^k est l'amplitude de la $k^{\text{ème}}$ harmonique ;

ϕ_q^k est la phase initiale de la $k^{\text{ème}}$ harmonique ;

N_q est le nombre des harmoniques ;

f_q dépend des éléments du roulement défaillant sous différentes charges et chemins de transmission.

3) Vibration induite par le roulement

$$x_{bs}(t) = \sum_{j=1}^{N_{bs}} A_{bs}^j e^{-\alpha_{bs}^j t} \cos(2\pi f_{bs}^j t + \phi_{bs}^j) \quad (3.64)$$

Où,

f_{bs}^j , A_{bs}^j , ϕ_{bs}^j , et α_{bs}^j sont respectivement, la fréquence de résonance, l'amplitude, la phase initiale et le facteur d'amortissement du $j^{\text{ème}}$ mode de vibration du roulement ;

N_{bs} est l'ordre modale du signal de vibration du roulement.

4) Vibration induite par la machine

$$x_s(t) = \sum_{n=1}^{N_s} A_s^n e^{-\alpha_s^n t} \cos(2\pi f_s^n t + \phi_s^n) \quad (3.65)$$

Où,

f_s^n , A_s^n , ϕ_s^n , et α_s^n sont respectivement, la fréquence de résonance, l'amplitude, la phase initiale et le facteur d'amortissement du $n^{\text{ème}}$ mode de vibration de la machine ;

N_s est l'ordre modale du signal de vibration de la machine.

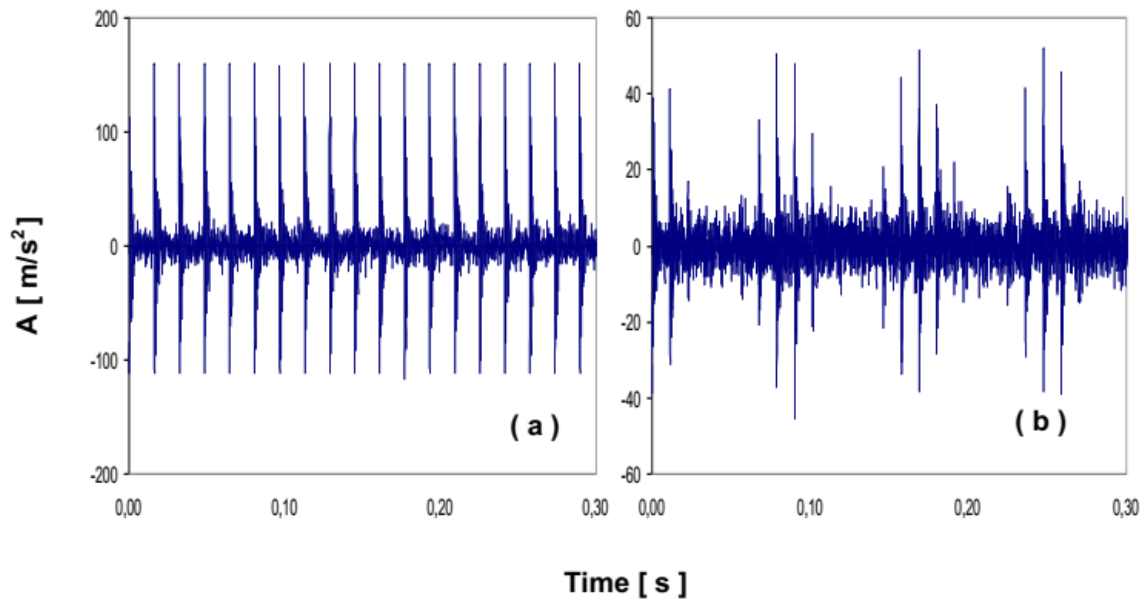


Figure 3.5 : Réponse temporelle d'un roulement endommagé du type SKF1210, à $v=11.6$ Hz.

(a) défaut localisé (1mm) sur la BE ; (b) défaut localisé sur la BI [144].

3.5.5 Prétraitements appliqués aux données

Les données utilisées dans la classification (diagnostic), sont des signaux représentatifs de différentes conditions de défauts de roulement. Ils sont mesurés à l'aide de capteurs accélérométriques placés au niveau des paliers. Ces signaux dynamiques sont généralement mesurés sous forme d'amplitude d'une grandeur physique variant en fonction du temps. Il y a donc deux échelles (axes) à considérer :

- L'amplitude instantanée du signal, représentative de la quantité physique (Accélération)

- Le Temps : l'axe des temps affecté.

Pour des considérations pratiques, généralement on procède à des arrangements de sorte que les expériences produisent des signaux stationnaires. Ainsi, on force la stationnarité en maintenant les conditions expérimentales constantes. A défaut d'y arriver, les analyses doivent se limiter à des portions de signal exploitables. Par ailleurs, le nombre N de mesures réalisables (la taille de l'ensemble des réalisations accessibles), ainsi que la durée T des signaux mesurés sont toujours finis. En conséquence, les valeurs moyennes mesurées ne sont que des estimations et les erreurs d'évaluation doivent être prises en compte. En plus de ces erreurs statistiques, il faut ajouter celles liées à la chaîne de mesure, issues des capteurs, du conditionnement des signaux, de la numérisation, etc.

Pour éviter que ces erreurs ne s'ajoutent encore à celles dues à la méthode de classification adoptée dans cette thèse, une procédure de prétraitement est employée. Pour chaque signal (segment ou trame), la valeur moyenne, aussi appelée composante continue, non nulle pour cause d'artéfact de la chaîne d'acquisition (offset électronique, tension de polarisation du capteur, etc.), doit donc être systématiquement soustraite. Ensuite, l'ensemble des signaux doivent être normalisés, avant d'être filtrés en dernier lieu.

3.5.6 Organisation et structuration des données

Il est primordial de disposer d'une base de données consistante pour valider les performances des algorithmes développés pour la classification. Les signaux sélectionnés sont traités et répartis en deux sous-ensembles, l'un est destiné à l'apprentissage, tandis que l'autre est réservé aux tests. Les détails de la base seront exposés dans la phase de validation expérimentale.

3.6 EXTRACTION DES INDICATEURS DISCRIMINANTS

Lors de cette étude deux approches distinctes ont été proposées. La première est basée sur les coefficients de prédiction Linéaire ou appelés aussi Codage Prédicatif Linéaire, LPC (Linear Predictive Coding). La seconde est basée sur les Coefficients Cepstraux à l'échelle de Fréquence Mel, MFCC (Mel Frequency Cepstral Coefficient).

3.6.1 Approche basée sur les LPCs

Dans cette partie, sont développés les codes numériques relatifs à la classification par les GMMS et basés sur les indicateurs de type LPC. L'ensemble des codes liés aux différentes étapes de la classification, forment l'algorithme de cette première approche de classification représenté par la figure 3.6.

Les modèles de calculs théoriques des attributs LPC et MFCC, sont développés ci-après. Concernant la modélisation en prédiction linéaire, un point intéressant sera consacré à l'estimation de l'erreur normalisée selon l'ordre du modèle.

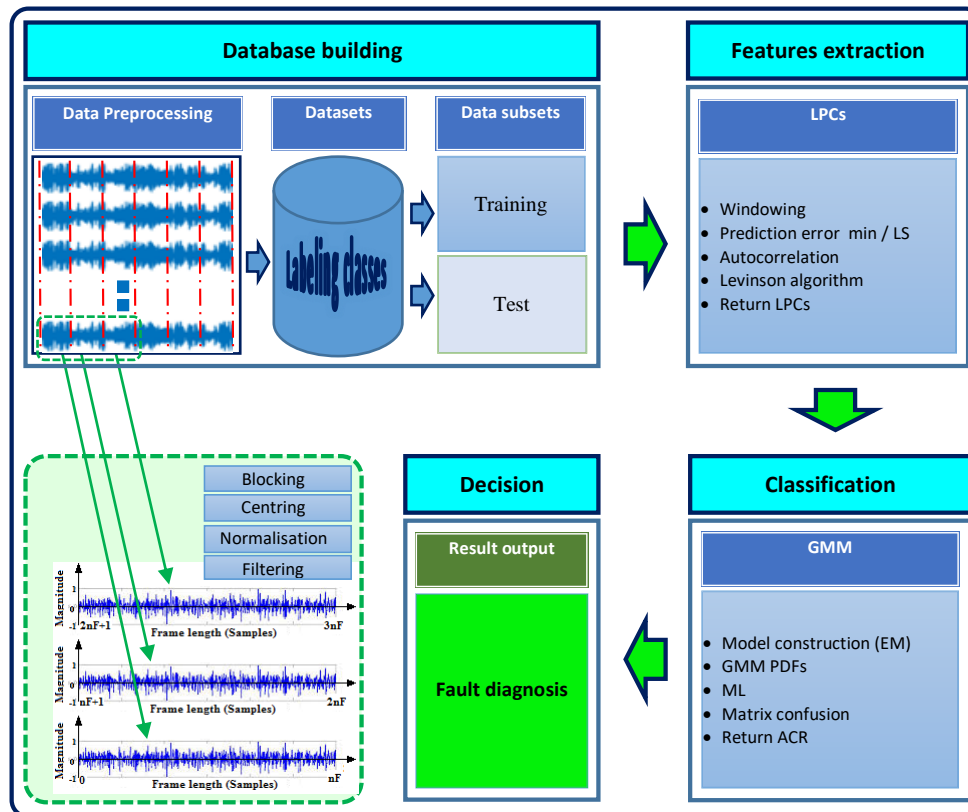


Figure 3.6 : Schéma de la procédure de classification avec les LPCs.

3.6.1.1 Calcul des LPCs

La prédiction linéaire est un outil puissant qui peut être utilisé pour l'analyse (modélisation) de séries temporelles telles que les signaux de vibration. Dans une telle approche, les coefficients de prédiction linéaire, sont considérés comme des vecteurs d'attributs (indicateurs) d'un problème de classification des défauts de roulement.

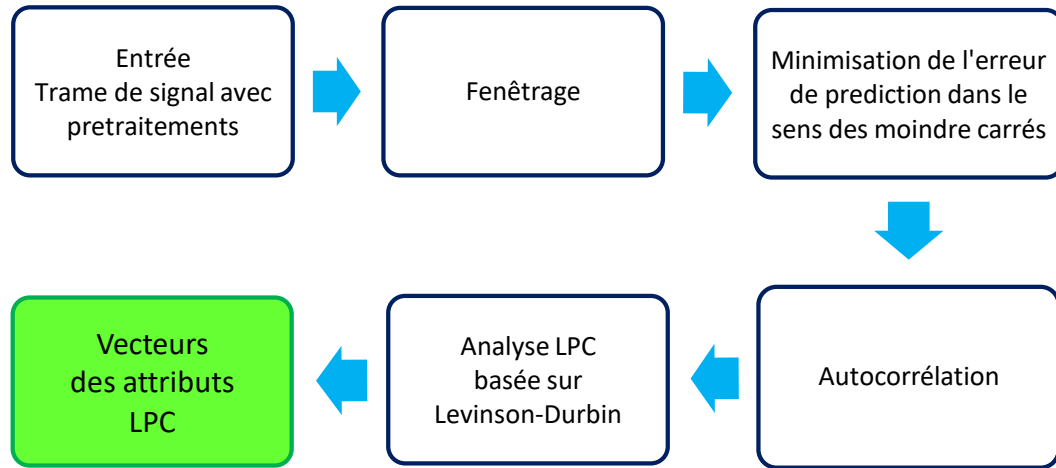


Figure 3.7 : Organigramme schématisant les étapes d'obtention des LPCs

Ils peuvent être obtenus à partir du développement d'un modèle de prédiction linéaire, en minimisant l'erreur de prédiction dans le sens des moindres carrés [149]. Les différentes étapes nécessaires à l'extraction de ces paramètres à partir des signaux de vibration sont montrées dans la figure 3.7, et décrites ci-dessous.

L'échantillon prédit $\hat{s}(n)$ considéré comme une combinaison linéaire de p échantillons précédents formant une analyse de prédiction linéaire, où p est l'ordre de prédiction, peut être représentée par :

$$\hat{s}(n) = -\sum_{k=1}^p a_k s(n-k) \quad (3.66)$$

Où,

$\{a_k\}_{k=1,\dots,p}$ sont les coefficients de prédiction linéaires

$s(n)$ est la séquence fenêtrée obtenue par la multiplication d'une trame du signal temporel par une fenêtre de Hamming ou similaire. Ainsi, on a :

$$s(n) = x(n)w(n) \quad (3.67)$$

Où,

$w(n)$ est la fenêtre de la séquence.

L'erreur de prédiction $e(n)$ peut être évaluée par la différence de l'échantillon actuel $s(n)$ et celui prédit (estimé) $\hat{s}(n)$. Soit :

$$e(n) = s(n) - \hat{s}(n) \quad (3.68)$$

$$e(n) = s(n) + \sum_{k=1}^p a_k s(n-k) \quad (3.69)$$

L'objectif de cette analyse de prédiction linéaire est de calculer les LPCs qui minimisent l'erreur de prédiction $e(n)$, par la méthode de l'autocorrélation aux moindres carrés. Ceci, peut être accompli par la minimisation de l'erreur de prédiction totale, représentée par :

$$E = \sum_{n=-\infty}^{\infty} e^2(n) \quad (3.70)$$

En utilisant sa valeur dans l'équation (3.69), on obtient :

$$E = \sum_{n=-\infty}^{\infty} [s(n) + \sum_{k=1}^p a_k s(n-k)]^2 \quad (3.71)$$

Les valeurs de $\{a_k\}_{k=1,\dots,p}$ qui minimisent l'erreur de prédiction totale E , peuvent être retrouvées de :

$$\frac{\partial E}{\partial a_k} = 0$$

Pour $k = 0, 1, 2, \dots, p$

Alors, on écrit :

$$\frac{\partial E}{\partial a_k} = \frac{\partial}{\partial a_k} \sum_{n=-\infty}^{\infty} [s(n) + \sum_{k=1}^p a_k s(n-k)]^2 = 0 \quad (3.72)$$

L'expression de différentiation peut être exprimée par :

$$\sum_{n=-\infty}^{\infty} s(n-i)s(n) = \sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s(n-i)s(n-k) \quad (3.73)$$

Où,

$i = 1, 2, 3, \dots, p$.

L'équation (3.66) peut être écrite en termes de séquence d'autocorrélation $R(i)$ suivante :

$$\sum_{k=1}^p a_k R(i-k) = R(i) \quad \text{pour } i = 1, 2, 3, \dots, p \quad (3.74)$$

Où,

La séquence d'autocorrélation utilisée dans l'équation (3.70) peut être écrite comme suit :

$$R(i) = \sum_{n=i}^{N-1} s(n)s(n-i) \quad (3.75)$$

Pour, $i = 1, 2, 3, \dots, p$ et N est la longueur de la séquence.

Sous la forme matricielle :

$$R.A = -r \quad (3.76)$$

Où,

R est la matrice symétrique $p * p$ composée des éléments $R(i, k) = R(|i - k|)$, ($1 \leq i, k \leq p$);

r est le vecteur colonne ayant les éléments $(R(1), R(2), \dots, R(p))$;

A est le vecteur colonne des coefficients LPC $(a(1), a(2), \dots, a(p))$.

Il peut être montré que R est une matrice de Toeplitz qui peut être représentée par :

$$R = \begin{bmatrix} R(1), R(2), R(3), & \dots & R(p) \\ R(2), R(1), R(2), & \dots & R(p-1) \\ R(3), R(2), R(1), & & R(p-2) \\ \vdots & \ddots & \vdots \\ R(p), R(p-1), R(p-2), & \dots & R(1) \end{bmatrix} \quad (3.77)$$

Les coefficients LPC peuvent être déterminés par :

$$A = -R^{-1}.r \quad (3.78)$$

Où,

R^{-1} est la matrice inverse de R .

Comme R a une structure de Toeplitz, son inverse R^{-1} peut être déterminé par l'algorithme de Levinson [150].

3.6.1.2 Erreur normalisée

L'erreur normalisée peut être définie comme le rapport entre l'erreur minimale totale et l'énergie totale du signal. Si E_p est l'erreur minimale totale obtenue pour un ordre de prédiction p et $R(1)$ est l'énergie totale du signal, alors l'erreur normalisée V_p peut être représentée par :

$$V_p = \frac{E_p}{R(1)} \quad (3.79)$$

En combinant les équations (3.71) et (3.73), l'expression de l'erreur totale minimum pour un ordre de prédiction donné, est donnée par :

$$E_p = \sum_{n=-\infty}^{\infty} s^2(n) + \sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s(n).s(n-k) \quad (3.80)$$

Soit, en termes de séquence d'autocorrélation, l'expression :

$$E_p = R(1) + \sum_{k=1}^p a_k \quad (3.81)$$

3.6.2 Approche basée sur les MFCCs

Dans cette partie, sont développés les codes numériques relatifs à la classification par les GMMS et basés sur les indicateurs de type MFCC. L'ensemble des codes liés aux différentes étapes de la classification, forment l'algorithme de cette première approche de classification représenté par la figure 3.8.

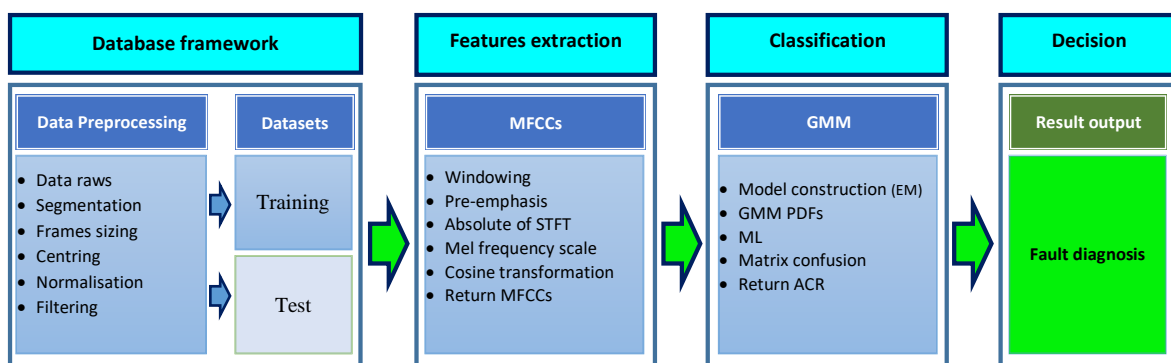


Figure 3.8 : Schéma de la procédure de classification avec les MFCCs.

Le processus consiste à transformer le signal du domaine temporel au domaine fréquentiel (Hertz), puis transformer ce dernier à l'échelle de Mel [151]. Pour le calcul des MFCC, il est nécessaire de suivre les étapes montrés dans la figure 3.9, et indiquées ci-dessous.

Le filtrage de préaccentuation, est un type spécial de filtre à réponse impulsionnelle finie (FIR). Il peut être utilisé pour compenser la fluctuation de l'énergie du spectre entre les basses fréquences et les hautes, fréquences.

Soit $x[n]$ le signal à l'échantillon n et $s[n]$ le signal après un filtrage passe-haut.

$$s[n] = x[n] - \alpha x[n - 1], \quad n = 1, 2, \dots, N \quad (3.82)$$

Où,

α est un paramètre de contrôle du filtrage, souvent choisi entre 0.95 et 1 en pratique.

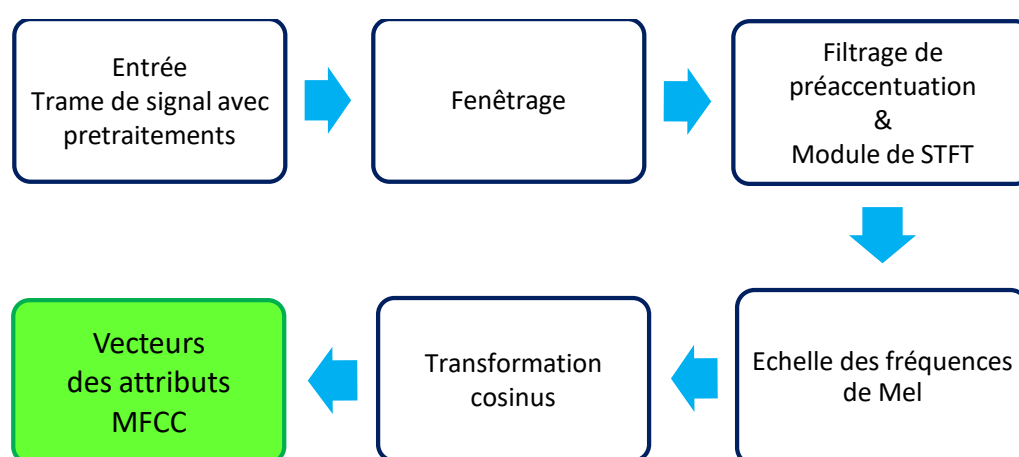


Figure 3.9 : Organigramme schématisant les étapes d'obtention des MFCCs

L'étape suivante consiste à transformer le signal du domaine temporel au domaine fréquentiel, en appliquant une transformation de Fourier à court terme (STFT). La fenêtre comme montré dans la figure 3.10, doit avoir une période suffisamment courte pour qu'elle puisse être considérée presque stationnaire, et donc être transformée en domaine fréquentiel.

