

définissons un seuil de détection s sur les p-valeurs identique pour les deux méthodes avec :

$$s = 0.07$$

Ce seuil permet la détection d'environ 85% des points atypiques de la base de validation, malgré une détection d'environ 95% des points décalés provenant des signatures vibratoires liées au N_1 . Les points normaux et le bruit sont faiblement détectés. Notre approche a pour but de mettre en évidence les signatures inusuelles, pour cela le seuil de détection ne doit pas entraîner un grand nombre de fausses détections pour éviter de masquer les signatures inusuelles sur le résultat final.

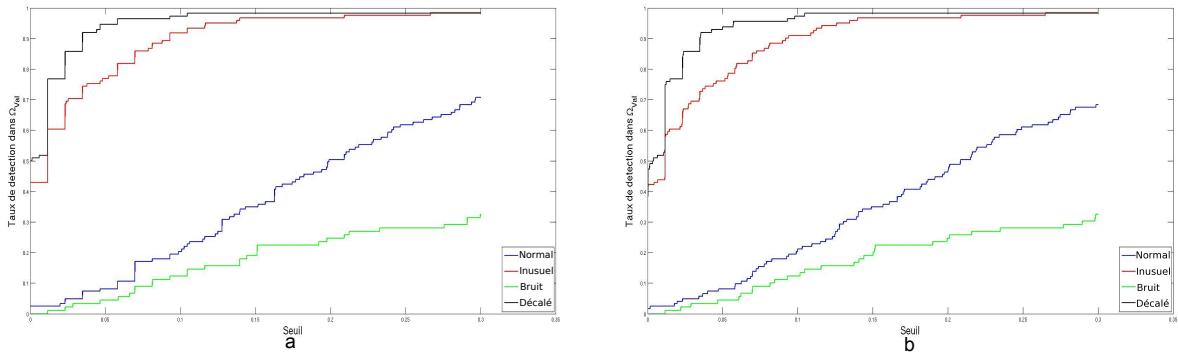


FIGURE 5.11 – Taux de détection en fonction des seuils de détection des différentes classes de points de la base de validation Ω_{Val} à partir des modèles de normalité définis par l'estimation de la distribution avec noyau gaussien en estimant l'échelle par la règle du pouce (a) et par le maximum de vraisemblance (b). Les résultats issus des 2 modèles de normalité sont similaires. Ils montrent un fort taux de bonnes detections des points inusuels pour des petites valeurs de seuils, les points normaux décalés sont également fortement détectés. Les points normaux liés aux raies N_2 et le bruit sont faiblement détectés pour ces valeurs de seuils, nous sélectionnons donc un seuil de détection au niveau de ces dernières.

5.4.4 Résultats sur la base de test Ω_{Test}

Comme dans la partie précédente, les résultats sont présentés sous 3 formats :

- à partir des points annotés de la base de test Ω_{Test} de \mathcal{B}_1 ,
- à partir du taux de points détectés appartenant aux zones atypiques extraites sur tous les patchs,
- à partir d'une visualisation des points détectés sur le patch.

Résultats de classification des points annotés de Ω_{Test}

Le tableau 5.3 fournit les taux de détection des différentes classes de points des patchs dans Ω_{Test} issus de la base \mathcal{B}_1 sans application du filtre pour différentes données d'apprentissage de la distribution. Le tableau 5.4 fournit les mêmes résultats avec l'application du filtre de voisinage.

TABLE 5.3 – Taux de détection à partir de la densité estimée par noyau gaussien à partir de différentes données d'apprentissage pour les différentes classes de points des données dans Ω_{Test} et les différentes méthodes de définition de l'échelle du noyau

Classe de points	Taux de détection - règle du pouce	Taux de détection - ML
Classe "normal"	$13.3\% \pm 1.1\%$	$13.3\% \pm 1.1\%$
Classe "inusuel"	$85.1\% \pm 1.3\%$	$83.3\% \pm 1\%$
Classe "bruit"	$3.12\% \pm 0\%$	$3.12\% \pm 0\%$
Classe "décalé"	100%	$99.5\% \pm 1.1\%$

TABLE 5.4 – Taux de détection à partir de la densité estimée par noyau gaussien à partir de différentes données d'apprentissage avec application du filtrage de voisinage pour les différentes classes de points des données dans Ω_{Test} et les différentes méthodes de définition de l'échelle du noyau

Classe de points	Taux de détection - règle du pouce	Taux de détection - ML
Classe "normal"	$10.3\% \pm 0\%$	$10.3\% \pm 0\%$
Classe "inusuel"	$75.4\% \pm 1.3\%$	$74.4\% \pm 0\%$
Classe "bruit"	$3.12\% \pm 0\%$	$3.12\% \pm 0\%$
Classe "décalé"	$91.8\% \pm 1.1\%$	$89.7\% \pm 1.8\%$

La méthode permet donc un bon niveau de détection des points atypiques mais également des points décalés, correspondant alors à de fausses détections. Les signatures liées au N_1 sont considérées comme plus atypiques que les points réellement inusuels. Cela provient du fait que ces raies sont d'intensités plus importantes que les signatures inusuelles, et leurs décalages entraînent leurs détections du fait de l'étude ponctuelle. Les points normaux et le bruit sont faiblement détectés conformément aux données de validation. Le filtrage réduit les fausses détections (principalement au niveau des points décalés) ainsi que la détection des points atypiques d'environ 10%. Ces résultats sont à combiner avec les détections visuelles des signatures inusuelles à partir de nos approches car ils correspondent à la détection de quelques points sélectionnés sur le patch uniquement. La détection visuelle permet d'avoir une appréciation du modèle sur l'ensemble du patch et de l'effet du filtrage par voisinage. Les modèles de normalité définis à partir de la règle du pouce et du maximum de vraisemblance sont très proches. Pour cette raison, dans la suite, nous présentons les résultats à partir de la règle du pouce uniquement.

Résultats de détection sur l'ensemble des patchs \mathcal{K}^{128} des spectrogrammes

Nous donnons les résultats de détection sur les différents patchs. Pour cela comme dans la partie précédente, nous calculons l'erreur E_j^i (3.15) correspondant à la proportion de points détectés appartenant à une zone atypique sur un patch contenant des signatures inusuelles et la proportion des points détectés sur les patchs ne contenant pas de zones atypiques. Le tableau 5.6 donne les résultats pour les deux modèles de normalité sur les différents patchs, il est subdivisé en sous-tableaux dont la structure est défini dans le tableau 5.5.

Plusieurs patchs possèdent un pourcentage important de points détectés faisant partie d'une

TABLE 5.5 – Structure des sous-tableaux de résultats pour le modèle défini par la distribution de chaque point estimé par noyau gaussien

Taux de bonnes détections du modèle de normalité défini par la règle du pouce pour les patchs contenant des signatures inusuelles	Taux de points détectés du modèle de normalité défini par la règle du pouce pour les patchs sans signatures inusuelles
Taux de bonnes détections du modèle de normalité défini par la règle du pouce avec filtrage du voisinage pour les patchs contenant des signatures inusuelles	Taux de points détectés du modèle de normalité défini par la règle du pouce avec filtrage pour les patchs sans signatures inusuelles

zone atypique tandis que d'autres possèdent un faible pourcentage. Cette variabilité provient de l'extraction des zones atypiques contenant plusieurs points normaux, ainsi l'intersection de certains patchs avec les zones atypiques ne contient pas de points inusuels. Les forts taux de détection sur les patchs contenant des signatures inusuelles signifient donc que parmi l'ensemble de la base de test sur ce patch, les points détectés se trouvent principalement aux alentours des signatures inusuelles permettant ainsi de les détecter. La proportion de points détectés dans les patchs sans signature atypique est faible comparativement à la taille du patch. La méthode est donc très intéressante pour détecter les points inusuels sur les spectrogrammes.

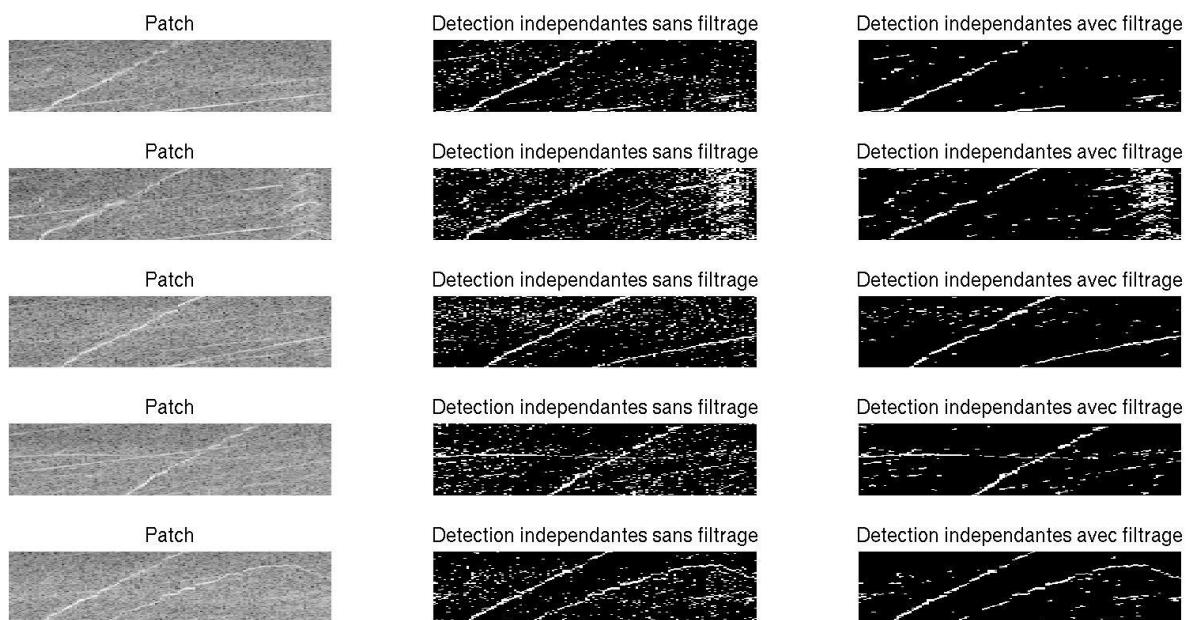


FIGURE 5.12 – Détection des points inusuels sur plusieurs patchs à partir du modèle de normalité définie par la distribution estimée par noyau avec (3ème colonne) et sans (2ème colonne) filtrage de voisinage. Le premier patch (1ère ligne) correspond à un patch sans signature inusuelle, les autres contiennent différentes signatures inusuelles.

TABLE 5.6 – Taux de détection sur l’ensemble des patchs de la base de test à partir du modèle de normalité défini par l’estimation de la distribution des points par noyau avec et sans filtrage par le voisinage.

<i>NA</i>	7.33% \pm 0.09%	<i>NA</i>	7.47% \pm 0.03%	0% \pm 0%	6.44% \pm 0.05%
<i>NA</i>	6.37% \pm 0.08%	<i>NA</i>	5.76% \pm 0.02%	0% \pm 0%	4.29 \pm 0.04%
<i>NA</i>	9.32% \pm 0.15%	<i>NA</i>	9.63% \pm 0.08%	73.3% \pm 0.5%	7.44% \pm 0.06%
<i>NA</i>	8.29% \pm 0.14%	<i>NA</i>	7.21% \pm 0.06%	72.5% \pm 0.9%	5.02% \pm 0.04%
23.1% \pm 0.1%	9.09% \pm 0.12%	3.1% \pm 0.2%	10.3% \pm 0.01%	22.5% \pm 0.4%	7.64% \pm 0.06%
23.7% \pm 0.4%	7.88% \pm 0.12%	4.3% \pm 0.3%	7.34% \pm 0.07%	21.0% \pm 0.4%	5.09% \pm 0.06
43.1% \pm 0.3%	8.1% \pm 0.09%	27.6% \pm 0.3%	9.48% \pm 0.08%	34.1% \pm 0.4%	7.37% \pm 0.15%
45.2% \pm 0.4%	7.0% \pm 0.08%	27.5% \pm 0.3%	6.35% \pm 0.06%	32.5% \pm 0.6%	3.84% \pm 0.11%
30.5% \pm 0.1%	9.75% \pm 0.11%	32.9% \pm 0.1%	12.5% \pm 0.26%	41.8% \pm 0.3%	10.35% \pm 0.1%
31.4% \pm 0.2%	8.69% \pm 0.09%	32.1% \pm 0.3%	9.03 \pm 0.23%	42.4% \pm 0.4%	6.15% \pm 0.08%
55.4% \pm 0.1%	8.41% \pm 0.11%	50.5% \pm 0.1%	9.4% \pm 0.19%	75.3% \pm 0.1%	8.3% \pm 0.31%
55.2% \pm 0.1%	7.25% \pm 0.1%	47.6% \pm 0.2%	5.64% \pm 0.15%	79.0% \pm 0.3%	4.75% \pm 0.2%
59.3% \pm 0.3%	7.1% \pm 0.15%	49.3% \pm 0.3%	6.97% \pm 0.04	60.4% \pm 0.1%	6.49% \pm 0.11%
59.3% \pm 0.4%	6.1% \pm 0.13%	46.3% \pm 0.5%	3.77% \pm 0.04%	64.3% \pm 0.2%	3.3% \pm 0.06%
28.2% \pm 2.3%	5.25% \pm 0.14%	45.3% \pm 0.1%	5.59% \pm 0.22%	74.2% \pm 0.2%	5.47% \pm 0.1%
25.4% \pm 0.6%	4.53% \pm 0.13%	48.0% \pm 0.7%	3.39% \pm 0.17%	70.5% \pm 0.4%	2.62% \pm 0.07%
6.1% \pm 0.1%	7.66% \pm 0.16%	25.2% \pm 0.5%	8.03% \pm 0.14%	47.2% \pm 0.3%	7.59% \pm 0.1%
6.3% \pm 0%	6.72% \pm 0.15%	22.2% \pm 0.5%	5.13% \pm 0.11%	53.3% \pm 0.9%	4.11% \pm 0.08%
<i>NA</i>	7.13% \pm 0.33%	<i>NA</i>	7.95% \pm 0.03%	35.1% \pm 0.1%	8.69% \pm 0.24%
<i>NA</i>	6.25% \pm 0.03%	<i>NA</i>	5.2% \pm 0.01%	31.8 \pm 0.6%	5.18% \pm 0.02%
12.8% \pm 0.1%	6.7% \pm 0.29%	<i>NA</i>	6.55% \pm 0.09%	23.8% \pm 0.1%	7.86% \pm 0.27%
13.3% \pm 0.1%	5.79% \pm 0.27%	<i>NA</i>	4.06% \pm 0.07%	21.7% \pm 0.3%	4.76% \pm 0.22%
6.9% \pm 0.1%	7.72% \pm 0.26%	13.5% \pm 0.1%	8.0% \pm 0.15%	13.2% \pm 0.3%	9.05% \pm 0.14%
5.7% \pm 0.2%	6.79% \pm 0.26%	12.7% \pm 0.6%	5.02% \pm 0.11%	10.0% \pm 1%	5.36% \pm 0.1%
8.0% \pm 0.2%	7.24% \pm 0.16%	20.4% \pm 0.3%	7.47% \pm 0.17%	36.4% \pm 0.6%	7.83% \pm 0.14%
7.4% \pm 0.4%	6.2% \pm 0.15%	20.9% \pm 0.8%	4.7% \pm 0.13%	31.8% \pm 0.6%	4.45% \pm 0.11%
<i>NA</i>	6.94% \pm 0.21%	2.2% \pm 0.2%	7.34% \pm 0.04%	51.7% \pm 1.1%	7.48% \pm 0.4%
<i>NA</i>	5.92% \pm 0.19%	1.0% \pm 0.5%	4.44% \pm 0.03%	66.8% \pm 0.9%	4.13% \pm 0.28%
<i>NA</i>	7.48% \pm 0.12%	41.7% \pm 1.2%	9.88% \pm 0.26%	50.2% \pm 0.5%	11.7% \pm 0.42%
<i>NA</i>	6.71% \pm 0.11%	47.3% \pm 0.4%	7.82% \pm 0.21%	58.4% \pm 0.5%	8.7% \pm 0.36%
20.0% \pm 0.1%	7.74% \pm 0.23%	31.5% \pm 0.2%	9.59% \pm 0.17%	40.3% \pm 0.2%	9.3% \pm 0.17%
20.1% \pm 0.3%	7.06% \pm 0.22%	34.4% \pm 0.3%	7.35% \pm 0.16%	49.1% \pm 0.1%	6.21% \pm 0.13%
83.6% \pm 0.1%	7.51% \pm 0.07%	30.0% + 0.3%	7.56% \pm 0.1%	14.5% \pm 0.1%	7.73% \pm 0.28%
84.5% \pm 0.1%	6.57% \pm 0.06%	37.0% \pm 0.5%	5.19% \pm 0.07%	16.4% \pm 0.3%	4.82% \pm 0.21%
22.9% \pm 0.2%	8.02% \pm 0.16%	3.0% \pm 0.1%	7.32% \pm 0.26%	27.2% \pm 0.2%	7.59% \pm 0.27%
22.8% \pm 0.5%	7.06% \pm 0.14%	4.1% \pm 0.1%	4.72% \pm 0.2%	28.4% \pm 0.1%	4.52% \pm 0.21%

Résultats visuels de détection sur la base de test Ω_{Test}

La figure 5.12 présente les points détectés à partir du modèle de normalité défini à partir de la règle du pouce avec et sans filtrage sur un patchs spécifique issu de la subdivision \mathcal{K}^{128} pour différents spectrogrammes, un dont le patch spécifique est normal (1ère ligne) et 4 spectrogrammes dont les patchs spécifiques contiennent différentes signatures atypiques. Les signatures inusuelles, ainsi que plusieurs points normaux, sont détectées sur les patchs. L'application du filtrage permet de visuellement réduire considérablement le nombre de points faussement détectés tout en conservant les points inusuels.

Le modèle de normalité pour chaque point du spectrogramme est défini sans aucun apriori à partir d'une approche non-paramétrique d'estimation de la densité par un noyau gaussien. La normalité de chaque point est alors caractérisée par cette distribution apprise et vérifiée à partir d'un test d'adéquation de ce point par rapport à la distribution estimée. Le modèle proposé permet la détection des points atypiques mais est bien plus sensible à l'intensité et la position des différents points que ne l'étaient les méthodes basées sur les dictionnaires entraînant une détection plus importante des signatures normales décalées. L'application du filtrage de voisinage est pertinente et permet de réduire les fausses détections.

5.5 Modèle de normalité défini à partir de l'estimation de densité par noyau gamma

Le choix du noyau gaussien permettait d'effectuer les calculs de manière simple. Cependant bien que le noyau n'influence pas fortement la distribution, ce dernier donne une densité aux valeurs négatives tandis que les intensités vibratoires appartiennent à $[0, +\infty[$. L'utilisation du noyau gaussien entraîne donc un biais à l'origine entraînant une perte de la densité à droite. Cela peut entraîner la détection de points normaux. Il faut donc utiliser un noyau n'allouant pas de poids aux valeurs négatives. Aucun noyau symétrique ne respecte ce critère par définition. Les distributions des points des spectrogrammes possèdent des allures proches des lois gamma (Figure 5.2), l'utilisation du noyau gamma pour l'estimation de la densité semble donc pertinent respectant ainsi la positivité des points des spectrogrammes.

5.5.1 Le modèle de normalité

Le modèle de normalité issu du noyau gamma (5.13) est similaire à celui défini à partir du noyau gaussien (5.9). Il est défini par la distribution indépendante estimée par noyau gamma sur

les points normaux d'apprentissage Ω_{App}^{f,N_2} pour chaque point (f, N_2) des spectrogrammes.

$$\forall (f, N_2) \quad \hat{f}_{f, N_2}^{K_\Gamma}(x | \Omega_{App}^{f, N_2}, h) = \frac{1}{\text{card}(\Omega_{App}^{f, N_2})} \sum_{i \in \Omega_{App}^{f, N_2}} K_{\rho_h(x), h}(S_{f, N_2}^i) \quad (5.13)$$

Estimation de l'échelle du noyau

Le seul paramètre à déterminer afin de définir intégralement le modèle de normalité est l'échelle h du noyau gamma. Nous calibrons ce paramètre à partir des deux mêmes approches développées pour l'échelle du noyau gaussien. La première méthode de calibration de l'échelle du noyau est la règle du pouce définie de manière identique à la section précédente :

$$\hat{h}_{RP} = 1.06\hat{\sigma} \text{card}(\Omega_{App})^{-\frac{1}{5}},$$

avec $\hat{\sigma}$ l'écart-type estimée sur les données.

