

Estimation théorique de la sévérité d'un défaut par un modèle analytique de la *DKL*

Jinane HARMOUCHE^{1,2}, Claude DELPHA¹, Demba DIALLO²,

¹Laboratoire des Systèmes et Signaux (CNRS, Supélec, Univ. Paris-Sud)
Supélec, 3 rue Joliot Curie, 91192 Gif-sur-Yvette, France

²Laboratoire de Génie Electrique de Paris (CNRS, Supélec, Univ. P. et M. Curie, Univ. Paris-Sud)
11 rue Joliot Curie, 91192 Gif sur Yvette, France

Jinane.Harmouche@supelec.fr, Claude.Delpha@lss.supelec.fr
Demba.Diallo@lgp.supelec.fr

Résumé – Ce papier propose la Divergence de Kullback-Leibler (*DKL*) pour détecter et estimer des défauts naissants de petites amplitudes (<1%) par une approche globale. Dans cette approche, l'indicateur de défaut est vu comme une mesure d'information et non plus de résidus. Une analyse théorique de l'expression de la divergence conduit à un modèle analytique ne dépendant que du paramètre du défaut. Partant du modèle établi, la *DKL* permet une estimation théorique de la sévérité du défaut. L'application sur un exemple numérique montre que ce modèle garantit une borne supérieure de l'estimation de la sévérité du défaut. Il fournit alors une marge de sécurité sur l'estimation du défaut, souvent nécessaire à des fins de surveillance.

Abstract – This paper addresses the problem of incipient fault detection and estimation. An informational measure, the Kullback-Leibler Divergence, is proposed as a fault indicator through a global approach. A theoretical analysis is conducted to establish an analytical model relating the divergence to the fault severity amplitude. The model allows a theoretical estimation of the fault amplitude depending on the divergence value. Once applied to a numerical example, the theoretical estimation turns out as an upper bound of the actual fault amplitude. While the usual statistics underestimate the fault severity, an upper bound is always necessary to guarantee a safety margin for the process.

1 Introduction

Beaucoup d'approches de diagnostic utilisent l'analyse en composantes principales (ACP) dans une démarche locale pour effectuer la détection d'un défaut [1, 2]. L'ACP, dans sa forme classique, se base sur les dépendances linéaires ou quasi-linéaires entre les m variables d'un processus pour définir, suite à une transformation orthogonale, l ($l < m$) nouvelles variables latentes, dites composantes principales, décrivant des directions principales non corrélées [3]. L'espace des variables est subdivisé en deux sous-espaces de dimensions réduites l et $(m - l)$, appelés sous-espace principal et résiduel. Ensuite, le suivi du fonctionnement du processus se fait en utilisant des tests statistiques calculés dans l'un ou/et l'autre des sous-espaces.

Dans ce cadre, les statistiques \mathbf{T}^2 de Hotelling et l'erreur quadratique de prédiction (Squared Prediction Error SPE) sont les indices typiques utilisés respectivement dans le sous-espace principal et résiduel [4]. Cependant, la détection dans le sous-espace principal est d'autant plus critique et non satisfaisante, même avec l'indice angulaire global [5], que celle dans le sous-espace résiduel, à cause des grandes variations normales qui masquent souvent les variations des projections des défauts. D'autre part, les cartes statistiques associées à ces indices ne permettent pas un suivi direct et exact de l'évolution de la sé-

vérité du défaut ; une lecture locale peut entraîner de fausses interprétations, à cause des fluctuations importantes des indicateurs.

Souvent, la signature du défaut est recherchée dans le sous-espace résiduel, le sous-espace principal étant concerné par les défauts sur les variables indépendantes [4]. Des approches de structuration des résidus par reconstruction des indicateurs, ont été proposées afin d'améliorer la sensibilité des indicateurs au défaut et de permettre sa localisation [6]. Cependant, il est bien établi que pour qu'un défaut soit totalement reconstruit par le SPE ou l'une de ses variantes, sa direction doit être totalement projetée dans le sous-espace résiduel [7]. Une perte de sensibilité à certaines parties du défaut aura lieu si la projection du défaut dans le sous-espace résiduel est partielle. Il en résulte une sous-estimation de l'amplitude du défaut.