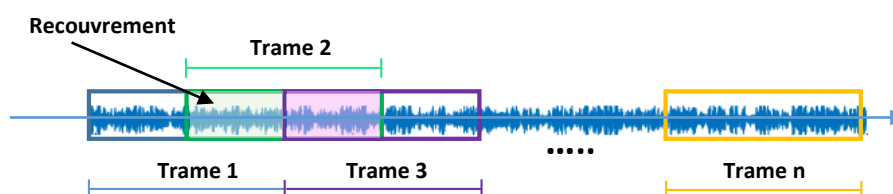


Figure 3.10 : Fenêtrage appliqué pour minimiser les discontinuités aux limites des trames

Cela peut être fait par :

$$X_a[k] = \sum_{n=0}^{N-1} s[n] \cdot w_a[n] \cdot e^{-\frac{i2\pi kn}{N}} = \sum_{n=0}^{N-1} s[n] \cdot w_a[n] \cdot e^{-i\omega k}, \quad 0 \leq k < N \quad (3.83)$$

Où,

$w_a[n]$ est la fonction de fenêtre, partout nulle sauf à l'intérieur de la fenêtre ;

i est le nombre imaginaire.

Pour garder les trames continues, une fenêtre de Hamming est préférée.

$$w_a[n] = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n < N-1, \quad \alpha = 0.54, \quad \beta = 1 - \alpha = 0.46 \quad (3.84)$$

Le fait que la bande de fréquence auditive humaine est sensible entre 20 et 1000 Hz, la même échelle est appliquée au signal aussi bien aux hautes que les basses, fréquences. Une projection des données doit être opérée entre l'échelle de Hertz et celle de Mel.

$$Mel = \begin{cases} f, & f \leq 1000 \\ 2595 \log_{10}\left(1 + \frac{f}{700}\right), & f > 1000 \end{cases} \quad (3.85)$$

En plus, son inverse est donné par :

$$f = \begin{cases} Mel, & Mel \leq 1000 \\ 700 \left(e^{Mel/2595} - 1\right), & Mel > 1000 \end{cases} \quad (3.86)$$

Pour une STFT d'une trame fenêtrée donnée $x_a[k]$, nous définissons une bande de M filtres ($m = 1, 2, \dots, M$) qui sont linéaires à l'échelle de Mel et non linéaires à l'échelle de Hz. Où,

m est un filtre triangulaire donné par :

$$M_m[k] = 1 - \left| \frac{k - \frac{N-1}{2}}{\frac{N-1}{2}} \right| \quad (3.87)$$

Avec, N est la longueur du filtre.

Il est à noter que ces filtres linéaires à l'échelle de Mel ont besoin d'être transformés en retour à l'échelle d'Hertz. D'où, le calcul du logarithme de l'énergie de chaque filtre, comme suit :

$$S[m] = \ln\left[\sum_{k=0}^{N-1} |X_a[k]|^2 M_m[k]\right], \quad 0 < m \leq M. \quad (3.88)$$

Finalement, Les coefficients cepstraux en fréquence de Mel (MFCCs) sont la transformée cosinus discrète des sorties des M filtres :

$$c[q] = \sum_{m=0}^{M-1} \left[S[m] \cos\left(\frac{\pi q \left(m - \frac{1}{2}\right)}{M}\right) \right], \quad 0 < m \leq M. \quad (3.89)$$

3.7 CONSTRUCTION DE MODELES

Cette étape est le cœur du processus de classification. Elle permet notamment de construire les modèles d'estimation des paramètres du mélange sur la base des données d'apprentissage et l'établissement des règles de décision conséquentes nécessaires à la classification des données de test.

La distribution de probabilité empirique des données échantillonnées peut être estimée par un GMM en utilisant une combinaison linéaire de distributions de gaussiennes. Il s'agit d'un modèle statistique qui utilise une somme pondérée des fonctions de densité de probabilité de multiples distributions gaussiennes pour approcher la distribution empirique des données échantillonnées. Dans notre cas comme le montre la figure 3.11, ces données sont formées par les attributs (LPC ou MFCC) concaténés extraits à partir de différentes trames de signaux de vibrations.

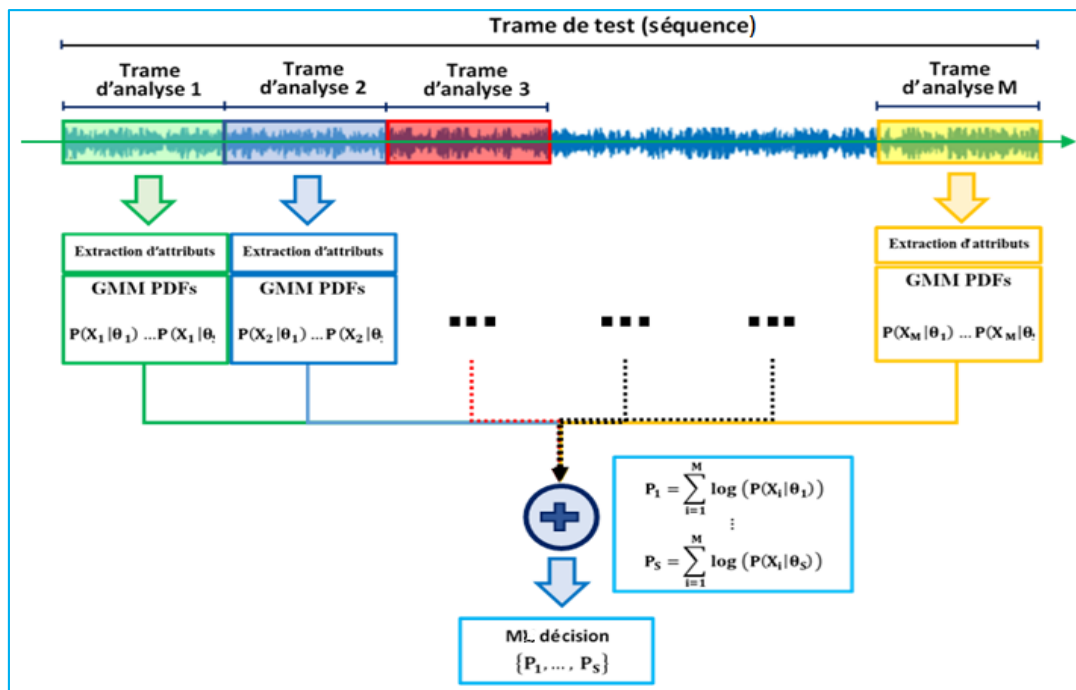


Figure 3.11 : Extraction des attributs à partir des trames d'analyse et de test

Un GMM est une somme pondérée de M composantes de densités gaussiennes. La densité du mélange gaussien $p(\mathbf{x}|\theta)$ est donnée par l'équation :

$$p(\mathbf{x}|\theta) = \sum_{i=1}^M w_i p_i(\mathbf{x}) \quad (3.90)$$

Pour un vecteur \mathbf{x} à D dimensions, la distribution gaussienne multivariée prend la forme :

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} \quad (3.91)$$

Où,

D est le nombre des attributs ou indicateurs (LPCs ou MFCCs).

Le modèle de mélange gaussien complet θ est paramétré par les poids du mélange w_i , les vecteurs de moyennes μ_i et les matrices de covariance Σ_i de toutes les densités des M composantes du mélange [132]. Ces paramètres sont représentés ensemble par la notation :

$$\theta = \{\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\} \quad (3.92)$$

Où,

θ est le modèle

$$\mathbf{w} = \{w_i\}, i = 1, 2, \dots, M$$

$$\boldsymbol{\mu} = \{\mu_i\}, i = 1, 2, \dots, M$$

$$\boldsymbol{\Sigma} = \{\Sigma_i\}, i = 1, 2, \dots, M$$

La fonction de densité est donc :

$$p_i(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) \quad (3.93)$$

La description du GMM est une superposition linéaire de gaussiennes,

$$p(\mathbf{x}) = \sum_{i=1}^M w_i \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) \quad (3.94)$$

Les conditions de normalisation et de positivité nécessitent :

$$\sum_{i=1}^M w_i = 1 \quad 0 \leq w_i \leq 1 \quad (3.95)$$

Pour une collection ou une série de T vecteurs d'apprentissage donnée, les paramètres de ce modèle sont estimés par l'algorithme Espérance-Maximisation (EM). Ce dernier permet de trouver les paramètres optimaux du modèle, de manière itérative. Ainsi, l'identification des paramètres du modèle qui maximisent la probabilité du GMM, compte tenu des données d'apprentissage, pour augmenter la probabilité du modèle estimé pour le vecteur de caractéristiques de défaut de roulement donné, s'obtient via l'estimation du maximum de vraisemblance.

En d'autres termes, le diagnostic de défaut pour chaque condition de défaut (classe) est représentée par un modèle GMM. Étant donné que les attributs des défauts sont extraits de la base d'apprentissage (entraînement) relative à chaque condition de défaut, l'objectif est d'estimer les paramètres GMM des modèles correspondants. Cela peut être réalisé en utilisant l'algorithme d'Espérance-Maximisation (EM), qui recherche une estimation du maximum de vraisemblance [132, 134].

$$\hat{s} = \arg \max_{1 \leq k \leq C} \sum_{t=1}^T \log p(\mathbf{x}_t | \theta_k) \quad (3.96)$$

Où,

k représente l'indice du type de défaut

C est le nombre total des conditions de défauts (classes) connues.

$\mathbf{x} = \{x_1, x_2, \dots, x_T\}$ est le vecteur des attributs inconnu relatif aux défauts

$p(x_t | \theta_k)$ est donnée par l'équation (3.91).

Une fois que les GMM ont été entraînés pour l'ensemble des conditions de défaut C , le diagnostic d'un vecteur d'attributs de défaut de test devient simple. Nous avons juste besoin de

trouver le modèle θ , qui a la probabilité a posteriori maximale pour un vecteur d'attributs de défaut de test donné.

3.8 TESTS DE PERFORMANCE

Le but de cette modélisation est d'estimer des métriques de performance indépendantes des attributs, des modèles d'apprentissages et du schéma de classification utilisé. La matrice de confusion au Tableau 3.1, est l'outil qui permet de définir les performances de classification [135]. Parmi ces performances, on y trouve la probabilité de classification correcte P_{cc} ou bien le taux de bonne classification moyen donnée par :

$$P_{cc} = \sum_{i=1}^K P(X_i)P_{cc}^i \quad (3.97)$$

Où,

P_{cc}^i est la probabilité de classification correcte de la $i^{\text{ème}}$ classe X_i .

$P(X_i)$ est la probabilité d'avoir la classe X_i .

P_{ij} est la probabilité d'avoir une fausse classification (classifier la $i^{\text{ème}}$ classe X_i comme la classe X_j).

Comme cela a été déjà mentionné antérieurement, la première partie de la base de données est utilisée pour l'apprentissage et la seconde partie est réservée aux calculs des performances de classification qui sont la matrice de confusion donnée dans le Tableau 1 et le taux de bonnes classifications moyen donné à l'équation (3.97).

Tableau 3.1 : Modèle de matrice de confusion d'un classifieur.

		Classes déclarées (obtenues)				
		X_1	X_2	X_K
Classes sous test	X_1	P_{cc}^1	P_{12}	P_{1K}
	X_2	P_{21}	P_{cc}^2	P_{2K}
	⋮	⋮	⋮	⋮		⋮
	⋮	⋮	⋮		⋮	⋮
	X_K	P_{K1}	P_{K2}	P_{cc}^K

La probabilité P_{ij} peut s'écrire par :

$$P_{ij} = \frac{d_{ij}}{M_i} \quad (3.98)$$

Où,

d_{ij} est le nombre de réalisations ayant la classe X_j en sortie sachant que la classe X_i était en entrée.

M_i est le nombre total des tests sur la classe X_i

Pour $j = i$, l'équation (3.97) estime la probabilité de classification correct de la classe X_i notée P_{cc}^i .

Plus la matrice de confusion est à diagonale dominante, plus la classification est performante. Souvent, on prend les classes $\{X_i\}_{i=1,\dots,K}$ équiprobables. Dans ce cas, le taux de bonnes classifications moyen est la moyenne des valeurs de la diagonale de la matrice de confusion.

3.9 CONCLUSION

Dans ce chapitre, nous avons abordé le problème de la classification des défauts de roulement par les modèles de mélange de gaussiennes. Nous avons tout d'abord présenté en revue des éléments théoriques fondamentaux sur les modèles de mélange gaussiens et leur contexte d'utilisation. Un accent particulier a été mis sur l'algorithme d'Espérance-Maximisation EM et son rôle principal dans l'estimation des paramètres des modèles par la maximisation de la fonction de vraisemblance. A partir de là, l'approche proposée de classification basée sur les GMMs, ne nécessitant pas l'élaboration de modélisations complexes, nous a été confirmée comme adéquate et la plus adaptée à la problématique de diagnostic des défauts de roulement dans un espace de données relativement réduit.

Le diagnostic basé sur la démarche proposée, nécessite plusieurs étapes, et ses dernières ont été largement décrites et discutées dans ce chapitre. La première étape consacrée à la préparation des données, débute avec un point qui montre la nature physique et l'origine des signaux et se termine par la construction d'une base de données étiquetée. La seconde est réservée à l'extraction des indicateurs spectraux à partir des trames de signaux. L'étape relative la sélection et réduction de l'espace des attributs, a été volontairement omise pour des raisons évidentes de simplicité et de l'unicité du type d'indicateur extrait. Une fois le vecteur d'entrée construit, l'étape suivante est la construction des modèles et l'établissement des règles de décision par la suite. La dernière étape consiste à l'élaboration des tests de performance nécessaires au jugement à porter sur la méthode de diagnostic proposée.

Chapitre 4

***EXPERIMENTATIONS ET RESULTATS
OBTENUS***

4.1. INTRODUCTION

Afin de tester expérimentalement les algorithmes développés suivant les deux approches proposées dans cette étude, il a fallu une base de signaux vibratoires de défauts de roulements, mesurés. En d'autres termes, il est nécessaire de disposer d'un banc d'essai dédié, pour la collecte des signaux représentatifs des conditions de défauts de roulement, au sein du laboratoire. L'étude de conception a été menée, mais les moyens financiers n'étaient pas au rendez-vous. Alors, le recours à des bases de données publiques, a été l'unique voie disponible.

L'ensemble de données fourni par le Case Western Reserve University (CWRU) Bearing Data Center a été utilisé. Cette base de données est devenue une référence standard dans le domaine du diagnostic des défauts de roulement. Une centaine, voire plus, d'articles de recherche qui sont publiés sur la base de ses données. Randall a estimé que le journal bien connu « Mechanical Systems and Signal Processing » a publié 41 articles en utilisant les données CWRU entre 2004 et début 2015. Ce qui représente une source crédible de données pour valider nos algorithmes développés.

Ce chapitre consacré à l'expérimentation des démarches de diagnostic proposées, présente en premier lieu une description sommaire des moyens expérimentaux et les conditions liées aux mesures effectuées des signaux de vibrations. Une attention particulière est réservée à l'organisation des signaux enregistrés au format « .mat » de Matlab, selon le type de défauts de roulement, selon la profondeur du défaut considéré et selon la charge (vitesse) appliquée. Par la suite, une sélection a été opérée sur les fichiers de données, selon les besoins de la validation. Enfin, les tests de validation, pour toutes les configurations envisagées de paramètres de classification. Les résultats obtenus ont été largement discutés.

4.2. EXPERIMENTATION

Les tests de validation sont basés sur la collection de données des défauts de roulements obtenue auprès du « CWRU (Case Western Reserve University) bearing data center » [152]. Cette banque de données public a été largement utilisée dans de très nombreux travaux. Ainsi, Randall a recensé des centaines d'articles publiés basés sur les données de cette source, durant une courte période [6]. Elle représente donc, l'une des principales références utilisées dans la validation de nouveaux algorithmes, malgré ses quelques inconvénients mis en évidence et soulevés par des spécialistes en la matière. Cependant, il est à souligner que l'usage de ses données demeure partiel, car les modes opératoires des essais sont variables. Ainsi une collection d'enregistrements a été utilisée basée sur une seule fréquence d'échantillonnage.

4.2.1 Description des moyens expérimentaux

Le banc d'essai utilisé dans la collecte des signaux de roulement par le « bearing data center », se compose d'un moteur de 2 CV (à gauche), d'un capteur de couple / encodeur (au centre) et d'un dynamomètre (à droite), comme indiqué sur la figure 4.1. Les roulements à billes à gorge profonde sont installés dans un système mécanique entraîné par moteur. Les roulements d'essai dont les caractéristiques dimensionnelles sont montrées dans le tableau 4.1, supportent donc, l'arbre du moteur.

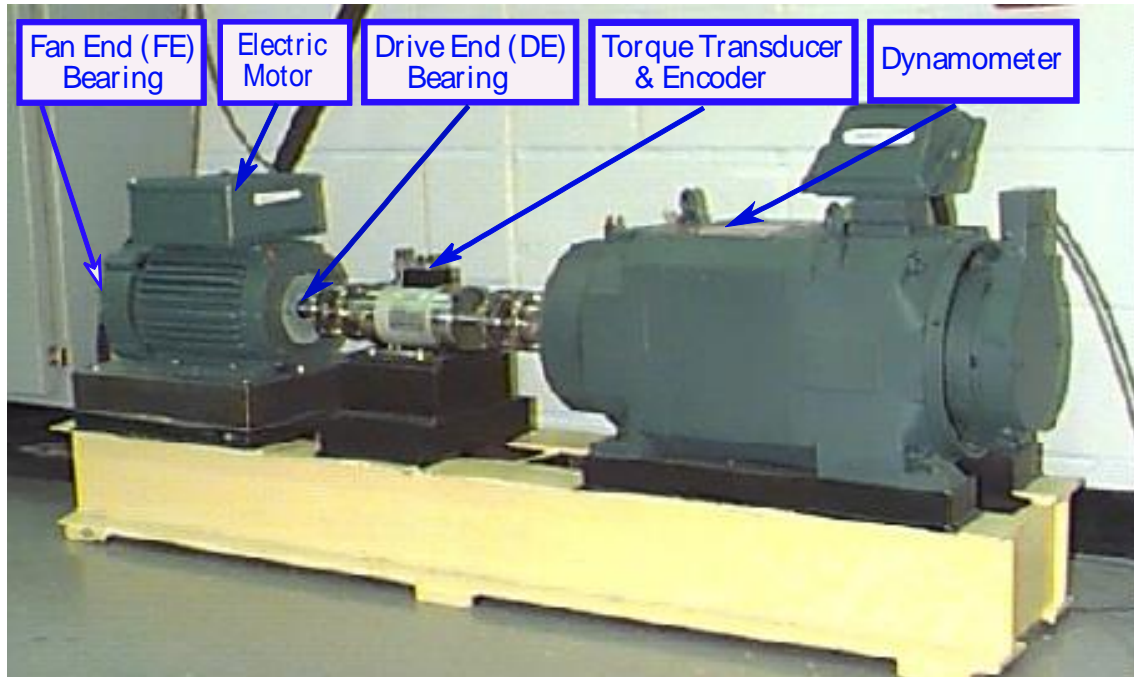


Figure 4.1 : Banc d'essai de vibration des roulements du CWRU [152]

Des défauts ponctuels ont été introduits dans les éléments de roulements d'essai au moyen de l'électroérosion par enfonçage, appelée aussi EDM (Electrical Discharge Machining). C'est un procédé d'usinage qui consiste à enlever de la matière dans une pièce en utilisant des décharges électriques. La précision de cette technique permet d'obtenir des défauts avec de très petites dimensions (diamètres de 0,007 à 0,040 pouces). Des accéléromètres sont fixés sur les supports rigides des paliers du côté entraînement (Drive End – DE), du côté ventilateur (Fan End – FE) et sur la base, permettent de mesurer toutes les forces réactives dues aux excitations vibratoires.

Les défauts ainsi créés ont été introduits séparément au niveau de la course de la bague intérieure, de l'élément roulant (la bille) et au niveau de la course de la bague extérieure des roulements d'essais. Les roulements défectueux ont été montés en alternance dans le moteur

d'essai et les données de vibration ont été enregistrées pour des charges de moteur allant de 0 à 3 chevaux (vitesses de moteur de 1797 à 1720 tr / min).

Tableau 4.1 : Caractéristiques dimensionnelles et cinématiques des roulements utilisés

Position sur le banc	Modèle de roulement	Dimensions Inches (mm)		Fréquences cinématiques (multiples de la vitesse de rotation de l'arbre en Hz)			
				BPMI	BPMO	FTF	BSR
Côté accouplement (Drive End – DE)	6205-2RS JEM SKF	ϕ_i	0.9843 (25)	5.4152	3.5848	0.3983	4.7135
		ϕ_e	2.0472 (52)				
		Epaisseur	0.5906 (15)				
		ϕ_b	0.3126 (08)				
		ϕ_m	1.537 (39)				
Côté Ventilateur (Fan End – FE)	6203-2RS JEM SKF	ϕ_i	0.6693 (17)	4.9469	3.0530	0.3817	3.9874
		ϕ_e	1.5748 (40)				
		Epaisseur	0.4724 (12)				
		ϕ_b	0.2656 (07)				
		ϕ_m	1.122 (28)				

4.2.2 Base de signaux de vibration sélectionnés

Les signaux de vibration sont enregistrés avec des noms de fichiers au format Matlab (xxx.mat). Le «xxx» signifie un nom de fichier sélectionné composé de trois chiffres, comme indiqué dans le tableau 4.2. Chaque fichier contient deux signaux enregistrés DE et FE (Drive End et Fan End). Le premier est mesuré au palier à roulement du côté entraînement et le deuxième concerne celui du palier à roulement du côté ventilateur.