La seconde méthode de calibration de l'échelle du noyau correspond au maximum de vraisemblance par rapport à l'échelle h estimée sur les points normaux de la base de validation Ω_{Val}^{f, N_2} ,

$$\hat{h}_{max}^\Gamma = \underset{h}{\operatorname{argmax}} \left\{ \prod_{j \in \Omega_{Val}^{f, N_2}} \hat{f}_{f, N_2}^\Gamma(S_{f, N_2}^j | \Omega_{App}^{f, N_2}, h) \right\}.$$

La vraisemblance en fonction de différentes valeurs de h est présentée sur la figure 5.13 pour différents points des spectrogrammes. Le maximum de la vraisemblance est atteint en un unique point qui correspond à \hat{h}_{max}^Γ . La valeur de l'échelle h du noyau est donc différente pour chaque point du spectrogramme et dépendante des données en ces coordonnées justifiant la définition d'un modèle de normalité pour chaque points des spectrogrammes.

5.5.2 Le score de détection

Le score pour chaque point du spectrogramme est défini à partir d'un test opposant l'hypothèse \mathcal{H}_0^{f, N_2} de normalité du point de coordonnées (f, N_2) à l'hypothèse \mathcal{H}_1^{f, N_2} caractérisant sa singularité. Identiquement au cas gaussien (5.10), ce test correspond à un test d'adéquation à la distribution estimée (5.14). La p-valeur correspondante à chacun de ces tests définit le score de détection permettant de classifier un point comme normal ou atypique.

$$\begin{aligned} pval_{f, N_2}^{i, K_\Gamma} &= \mathbb{P}_{\mathcal{H}_0^{f, N_2}}(X > S_{f, N_2}^i) \approx \mathbb{P}_{f_{f, N_2}^{K_\Gamma}}(X > S_{f, N_2}^i) \\ &\approx \int_{S_{f, N_2}^i}^{+\infty} \hat{f}_{f, N_2}^{K_\Gamma}(x | \Omega_{App}^{f, N_2}, \hat{h}) dx = \frac{1}{\text{card}(\Omega_{App}^{f, N_2})} \sum_{j \in \Omega_{App}^{f, N_2}} \int_{S_{f, N_2}^i}^{+\infty} K_{\rho_{\hat{h}}(x), \hat{h}}(S_{f, N_2}^j) dx \end{aligned} \quad (5.14)$$

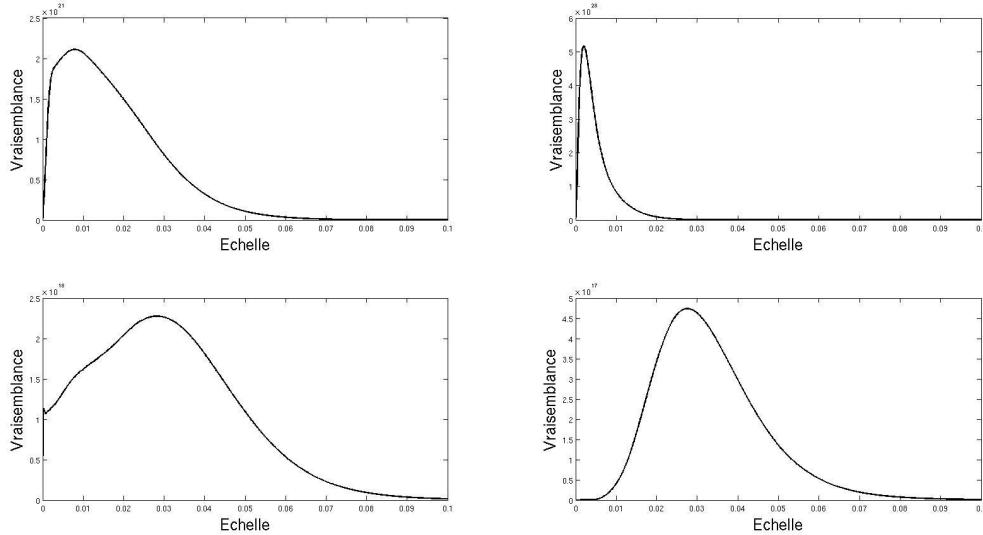


FIGURE 5.13 – Vraisemblance sur Ω_{Val} de la distribution estimée par noyau gamma en fonction de différentes valeurs de h et pour différents points du spectrogramme. Le maximum est atteint en un unique point définissant l'échelle optimale. La valeur du maximum de vraisemblance est différente selon le point étudié, la définition d'un modèle de normalité pour chaque point est donc pertinente.

Le noyau gaussien permettait de se servir de la symétrie afin de considérer le noyau comme la densité d'une gaussienne aussi bien centrée en S_{f,N_2}^j qu'en x . Cela permettait de faciliter le calcul de la p-valeur et de la décomposer comme la moyenne des p-valeurs de chaque noyau centré en la variable S_{f,N_2}^j . Le noyau gamma ne possède pas cette symétrie, entraînant un calcul d'intégrale ne possédant pas de forme fermée. Il n'est pas possible de calculer cette p-valeur par décomposition de la densité estimée en fonction de ses noyaux.

Pour calculer cette p-valeur, nous utilisons la positivité du domaine de définition et estimons $F_{\hat{f}_{f,N_2}^{K_\Gamma}}(S_{f,N_2}^i) = \mathbb{P}_{\hat{f}_{f,N_2}^{K_\Gamma}}(X \leq S_{f,N_2}^i)$, la fonction de répartition de la densité estimée, plutôt que $\mathbb{P}_{\hat{f}_{f,N_2}^{K_\Gamma}}(X > S_{f,N_2}^i)$. Nous utilisons l'approximation de l'intégrale par la méthode des trapèzes (5.15) avec un pas de discrétisation $\xi = 10^{-4}$ équivalant à celui des intensités des spectrogrammes. La densité est ainsi calculée pour toutes les valeurs possibles des spectrogrammes. Nous définissons alors la suite $(x_k)_{k \in \mathbb{N}}$ comme la discrétisation de \mathbb{R}_+ par pas de ξ ($x_{k+1} = x_k + \xi$) sur laquelle la fonction de répartition est estimée.

La méthode des trapèzes ne donne pas une intégrale unitaire sur \mathbb{R}_+ . Afin d'obtenir des p-valeurs bien comprises entre 0 et 1, nous normalisons l'estimation de la fonction de répartition par M , l'estimation de l'intégrale de la densité par la méthode des trapèzes appliquée à l'ensemble de l'espace.

$$\mathbb{P}_{\hat{f}_{f,N_2}^{K_\Gamma}}(X \leq S_{f,N_2}^i) \approx \frac{1}{M} \sum_{k=1}^{k:x_k=S_{f,N_2}^i} \frac{\hat{f}_{f,N_2}^{K_\Gamma}(x_{k+1} | \Omega_{App}^{f,N_2}, \hat{h}) - \hat{f}_{f,N_2}^{K_\Gamma}(x_k | \Omega_{App}^{f,N_2}, \hat{h})}{2} \xi \quad (5.15)$$

avec

$$M = \sum_{k=1}^{+\infty} \frac{\hat{f}_{f,N_2}^{K_\Gamma}(x_{k+1}|\Omega_{App}^{f,N_2}, \hat{h}) - \hat{f}_{f,N_2}^{K_\Gamma}(x_k|\Omega_{App}^{f,N_2}, \hat{h})}{2} \xi$$

Cette approche de définition du modèle de normalité par le noyau gamma possède un temps de calcul 10 fois plus important que pour le modèle défini à partir du noyau gaussien. Dans le cas gaussien, la simplification des calculs permettait d'accélérer le processus et d'obtenir des résultats dans des temps convenables, cette approche ne permet pas de simplification et entraîne plusieurs estimations pour chaque point du spectrogramme.

Nous avons alors :

$$pval_{f,N_2}^{i,K_\Gamma} \approx \mathbb{P}_{\hat{f}_{f,N_2}^{\Gamma}}(X > S_{f,N_2}^i) \approx 1 - \mathbb{P}_{\hat{f}_{f,N_2}^{\Gamma}}(X \leq S_{f,N_2}^i).$$

La décision de singularité des points s'effectue par la comparaison des p-valeurs à un seuil de détection s (5.16) identique pour tous les points (par manque d'annotations sur tous les points du spectrogramme) avec une application du filtrage de voisinage \mathcal{V}_{f,N_2} identique à (5.12).

$$\tilde{Y}_{f,N_2}^{i,K_\Gamma} = \mathbb{1}\{pval_{f,N_2}^{i,K_\Gamma} \leq s\} \quad (5.16)$$

5.5.3 Calibration des seuils de détection

La figure 5.14 présente les taux des détections à partir des modèles définis par l'estimation de la distribution par noyau gamma des différentes classes de points des données de validation Ω_{Val} de la base \mathcal{B}_1 en fonction de différents seuils de détection. La calibration a été effectuée pour les 2 méthodes de définition de l'échelle.

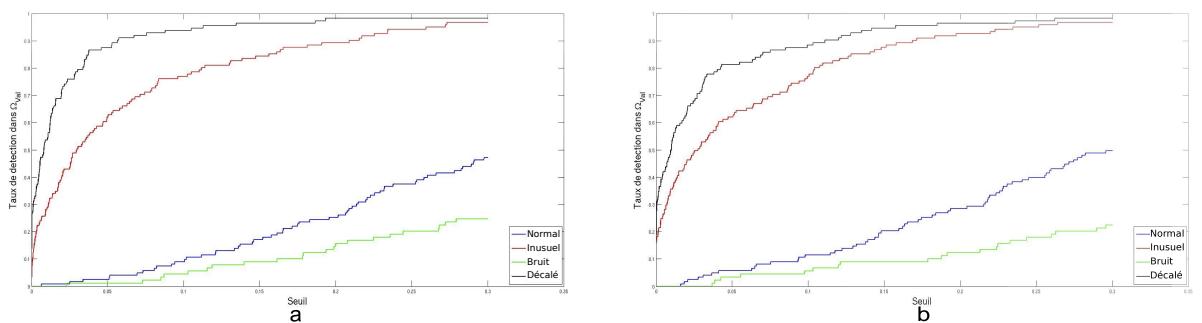


FIGURE 5.14 – Taux de détection en fonction du seuil de détection des différentes classes de points de la base de validation Ω_{Val} à partir des modèles de normalité définis par l'estimation de la distribution avec noyau gamma en estimant l'échelle par la règle du pouce (a) et par le maximum de vraisemblance (b). Les résultats entre les deux modèles sont proches. Les points atypiques (rouge) et les points normaux décalés (noir) sont fortement détectés pour de petites valeurs de seuil tandis que les points normaux (bleu) et le bruit (vert) le sont faiblement. Le seuil de détection est donc défini au niveau de ces valeurs.

Les deux méthodes de définition d'échelle donnent des résultats relativement proches. Les pentes des détections des points normaux et du bruit (toutes deux concernant des points normaux) sont plus faibles que pour le cas gaussien (il en est de même pour les points inusuels et les points décalés). Nous prenons donc un seuil de détection plus important pour le modèle défini dans cette section avec $s = 0.15$. Sur la base de validation ce seuil permet un bon niveau de détection des points inusuels d'environ 80% mais également des points décalés comme les méthodes précédentes.

5.5.4 Résultats sur la base de test Ω_{Test}

Les temps de calcul des modèles définis à partir du noyau gamma sont trop longs. Nous présentons uniquement les résultats sur les données tests Ω_{Test} du patch spécifique de la base \mathcal{B}_1 sous la forme du taux de détection des points annotés et de la représentation visuelle des points détectés sur l'intégralité du patch. Nous ne donnons pas les résultats sur l'ensemble des patchs à partir de l'erreur E_j^i .

Résultats de classification des points annotés de Ω_{Test}

Le tableau 5.7 présente les taux de détections des différentes classes de points de la base \mathcal{B}_1 à partir du modèle de normalité défini par le noyau gamma pour les deux méthodes d'estimation de l'échelle. Le modèle est défini pour plusieurs ensembles d'apprentissage. Le noyau gamma apporte une bonne caractérisation des distributions ponctuelles des points entraînant une très bonne détection des points atypique de plus de 90%. Les points décalés sont également fortement détectés. Plusieurs points normaux sont également détectés à tort par l'approche, cette fausse détection peut être réduite à partir du filtrage de voisinage. Les effets du filtrage de voisinage sont observables sur la représentation visuelle sur l'ensemble du patch des résultats de détection (Figure 5.15). Les 2 méthodes de définition d'échelle donnent des résultats proches, la définition du seuil à partir du maximum de vraisemblance est légèrement supérieure sur les points annotés. Ces résultats de détection sont légèrement supérieurs au modèle de normalité défini à partir du noyau gaussien, cependant ce dernier est 10 fois plus rapide que le modèle défini à partir du noyau gamma tout en donnant des résultats proches.

TABLE 5.7 – Taux de détection à partir de la densité estimée par noyau gamma à partir de différentes données d'apprentissage pour les différentes classes de points des données dans Ω_{Test} et les différentes méthodes de définition de l'échelle du noyau

Classe de points	Taux de détection - règle du pouce	Taux de détection - ML
Classe "normal"	$18.5\% \pm 1.1\%$	$25.1\% \pm 1.1\%$
Classe "inusuel"	$90.7\% \pm 0\%$	$96.3\% \pm 1.3\%$
Classe "bruit"	$6.25\% \pm 0\%$	$6.25\% \pm 0\%$
Classe "décalé"	$98.5\% \pm 1.4\%$	$97.9\% \pm 1.1\%$

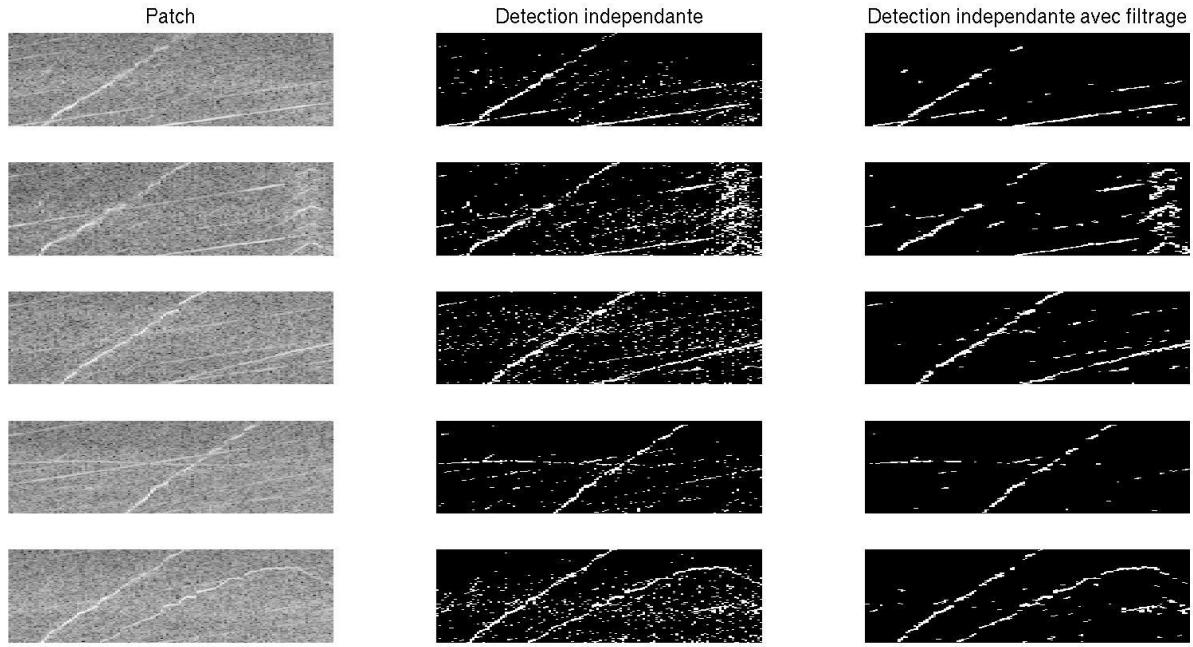


FIGURE 5.15 – Détection des points atypiques sur plusieurs patchs avec et sans (1ère ligne) signature atypique à partir du modèle de normalité définie par la distribution estimée par noyau gamma avec (3ème colonne) et sans (2ème colonne) filtrage de voisinage. Les signatures atypiques sont mises en évidence sur les patchs. Les signatures liées au N_1 et au N_2 de fortes intensités sont également détectés. Le filtrage de voisinage permet de réduire les fausses détections.

Résultats visuels de détection sur la base de test Ω_{Test}

La figure 5.15 présente l'intégralité des points détectés pour plusieurs patchs, le premier ne contenant pas de signature inusuelle et les suivants contenant différents types de signatures atypiques (Figure 2.3). Les résultats sont présentés avec et sans filtrage de voisinage. Les signatures sont mises en évidence dans l'application du modèle sans filtrage sur les patchs contenant des signatures inusuelles. Cependant plusieurs points normaux sont également détectés par l'approche, principalement les points normaux décalés illustrant la sensibilité de cette dernière aux positions des signatures. L'utilisation du filtrage de voisinage permet de réduire efficacement ces fausses détections tout en conservant la détection visuelle des signatures atypiques, permettant ainsi une meilleure mise en évidence des points inusuels. Cependant, comme nous pouvons le voir aussi bien sur le patch normal que les patchs inusuels, cette approche entraîne la détection des points normaux issus des signatures liées au N_2 . Elle reste donc très sensible à l'intensité vibratoire et peut entraîner des détections aux niveaux de raies normales de fortes intensités.

La caractérisation de la normalité à partir de la distribution de chaque point estimée avec des noyaux gamma permet une détection intéressante des signatures atypiques présentes sur les patchs pour les mettre en évidence. Cela entraîne également la détection des points sur les signatures vibratoires liées au N_1 et au N_2 , montrant une forte sensibilité aux intensités vibratoires et à leurs positions. L'application du filtrage de voisinage réduit le taux de fausses détections visuellement.

5.6 Conclusions

Dans ce chapitre, nous avons présenté la mise en place d'une détection ponctuelle des points des spectrogrammes à travers un modèle de normalité caractérisant chacun de ces points de manière indépendante. Ce modèle de normalité est défini pour chaque point à travers la distribution des points normaux. Bien que le modèle soit défini de manière ponctuelle, nous avons gardé la structure en patch afin de présenter les résultats.

La détection d'anomalies s'effectue à travers de tests statistiques d'adéquation par rapport au modèle de normalité. Cela entraîne donc un nombre de tests égal au nombre de points du spectrogramme et donc des problématiques de tests multiples. Cette problématique est résolue dans la littérature en considérant l'ensemble des tests réalisés globalement comme un unique ensemble à contrôler au niveau α et non unitairement. Les tests multiples sont cependant trop conservateurs et ont tendance à ne pas rejeter l'hypothèse de normalité entraînant une forte proportion de non détections. Cet effet est contradictoire avec le but recherché de détecter toutes les signatures atypiques présentes sur les spectrogrammes. Nous avons donc écarté l'approche basée sur les tests multiples et avons opté sur l'utilisation d'un filtrage basé sur les voisinages des points pour limiter les fausses détections. Un point est réellement détecté uniquement si 2 de ses voisins directs sont détectés également. Ainsi les points isolés ne sont pas détectés.

