

Un Nouveau Modèle Probabiliste pour l'Identification Modale sur un Diagramme de Campbell

Alexandre BEREZIN^{1,4} Yann ROTROU⁴ Jean-Yves TOURNERET^{2,3} François VINCENT¹

¹Laboratoire TéSA, 7 boulevard de la gare 31500 Toulouse

²INP-ENSEEIHRT IRIT, 2 Rue Charles Camichel 31000 Toulouse

³ISAE-Supaero, 10 avenue Marc Pégélin 31400 Toulouse

⁴Safran Aircraft Engines, Rond-point René Ravaud 77550 Moissy-Cramayel

Résumé – L'identification modale des machines tournantes constitue un enjeu majeur pour la conception industrielle, notamment dans le domaine aéronautique. Dans ce contexte, la présence d'harmoniques et la dépendance des modes au régime compliquent l'analyse des signaux. Cet article propose un modèle probabiliste original pour détecter les modes propres à partir d'une carte de détection construite sur un diagramme de Campbell. Deux variantes de l'algorithme EM sont développées : une version classique et une version à classification sous contrainte (CEM) garantissant l'unicité des affectations le long des harmoniques. Les performances sont évaluées à l'aide de données simulées et comparées à la méthode RANSAC. Les résultats montrent que l'approche proposée améliore la précision d'identification et la robustesse aux outliers.

Abstract – Modal identification of rotating machinery is a key challenge in industrial design, particularly in the aerospace sector. In this context, the presence of harmonics and the regime-dependent nature of modes complicate vibration signal analysis. This paper introduces a novel probabilistic model for identifying natural modes from a detection map constructed on a Campbell diagram. An Expectation-Maximization algorithm and a constrained classification EM (CEM) that enforces uniqueness of mode assignments along harmonics are investigated. The performance of the proposed methods is assessed on simulated and benchmark data and compared to the RANSAC algorithm. Results show significant improvements in identification accuracy and robustness to outliers.

1 Introduction

La maîtrise des vibrations est un enjeu crucial dans le domaine des essais aéronautiques. Des vibrations excessives peuvent engendrer des nuisances acoustiques ou mécaniques pour les passagers, altérer le fonctionnement de certains sous-systèmes, voire, dans les cas extrêmes, provoquer des défaillances structurales. Pour anticiper ces risques, le comportement dynamique des composants est d'abord modélisé par des simulations numériques, puis validé au moyen d'essais physiques. L'analyse des signaux vibratoires issus de ces essais est alors déterminante pour vérifier que le moteur fonctionne dans des plages de tolérance acceptables. Dans ce contexte, on propose une méthode permettant de faciliter l'analyse de ces données, en particulier à travers l'identification des modes propres du système à partir de mesures issues d'essais sur des moteurs en rotation.

L'analyse modale consiste à étudier les propriétés dynamiques d'un système dans le domaine fréquentiel, en particulier pour déterminer ses modes propres et leurs caractéristiques. L'analyse modale opérationnelle (OMA) se concentre plus spécifiquement sur ces propriétés lorsque les forces d'excitation ne sont pas connues [1]. Pour ce faire, divers algorithmes sont utilisés, aussi bien dans le domaine temporel que dans le domaine fréquentiel [2]. Toutefois, l'OMA repose sur l'hypothèse forte d'un bruit blanc en entrée, qui n'est pas respectée dans le cas des machines en rotation du fait de la présence d'harmoniques. Ces harmoniques ou ordres dont la fréquence est multiple de la fréquence de rotation proviennent notamment des tolérances mécaniques du système et de l'in-

teraction entre la machine et le fluide environnant [3]. Les algorithmes d'OMA peuvent confondre ces harmoniques avec des modes [4]. Pour surmonter ce problème, une approche récente consiste à extraire le signal associé à ces harmoniques qui balayent le spectre lors d'une montée en régime, puis à les utiliser comme données d'entrée pour l'identification modale. Regroupées sous le nom d'Order-Based Modal Analysis (OBMA), ces méthodes permettent d'éliminer les faux modes liés aux harmoniques [1].

Cet article présente une méthode permettant de créer une carte de détection inspirée de l'OBMA et étudie deux approches probabilistes pour l'identification modale : un algorithme Expectation-Maximization (EM) basé sur une estimation des responsabilités [5] et une variante appelée Classification EM (CEM) [6] reposant sur une affectation optimale contrainte. L'utilisation conjointe de la carte de détection et de l'algorithme EM permet la détection des modes dans un système en rotation soumis à des excitations harmoniques prenant en compte l'évolution de la fréquence propre des modes.

