

# UNSUPERVISED CONDITION CHANGE DETECTION IN LARGE DIESEL ENGINES

Niels Henrik Pontoppidan and Jan Larsen  
Informatics and Mathematical Modelling, Technical University of Denmark  
Richard Petersens Plads, Building 321, DK-2800 Lyngby, Denmark  
Phone: +45 45253899, Fax: +45 45872599  
E-mail: nhp,jl@imm.dtu.dk, Web: isp.imm.dtu.dk

**Abstract.** This paper presents a new method for unsupervised change detection which combines independent component modeling and probabilistic outlier detection. The method further provides a compact data representation, which is amenable to interpretation, i.e., the detected condition changes can be investigated further. The method is successfully applied to unsupervised condition change detection in large diesel engines from acoustical emission sensor signal and compared to more classical techniques based on principal component analysis and Gaussian mixture models.

## INTRODUCTION

Identification of engine conditions and faults is important for automatic monitoring of critical failures in large marine diesel engines and stationary power plants. The possibility of early detecting small defects prior to evolving into serious breakdowns often reduce the costs for repair significantly. While the long term objective is to classify engine conditions into known fault types, this work focuses merely on the detection of condition changes.

The literature suggests that monitoring based on acoustical emission (AE) offers advantages over sensor techniques such as pressure and vibration [19, 20]. The signal-to-noise ratio is typically better for AE sensor signals, and further a system based on AE is more suitable from an operational point of view. Previous work on adaptive signal processing and machine learning [4, 5, 7, 6, 13, 21, 22] has mainly focused on supervised learning from sensor data and known faults. This paper focuses on unsupervised learning for significant detection of changes in measured AE signals, that is, modeling the probability density of the AE signal. Since AE data are abundant we focus on models, which also offers compact data representation, such as the Independent Component Analysis (ICA), Principal Component Analysis (PCA) and Unsupervised Gaussian Mixture (UGM) models in combination with PCA. The probability density associated with the trained ICA, PCA

and UGM models [9, 18, 16, 12, 15] can be used to identify events which do not conform to the model assumptions [15, 3] and thus represent a significant change in engine condition.

The next section presents the modeling framework and a novel change detection algorithm based on ICA and/or PCA models. The results of a comparative analysis using Bayesian Information Criterion (BIC) and Receiver Operator Characteristics (ROC) is followed by the description of data acquisition, experimental setup, ending with the concluding remarks.

## MODELING FRAMEWORK

Feature vectors from  $N$  examples (revolutions) are assembled into a training data matrix  $\mathbf{X}_{\mathcal{T}}$  of size  $d \times N$   $\mathbf{X}_{\mathcal{T}} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ . Unsupervised modeling considers modeling the probability density  $p(\mathbf{x}, \boldsymbol{\theta})$ , where  $\boldsymbol{\theta}$  is a parameter vector. The model parameters estimated from available training data  $\mathbf{X}_{\mathcal{T}}$  are denoted by  $\hat{\boldsymbol{\theta}}$ .

### Principal Component Analysis Model (PCA)

Since  $d$  is typically larger than  $N$  we will invoke the PCA model [9, 18], which considers a  $K$  dimensional  $K \ll d$  signal space with rank  $K \leq \min(d, N)$  covariance, and an additive isotropic noise,  $\tilde{\mathbf{x}} = \mathbf{s} + \mathbf{v}$ ,  $p(\tilde{\mathbf{x}}|\boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\tilde{\mathbf{x}}})$ , where  $\tilde{\mathbf{x}} = \mathbf{x} - E\{\mathbf{x}\}$  and  $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\tilde{\mathbf{x}}})$  is the zero mean Gaussian distribution with covariance matrix  $\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}} = \boldsymbol{\Sigma}_{\mathbf{s}} + \sigma_{\varepsilon}^2 \mathbf{I}$ ,  $\boldsymbol{\Sigma}_{\mathbf{s}}$  has rank  $K$ .