Le modèle de normalité a été défini dans un premier temps de manière paramétrique à partir de la distribution gamma. Cependant, cela supposait une distribution unimodale des points ce qui n'est pas nécessairement le cas dans une base de données volumineuse. Nous avons donc opté pour une caractérisation non-paramétrique sans a priori de la distribution des points du spectrogramme à partir de l'estimation de la densité par noyau. Deux noyaux ont été étudiés lors du développement, un noyau gaussien permettant une facilité de calcul mais ayant également une propriété de symétrie engendrant une densité non nulle pour les valeurs négatives, un noyau gamma permettant de pallier cela mais complexifiant les calculs et entraînant une impossibilité de mise à jour efficace du modèle avec de nouvelles données normales. Les temps de calcul des modèles basés sur les noyaux gammas sont 10 fois plus longs que ceux à partir des noyaux gaussiens pour des résultats proches. Il est donc plus pertinent d'utiliser le noyau gaussien dans un cadre industriel. Les modèles développés permettent une détection très efficace des points atypiques potentiellement présents sur les patchs et donc des signatures inusuelles. Cependant cette approche est fortement sensible à l'intensité des points et au décalage de ces derniers entraînant de fausses détections (faux positifs) de ces points. Il serait plus pertinent de définir le seuil de normalité sur chacun des points séparément cependant cela requiert une annotation en classes de l'ensemble des points du spectrogramme.

La méthode de détection développée dans ce chapitre considère d'abord les points comme indépendants puis tient compte du voisinage des points détectés via le filtrage. Dans la suite, nous ajoutons la dépendance des points voisins dans la définition du modèle de normalité en le définissant d'emblée par la distribution conditionnelle des points en fonction de leurs voisinsages.

Chapitre 6

Analyse ponctuelle conditionnellement au voisinage des spectrogrammes

6.1 Considération des points voisins

6.1.1 Dépendance par rapport au voisinage des points

Dans ce chapitre, nous décidons d'inclure la dépendance entre les points directement dans le modèle de normalité à travers la définition de la distribution des points et de leur voisinage. Il ne semble pas pertinent d'utiliser la dépendance du point par rapport à l'ensemble du reste du spectrogramme ou du patch. La figure 6.1 présente la corrélation entre les différents points d'un patch et leur voisins de rayon maximal 3 (soient 48 voisins dont la distance en norme l_∞ entre le point considéré et les voisins est au plus 3). Les différents voisins ont été ordonnés selon leurs distances en norme infinie par rapport au point, les premiers (ordre 1) sont les voisins de la première couronne autour du point, les suivants (ordre 2) correspondent à la seconde couronne et les derniers (ordre 3) à la troisième couronne. Nous pouvons remarquer que les corrélations sont plus importantes pour les voisins d'ordre 1. Certains voisins éloignés d'ordre 2 et 3 présentent des corrélations significatives avec le point considéré, cependant la plupart des points présents dans ces couronnes ont des corrélations proches de 0. Nous pouvons donc supposer que plus nous nous éloignons du point considéré plus les corrélations avec la majorité des points éloignés deviennent nulles. Il n'est donc pas utile de considérer des voisins trop éloignés du point étudié dans la définition du modèle. De plus, en augmentant le rayon de voisinage, le nombre de voisins à considérer augmente fortement pouvant entraîner des complications de calcul dû au nombre important de variables. Il est donc important de limiter ce rayon de voisinage.

Nous remarquons également que ces corrélations calculées pour tous les points du patch ne font pas apparaître de corrélation négative entre les points. Cela conforte l'idée selon laquelle les points du spectrogramme peuvent interagir entre eux de manière positive mais pas négative.

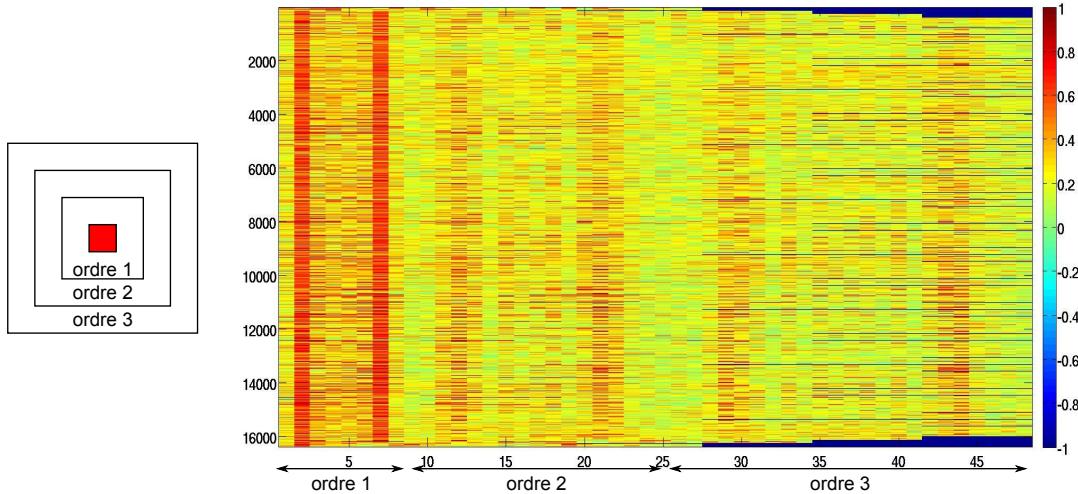


FIGURE 6.1 – Corrélation entre les points du patch et leurs voisins respectifs d’ordre 3 au maximum. Schéma de la signification des voisins des différents ordres. Les corrélations diminuent lorsque la distance entre les points augmente. Les majorité des points voisins d’ordre 3 ont des corrélations proches de 0 avec le point étudié.

6.1.2 Modèle de normalité par rapport au voisinage

Le modèle de normalité correspond à la distribution estimée de chaque point du spectrogramme sachant la distribution des points voisins. La classification des points s’effectue comme précédemment par rapport à la p-valeur du test d’adéquation par rapport à la distribution, cela équivaut à déterminer un classifieur binaire sur le point étudié avec comme variables explicatives les points voisins \mathcal{V}_{f,N_2} .

$$\hat{Y}_{f,N_2} = g(S_{f,N_2} | S_{\mathcal{V}_{f,N_2}}) = g(S_{f,N_2} | \mathcal{V}_{f,N_2})$$

avec g le classifieur et \mathcal{V}_{f,N_2} les coordonnées des voisins du point (f, N_2) avec un rayon de voisinage à définir et par abus de langage, pour simplifier les notations, les intensités vibratoires à ces mêmes coordonnées (donc $\mathcal{V}_{f,N_2} \equiv S_{\mathcal{V}_{f,N_2}}$). Les critères d’erreur sont les mêmes que ceux définis dans les chapitres précédents.

Nous sommes toujours dans un cadre de détection de nouveautés, cette fonction g est donc définie à partir de la distribution conditionnelle de normalité du point par rapport à son voisinage. Différents paramètres sont à prendre en compte dans l’estimation de cette distribution ainsi que dans la structure du voisinage. En effet, cette dernière ne correspond pas nécessairement à un cercle autour du point mais peut provenir d’information a priori.

L’intérêt d’étudier les probabilités conditionnelles en fonction du voisinage dans un cadre de détection de nouveautés provient de la structure de nos signatures atypiques. Ces dernières correspondent à des raies sur les spectrogrammes, ainsi plusieurs points voisins des spectrogrammes sont donc à détecter. La distribution conditionnelle des points est estimée une nouvelle fois par des méthodes d’estimation de densité par noyau [11, 62].

6.2 Estimation de densité conditionnelle par noyau

6.2.1 L'estimation de densité par noyau dans un cadre multidimensionnel

La formule générale d'estimation de densité par noyau est proche en d dimensions de la formule en 1 dimension, la principale différence provient de l'utilisation d'une matrice pour définir le paramètre d'échelle. Pour un jeu de données $X = (X_1, \dots, X_n)$ avec $X_i = (X_i^1, \dots, X_i^d)$, l'estimation de la densité multidimensionnelle de ces données au point $x = (x_1, \dots, x_d)$ est donnée par :

$$\hat{f}_{X,H}(x = (x_1, \dots, x_d)) = \frac{1}{n} \sum_{i=1}^n K_H(x - X_i) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\det(H)^{\frac{1}{2}}} K\left(H^{-\frac{1}{2}}(x - X_i)\right), \quad (6.1)$$

avec K le noyau multidimensionnel et H la matrice d'échelle symétrique définie positive.

La méthode est similaire à son homologue unidimensionnel, chaque observation du jeu de données multidimensionnelles intervient dans l'estimation de la densité en tout point $x \in \mathbb{R}^d$ avec des contributions plus ou moins importantes suivant la distance entre ce points et les observations.

Si le noyau K respecte les propriétés de densité (intégrale du noyau unitaire), de positivité, de symétrie, de moment d'ordre 2 fini alors l'estimateur est consistant en tout point x et converge en probabilité vers la densité à estimer. Le biais et la variance de cet estimateur sont donnés par :

$$\begin{aligned} \text{biais}\left(\hat{f}_{X,H}(x)\right) &= \frac{1}{2}\mu_2(K) \operatorname{tr}(H^T \mathcal{H}_f(x) H), \\ \text{var}\left(\hat{f}_{X,H}(x)\right) &= \frac{1}{n \det(H)} \|K\|^2 f(x), \end{aligned}$$

avec \mathcal{H}_f la hessienne de la densité f à estimer et $\mu_2(K) = \int x x^T K(x) dx$.

Le noyau gaussien multidimensionnel est le plus utilisé dans la littérature pour ses propriétés intéressantes de symétrie et de facilité de calcul.

$$K_{\mathcal{N}(\mu, \Sigma)}(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

La matrice H de l'échelle correspond au paramètre de la méthode d'estimation de densité multivariée par noyau. La forme de la matrice est le premier élément à définir. Elle peut être diagonale avec (h_1, \dots, h_d) les éléments diagonaux, donc de la forme :

$$H = \begin{pmatrix} h_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & h_d \end{pmatrix}.$$

Dans ce cas, le noyau symétrique peut s'écrire sous une forme multiplicative à partir de (6.1),

où chaque terme du produit correspond à un noyau unidimensionnel appliqué à la dimension correspondante [112],

$$K_H(x) = K_{h_1}(x_1) \times K_{h_2}(x_2) \times \dots \times K_{h_d}(x_d).$$

Le cas où tous les h_i sont égaux correspond à une sphère. Il est également possible de considérer une matrice d'échelle H pleine,

$$H = \begin{pmatrix} h_{11} & \dots & h_{1d} \\ \vdots & \ddots & \vdots \\ h_{d1} & \dots & h_{dd} \end{pmatrix}.$$

Dans ce cas, le noyau possédera une orientation dépendante de la matrice H . Ce cadre n'a pas d'analogie avec le cas univarié. Cette approche correspond à préorienter les données et utiliser un noyau diagonal. Une fois la forme de la matrice définie, il reste à estimer les différentes valeurs de cette dernière.

6.2.2 Calibration de la matrice d'échelle

L'estimation de la matrice d'échelle H est plus complexe que pour son homologue univarié, ce qui l'a rendu bien moins intéressante que son homologue durant plusieurs années. La structure de la matrice d'échelle H étant définie, différentes approches ont été mises en place pour définir les valeurs. Les différentes mesures de l'erreur d'estimation sont identiques au cas unidimensionnel tel que MISE et son approximation asymptotique AMISE. Il s'agit des deux grandeurs les plus couramment utilisées pour déterminer la matrice d'échelle H .

Il est possible d'utiliser la règle du pouce de manière analogue au cas univarié [104] en prenant une matrice d'échelle dépendante de la dimension d , du nombre d'observations n pour l'estimation de la densité et de la matrice de covariance estimée sur les données $\hat{\Sigma}$,

$$\hat{H} = n^{-\frac{1}{d+4}} \hat{\Sigma}^{\frac{1}{2}} \quad (6.2)$$

Une seconde estimation de H basée sur la règle du pouce est donnée dans [58] dans un cadre gaussien entraînant cependant un lissage trop important de la distribution.

$$\hat{H} = \left(\frac{4}{n(d+2)} \right)^{\frac{1}{d+4}} \hat{\Sigma}^{\frac{1}{2}}.$$

Des méthodes plus sophistiquées similaires à l'estimation de densité par noyau en 1 dimension ont également été mises en place. Elles se regroupent suivant 2 différentes approches, les méthodes plug-in et les méthodes de validation croisée. La première approche permet de trouver une estimation H en approximant les termes dépendant de la densité à estimer dans AMISE [112, 24]. La seconde approche consiste à appliquer la méthode de validation croisée "leave one out" de manière analogue au cas univarié pour trouver le minimum des fonctionnelles MISE ou

AMISE [44]. Une autre catégorie d'approche consiste à utiliser le maximum de vraisemblance pour trouver la matrice H optimale sur les données [119, 74].

6.2.3 L'estimation de la densité conditionnelle par noyau

L'estimation de la densité conditionnelle par noyau provient de l'application de l'estimation par noyau multidimensionnel à la formule de base de la densité conditionnelle [62]

$$f_{Y|X} = \frac{f_{XY}}{f_X} \approx \frac{\hat{f}_{XY}}{\hat{f}_X},$$

$f_{Y|X}$, f_{XY} et f_X sont respectivement la densité conditionnelle de $Y|X$, la densité jointe du couple (X, Y) et la densité de la variable aléatoire X .

Il s'agit donc d'utiliser l'estimation de la densité multidimensionnelle par noyau pour estimer aussi bien f_{XY} et f_X . L'estimateur de $f_{Y|X}(Y|X)$ à partir des données $(X, Y) = \{(X_i, Y_i) \in \mathbb{R}^d \times \mathbb{R}\}_{i=1, \dots, n}$ avec $X_i = (X_i^1, \dots, X_i^d)$ est donc

$$\hat{f}_{Y|X}(y|x) = \frac{\sum_{i=1}^n K_{H_{XY}}((x, y) - (X_i, Y_i))}{\sum_{i=1}^n K_{H_X}(x - X_i)}.$$

$K_{H_{XY}}$ correspond au noyau permettant d'estimer la loi jointe de (X, Y) en dimension $d + 1$ et K_{H_X} correspond au noyau utilisé pour estimer la densité de la variable X en dimension d .

Dans la littérature, les noyaux utilisés sont souvent d'échelle diagonale permettant de les mettre sous leurs formes multiplicatives, les noyaux définis au dessus se mettent donc sous la forme :

$$K_{H_{XY}}(x^1, \dots, x^d, y) = K_{H_Y}(y) \prod_{i=1}^d K_{H_{X^i}}(x^i) \quad K_{H_X}(x^1, \dots, x^d) = \prod_{i=1}^d K_{H_{X^i}}(x^i)$$

avec $K_{H_{X^i}}$ le noyau monodimensionnel avec pour valeur d'échelle $H_{X^i}(i, i)$ l'élément diagonale i de la matrice d'échelle. L'estimation de la densité conditionnelle (6.3) se reformule donc dans ce cadre :

$$\begin{aligned} \hat{f}_{Y|X}(y|x) &= \frac{\sum_{i=1}^n K_{H_Y}(y - Y_i) \prod_{j=1}^d K_{H_{X^j}}(x^j - X_i^j)}{\sum_{i=1}^n \prod_{j=1}^d K_{H_{X^j}}(x^j - X_i^j)} \\ &= \sum_{i=1}^n w_i(x|X_i^1, \dots, X_i^d) K_{H_Y}(y - Y_i) \end{aligned} \quad (6.3)$$

avec

$$w_i(x|X_i^1, \dots, X_i^d) = \frac{\prod_{j=1}^d K_{H_{X^j}}(x^j - X_i^j)}{\sum_{i=1}^n \prod_{j=1}^d K_{H_{X^j}}(x^j - X_i^j)} \quad (6.4)$$

Il s'agit donc d'une somme pondérée sur toutes les observations utilisées pour calibrer la distribution du noyau unidimensionnel appliqué à la variable cible Y par rapport à la contribution de

chaque observation w_i des noyaux associés aux variables conditionnelles X .

Le noyau le plus couramment utilisé pour estimer la densité conditionnelle est le noyau gaussien. Informatiquement il permet un calcul rapide de la densité conditionnelle en utilisant directement la densité gaussienne multidimensionnelle. H correspondant alors à la matrice de covariance de la distribution.

6.3 Le modèle de normalité défini à partir du voisinage d'ordre 1

Le modèle correspond à la densité conditionnelle du point en fonction de ses voisins. Il est défini sur les données normales pour chaque point du spectrogramme en considérant l'ensemble des points au voisinage du point étudié défini par une structure. Deux éléments sont donc à prendre en compte dans la définition de ce modèle de normalité :

- les points voisins utilisés comme paramètres de mise en place du modèle, il s'agit de la structure du voisinage prise en compte,
- la structure de la matrice d'échelle et l'estimation de ses coefficients.

6.3.1 La structure du voisinage

La figure 6.1 présente les corrélations entre un point étudié et son voisinage de rayon allant de 1 à 3. Les corrélations entre la majorité des points voisins et le point considéré diminuent en fonction de leurs distances l_∞ . Les intensités vibratoires ne sont pas toutes liées les unes aux autres, une intensité vibratoire à fréquence et régime donnés n'a aucune relation avec les intensités à des fréquences et régimes complètement différents. Il est donc logique de se limiter au voisinage proche (en norme l_∞) du point étudié. Nous ne disposons pas d'un nombre important de données pour estimer la densité de nos données. Cependant l'estimation de cette dernière impose l'estimation de la densité jointe dans un espace de grande dimension qui augmente en fonction du rayon de voisinage. Plus le nombre de variables considérées est important, plus le nombre de données nécessaires pour avoir une estimation fiable de la densité jointe augmente. Le nombre de variables explicatives est de 8 pour un rayon de voisinage de 1, 24 pour un rayon de 2 et 48 pour un rayon de convergence de 3. Ainsi pour faciliter et rendre pertinente l'estimation de la distribution conditionnelle, nous utilisons les voisins du premier ordre pour le modèle de normalité (Figure 6.2). Les variables de rayon de voisinage supérieur ou égal à 2 sont moins corrélées au point étudié (sur plusieurs directions) et apportent un nombre trop important de variables explicatives.

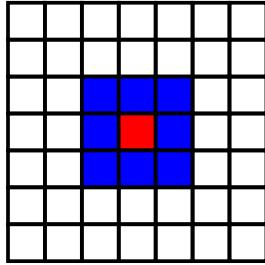


FIGURE 6.2 – Structure du voisinage direct (bleu) d'ordre 1 pour le point étudié (rouge)

6.3.2 Le modèle de normalité

Le modèle provient de l'estimation de la densité conditionnelle du point étudié sachant les intensités vibratoires de ses voisins directs. Cette distribution est estimée uniquement sur les spectrogrammes dont le patch est considéré comme normal afin que tous les points considérés dans la mise en place du modèle soient normaux. Les points atypiques ne sont pas isolés sur les spectrogrammes, ils font partie d'une structure sous forme de raies. Autour d'un point inusuel étudié se trouvent plusieurs points inusuels. Ces points ne sont pas utilisés lors de l'estimation de la densité entraînant des densités jointes très faibles se situant dans les extrêmes de la distribution. La densité jointe des variables explicatives et de la variable étudié est bien plus faible que la densité jointe des variables explicatives uniquement pour un point atypique alors que ces deux densités jointes seront du même ordre de grandeur pour un point normal.

Le modèle de normalité est donc défini par rapport à la densité conditionnelle en tout point (f, N_2) du patch par rapport au voisinage \mathcal{V}_{f, N_2} du point estimée à partir de l'ensemble des points normaux $S_{f, N_2}^{\Omega_{App}} = \{S_{f, N_2}^i\}_{i \in \Omega_{App}^{f, N_2}}$ et de leur voisinage associé $\mathcal{V}_{f, N_2}^{\Omega_{App}} = \{\mathcal{V}_{f, N_2}^i\}_{i \in \Omega_{App}}$ (6.5).

$$\forall (f, N_2) \quad \hat{f}_{(f, N_2)} | \mathcal{V}_{f, N_2} \left(x | \Omega_{App}^{f, N_2}, H_{SV}, H_V, \mathcal{V}_{f, N_2} \right) = \frac{\sum_{i \in \Omega_{App}^{f, N_2}} K_{H_{SV}}(x - S_{f, N_2}^i, \mathcal{V}_{f, N_2}^x - \mathcal{V}_{f, N_2}^i)}{\sum_{i \in \Omega_{App}^{f, N_2}} K_{H_V}(\mathcal{V}_{f, N_2}^x - \mathcal{V}_{f, N_2}^i)} \quad (6.5)$$

\mathcal{V}_{f, N_2}^x correspond au voisinage du point x étudié, par abus de langage et pour simplifier les notations, nous considérons également \mathcal{V}_{f, N_2}^x comme les intensités vibratoires des voisins. H_{SV} et H_V correspondent respectivement aux matrices d'échelle du noyau utilisé pour estimer la loi jointe du point étudié avec son voisinage et la loi jointe du voisinage. La seconde matrice peut être extraite de la première ou peut être estimée. Le noyau gaussien a été choisi pour estimer la densité conditionnelle pour accélérer les calculs malgré le biais à l'origine.