La section 2 introduit le problème de l'identification modale en décrivant la construction de la carte de détection et le nouveau modèle probabiliste proposé pour l'estimation modale. La section 3 détaille les algorithmes EM et sa variante contrainte CEM permettant l'estimation modale. La section 4 présente les résultats expérimentaux obtenus à partir de données simulées, comparant les performances des approches proposées à celles de la méthode RANdom SAmple Consensus (RANSAC). Enfin, la section 5 conclut et propose des pistes d'amélioration et d'extension de ce travail.

2 Modélisation probabiliste

2.1 Carte de détection

Les modes propres d'une machine en rotation dépendent en première approximation uniquement du régime. Une représentation pertinente pour l'analyse des systèmes en rotation est donc une représentation dans un plan régime moteur (N en tour/min) - fréquence de vibration (f en Hz). Ce type de représentation, appelé diagramme de Campbell, représente les modes par des courbes et les harmoniques par des droites obliques. On peut modéliser les signaux mesurés comme étant la sortie de filtres résonants inconnus et variant dans le temps que l'on cherche à identifier, excités par le faisceau des harmoniques du régime moteur. Ainsi, lors d'une analyse fréquentielle, comme représenté sur la figure 1, les pics de résonance se trouvent aux intersections entre le faisceau de droites harmoniques connu et les droites quasi-horizontales caractéristiques des modes propres que l'on cherche. Deux algorithmes sont utilisés dans l'étape de création d'une carte de détection : le suivi d'ordre et la détection de pics. Le suivi d'ordre consiste à extraire l'amplitude d'un ordre dans une représentation amplitude - temps. Ce suivi d'ordre peut être effectué de différentes manières en connaissant le régime N , la plus simple étant d'extraire la norme de la FFT le long de l'harmonique considérée. D'autres méthodes existent comme la méthode du maximum de vraisemblance [7] ou le suivi d'ordre par la méthode Vold Kalman [8] qui sera utilisée dans cet article. Le deuxième algorithme effectue une détection de pics en analysant les variations du signal extrait par suivi d'ordre (e.g., en utilisant la fonction `peak_finder` de `scipy` [9]).

On peut ensuite représenter une "carte de détection" qui résume les informations extraites d'un capteur et qui sera l'entrée du modèle probabiliste. Un exemple de carte de détection est donnée sur la figure 1 avec en rouge les points représentant les intersections détectées. On considère un ensemble de K harmoniques $f = kN/60, \forall k = 1, \dots, K$, la carte de détection est un ensemble de points qui se répartissent sur chaque harmonique. Le k ème harmonique possède un ensemble de $N_k \geq 0$ pics détectés dont l'abscisse est notée $(n_{k,1}, \dots, n_{k,N_k}) \in \mathbb{R}^{N_k}$. Les modes propres sont quant à eux représentés par un ensemble de droites affines d'équations inconnues :

$$D_m : f = a_m N + b_m, \forall m = 1, \dots, M.$$

L'objectif de ce travail est d'estimer les coefficients caractéristiques (a_m et b_m) de ces droites à partir des détections précédentes, comme représenté par les droites rouges sur la figure 1.

2.2 Vraisemblance

Pour définir une fonction de vraisemblance associée au problème d'analyse modale, on introduit une variable "latente" pour chaque point $n_{k,i}$ notée $z_{k,i} \in \{0, \dots, M\}$ indiquant la classe de ce point :

- $z_{k,i} = m$ signifie que le i ème point détecté sur l'harmonique k d'équation $f = kN/60$ correspond au mode propre d'équation D_m ,
- $z_{k,i} = 0$ signifie que le point $n_{k,i}$ est un outlier qui ne correspond à aucun mode.