The model is estimated from data using a singular value decomposition of centered data  $\tilde{\mathbf{X}}_{\mathcal{T}}$ ,  $\tilde{\mathbf{x}}_n = \mathbf{x}_n - \hat{\boldsymbol{\mu}}_{\mathbf{x}}$  and  $\hat{\boldsymbol{\mu}}_{\mathbf{x}} = N^{-1} \sum_{n=1}^N \mathbf{x}_n$ . Assuming  $d \geq N$ , the SVD is given as  $\tilde{\mathbf{X}}_{\mathcal{T}} = \mathbf{U} \mathbf{D} \mathbf{V}^{\top}$ , where  $\mathbf{U}$  is  $d \times N$  and  $\mathbf{V}$   $N \times N$  are left and right eigenvectors and  $\mathbf{D}$  is the  $N \times N$  diagonal matrix of decreasing singular values. Define  $\tilde{\mathbf{U}}$  as the first  $K$  columns of  $\mathbf{U}$ , then for specific choice  $K$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{s}} = \frac{\tilde{\mathbf{U}} \text{diag}(D_1^2 - \hat{\sigma}_{\varepsilon}^2, \dots, D_K^2 - \hat{\sigma}_{\varepsilon}^2) \tilde{\mathbf{U}}^{\top}}{N}, \quad \hat{\sigma}_{\varepsilon}^2 = \frac{1}{N(d-K)} \sum_{i=K+1}^N D_i^2. \quad (1)$$

In order to estimate the optimal model complexity,  $K_{\text{opt}}$ , we use the Bayesian information criterion (BIC) [17, 10, 18]. BIC is an estimate of model evidence given by

$$p(\tilde{\mathbf{X}}|K) \approx p(\tilde{\mathbf{X}}|\hat{\boldsymbol{\theta}}) \cdot p(\hat{\boldsymbol{\theta}}) \cdot (2\pi/N)^{\dim(\boldsymbol{\theta})/2} \quad (2)$$

Here  $p(\tilde{\mathbf{X}}|\hat{\boldsymbol{\theta}})$  is the likelihood on training data and  $p(\hat{\boldsymbol{\theta}})$  is the prior on parameters. When no explicit prior is available, we use an improper uniform prior. In the case of the PCA model the parameter vector is  $\boldsymbol{\theta} = (\boldsymbol{\mu}, \tilde{\mathbf{U}}, \sigma_{\varepsilon}^2, D_1, \dots, D_K)$ . That is,  $\dim(\boldsymbol{\theta}) = d + K(2d - K + 1)/2 + 1 + K$ .

The PCA model can also be written as

$$p(\mathbf{x}|\boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\varepsilon}|\boldsymbol{\theta}) \quad (3)$$

where  $\mathbf{y} = \tilde{\mathbf{U}}^\top \mathbf{x}$  is  $K$  dimensional signal space and  $\boldsymbol{\varepsilon} = \mathbf{U}_\varepsilon^\top \mathbf{x}$  is the  $d - K$  dimensional noise space with diagonal covariance structure  $\sigma_\varepsilon^2 \mathbf{I}$ . Here  $\mathbf{U}_\varepsilon$  are the last  $d - K$  columns of  $\mathbf{U}$ . Under the model,  $\tilde{\mathbf{x}}$  is estimated by  $\tilde{\mathbf{U}}\mathbf{y}$ . That is, the columns of  $\tilde{\mathbf{U}}$  can be interpreted as  $K$  AE signatures which describe the  $\mathcal{H}_0$  condition. The principal components (sources)  $\mathbf{y}$  express the strength of each signature.

