COUPLED GAUSSIAN MIXTURES FOR MODAL ANALYSIS: EM INFERENCE AND CRAMÉR-RAO BOUNDS

Alexandre Berezin[1,2,3]*, Yann Rotrou*[2]*, Jean-Yves Tourneret*[3,1]*, François Vincent*[4]

[1] TéSA laboratory, Toulouse, France. alexandre.berezin@tesa.prd.fr
[2] Safran Aircraft Engines, Moissy-Cramayel, France. yann.rotrou@safrangroup.com
[3] INP-ENSEEIHT IRIT, Toulouse, France. jean-yves.tourneret@toulouse-inp.fr
[4] ISAE-SUPAERO, Toulouse, France. francois.vincent@isae-supaero.fr

ABSTRACT

Order-Based Modal Analysis estimates resonances at frequencies that are integer multiples of a rotating machine's speed. These resonances are represented as a cloud of frequency-versus-speed intersections revealing the natural modes of the mechanical structure. This paper shows that grouping these intersections can be cast as inference in a coupled affine Gaussian mixture model where each mode is represented by a straight line shared across all harmonic orders, while a uniform component captures outliers. A dedicated expectation maximisation (EM) algorithm is investigated for this model, estimating mixture weights in closed form and the other model parameters through a one-dimensional search. Cramér–Rao lower bounds are derived for the joint estimation of slopes, intercepts and mixing proportions in the proposed statistical model allowing performance of the estimators of the unknown parameters to be studied. Monte-Carlo simulations illustrate how the variances of EM estimates approach those bounds. Applied to data from an industrial turbomachine, the method extracts modal lines whose characteristics agree with historical benchmarks, despite strong deterministic harmonics and regime-dependent drifts.

Index Terms— Gaussian mixture models, expectation maximisation, Cramér–Rao lower bounds, order-based modal analysis.

1. INTRODUCTION

Time–frequency techniques have become a cornerstone of modern signal processing, providing a joint description of how spectral content evolves over time. These representations are routinely exploited to *detect salient patterns* such as horizontal ridges (quasi-stationary sinusoids), chirps or other sweeps, and more intricate modulations such as time-varying amplitude or frequency components [2]. Depending on the compromise between time and frequency resolutions, practitioners may favour the short-time Fourier transform, the continuous wavelet transform, or high-concentration bilinear distributions [9]. Such tools have been used successfully in many applications ranging from radar tracking and bio-acoustics to structural health monitoring and *operational modal analysis* [12, 13, 16].

Modal analysis leverages prior knowledge about a physical system subjected to vibrations to explain the measured response, with the objective of estimating key parameters such as the natural frequencies and damping ratios of its modes. For turbomachinery blades, accurately identifying these modes and their amplitudes is essential for monitoring R&T tests and for certification. Two well-known obstacles complicate this identification [10]: 1) Strong parasite harmonics: Shaft rotation induces deterministic responses at integer and sometimes fractional orders of the speed, masking the weaker natural modes, 2) Regime-dependent filtering: Blade dynamics vary with operating point making modal frequency drifts during tests with varying speeds. Traditional Operational Modal Analysis (OMA) is not adapted to rotating machinery, as it relies on the assumption that the input excitation is a broadband stochastic process. This assumption is violated in rotating systems, where the excitation is dominated by a small number of deterministic harmonics tied to the shaft rotation. Recent approaches developed for such systems aim to exploit the trajectories of these harmonic orders to infer potential resonances at their intersection with structural modes. These methods are collectively referred to as *Order-Based Modal Analysis* (OBMA) [14]. Our specific contribution is the post-processing step that turns OBMA's raw output, a scatter of discrete (frequency, order) points that are expected to align along *affine* curves corresponding to the natural modes of the structure. Extracting these curves amounts to a geometric clustering problem in the (N, f) (regime, frequency) plane, made difficult by measurement errors and outliers.

The OMA problem can be formulated as inference in a *coupled affine Gaussian mixture model* (C-AGMM), in which the distribution of observed frequencies for each harmonic is a mixture of Gaussian components and outliers. Each Gaussian component corresponds to a candidate for an intersection between an harmonic and a mode, and its parameters—namely, the slope and intercept of the affine relation $f = a_m N + b_m$—are shared across all harmonics, which introduces a constraint in the frequency regime analysis. This coupling enforces global consistency of the modal structure across the dataset. An additional outlier component accounts for spurious detections unrelated to the natural modes. The parameters of the C-AGMM are estimated with an expectation maximisation (EM) algorithm especially tailored to the affine coupling structure.