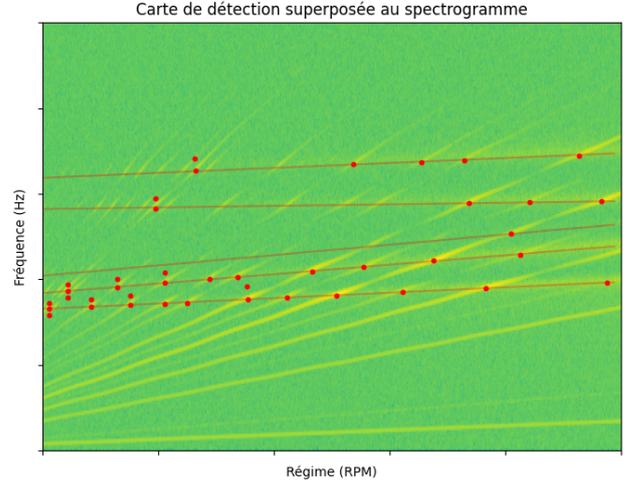


FIGURE 1 : Exemple de carte de détection superposée au spectrogramme d'un signal réel. Les modes propres correspondent aux droites en rouge, les harmoniques aux droites obliques et les points en rouge aux $n_{k,i}$ issus de la carte de détection. Les unités sont masquées pour des raisons de confidentialité.

On notera que si $z_{k,i} = m$, on a $f_{k,i} = a_m n_{k,i} + b_m$ et $f_{k,i} = kn_{k,i}/60$, soit $n_{k,i} = b_m / (k/60 - a_m)$. Ainsi en supposant une loi de probabilité Gaussienne sur les erreurs de détection, on peut calculer la densité de probabilité d'un point en fonction des paramètres du modèle :

$$n_{k,i} | z_{k,i} = m \sim \mathcal{N} \left(\frac{b_m}{k/60 - a_m}, \sigma_k^2 \right),$$

où σ_k^2 est une variance dépendant de l'algorithme de détection des pics et de l'harmonique considérée¹. Si $z_{k,i} = 0$, alors le point est tiré aléatoirement le long de l'harmonique, i.e., $n_{k,i} | z_{k,i} = 0$ suit une loi uniforme sur l'intervalle $]N_k^{\min}, N_k^{\max}[$, avec $L_k = N_k^{\max} - N_k^{\min}$ la "longueur" de l'harmonique mesurée dans la fenêtre.

On peut ainsi définir la vraisemblance complète du modèle :

$$L(N, \mathbf{Z} | \boldsymbol{\theta}) = \prod_{k=1}^K \prod_{i=1}^{N_k} p(n_{k,i}, z_{k,i} | \boldsymbol{\theta}),$$

Où $p(n_{k,i}, z_{k,i} | \boldsymbol{\theta})$ est la loi jointe des observations et des données cachées :

$$p(n_{k,i}, z_{k,i} | \boldsymbol{\theta}) = p(n_{k,i} | z_{k,i}, \boldsymbol{\theta}) p(z_{k,i} | \boldsymbol{\theta}).$$

Et où $\boldsymbol{\theta} = [\mathbf{a}, \mathbf{b}, \boldsymbol{\pi}]^T$ regroupe les paramètres du modèle avec $\mathbf{a} = (a_1, \dots, a_M)^T$, $\mathbf{b} = (b_1, \dots, b_M)^T$ et $\boldsymbol{\pi} = (\pi_0, \dots, \pi_M)^T$. On notera $\pi_m = P(z_{k,i} = m | \boldsymbol{\theta})$ la probabilité a priori que le point appartienne à la droite D_m et π_0 la probabilité qu'un point détecté soit un outlier avec $\sum_{m=0}^M \pi_m = 1$, d'où

$$p(n_{k,i}, z_{k,i} | \boldsymbol{\theta}) = \left(\frac{\pi_0}{L_k} \right)^{\mathbf{1}_{\{z_{k,i}=0\}}} \times \prod_{m=1}^M [p(n_{k,i} | \boldsymbol{\theta}, z_{k,i} = m) \pi_m]^{\mathbf{1}_{\{z_{k,i}=m\}}}.$$

¹On suppose que l'erreur de mesure se fait au niveau de la fréquence, c'est-à-dire que l'erreur suit une loi normale mais le long de l'axe f . Cette erreur doit être projetée sur l'axe N du plan. Une variance de σ^2 en fréquence conduit à une variance le long de l'axe N égale à $\sigma_k^2 = (60\sigma/k)^2$.

3 Algorithme EM

L'algorithme Expectation Maximization (EM) [5] est un algorithme itératif qui se compose de deux étapes complémentaires notées (E) et (M) qui sont répétées jusqu'à convergence des paramètres, assurant ainsi l'obtention d'une solution égale à un maximum local de la vraisemblance.