Dans cette section, nous considérons deux matrices d'échelle H , la première diagonale permettant une simplification des calculs et la seconde pleine permettant de donner une orientation spécifique au noyau. Cette seconde approche est intéressante car un point se trouvant sur une raie attendue devrait avoir son noyau associé avec une orientation similaire à cette raie. Cependant cette approche complexifie le calcul empêchant toute simplification et impose certaines approximations. Dans la suite nous donnons les résultats pour ces deux types de matrices d'échelle.

La matrice d'échelle doit être estimée pour chaque point du spectrogramme entraînant un grand nombre d'estimations. Les différentes méthodes avancées proposées nécessitent une optimisation pour trouver la matrice H optimale. Ces méthodes sont difficilement applicables pour tous les points. Nous avons donc opté pour la méthode du pouce défini en (6.2) afin de donner une estimation rapide et efficace des matrices d'échelle pour chaque point du spectrogramme.

6.3.3 Le score de détection

Le score de détection des points inusuels s'effectue à partir d'un test statistique opposant l'hypothèse \mathcal{H}_0^{f,N_2} : le point aux coordonnées (f, N_2) est normal sachant son voisinage d'ordre 1 contre l'hypothèse \mathcal{H}_1^{f,N_2} : le point aux coordonnées (f, N_2) est atypique sachant son voisinage d'ordre 1. La distribution sous \mathcal{H}_0^{f,N_2} est approximée par la loi de normalité estimée, ce test revient à un test d'adéquation des données à cette distribution. Ce test est appliqué pour chaque point des spectrogrammes. La décision est alors effectuée sur les p-valeurs (6.6) associées à ces tests.

$$pval_{f,N_2}^{i,K_{\mathcal{V}}} = \mathbb{P}_{\mathcal{H}_0^{f,N_2}} (X > S_{f,N_2}^i | \mathcal{V}_{f,N_2}^i) \approx \mathbb{P}_{\hat{f}_{(f,N_2)}|\mathcal{V}_{f,N_2}} (X > S_{f,N_2}^i | \mathcal{V}_{f,N_2}) \quad (6.6)$$

Le calcul de cette p-valeur est différent suivant la matrice d'échelle H diagonale ou pleine.

Matrice H diagonale

Dans ce cas, le noyau peut s'écrire sous sa forme multiplicative :

$$K_H(x) = \prod_{i=1}^d K_{H_i}(x^i) = \prod_{i=1}^d \frac{1}{H_i} K_{\mathcal{N}}\left(\frac{x^i}{H_i}\right),$$

avec H_i l'élément diagonal i de la matrice d'échelle H et x^i la i ème composante de x .

Le calcul de la p-valeur correspond alors à :

$$\begin{aligned}
pval_{f,N_2}^{i,K_V} &= \int_{S_{f,N_2}^i}^{+\infty} \hat{f}_{(f,N_2)} | \mathcal{V}_{f,N_2} (x | \Omega_{App}^{f,N_2}, H_{SV}, H_V, \mathcal{V}_{f,N_2}) dx \\
&= \int_{S_{f,N_2}^i}^{+\infty} \frac{\sum_{j \in \Omega_{App}^{f,N_2}} K_{H_{SV}}(x - S_{f,N_2}^j, \mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j)}{\sum_{j \in \Omega_{App}^{f,N_2}} K_{H_V}(\mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j)} dx \\
&= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \int_{S_{f,N_2}^i}^{+\infty} K_{H_{SV}}(x - S_{f,N_2}^j, \mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j) dx \\
&= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \int_{S_{f,N_2}^i}^{+\infty} \frac{1}{H_{d+1}} K_N \left(\frac{x - S_{f,N_2}^j}{H_{d+1}} \right) \prod_{k=1}^d \frac{1}{H_k} K_N \left(\frac{\mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j}{H_k} \right) dx \\
&= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \underbrace{\prod_{k=1}^d \frac{1}{H_k} K_N \left(\frac{\mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j}{H_k} \right)}_{\mathcal{N}_{\mathcal{V}_{f,N_2}^j, \bar{H}_{SV} \bar{H}_{SV}^T} (\mathcal{V}_{f,N_2}^x)} \underbrace{\int_{S_{f,N_2}^i}^{+\infty} \frac{1}{H_{d+1}} K_N \left(\frac{x - S_{f,N_2}^j}{H_{d+1}} \right) dx}_{\mathcal{N}_{S_{f,N_2}^j, H_{d+1}^2} (x)} \\
&= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \mathcal{N}_{\mathcal{V}_{f,N_2}^j, \bar{H}_{SV} \bar{H}_{SV}^T} (\mathcal{V}_{f,N_2}^x) \times \mathbb{P}_{\mathcal{N}_{S_{f,N_2}^j, H_{d+1}^2}} (X > S_{f,N_2}^i)
\end{aligned}$$

avec \bar{H}_{SV} la matrice extraite de H_{SV} correspondant aux points du voisinage de dimension $d \times d$, $\mathcal{N}_{\mathcal{V}_{f,N_2}^j, \bar{H}_{SV} \bar{H}_{SV}^T}$ la densité de la gaussienne de moyenne \mathcal{V}_{f,N_2}^j et de matrice de covariance $\bar{H}_{SV} \bar{H}_{SV}^T$, H_k le k ème élément diagonal de la matrice H_{SV} et

$$M_V = \sum_{j \in \Omega_{App}^{f,N_2}} K_{H_V} \left(\mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j \right) = \sum_{j \in \Omega_{App}^{f,N_2}} \mathcal{N}_{\mathcal{V}_{f,N_2}^j, H_V H_V^T} (\mathcal{V}_{f,N_2}^x).$$

Donc

$$pval_{f,N_2}^{i,K_V} = \sum_{j \in \Omega_{App}^{f,N_2}} w_j \times \mathbb{P}_{\mathcal{N}_{S_{f,N_2}^j, H_{d+1}^2}} (X > S_{f,N_2}^i) = \sum_{j \in \Omega_{App}^{f,N_2}} w_j \times \left(1 - F_{\mathcal{N}_{S_{f,N_2}^j, H_{d+1}^2}} (S_{f,N_2}^i) \right) \quad (6.7)$$

avec

$$w_j = \frac{\mathcal{N}_{\mathcal{V}_{f,N_2}^j, \bar{H}_{SV} \bar{H}_{SV}^T} (\mathcal{V}_{f,N_2}^x)}{\sum_{j \in \Omega_{App}^{f,N_2}} \mathcal{N}_{\mathcal{V}_{f,N_2}^j, H_V H_V^T} (\mathcal{V}_{f,N_2}^x)}$$

Par abus de langage, nous notons la densité de la loi $\mathcal{N}_{\mu, \sigma^2}$ par ce terme également.

La p-valeur dans le cas d'une matrice d'échelle diagonale correspond donc à la moyenne du complémentaire de la fonction de répartition des lois normales $\left\{ \mathcal{N}_{S_{f,N_2}^j, H_{d+1}^2} \right\}_{j \in \Omega_{App}^{f,N_2}}$ au point S_{f,N_2}^i pondérée par le ratio entre le noyau appliqué au voisinage d'une des observations et la somme des noyaux appliqués aux voisinages de toutes les observations. Les w_j peuvent ne pas sommer à 1 du fait de la différence entre les matrices d'échelle estimées en $d+1$ et d dimensions. Cependant cet écart n'est pas significatif. La pondération permet de donner plus de poids aux

données d'apprentissage dont les voisins sont proches des voisins des données étudiées.

En divisant par la densité jointe des points voisins, il est possible que ce dénominateur soit numériquement nul entraînant une division par 0. Ce phénomène se limite à une faible nombre de points. Le voisinage direct du point étudié contient donc des points fortement éloignés de la normalité, nous avons alors choisi de déclarer ces points comme inusuels.

Matrice H pleine

Dans ce cas de figure, il n'existe pas de simplification de la formulation de la densité conditionnelle. De plus le calcul de la p-valeur n'est pas simplifiée par l'utilisation du noyau gaussien étant donné que nous cherchons à estimer la fonction de répartition sur une seule des dimensions tout en tenant compte des produits croisés entre les différents directions.

$$\begin{aligned}
 pval_{f,N_2}^{i,K_V} &= \int_{S_{f,N_2}^i}^{+\infty} \hat{f}_{(f,N_2)|\mathcal{V}_{f,N_2}}(x|\Omega_{App}^{f,N_2}, H_{SV}, H_V, \mathcal{V}_{f,N_2}) dx \\
 &= \int_{S_{f,N_2}^i}^{+\infty} \frac{\sum_{j \in \Omega_{App}^{f,N_2}} K_{H_{SV}}(x - S_{f,N_2}^j, \mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j)}{\sum_{j \in \Omega_{App}^{f,N_2}} K_{H_V}(\mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j)} dx \\
 &= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \int_{S_{f,N_2}^i}^{+\infty} K_{H_{SV}}(x - S_{f,N_2}^j, \mathcal{V}_{f,N_2}^x - \mathcal{V}_{f,N_2}^j) dx \\
 &= \frac{1}{M_V} \sum_{j \in \Omega_{App}^{f,N_2}} \underbrace{\int_{S_{f,N_2}^i}^{+\infty} \mathcal{N}\left(\frac{\mathcal{V}_{f,N_2}^j}{S_{f,N_2}^j}, H_{SV} H_{SV}^T\right)(x, \mathcal{V}_{f,N_2}^x) dx}_{(*)}
 \end{aligned}$$

Il s'agit donc de calculer le complémentaire de la fonction de répartition mais uniquement sur une dimension de la densité de la distribution normale. Cette densité n'est pas séparable, il est donc impossible de donner la valeur réelle de cette intégrale. Nous en calculons donc une estimation.

Nous estimons l'intégrale à partir de la méthode des trapèzes utilisée pour estimer les intégrales de Riemann (6.8). Nous discrétisons l'espace d'intensité des points étudiés par la suite $(x_k)_{k \in \mathbb{N}}$ telle que $x_{k+1} = x_k + \xi$, avec $\xi = 10^{-4}$. Cette méthode permet de donner une estimation de l'intégrale qui ne somme pas à 1 sur l'ensemble de l'espace pour la fonction de densité, nous normalisons donc l'estimation de l'intégrale obtenue par M , l'estimation de cette même intégrale sur \mathbb{R}_+ . L'infini considéré dans le calcul de l'intégrale correspond à une valeur suffisamment

grande pour considérer la densité comme nulle.

$$\begin{aligned}
 (\star) &= \frac{\xi}{2M} \sum_{k|x_k=S_{f,N_2}^i}^{+\infty} \left\{ \mathcal{N}_{\begin{pmatrix} \mathcal{V}_{f,N_2}^j \\ S_{f,N_2}^j \end{pmatrix}, H_{S\mathcal{V}}H_{S\mathcal{V}}^T}^{(x_{k+1}, \mathcal{V}_{f,N_2}^x)} + \mathcal{N}_{\begin{pmatrix} \mathcal{V}_{f,N_2}^j \\ S_{f,N_2}^j \end{pmatrix}, H_{S\mathcal{V}}H_{S\mathcal{V}}^T}^{(x_k, \mathcal{V}_{f,N_2}^x)} \right\} \quad (6.8) \\
 M &= \frac{\xi}{2} \sum_{k=1}^{+\infty} \left\{ \mathcal{N}_{\begin{pmatrix} \mathcal{V}_{f,N_2}^j \\ S_{f,N_2}^j \end{pmatrix}, H_{S\mathcal{V}}H_{S\mathcal{V}}^T}^{(x_{k+1}, \mathcal{V}_{f,N_2}^x)} + \mathcal{N}_{\begin{pmatrix} \mathcal{V}_{f,N_2}^j \\ S_{f,N_2}^j \end{pmatrix}, H_{S\mathcal{V}}H_{S\mathcal{V}}^T}^{(x_k, \mathcal{V}_{f,N_2}^x)} \right\}
 \end{aligned}$$

Nous avons donc des formules des p-valeurs pour les modèles de normalité basés sur les distributions conditionnelles dans le cas des matrices d'échelle diagonales et pleines. La détection des points inusuels s'effectue à partir d'un seuil s sur ces p-valeurs.

$$\tilde{Y}_{f,N_2}^{i,K_{\mathcal{V}}} = \mathbb{1} \left\{ pval_{f,N_2}^{i,K_{\mathcal{V}}} \leq s \right\}$$

Le filtrage de voisinage (5.12) reste appliqué sur les détection de manière identique aux approches précédentes.

6.3.4 Calibration du modèle

Convergence des p-valeurs

Nous estimons des densités multidimensionnelles pouvant aller jusqu'à la dimension 9 avec un nombre limité d'observations pour les estimer. Nous vérifions donc la convergence des p-valeurs pour le patch contenant le moins de données normales. La figure 6.3 présente la différence d'estimation entre la p-valeur estimée à partir de $r+1$ données et celle estimée à partir de r données pour 2 points du spectrogramme et différentes données d'apprentissage. La convergence a été étudiée pour les 2 modèles de normalité définis avec différentes structures d'échelle. Nous pouvons voir la différence des estimations entre 2 nombres d'observations successifs diminuer en fonction de ce nombre et converger vers 0. De plus, les estimations restent rapidement dans les mêmes ordres de grandeur permettant une classification des points.

Estimation des seuils de détection

Les seuils de détection sont calibrés sur les données de validation de la base \mathcal{B}_1 annotées ponctuellement. Les taux de détection sont calculés sur les différentes classes de points pour différentes valeurs de seuil. Seuls les points inusuels doivent être détectés. La figure 6.4 présente ces taux de détection pour les modèles de normalité définis respectivement avec une matrice

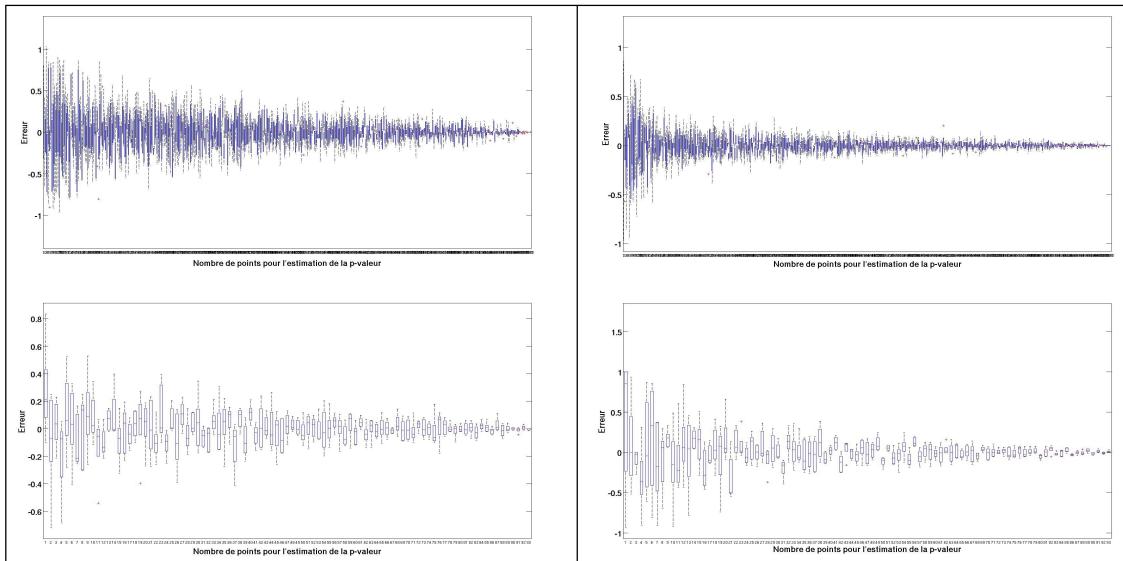


FIGURE 6.3 – Convergence des p-valeurs par rapport au nombre de points utilisés pour estimer la distribution conditionnelle à partir des deux modèles de normalité (matrice d'échelle diagonale (gauche) et pleine (droite)) pour deux points du spectrogramme. Les différentes de p-valeurs diminuent avec le nombre d'observations et convergent vers 0. Elles restent rapidement dans les mêmes ordres de grandeur permettant la classification des patchs.

d'échelle diagonale (à gauche) et une matrice d'échelle pleine (à droite).

Cette approche entraîne une bonne détection des points atypiques mais également la fausse détection des points normaux situés sur des raies se décalant sur les spectrogrammes (les points décalés). Les niveaux de détection sont différents pour les deux approches, l'utilisation d'une matrice d'échelle du noyau diagonale ou pleine modifiant la distribution estimée. Avec un seuil adapté, les méthodes basées sur le modèle de normalité conditionnel permettent toutes les deux de donner un fort niveau de bonnes détections. La figure 6.4 montre également que pour obtenir

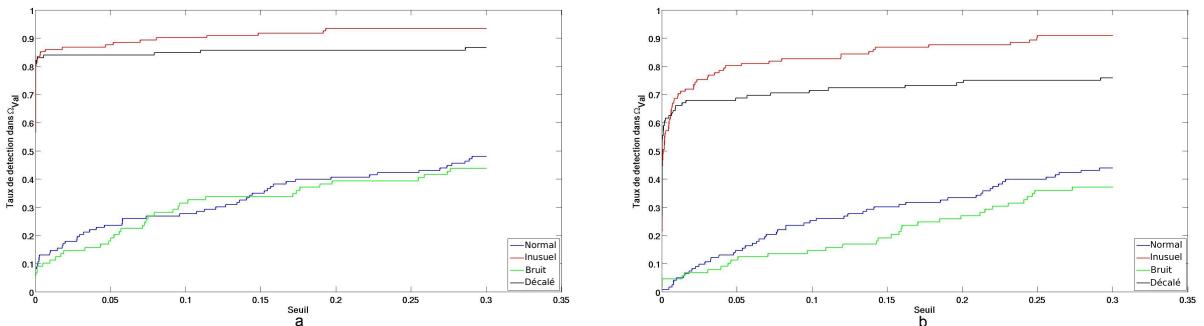


FIGURE 6.4 – Taux de détection des différentes classes de points de la base de validation Ω_{Val} à partir du modèle de normalité défini avec une matrice d'échelle diagonale (a) et une matrice d'échelle pleine (b). Les résultats établis à partir des deux matrices d'échelle sont différents. Les deux modèles permettent la détection des points atypiques (rouge) et des points normaux décalés (noir) pour de petites valeurs de seuil avec de faibles taux de détection des points normaux (bleu) et du bruit (vert). Les seuils de détection optimaux appartiennent donc à cette plage de valeurs et sont différents pour les deux approches.

le même niveau de bonnes détections, le seuil défini pour le modèle de normalité avec une matrice d'échelle diagonale doit être inférieur à celui pour la matrice d'échelle pleine. Nous avons donc défini deux seuils différents, s_{diag} pour le modèle défini avec une matrice diagonale et s_{pleine} pour le modèle avec une matrice pleine.

$$s_{diag} = 10^{-3} \quad s_{pleine} = 0.05$$

Ces seuils entraînent de bons taux de détection à plus de 80% sur la base de validation Ω_{Val} de \mathcal{B}_1 et limitent les fausses détections des points normaux et du bruit (autour de 15%). Les points décalés restent fortement détectés. Les seuils de détection sont les mêmes pour tous les points par manque d'annotations ponctuelles. L'approche peut être améliorée en calibrant un seuil de détection sur les p-valeurs pour chaque point.

6.3.5 Résultats sur la base de test Ω_{Test}

Dans ce chapitre, les résultats sont donnés uniquement sur le patch spécifique de la base \mathcal{B}_1 . Nous ne donnons pas les résultats sur l'ensemble des patchs comme pour les approches précédentes. Les résultats sont présentés sous la forme du taux de détection des points annotés de la base de test de \mathcal{B}_1 à partir de différentes données d'apprentissage du modèle de normalité et sous la forme de représentation visuelle de la détection sur l'ensemble du patch. Les effets du filtrage sont illustrés sur ce dernier résultat.