The contributions of this work are: 1) Formulation of OBMA clustering using a coupled affine Gaussian mixture model, 2) Derivation of an EM algorithm that exploits the affine coupling between modes to achieve computationally efficient parameter estimation and 3) Theoretical performance analysis based on the the Cramér–Rao lower bounds (CRLB), highlighting the influence of model parameters like the noise or the distance between modes for instance. The paper is organized as follows: Section 2 introduces a probabilistic model adapted to OBMA and investigates an EM algorithm to estimate its parameters. Section 3 derives the Fisher information matrix and the associated CRLBs. Section 4 validates the estimator against these bounds through Monte-Carlo simulations and a run-up test on an industrial turbofan.

2. ORDER-BASED MODEL ANALYSIS

2.1. Probabilistic model

To a first approximation, the natural modes of a rotating machine depend only on the operating regime, which suggests a relevant representation expressing the vibration frequency f (in Hz) versus the engine speed N (in rpm). In such a representation known as Campbell diagram, modes appear as (nearly) horizontal curves, while harmonics are characterized by oblique straight lines. Measurements from strain gauge can be modeled as the output of unknown resonant filters—corresponding to the system's modes—excited by the harmonic content of the rotating machine. These filters are time-varying with coefficients depending on the rotational regime, which itself evolves over time during a test. As a result, resonance peaks in a regime-frequency analysis are located at the intersections between the known harmonic trajectories and the quasi-horizontal modal lines that need to be identified, as illustrated by the white points in Fig. 1 that also contain outliers. The detection map built for each sensor gathers all the local maxima identified along the chosen harmonics. These maxima serve as candidate intersection points for the probabilistic model developed in this paper. Precisely, we consider a set of K harmonics defined by the frequencies $f = d_k N$, $k = 1, \ldots, K$ (with $d_k = k/60\,\text{Hz}\cdot\text{RPM}^{-1}$). For the kth harmonic, $N_k \geq 0$ peaks are detected at locations $(n_{k,1}, \ldots, n_{k,N_k}) \in \mathbb{R}^{N_k}$. The natural modes in the frequency-regime plane are modelled as M affine lines defined by:

$$D_m : \; f = a_m N + b_m, \qquad m = 1, \ldots, M.$$

The goal of the suggested OBMA model is to estimate the modal parameters (a_m, b_m) from the detected points associated with the white lines in Fig. 1.

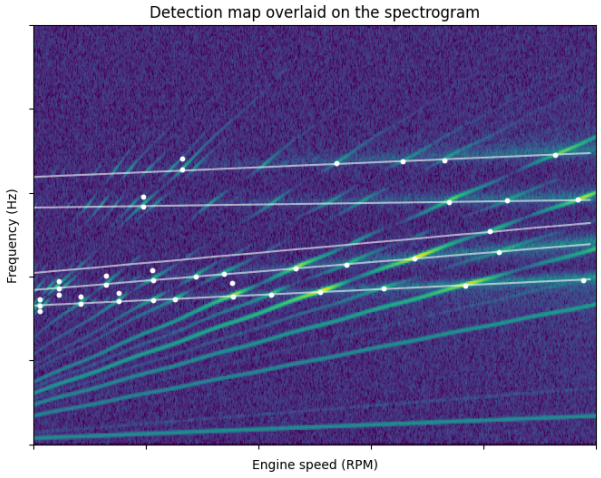

Fig. 1: Example of detection map superimposed on a simulated spectrogram. Horizontal white lines are the natural modes, oblique lines are the harmonics and white points are the detected local maxima $n_{k,i}$ along the hamonics. Units are hidden for confidentiality.