3.1 Étape E : Calcul des Responsabilités

L'étape E évalue, pour chaque observation $n_{k,i}$, la probabilité qu'elle soit générée par l'un des modes ou qu'elle constitue un outlier en fonction des paramètres à l'étape courante $\theta^{(t)}$. La responsabilité $\gamma_{k,i,m}^{(t)}$, qui représente la probabilité conditionnelle d'appartenance du point $n_{k,i}$ à la droite D_m (mode m), est définie par le théorème de Bayes :

$$\begin{aligned} \gamma_{k,i,m}^{(t)} &= p(z_{k,i} = m \mid n_{k,i}, \theta^{(t)}) \\ &= \frac{p(n_{k,i} \mid z_{k,i} = m, \theta^{(t)}) \pi_m^{(t)}}{\sum_{m=1}^M p(n_{k,i} \mid z_{k,i} = m, \theta^{(t)}) \pi_m^{(t)}}. \end{aligned}$$

3.2 Étape M

Une fois les responsabilités calculées, l'étape M vise à mettre à jour les paramètres du modèle en maximisant la vraisemblance complète. Cette étape se divise en deux parties :

Mise à jour des probabilités a priori π_m : Les coefficients π_m sont réajustés en fonction de la contribution relative de chaque mode dans l'ensemble des observations. On obtient :

$$\pi_m^{(t+1)} = \frac{1}{N_{\text{tot}}} \sum_{k=1}^K \sum_{i=1}^{N_k} \gamma_{k,i,m}^{(t)}, \quad \text{avec } N_{\text{tot}} = \sum_{k=1}^K N_k.$$

Cette mise à jour garantit que $\sum_{m=0}^M \pi_m^{(t+1)} = 1$ et reflète la proportion d'observations assignées à chaque mode.

Mise à jour des paramètres des modes (a_m, b_m) : Pour chaque mode m , nous cherchons à minimiser la fonction de coût pondérée :

$$\Phi_m(a_m, b_m) = \sum_{k=1}^K \sum_{i=1}^{N_k} k^2 \gamma_{k,i,m}^{(t)} \left(n_{k,i} - \frac{b_m}{k/60 - a_m} \right)^2.$$

La fonction Φ_m étant quadratique par rapport aux coefficients b_m , pour un paramètre a_m fixé, le problème se ramène à une régression par moindres carrés pondérée qui admet la solution

$$\hat{b}_m(a_m) = \frac{\sum_{k=1}^K \sum_{i=1}^{N_k} k^2 \gamma_{k,i,m}^{(t)} \frac{n_{k,i}}{k/60 - a_m}}{\sum_{k=1}^K \sum_{i=1}^{N_k} k^2 \gamma_{k,i,m}^{(t)} \frac{1}{(k/60 - a_m)^2}}.$$

La substitution de $\hat{b}_m(a_m)$ dans Φ_m permet de réduire le problème à une optimisation unidimensionnelle par rapport à a_m . La fonction à minimiser s'exprime alors par :

$$\Phi_m(a_m) = \sum_{k=1}^K \sum_{i=1}^{N_k} k^2 \gamma_{k,i,m}^{(t)} \left(n_{k,i} - \frac{\hat{b}_m(a_m)}{k/60 - a_m} \right)^2.$$

L'optimisation de cette fonction par rapport aux inconnues a_m est ensuite réalisée par une méthode numérique (fonction `minimize_scalar` de `scipy`).

3.3 Classification EM

En théorie, la carte de détection est telle qu'un point détecté correspond à au plus un mode. Cette contrainte va à l'encontre de l'indépendance du tirage des points dans le modèle de mélange précédemment décrit. Afin de garantir que deux pics détectés sur une même harmonique ne soient pas associés au même mode propre, on peut modifier la vraisemblance complète :

$$L(N, \mathbf{Z} \mid \theta) \propto \begin{cases} 0 & \text{si } \exists k, i, j \text{ tel que } z_{k,i} = z_{k,j} \neq 0 \\ \prod_{k=1}^K \prod_{i=1}^{N_k} f(n_{k,i} \mid \theta, z_{k,i}) f(z_{k,i} \mid \theta) & \text{sinon.} \end{cases}$$

