

à la probabilité que le cardinal de cette intersection (le nombre de faux rejets) soit supérieur ou égal à 1 (A.2).

$$FWER(R) = \mathbb{P}(\text{card}(R \cap \mathbf{H}_0) \geq 1) = \mathbb{P}(V(R) \geq 1) = \mathbb{P}\left(\inf_{i \in \mathbf{H}_0} \{p_i(X)\} \leq t_{proc}\right) \quad (\text{A.2})$$

Le seuil de décision  $t$  est équivalent au rejet effectué à partir de ce seuil, donc  $V(R) = V(t)$ .

Ainsi en contrôlant cette probabilité à un niveau  $\alpha$ , nous nous assurons de commettre 0 fausse détection avec probabilité  $1 - \alpha$ . La procédure de Bonferroni [59] consiste à définir le seuil  $t$  de détection par  $\alpha/m$  avec  $\alpha$  le niveau souhaité.

$$t_{Bonferroni} = \frac{\alpha}{m}$$

La procédure de Bonferroni permet de contrôler le FWER à un niveau  $\alpha$ .

$$R_{Bonferroni} = \{i \in \{1, \dots, m\} : p_i(X) \leq t_{Bonferroni}\}$$

$$\begin{aligned} FWER(R_{Bonferroni}) &= \mathbb{P}\left(\bigcup_{i \in \mathbf{H}_0} \{p_i(X) \leq t_{Bonferroni}\}\right) \\ &\leq \sum_{i=1}^m \mathbb{1}\{i \in \mathbf{H}_0\} \mathbb{P}(p_i(X) \leq t_{Bonferroni}) = \frac{m_0}{m} \alpha \leq \alpha \end{aligned}$$

Plus le nombre de tests est important, plus le seuil défini sur les p-valeurs sera très faible. Cette procédure est très conservatrice, elle a donc fortement tendance à éviter de rejeter l'hypothèse  $\mathcal{H}_0$ .

Pour relaxer cette forte contrainte il est possible de définir le k-FWER [71] permettant de contrôler les tests multiples en autorisant avec probabilité  $1 - \alpha$ , un nombre de fausses détections d'au plus  $k - 1$ .

$$k - FWER = \mathbb{P}(\text{card}(R \cap \mathbf{H}_0) \geq k) = \mathbb{P}(V(R) \geq k)$$

Dans un cadre indépendant, la majoration de la probabilité de l'union par la somme des probabilités est pertinente, cependant si les tests sont fortement dépendants cette majoration est très forte entraînant un niveau réel nettement inférieur à  $\alpha$ , dans ce cas il est possible de contrôler le FWER comme un quantile de niveau  $\alpha$  de la variable aléatoire  $\inf_{i \in \mathbf{H}_0} \{p_i(X)\}$  [95]. Une procédure simple permettant de contrôler cette grandeur est de définir le seuil sur une base de données de validation comme étant la plus petite p-valeur dont l'hypothèse  $\mathcal{H}_0^i$  est vraie. Cependant dans notre cas d'étude, cette approche n'est pas pertinente due aux différentes classes des points sous l'hypothèse  $\mathcal{H}_0$  de normalité. Les p-valeurs des points normaux de la classe 4 (c'est-à-dire les points décalés dans les différents spectrogrammes) ont le même ordre de grandeur que les signatures inusuelles. Le seuil défini serait alors trop faible et entraînerait le non-rejet de l'hypothèse  $\mathcal{H}_0$  pour de nombreux points inusuels. Afin de définir un seuil de cette manière,

il est plus pertinent de le fixer uniquement sur les points des classes 1 (normaux) et 3 (bruit), augmentant ainsi la détection au niveau de la classe 4.