Tableau 4.2 : Nomenclature de fichiers des défauts sélectionnés (roulement DE) à 48KHz.

Charge moteur (HP)	Vitesse moteur (tr/min)	Largeur défaut 0.007 in (0.18 mm)			Etat normal (sain)
		BI	Bille	BE centrée	
0	1797	109	122	135	97
1	1772	110	123	136	98
2	1750	111	124	137	99
3	1730	112	125	138	100

Les signaux sélectionnés pour notre classification concernent à la fois des roulements sains et défectueux au niveau de la bague intérieure, de la bague extérieure (en position centrée) et de la bille. La fréquence d'échantillonnage pour l'ensemble des signaux de vibration sélectionnés est de 48 kHz.

4.2.3 Mise en œuvre de la procédure de classification

4.2.3.1 Création de base de données pour la classification

A partir de la topologie des données sélectionnées dans la sous-section précédente, quatre classes ont été définies pour représenter trois cas de défauts respectivement BPFI (Ball Pass Frequency, Inner race), BPFO (Ball Pass Frequency, Outer race), BSF (Ball Spin Frequency) et une quatrième classe représentative de l'état sain (aussi appelé normal lors de cette étude). En conséquence, la base de données étiquetée sera créée autour de ces quatre classes prédéfinies.

Globalement, l'ensemble de données sélectionnées, couvre 16 scénarios couvrant les quatre conditions de défaut considérées (04 classes). Chacune des conditions représente quatre cas de chargement (quatre vitesses différentes). De plus, avec la prise en considération des deux signaux relatifs aux points de mesure (côté entraînement - DE - et côté ventilateur – FE - de chaque scénario), nous obtenons un total de 32 signaux à traiter, dans notre schéma de classification.

Tableau 4.3 : Partition des données entre l'apprentissage et le test selon la taille de la trame

Longueur de trame	Nombre de trames d'apprentissage	Nombre de trames de test	Total
64	28672	21504	50176
128	14336	10752	25088
256	7168	5376	12544
512	3584	2688	6272
1024	1792	1344	3136
2048	896	672	1568
4096	448	336	784
8192	224	168	392
16384	112	84	196

La création de la base de données étiquetée a commencé par remodeler tous les signaux de vibration de l'ensemble de données car leurs tailles dans leurs formes originales sont différentes. Ainsi, cette première opération a généré 14 blocs de tailles uniformes de 229376 échantillons, pour chaque classe. Chaque bloc est divisé en trames d'analyse (ou segments) de différentes longueurs pour tester celles qui donnent le taux de classification le plus élevé. En conséquence, neuf longueurs de trame comme indiqué dans le tableau 4.3, ont été explorées.

Afin de créer un sous-ensemble d'apprentissage et un autre de test à partir de l'ensemble des données, les huit premiers blocs ont été consacrés à l'apprentissage et à la modélisation ($\approx 58\%$), tandis que le reste ($\approx 42\%$) a été conservé pour tester la méthode proposée.

Avant de réaliser l'extraction des indicateurs à partir de l'ensemble des trames de la base de données, il convient de leur faire subir des prétraitements nécessaires. Dans une première étape l'ensemble des signaux sont centrés et normalisés pour réaliser une classification correcte. Dans une deuxième étape, un filtrage du bruit à l'aide d'un filtre de prédiction linéaire ayant un ordre élevé. En dernier, les signaux traités, étiquetés selon les quatre classes prédéfinies sont affectés aux partitions d'entraînement et de test selon les proportions fixées ci-dessus. Des prototypes des signaux ayant subits des prétraitements et représentatifs de chacune des quatre classes, sont illustrés dans la figure 4.2.

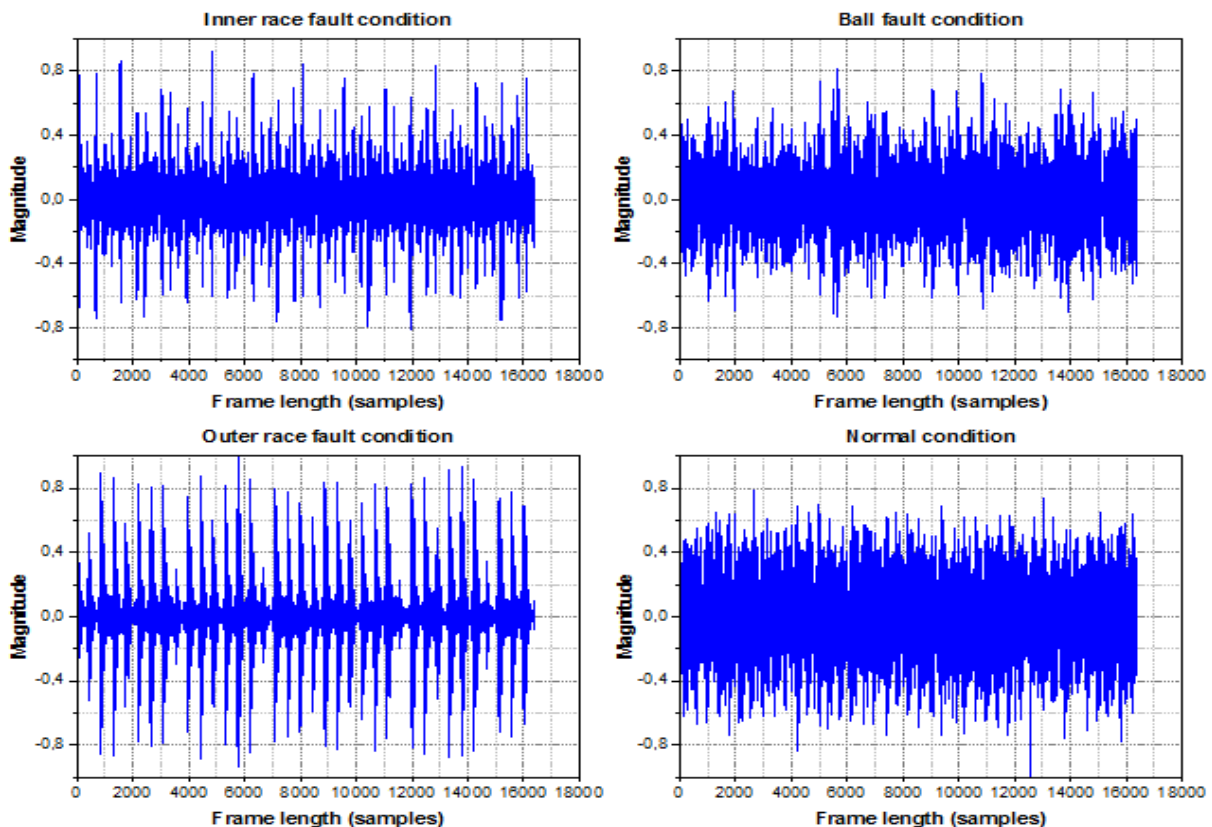


Figure 4.2 : Signaux relatifs aux quatre classes considérées après prétraitements.

4.2.3.2 Extraction des attributs LPCs

Après les prétraitements appliqués aux données la seconde étape de la classification consiste à l'extraction des indicateurs LPC, à la fois aux données d'apprentissage et de test. Ce processus calcule à sa fin le vecteur LPC pour chaque trame. Les courbes des LPCs extraits des trames représentatives (modèles) des quatre conditions de défauts sont montrées dans la figure 4.3.

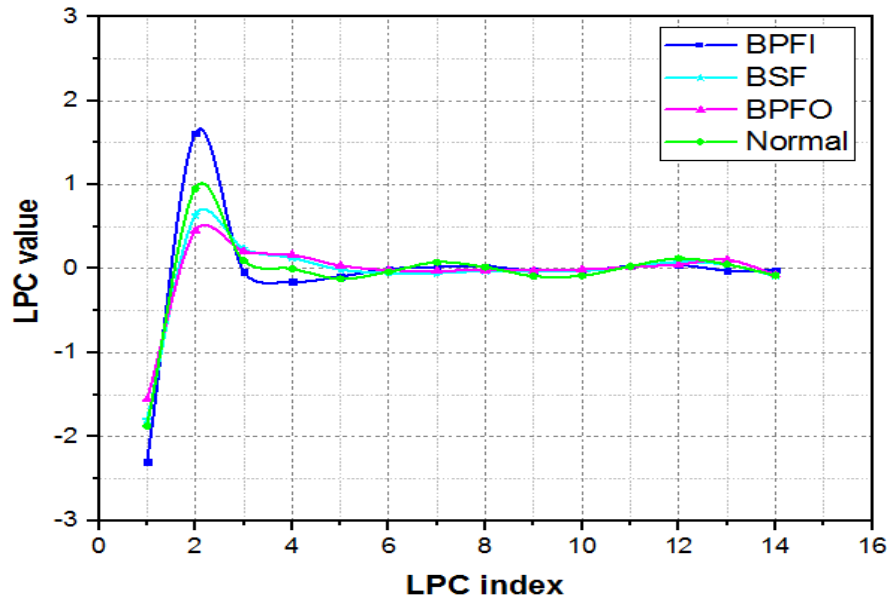


Figure 4.3 : Valeurs des LPC pour quatre trames représentatives des conditions de défaut

Les caractéristiques extraites pour les trames d'apprentissage de toutes les conditions de défaut (04 classes) sont illustrées à la Figure 4.4. Sur l'axe X est représenté l'indice de chaque coefficient parmi les quatorze (14 LPC) extraits de chaque trame. L'axe des Y montre le nombre de trames d'apprentissage (Training Frames Number - TFN) utilisées. Dans ce cas, avec une longueur de trame de 1024 échantillons, le TFN max est de 1792 trames. L'axe Z montre les valeurs des LPC. Les longueurs des fenêtres choisies (durées de temps) étant fonction de la longueur des trames d'analyse et de la fréquence d'échantillonnage. Le résultat conduit à des durées variables. L'extraction des fonctionnalités se fait avec un nombre de LPC variable. Cette solution nous permettra de tester différentes combinaisons en termes de nombre de LPC, de longueur de trame d'analyse et de test, et le nombre de composants dans le mélange. Ainsi on se fixera sur le (ou les) paramètre(s) à choisir qui donne(nt) le taux de classification moyen le plus élevé (le meilleur).

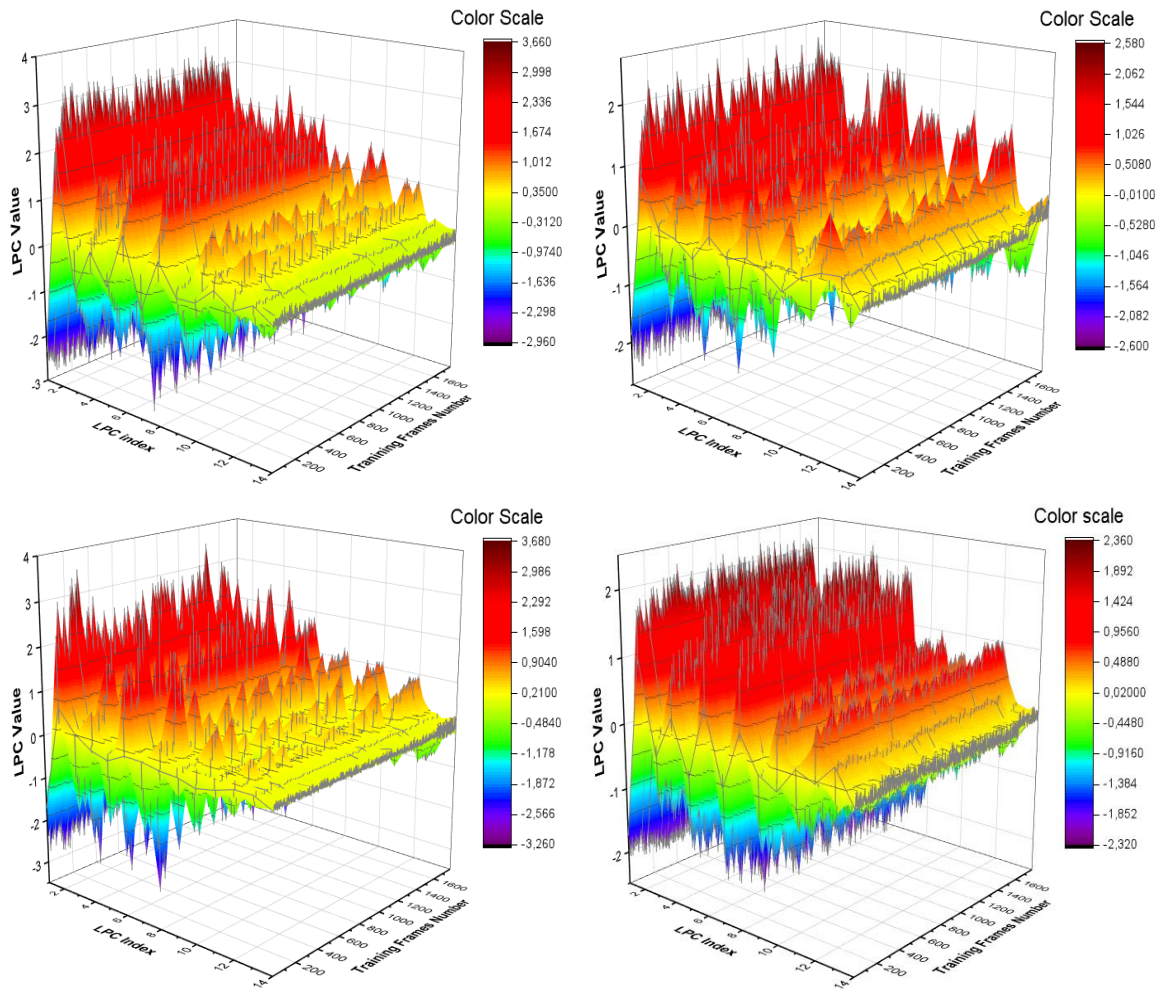


Figure 4.4 : Représentation graphique de la base d'apprentissage des LPC (chaque condition de défaut correspondant à 1792 trames de 1024 échantillons)

4.2.3.3 Extraction des attributs MFCCs

Contrairement à la méthode précédente (basée sur les LPCs), la classification avec les MFCCs n'est possible que pour les trames ayant des longueurs de 512 échantillons et plus. Les indicateurs extraits des trames représentatives (prototypes) des quatre conditions de défauts sont montrées dans la figure 4.5.

Les longueurs des fenêtres choisies (durées de temps) étant fonction de la taille des trames d'analyse et de la fréquence d'échantillonnage, le résultat conduit à des durées variables. Cette disposition permet d'explorer différentes situations pour chercher les configurations qui donnent les classifications les plus performantes.

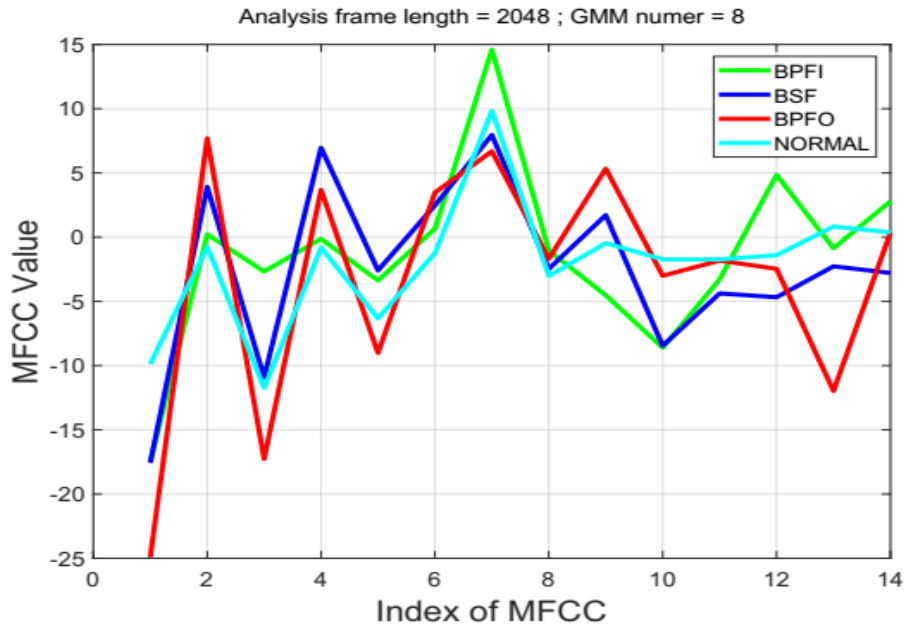


Figure 4.5 : Valeurs de MFCC correspondant aux différentes conditions de défauts

Les caractéristiques extraites pour les trames d'apprentissage de toutes les conditions de défaut (04 classes) sont illustrées à la Figure 4.6. Sur l'axe X est représenté l'indice de chaque coefficient parmi les quatorze (14 MFCC) extraits de chaque trame. L'axe des Y montre le nombre de trames d'apprentissage (Training Frames Number - TFN) utilisées. Dans ce cas, avec une longueur de trame de 2048 échantillons, le TFN max est de 896 trames. L'axe Z montre les valeurs des MFCC.

Les longueurs des fenêtres choisies (durées de temps) étant fonction de la longueur des trames d'analyse et de la fréquence d'échantillonnage. Le résultat conduit à des durées variables. L'extraction des fonctionnalités se fait avec un nombre de MFCC variable. De cette manière, il est possible de tester différentes combinaisons en termes de nombre de MFCC, de longueur de trame d'analyse et de test, et le nombre de composantes dans le mélange. Ainsi on se fixera sur le (ou les) paramètre(s) à choisir qui donne(nt) le taux de classification moyen le plus élevé (le meilleur).

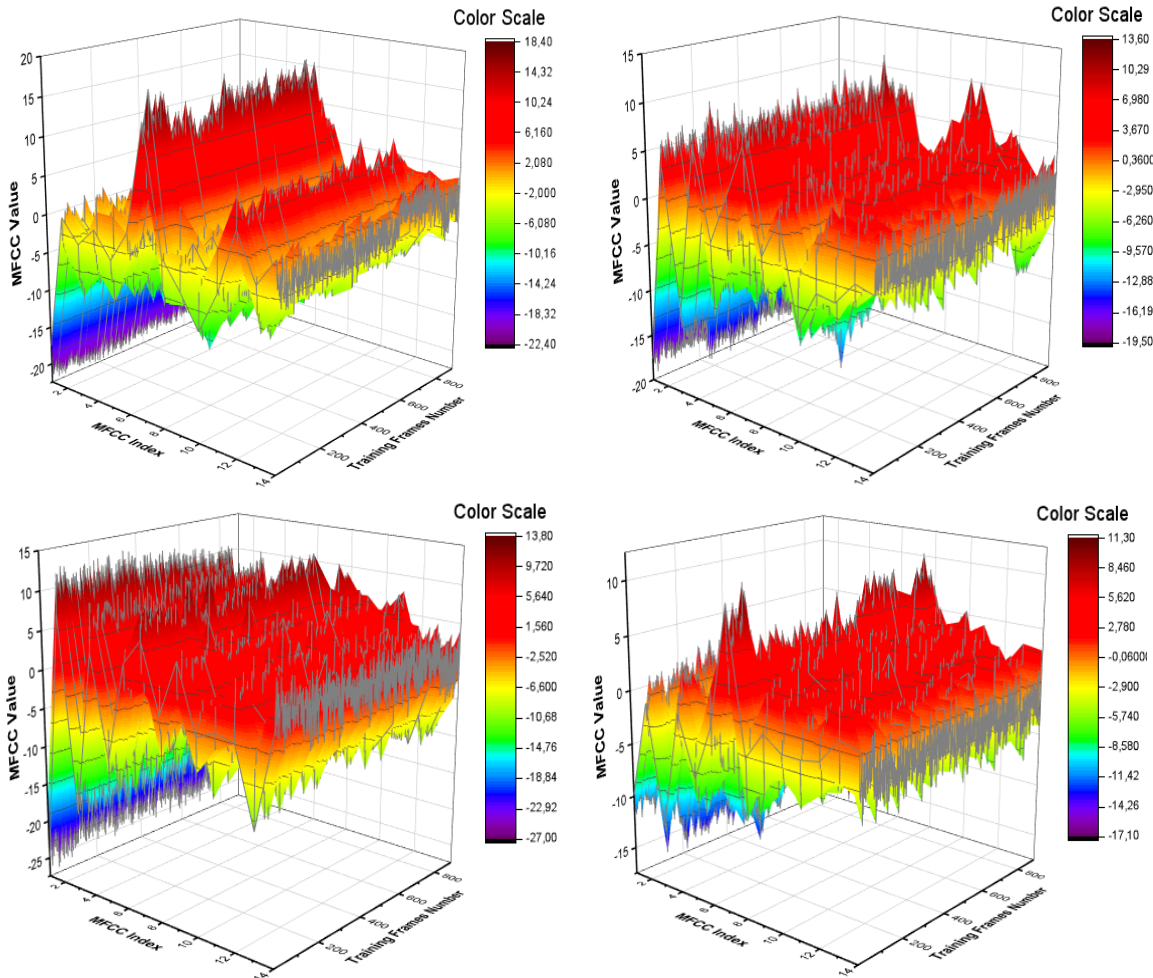


Figure 4.6 : Représentation graphique de la base d'apprentissage des MFCC (chaque condition de défaut correspondant à 896 trames de 2048 échantillons)

4.3 RESULTATS DE LA CLASSIFICATION BASEE SUR LES LPCs

L'évaluation de la performance de classification est effectuée par une estimation du taux de classification moyen ACR (Average Classification Rate), obtenu à partir de la matrice de confusion. L'ensemble des procédures de classification, vise à trouver une valeur optimale pour chacun des paramètres d'entrée du classificateur comme le nombre de composants dans le mélange de gaussiennes, le nombre de LPC ou de MFCC, la longueur de la trame d'analyse et la longueur de la séquence de test. Par conséquent, une gamme de valeurs possibles pour chacun des paramètres de la classification a été testée.