Afin de prendre en compte ces contraintes, nous proposons une variante de l'algorithme EM, appelée Classification EM (CEM) [6], qui repose sur une affectation explicite de chaque point à un mode (ou à la classe des outliers). Plutôt que de calculer une espérance pondérée, comme dans la méthode EM classique, la méthode CEM choisit directement la meilleure affectation des observations en maximisant la vraisemblance. Pour cela, on résout pour l'harmonique k le problème suivant :

$$\hat{z}_k = \arg \max_{z_k} \prod_{i=1}^{N_k} p(n_{k,i} \mid \theta, z_{k,i}) p(z_{k,i} \mid \theta),$$

où $z_k = (z_{k,1}, \dots, z_{k,N_k})^T$. Cette formulation permet d'ajouter simplement notre contrainte et de transformer le problème en problème d'affectation. Ce problème peut être formulé comme un problème d'assignation dans un graphe biparti. D'un côté, on trouve les observations $n_{k,i}$ sur une harmonique, de l'autre les modes possibles (ainsi que plusieurs nœuds "outlier"). Le coût d'association $s_j^{k,i}$ entre le point $n_{k,i}$ et le mode j est défini par :

$$s_j^{k,i} = -\log[p(n_{k,i} \mid \theta, z_{k,i}) p(z_{k,i} \mid \theta)].$$

Le but est de trouver l'affectation qui minimise le coût total sous la contrainte d'unicité. On résout ce problème à l'aide de l'algorithme hongrois qui nous donne la partition de coût optimal \hat{z}_k . L'étape M reste identique à celle de l'algorithme EM classique : les paramètres θ sont mis à jour à partir des affectations obtenues. On note simplement que les responsabilités deviennent binaires : $\gamma_{k,i,j} = 1$ si $z_{k,i} = j$ et 0 sinon.

4 Expérimentations

La carte de détection générée simule de façon réaliste les intersections entre harmoniques et modes suivant les paramètres du modèle : on définit une certaine probabilité d'intersection entre les harmoniques et les modes ainsi qu'une erreur de mesure. Un certain nombre d'outliers est tirée de façon uniforme sur les harmoniques selon un ratio à ajuster. Pour estimer les paramètres des modes, les méthodes EM et CEM ont été comparées avec l'algorithme RANSAC [10], qui est une méthode robuste permettant d'identifier les modes en éliminant les outliers avec un système de vote. Nous comparons les erreurs quadratiques moyennes des paramètres à estimer (notées e_a , e_b et e_π), ainsi que les valeurs moyennes résiduelles des Log-vraisemblance, après convergence. Les erreurs sont normalisées par la norme des vecteurs de vérité correspondants, ce qui permet d'exprimer l'écart en proportion de la valeur réelle.

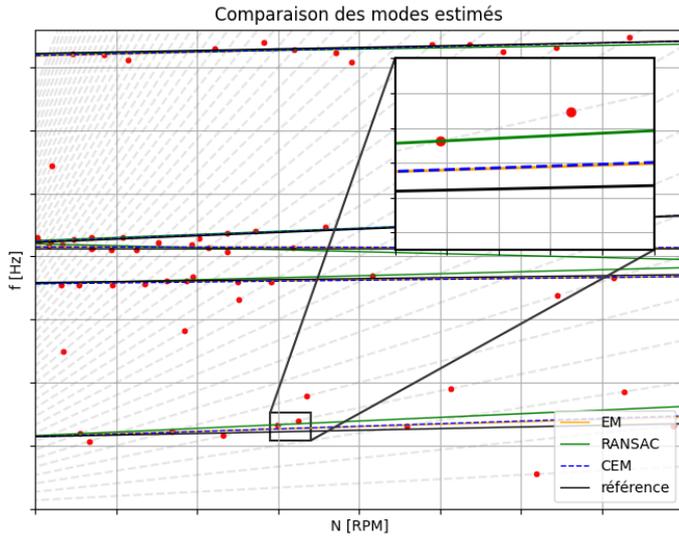


FIGURE 2 : Exemples de modes estimés avec les différentes méthodes.

TABLE 1 : Comparaison des performances entre RANSAC, EM et CEM pour différents ratios d’outliers (utilisant les RMSE relatifs). Le ratio correspond au nombre d’outliers rapporté au nombre de points réels. Les meilleurs résultats sont en gras.