L'approche globale consiste à opter pour le sous-espace de l'information, c'est-à-dire le sous-espace principal, afin d'améliorer la qualité de détection. Elle consiste à extraire la signature du défaut de l'ensemble de l'information contenue dans les données. Pour ce faire, on propose comme outil de détection une mesure probabiliste issue de la théorie de l'information, qui est la Divergence de Kullback-Leibler. Après application de l'ACP sur un exemple numérique, le calcul de la divergence a montré dans [8, 9] ses performances en termes de

détection et de localisation des défauts. Sa pertinence vient du fait que les distributions statistiques des composantes principales le long de leurs axes, sont révélatrices de l'information latente des données. Ainsi, l'apparition d'un défaut provoque des changements dans la distribution des composantes principales, pourtant inaperçus à l'oeil nu, mais qui sont détectables par la divergence. Suite à ces résultats, il nous semble intéressant de développer un modèle de la divergence, conduisant à une estimation théorique de l'amplitude du défaut.

2 Estimation de l'amplitude du défaut par le modèle théorique de la DKL

La Divergence de Kullback-Leibler (*DKL*) entre deux distributions de probabilité continues $f(x)$ et $g(x)$ d'une variable aléatoire x , est définie comme étant la version symétrique de l'Information de Kullback-Leibler I entre ces deux distributions [10],

$$I(f||g) = \int f(x) \log \frac{f(x)}{g(x)} dx, \quad (1)$$

et

$$DKL(f, g) = I(f||g) + I(g||f). \quad (2)$$

Elle est définie si $\int_x f = 1$, $\int_x g = 1$ et si les deux distributions partagent le même support. Dans le cas de distributions normales, $f \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $g \sim \mathcal{N}(\mu_2, \sigma_2^2)$, il est possible de montrer que la *DKL* s'écrit :

$$DKL(f, g) = \frac{1}{2} \left[\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} + (\mu_1 - \mu_2)^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) - 2 \right]. \quad (3)$$

Dans notre approche, l'hypothèse de normalité sera appliquée sur les composantes principales. Cette hypothèse reposera sur le fait que la solution au problème de l'ACP est donnée sous forme de décomposition en valeurs et vecteurs propres de la matrice de covariance. Les composantes principales, obtenues par combinaison linéaire des variables, sont alors également à distributions normales. Cependant, dans notre étude, la divergence ne sera pas appliquée aux composantes décrivant l'espace résiduel : celles-ci risquent d'avoir leurs distributions dégénérées. A partir de ces hypothèses, nous nous appuyons sur (3) pour établir le modèle analytique de la divergence en fonction des paramètres du défaut, à partir des composantes du sous-espace principal.

2.1 Modèle analytique proposé

Considérons la matrice $X_{[N \times m]}$ des données collectées sur un système à contrôler, tel que $\bar{X} = (\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_m) = (x_{ij})_{i,j}$, où $\mathbf{x}_j = [x_{1j} \dots x_{Nj}]^T$ est le vecteur colonne de N mesures sur la $j^{\text{ème}}$ variable. La matrice des données centrées (colonnes à moyennes nulles) est désignée par $\bar{X}_{[N \times m]}$. La $j^{\text{ème}}$ colonne de \bar{X} s'écrit $\bar{\mathbf{x}}_j = \mathbf{x}_j - \mu_j$, où μ_j désigne la moyenne du vecteur \mathbf{x}_j . Soit S la matrice de covariance des données. Soit $P_{[m \times l]}$ la matrice des vecteurs propres générant le sous-espace principal de dimension l , $P = (p_1, \dots, p_l)$. Soit