For each point $n_{k,i}$, a latent variable $z_{k,i} = 0, \ldots, M$ is introduced to indicate if the ith point on harmonic $\#k$ belongs to mode D_m ($z_{k,i} = m$ with $m = 1, ..., M$) or if the ith point is an outlier ($z_{k,i} = 0$). If $z_{k,i} = m$, then $f_{k,i} = a_m n_{k,i} + b_m$ and $f_{k,i} = d_k n_{k,i}$ leading to $n_{k,i} = b_m/(d_k - a_m)$. Assuming Gaussian detection errors, the following result is obtained:

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

where $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian distribution with mean μ and variance σ^2. Note that the variance σ_k^2 depends on the peak detection algorithm. If $z_{k,i} = 0$, the point is uniformly drawn along the kth harmonic, i.e. $n_{k,i} \mid z_{k,i} = 0 \sim \mathcal{U}(N_k^{\min}, N_k^{\max})$ where $N_k^{\min}$ and $N_k^{\max}$ are defined by the frequency-regime window. The complete likelihood for the suggested OBMA probabilistic model is:

$$L(\mathbf{n}, \mathbf{z} \mid \boldsymbol{\theta}) = \prod_{k=1}^{K} \prod_{i=1}^{N_k} p(n_{k,i}, z_{k,i} \mid \boldsymbol{\theta}),$$

with the joint distribution

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

and the unknown parameter vector $\boldsymbol{\theta} = [\mathbf{a}^\top, \mathbf{b}^\top, \boldsymbol{\pi}^\top]^\top$, where $\mathbf{a} = (a_1, \ldots, a_M)^\top$, $\mathbf{b} = (b_1, \ldots, b_M)^\top$, and $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_M)^\top$ with $\pi_m = P(z_{k,i} = m \mid \boldsymbol{\theta})$, $m = 1, ..., M$, π_0 is the outlier prior and $\sum_{m=0}^{M} \pi_m = 1$.

2.2. EM Algorithm

The EM algorithm is an iterative scheme for maximum–likelihood estimation when part of the data is latent or missing. Starting from an initial parameter vector $\boldsymbol{\theta}^{(0)}$, each iteration alternates between two complementary steps: the *E-step* that computes the expected complete-data log-likelihood

$$Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(t)}) = \mathbb{E}_{\mathbf{z}|\mathbf{n}, \boldsymbol{\theta}^{(t)}} \left[\log p(\mathbf{n}, \mathbf{z} \mid \boldsymbol{\theta})\right],$$

defined as an expectation with respect to the posterior of the latent-variables under the current parameters, and the *M-step* that updates the parameters by maximizing this expectation:

$$\boldsymbol{\theta}^{(t+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(t)}). \tag{1}$$

The E and M steps are repeated until convergence. Note that the observed-data likelihood is increasing at each iteration and that the sequence $L(\boldsymbol{\theta}^{(t)}) = p(\mathbf{n} \mid \boldsymbol{\theta}^{(t)})$ increases monotonically toward a local maximum of the model likelihood [3, 11]. The E and M steps for the suggested C-AGMM model are detailed below.

E step. For each observation $n_{k,i}$, the E-step computes the posterior probability that it was generated by each mode (or that it is an outlier) given the current parameters $\boldsymbol{\theta}^{(t)}$:

$$\begin{aligned} \gamma_{k,i,m}^{(t)} \;&=\; p\big(z_{k,i} = m \mid n_{k,i}, \boldsymbol{\theta}^{(t)}\big) \\ &= \frac{p(n_{k,i} \mid z_{k,i} = m, \boldsymbol{\theta}^{(t)})\, \pi_m^{(t)}}{\sum_{m'=0}^{M} p(n_{k,i} \mid z_{k,i} = m', \boldsymbol{\theta}^{(t)})\, \pi_{m'}^{(t)}}. \end{aligned}$$

M step. After computing the co-called responsabilities $\gamma_{k,i,m}^{(t)}$ and using the notation $N_{\text{tot}} = \sum_{k=1}^{K} N_k$, the priors π_m are updated as:

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

The maximisation of the Q function in (1) is equivalent to minimise:

$$\Phi_m(a_m, b_m) = \sum_{k=1}^{K} \sum_{i=1}^{N_k} \frac{\gamma_{k,i,m}^{(t)}}{\sigma_k^2} \Big(n_{k,i} - \frac{b_m}{d_k - a_m}\Big)^2, \tag{2}$$

allowing the modal parameters (a_m, b_m) to be estimated for each mode m. For a fixed a_m, the value of b_m minimising (2) is:

$$\hat{b}_m(a_m) = \frac{\displaystyle\sum_{k,i} \frac{\gamma_{k,i,m}^{(t)}}{\sigma_k^2} \frac{n_{k,i}}{d_k - a_m}}{\displaystyle\sum_{k,i} \frac{\gamma_{k,i,m}^{(t)}}{\sigma_k^2} \frac{1}{(d_k - a_m)^2}}.$$

Substituting $\hat{b}_m(a_m)$ into (1) leads to a 1D minimisation w.r.t. a_m that can be solved numerically (SciPy's `minimize_scalar`). Note that σ_k is not estimated since the uncertainty related to the n_{ki} values can be known (it depens on the detection method).