La méthode de Romano-Wolf [95] permet de contrôler le FWER à un niveau  $\alpha$  dans un cadre dépendant en appliquant la procédure de Bonferroni de manière itérative sur des sous-ensembles de tests. A la première itération, la procédure de Bonferroni est appliquée sur l'ensemble des tests, le sous-ensemble récupéré  $R_1^c$  (complémentaire des points rejetés) correspond à l'ensemble des tests qui n'ont pas rejetés l'hypothèse  $\mathcal{H}_0$ , sur lequel on applique une nouvelle fois la procédure de Bonferroni en considérant uniquement ces tests. La procédure continue rejetant de nouveaux tests  $R_k$  jusqu'à l'itération  $\hat{k}$  où les tests rejetant l'hypothèse  $\mathcal{H}_0$  soient les mêmes sur deux itérations successives.

$$R_k = \left\{ i \in \{1, \dots, m\} : p_i(X) \leq t_{Bonferroni}^{R_{k-1}^c} \right\}$$

$$\hat{k} = \min \{k \geq 1 : R_{k+1} = R_k\}$$

avec  $t_{Bonferroni}^{R_{k-1}^c}$  le seuil de la procédure de Bonferroni défini à partir des tests dans  $R_{k-1}^c$ .

Il existe également des procédures permettant de contrôler le FWER lorsque les dépendances entre les différents tests sont connues [50].

## False discovery rate (FDR)

Le FWER contrôle le nombre de fausses détections effectuées par les  $m$  tests, la grandeur ne tient pas compte du nombre entier d'hypothèses  $\mathcal{H}_0^i$  rejetées par la procédure.  $V(R) = 1$  hypothèse rejetée à tort parmi  $R = 10$  hypothèses rejetées est plus grave que parmi  $R = 100$  hypothèses rejetées. Pour le FWER, il n'y a pas de différence entre ces deux résultats. Le False discovery proportion (FDP) prend en compte cette différence en considérant comme grandeur pour contrôler les tests multiples le ratio entre le nombre de rejets d'hypothèses  $\mathcal{H}_0^i$  à tort  $V(R)$  et le nombre total  $R$  de rejets d'hypothèse  $\mathcal{H}_0^i$ .

$$FDP(R) = \frac{\text{card}(R \cap \mathbf{H}_0)}{\max(R, 1)} = \frac{V(R)}{\max(R, 1)}$$

Le FDP est une variable aléatoire et ne permet donc pas de définir un niveau pour contrôler les tests multiples, le FDR (A.3) est la grandeur issue du FDP permettant de contrôler les tests multiples à un niveau  $\alpha$ . Elle correspond à une tolérance moyenne de fausses détections parmi l'ensemble des détections.

$$FDR(R) = \mathbb{E}(FDP(R)) = \begin{cases} \mathbb{E}\left(\frac{V(R)}{R}\right) & \text{si } R > 0 \\ 0 & \text{si } R = 0 \end{cases} \quad (\text{A.3})$$

La procédure de BH [14] définie par le seuil (A.4) permet de contrôler le FDR à un niveau  $\alpha$

dans le cas où les tests sont indépendants ou positivement dépendants.

$$t_{BH} = \frac{\alpha \hat{l}}{m} \text{ avec } \hat{l} = \max \left\{ l \in \{0, 1, \dots, m\} : p_{(l)} \leq \frac{\alpha l}{m} \right\} \quad (\text{A.4})$$

avec  $p_{(l)}$  les p-valeurs  $(p_1, \dots, p_m)$  ordonnées par ordre croissant.

$$R_{BH} = \{i \in \{1, \dots, m\} : p_i(X) \leq t_{BH}\} \Leftrightarrow \{p_{(1)}, \dots, p_{(\hat{l})}\}$$

$$FDR(R_{BH}) \leq \frac{m_0}{m} \alpha \leq \alpha$$

Cette procédure consiste à définir comme seuil le dernier point d'intersection entre la courbe des p-valeurs ordonnées et la droite d'équation  $l \mapsto \alpha l/m$ . Cette procédure est invariante par échelle lorsque le nombre de tests  $m$  croît. La procédure BH peut être considérée comme une procédure "step-up" car le seuil est défini comme le point maximal d'intersection, il est également possible de prendre le point d'intersection minimal (le premier point où les 2 courbes se croisent) ou un point intermédiaire. Il existe d'autres procédures permettant de contrôler le FDR [96].

Dans le cadre de nos données de spectrogrammes, les différents points sur les patches peuvent être considérés comme indépendants ou comme étant positivement dépendants. Une intensité vibratoire à régime et fréquence donnés n'influent pas négativement avec les autres points du voisinage, mais peut impacter positivement les valeurs du voisinage, des points éloignés dans l'espace régime/fréquence peuvent être considérés comme indépendants.