$\lambda_1, \dots, \lambda_l$ tel que $\lambda_1 > \lambda_2 > \dots > \lambda_l$ les valeurs propres associées, aussi définies comme étant les variances des composantes principales. Considérons la matrice des composantes principales $T_{[N \times l]}$, où $T = (\mathbf{t}_1, \dots, \mathbf{t}_l) = \bar{X}P$. Le défaut affectant la variable \mathbf{x}_j est un biais additif dépendant de la variable \mathbf{x}_j et caractérisé par un facteur multiplicatif d'amplitude a affectant les mesures dans l'intervalle $[b, c]$. Un défaut simple est considéré, tel que la variable \mathbf{x}_r , $r \neq j$, est supposée sans défaut. Le vecteur des mesures de \mathbf{x}_j s'écrit alors $\mathbf{x}_j = \mathbf{x}_j^* + \mathbf{F}_j$, où l'étoile (*) désigne les mesures sans défauts; $\mathbf{x}_j^* = [x_{1j}^* \dots x_{Nj}^*]^T$ et $\mathbf{F}_j = a \times [0 \dots x_{bj}^* \dots x_{cj}^* \dots 0]^T$. La $j^{\text{ème}}$ colonne de la matrice centrée $\bar{X}_{[N \times m]}$ s'écrit, en passant par la moyenne du vecteur \mathbf{x}_j , tel que $\bar{\mathbf{x}}_j = \bar{\mathbf{x}}_j^* + \bar{\mathbf{F}}_j$ où $\bar{\mathbf{F}}_j = \mathbf{F}_j - a \times \frac{1}{N} \sum_{i=b}^c x_{ij}^* \mathbf{1}$, $\mathbf{1}$ est un vecteur colonne de N unités.

La normalité des données implique que la $i^{\text{ème}}$ ligne de \bar{X} est un vecteur aléatoire d'une distribution normale $\mathcal{N}(0, \Gamma)$. Un estimateur non biaisé de Γ est donné par $S = \frac{1}{N} \bar{X}^T \bar{X}$. D'autre part, et sous la même hypothèse, chaque composante principale \mathbf{t}_k , (avec $k = \{1, 2, \dots, l\}$), suit une loi normale $f \sim \mathcal{N}(0, \lambda_k)$. On propose de comparer f à sa référence f^* , telle que $f^* \sim \mathcal{N}(0, \lambda_k^*)$. Les deux distributions, dont les moyennes sont nulles, sont totalement décrites par leurs variances. La divergence est *a priori* dédiée à la détection des défauts de petites amplitudes ne modifiant pas le centre du sous-espace principal. Ainsi peut-on écrire $\lambda_k = \lambda_k^* + \Delta\lambda_k$, où $\Delta\lambda_k$ est le biais de variance dû à l'occurrence du défaut. D'après (3), la divergence entre f et f^* s'écrit :

$$DKL(f, f^*) = \frac{1}{2} \left[\frac{\Delta\lambda_k^2}{\lambda_k^*(\lambda_k^* + \Delta\lambda_k)} \right]. \quad (4)$$

Supposons ensuite que λ_k est une fonction du paramètre du défaut a , et que cette fonction est infiniment différentiable au voisinage de $a = 0$, le développement de Taylor de λ_k autour de $a = 0$ donne :

$$\lambda_k = \lambda_k^* + \frac{\partial\lambda_k}{\partial a} a + \frac{1}{2} \frac{\partial^2\lambda_k}{\partial a^2} a^2 + \frac{1}{3!} \frac{\partial^3\lambda_k}{\partial a^3} a^3 + \dots \quad (5)$$

La dérivée d'ordre n de la valeur propre associée au vecteur propre p_k^* de la matrice S , où $p_k^* = [p_{1k} \dots p_{mk}]^T$, s'exprime d'après [11] par :

$$\frac{\partial^n \lambda_k}{\partial a^n} = p_k^{*\top} \frac{\partial^n S}{\partial a^n} p_k^*. \quad (6)$$

Il s'agit enfin d'explicitier la matrice S en fonction du paramètre du défaut, et de calculer ensuite ses dérivées. Il en résulte que les dérivées d'ordre supérieur de S ($n > 2$) sont nulles, tandis que les dérivées première et seconde donnent :

$$\begin{cases} \frac{\partial\lambda_k}{\partial a} = p_k^{*\top} \frac{\partial S}{\partial a} p_k^* = \frac{2}{N} \left(p_{jk} \sum_{r=1}^m p_{rk} \delta_r + p_{jk}^2 \sigma \times a \right) \\ \frac{\partial^2\lambda_k}{\partial a^2} = p_k^{*\top} \frac{\partial^2 S}{\partial a^2} p_k^* = \frac{2}{N} p_{jk}^2 \sigma \end{cases} \quad (7)$$

$$DKL(f, f^*) = \frac{2}{N^2} \frac{\left(p_{jk} \sum_{r=1}^m p_{rk} \delta_r \times a + (3/2p_{jk}^2) \sigma \times a^2 \right)^2}{\lambda_k^* \left(\lambda_k^* + \frac{2}{N} (p_{jk} \sum_{r=1}^m p_{rk} \delta_r) \times a + \frac{3}{N} (p_{jk}^2) \sigma \times a^2 \right)} \quad (12)$$

ou δ_r , δ_j et σ sont des constantes calculées à partir des mesures sans défauts. Cependant, elles dépendent de l'intervalle $[b, c]$ que l'on suppose être connu.