3. PERFORMANCE ANALYSIS

The Fisher information matrix (FIM) of a probabilistic model is:

$$\mathcal{I}(\boldsymbol{\theta}) \;=\; \mathbb{E}_{\mathbf{y}\sim p(\cdot|\boldsymbol{\theta})}\big[-\nabla^2_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta};\mathbf{y})\big] \;=\; \mathbb{E}\big[S(\boldsymbol{\theta})\,S(\boldsymbol{\theta})^\top\big],$$

where $\ell(\boldsymbol{\theta};\mathbf{y})$ is the model likelihood, $S(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta};\mathbf{y})$ is the score vector and $\boldsymbol{\theta} \in \mathbb{R}^p$ is the parameter vector to be estimated. For an unbiased estimator $\hat{\boldsymbol{\theta}}$, the variance of each individual parameter is bounded below by the corresponding Cramér–Rao lower bound (CRLB), which is the diagonal entry of the inverse FIM:

$$\mathrm{Var}\big(\hat{\theta}_i\big) \;\geq\; \big[\mathcal{I}^{-1}(\boldsymbol{\theta})\big]_{ii}, \qquad i = 1, \ldots, p,$$

Note that a steeper log-likelihood curvature (larger information) implies tighter confidence intervals around every component $\hat{\theta}_i$. To evaluate $\mathcal{I}(\boldsymbol{\theta})$, the score-outer-product identity can be used, which avoids the use of second-order derivatives and requires the pointwise score vectors to be computed. Because the incomplete log-likelihood $L(\mathbf{N} \mid \boldsymbol{\theta})$ is a sum of 1D Gaussian terms, computing its analytic gradient is straightforward. This computation is simplified by introducing an unconstrained vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)^\top$, which is mapped into the simplex with the soft-max transform

$$\pi_0(\boldsymbol{\eta}) = \frac{1}{Z}, \qquad \pi_m(\boldsymbol{\eta}) = \frac{e^{\eta_m}}{Z}\;, \qquad Z = 1 + \sum_{j=1}^{M} e^{\eta_j},$$

so that $\pi_m \in (0,1)$ and $\sum_{m=0}^{M} \pi_m = 1$ hold automatically. Using the relation $\gamma_{k,m} = \pi_m f_{k,m}/p_k$, the residuals $\Delta_{k,i,m} = n_{k,i} - \mu_{k,m}$ yield the following scores:

$$S_{\eta_r}(n_{k,i}) = \gamma_{k,r} - \pi_r,\; S_{a_m}(n_{k,i}) = \gamma_{k,m}\frac{b_m \Delta_{k,i,m}}{\sigma_k^2(d_k - a_m)^2},$$

$$S_{b_m}(n_{k,i}) = \gamma_{k,m}\frac{\Delta_{k,i,m}}{\sigma_k^2(d_k - a_m)}.$$

A summation over the N_k samples of each harmonic leads to:

$$I_{rs}(\boldsymbol{\theta}) = \sum_{k=1}^{K}\sum_{i=1}^{N_k} \mathbb{E}\big[S_r(n_{k,i})\,S_s(n_{k,i})\big],$$

whose computation is explained below.

Block $\boldsymbol{\eta}$ (mixture weights). The relation $\mathbb{E}_{p_k}[\gamma_r] = \pi_r$ yields:

$$I_{\eta_r,\eta_s} = N_{\mathrm{tot}}\big(\mathbb{E}[\gamma_r\gamma_s] - \pi_r\pi_s\big).$$

Block (a,b) and cross–term. Define the constants $C_{k,m} = b_m/[\sigma_k^2(d_k - a_m)^2]$, $D_{k,m} = 1/[\sigma_k^2(d_k - a_m)]$ and the two integrals

$$M_k(m,s) = \int \frac{\pi_m\pi_s\, f_{k,m}f_{k,s}}{p_k}(n - \mu_{k,m})(n - \mu_{k,s})\,\mathrm{d}n,$$

$$A_k(m,r) = \int \frac{\pi_m\pi_r\, f_{k,m}f_{k,r}}{p_k}(n - \mu_{k,m})\,\mathrm{d}n.$$