Taux de détection des points annotés

Le tableau 6.1 contient les taux de détections des différentes classes de points déterminés à partir de différentes données d'apprentissage des modèles de normalité avec matrice d'échelle diagonale et pleine. Les taux de détection des points atypiques sont très satisfaisants (à plus de 80%). De plus les points décalés sont moins détectés par rapport aux approches précédentes, de même pour les autres classes de points normaux. La prise en compte du voisinage directement dans le modèle de normalité basé sur l'estimation de la densité conditionnelle est donc pertinente pour détecter les signatures inusuelles tout en limitant les fausses détections.

Les résultats sur les deux approches sont assez similaires (hormis pour la classe "décalé"), les deux méthodes détectent de manière efficace les points inusuels. Le fait que certains points ne sont pas détectés par l'approche ne signifie pas une impossibilité de détecter les signatures atypiques correspondantes. Les signatures inusuelles consistent en des raies sur les spectrogrammes et constituent donc un ensemble de points à détecter, la détection de la majorité de ces points est suffisante pour détecter les signatures atypiques.

TABLE 6.1 – Taux de détection calculés à partir du modèle de normalité défini par la densité conditionnelle estimée à partir de différentes données d'apprentissage avec matrice d'échelle diagonale et pleine pour les différentes classes de points des données dans Ω_{Test}

Classe de points	Matrice d'échelle diagonale	Matrice d'échelle pleine
Classe "normal"	8.7% \pm 1.4%	12.3% \pm 2.2%
Classe "inusuel"	82.8% \pm 2.1%	80.5% \pm 2.1%
Classe "bruit"	5.6% \pm 1.4%	3.75% \pm 1.4%
Classe "décalé"	84.6% \pm 2.6%	70.8% \pm 2.3%

Résultats visuels de détection sur la base de test Ω_{Test}

Nous représentons sur la figure 6.5 les détections sur l'intégralité des patchs de plusieurs données tests à partir des deux modèles de normalité basés sur la distribution conditionnelle avec et sans filtrage. Le premier patch correspond à un patch normal (1ère ligne), les suivants contiennent différentes signatures atypiques. Il s'agit des mêmes patchs présentant les résultats dans les chapitres précédents (Figure 2.3). Les signatures inusuelles sont bien mises en évidence sur les différents patchs. Pour le patch ne contenant pas de signature inusuelle, les approches détectent principalement la raie N_1 avec quelques points isolés qui sont principalement retirés lors du filtrage par voisinage. La raie N_1 est également détectée sur les patchs atypiques. L'approche basée sur la matrice diagonale semble donner une détection plus forte avec la détection de quelques points voisins des signatures inusuelles mettant plus efficacement en évidence ces signatures, tandis que le modèle de normalité basé sur une matrice d'échelle pleine semble plus focalisé sur la détection des points uniquement sur les raies atypiques. Les deux approches entraînent des détections de nombreux points normaux isolés illustrant la sensibilité des approches aux intensités vibratoires. Ces fausses détections sont en majorité supprimées par le filtrage de voisinage (5.12).

Les temps de calcul des deux approches sont également complètement différents. L'utilisation des matrices diagonales permet d'avoir une formule explicite des p-valeurs à partir de la densité gaussienne, tandis que le calcul des p-valeurs dans le cas des matrices pleines passe par l'estimation de la densité entière. La méthode basée sur l'estimation de densité conditionnelle par noyau en utilisant des matrices diagonales est nettement plus rapide (environ 4 fois plus rapide) et donne des résultats proches de ceux du modèle de normalité utilisant des matrices d'échelle pleines. Dans la suite nous utilisons donc l'estimation de la densité conditionnelle à partir de matrices d'échelle diagonales.

Le modèle de normalité basé sur la densité conditionnelle de chaque point en fonction de tout le voisinage est donc efficace pour détecter les signatures vibratoires atypiques sur les patchs et permet la prise en compte du voisinage dans le modèle directement. Cette approche reste cependant sensible aux intensités vibratoires importantes et entraîne la détection des signatures liées au N_1 également. Malgré un modèle défini à partir du voisinage, l'utilisation du filtre de voisinage reste indispensable pour réduire les fausses détections tout en conservant une bonne mise en évidence des signatures inusuelles.

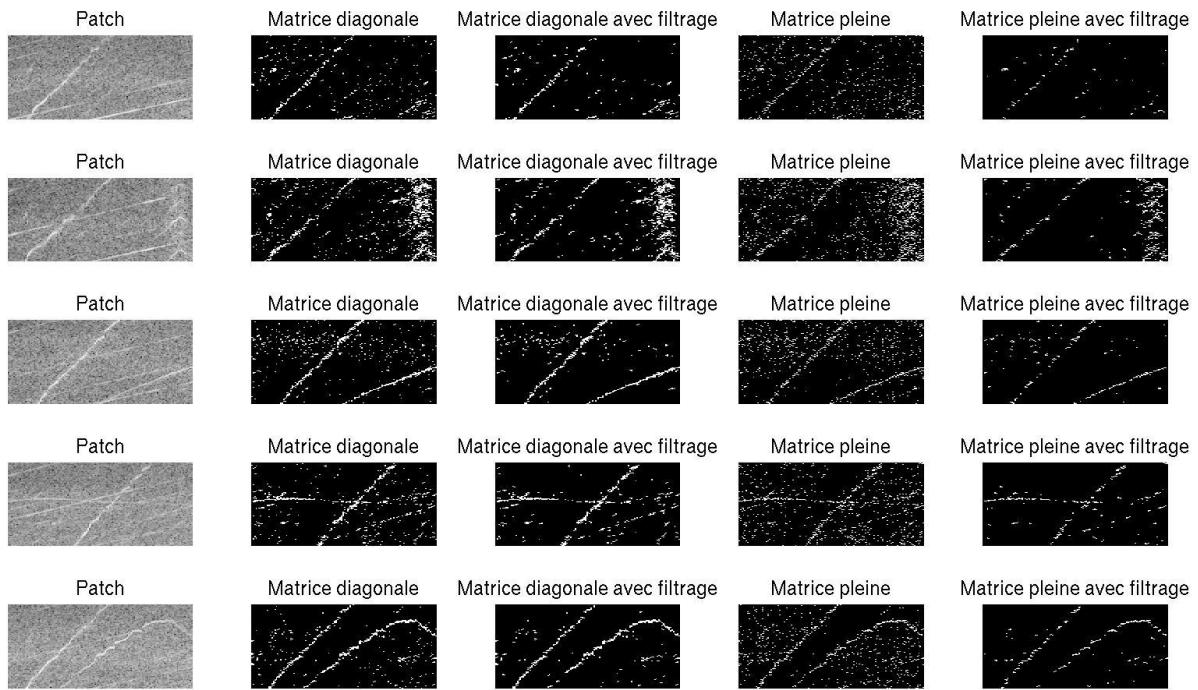


FIGURE 6.5 – Détections sur les patchs entiers à partir du modèle de normalité défini par les distributions conditionnelles des points avec des matrices d'échelles diagonale ou pleine et avec ou sans application du filtrage de voisinage. le premier patch (1ère ligne) correspond est normal, les suivants contiennent différentes signatures inusuelles. Les signatures atypiques sont mises en évidence à partir des deux modèles. Les raies vibratoires liées au N_1 sont également détectées, aucune signature liée au N_2 n'est détectée. Le filtrage de voisinage permet de réduire les fausses détections tout en conservant la détection des signatures atypiques.

6.4 Le modèle de normalité défini à partir du voisinage dans différentes directions

6.4.1 Le voisinage directionnel

Le modèle précédent tenait compte uniformément des voisins du premier ordre autour du point étudié pour trouver les points atypiques. Cependant, les signatures inusuelles que nous cherchons à détecter se présentent sous la forme de raies et ont donc pour attribut une direction. Ainsi en tenant compte de tous les voisins du premier ordre uniformément, nous masquons cette direction et prenons en considération plusieurs points non pertinents (les points voisins n'étant pas dans la direction de la signature inusuelle).

Dans cette section, la direction ainsi que le voisinage sont inclus dans le modèle de normalité. La direction des signatures atypiques n'est pas la même suivant le type d'anomalie sur les spectrogrammes et ne disposant pas d'information a priori sur ces signatures, ou leur localisation, nous étudions les différentes directions possibles pour chaque point. Pour cela, chaque point du

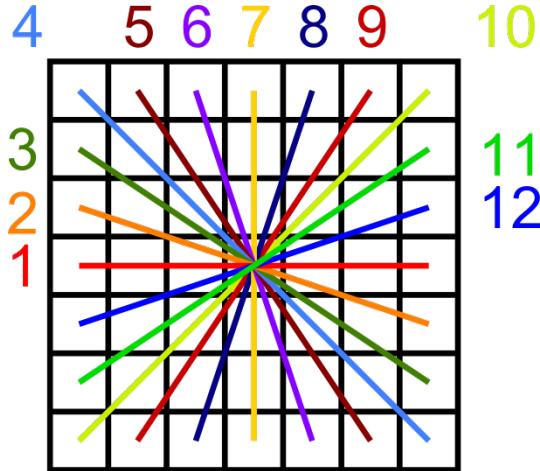


FIGURE 6.6 – Différentes directions linéaires de voisinage possibles en considérant un rayon de voisinage de 3. Chaque valeur correspond à une direction de voisinage. Nous considérons uniquement des voisinages linéaire car nous sommes sur de petites échelles où il est possible de considérer les formes courbes des spectrogrammes comme des droites.

spectrogramme est étudié par rapport à l'ensemble de ses voisins dans différentes directions $\vec{\mathcal{V}}$.

$$\vec{\mathcal{V}}_{f,N_2} = \left\{ \mathcal{V}_{f,N_2}^k, k = 1, \dots, \text{card}(\vec{\mathcal{V}}) \right\}$$

$\vec{\mathcal{V}}_{f,N_2}$ correspond à différents voisinages autour du point de coordonnées (f, N_2) suivant différentes directions paramétrées par k (Figure 6.6), et donc \mathcal{V}_{f,N_2}^k correspond au voisinage suivant la direction k . À la différence du voisinage non directionnel étudié dans la section précédente, il est possible de prendre en compte les voisins d'ordre supérieur à 1 car tous les voisins du même ordre ne sont pas pris en compte simultanément. Nous utilisons les voisins d'ordre inférieur ou égal à 3 afin de pouvoir caractériser la structure locale des raies. En prenant en compte les voisins d'ordre maximal 3, nous restons sur de petites échelles, il est donc possible de considérer chaque raie comme une forme linéaire (comme pour les curvelets). Ce niveau de voisinage permet de considérer 12 directions linéaires différentes représentées sur la figure 6.6. Chaque direction considérée prend en compte moins de points que le voisinage direct d'ordre 1 utilisé dans la section 6.3.

6.4.2 Le modèle de normalité

Les différentes directions considérées sont étudiées de manière indépendante. La distribution conditionnelle (6.9) des points le long de chaque direction est estimée à partir de la méthode à noyau appliquée aux données normales des mêmes points. On obtient ainsi un vecteur de densité où chaque composante correspond à la densité conditionnelle estimée dans une des directions de voisinage étudiées.

$$\varphi_{f,N_2}(x, \vec{\mathcal{V}}_{f,N_2}^x) = \left\{ \hat{f}_{f,N_2|\mathcal{V}_{f,N_2}^k}^k(x|\Omega_{App}^{f,N_2}, H_{S\mathcal{V}^k}, H_{\mathcal{V}^k}, \mathcal{V}_{f,N_2}^k) \right\}_{k=1,\dots,12} \quad (6.9)$$

avec k une des directions de voisinage présentées dans la figure 6.6, $\vec{\mathcal{V}}_{f,N_2}^x$ les voisinages selon les différentes directions associées à la donnée dont x est issu, \mathcal{V}_{f,N_2}^k le voisinage selon la direction k aux coordonnées (f, N_2) , $H_{S\mathcal{V}^k}$ et $H_{\mathcal{V}^k}$ sont les matrices d'échelles utilisées pour estimer respectivement la densité jointe du point étudié avec son voisinage dans la direction k et la densité jointe du voisinage dans la direction k . $\hat{f}_{f,N_2|\mathcal{V}_{f,N_2}^k}^k$ est alors la densité conditionnelle estimée du point par rapport à son voisinage dans la direction k . La densité conditionnelle est estimée de manière identique au chapitre précédent :

$$\hat{f}_{f,N_2|\mathcal{V}_{f,N_2}^k}^k(x|\mathcal{V}_{f,N_2}^{x,k}) = \frac{\hat{f}_{f,N_2,\mathcal{V}_{f,N_2}^k}^k(x, \mathcal{V}_{f,N_2}^{x,k})}{\hat{f}_{\mathcal{V}_{f,N_2}^k}^k(\mathcal{V}_{f,N_2}^k)}.$$

Chaque point du spectrogramme est alors caractérisé par sa vraisemblance par rapport à la densité conditionnelle estimée sur des données normales dans les différentes directions de voisinage. Un point inusuel appartient à une raie atypique et est donc lié à une direction. La vraisemblance de ce point dans la direction de la signature atypique est donc faible.

A la différence de la section précédente, plusieurs distributions doivent être estimées pour chaque point du spectrogramme. Ainsi, afin d'effectuer ces estimations dans un temps raisonnable, nous utilisons la méthode d'estimation de densité par noyau gaussien avec une matrice d'échelle diagonale estimée à partir de la règle du pouce. Le choix d'utiliser une matrice d'échelle pleine dans la section 6.3 venait d'une volonté de tenir compte de l'orientation des dépendances des points dans l'estimation. Cette prise en compte est effectuée dans ce cas directement à partir de la sélection du voisinage.

Le modèle de normalité correspond donc à l'estimation de la densité conditionnelle pour chaque point du spectrogramme par rapport à leurs voisins suivant les différentes directions à partir de la méthode à noyau gaussien appliquée avec des matrices d'échelle diagonales.

6.4.3 Le score de détection

La détection des points atypiques s'effectue à partir de la p-valeur des tests statistiques calculés à partir des distributions estimées. Un test statistique est réalisé au niveau de chaque direction de voisinage opposant l'hypothèse $\mathcal{H}_{0,k}^{f,N_2}$ de normalité du point suivant la direction étudiée à l'hypothèse $\mathcal{H}_{1,k}^{f,N_2}$ d'irrégularité de ce point sur cette même direction. Le calcul de la p-valeur s'effectue de manière analogue à la section précédente (6.7) comme un test d'adéquation par rapport aux distributions des différentes directions.

$$pval_{f,N_2}^{i,K_{\mathcal{V}^k}} = \mathbb{P}_{H_{0,k}^{f,N_2}}(X > S_{f,N_2}^i | \mathcal{V}_{f,N_2}^{i,k}) \approx \mathbb{P}_{\hat{f}_{f,N_2|\mathcal{V}_{f,N_2}^k}^k}(X > S_{f,N_2}^i | \mathcal{V}_{f,N_2}^{i,k}) \quad (6.10)$$

Ainsi, pour chaque point du spectrogramme nous disposons d'un indicateur ϕ_{f,N_2}^i (6.11)

correspondant aux p-valeurs (6.10) des tests statistiques effectués dans les différentes directions.

$$\phi_{f,N_2}^i \left(S_{f,N_2}^i, \vec{\mathcal{V}}_{f,N_2}^i \right) = \left\{ pval_{f,N_2}^{i,K_{\mathcal{V}_k}} \right\}_{k=1,\dots,12} \quad (6.11)$$

Le classifieur correspond alors à la comparaison de la **p-valeur minimale** de cet indicateur par rapport à un seuil de détection (6.12), nous associons cette p-valeur minimale à la direction selon laquelle la signature inusuelle est orientée. Si la p-valeur minimale est supérieure au seuil de détection s , cela signifie que dans toutes les différentes directions, le point peut être considéré comme normal. A l'inverse il suffit qu'elle soit inférieure au seuil de détection sur une seule direction pour que le point soit considéré comme atypique.

$$\tilde{Y}_{f,N_2}^{i,K_{\vec{\mathcal{V}}}} = \begin{cases} 1 & \text{si } \min \phi_{f,N_2}^i \left(S_{f,N_2}^i, \vec{\mathcal{V}}_{f,N_2}^i \right) \leq s \\ 0 & \text{sinon} \end{cases} \quad (6.12)$$

Le filtrage de voisinage (5.12) est également appliqué pour supprimer la détection de points isolés. Ce filtrage est intéressant dans ce cas de figure particulièrement car la prise en compte du voisinage directionnel entraîne la détection des voisins des points inusuels également.

$$\hat{Y}_{f,N_2}^{i,K_{\vec{\mathcal{V}}}} = \begin{cases} 1 & \text{si } \tilde{Y}_{f,N_2}^{i,K_{\mathcal{V}_k}} = 1 \& \sum_{l \in \mathcal{V}_{f,N_2}} \tilde{Y}_l^{i,K_{\mathcal{V}_k}} \geq 2 \\ 0 & \text{sinon} \end{cases}$$

6.4.4 La caractérisation des signatures inusuelles

La détection d'anomalies prend en compte la p-valeur minimale de l'indicateur (6.11) que nous associons à la direction de la signature inusuelle. Ainsi, pour chaque point nous disposons également de la direction estimée de la signature atypique à laquelle il appartient. Cependant, plusieurs composantes de l'indicateur des p-valeurs peuvent être très faibles numériquement et considérées comme nulles dû à la proximité des différentes directions. Elles ne permettent pas d'estimer la direction mais permettent tout de même de déclarer le point comme inusuel. Ainsi l'estimation de la direction de la signature atypique ne s'effectue pas sur l'indicateur des p-valeurs $\phi_{f,N_2}^i \left(S_{f,N_2}^i, \vec{\mathcal{V}}_{f,N_2}^i \right)$, mais correspond au minimum des densités conditionnelles dans les différentes directions $\varphi_{f,N_2} \left(S_{f,N_2}^i, \vec{\mathcal{V}}_{f,N_2}^i \right)$ (6.9). Cela correspond à sélectionner la direction selon laquelle le point étudié sachant les voisins dans cette même direction est le moins vraisemblable (6.13).

$$\vec{d}_{f,N_2}^i = \operatorname{argmin}_k \left(\varphi_{f,N_2} \left(S_{f,N_2}^i, \vec{\mathcal{V}}_{f,N_2}^i \right) \right) \quad (6.13)$$

Ces directions caractérisent donc les signatures inusuelles car tous les points issus de la même signature atypique et détectés par l'approche possèdent la même direction estimée (ou une des directions voisines) associée à l'orientation de la raie. Cette raie correspond également à une forme

connexe sur le spectrogramme. Une signature inusuelle détectée par notre approche correspond donc à une forme connexe dans l'espace des détections (espace de la dimension du patch dont chaque élément vaut 0 ou 1 suivant que le point est détecté ou non).

Les composantes connexes correspondent à un ensemble de points voisins détectés par notre approche, elles sont donc associées à une signature inusuelle. Nous considérons les directions estimées pour chaque point des composantes connexes comme un indicateur caractérisant les signatures atypiques. Chaque composante est alors caractérisée par un vecteur de dimension égale au nombre de points dans la composante. Chaque élément du vecteur correspond à la direction estimée du point associé. Nous caractérisons alors la signature inusuelle associée à la composante connexe par l'histogramme des différentes directions prises par les points la composant. Le mode de cet histogramme correspond à l'orientation principale de la signature atypique. Seules les composantes connexes contenant suffisant de points sont considérées.

6.4.5 Calibration du seuil de détection

Le seuil de détection est estimé à partir des taux de détection des différentes classes de points annotés des données de validation Ω_{Val} de la base \mathcal{B}_1 pour différentes valeurs de seuil de détection (Figure 6.7). Cette approche permet une très forte détection à plus de 85% des points atypiques ainsi que des points normaux décalés à des seuils de détection très faibles tandis que les autres points normaux restent très faiblement détectés à moins de 10%. Les probabilités conditionnelles sont d'intensités très faibles (principalement dans les directions des signatures inusuelles) entraînant un seuil de détection très faible et proche de 0. Afin d'avoir une meilleure visualisation du seuil de détection, nous représentons sur la figure 6.8 les points détectés sur 4 patchs de la base de validation à différents niveaux de seuils.