Les résultats obtenus et les influences sur le processus de classification sont présentés dans cette section. Ainsi, à la fin des deux méthodes de classification, nous aurions atteint les deux configurations souhaitées de paramètres optimaux qui garantissent une classification efficace (ACR élevé).

4.3.1 Erreur estimée et ordre de prédiction

4.3.1.1 Erreur estimée du prédicteur linéaire

Le premier résultat à afficher concerne l'erreur estimée du prédicteur linéaire. Une série de données utilisant un prédicteur direct d'ordre élevé (quatre-vingtième), est estimée et comparée au signal d'origine. Il s'en suit dans la première représentation montrée sur la figure 4.7, que les courbes représentatives des signaux originaux et estimés sont pratiquement superposées et confondues. Cela suggère que l'erreur d'estimation est significativement réduite. Dans les deuxième et troisième représentations de la même figure, sont montrées les courbes de l'évolution de l'erreur de prédiction et de l'autocorrélation.

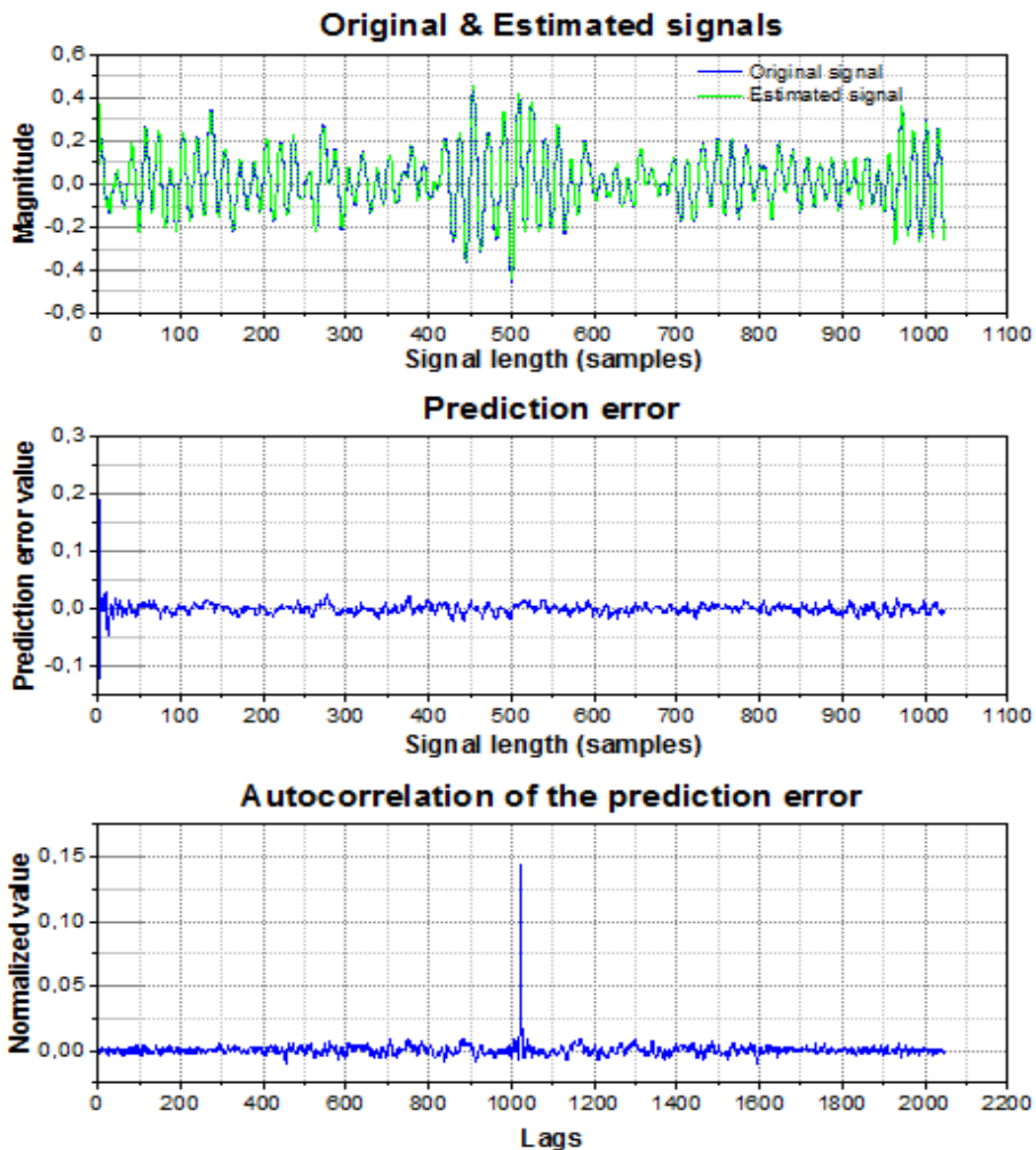


Figure 4.7 : Comparaison des signaux original et estimé du modèle de prédiction

4.3.1.2 Ordre de prédiction théorique optimal

La première contribution de ce travail consiste à déterminer l'ordre de prédiction théorique optimal à utiliser dans l'extraction des indicateurs (caractéristiques). Pour ce faire, nous avons dû tracer la courbe de l'erreur normalisée en fonction de l'ordre de prédiction, obtenue à partir de l'équation (3.79). Ainsi, en utilisant la même équation, quatre courbes représentatives de chacun des quatre conditions de défauts, ont été déduites.

On observe sur la figure 4.8, que les quatre courbes ont des allures similaires. Dans chacune, l'erreur normalisée diminue fortement pour un ordre de prédiction allant jusqu'à 4. Par la suite la décroissance devient progressive pour des valeurs de l'ordre de prédiction comprises entre 4 et 14. Enfin et pour celles qui sont supérieures à cette dernière valeur, les quatre courbes semblent linéaires et la décroissance de l'erreur normalisée devient négligeable. Ces résultats sont affichés partiellement dans le tableau 4.4.

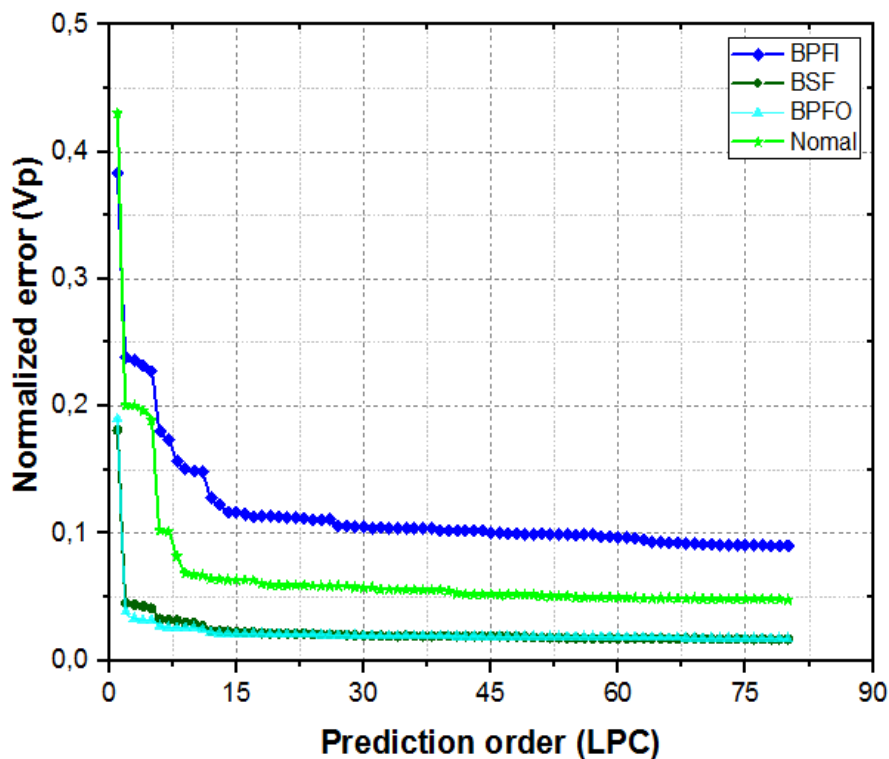


Figure 4.8 : Courbes d'erreur normalisée relatives aux défauts de roulement

Les valeurs d'erreur normalisées sont calculées pour des ordres de prédiction variant de 1 à 80. Le constat fait à partir des allures de ces quatre courbes, est qu'au-delà de l'ordre $P = 14$, la réduction de l'erreur normalisée n'est pas significative et devient donc pénalisante en

termes de temps de calcul. En conséquence, nous pouvons conclure que $P = 14$ est l'ordre optimal recherché pour cette analyse LP.

Tableau 4.4 : Influence de l'ordre de prédiction sur l'erreur normalisée

Ordre P (BPFI)	1	10	12	14	16	20	50	80
Erreur normalisée	0.38	0.15	0.13	0.12	0.11	0.11	0.10	0.09
Réduction %	100	38.93	33.45	30.47	30.02	29.65	25.97	23.62

4.3.2 Choix du nombre de LPC

Dans cette partie, les résultats obtenus confirment bien ceux de la théorie obtenus dans la sous-section ci-dessous. La figure 4.9 montre la courbe du taux de classification moyen (ACR) en fonction du nombre de coefficients LPC. Il varie entre une valeur minimale de 75% correspondant à un nombre LPC de 2, et une valeur maximale de 100% pour un nombre LPC de 14. Dans la plage allant de 14 à 20 coefficients, il demeure constant et équivaut à 100% pour une trame de 1024 points. Au-delà de cette valeur, et malgré ses légères fluctuations, il reste dans une gamme supérieure à 85% jusqu'à l'ordre le plus élevé du test (80).

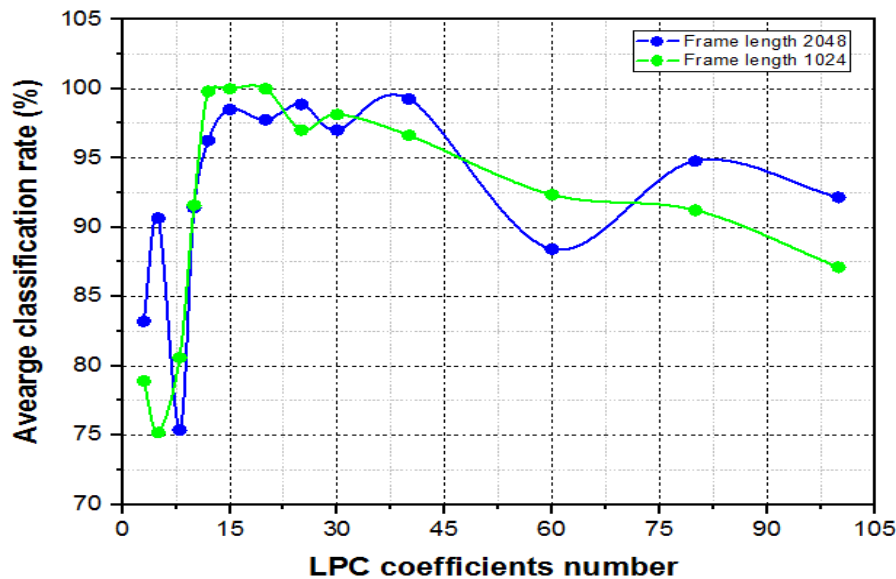


Figure 4.9 : Variation du taux de l'ACR en fonction du nombre de LPCs

Il est difficile d'expliquer précisément les causes de ces fluctuations. En ce qui concerne le temps de calcul, on peut avancer une première conclusion que le nombre LPC optimal est 14.

4.3.3 Effet de la longueur de la trame d'analyse

Ayant fixé au préalable le nombre de LPC à sa valeur optimale de 14, obtenue dans la sous-section ci-dessus, l'effort est maintenant concentré sur l'effet de la longueur des trames d'analyse sur la classification. Ainsi, la courbe de variation du taux de classification moyen (ACR) en fonction de la longueur de la trame d'analyse est illustrée à la figure 4.10.

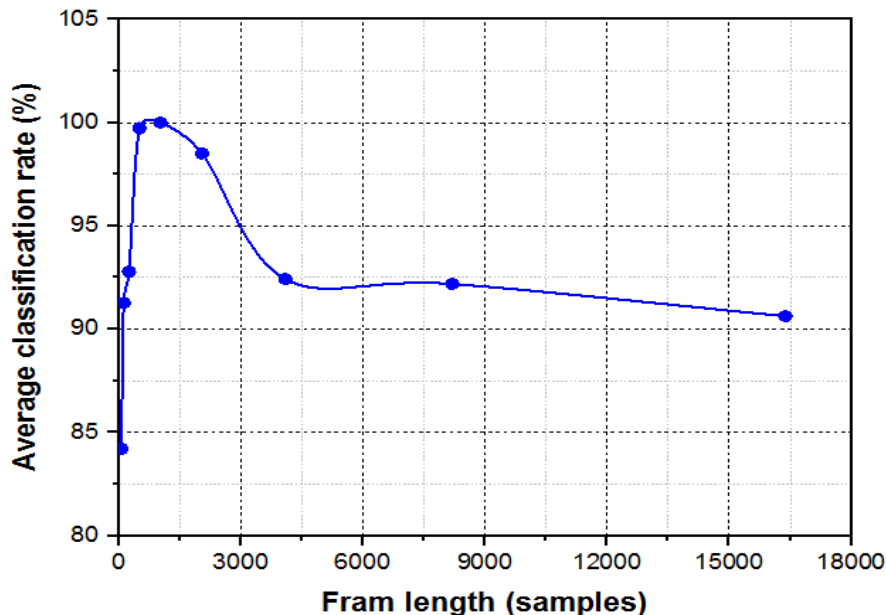


Figure 4.10 : Variation du taux de l'ACR en fonction de la longueur de la trame d'analyse

On peut aisément voir que la meilleure précision est obtenue pour une longueur de trame de 1024 points. En prenant en considération la valeur de la fréquence d'échantillonnage des signaux de 48 KHz, cette longueur de trame correspond également à une durée de 21,33 ms. De plus, on peut noter que pour les longueurs supérieures à 512 échantillons (allant jusqu'à 16384 échantillons), l'ACR est supérieur à 90%. Cela signifie que pour toute cette gamme de longueurs de trame sous étude, les LPC extraits reproduisent fidèlement les informations contenues dans ces trames. Par conséquent, la discrimination entre les classes devient facile. À 1024 échantillons, les résultats montrent que la classification est encore nettement meilleure. Ainsi, la longueur de 1024 échantillons est prise comme valeur de trame d'analyse optimale.

4.3.4 Effet de la longueur de la séquence de test

Comme prévu, l'augmentation de la longueur de la séquence de test (trame de test) améliore le taux de reconnaissance global. L'allure de la courbe de la figure 4.11, montre bien que le taux de classification moyen augmente avec le nombre de trames d'analyse incluses dans la séquence de test.

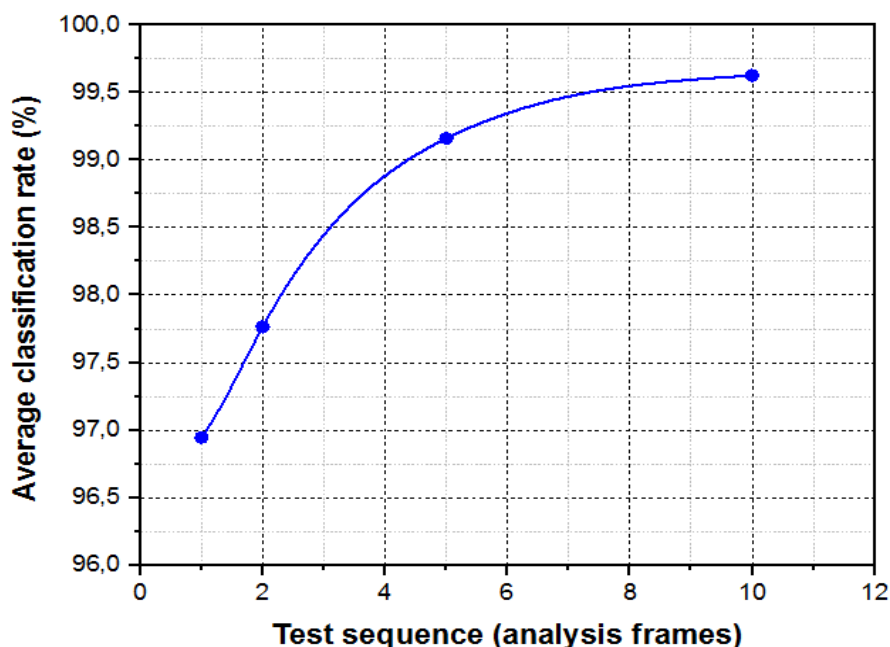


Figure 4.11 : Variation du taux de l’ACR en fonction de la longueur de la séquence de test

Il est important de signaler que cette variation du résultat de classification comprise dans cette plage de test, est de l’ordre de 2,5 %. Ceci signifie que le taux de classification reste élevé même avec une seule trame d’analyse dans une séquence de test. Néanmoins, on peut conclure que l’estimateur du prédicteur linéaire est meilleur avec des séquences de test ayant 10 trames d’analyse qu’avec celles ayant un nombre plus petit.

Tableau 4.5 : Matrices de confusion : NLPC=14 ; NGMM=2 ; Trame d’analyse=1024 pts.

Séquence de test = 1 trame d’analyse				Séquence de test = 2 trames d’analyse					
ACR = 96.9449 %				ACR = 97.7645 %					
	BPFI	BSF	BPFO	Normal		BPFI	BSF	BPFO	Normal
BPFI	0.9694	0.0306	0	0	BPFI	0.9866	0.0134	0	0
BSF	0.0633	0.9106	0.0261	0	BSF	0.0551	0.9240	0.0209	0
BPFO	0	0.0022	0.9978	0	BPFO	0	0	1.0000	0
Normal	0	0	0	1.0000	Normal	0	0	0	1.0000
Séquence de test = 5 trames d’analyse				Séquence de test = 10 trames d’analyse					
ACR = 99.1604 %				ACR = 99.6269 %					
	BPFI	BSF	BPFO	Normal		BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0	BPFI	1.0000	0	0	0
BSF	0.0261	0.9664	0.0075	0	BSF	0.0075	0.9851	0.0075	0
BPFO	0	0	1.0000	0	BPFO	0	0	1.0000	0
Normal	0	0	0	1.0000	Normal	0	0	0	1.0000

Les matrices de confusion présentées dans le tableau 4.5, montrent plus de détails dans les résultats de la classification. Il est clair qu'avec 10 trames d'analyse dans une séquence de test, on a obtenu le taux de classification moyen le plus élevé sur la courbe (99.6269). Avec 5 trames d'analyse dans une trame de test, le score est approximativement le même (99.1604), soit à 0.5 % moins élevé que le précédent. Cependant, en termes de temps de calcul, il est bien meilleur car la longueur de la séquence à 10 trames est deux fois plus longue que celle à 5 trames. Par conséquent, il est plus judicieux de prendre 5 trames d'analyse dans une séquence de test, comme valeur optimale.

4.3.5 Effet du nombre de composantes du GMM

Dans la littérature, l'influence du nombre de composantes du mélange sur le processus de classification par GMM, est un sujet qui occupe une place importante. Par conséquent, l'estimation du nombre optimal de paramètres est essentielle car les paramètres avec des composantes inappropriées peuvent ne pas évaluer le modèle de mélange avec précision. C'est une tentative difficile de décider du nombre optimal de composantes. Pour estimer l'ordre du modèle et les composants ensemble, certaines techniques et certains critères ont été développés (critère d'information d'Akaike (AIC), longueur minimale de description (MDL), critère d'inférence bayésien (BIC), entre autres) [132, 153, 154]. Cependant, dans ce travail, cet aspect n'a pas été développé, car il représente en soi une étude de dimension entière.