For each harmonic, the following results are obtained:

$$I^{(k)}_{a_m,a_s} = N_k\, C_{k,m}C_{k,s}M_k(m,s), I^{(k)}_{b_m,b_s} = N_k\, D_{k,m}D_{k,s}M_k(m,s),$$

with analogous formulas for $I^{(k)}_{a_m,b_s}$ and the cross–block entries $I^{(k)}_{a_m,\eta_r} = N_k C_{k,m}A_k(m,r)$ and $I^{(k)}_{b_m,\eta_r} = N_k D_{k,m}A_k(m,r)$.

Approximations. For the separation approximation, the following approximation holds: $\gamma_{k,m}\gamma_{k,s} \approx \delta_{ms}$ (with $\delta_{ms} = 1$ if $m = s$ and 0 else), which implies $M_k(m,s) \approx \delta_{ms}\sigma_k^2$ and $A_k(m,r) \approx 0$. Thus, the diagonal terms simplify to:

$$\begin{aligned} I^{(k)}_{a_m,a_m} &\approx \frac{N_k\pi_m b_m^2}{\sigma_k^2(d_k - a_m)^4}, I^{(k)}_{b_m,b_m} \approx \frac{N_k\pi_m}{\sigma_k^2(d_k - a_m)^2}, \\ I^{(k)}_{a_m,b_m} &\approx \frac{N_k\pi_m b_m}{\sigma_k^2(d_k - a_m)^3}, \end{aligned} \tag{B5}$$

and all coupling terms depending on $\boldsymbol{\eta}$ are close to 0. The CRLB is therefore *block-diagonal* in this regime: frequency–shape parameters (a,b) decouple from mixing weights. This is a theoretical justification for estimating the vectors (a_m, b_m) and $\boldsymbol{\eta}$ separately in the case of well-resolved modal peaks. Similar results are obtained with the modified CRLB in the presence of latent variables [5].

4. SIMULATION RESULTS

This section evaluates the performance of the EM algorithm of Section 2.2 (that estimates the C-AGMM parameters) and verifies that the Cramér–Rao bounds derived in Section 3 are consistent with the mean square errors (MSEs) of the EM estimates computed using Monte Carlo simulations. Two experiments are considered studying the influence of the noise variance and the effect of the overlap between two modes. For both experiments, the mixture probability vector is $\boldsymbol{\pi} = [0.05,\ 0.325,\ 0.625]$ (5% of the detected points are outliers), the number of harmonics satisfies $K \in \{3, ..., 11\}$ and each harmonic intersects each mode. The sample size per harmonic is $N_k = 4$, which agrees with real data. The MSEs of the EM estimates are computed using 1000 independent Monte-Carlo runs initialized with the RANSAC algorithm [7]. They are compared with those obtained using a Hough algorithm for line detection [6].

Experiment 1: varying noise variance. This scenario considers two modes with parameters $(a_1, b_1) = (0\,\text{Hz.RPM}^{-1},\ 150\ \text{Hz})$ and $(a_2, b_2) = (25.10^{-4}\,\text{Hz.RPM}^{-1},\ 300\ \text{Hz})$ and the noise standard deviation $\sigma_k = \sigma$ increases logarithmically in the interval $[0.1, 100]$Hz. Figure 2 displays for each value of σ the empirical RMSEs of the EM estimates of b_1 and b_2 and the corresponding square-root of the CRLBs (dashed lines). The MSE of b_2 is closer to the CRLB than that of b_1. This is expected because the CRLB is attained only asymptotically, and each Monte-Carlo run provides nearly twice as many effective samples for b_2 for b_1 ($\pi_2 \approx 2\pi_1$). Note that the RMSEs obtained using the Hough and EM algorithms are similar for small noise levels, with a slightly better performance for the Hough algorithm for large noise variances.