Les signatures inusuelles restent détectables pour toutes les valeurs de seuil étudiées malgré leurs très faibles valeurs. Nous pouvons remarquer que les valeurs de seuil de 10^{-6} et 10^{-8} entraînent de nombreuses détections de points normaux. La décision pour un seuil de 10^{-15} est intéressante mais ce seuil est numériquement très faible pouvant manquer certaines signatures. Nous avons donc opté pour un seuil plus tolérant malgré un nombre plus important de fausses détections dont la majorité pourra être évitée à partir du filtrage de voisinage.

$$s = 10^{-10}$$

6.4.6 Résultats sur la base de test Ω_{Test}

Taux de détection des points annotés

Le tableau 6.2 présente les taux de détection des différentes classes de points de la base de test Ω_{Test} de \mathcal{B}_1 à partir de notre approche avec et sans filtrage. Nous remarquons un important

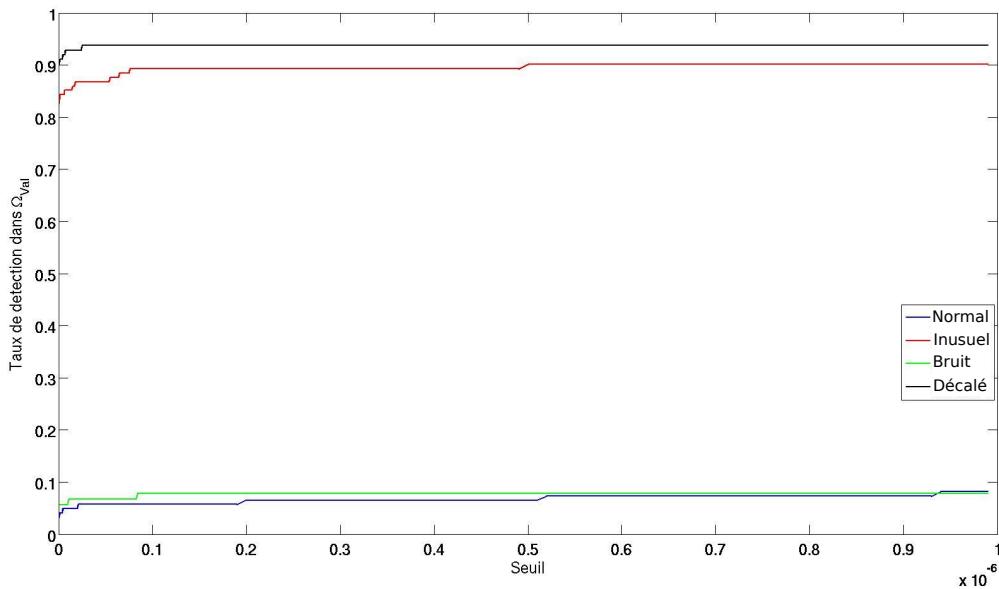


FIGURE 6.7 – Taux de détection des différentes classes de points de base de validation annotée ponctuellement Ω_{Val} en fonction du seuil de détection. Les points atypiques (rouge) et les points normaux décalés (noir) sont très fortement détectés à plus de 85% pour des valeurs de seuils très faibles tandis que les points normaux (bleu) et le bruit (vert) sont très faiblement détectés à moins de 10%. Nous choisissons donc un seuil de détection très faible.

taux de détection avec et sans filtrage des points atypiques attestant des bonnes performances du modèle de normalité basé sur la distribution conditionnelle des points en fonction des voisins dans différentes directions. Les points normaux et le bruit sont faiblement détectés, mais comme pour les autres approches, les points normaux décalés sur les spectrogrammes restent très fortement détectés. L'utilisation du filtrage de voisinage réduit légèrement les taux de détection mais son effet reste plus visible sur les représentations visuelles des résultats.

TABLE 6.2 – Taux de détection calculés à partir du modèle de normalité défini par les densités conditionnelles dans les différentes direction estimées à partir de différentes données d'apprentissage avec et sans filtrage pour les différentes classes de points des données dans Ω_{Test}

Classe de points	Détection sans filtrage	Détection avec filtrage
Classe "normal"	$10.3\% \pm 0\%$	$10.3\% \pm 0\%$
Classe "inusuel"	$88.4\% \pm 2.3\%$	$82.8\% \pm 2.7\%$
Classe "bruit"	$3.8\% \pm 1.4\%$	$3.1\% \pm 0\%$
Classe "décalé"	$93.3\% \pm 2.6\%$	$91.8\% \pm 2.3\%$

Résultats visuels de détection

Les figures 6.9, 6.10, 6.11, 6.12 et 6.13 présentent les résultats de détection à partir du modèle de normalité défini par la densité conditionnelle directionnelle sur tout le patch spécifique de la base \mathcal{B}_1 pour 5 différents spectrogrammes avec et sans application du filtrage de voisinage. Nous présentons également pour chacun d'entre eux les 4 composantes connexes les plus importantes

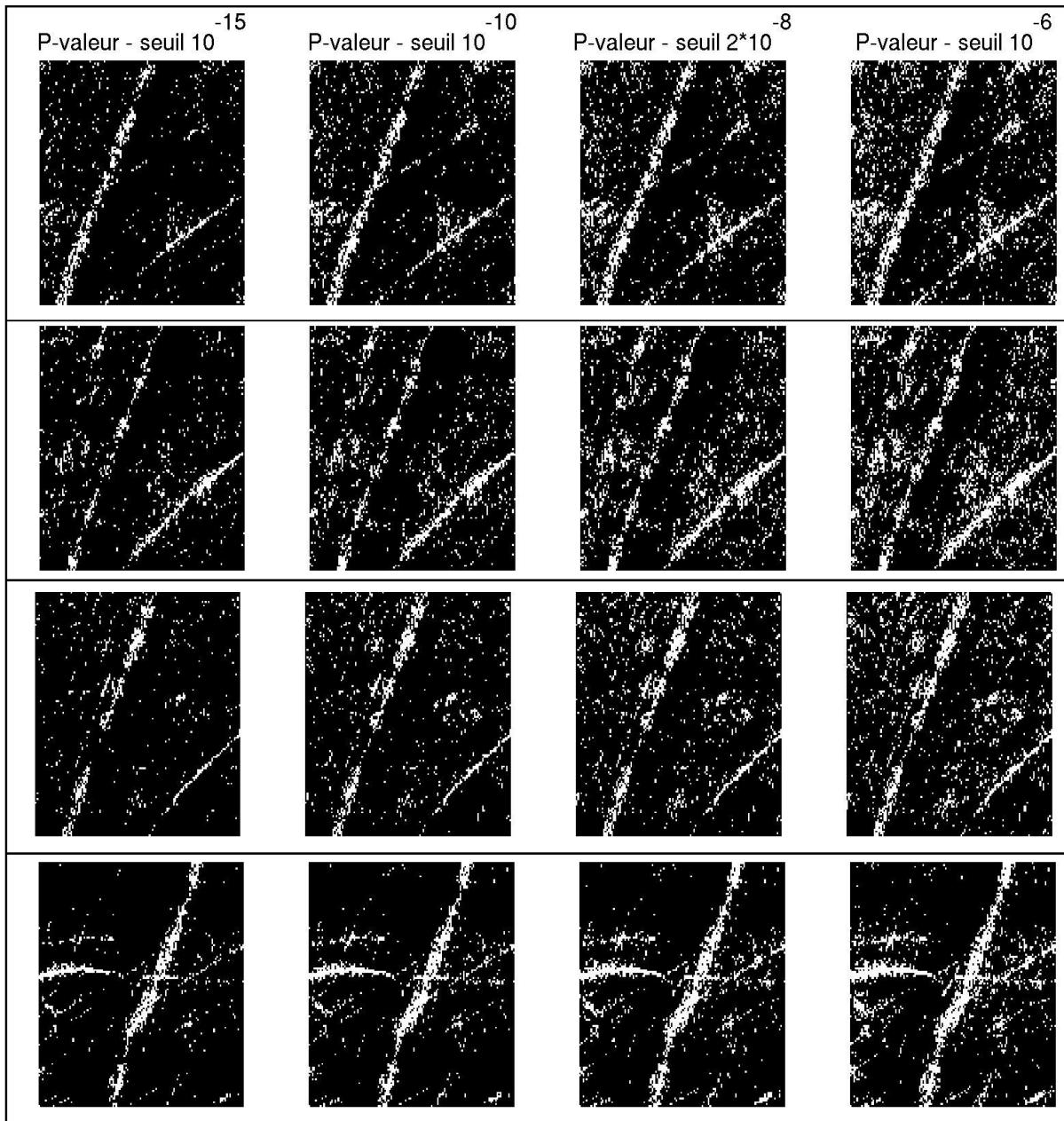


FIGURE 6.8 – Détection des points inusuels sur 4 différents patchs de Ω_{Val} pour différentes valeurs de seuil. Les signatures atypiques sont détectées pour les différentes valeurs de seuils sur les 4 patchs, la détection à 10^{-15} est intéressante. Nous avons choisi un seuil de 10^{-10} plus tolérant pour assurer la détection des signatures atypiques. Les seuils de 2×10^{-8} et 10^{-6} entraînent trop de fausses détections.

(celles contenant le plus de points) dans l'espace des détections avec les histogrammes des directions pour chacune de ces composantes connexes. La figure 6.9 concerne un patch normal sans signature inusuelle et les figures 6.10, 6.11, 6.12, 6.13 contiennent différentes signatures atypiques (Figure 2.3).

Le premier élément à constater sur ces résultats concerne les effets du filtrage. L'approche sans application du filtrage entraîne un certain nombre de fausses détections qui sont fortement réduites par son application permettant de ne laisser principalement que les signatures atypiques et celles décalées liées au N_1 . Le filtrage de voisinage permet donc de diminuer le nombre de fausses détections tout en laissant intact les détections des signatures inusuelles. Les différentes signatures atypiques sont bien mises en évidence dans les différents patchs contenant ces signatures. Nous pouvons constater que ces signatures sont visuellement très soulignées du fait de la détection des points voisins autour de ces dernières. Il s'agit tout de même de fausses détections, pouvant masquer les formes des signatures atypiques (Figure 6.10). Caractériser la normalité à partir de la densité conditionnelle de chaque point par rapport à leur voisinage dans différentes directions est donc une approche pertinente pour détecter les signatures atypiques.

Les éléments détectés sur le patch normal sont également très intéressants. Nous pouvons voir que le nombre de points détectés sur ce patch est nettement inférieur aux autres données étant donné qu'il ne possède pas de signature inusuelle. Seule la raie liée au N_1 est cataloguée comme inusuelle. Certaines sous zones de petite taille sont également détectées. Elles constituent une infime partie du patch étudié. Nous pouvons également voir que les raies liées au N_2 ne sont pas détectées par l'approche et sont donc bien caractérisées par le modèle de normalité.

Les plus grandes composantes connexes sur les différents patchs concernent les signatures liées au N_1 et les signatures atypiques. Ces différentes signatures peuvent se composer d'une ou plusieurs composantes. Il s'agit donc bien des éléments détectés par l'approche. Logiquement, les composantes connexes détectées sont les mêmes pour les détections avec et sans filtrage, ce qui permet d'affirmer l'inaction du filtrage sur les signatures inusuelles. Chaque composante connexe est associée à un histogramme de la même couleur dont les bins correspondent aux numéros des directions présentés en figure 6.6. Les histogrammes associés à ces composantes possèdent des modes au niveau de la direction de la signature, cependant la détection des points voisins entraînent la considération d'autres directions également. Pour le patch normal (Figure 6.9), seule la plus grande composante contenant le plus de points est intéressante (les autres ne contenant pas suffisamment de points pour caractériser les signatures). L'histogramme associé à cette composante connexe possède un mode correspondant exactement à la direction de la signature. La détection des signatures inusuelles au niveau du second patch (Figure 6.10) concerne de nombreux points voisins également masquant les directions des signatures. L'histogramme de la composante au niveau de la raie N_1 (jaune) possède un mode au niveau de la direction de la signature. Les composantes des autres patchs possèdent des histogrammes dont les modes sont proches de l'orientation de la signature. La direction des points détectés par l'approche sur les patchs permet bien de caractériser plus finement les signatures atypiques avec la détection des points et la direction de la signature à laquelle ils appartiennent.

Le modèle de normalité basé sur la densité conditionnelle de chaque point en fonction des voisins dans différentes directions est donc efficace pour détecter les points atypiques sur les spectrogrammes à partir de la p-valeur minimale malgré la détection des points voisins masquant certaines signatures. Les densités associées à ces p-valeurs permettent d'estimer la direction de la signature inusuelle. L'étude des composantes connexes dans l'espace de détections permet d'obtenir un histogramme des directions caractérisant les signatures atypiques qui pourrait permettre de labéliser les signatures à partir d'un clustering.

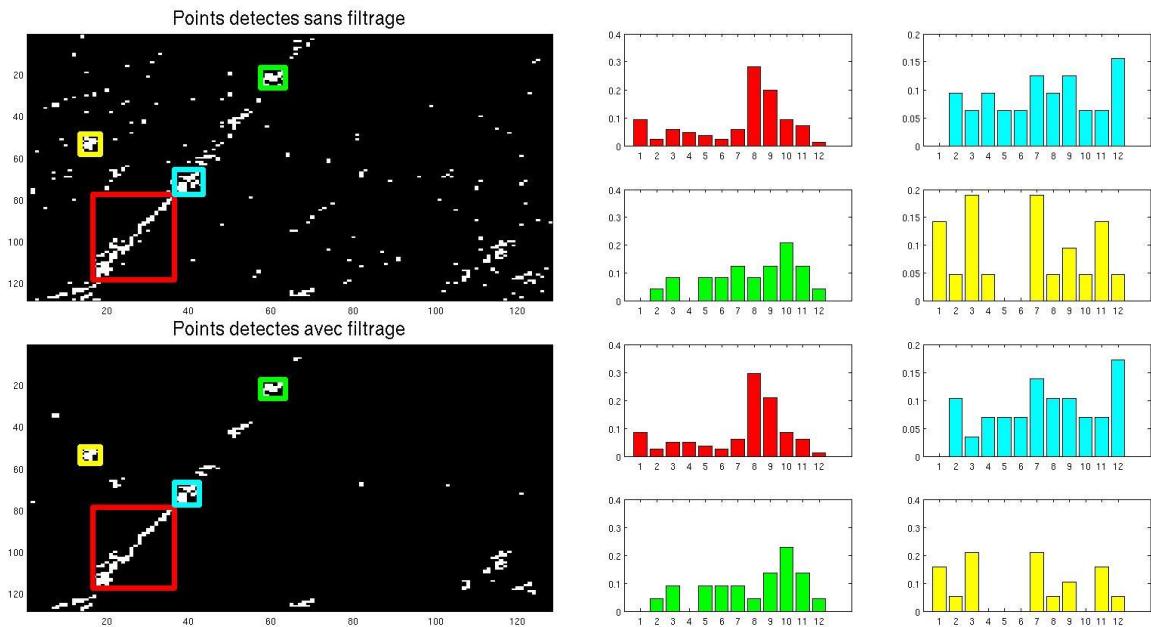


FIGURE 6.9 – Détection des points inusuels sur un patch sans signature inusuelle avec (partie inférieure) et sans (partie supérieure) application du filtrage de voisinage. Les 4 composantes connexes sur les espaces de détections sont extraites (cadres de couleur) et associées à un histogramme de détection (à droite dans la même couleur) pour les 2 images de détections, l'abscisse des histogrammes correspond aux différentes directions de voisinage étudiées. Le filtrage de voisinage (partie inférieure) permet de réduire les mauvaises détections. Les composantes connexes correspondent à des éléments de la raie N_1 , l'histogramme de la composante connexe principale possède un modèle correspondant à l'orientation de la signature.

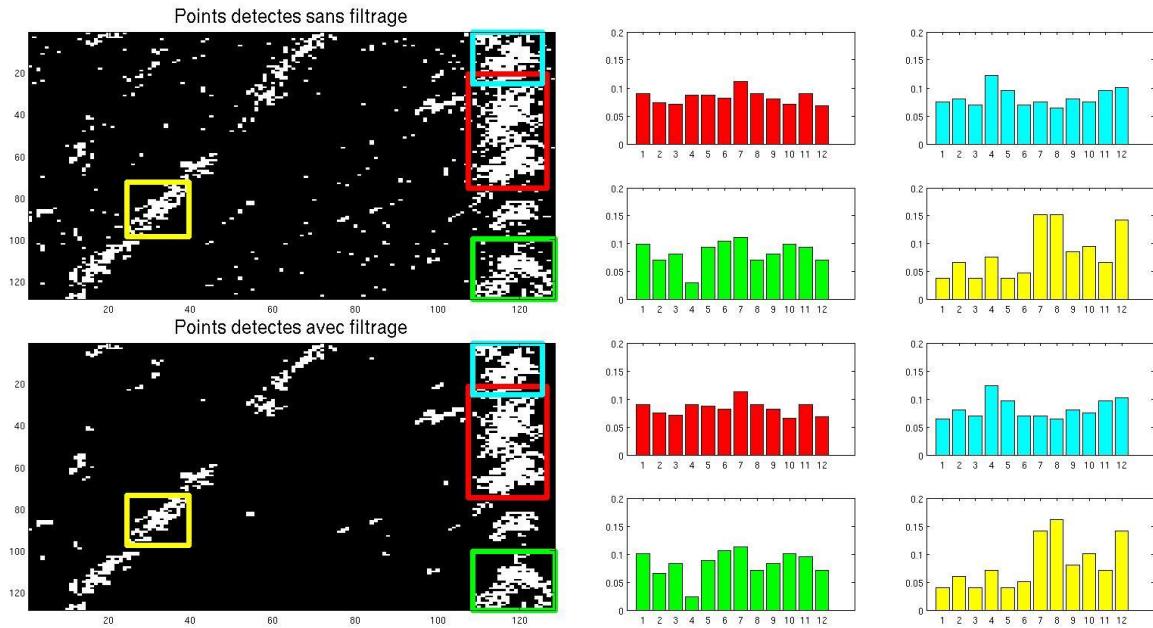


FIGURE 6.10 – Détection des points inusuels sur un patch contenant un premier type de signature inusuelle. La structure est identique à la figure 6.9.

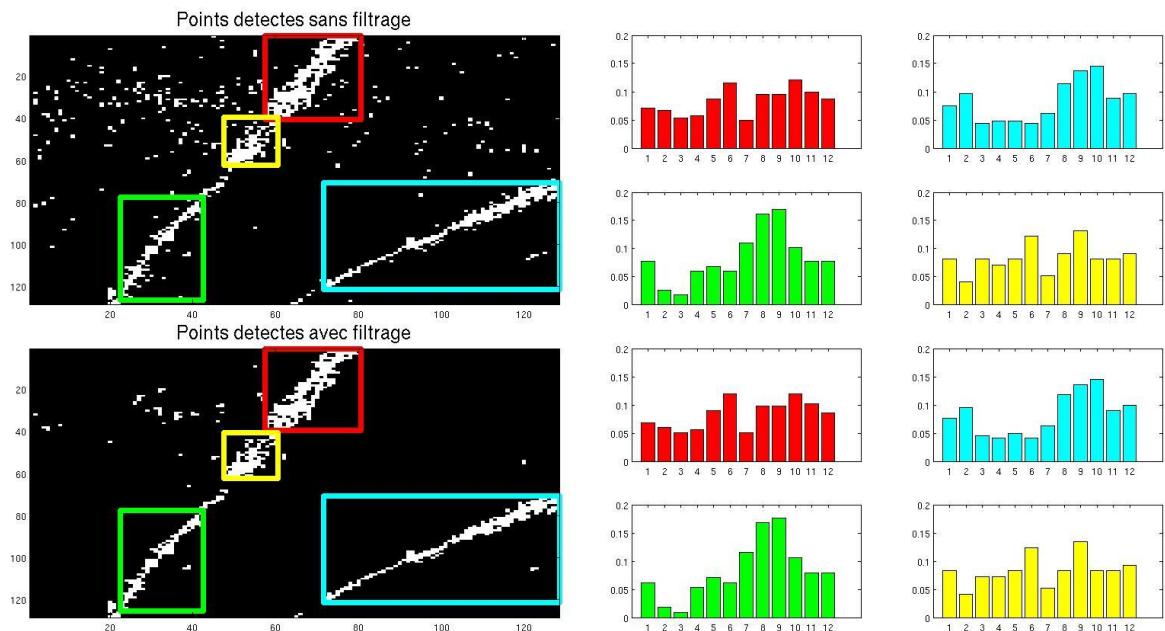


FIGURE 6.11 – Détection des points inusuels sur un patch contenant un second type de signature inusuelle. La structure est identique à la figure 6.9.