$$\delta_r = \sum_{i=b}^c (x_{ir} - \mu_r) x_{ij}^* - \frac{1}{N} \sum_{i=b}^c x_{ij}^* \times \sum_{i=1}^N (x_{ir} - \mu_r), \quad (8)$$

$$\delta_j = \sum_{i=b}^c (x_{ij}^* - \mu_j^*) x_{ij}^* - \frac{1}{N} \sum_{i=b}^c x_{ij}^* \times \sum_{i=1}^N (x_{ij}^* - \mu_j^*) \quad (9)$$

$$\sigma = [N - (c - b + 1)] \left(\frac{1}{N} \sum_{i=b}^c x_{ij}^* \right)^2 + \sum_{q=b}^c \left(x_{qj}^* - \frac{1}{N} \sum_{i=b}^c x_{ij}^* \right)^2 \quad (10)$$

En appliquant (5) au cas considéré, on obtient :

$$\lambda_k = \lambda_k^* + \frac{2}{N} \left(p_{jk} \sum_{r=1}^m p_{rk} \delta_r \right) \times a + \frac{3}{N} (p_{jk}^2) \sigma \times a^2. \quad (11)$$

Enfin, l'expression de la divergence en fonction de l'amplitude du défaut est donnée à partir de (4) par (12).

Par conséquent, si l'on note $\alpha_1 = p_{jk} \sum_{r=1}^m p_{rk} \delta_r$ et $\alpha_2 = 3/2p_{jk}^2 \sigma$, l'estimation théorique de l'amplitude du défaut s'obtient à partir de la DKL et de son expression (12) par :

$$\hat{a} = \frac{-\alpha_1 + \sqrt{\alpha_1^2 + 2\alpha_2 N \lambda_k^* (DKL + \sqrt{(DKL^2 + 2DKL)})}}{2\alpha_2} \quad (13)$$

2.2 Validation et interprétation du modèle

Considérons un exemple numérique décrivant le fonctionnement d'un système [12] avec $m = 7$ variables explicatives, définies à l'instant i par les équations suivantes :

$$\begin{aligned} x_1(i) &\propto \mathcal{N}(0, 1), & x_2(i) &= -x_1(i) \\ x_3(i) &= x_1(i) - x_2(i), & x_4(i) &= x_1(i) - 3x_2(i) \\ x_5(i) &= 0.5x_2(i) + x_3(i), & x_6(i) &= x_1(i) - 4x_2(i) \\ x_7(i) &= 0.2x_2(i) + x_3(i) \end{aligned}$$

La matrice de données X formée de N observations donne une seule composante principale $t_1^* = \bar{X} p_1^*$, dont la direction est déterminée par le vecteur propre $p_1^* = [0.138 \ -0.138 \ 0.2761 \ 0.5521 \ 0.2070 \ 0.6901 \ 0.2484]^T$ et de variance $\lambda_1^* = 52.457$.

La Fig.1 trace l'évolution de la DKL théorique (12) et celle estimée (2) en fonction de l'amplitude du défaut $a \in [0, 1]\%$, tenant compte que le modèle théorique est dédié aux défauts naissants, c'est-à-dire de très faible amplitude. Différentes valeurs de N sont considérées et l'intervalle $[b, c]$ est ajusté tel que la taille de l'échantillon en défaut correspond à 1% de N .