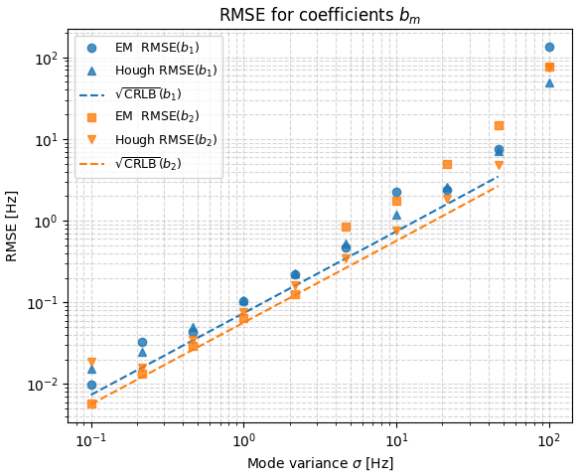

Fig. 2: Experiment 1: Influence of noise variance. RMSEs of EM and Hough estimates for b_1 and b_2 (triangles, squares and circles) and corresponding $\sqrt{\text{CRLBs}}$ (dashed lines).

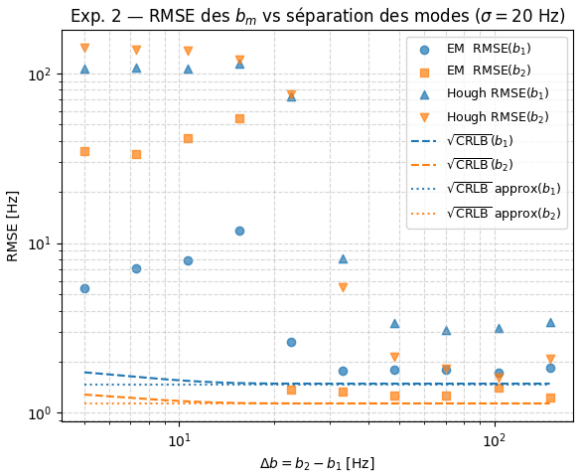

Fig. 3: Experiment 2 — Influence of mode separation. RMSEs of EM and Hough estimates (triangles, squares and circles) and $\sqrt{\text{CRLB}}$ (dashed lines) for different separations ($\sigma_k = 20$ Hz).

Experiment 2: mode separation. This experiment studies the impact of the difference between the two frequencies b_1 and b_2 on the estimation performance (see [1] for a similar analysis). For a fixed noise variance $\sigma = 20$Hz, the second mode is moved toward the first mode, i.e. $b_1 = 150$ Hz and $b_2 = b_1 + \Delta b$ with $\Delta b \in [5, 150]$ Hz. Figure 3 shows the RMSEs and square roots of the CRLBs of b_1 and b_2 as the two modes get close to each other. The MSEs of the EM estimates are close to the CRLBs for large values of Δ, as expected. For small values of Δ, the performance of the EM algorithm drops (higher MSEs and CRLBs) since the high SNR region is not reached and the estimator is possibly biased [15, p19]. However, the EM algorithm performs better than the Hough transform for close modes.This experiment provides a practical lower bound on the minimum separation—and thus on the attainable precision—when analysing clusters of closely spaced modes. It confirms that the EM procedure is nearly efficient in well-conditioned regimes, while also highlighting the practical limits set by high noise level or strong spectral overlap, situations where the FIM approaches singularity and the performance of the EM estimator decreases significantly.

Experiment 3: real dataset. The method was finally applied to a real-world signal acquired during a fan-blade test campaign on a turbofan engine. The detection map was generated in two stages. First, order tracking was performed with the Vold–Kalman filter [8], an algorithm widely used in industry (see [4] for alternative). Second, spectral peaks were identified by estimating the local maxima along the estimated orders. The EM Algorithm was then run with a RANSAC initialization. The Akaike Information Criterion (AIC) was applied to each candidate model in order to estimate the unknown number of modes. The order associated with the lowest AIC score was $M = 18$ with estimated modes shown in Fig. 4. In the frequency band of interest, the modal lines coincide very closely with those identified by domain experts with a much faster processing.

5. CONCLUSION & PERSPECTIVES

This work formulated order-based modal analysis with regime dependent modes as an inference problem in a coupled affine Gaussian mixture model. A dedicated EM algorithm was then derived to compute the maximum-likelihood estimator of the model parameters. Cramér–Rao lower bounds were finally derived for the parameters of this model allowing optimal performance to be determined. Several experiments confirmed that the suggested EM estimator yields parameter mean square errors close to these bounds, when modal peaks are sufficiently separated and noise is moderate. The EM algorithm performs similarly to the Hough method in easy regimes but remains more accurate for small inter-mode separations. Moreover, the EM algorithm provides association probabilities that can be used for uncertainty quantification, which is important for practical applications. The application to data from an industrial turbofan showed that the method extracts modal lines whose characteristics are very close to those provided by experts, which is very encouraging. Future work includes a fully Bayesian formulation with priors informed by mechanical finite-element predictions, which would deliver posterior uncertainties and impose natural regularisation.