Outl. ratio	Méthode	e_a	e_b	e_π
0.05	RANSAC	1.0546	0.0098	0.3844
	EM	0.3102	0.0097	0.1235
	CEM	0.3231	0.0093	0.1040
0.15	RANSAC	0.8430	0.0256	0.4073
	EM	0.5337	0.0169	0.2822
	CEM	0.5508	0.0156	0.1494
0.30	RANSAC	0.5669	0.0158	0.3573
	EM	0.2856	0.0091	0.2102
	CEM	0.3155	0.0096	0.1007
0.50	RANSAC	0.6271	0.2007	0.1580
	EM	0.4462	0.1932	0.1942
	CEM	0.4096	0.1987	0.1229

L’analyse des résultats présentés dans le tableau 1 met en évidence les performances supérieures des algorithmes EM et CEM par rapport à RANSAC, notamment en termes d’erreurs quadratiques moyennes sur les paramètres estimés pour des ratios d’outliers faibles à modérés (jusqu’à 0.5). Les méthodes EM et CEM parviennent à estimer les paramètres avec une grande précision, avec des erreurs nettement inférieures à celles de RANSAC. L’algorithme CEM se distingue par une meilleure robustesse dans la gestion des outliers, en particulier sur e_π , ce qui suggère que l’ajout de la contrainte d’unicité améliore la cohérence de l’affectation des modes.

5 Conclusion

Cet article présente un cadre probabiliste original pour l’identification modale à partir de signaux issus de systèmes en rotation, en particulier dans le contexte aéronautique. En combinant une représentation pertinente sous forme de carte de détection avec des approches probabilistes de type EM et CEM,

la méthode proposée permet d’identifier de manière robuste les modes propres malgré la présence d’outliers.

Les résultats expérimentaux montrent que cette approche surpasse les méthodes classiques de type RANSAC, notamment en termes de vraisemblance et d’erreurs quadratiques moyennes tant que le taux d’outlier est acceptable. En particulier, l’algorithme CEM, en intégrant une contrainte physique réaliste d’unicité sur les affectations permet d’avoir des erreurs réduites. Au-delà de ses performances, cette méthode offre une interprétation probabiliste directe de la carte de détection, un atout essentiel pour des applications industrielles où la fiabilité de l’information extraite est critique.

Plusieurs pistes d’amélioration sont envisagées. D’une part, l’utilisation de méthodes de type Constrained EM pourrait offrir des garanties supplémentaires de convergence et une meilleure exploration de l’espace des solutions. D’autre part, l’enrichissement du modèle par la prise en compte de l’amplitude des pics ou de la cohérence entre capteurs permettrait de renforcer encore la robustesse de l’identification, en particulier dans des environnements fortement bruités.

Références

- [1] G. Sternharz and T. Kalganova, “Current methods for operational modal analysis of rotating machinery and prospects of machine learning,” in *Proc. Conf. Exposition on Structural Dynamics*, (Houston, Texas, USA), pp. 155–163, Feb. 2020.
- [2] F. B. Zahid, Z. C. Ong, and S. Y. Khoo, “A review of operational modal analysis techniques for in-service modal identification,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 42, no. 8, p. 398, 2020.
- [3] M. P. Boyce, *Rotor dynamics*. Oxford, UK : Butterworth-Heinemann, 2012.
- [4] E. Di Lorenzo *et al.*, “Dynamic characterization of wind turbine gearboxes using order-based modal analysis,” in *Proc. Int. Conf. on Noise and Vibration Engineering (ISMA’2014) and Int. Conf. on Uncertainty in Structural Dynamics (USD’2014)*, (Leuven, Belgium), pp. 4349–4362, 2014.
- [5] T. K. Moon, “The expectation-maximization algorithm,” *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, 1996.
- [6] G. Celeux and G. Govaert, “A classification EM algorithm for clustering & two stochastic versions,” *Computational Statistics and Data Analysis*, vol. 14, no. 3, pp. 315–332, 1992.
- [7] A. Emelchenkov, M. Fontaine, Y. Grenier, H. Mahé, and F. Roueff, “Multifrequency highly oscillating aperiodic amplitude estimation for nonlinear chirp signal,” in *Proc. European Signal Processing Conference (EUSIPCO)*, (Lyon, France), 2024.
- [8] H. Herlufsen, S. Gade, H. Konstantin-Hansen, and H. Vold, “Characteristics of the Vold-Kalman order tracking filter,” in *Proc. Int. Conf. Acoustics, Speech, and Signal Process.*, (Istanbul, Turkey), 2000.
- [9] P. Virtanen *et al.*, “SciPy 1.0 : Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [10] K. G. Derpanis, “Overview of the RANSAC algorithm,” *Image Rochester NY*, vol. 4, no. 1, pp. 2–3, 2010.