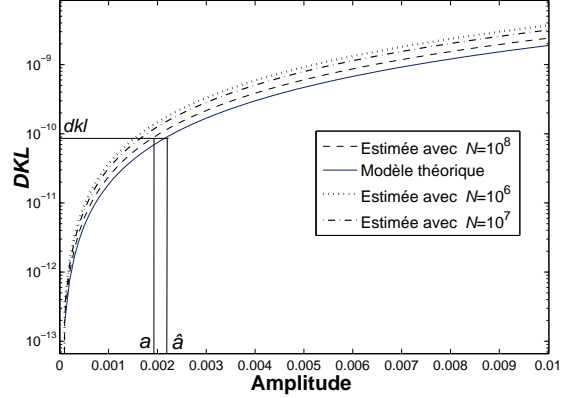


FIGURE 1 – Evolution de la DKL

La divergence estimée tend vers la divergence théorique quand N augmente. En effet et d'après le théorème de la limite centrale, l'hypothèse de la normalité devient valide pour N assez large. Dans ce cas, la divergence théorique donne une estimation exacte de la sévérité du défaut. Le modèle analytique détermine une surestimation (\hat{a}) de l'amplitude réelle (a) du défaut. En ce sens, cette borne supérieure du défaut représente une marge de sécurité.

L'erreur relative, notée E_r , d'estimation de a par la valeur théorique \hat{a} donnée par (13) s'écrit :

$$E_r = \frac{|a - \hat{a}|}{a} \quad (14)$$

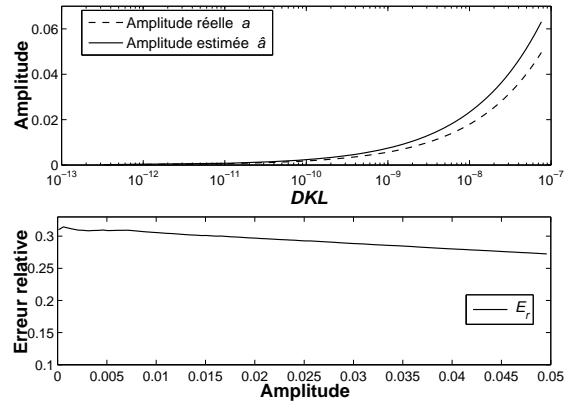


FIGURE 2 – Surestimation de l'amplitude du défaut et erreur relative correspondante ; cas 1% de N en défaut

Considérons 1 million d'échantillons, soit $N = 10^6$, que l'on peut effectivement obtenir avec une fréquence d'échantillonnage f_s de 10kHz pendant une durée d'acquisition de $1\text{mn}40\text{s}$. La taille de l'échantillon en défaut étant toujours de 1% de N , la Fig.2 trace l'amplitude réelle du défaut ainsi que

celle théoriquement estimée par la DKL , et l'erreur relative qui en résulte. L'erreur absolue $|a - \hat{a}|$ augmente presque proportionnellement à l'amplitude du défaut conduisant à un pourcentage d'erreur de $E_r = 30\%$.

Un pourcentage d'erreur de 30% peut-être jugé acceptable pour des amplitudes faibles du défaut ($a < 1\%$), qui est l'hypothèse sous laquelle (13) est établie. Cependant, l'erreur relative dépend fortement de la taille de l'échantillon en défaut : elle diminue quand la proportion de l'échantillon en défaut augmente. Pour une proportion allant de 1% jusqu'à 20% de N on trace sur la Fig.3 l'évolution de l'erreur relative commise sur l'estimation de l'amplitude du défaut.

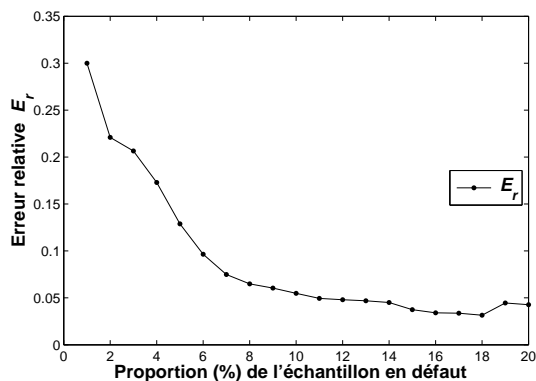


FIGURE 3 – Erreur relative de surestimation du défaut en fonction de la proportion (%) de l'échantillon en défaut.