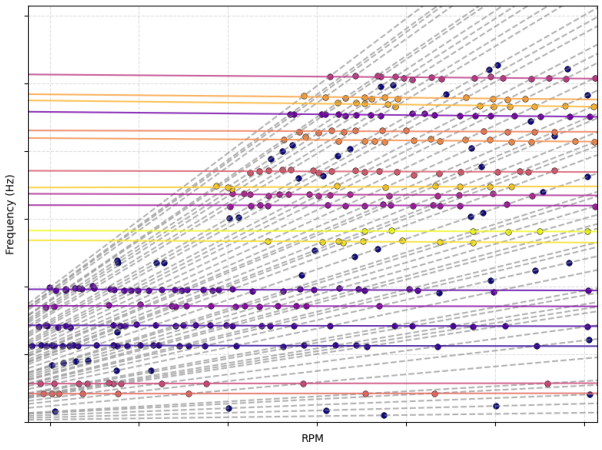

Fig. 4: Detection map from a real engine run-up test. Each dot represents a detected peak: black dots correspond to outliers, while colored dots indicate assignment to one of the identified modes.

6. REFERENCES

[1] R. Badeau, B. David, and G. Richard. Cramér–Rao bounds for multiple poles and coefficients of quasi-polynomials in colored noise. *IEEE Trans. on Signal Processing*, 56(8):3458–3467, August 2008.

[2] R. A. Carmona, W. L. Hwang, and B. Torresani. Characterization of signals by the ridges of their wavelet transforms. *IEEE Trans. Signal Process.*, 45(10):2586–2590, 1997.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 39:1–22, September 1977.

[4] A. Emelchenkov et al. Multifrequency highly oscillating aperiodic amplitude estimation for nonlinear chirp signal. In *Proc. Int. Conf. European Signal Processing Conference*, Lyon, France, August 2024.

[5] C. Herzet et. al. A Cramér-Rao bound characterization of the em-algorithm mean speed of convergence. *Transactions on Signal Processing*, 56(6), June 2008.

[6] R. Duda et. al. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, January 1972.

[7] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.

[8] H. Herlufsen, S. Gade, H. Konstantin-Hanse, and H. Vold. Characteristics of the vold-kalman order tracking filter. In *Proc. Int. Conf. Acoustics, Speech, and Signal Processing. Proceedings (ICASSP '00)*, volume 6, page 3895–3898, Istanbul, Turkey, April.

[9] F Hlawatsch. *Time-Frequency Analysis and Synthesis of Linear Signal Spaces*. Springer, MA, USA, 1998.

[10] E. Di Lorenzo et al. Dynamic characterization of wind turbine gearboxes using order-based modal analysis. In *Proc. Int. Conf. Noise Vibration Engineering (ISMA2014) and Int. Conf. Uncertainty Structural Dynamics (USD2014)*, pages 4349–4362, Leuven, Belgium, Sep. 15-17, 2014.

[11] T. K. Moon. The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, 13(6):47–60, 1996.

[12] A. Silik, M. Noori, Z. Wu, W. A. Altabey, J. Dang, and N. S. D. Farhan. Wavelet-based vibration denoising for structural health monitoring. *Urban Lifeline*, 2(15), Nov. 2024.

[13] W. J. Staszewski and A. N Robertson. Time–frequency and time–scale analyses for structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):449–477, Dec. 2006.

[14] G. Sternharz, T. Kalganova, C. Mares, and M. Meyeringh. Comparative performance assessment of methods for operational modal analysis during transient order excitation. *Mechanical Systems and Signal Processing*, 169:108719, April 2022.

[15] Harry L. Van Trees and Kristine L. Bell. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley-IEEE Press, 2007.

[16] Y. Zeng et al. An interpretable method for operational modal analysis in time-frequency representation and its applications to railway sleepers. *Structural Control and Health Monitoring*, 2023:1–26, July 2023.