La contribution de ce travail vise à estimer, l'influence du nombre de composantes (GMM) sur l'évolution du taux de classification moyen. C'est pourquoi différentes valeurs du nombre de composantes du mélange ont été testées afin chercher celles qui maximisent notre critère d'évaluation (ACR). Ainsi, les tests ont été réalisés sur des nombres de composantes du mélange allant de 2 à 12 (avec un pas de 2). Pour mieux cerner cette variation, elle a été testée pour trois valeurs du nombre de LPCs. Les résultats sont représentés sur la figure 4.12.

À partir des allures des trois courbes, il convient de constater que le nombre de LPC a une grande influence sur les courbes du taux de classification moyen par rapport au nombre de composantes du mélange. L'analyse des résultats présentés, permet d'observer deux situations remarquables.

La première constatation est que les courbes représentées en trois dimensions, ont des allures décroissantes dans la direction du nombre de composantes de GMM, et croissantes dans la direction du nombre de LPCs. Pour les nombres de LPCs correspondant à 8 et 10, les courbes représentatives de l'ACR ont des valeurs maximales pour le nombre de GMM égal à 2, et des valeurs minimales pour le nombre élevé de GMM égal à 12.

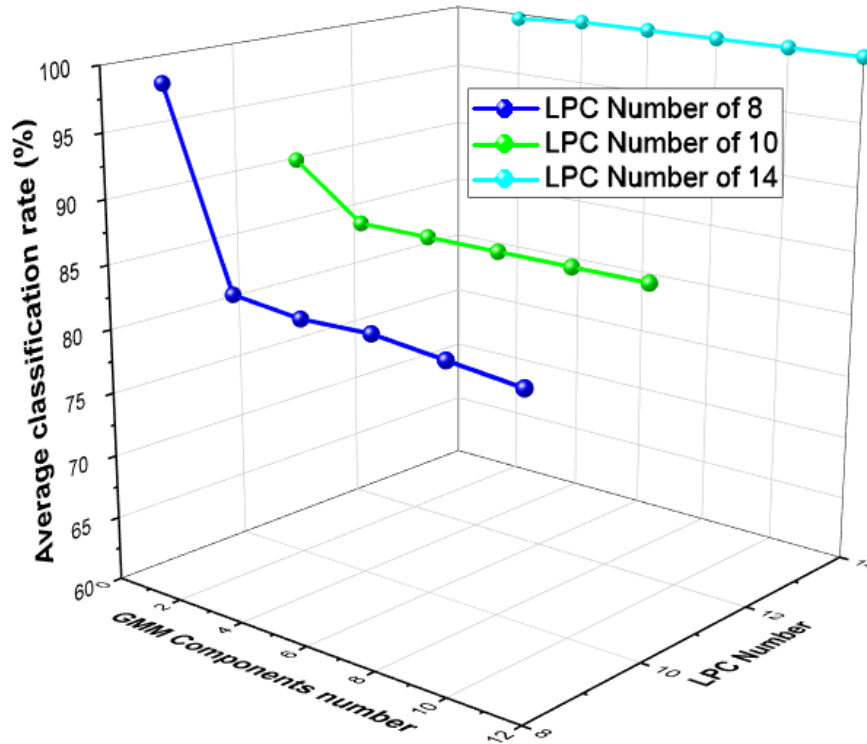


Figure 4.12 : Variation du taux de l'ACR en fonction du nombre de composantes du mélange

La seconde situation est liée au nombre de LPCs correspondant à 14 (sa valeur optimale estimée). Pour ce cas, la courbe de l'ACR est constante et ses valeurs sont toutes maximales (100%), indépendamment du nombre de composantes du mélange. Il semble donc que l'erreur d'ordre du classificateur soit négligeable par rapport à l'erreur de l'estimateur de prédiction linéaire. Une valeur optimale dans l'ordre de ce dernier élève le taux de classement moyen à sa valeur maximale, quel que soit l'ordre du classificateur.

Ces résultats obtenus mettent en valeur l'effet du bruit. Plus de composantes GMM sont ajoutées, plus le bruit dans les mesures est modélisé. Ainsi, en augmentant l'ordre GMM, les performances diminuent. Ce cas visible est pour les cas où le nombre de LPCs est inférieur à 14. En conséquence, le nombre de composantes GMM correspondant à 2 devient la dernière valeur optimale dans ce processus de choix des paramètres de la classification.

4.4 RESULTATS DE LA CLASSIFICATION BASEE SUR LES MFCCs

Le long de ce processus, une meilleure classification signifie un taux de classification moyen maximum (ACR). Par conséquent, une valeur de 100% est considérée comme idéale. Il est obtenu à partir de la moyenne des termes diagonaux de la matrice de confusion, à la fin de chaque test de classification. Il s'agit du critère de référence dont la valeur juge l'optimalité des autres paramètres impliqués dans le processus de classification comme le nombre de

composantes dans le mélange gaussien, le nombre de MFCC, la longueur de la trame d'analyse et la longueur de la séquence de test.

L'objectif donc, est de trouver une combinaison des paramètres optimaux permettant une meilleure classification. Pour ce faire, différentes valeurs de chacun des paramètres de classification ont été testées. Les résultats obtenus et les influences de chaque paramètre sur le processus de classification sont présentés ci-dessous [155].

4.4.1 Effet du nombre de composantes du GMM

D'une manière similaire à la procédure employée dans la sous-section précédente, l'objectif visé, est de trouver le nombre de composantes du mélange, qui maximisent le taux de classification moyen (ACR). Par conséquent, nous essayons d'augmenter le nombre de composantes du mélange de 2 à 12 (avec un pas de 2). Deux cas ont été étudiés. Le premier illustré sur la figure 4.13, concerne une trame de 2048 échantillons. Les nombres de MFCC testés sont 10, 12, 13 et 14. A ce stade, nous ne connaissons pas encore le nombre optimal nécessaire. Nous avons utilisé les nombres les plus utilisés dans la littérature. Au final, le nombre de huit (8) composantes a été obtenu et satisfait au taux ACR le plus élevé, quel que soit le nombre de MFCC. Le deuxième cas illustré sur la figure 4.14, a pris en charge le choix du premier, car la trame de longueur de 1024 échantillons rend le nombre de composantes dépendant du nombre de MFCC. Il est également souligné que les taux de classification dans ce cas sont relativement inférieurs aux premiers.

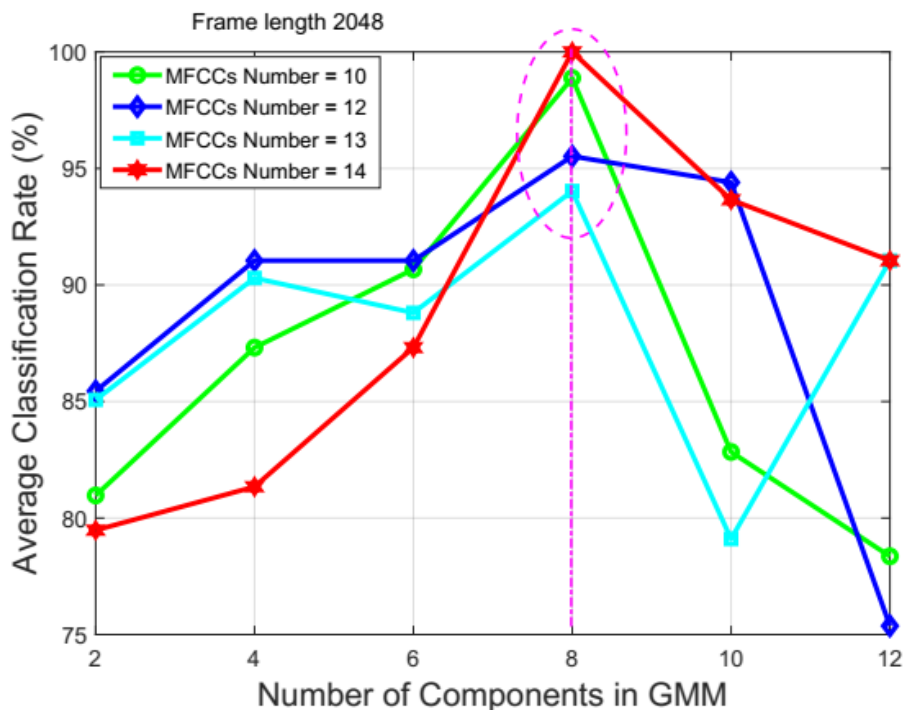


Figure 4.13 : Evolution de l'ACR en fonction du nombre de composantes GMM (2048 pts)

D'autre part, il convient de noter que le choix des trames de longueur de 2048 et 1024 échantillons est fortuit à ce stade de l'étude. Elle est principalement justifiée par leurs positions moyennes dans une gamme de trames sous étude comprises entre 512 et 16384 échantillons.

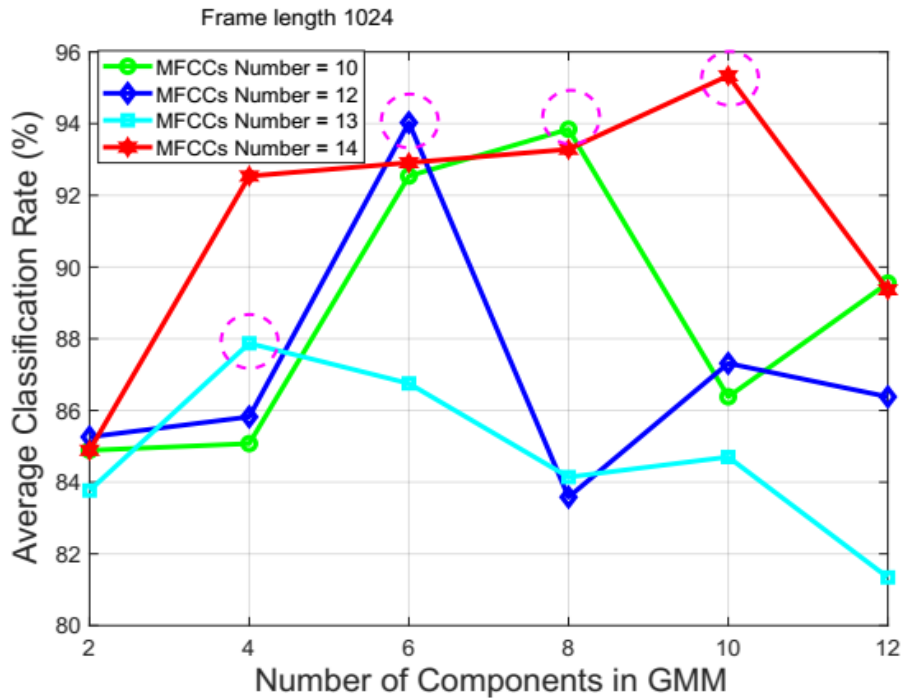


Figure 4.14 : Evolution de l'ACR en fonction du nombre de composantes GMM (1024 pts)

En plus, l'usage de ces deux longueurs dans la littérature est très répandu, sans qu'il soit expérimentalement justifié. En conséquence et pour tous ces arguments et considérations expérimentales, le nombre fixé de composantes à considérer dans toutes les exécutions de l'algorithme EM par le classificateur GMM est de huit (8). Cela signifie qu'avec cette valeur, le processus EM donne la meilleure estimation des paramètres du modèle, et donc la meilleure précision de classification des conditions de défaut de roulement.

4.4.2 Effet du nombre de MFCC

Le choix du nombre de MFCC à inclure dans le processus de classification est largement empirique [28]. Par conséquent, la méthode a été testée avec deux combinaisons de paramètres. Le processus d'extraction des indicateurs a utilisé des longueurs de trame de 1024 échantillons et 2048 échantillons dont les durées sont respectivement égales à 21,3 ms et 42,7 ms avec un chevauchement de 50%. Le nombre de MFCC testés était compris entre 2 et 20. En essayant d'améliorer les performances de classification, nous étudions l'effet de la modification du nombre de MFCC extraits de chaque trame et comparons les résultats de la classification comme indiqué sur la figure 4.15. Il est rapidement devenu évident que les nombres de

coefficients cepstraux très élevés n'étaient pas utiles pour la classification, pas plus que les très faibles. Pour plus de 14 coefficients, les performances se dégradent rapidement. En effet, trop de coefficients tuent la classification.

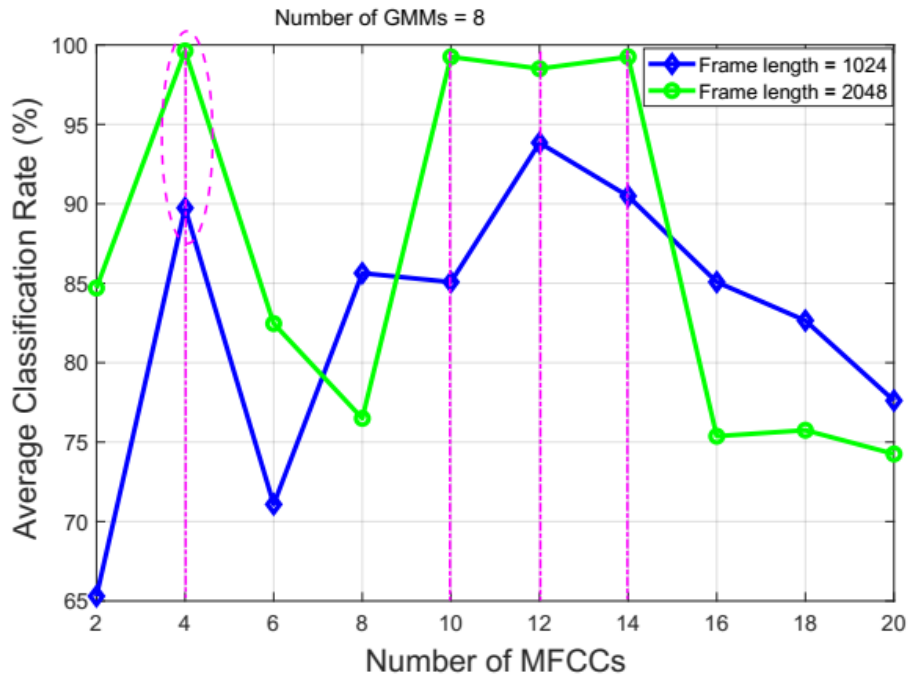


Figure 4.15 : Evolution de l'ACR en fonction du nombre de MFCCs ($N_{GMMs} = 8$)

Deux domaines sont particulièrement intéressants. Le premier est le nombre de MFCC égal à 4 avec l'ACR le plus élevé des courbes. Le second est une plage de numéros MFCC entre 10 et 14. Pour ce dernier, l'ACR est également élevé. Cela correspond bien aux résultats de la littérature. Le nombre de 12 ou 13 sont les plus cités et les plus utilisés.

En revanche, la trame de 2048 échantillons semble une fois de plus avoir un taux de classement supérieur à celui de 1024 échantillons. En revanche, ce dernier a un ACR élevé avec un nombre de MFCC de 4 dans la première zone, mais relativement fluctuant dans la gamme des MFCC entre 10-14.

Il s'ensuit que le choix du nombre de MFCC sera de quatre (4), pour des raisons évidentes d'économie en de temps de calcul. Reste à savoir si la trame de 2048 restera toujours la meilleure ou une autre pourra éventuellement la remplacer. La réponse à cette question sera discutée dans la sous-section suivante, qui discutera de l'influence de la longueur de la trame d'analyse sur le taux de classification. Dans tous les cas, le nombre de MFCC est désormais fixé à la valeur de quatre (04) et sa matrice de confusion indiquée dans le tableau 4.6.

Tableau 4.6 : Matrice de confusion dans le cas : $N_{MFCC} = 4$;

séquence de test (Trame de test) = 10 trames d'analyse				
$N_{GMM} = 8$; longueur de trame = 2048 échantillons.				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0
Normal	0	0.0299	0	0.9701

De toute façon, le nombre optimal de coefficients dépend de la quantité de données d'apprentissage, des détails de l'algorithme d'apprentissage (en particulier de la façon dont les PDF peuvent être modélisés à mesure que la dimensionnalité de l'espace des fonctionnalités augmente), du mélange gaussien, du bruit de fond caractéristiques, et peut-être des ressources informatiques disponibles. Si nous évoquons les limites de MFCC, l'une des techniques sérieuses d'extraction des fonctionnalités MFCC est que la bande passante du filtre n'est pas un paramètre de conception indépendant mais est plutôt déterminée par la plage de fréquences de la banque de filtres, le nombre de filtres utilisés et le type de fenêtre [156].

4.4.3 Effet de la longueur de la trame d'analyse

Dans cette procédure de recherche de paramètres optimaux nécessaires à la plus efficace classification possible, l'objectif initial était de segmenter les signaux de vibration en trames ayant les longueurs : 64, 128, 256, 512, 1024, 2048, 4096, 8192 et 16384 points. Le constat est que pour les trames inférieures à 512 (256, 128 et 64), le processus ne converge pas avec les paramètres initialement définis pour la classification automatique des défauts. En conséquence, ces trois trames ont été supprimées car ils nécessitent des précautions particulières. Les trames étudiées varient donc entre 512 et 16384 échantillons. Chaque trame est ensuite multipliée par une fenêtre de Hamming de même longueur.

Comme le montre la figure 4.16, l'ACR le plus élevé est obtenu pour la longueur de trame avec 2048 échantillons. Si nous considérons également le nombre de MFCC de quatre (4) comme étant le meilleur, que le reste indiqué sur les courbes. Ce résultat confirme celui obtenu précédemment dans la sous-section ci-dessus. En conséquence à tous ces résultats graphiques, la valeur optimale de la longueur de la trame d'analyse est prise 2048 échantillons.

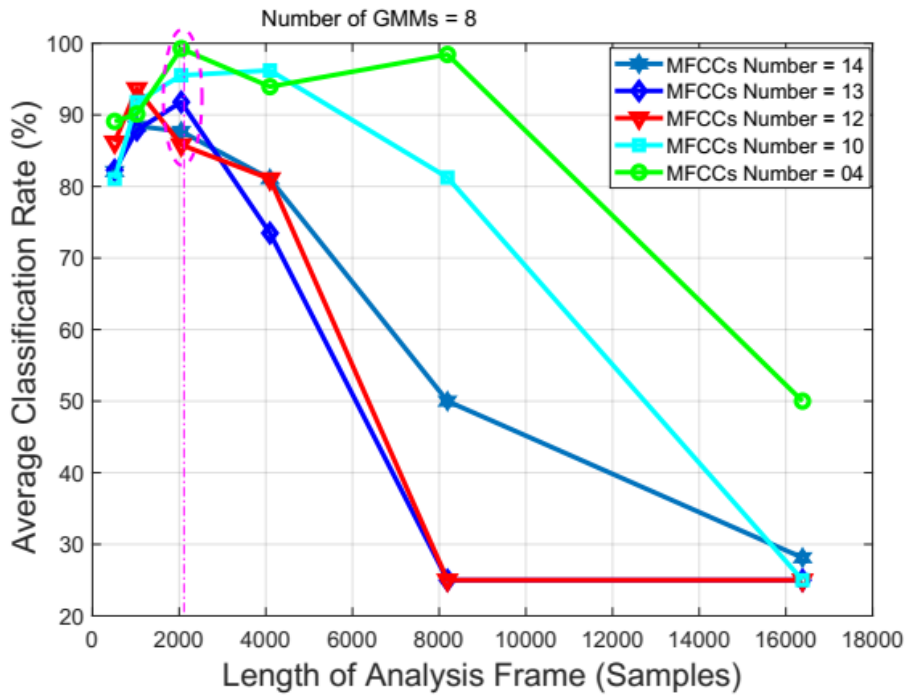


Figure 4.16 : Variation du taux ACR en fonction de la longueur de la trame d'analyse

4.4.4 Effet de la longueur de la séquence de test

Le schéma de principe illustrant l'influence sur la classification du nombre de trames d'analyse, présentes dans chaque séquence de test, fait l'objet de cette exploration. Dans tous les tests effectués auparavant, la séquence de test a toujours été fixée à dix (10) trames d'analyse.