Le fait de contrôler le FDR à un niveau  $\alpha$  ne signifie pas que le FDP est lui aussi contrôlé à ce même niveau, cela est uniquement le cas asymptotiquement lorsque les dépendances entre les tests sont faibles. Il existe donc des procédures visant à contrôler le FDP [94] via les grandeurs suivantes :

$$\mathbb{P}(FDP(R) > \alpha) \leq \zeta,$$

$$\limsup_m \mathbb{P}(FDP(R) > \alpha) \leq \zeta.$$

Une approche différente proposée dans [106] considère une estimation du FDR à partir de la région de rejet, le FDR n'est plus défini à partir du seuil de décision à partir duquel la région de rejet est calculé, la région de rejet est prédéfinie et le niveau du FDR est calculé en conséquence. Pour cela différentes régions de rejet sont testées sur les p-valeurs ordonnées, et une estimation du FDR est apportée pour chacune d'entre elles, il s'agit alors de sélectionner la région apportant une estimation convenable du FDR. Cette approche permet d'estimer une variation du FDR, le positive false discovery rate (pFDR).



**Titre :** Représentations pour la détection d'anomalies — Application aux données vibratoires des moteurs d'avions

**Mots clés :** apprentissage de dictionnaire, curvelets, estimation de densité par noyau, détection de nouveautés, détection d'anomalies, vibrations

**Résumé :** Les mesures de vibrations sont l'une des données les plus pertinentes pour détecter des anomalies sur les moteurs. Les vibrations sont acquises sur banc d'essai en phase d'accélération et de décélération pour assurer la fiabilité du moteur à la sortie de la chaîne de production. Ces données temporelles sont converties en spectrogrammes pour permettre aux experts d'effectuer une analyse visuelle de ces données et de détecter les différentes signatures atypiques. Les sources vibratoires correspondent à des raies sur les spectrogrammes. Dans cette thèse, nous avons mis en place un outil d'aide à la décision automatique pour analyser les spectrogrammes et détecter tout type de signatures atypiques, ces signatures ne proviennent pas nécessairement d'un endommagement du moteur. En premier lieu, nous avons construit une base de données numérique de spectro-

grammes annotés. Il est important de noter que les signatures inusuelles sont variables en forme, intensité et position et se trouvent dans un faible nombre de données. Par conséquent, pour détecter ces signatures, nous caractérisons les comportements normaux des spectrogrammes, de manière analogue aux méthodes de détection de nouveautés, en représentant les patches des spectrogrammes sur des dictionnaires comme les curvelets et la Non-negative matrix factorization (NMF), ainsi qu'en estimant la distribution de chaque point du spectrogramme à partir de données normales dépendamment ou non de leur voisinage. La détection des points atypiques est réalisée par comparaison des données tests au modèle de normalité estimé sur des données d'apprentissage normales. La détection des points atypiques permet la détection des signatures inusuelles composées par ces points.

**Title :** Representations for anomaly detection — Application to aircraft engines' vibration data

**Keywords :** dictionary learning, curvelets, kernel density estimation, novelty detection, anomaly detection, vibrations

**Abstract :** Vibration measurements are one of the most relevant data for detecting anomalies in engines. Vibrations are recorded on a test bench during acceleration and deceleration phases to ensure the reliability of every flight engine at the end of the production line. These temporal signals are converted into spectrograms for experts to perform visual analysis of these data and detect any unusual signature. Vibratory signatures correspond to lines on the spectrograms. In this thesis, we have developed a decision support system to automatically analyze these spectrograms and detect any type of unusual signatures, these signatures are not necessarily originated from a damage in the engine. Firstly, we have built a numerical spectrograms database with annota-

ted zones, it is important to note that data containing these unusual signatures are sparse and that these signatures are quite variable in shape, intensity and position. Consequently, to detect them, like in the novelty detection process, we characterize the normal behavior of the spectrograms by representing patches of the spectrograms in dictionaries such as the curvelets and the Non-negative matrix factorization (NMF) and by estimating the distribution of every points of the spectrograms with normal data depending or not of the neighborhood. The detection of the unusual points is performed by comparing test data to the model of normality estimated on learning normal data. The detection of the unusual points allows the detection of the unusual signatures composed by these points.