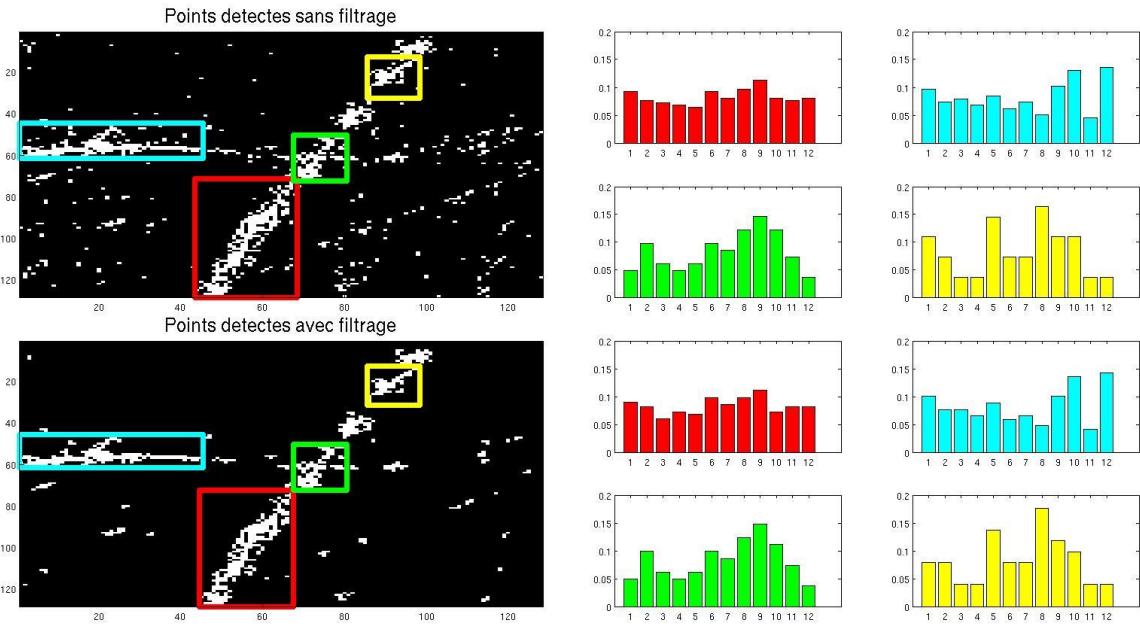


FIGURE 6.12 – Détection des points inusuels sur un patch contenant un troisième type de signature inusuelle. La structure est identique à la figure 6.9.

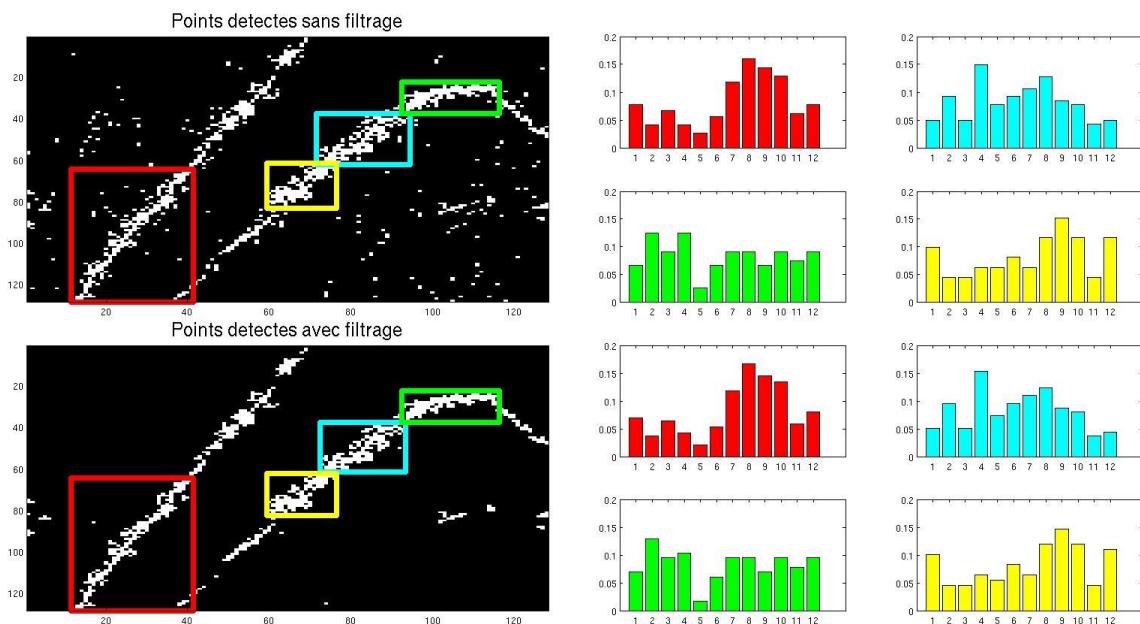


FIGURE 6.13 – Détection des points inusuels sur un patch contenant un quatrième type de signature inusuelle. La structure est identique à la figure 6.9.

6.5 Fusion des différentes approches

6.5.1 Comparaison des approches par dictionnaire et ponctuelles

Les approches par dictionnaire définissent un modèle de normalité globalement sur le patch, tandis que les approches ponctuelles définissent des modèles de normalité localement pour chaque point. Les approches par dictionnaire sont donc plus à même de capturer la structure entière du patch, tandis que les approches ponctuelles sont plus efficaces pour étudier les comportements locaux. Dans l'étude ponctuelle, la détection s'effectue uniquement par comparaison avec les mêmes points sur les patchs normaux entraînant une sensibilité aux intensités vibratoires et à la position des raies. Pour les modèles définis sur les patchs, la détection est effectuée à travers les résidus de reconstruction intra-patchs. Le seuil de détection est différent sur chaque patch entraînant une détection systématique de plusieurs points, y compris pour les patchs sans signature atypique. Les approches ponctuelles définissant un modèle pour chaque point sont plus coûteuses et prennent plus de temps que les approches par patch. La définition du modèle de normalité des approches par dictionnaire est de l'ordre de quelques secondes (entre 5 secondes pour les curvelets et 10 secondes pour la NMF). Pour les approches ponctuelles, la définition est de l'ordre de plusieurs minutes (2 minutes pour les approches indépendantes, 5 minutes pour les approches conditionnelles aux voisins d'ordre 1 et une dizaine de minutes pour les approches conditionnelles directionnelles). Ces temps de calcul restent convenables dans un cadre industriel. La décision pour ces différents modèles s'effectuent en des temps très court.

La comparaison des approches montrent un avantage pour les approches ponctuelles par rapport aux approches par dictionnaire. Elles proposent un bon niveau de détection tout en limitant les fausses détections des points liés au N_2 et du bruit. Les points liés au N_1 restent détectés par les 2 types d'approches. Visuellement les deux approches permettent la détection des signatures atypiques sur les spectrogrammes. La différence provient des points détectés à tort. Les approches ponctuelles détectent à tort des points isolés tandis que les approches par dictionnaire détectent à tort des morceaux de raies. Les résultats numériques sont à nuancer car nous utilisons une base annotée ponctuellement de quelques points afin d'établir les seuils et de présenter les résultats. Il serait plus efficace de définir un seuil sur chaque point séparément afin d'optimiser la détection mais cela nécessite une base de données intégralement annotées ponctuellement afin de donner des résultats généraux.

6.5.2 Fusion des approches par dictionnaire et ponctuelles

Dans la section 4.5.5, les détections des 2 approches par dictionnaire avaient été fusionnées afin d'obtenir des résultats plus robustes et de réduire les fausses détections. Cela était possible car les différentes approches permettaient la détection des signatures inusuelles mais également de différents points normaux. Les approches ponctuelles permettent également la détection des points inusuels. Il est alors possible d'utiliser les différents résultats de détection des différentes

approches afin de rendre la méthode de détection plus robuste. De plus, il est intéressant de combiner les approches par dictionnaire définissant des modèles sur le patch global et les approches ponctuelles définissant des modèles localement sur les patchs.

Les différentes approches combinées sont :

- la NMF (\hat{Y}_{f,N_2}^{NMF})
- les curvelets (\hat{Y}_{f,N_2}^C)
- l'estimation de densité indépendante par noyau gaussien ($\hat{Y}_{f,N_2}^{K_N}$)
- l'estimation de densité conditionnelle par rapport à tous les voisins d'ordre 1 ($\hat{Y}_{f,N_2}^{K_V}$)
- l'estimation de densité conditionnelle par rapport au voisinage directionnel ($\hat{Y}_{f,N_2}^{K_{\vec{V}}}$).

Nous définissons *methods* comme étant l'ensemble des différentes méthodes énoncées ci-dessus.

La décision à partir de cette combinaison d'approche correspond à un vote, un point est considéré comme inusuel si la majorité des approches le considère comme tel.

$$\hat{Y}_{f,N_2}^i = \begin{cases} 1 & \text{si } \frac{1}{\text{card}(\text{methods})} \sum_{k \in \text{methods}} \hat{Y}_{f,N_2}^{i,k} \geq 0.5 \\ 0 & \text{sinon} \end{cases}$$

Les figures 6.14, 6.15, 6.16, 6.17 et 6.18 donne les résultats de détections des différentes approches et de l'approche combinée pour les patchs sur lesquels les résultats ont été illustrés tout au long du mémoire. Les différentes sous-images de chacune de ces figures correspondent au patch, à la décision effectuée à partir de chacune des approches ainsi qu'à la décision effectuée à partir d'un vote sur ces différentes approches. La figure 6.14 concerne un patch normal tandis que les autres figures contiennent différents types de signatures atypiques.

La combinaison des différentes approches de détection permet d'affiner la décision finale. Les différentes méthodes ne détectent pas à tort les mêmes points. Les signatures atypiques restent détectées par les différentes approches, mais des éléments supplémentaires sont également détectés comme des portions de raies pour la NMF ou des points mal représentés par les autres modèles tels que les points voisins aux signatures atypiques. Ce mélange des différentes méthodes permet de détecter plus précisément les points inusuels ainsi que la structure des signatures atypiques. Il semble être prometteur pour la détection des signatures atypiques sur les spectrogrammes.

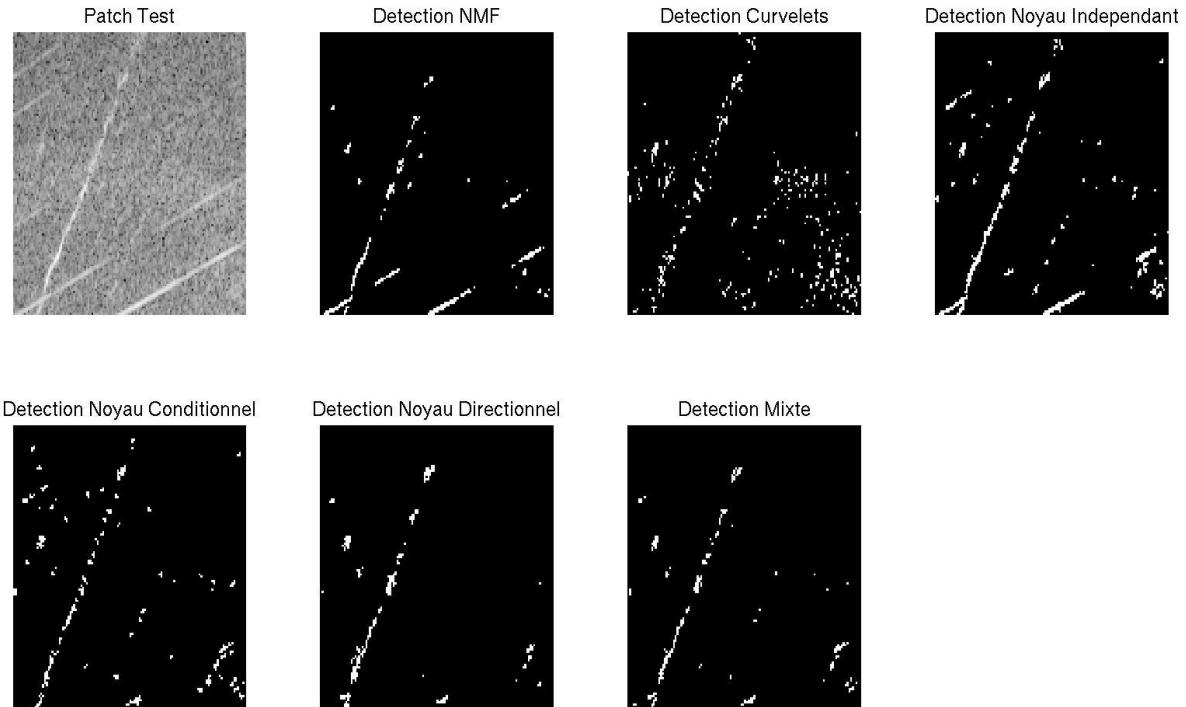


FIGURE 6.14 – Détection du patch sans signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{V}}}$, ainsi que de l'approche mixte \hat{Y} . La seule signature détectée sur ce patch correspond à la raie N_1 se décalant sur les différents spectrogrammes.

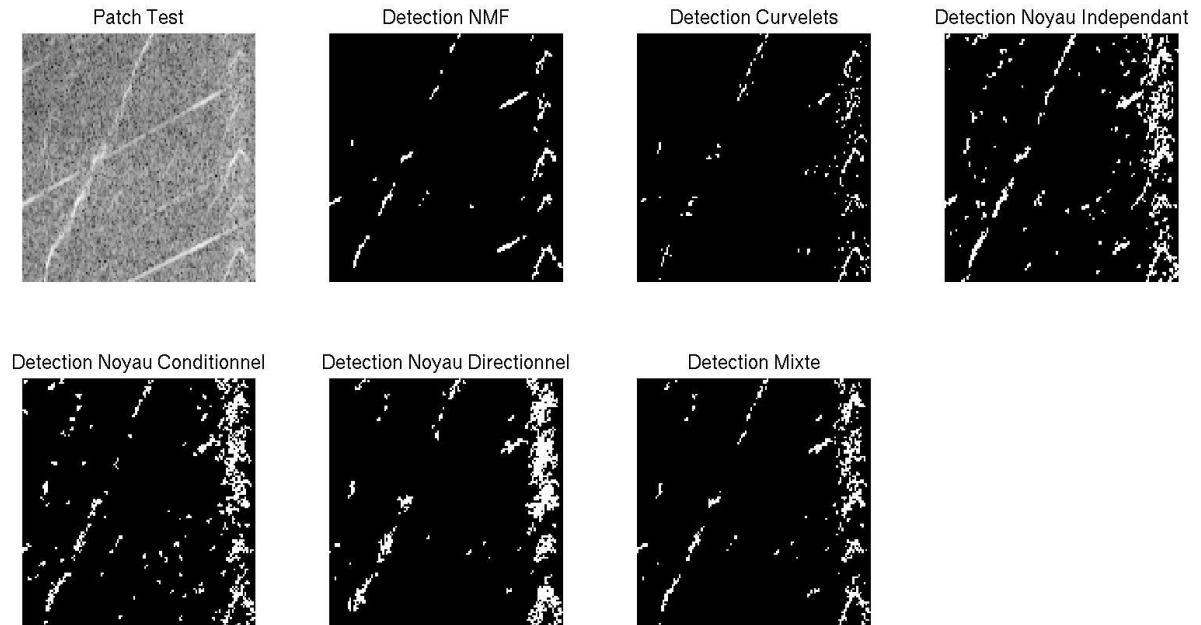


FIGURE 6.15 – Détection du patch contenant un premier type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{V}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

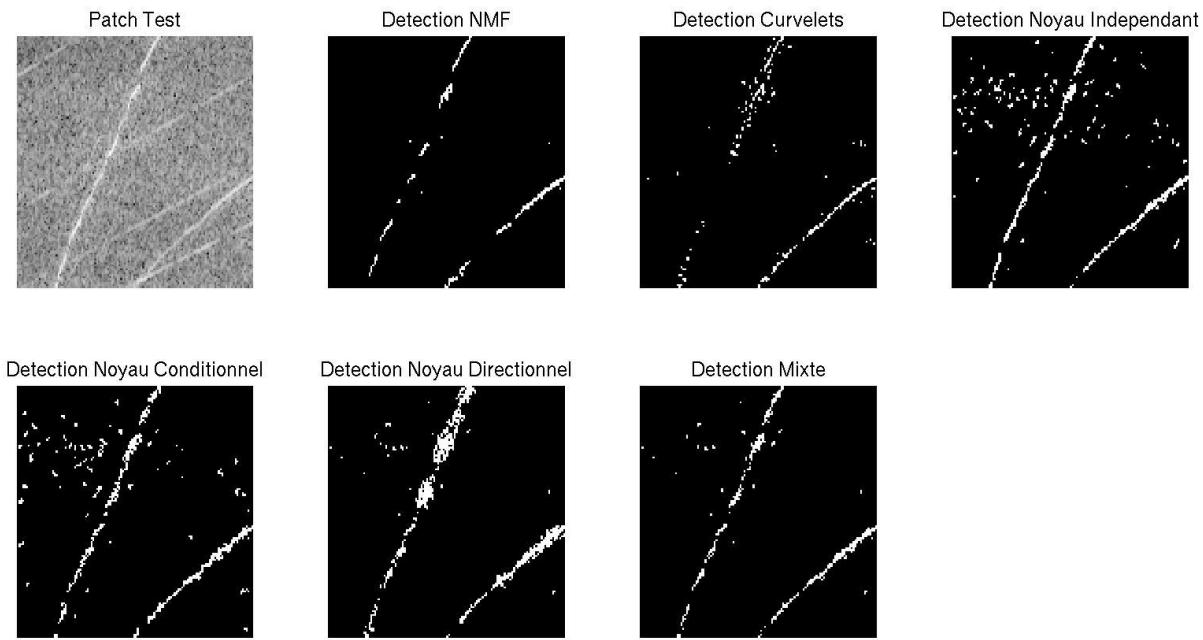


FIGURE 6.16 – Détection du patch contenant un second type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{V}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

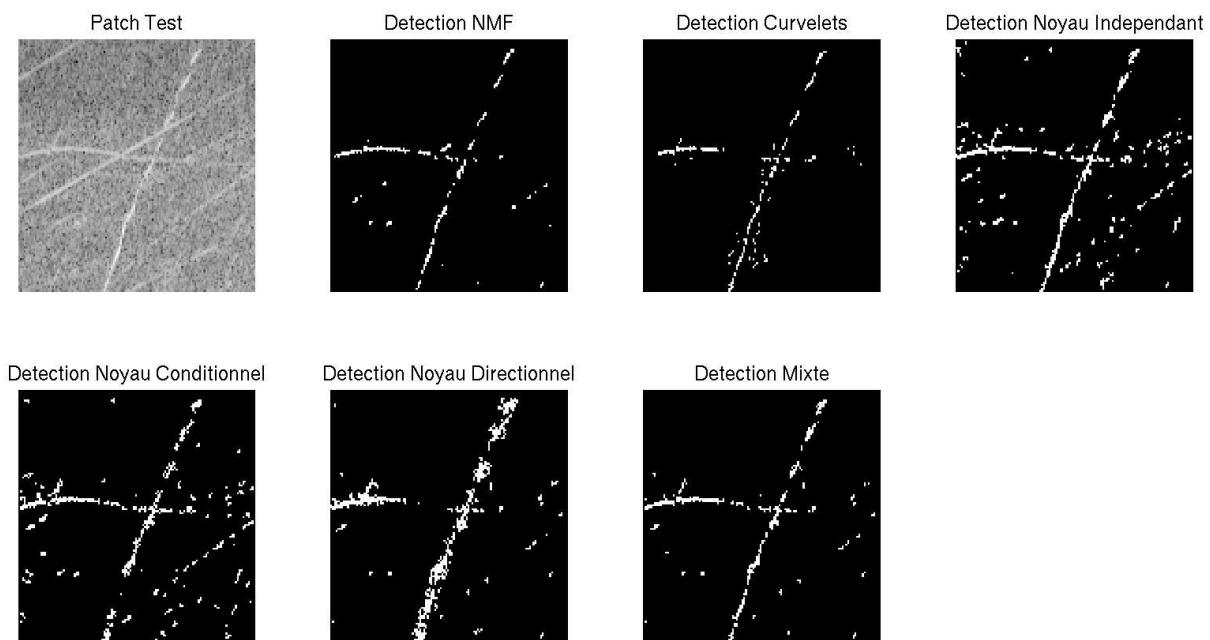


FIGURE 6.17 – Détection du patch contenant un troisième type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{V}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