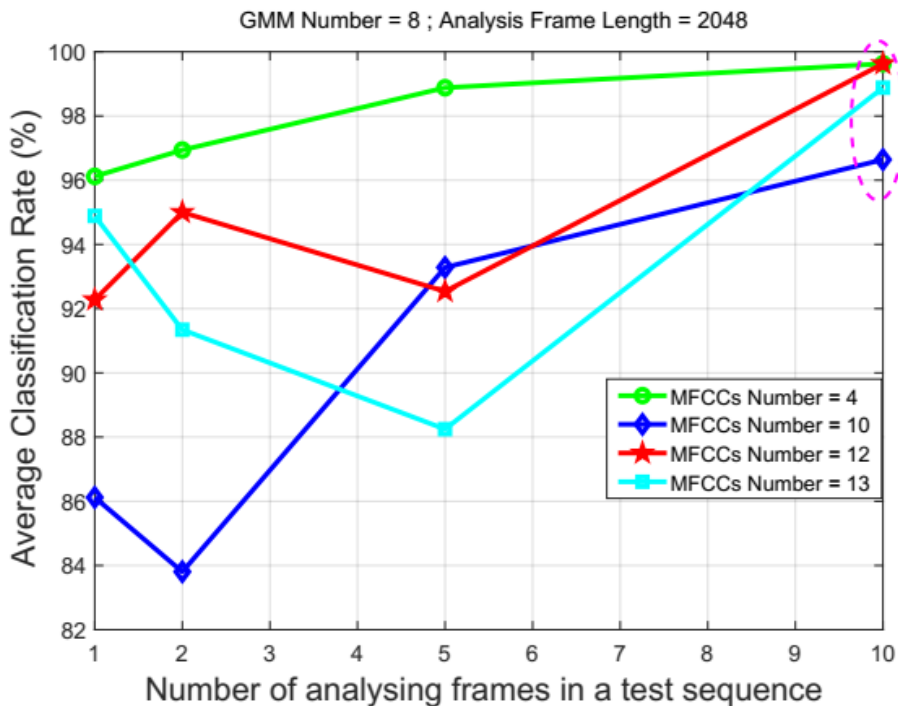


Figure 4.17 : Variation du taux ACR en fonction de la longueur de la séquence de test

Dans cette sous-section, nous avons évalué trois autres possibilités en termes de nombre de trames d'analyse contenues dans une séquence de test (test de trame): cinq (5), deux (2) et une (1) trame. Le résultat de la variation du taux de classification moyen (ACR) en fonction de la longueur de la séquence de test, est illustré dans la Fig. 4.17. Les courbes relatives aux différents nombres de MFCC testés, ont des toutes des allures ascendantes. L'ACR est maximisé à la valeur maximale de la séquence de test. En d'autres termes, l'ACR le plus élevé est attribué à la séquence de test la plus longue (celle qui contient dix (10) trames d'analyse).

Tableau 4.7 : Matrices de confusion : $N_{MFCC} = 4$; $N_{GMM} = 8$; Trame d'analyse=2048 points

Séquence de test = 1 trame d'analyse				Séquence de test = 2 trames d'analyse					
ACR = 94.8881 %				ACR = 96.3433 %					
	BPFI	BSF	BPFO	Normal		BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0	BPFI	1.0000	0	0	0
BSF	0	0.9030	0	0.0970	BSF	0	0.9493	0	0.0507
BPFO	0.0045	0	0.9955	0	BPFO	0.0030	0	0.9970	0
Normal	0	0.1030	0	0.8970	Normal	0	0.0925	0	0.9075

Séquence de test = 5 trames d'analyse				Séquence de test = 10 trames d'analyse					
ACR = 96.9449 %				ACR = 99.1604 %					
	BPFI	BSF	BPFO	Normal		BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0	BPFI	1.0000	0	0	0
BSF	0	0.9552	0	0.0448	BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0	BPFO	0	0	1.0000	0
Normal	0	0.0299	0	0.9701	Normal	0	0.0299	0	0.9701

Cela signifie que l'estimation est meilleure avec les échantillons plus longs que les moins longs. Les matrices de confusion de chaque cas A, B, C et D sont présentées dans le tableau 4.7. La courbe correspondant au nombre de quatre (4) MFCC a le taux de classification le plus élevé. Cela représente à nouveau une confirmation des résultats obtenus précédemment. Les autres courbes ont été tracées pour montrer ces différences notoires entre les ACR liés à chaque nombre de MFCC utilisé.

4.5 DISCUSSIONS DES RESULTATS

4.5.1 Synthèse des performances de la classification avec les LPCs

En résumé à cette étude paramétrique de classification basée sur les descripteurs LPC, les paramètres de la classification obtenus lors des différents tests, sont finalement fixés à leurs valeurs optimales finales. Ainsi, le nombre de composants du mélange est fixé à 2, le nombre de LPC est fixé à 14, la taille de la trame d'analyse est fixée à 1024 échantillons et la taille de la séquence de test est fixée à 10 trames d'analyse.

Les résultats du test réalisé avec ces paramètres optimaux de la classification, sont représentés par la matrice de confusion du tableau 4.8, les nuages de points en 2D (deux dimensions) relatifs à chaque condition de défaut, de la figure 4.18 et le nuage de points en 3D (trois dimensions) de la figure 4.19.

L'ensemble de ces figures, montre bien une dispersion avérée dans les nuages de points relatifs aux différentes classes. Elle est due éventuellement à la modélisation du bruit modélisé ou aux irrégularités liées aux structures des matrices de covariance estimées par l'algorithme EM. Le taux de classement moyen atteint 99,8134 %. C'est un bon résultat obtenu par la méthodologie développée lors de cette étude expérimentale.

Tableau 4.8 : Matrice de confusion : $N_{LPC}=14$; $N_{GMM}=2$; longueur de trame =1024 pts

Séquence de test =10 trames d'analyse				
ACR = 99. 8134 %				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0.0075	0.9925	0	0
BPFO	0	0	1.0000	0
Normal	0	0	0	1.0000

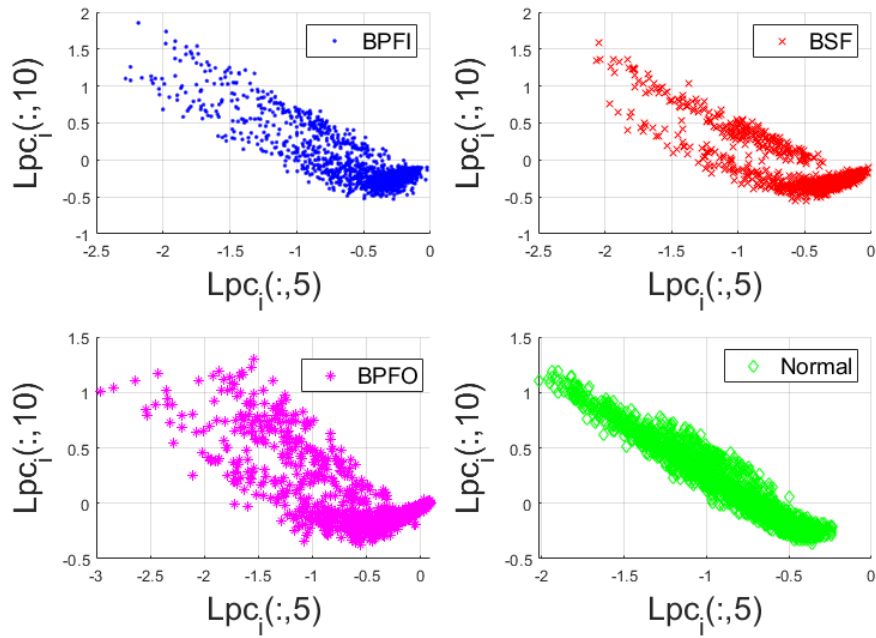


Figure 4.18 : Nuages de points 2D pour les LPCs relatifs à chaque condition de défaut

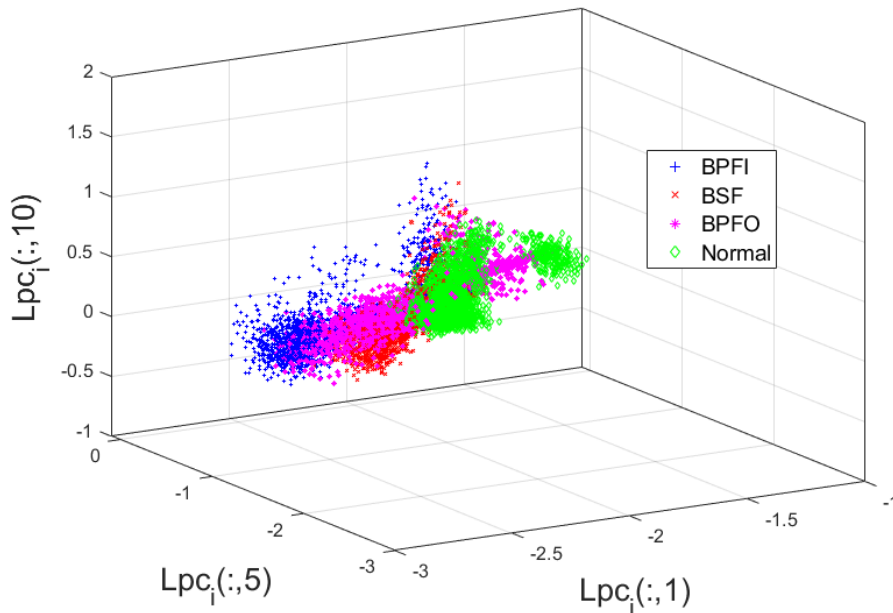


Figure 4.19 : Nuage de points en 3D pour les LPCs relatifs aux conditions de défaut

4.5.2 Synthèse des performances de la classification avec les MFCCs

Au terme de cette étude, les paramètres de la classification optimisée lors des différents tests sont finalement fixés à leurs valeurs optimales finales. Ensuite, le nombre de composants du mélange est fixé à 8, le nombre de MFCC est fixé à 4, la taille de la trame d'analyse est fixée à 2048 échantillons et la taille de la séquence de test est fixée à 10 trames d'analyse.

Les résultats de ce dernier test réalisé avec ces paramètres optimaux ci-dessous sont représentés par la matrice de confusion du tableau 4.9, le nuage de points 2D (deux dimensions) de la figure 4.20 et le nuage de points 3D (trois dimensions) de la figure 4.21. Sur ces figures, on peut facilement remarquer une certaine dispersion dans les classes BPFO et BSF. Elle est probablement due au bruit modélisé ou aux irrégularités liées aux structures des matrices de covariance estimées par l'algorithme EM. Le taux de classement moyen atteint 99,6269%. C'est un bon résultat obtenu par la méthodologie développée lors de cette étude expérimentale.

Tableau 4.9 : Matrice de confusion : $N_{MFCC}=4$; $N_{GMM}=8$; longueur de trame =2048 pts

Séquence de test =10 trames d'analyse				
ACR = 99. 6269 %				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0
Normal	0	0	0	1.0000

Le processus d'optimisation des paramètres dans la méthodologie proposée a rendu le duo GMM-MFCC suffisamment puissant pour fournir une bonne classification. Cependant, il y a encore deux aspects qui n'ont pas été abordés au cours de ces travaux et qui nécessitent des perspectives d'exploration.

Le premier est lié à l'extraction des indicateurs (caractéristiques ou attributs) grâce à une optimisation de la proportion de données nécessaires à l'apprentissage et à l'étude de l'influence du nombre de filtres utilisés pour l'extraction des MFCC. Cette dernière est liée aux fenêtres utilisées, à la taille de la transformée de Fourier et aux différentes bandes du spectre de fréquence. En effet, différentes possibilités doivent être testées et différentes relations doivent encore être établies pour optimiser ce paramètre, augmenter la gamme des MFCC pour une bonne classification.

Le deuxième aspect est lié au classificateur GMM formé de composantes de distribution gaussiennes multivariées. Chaque composante du mélange est définie par sa moyenne, sa covariance, et un vecteur de proportions de mélange (poids). Ces paramètres sont utilisant l'estimation du maximum de vraisemblance (ML) obtenue de manière itérative en utilisant

l'algorithme EM (Expectation-Maximization). Les modèles de mélange gaussiens nécessitent de spécifier un certain nombre de composantes avant d'être appliqués aux données.

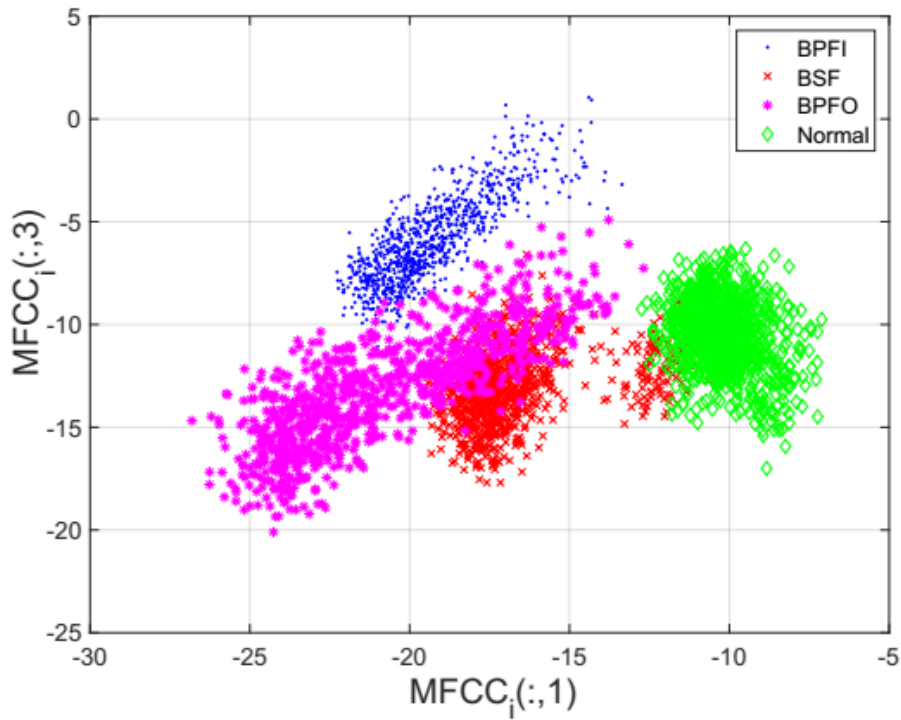


Figure 4.20 : Nuage de points en 2D pour les MFCCs relatifs aux conditions de défaut

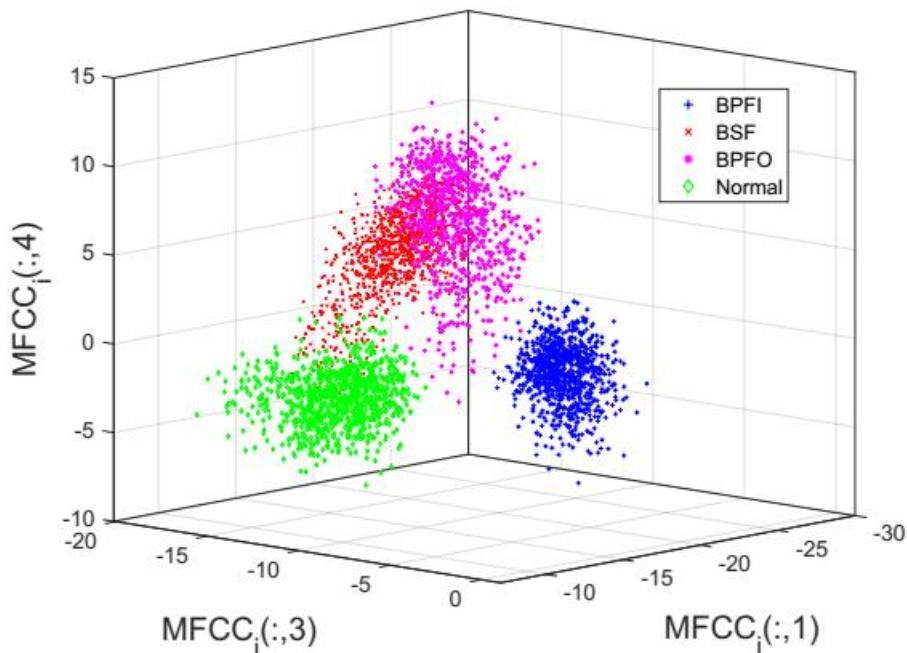


Figure 4.21 : Nuage de points en 3D pour les MFCCs relatifs aux conditions de défaut

Dans de nombreux cas, il peut être difficile de connaître le nombre approprié de composantes. Les critères AIC et BIC, sont des ajustements statistiques qui peuvent aider à choisir le modèle de mélange gaussien le mieux adapté à un nombre variable de composantes.

Ils peuvent être utilisés pour comparer l'ajustement de plusieurs modèles aux mêmes données. En tant que scalaires, des valeurs plus petites d'entre eux permettent un meilleur modèle. La log-vraisemblance négative, le nombre d'itérations dans l'algorithme Expectation-Maximization (EM), la valeur du paramètre de régularisation et l'effet de la tolérance pour les probabilités postérieures sont les autres paramètres qui peuvent être analysés pour le problème d'optimisation dans les explorations de perspectives.

4.5.3 Evaluation des approches utilisées

L'évaluation des approches proposées de classification par les GMM montre l'intérêt du choix à la fois des descripteurs et des paramètres intrinsèques liés au processus de la classification. Les résultats obtenus pour les paramètres optimaux sont résumés dans le Tableau 4.10. L'analyse des résultats passe par l'analyse des performances de chacune des deux approches utilisées. Le taux de classification moyen pour la classification basée sur les LPC est estimé à 99,8134 %. Tandis que celui obtenu grâce aux MFCC atteint 99.6269 %. Ces deux bonnes performances prouvent donc l'usage des descripteurs spectraux, réputés pour leur efficacité dans le traitement de la parole et autres domaines variés, dans le diagnostic des signaux vibratoires.

Tableau 4.10 : Résumé des résultats des deux approches de classification utilisées.

Paramètres optimaux	Classification / LPC	Classification / MFCC
Nombre de composantes GMM	2	8
Nombre de coefficients	14	4
Longueur de trame d'analyse	1024	2048
Longueur de séquence de test	10	10
Taux de classification	99.8134 %	99.6269 %

Cependant, il est à noter que les paramètres de classification sont différents pour les deux approches proposées. Le nombre de composantes du mélange ou nombre de distributions de Gauss, relatif à la classification basée sur les LPCs, est de l'ordre de 2 seulement. Tandis qu'il vaut 8 pour celle basée sur les MFCCs. Ce qui signifie un gain en temps de calcul, vu sous cet angle seulement. Ces nombres de composantes ont été obtenus de façon empirique et loin de toute utilisation de critères tels que AIC, BIC, ou autres.

Sous un autre angle, un autre paramètre de classification important permet de faire la comparaison entre les méthodes proposées. Il s'agit de la longueur des trames d'analyse utilisée. La classification basée sur les LPCs permet l'usage de trames de longueurs variées (de 32 à 16384 échantillons). Alors que celle basée sur les MFCCs est relativement réduite, et réservée aux trames de longueurs élevées (de 512 à 16384 échantillons). Ceci représente encore un avantage de la première approche, vu encore sous cet angle seulement.

Dans tous les cas, et bien que la classification basée sur les LPCs, semble avantageuse en termes des deux critères de comparaison énumérés précédemment, et a présenté un ACR légèrement plus élevé que son homologue basée sur les MFCCs lors de cette étude, elle demeure encore incertaine. La nature des signaux de roulement fait qu'une partie du bruit est toujours modélisée et crée des dispersions dans les classes. Elle est théoriquement plus adaptée aux signaux synchrones (engrenages, balourds, ...). Néanmoins, son application à d'autres bases de données de roulements (signaux de vibrations), peut nous aider d'avantage à statuer sur son application dans le domaine du diagnostic des défauts de roulements.

4.6 CONCLUSION

A travers ce chapitre, l'aspect expérimental lié à la validation des algorithmes développés, a été abordé. L'optimisation de ses paramètres d'entrée de la classification, autre contribution principale de cette étude, à savoir le nombre de LPCs ou de MFCCs, la taille de la trame d'analyse, le nombre de composantes du mélange, etc., a permis d'atteindre au final des Taux de Classification Moyen (ACR) élevés. De plus, et bien que la classification par les GMMs ait été appliquée à une base de données réputée bruyante avec certaines caractéristiques non stationnaires, elle s'est avérée globalement efficace.

A partir des résultats obtenus avec les deux approches proposées, en termes de type d'attributs utilisés, l'approche à base de MFCCs, nous a semblée comme la plus adéquate et la plus adaptée à la problématique de classification (diagnostic de défauts de roulements). L'approche basée sur les LPCs demeure linéaire, et vue la nature des signaux de roulement, elle laisse passer le bruit quelque soient les moyens de filtrage utilisés en prétraitements. Le bruit modélisé laisse apparaitre une dispersion dans le regroupement des classes et réduit les probabilités de discrimination entre celles-ci.

CONCLUSION GENERALE

CONCLUSION GENERALE

Le travail présenté dans cette thèse contribue essentiellement à la reconnaissance intelligente de défauts, par méthode d'apprentissage machine (Machine Learning), dans un contexte de diagnostic des roulements. C'est une méthodologie de diagnostic automatisée des défauts basée sur une approche de classification.

La première contribution est relative à l'utilisation du classifieur GMM dans le diagnostic des défauts mécaniques. En effet, les modèles de mélange gaussiens sont rarement utilisés dans ce domaine, malgré leur succès fabuleux dans les domaines de traitement de la voix, de l'astronomie, de la biologie, de l'économie, des sciences de l'ingénieur, du marketing, de la reconnaissance d'images, et beaucoup d'autres. Plus encore, ils sont souvent classés et utilisés comme méthode non supervisée (clustering). En plus, dans cette étude, c'est leur autre mode de classification supervisée qui est mis en valeur, avec une efficacité prouvée. En conséquence, notre choix de cette méthode plutôt qu'une autre parmi les plus utilisées (ANN, SVM, etc.) dans le diagnostic des défauts, est conforté non seulement par leur simplicité d'application et leur flexibilité d'adaptation à la diversité des données d'analyse, mais aussi par leur double mode d'application en modes, supervisé ou non.

La seconde est liée à l'application avec succès, des attributs dits spectraux (LPC et MFCC) issus du domaine de traitement de la voie en particulier, dans la classification des défauts mécaniques de roulement. Ils ont été volontairement utilisés indépendamment, pour mieux cerner leurs avantages et leurs inconvénients respectifs.