Pour un échantillon en défaut qui représente 10% de l'échantillon total (une proportion souvent rencontrée dans la littérature), l'erreur relative de surestimation par la DKL tombe à 5%. Sur la Fig. 4 sont représentées pour ce cas les amplitudes réelle et celle estimée du défaut en fonction de la DKL pour $a \in [0, 5]\%$. Ainsi, grâce à la sensibilité de la DKL aux pe-

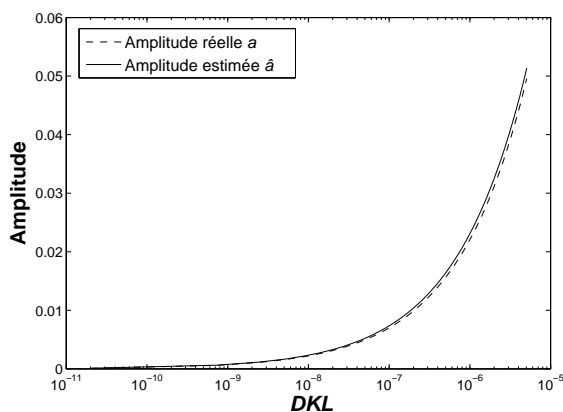


FIGURE 4 – Surestimation de l'amplitude du défaut par la DKL , cas 10% de N en défaut

tités variations et donc sa capacité à détecter des défauts naissants, une erreur relative de 0.05 sur l'estimation de l'amplitude

faible du défaut est considérée très satisfaisante. En plus, l'erreur, conduisant ici à une surestimation, représente une marge de sécurité du système en surveillance.

3 Conclusion

Contrairement aux indicateurs de défauts usuels, la Divergence de Kullback-Leibler permet l'extraction de la signature du défaut non plus de l'espace des résidus, mais de l'espace de l'information décrit par les données. La DKL montre une plus grande sensibilité aux défauts de faibles amplitudes vue quelle quantifie les distortions globales dues à l'occurrence du défaut. En se basant sur la divergence calculée dans le sous-espace principal défini dans le cadre de l'ACP, un modèle analytique de la DKL permet une estimation théorique de la sévérité du défaut. La valeur estimée constitue une surestimation satisfaisante de la sévérité actuelle du défaut : tenant compte de la sensibilité de la DKL aux défauts naissants, l'erreur de surestimation est jugée faible et constitue une marge de sécurité intéressante pour la surveillance.

Références

- [1] J. F. MacGregor and T. Kourti, "Statistical process control of multivariate processes," *Control Engineering Practice*, vol. 3, no. 3, pp. 403–414, 1995.
- [2] C. Delpha, D. Diallo, M.E.H. Benbouzid, and C. Marchand, "Application of classification methods in fault detection and diagnosis of inverter fed induction machine drive : A trend towards reliability," *European Physical Journal of Applied Physics*, vol. 43, pp. 245–251, 2008.
- [3] J.E. Jackson, *A User's Guide to Principal Components*, New York : Wiley, 1991.
- [4] U. Kruger and L. Xie, *Advances in Statistical Monitoring of Complex Multivariate Processes*, New York : Wiley, 2012.
- [5] A. Raich and A. Cinar, "Diagnosis of process disturbances by statistical distance and angle measures," *Comp. & Chem. Eng.*, vol. 21, pp. 661–673, 1997.
- [6] S.J. Qin, "Statistical process monitoring : basics and beyond," *Journal of Chemometrics*, vol. 17, pp. 480–502, 2003.
- [7] B. Mnassri, E.M. El Adel, S.J., and M. Ouladsine, "Unified sufficient conditions for pca-based fault detectability and isolability," in *8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, vol. 8, pp. 421–426, 2012a.
- [8] J. Harmouche, C. Delpha, and D. Diallo, "Faults diagnosis and detection using principal component analysis and kullback-leibler divergence," *IEEE IECON 2012*, pp. 3907–3912, 2012.
- [9] J. Harmouche, C. Delpha, and D. Diallo, "Incipient fault detection and diagnosis based on kullback-leibler divergence using principal component analysis : Part I," *Elsevier Signal Processing*, 2013, to be published.
- [10] S. Kullback and R.A. Leibler, "On information and sufficiency," *The Ann. of Mat. Stat.*, vol. 22, no. 1, pp. 79–86, 1951.
- [11] N.P. Van Der Aa, H.G. Ter Morsche, and R.R.M. Mattheij, "Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem," *Electronic Journal of Linear Algebra ELA*, vol. 16, pp. 300–314, Oct. 2007.
- [12] M.F. Harkat, G. Mourot, and J. Ragot, "Détection de défauts à l'aide d'une analyse en composantes principales robuste," *Sixième Conf. Int. Franc. d'Aut., CIFA*, Juin 2010.