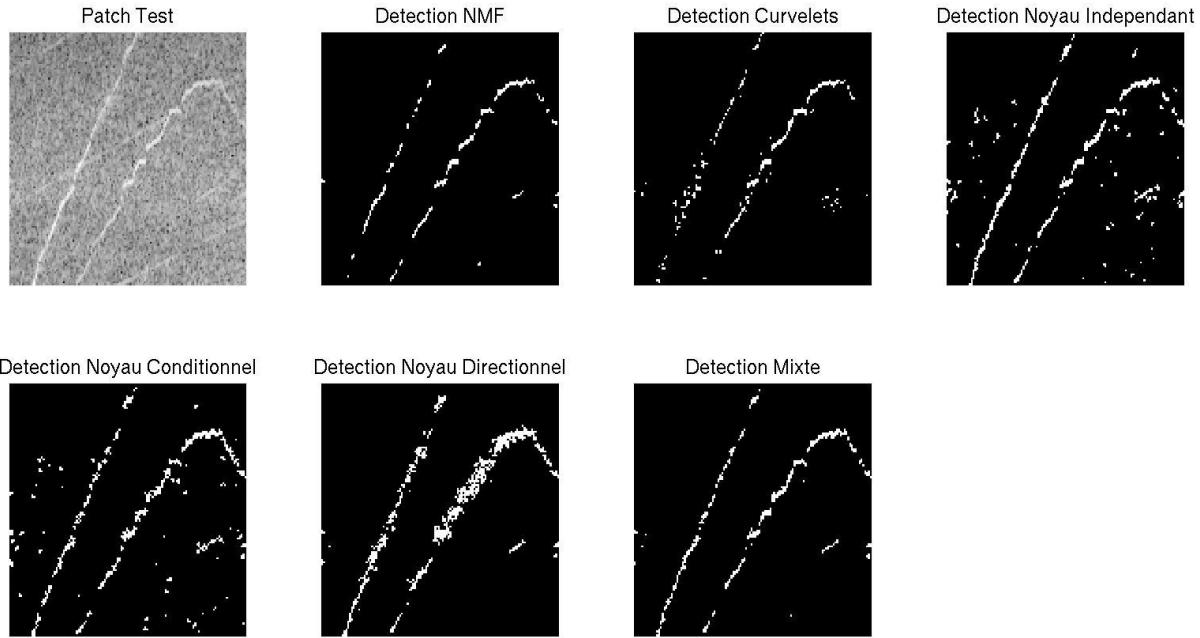


FIGURE 6.18 – Détection du patch contenant un quatrième type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{KN} , \hat{Y}^{KV} et $\hat{Y}^{K\vec{v}}$, ainsi que de l’approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

6.6 Conclusions

Les approches conditionnelles tiennent compte de la structure des signatures inusuelles. En effet, les signatures atypiques recherchées correspondent à des raies, les points voisins ont donc leur importance et il est donc pertinent de tenir compte du voisinage dans la mise en place du modèle de normalité. Cela permet de considérer la structure locale des données. Le rayon de voisinage ne doit pas être trop important car l’influence du voisinage sur un point à fréquence et régime donnés diminue en fonction de la distance au point étudié. La structure du voisinage est également à prendre en compte. Nous avons étudié le voisinage direct (d’ordre 1) où chaque intensité vibratoire voisine était considérée de manière uniforme. Cependant cette structure n’est pas en adéquation avec la structure du patch étant donné que les signatures correspondent à des raies et possèdent donc une orientation. La prise en compte de tous les voisins entraîne l’utilisation de voisins non pertinents. Nous avons également étudié un voisinage directionnel autour du point étudié où le voisinage correspond à des formes linéaires ayant une certaine orientation.

Pour les deux approches, la détection des points inusuels s’effectue à partir de tests statistiques d’adéquation des points par rapport à la distribution. Cette distribution est définie pour chaque point à partir d’une estimation de densité conditionnelle par noyau gaussien calibrée sur des points normaux. La distribution estimée correspond alors au modèle de normalité sur lequel nous nous basons pour détecter les points atypiques. Le score de détection correspond à la p-valeur du test statistique dans le cas du voisinage direct et la p-valeur minimale suivant les différentes orientations pour le voisinage directionnel. L’anomalie est détectée lorsque ce score passe sous le seuil de détection. Un filtrage par voisinage complète l’approche pour filtrer les

points isolés faussement détectés.

Les résultats produits par cette approche sont très intéressants et permettent la détection des signatures inusuelles, mais également des raies liées à l'arbre BP (N_1). Les signatures liées à l'arbre HP et le bruit sont faiblement détectées. Ces deux approches restent tout de même sensibles aux intensités vibratoires fortes du patch. La prise en compte du voisinage directionnel permet de considérer des points pertinents vis à vis de la structure des signatures atypiques. Cependant elle entraîne également la détection des points voisins pouvant masquer la structure des signatures inusuelles, mais permettant la détection des zones du patch contenant des signatures atypiques.

La détection à partir du voisinage directionnel permet de trouver l'orientation principale de chaque point à partir de la densité conditionnelle minimale associée. Ainsi en analysant les composantes connexes des points détectés, il est possible d'obtenir l'histogramme des directions des points de ces composantes. Ces histogrammes obtenus caractérisent donc les structures connexes correspondant à l'ensemble ou une partie des signatures détectées. Pour de futurs travaux, il est possible d'étudier cet indicateur pour classifier les différentes signatures détectées.

Les résultats obtenus par cette approche sont donc très prometteurs et permettent une aide à la décision pour l'analyse des spectrogrammes vibratoires. Il est également possible d'utiliser la densité jointe du point et de son voisinage comme modèle de normalité plutôt que la densité conditionnelle.

Conclusions et perspectives

Conclusions

L'objectif de cette thèse a été de mettre en place un outil d'aide à la décision automatique pour la détection des signatures atypiques sur les spectrogrammes vibratoires des moteurs d'avions. Nous avons aborder cette étude en considérant la représentation de la normalité au sein de ces spectrogrammes. Les signatures vibratoires sont représentées par des raies sur les spectrogrammes. Les éléments à détecter sont le plus fréquemment caractérisés par la présence de raies non présentes sur les spectrogrammes normaux. A l'heure actuelle, les spectrogrammes sont analysés visuellement et annotés manuellement par les experts de Safran Aircraft Engines dans un format image des données.

Notre travail a débuté dans un premier temps par associer les spectrogrammes numériques aux annotations manuelles des experts (format image) afin de construire une base de données numérique de spectrogrammes annotés. Les annotations correspondant aux différentes zones atypiques présentes sur les spectrogrammes et contenant une signature inusuelle. Un aboutissement de ce travail a été la consolidation d'une base annotée unique, regroupant l'expertise issue des analyses précédentes dans un format numérique. Nous avons également catégorisé différents points d'une soixantaine de spectrogrammes en 4 différentes classes, les points situés sur les signatures vibratoires normales, les points situés sur les signatures atypiques, les points correspondant à du bruit et les points situés sur des signatures vibratoires normales mais décalés sur les différents spectrogrammes. Ceci a permis de créer une seconde base, plus restreinte que la première mais plus fine également (la première contenant une annotation par zone), contenant une annotation ponctuelle des spectrogrammes. Cette base a été utilisée pour présenter les résultats, mais également pour paramétriser les seuils de décision.

Les spectrogrammes correspondent à des données en très grande dimension où chaque point correspond à la mesure vibratoire à une certaine fréquence f et un certain régime N_2 . Les signatures atypiques au sein de ces spectrogrammes consistent en une partie du spectrogramme de dimension bien plus petite que ce dernier. Analyser les spectrogrammes dans leur globalité ne permet pas de mettre en évidence les potentielles signatures inusuelles car ces dernières se retrouvent noyées par les informations normales présentes dans les données. Ces dernières sont structurées avec des informations normales identiques physiquement sur les différents spectro-

grammes. Sur l'intégralité du spectrogramme, le rapport points atypiques sur points normaux est très faible, ne permettant pas une analyse globale du spectrogramme. Nous avons donc subdivisé le spectrogramme en patchs définis par des plages de fréquences et de régimes, sur lesquels les différentes approches de détection d'anomalies ont été appliquées de manière indépendante. La proportion d'information inusuelle potentiellement présente sur un patch est bien plus importante qu'à l'échelle du spectrogramme. Cette approche est similaire à l'analyse visuelle des experts analysant les spectrogrammes par plage de fréquences. Cependant le nombre de données atypiques au niveau de ces différents patchs n'est pas suffisant pour mettre en place des modèles d'anomalies sur chacun d'entre eux. De plus les signatures inusuelles présentes sur les différents patchs sont trop variables pour pouvoir caractériser chaque type de signature par un modèle mathématique précis. Nous nous sommes donc portés sur des approches de détection de nouveautés, détection d'anomalies qui permettent de classifier des données lorsqu'un des labels des données (généralement normal) est fortement prédominant dans la base de données. Nous avons donc caractérisé les comportements normaux de ces patchs, c'est-à-dire les signatures vibratoires normales.

Nous avons étudié la normalité au sein des patchs sous différents aspects. Dans un premier temps, globalement sur le patch en mettant en place des modèles de normalité caractérisant l'intégralité des patchs à partir des approches par dictionnaire. Puis, dans un second temps localement en définissant un modèle de normalité pour chaque point des patchs à travers la distribution des points normaux. La caractérisation de cette normalité permet de bonnes performances de détection des signatures atypiques sans aucun a priori sur ces dernières répondant ainsi à la problématique énoncée. Ces méthodes ont donné lieu à un programme informatique (en MATLAB) d'outil d'aide à la décision pour analyser les spectrogrammes sans a priori sur les signatures atypiques. Ce programme fonctionne en prenant en compte les résultats des différentes méthodes mises en place au cours de cette thèse.

Représentation globale par dictionnaire

La représentation de la normalité des patchs des spectrogrammes a été étudiée pour deux dictionnaires différents. Le premier correspond à un dictionnaire fixe (non-adaptatif comme les atomes du dictionnaire sont définis par des fonctions et restent inchangés par rapport aux données), les curvelets. La normalité dans ce dictionnaire est caractérisée par une sélection des atomes des curvelets activés régulièrement sur les patchs normaux. La reconstruction du patch est effectuée uniquement à partir des coefficients associés à ces atomes. Le dictionnaire des curvelets caractérise localement les signatures sur le patch. Le second dictionnaire est cette fois-ci data-driven (adaptatif avec des atomes dépendants des données), la NMF. Les atomes du dictionnaire sont directement appris sur les données normales. Ils caractérisent globalement les signatures normales des patchs uniquement dues aux propriétés linéaires et positives de ce dictionnaire. La dimension du dictionnaire est calibrée sur les données normales comme le rang à partir duquel le gain de reconstruction n'est plus significatif. Ces deux dictionnaires donnent une représentation normale des patchs qui permet une reconstruction de ces derniers sans signature atypique. Il

s’agit alors d’une estimation normale des patchs.

Les reconstructions à partir de ces dictionnaires caractérisant la normalité donnent des estimations des caractères normaux des patchs des spectrogrammes. Les signatures atypiques sont grossièrement reconstruites à partir du dictionnaire des curvelets dus aux atomes basses fréquences avec des artefacts autour des signatures atypiques. Pour le dictionnaire appris de la NMF, les signatures inusuelles ne sont pas du tout reconstruites. Les approches par dictionnaire permettent de représenter les signatures atypiques au niveau des résidus de reconstruction. Il s’agit des points mal représentés par les dictionnaires et correspondent alors aux valeurs extrêmes positives au sein des résidus. Ces approches entraînent des fausses détections pour les deux dictionnaires qui sont limitées par l’application du filtrage de voisinage tenant compte de la décision au niveau des points voisins. Les signatures vibratoires liées à l’arbre basse pression sont en partie également détectées du fait de leur décalage sur les différents spectrogrammes. La combinaison des deux dictionnaires permet une meilleure considération visuelle des signatures atypiques.

La représentation de la normalité des patchs dans des dictionnaires avec une décision basée sur l’erreur de reconstruction qui en est issue permet une bonne mise en évidence des signatures inusuelles. Les différentes signatures atypiques sont détectées indépendamment du type de signatures. Les approches par dictionnaires permettent également un débruitage des spectrogrammes au niveau de la reconstruction.

La détection est effectuée à partir des valeurs extrêmes des résidus intra-patch sans tenir compte du même patch des autres données. Cela entraîne une détection systématique de points sur tous les patchs indépendamment de la présence de signatures atypiques sur ces derniers.

Représentation ponctuelle

La représentation des patchs correspond à la distribution estimée à partir des points normaux pour chaque point du patch. Cette distribution correspond alors au modèle de normalité établie pour chaque point. L’estimation de la distribution est effectuée de manière non paramétrique à partir des méthodes à noyau. L’utilisation des méthodes non paramétriques permet de ne donner aucun a priori sur la distribution des points. Trois différentes approches ont été étudiées dans ce cadre. La première considérant tous les points du patch comme mutuellement indépendants. Cependant les points voisins des spectrogrammes sont corrélés et deux points appartenant à la même raie vibratoire sont dépendants. La seconde approche tient donc compte uniformément du voisinage du premier ordre autour du point étudié. La dernière approche considère le voisinage du point sous différentes directions afin de tenir compte des voisins liés à une certaine orientation uniquement. Lors de l’étude, nous avons étudié le noyau gaussien permettant une simplification des calculs mais attribuant une densité non nulle aux points négatifs (impossible pour les intensités vibratoires) et le noyau gamma, permettant de ne considérer que les points positifs mais complexifiant les calculs. Les résultats à partir des deux noyaux sont proches dans le cas indé-

pendant. Nous avons donc opté pour le noyau gaussien dans les approches conditionnelles pour la simplification des calculs.

Les points atypiques correspondent à des intensités vibratoires importantes par rapport aux mêmes vibrations exactement sur les autres données. Il s'agit donc des valeurs extrêmes de la distribution en chaque point. La détection est effectuée à partir de la p-valeur d'un test statistique d'adéquation à la loi estimée dans les cas indépendant et tenant compte uniformément du voisinage. La décision est donnée pour chaque point du spectrogramme comparativement à un seuil de décision. Dans le cas du voisinage directionnel, un test statistique d'adéquation est réalisé suivant chacune des directions, et donc a fortiori une p-valeur est calculée pour chacune d'entre elles. La décision est établie en considérant la p-valeur minimale suivant ces différentes directions. Il suffit donc de détecter une irrégularité sur une seule direction pour que le point soit considéré comme atypique. Ces différentes approches entraînent également de fausses détections qui sont réduites par l'application du filtrage de voisinage. Les signatures vibratoires liées à l'arbre basse pression sont également détectées.

Ces méthodes définies localement permettent d'obtenir de bons résultats de détections des points atypiques avec une mise en évidence de ces signatures. De plus, la considération du voisinage sous différentes directions permet d'associer une orientation aux points détectés correspondant à la vraisemblance minimale du point par rapport aux différents voisinages. L'histogramme des directions des composantes connexes de points détectés permet d'avoir une caractérisation des signatures atypiques pour possiblement les classifier par la suite.

Ces approches sont plus sensibles aux intensités vibratoires que les approches par dictionnaire. Cela se voit aussi bien visuellement que numériquement avec une détection plus importante des points issus des raies N_1 . Cependant, comme la détection s'effectue par comparaison avec les mêmes points des autres données, il n'y a pas de détection systématique pour chaque patch des spectrogrammes à la différence des approches par dictionnaire.

Fusion des approches

Les différentes approches proposées, aussi bien globalement sur les patchs avec les dictionnaires que localement à partir de la modélisation ponctuelle des points, permettent toutes de détecter efficacement les signatures atypiques sur les patchs des spectrogrammes. Les points faussement détectés engendrés sont dépendants des différentes approches.

Afin de les limiter, nous avons associé les détections des différentes approches par un système de vote. Un point est considéré comme atypique si plus de la moitié des approches le considèrent comme tel. Cette fusion permet de réduire les fausses détections tout en consolidant les bonnes détections. La combinaison des différentes approches sous la forme de vote permet donc d'obtenir un bon outil de détection automatique des signatures atypiques.

Cet outil permet la détection des signatures atypiques, y compris celles ayant des intensités faibles. Cependant les signatures inusuelles ayant des intensités vibratoires équivalentes au bruit sont difficilement détectables car difficilement observables. De plus, les signatures normales liées au N_1 restent détectées à tort. Les différentes approches sont sensibles aux intensités vibratoires et à la position des signatures. Pour le moment, nous détectons uniquement la présence des signatures atypiques sur les patchs, nous ne les classifions pas selon les différents types.

Perspectives

Les principales fausses détections dans les méthodes développées concernent les signatures liées au N_1 qui sont décalées sur les différents spectrogrammes. Du fait de ce décalage, les dictionnaires et les approches ponctuelles caractérisent mal ce type de signature, ce qui entraîne sa détection.

Une première approche permettant de palier le problème lié au décalage des raies N_1 est d'effectuer un prétraitement sur les données à travers un clustering du comportement de la relation entre le N_1 et le N_2 sur le spectrogramme. Ainsi chaque spectrogramme au sein des clusters possèderait des relations N_1/N_2 proches et donc des raies N_1 proches. Les méthodes mises en place seraient alors appliquées dans chaque cluster séparément limitant ainsi les fausses détections liées au N_1 . Cette approche nécessite cependant une base de données plus conséquente que la notre afin d'établir un clustering pertinent des spectrogrammes.

Une seconde approche serait d'effectuer la même étude sur des spectrogrammes échantillonnés en N_1 (au lieu du N_2). Sur ces spectrogrammes les signatures liées au N_1 ne se décalent plus et ne sont donc pas détectées. Les signatures N_2 ont de fortes chances d'être détectées à tort dus à leurs décalages. En fusionnant les deux types d'approches en associant les points détectés du spectrogramme en N_2 avec les points détectés du spectrogramme en N_1 , il est possible de définir un point atypique comme un point détecté sur les deux types de spectrogrammes.

Tout au long de cette thèse, nous avons étudié les spectrogrammes par patch en comparant exactement les mêmes patchs sur les différents spectrogrammes afin d'en respecter la physique. Il est possible d'augmenter les données en considérant les patchs comme des données indépendantes avec un recouvrement de 50% des patchs dans la subdivision. Les labels "normal" et "atypique" sont alors récupérés pour chaque patch à partir de l'extraction des zones atypiques de la base de données. En considérant l'ensemble de ces patchs comme indépendants, nous nous retrouvons avec une grande volumétrie de données pouvant être considérées comme des images de taille 128×128 avec des labels. A but exploratoire, nous avons testé un réseau de neurones 5 couches [66] pour classifier les patchs et nous avons comparé les taux d'erreurs de classification à un SVM appliqué aux indicateurs images HoG (Histograms of Gradients) [36], SIFT (Scale-Invariant Feature Transform) [76] et LBP (Local Binary Patterns) [86] produisant des descripteurs en grande dimension des points de l'image en fonction de leur voisinage. Sur une base de données

de tests, nous avons obtenu un taux de mauvaise classification d'environ 20% pour les réseaux de neurones et de 25% pour le SVM. Nous n'avons pas poussé les études plus loin. Il est encore possible d'améliorer les résultats à travers des modifications des paramètres du réseau ou des structures de réseau plus complexes. Les principaux défauts de ce type d'approche sont la non-prise en compte de la physique en considérant les patchs d'un même spectrogramme comme des images indépendantes, ainsi que leur aspect "boîte noire".