La troisième contribution principale de cette étude, est l'optimisation expérimentale des paramètres d'entrée de la classification à savoir, le nombre de coefficients nécessaires (LPCs ou de MFCCs), la taille de la trame d'analyse, la longueur de la séquence de test et le nombre de composantes du mélange, a permis d'atteindre au final des Taux de Classification Moyen (ACR) élevés. De plus, et bien que la classification par les GMMs ait été appliquée à une base de données réputée bruyante avec certaines caractéristiques non stationnaires, elle s'est avérée globalement efficace.

Cependant, il est à rappeler que les modèles de mélanges de gaussiennes, font partie des méthodes statistiques, et de ce fait, chaque processus de classification doit donc impérativement être répété plusieurs fois (trois dans notre cas) avant de prendre la moyenne des résultats

obtenus. En conséquence, ces répétitions engendrent des coups en termes de temps de calcul et des incertitudes sur la précision des valeurs du taux de classification moyen obtenues.

A partir des résultats obtenus avec les deux approches proposées, en termes de type d'attributs utilisés, l'approche à base de MFCCs, nous a semblée comme la plus adéquate et la plus adaptée à la problématique de classification (diagnostic de défauts de roulements). L'approche basée sur les LPCs demeure linéaire, et vue la nature des signaux de roulement, elle laisse passer le bruit quelque soient les moyens de filtrage utilisés en prétraitements. Le bruit modélisé laisse apparaitre une dispersion dans le regroupement des classes et réduit les probabilités de discrimination entre celles-ci. Cependant, cette approche pourra éventuellement s'adapter à des défauts de type alignement, balourds ou même engrenages, car sa linéarité s'adaptera d'avantage avec les signaux à composantes synchrones. Cela peut ouvrir de nouvelles possibilités pour sa généralisation à d'autres types de défaillances présentes dans les machines tournantes.

Comme perspectives, et pour une éventuelle continuité de ce travail, on propose :

- Augmentation du nombre de classes ;
- Amélioration des prétraitements par usage de filtres adaptatifs ;
- Usage de la cyclostationarité ;
- Usage de méthodes du Deep Learning.

REFERENCES BIBLIOGRAPHIQUES

- [1] **C. Scheffer, P. Girdhar**, Practical Machinery Vibration Analysis and Predictive Maintenance, Newnes An imprint of Elsevier, 2004.
- [2] **O. Djebili**, Contribution à la maintenance prédictive par analyse vibratoire des composants mécaniques tournants. Application aux butées à billes soumises à la fatigue de contact de roulement, Thèse de doctorat, Université De Reims Champagne Ardenne, 2013.
- [3] **A. Boulenger, C. Pachaud**, Analyse vibratoire en maintenance Surveillance et diagnostic des machines, Dunod, L'Usine Nouvelle, Collection Technique et ingénierie - Gestion industrielle, 2007.
- [4] **P. D. McFadden, J.D. Smith**, Vibration monitoring of roller element bearing by the high-frequency resonance technique – a review, Tribology International, vol. 17, no. 1, pp 3–10, 1984.
[https://doi.org/10.1016/0301-679X\(84\)90076-8](https://doi.org/10.1016/0301-679X(84)90076-8)
- [5] **J. P. Dron, F. Bolaers, L. Rasolofondraibe**, Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction : application to the detection of defects in ball bearings, Journal of Sound and Vibration, vol. 270, no. 1-2, pp. 61-73, 2004.
- [6] **R. B. Randall, J. Antoni**, Rolling element bearing diagnostics – A tutorial, Mechanical Systems and Signal Processing, vol. 25, pp 485-520, 2011.
- [7] **G. Yu**, A Concentrated Time–Frequency Analysis Tool for Bearing Fault Diagnosis, IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 2, pp 371 – 381, 2019.
<https://doi.org/10.1109/TIM.2019.2901514>
- [8] **I. R. Quinde, J. C. Sumba, L. E. Ochoa, A. Jr. V. Guevara, R. M. Menendez**, Bearing Fault Diagnosis Based on Optimal Time-Frequency Representation Method, IFAC Papers On Line vol. 52, no. 11, pp 194–199, 2019.
<https://doi.org/10.1016/j.ifacol.2019.09.140>
- [9] **H. Li, H. Zheng, L. Tang**, Wigner-Ville Distribution Based on EMD for Faults Diagnosis of Bearing, International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2006: Lecture Notes in Computer Science, vol. 4223, Springer, Berlin, Heidelberg, pp 803-812, 2006.
https://doi.org/10.1007/11881599_99
- [10] **Y. Bella, A. Oulmane, M. Mostefai**, Industrial Bearing Fault Detection Using Time-Frequency Analysis, Engineering, Technology & Applied Science Research, Vol. 8, No. 4, pp 3294-3299, 2018.
- [11] **V. Purushotham, S. Narayanan, S.A.N. Prasad**, Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, NDT&E International, vol. 38, pp 654–664, 2005.
<https://doi.org/10.1016/j.ndteint.2005.04.003>

- [12] **J. Tian, C. Morillo, M. H. Azarian, M. Pecht**, Motor bearing fault detection using spectral Kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis,” IEEE Trans. Ind. Electron., vol. 63, no. 3, pp. 1793-1803, 2016.
<https://doi.org/10.1109/TIE.2015.2509913>
- [13] **T. Benkedjouh, T. Chettibi, Y. Saadouni, M. Afroun**, Gearbox Fault Diagnosis Based on Mel-Frequency Cepstral Coefficients and Support Vector Machine, Computational Intelligence and Its Applications: 6th IFIP TC 5 International Conference, CIIA, 2018.
https://doi.org/10.1007/978-3-319-89743-1_20
- [14] **D. H. Pandya, S.H. Upadhyay, S.P. Harsha**, ANN based fault diagnosis of rolling element bearing using time-frequency domain feature, International Journal of Engineering Science and Technology (IJEST), Vol. 4, no.06, 2012.
- [15] **G. Yu, C. Li, J. Sun**, Machine fault diagnosis based on Gaussian mixture model and its application, Int J Adv Manuf Technol, Vol. 48, pp 205–212, 2010.
<https://doi.org/10.1007/s00170-009-2283-5>
- [16] **J. Yu**, Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models, Mech. Syst. Signal Process. Vol. 25, no. 7, pp 2573–2588, 2011.
<https://doi.org/10.1016/j.ymssp.2011.02.006>
- [17] **F. V. Nelwamondo, T. Marwala, U. Mahola**, Early classifications of bearing faults using hidden Markov models, Gaussian mixture models, Mel-frequency cepstral coefficients and fractals, International Journal of Innovative Computing, Information and Control Vol. 2, no. 6, pp 1281-1299, 2006.
- [18] **X. Guo, L. Chen, C. Shen**, Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis, Meas., Vol. 93, pp 490 - 502, 2016.
<https://doi.org/10.1016/j.measurement.2016.07.054>
- [19] **Z. Zhuang, Q. Wei**, Intelligent fault diagnosis of rolling bearing using one-dimensional multi-scale deep convolutional neural network based health state classification, in Proc. 2018 IEEE 15th Int. Conf. Netw., Sens. & Control (ICNSC), Zhuhai, China, pp 1 – 6, 2018.,
<https://doi.org/10.1109/ICNSC.2018.8361296>
- [20] **W. Mao, J. He, Y. Li, Y. Yan**, Bearing fault diagnosis with auto-encoder extreme learning machine: A comparative study, Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci., Vol. 231, no. 8, pp 1560 - 1578, 2016.
<https://doi.org/10.1177/0954406216675896>
- [21] **K. Zhao, H. Shao**, Intelligent Fault Diagnosis of Rolling Bearing Using Adaptive Deep Gated Recurrent Unit, Neural Processing Letters, 51, 1165–1184, 2020.
<https://doi.org/10.1007/s11063-019-10137-2>

- [22] **H. Schneider**, Technique d'équilibrage, Traduction de la 6e édition allemande, Springer-Verlag France, 2006.
- [23] **N. Feki**, Modélisation électro-mécanique de transmissions par engrenages – Applications à la détection et au suivi des avaries. Thèse de Doctorat, INSA de Lyon, 2012.
- [24] <http://www.novexa.com/engrenage-defauts.php> , date d'accès 07/04/2020.
- [25] http://www.vibrexon.com/surveillance_engrenages.php , date d'accès 07/04/2020.
- [26] **Niabot**, Rolling-element bearing (numbered) - Wikimedia Commons,
[http://commons.wikimedia.org/wiki/File:Rolling
element_bearing_\(numbered\).png?uselang=fr#filelinks](http://commons.wikimedia.org/wiki/File:Rolling_element_bearing_(numbered).png?uselang=fr#filelinks), date d'accès 07/04/2020.
- [27] **M. Mauret**, Classement et caractéristiques des roulements, Techniques de l'ingénieur, 1990.
- [28] **F. Bolaers**, Contribution à l'étude et au développement d'un système intégré de suivi de l'endommagement de composants mécaniques sur machines tournantes. Thèse, Université de Reims, 2002.
- [29] <https://www.emersonbearing.com/technical-toolbox.html>, date d'accès 08/04/2020.
- [30] <https://www.nskeurope.fr/fr/services/troubleshooting/damage-by-type/seizure.html#>, date d'accès 09/04/2020.
- [31] **M. Hafida**, Analyse non stationnaire des signaux vibratoires dans la surveillance des machines et la prévention des défaillances, Thèse de Doctorat, Université Ferhat Abbas Setif1 – Setif, 2013.
- [32] [https://www.skf.com/fr/news-and-media/news-search/2017-01-22-
5_most_common_causes_of_bearing_failures_in_electric_motors_and_what_to_do_about_them.
html](https://www.skf.com/fr/news-and-media/news-search/2017-01-22-5-most-common-causes-of-bearing-failures-in-electric-motors-and-what-to-do-about-them.html), date d'accès 01/04/2020.
- [33] **H. Ocak, K. A. Loparo**, Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data, Mechanical Systems and Signal Processing, Vol. 18, pp 515–533, 2004
- [34] **A. Ibrahim**, Contribution au diagnostic de machines électromécaniques : Exploitation des signaux électriques et de la vitesse instantanée, thèse de Doctorat, Université Jean Monnet, Roanne, 2009.
- [35] **H. Yang**, Automatic Fault Diagnosis of Rolling Element Bearings Using Wavelet Based Pursuit Features, Doctor of Philosophy thesis, Queensland University of Technology, 2004.
- [36] **P. Stepanic, I.V. Latinovic, Z. Djurovic**, A new approach to detection of defects in rolling element bearings based on statistical pattern recognition, International Journal of Advanced Manufacturing Technology, Vol 45, pp 91–100, 2009.
- [37] **B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi**, Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection, Engineering Applications of Artificial Intelligence, Vol 16, pp 657–665, 2003.

- [38] **Y. Zhang, H. Zuo, F. Bai**, Classification of fault location and performance degradation of a roller bearing, *Measurement*, Vol 46, pp 1178–1189, 2013.
- [39] **R. Ziani**, Contribution à l'analyse et à la classification automatique des défauts mécaniques, Thèse de Doctorat, Université Ferhat Abbas Setif1– Sétif, 2015.
- [40] **R. Bigret, J.L. Féron, C. Pachaud**, Diagnostic, maintenance, disponibilité des machines tournantes - modèles, mesurages, analyses des vibrations, Masson, Paris, 1995.
- [41] **H. Li, X. Lian, C. Guo, P. Zhao**, Investigation on early fault classification for rolling element bearing based on the optimal frequency band determination, *Journal of Intelligent Manufacturing*, 2013.
<https://doi.org/10.1007/s10845-013-0772-8>
- [42] **F. Bolaers, O. Cousinard, P. Marconnet, L. Rasolofondraibe**, Advanced detection of rolling bearing spalling from de-noising vibratory signals, *Control Engineering Practice*, vol. 12, no. 2, pp. 181-190, 2004.
- [43] **A. Sturm, D. Kinsky**, Diagnostics of rolling-element bearing condition by means of vibration monitoring under operating conditions. *Measurement*, vol. 2, no. 2, pp 58-62, 1984.
[https://doi.org/10.1016/0263-2241\(84\)90033-2](https://doi.org/10.1016/0263-2241(84)90033-2)
- [44] **H.R. Martin, F. Honarvar**, Application of statistical moments to bearing failure detection. *Applied acoustics*, vol. 44, no. 1, pp 67-77, 1995.
[https://doi.org/10.1016/0003-682X\(94\)P4420-B](https://doi.org/10.1016/0003-682X(94)P4420-B)
- [45] **F. Honarvar, H.R. Martin**, New statistical moments for diagnostics of rolling element bearings. *Journal of Manufacturing Science and Engineering*, vol. 119 no. 3, pp 425-432, 1997.
<https://doi.org/10.1115/1.2831123>
- [46] **R. B. W. Heng, M. J. M. Nor**, Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition, *Applied Acoustics*, Vol. 53, no. 1, pp 211-226, 1998.
[https://doi.org/10.1016/S0003-682X\(97\)00018-2](https://doi.org/10.1016/S0003-682X(97)00018-2)
- [47] **S. Braun**, The extraction of periodic waveforms by time domain averaging, *Acoustica* vol. 23 no. 2, pp 69–77, 1975.
- [48] **P.D. McFadden**, A revised model for the extraction of periodic waveforms by time domain averaging, *Mechanical systems and signals processing*, vol. 1, no. 1, pp 83-95, 1987.
- [49] **N.C. Kongom, N. Mureithi, L. Aouni, T. Marc**, 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*, Montreal pp. 610–624, 2007.
- [50] **W. Wang, N. Sawalhi, A. Becker**, Size Estimation for Naturally Occurring Bearing Faults Using Synchronous Averaging of Vibration Signals, *Transactions of the ASME, Journal of Vibration and Acoustics*, Vol. 138, 2016.

- [51] **A. Felkaoui**, Diagnostic vibratoire au service de la maintenance : de l'acquisition à la décision, Thèse de Doctorat, Université Ferhat Abbas Setif1– Sétif, 2007.
- [52] **S. Fortier**, Les modèles MA, AR et ARMA multidimensionnels : estimation et causalité, CAMUS, 4, pp 112 – 136, Université de Sherbrooke, 2013.
- [53] **H. Wold**, A study in the Analysis of Stationary Time Series, Almgvist and Wiksell, 1938.
- [54] **G. Box, G. Jenkins**, Time series analysis : Forecasting and Daviscontrol, Holden Day, 1970.
- [55] **C.K. Mechefske, J. Mathew**, Fault detection and diagnosis in low speed rolling element bearings. I: The use of parametric spectra, Mechanical Systems and Signal Processing, Vol. 6, no. 4, pp 297–307, 1992.
[https://doi.org/10.1016/0888-3270\(92\)90032-E](https://doi.org/10.1016/0888-3270(92)90032-E)
- [56] **C. K. Mechefske, J. Mathew**, Parametric spectral estimation to detect and diagnose faults in low speed rolling element bearings: Preliminary investigations, Mechanical Systems and Signal Processing, Vol. 7, no. 1, pp 1–12, 1993.
[https://doi.org/10.1016/0888-3270\(93\)90001-D](https://doi.org/10.1016/0888-3270(93)90001-D)
- [57] **D.C. Baillie, J. Mathew**, A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings, Mechanical Systems and Signal Processing, Vol. 10, pp 1–17, 1996.
- [58] **R. B. Randall**, Frequency Analysis, Naerum, Denmark: Brüel & Kjaer, 1987.
- [59] **J. Harmouche, D. Diallo, C. Delpha**, Discrimination des défauts de roulements par une analyse spectrale globale, Symposium de Génie Électrique 2014, Cachan, France, 2014.
- [60] **S. Tyagi**, A Comparative Study of SVM Classifiers and Artificial Neural Networks Application for Rolling Element Bearing Fault Diagnosis using Wavelet Transform Preprocessing, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering Vol. 2, no. 7, 2008.
- [61] **Y. Lei, Z. He, Y. Zi, X. Chen**, New clustering algorithm-based fault diagnosis using compensation distance evaluation technique, Mechanical Systems and Signal Processing, Vol. 22, no. 2, pp 419-435, 2008.
<https://doi.org/10.1016/j.ymssp.2007.07.013>
- [62] **J-H Cai**, Fault diagnosis of rolling bearing based on empirical mode decomposition and higher order statistics, J Mechanical Engineering Science, Vol. 229, no. 9, pp 1630–1638, 2015.
<https://doi.org/10.1177/0954406214545820>
- [63] **K.C. Gryllias, IA Antoniadis**, A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments, Engineering applications of Artificial Intelligence, Vol. 25, no. 2, pp 326 - 344, 2012.
<https://doi.org/10.1016/j.engappai.2011.09.010>

- [64] **I. Howard**, A Review of rolling element bearing vibration - detection, diagnosis and prognosis, DSTO Aeronautical and Maritime Research Laboratory, DSTO-RR-0013, Australia, 1994.
- [65] **M.S. Darlow, R.H. Badgley, G.W. Hogg**, Application of high frequency resonance techniques for bearing diagnostics in helicopter gearboxes, US Army Air Mobility Research and Development Laboratory Technical Report-74-77, 1974.
- [66] **H. Prashad, M. Ghosh, S. Biswas**, Diagnostic monitoring of rolling-element bearings by high-frequency resonance technique, ASLE Transactions, Vol. 28, no. 4, pp 439-448, 1985.
<https://doi.org/10.1080/05698198508981641>
- [67] **P.W. Tse, Y. H. Peng, R. Yam**, Wavelet Analysis and Envelope Detection For Rolling Element Bearing Fault Diagnosis - Their Effectiveness and Flexibilities, Journal of Vibration and Acoustics, Vol. 123, no. 3, pp 303-310, 2001.
<https://doi.org/10.1115/1.1379745>
- [68] **Y-T. Sheen**, An envelope analysis based on the resonance modes of the mechanical system for the bearing defect diagnosis, Measurement, Vol. 43, pp 912 - 934, 2010.
<https://doi.org/10.1016/j.measurement.2010.03.011>
- [69] **D. Ho, R.B. Randall**, Optimisation of bearing diagnostic Techniques using simulated and actual Bearing fault signals, Mechanical Systems and Signal Processing, Vol.14, no. 5, 763-788, 2000.
<https://doi.org/10.1006/mssp.2000.1304>
- [70] **J. Antoni, M. Sidahmed**, Contrôle et diagnostic à partir des signaux acoustiques et vibratoires, Acoustique & Techniques Vol. 38, pp 9-15, 2004.
- [71] **C-S. Park, Y- C. Choi, Y- H. Kim**, Early fault detection in automotive ball bearings using the minimum variance cepstrum. Mechanical Systems and Signal Processing, Vol. 38, no. 2, pp 534-548, 2013.
<https://doi.org/10.1016/j.ymssp.2013.02.017>
- [72] **P. Borghesani, P. Pennacchi, R.B. Randall, N. Sawalhi, R. Ricci**, Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions. Mechanical Systems and Signal Processing, Vol. 36, 2, pp 370-384, 2013.
<https://doi.org/10.1016/j.ymssp.2012.11.001>
- [73] **X. Zhang, , N. Hu, , L. Hu, B. Fan, Z. Cheng**, Enhanced Fault Detection of Rolling Element Bearing Based on Cepstrum Editing and Stochastic Resonance, In Journal of Physics, Conference Series, Vol. 364, no. 1, IOP Publishing, 2012.
<https://doi.org/10.1088/1742-6596/364/1/012029>
- [74] **S. Chakkor**, E-diagnostic de processus physiques à base des méthodes de haute résolution Application : machines éoliennes, Thèse de Doctorat, Université Abdelmalek Essaâdi, Tétouan, Maroc, 2015.

- [75] **S. Agrawal, S.R. Mohanty, V. Agarwal**, Bearing Fault Detection Using Hilbert and High Frequency Resolution Techniques, *IETE Journal of Research*, vol. 61, no. 2, pp 99-108, 2015.
<https://doi.org/10.1080/03772063.2015.1009398>
- [76] **A. Garcia-Perez, R. J Romero-Troncoso, E. Cabal-Yepez, R.A. Osornio-Rios**, The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors, *IEEE transactions on industrial electronics*, vol. 58, no. 5, 2011.
<https://doi.org/10.1109/TIE.2010.2051398>
- [77] **L. Cohen**, Time–frequency distribution - A review, *Proceedings of the IEEE*, Vol. 77, no. 7, pp 941-981, 1989.
<https://doi.org/10.1109/5.30749>
- [78] **L. Cohen**, *Time-frequency analysis*, Prentice Hall PTR, 1995.
- [79] **R. B. Randall**, *Vibration-based condition monitoring : industrial, aerospace and automotive applications*. John Wiley & Sons, Ltd., 2011.
- [80] **D. Liu, W. Cheng, W. Wen**, Rolling bearing fault diagnosis via STFT and improved instantaneous frequency estimation method, *Procedia Manufacturing*, Vol. 49, 166-172, 2020.
<https://doi.org/10.1016/j.promfg.2020.07.014>
- [81] **K.M. Arun Kumar, T.C. Manjunath, G. Arun Kumar**, Bearing Fault Diagnosis in IM Using STFT and J-48 Algorithm based on Vibration Signals in Dynamic Machines, *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 7, no 5S2, 2019.
- [82] **B. Boashash**, *Time-Frequency Signal Analysis and Processing - A Comprehensive Reference*, Academic Press, 2003.
- [83] **P. Singru, V. Krishnakumar, D. Natarajan, A. Raizada**, Bearing failure prediction using Wigner-Ville distribution, modified Poincare mapping and fast Fourier transform, *Journal of vibroengineering*, vol. 20, no. 1, 1392-8716, 2018.
<https://doi.org/10.21595/jve.2017.17768>
- [84] **Y. Zhou, J. Chen, G.M. Dong, W.B. Xiao, Z.Y. Wang**, Wigner–Ville distribution based on cyclic spectral density and the application in rolling element bearings diagnosis, *Proc. IMechE Vol. 225 Part C: Journal of Mechanical Engineering Science*, Vol. 225 no. 12, pp 2831-2847, 2011.
<https://doi.org/10.1177/0954406211413215>
- [85] **H.-I. Choi, W.J. Williams**, Improved time-frequency representation of multicomponent signals using exponential kernels, *IEEE Trans. Acoust. Speech Signal Process*, Vol. 37, no. 6, pp 862-871, 1989.
<https://doi.org/10.1109/ASSP.1989.28057>

- [86] **X.C. Wei, Y.L. Tang, T. Chen**, Research of Rolling Bearing Fault Feature Extraction Based on EMD and Choi-Williams, *Advanced Materials Research (Volumes 694-697)*, pp 1377-1381, 2013.
<https://doi.org/10.4028/www.scientific.net/AMR.694-697.1377>
- [87] **S. Gu, J. Ni, J. Yuan**, Non-stationary signal analysis and transient machining process condition monitoring, *International Journal of Machine Tools & Manufacture* Vol. 42, pp 41–51, 2002.
- [88] **J. Antoni**, The spectral kurtosis: a useful tool for characterizing non-stationary signals. *Mechanical Systems and Signal Processing*, Vol 20, no. 2, pp 282-307, 2006.
- [89] **S.G. Mallat**, A Theory of Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern and Machine Intelligence* Vol. 11 no.7, pp 674-693, 1989.
- [90] **I. Daubechies**, Ten lectures on wavelets, Philadelphia, society for industrial and applied Mathematics, SIAM, 1992.
- [91] **H. Qiu, J. Lee, J. Lin, G. Yu**, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics, *Journal of Sound and Vibration* , Vol 289, pp 1066–1090, 2006.
<https://doi.org/10.1016/j.jsv.2005.03.007>
- [92] **M. Deriche**, Bearing fault diagnosis using wavelet analysis, 1st International Conference on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering, Kuala Lumpur, Malaysia, IEEE, 197 – 201, 2005.
<https://doi.org/10.1109/CCSP.2005.4977189>
- [93] **N.G. Nikolaou, I.A. Antoniadis**, Rolling element bearing fault diagnosis using wavelet packets, *NDT&E International*, Vol. 35, 197–205, 2002.
[https://doi.org/10.1016/S0963-8695\(01\)00044-5](https://doi.org/10.1016/S0963-8695(01)00044-5)
- [94] **N.E. Huang, Z. Shen, S.R Long**, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis in: *Proceedings of the Royal Society of London Series*, Vol 454, 903-995, 1998.
- [95] **F. Tavakkoli, M. Teshnehlab**, A ball bearing fault diagnosis method based on wavelet and EMD energy entropy mean, *IEEE 2007 International Conference on Intelligent and Advanced Systems (ICIAS)*, Kuala Lumpur, 1210-1212, 2007.
<https://doi.org/10.1109/ICIAS.2007.4658576>
- [96] **X. Zhang , J. Kang , L. Hao, L. Cai , J. Zhao**, Bearing fault diagnosis and degradation analysis based on improved empirical mode decomposition and maximum correlated kurtosis deconvolution, *Journal of Vibroengineering JVE*, Vol. 17, no 1, 243-260, 2015.
- [97] **N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, K. L. Fan**, A confidence limit for the empirical mode decomposition and Hilbert spectral analysis, *Proceedings*

- of the Royal Society A Mathematical Physical & Engineering Sciences, Vol 459, no 2037, 2317–2345, 2003.
<https://doi.org/10.1098/rspa.2003.1123>
- [98] **A. Utpat, R. B. Ingle, M. R. Nandgaonkar**, A Model for Study of the Defects in Rolling Element Bearings at Higher Speed by Vibration Signature Analysis. World Academy of Science, Engineering and Technology, International Journal of Mechanical and Mechatronics Engineering, Vol 3, no 8, 879-885, 2009.
- [99] **Z. Wu, N. E. Huang**, Ensemble empirical mode decomposition : A noise-assisted data analysis method, Advances in Adaptive Data Analysis, Vol. 1, no 1, 1-41, 2009.
- [100] **D. H. Lee, J. H. Ahn, B. H. Koh**, Fault Detection of Bearing Systems through EEMD and Optimization Algorithm, Sensors, vol 17, no 11, 2017.
<https://doi.org/10.3390/s17112477>
- [101] **X. Qin, Q. Li, X. Dong, S. Lv**, The Fault Diagnosis of Rolling Bearing Based on Ensemble Empirical Mode Decomposition and Random Forest, Hindawi Shock and Vibration, Vol 2017, 2017.
<https://doi.org/10.1155/2017/2623081>
- [102] **K. Dragomiretskiy, Z. Dominique**, Variational mode decomposition, IEEE Transactions on Signal Processing, Vol. 62, 531-544, 2014.
- [103] **S. Mohanty, K. K. Gupta, K. S. Raju**, Bearing fault analysis using variational mode decomposition, 2014 9th International Conference on Industrial and Information Systems (ICIIS), Gwalior, IEEE, 1-6, 2014.
<https://doi.org/10.1109/ICIINFS.2014.7036617>
- [104] **M. Firdaus Isham, M. Salman Leong, M. Hee Lim, Z. Asrar Ahmad**, Variational mode decomposition: mode determination method for rotating machinery diagnosis, Journal of Vibroengineering JVE, Vol. 20, no 7, 2604-2621, 2018.
<https://doi.org/10.21595/jve.2018.19479>
- [105] **J. Antoni**, Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions, Journal of Sound and Vibration 304, 497–529, 2007.
- [106] **R.B. Randall, J. Antoni, S. Chobsaard**, The relationship between spectral correlation and envelope analysis in the diagnosis of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, Vol 15, pp 945-962, 2001.
- [107] **C. M. Bishop**, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.
- [108] **S. Samarasinghe**, Neural Networks for Applied Sciences and Engineering from fundamentals to Complex Pattern Recognition, Auerbach Publication, Taylor & Francis Group, LLC, 2006.

- [109] **B. Samanta, K. R. Al-Balushi, S. A. Al-Araimi**, Bearing Fault Detection Using Artificial Neural Networks and Genetic Algorithm, *EURASIP Journal on Applied Signal Processing*, Hindawi Publishing Corporation, Vol. 3, 366-377, 2004.
- [110] **O. J. lara castro, C. C. Sisamón, J. C. García Prada**, Bearing Fault Diagnosis based on Neural Network Classification and Wavelet Transform, *Proceedings of the 6th WSEAS International Conference on Wavelet Analysis & Multirate Systems*, Bucharest, Romania, 2006.
- [111] **B. Badri, M. Thomas, S. Sassi**, Etude et développement d'un système expert basé sur les réseaux de neurones pour le diagnostic des défauts de roulements, *Proceedings of the 24th seminar on machinery vibration*, ISBN-2-921145-61-8, Montréal, 2006.
- [112] **A. Djeflal**, Utilisation des méthodes Support Vector Machine (SVM) dans l'analyse des bases de données, Thèse de doctorat, Université Mohamed Khider Biskra, 2012.
- [113] **C. Rajeswari, B. Sathiyabhama, S. Devendiran, K. Manivannan**, Bearing Fault Diagnosis using Multiclass Support Vector Machine with efficient Feature Selection Methods, *International Journal of Mechanical & Mechatronics Engineering IJMME-IJENS Vol.15, no 01*, 2015.
- [114] **C. J. Lin, W.-L. Chu, C.-C. Wang, C.-K. Chen, I-T. Chen**, Diagnosis of ball-bearing faults using support vector machine based on the artificial fish-swarm algorithm, *Journal of Low Frequency Noise, Vibration and Active Control*, SAGE, Vol.0, no 0, 1-14, 2019.
<https://doi.org/10.1177/1461348419861822>
- [115] **R. Semil, P. Jaiswal**, Bearing Fault Diagnosis using Support Vector Machine with Genetic Algorithms Based Optimization and K Fold Cross-Validation Method, *International Journal of Recent Technology and Engineering (IJRTE)*, Vol.8, no 2, 2019.
<https://doi.org/10.35940/ijrte.B2828.078219>
- [116] **A. Sharma, R. Jigyasu, L. Mathew, S. Chatterji**, Bearing Fault Diagnosis Using Weighted K-Nearest Neighbor, *Proceedings of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, IEEE, 2018.
- [117] **P. Agrawal, P. Jayaswal**, Selection of Best Classification Algorithm for Fault Diagnosis of Bearing using Vibration Signature Analysis, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, Vol. 8, no 5, 2019.
- [118] **M. Boumahdi, J. P. Dron, S. Rechak, O. Cousinard**, On the extraction of rules in the identification of bearing defects in rotating machinery using decision tree, *Expert Systems with Applications Vol. 37*, 5887-5894, 2010.
<https://doi.org/10.1016/j.eswa.2010.02.017>
- [119] **HH. Lee, NT. Nguyen, JM. Kwon**, Bearing Diagnosis Using Time-Domain Features and Decision Tree. In: Huang DS., Heutte L., Loog M. (eds) *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. ICIC 2007. Lecture Notes in Computer Science*, vol 4682. Springer, Berlin, Heidelberg, 2007.

- https://doi.org/10.1007/978-3-540-74205-0_99
- [120] **V. Muralidharana, V. Sugumaran**, A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis, *Applied Soft Computing* Vol.12, 2023-2029, 2012.
<https://doi.org/10.1016/j.asoc.2012.03.021>
- [121] **H. Kumar, T.A. Ranjit Kumar, M. Amarnath, V. Sugumaran**, Fault diagnosis of bearings through vibration signal using Bayes classifiers, *International Journal of Computer Aided Engineering and Technology*, Vol. 6, no 1, 2014.
<https://doi.org/10.1504/IJCAET.2014.058002>
- [122] **X. Lou, K. A. Loparo**, Bearing fault diagnosis based on wavelet transform and fuzzy inference, *Mechanical Systems and Signal Processing*, Vol. 18, no 5, 2004.
[https://doi.org/10.1016/s0888-3270\(03\)00077-3](https://doi.org/10.1016/s0888-3270(03)00077-3)
- [123] **R. Tiwari, V. K. Gupta, P.K. Kankar**, Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier, *Journal of Vibration and Control JVC*, Vol. 21, no3, 461-467, 2015.
<https://doi.org/10.1177/1077546313490778>
- [124] **C. Lu, Z. Wang, B. Zhou**, Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification, *Adv. Eng. Informat.*, vol. 32, pp. 139-157, 2017.
<https://doi.org/10.1016/j.aei.2017.02.005>
- [125] **L. Wen, X. Li, L. Gao, Y. Zhang**, A new convolutional neural network-based data-driven fault diagnosis method, *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990-5998, 2018.
<https://doi.org/10.1109/TIE.2017.2774777>
- [126] **F. Wang, B. Dun, G. Deng, H. Li, Q. Han**, A deep neural network based on kernel function and auto-encoder for bearing fault diagnosis,” in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Houston, TX, USA, pp. 1–6, 2018.
- [127] **C. Li, W. Zhang, G. Peng, S. Liu**, Bearing fault diagnosis using fully-connected winner-take-all autoencoder,” *IEEE Access*, vol. 6, pp. 6103-6115, 2018.
- [128] **L. Guo, N. Li, F. Jia, Y. Lei, J. Lin**, A recurrent neural network based health indicator for remaining useful life prediction of bearings,” *Neurocomput.*, vol. 240, pp. 98-109, 2017.
- [129] **Z. Chen, W. Li**, Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693-1702, 2017.
- [130] **W. Mao, Y. Liu, L. Ding, Y. Li**, Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: A comparative study,” *IEEE Access*, vol. 7, pp. 9515–9530, 2019.

- [131] **C. Bouveyron**, Modélisation et classification des données de grande dimension application à l'analyse d'images, Thèse de doctorat, Université Joseph Fourier - Grenoble 1, 2006.
- [132] **C. M. Bishop**, Pattern Recognition and Machine Learning, Springer science + Business Media, 2006.
- [133] **R. Fletcher**, Practical Methods of Optimization, Second edition, Wiley, 1987.
- [134] **A. P. Dempster, N. M. Laird, D. B. Rubin**, Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1, pp. 1-38, 1977.
- [135] **A. Mesloub, K. Abed-Meraim, A. Belouchrani**, Ground Moving Target Classification Based on Micro-Doppler Signature Using Novel Spectral Information Features, European Radar Conference (EURAD), Nuremberg, October 11-13, 2017, IEEE Xplore, 2018.
- [136] **S. G. Koolagudi, Y. V. Srinivasa Murthy, S. P. Bhaskar**, Choice of a classifier, based on properties of a dataset: case study - speech emotion recognition, International Journal of Speech Technology, 2018.
- [137] **D. A. Reynolds, T. F. Quatieri, R. B. Dunn**, Speaker verification using adapted gaussian mixture models. Digital SignalProcessing, Vol. 10, no. 1, 19-41, 2000.
- [138] **G.J. McLachlan, D. Peel**, Finite Mixture Models. New York: John Wiley & Sons, 2000.
- [139] **M. Aublin, R. Boncompain, M. Boulaton, D. Caron, E. Jeay, B. Lacage, J. Réa**, Systèmes mécaniques, Théorie et dimensionnement, Dunod, 1992.
- [140] **H. Hertz**, On the contact of elastic solids, Journal für die reine und angewandte Mathematik, Vol.92, pp156-171, 1881.
- [141] **A. Palmgren**, Ball and roller bearing engineering, 3rd Edition, S.H. Burbank & Co, Philadelphia, USA, 1959.
- [142] **J.M. De Mul, J.M. Vree, D.A. Maas**, Equilibrium and associated load distribution in ball and roller bearings loaded in five degrees of freedom while neglecting friction : Part I - General Theory and Application to Ball Bearings, Journal of Tribology, Vol. 111, no. 1, pp 142- 148, 1989.
<https://doi.org/10.1115/1.3261864>
- [143] **A. Bourdon, J.F. Rigal, D. Play**, Static rolling bearing models in a C.A.D. Environment for the study of complex mechanisms : Part I - Rolling bearing model, Journal of Tribology, Vol. 121, no. 2, pp 205-214, 1999.
<https://doi.org/10.1115/1.2833923>
- [144] **S. Sassi, B. Badri, M. Thomas**, A Numerical Model to Predict Damaged Bearing Vibrations, Journal of Vibration and Control, Vol 13, no. 11, 2007.
<https://doi.org/10.1177/1077546307080040>

- [145] **N.G. Nikolaou, I.A. Antoniadis**, Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets, *Mechanical Systems and Signal Processing*, Vol. 16, no. 4, 677–694, 2002.
- [146] **B. J. Hamrock**, *Fundamentals of Fluid Film Lubrication*, McGraw-Hill, Singapore, 1994.
- [147] **S. Ericsson, N. Grip, E. Johansson, L.E. Persson, R. Sjoberg, J.O. Stromberg**, Towards automatic detection of local bearing defects in rotating machines, *Mechanical Systems and Signal Processing*, vol.19, pp 509-535, 2004.
- [148] **R.O. Duda, P. E. Hart**, *Pattern Classification and Scene Analysis*. John Wiley and Sons, Inc. 1973.
- [149] **W. Wang**, Autoregressive model-based diagnostics for gears and bearings, *Insight - Non-Destructive Testing and Condition Monitoring*, Vol. 50, no. 8, pp 414–418, 2008.
<https://doi.org/10.1784/insi.2008.50.8.414>
- [150] **F.E.C. Ruiz-carcel, D. Mba, P. Chandra**, A Comparative Study of the Effectiveness of Adaptive Filter Algorithms , Spectral Kurtosis and Linear Prediction in Detection of a Naturally Degraded Bearing in a Gearbox, Vol. 14, pp 623–636, 2014.
<https://doi.org/10.1007/s11668-014-9857-8>
- [151] **L. R. Rabiner, R. W. Schafer**, *Theory and Applications of Digital Speech Processing*. Upper Saddle River, NJ: Pearson, 2010.
- [152] **K. A. Loparo**, *Bearings Vibration Data Sets*, Case Western Reserve University, 2012.
<http://csegroups.case.edu/bearingdatacenter/home>.
- [153] **H. Akaike**, A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, Vol. 19, no. 6, pp 716–723, 1974.
<https://doi.org/10.1109/TAC.1974.1100705>
- [154] **P. McKenzie, M. Alder**, Selecting the optimal number of components for a Gaussian mixture model, *Proc. 1994 IEEE Int. Symp. Inf. Theory*, pp 393-, 1994.
<https://doi.org/10.1109/ISIT.1994.394626>
- [155] **Y. ATMANI, S. RECHAK, A. MESLOUB, L. HEMMOUCHE**, Enhancement in Bearing Fault Classification Parameters Using Gaussian Mixture Models and Mel Frequency Cepstral Coefficients Features, *ARCHIVES OF ACOUSTICS*, Vol. 45, N°2, pp. 283- 295, 2020.
<https://doi.org/10.24425/aoa.2020.133149>
- [156] **X. Huang, A. Acero, H. Hon**. *Spoken Language Processing: A guide to theory, algorithm, and system development*. Prentice Hall, 2001.

ANNEXE

NOMENCLATURE DES FICHIERS DE LA BASE DE DONNEES DES ROULEMENTS

Les données utilisées pour la validation de cette étude ont été extraites du site public du Bearing Data Center de l'université américaine CWRU [152]. La nomenclature des fichiers (enregistrements) correspondants aux conditions des défauts induits aux roulements durant la campagne des essais, est répartie dans les tableaux ci-après.

Tableau A.1 : Données relatives à l'état normal (sain) ; $f_s = 48$ kHz ;

Charge moteur Cv	Vitesse tr/min	Nom du fichier de données
0	1797	97
1	1772	98
2	1750	99
3	1730	100

Tableau A.2 : Données relatives au défaut de roulement côté entraînement (DE) ; $f_s = 12$ kHz ;

Dimension du défaut en pouces (mm)	Charge moteur Cv	Vitesse tr/min	Nom du fichier de données pour chaque type de défaut				
			BPFI	BSF	BPFO centré	BPFO Orthogonal	BPFO Opposé
0.007 (0.18)	0	1797	105	118	130	144	156
	1	1772	106	119	131	145	158
	2	1750	107	120	132	146	159
	3	1730	108	121	133	147	160
0.014 (0.36)	0	1797	169	185	197	-	-
	1	1772	170	186	198	-	-
	2	1750	171	187	199	-	-
	3	1730	172	188	200	-	-
0.021 (0.53)	0	1797	209	222	234	246	258
	1	1772	210	223	235	247	259
	2	1750	211	224	236	248	260
	3	1730	212	225	237	249	261
0.028 (0.71)	0	1797	3001	3005	-	-	-
	1	1772	3002	3006	-	-	-
	2	1750	3003	3007	-	-	-
	3	1730	3004	3008	-	-	-

Tableau A.3 : Données relatives au défaut de roulement côté ventilateur (FE) ; $f_s = 12$ kHz ;

Dimension du défaut en pouces (mm)	Charge moteur Cv	Vitesse tr/min	Nom du fichier de données pour chaque type de défaut				
			BPFI	BSF	BPFO centré	BPFO Orthogonal	BPFO Opposé
0.007 (0.18)	0	1797	278	282	294	298	302
	1	1772	279	283	295	299	305
	2	1750	280	284	296	300	306
	3	1730	281	285	297	301	307
0.014 (0.36)	0	1797	274	286	313	310	-
	1	1772	275	287	-	309	-
	2	1750	276	288	-	311	-
	3	1730	277	289	-	312	-
0.021 (0.53)	0	1797	270	290	315	-	-
	1	1772	271	291	-	316	-
	2	1750	272	292	-	317	-
	3	1730	273	293	-	318	-

Tableau A.4 : Données relatives au défaut de roulement côté entraînement (DE) ; $f_s = 48$ kHz ;

Dimension du défaut en pouces (mm)	Charge moteur Cv	Vitesse tr/min	Nom du fichier de données pour chaque type de défaut				
			BPFI	BSF	BPFO centré	BPFO Orthogonal	BPFO Opposé
0.007 (0.18)	0	1797	109	122	135	148	161
	1	1772	110	123	136	149	162
	2	1750	111	124	137	150	163
	3	1730	112	125	138	151	164
0.014 (0.36)	0	1797	174	189	201	-	-
	1	1772	175	190	202	-	-
	2	1750	176	191	203	-	-
	3	1730	177	192	204	-	-
0.021 (0.53)	0	1797	213	226	238	250	262
	1	1772	214	227	239	251	263
	2	1750	215	228	240	252	264
	3	1730	217	229	241	253	265