### Noise Free Independent Component Analysis Model (ICA-BS)

Assume the noise free ICA model [16]  $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{s}$ , where  $\mathbf{A}$  is a  $d \times K$  mixing matrix and  $\mathbf{s}$  the  $K$  dimensional source vector with statistically independent components. The non-quadratic noise free ICA can be performed in two steps by decomposing  $\mathbf{A} = \tilde{\mathbf{U}}\boldsymbol{\Phi}$ , where  $\tilde{\mathbf{U}}$  is  $d \times K$  projection matrix onto the  $K$  subspace spanned by the sources, and  $\boldsymbol{\Phi}$  is  $K \times K$  mixing matrix. If the source space is  $K$ -dimensional and second order moments of  $\tilde{\mathbf{x}}$  exist, then the projection matrix can be obtained from an SVD projection as described in the previous subsection.

We will use the Infomax algorithm [2] with classical  $\tanh(\cdot)$  nonlinearity<sup>1</sup>. The deployed implementation of the algorithm can be obtained from ICA-ML *DTU:toolbox* [14].

For model selection we will use BIC Eq. (2) with the assumption of independent signal and noise spaces as in Eq. (3), i.e.,

$$p(\mathbf{x}|\boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\Phi}) \cdot p(\boldsymbol{\varepsilon}|\sigma_\varepsilon^2) \quad (4)$$

where  $p(\mathbf{y}|\boldsymbol{\Phi})$  is the Infomax likelihood [16]

$$p(\mathbf{y}|\boldsymbol{\Phi}) = |\det(\boldsymbol{\Phi})|^{-1} \cdot p_{\mathbf{s}}(\boldsymbol{\Phi}^{-1}\tilde{\mathbf{U}}^\top \mathbf{x}) \quad (5)$$

with  $p_{\mathbf{s}}(\mathbf{s}) = \prod_i 1/\pi \cosh(s_i)$ . The noise likelihood function is [10, Eq. (12)]

$$p(\boldsymbol{\varepsilon}|\hat{\sigma}_\varepsilon^2) = (2\pi\hat{\sigma}_\varepsilon^2)^{-N(d-K)/2} \cdot \exp(-N(d-K)/2) \quad (6)$$

Since  $\boldsymbol{\theta} = (\boldsymbol{\mu}, \tilde{\mathbf{U}}, \sigma_\varepsilon^2, \boldsymbol{\Phi})$ , the total number of parameters are  $\dim(\boldsymbol{\theta}) = d + K(2d - K + 1)/2 + 1 + K^2$ .

### Noisy Independent Component Analysis Model (ICA-MF)

An advanced Bayesian ICA using mean field training [12] enables the training of an ICA model with noise,  $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{e}$ , under flexible source distributions and possible priors on the mixing matrix. The noise is assumed Gaussian,

<sup>1</sup>Corresponding to identical source priors  $p_{s_i}(s_i) = 1/\pi \cosh(s_i)$ .

independent of the sources, and with diagonal covariance matrix. The preprocessing SVD projection step is not exact in the case of noise, i.e., the estimation procedure estimates the  $d \times K$  mixing matrix  $\mathbf{A}$  directly.

As described above, the columns of  $\mathbf{A}$  correspond to AE RMS signatures associated with individual sources, which consequently are non-negative. We therefore invoke a non-negativity prior constraint on the mixing matrix. The activation of these signatures should also be non negative, i.e., source should be non-negative and consequently we use an exponential prior source distribution. The noisy ICA model is estimated using the the ICA-MF *DTU:toolbox* code [14].

The number of sources is also in this case estimated using BIC, Eq. (2).

### Unsupervised Gaussian Mixture Model (UGM)

For comparison we also consider the Gaussian mixture model with SVD signal space preprocessing [11]. Thus as in Eq. (3) we assume  $p(\mathbf{x}|\boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\varepsilon}|\sigma_\varepsilon^2)$ , where  $p(\mathbf{y}|\boldsymbol{\theta})$  is the Gaussian mixture density  $p(\mathbf{y}|\boldsymbol{\theta}) = \sum_{c=1}^C P(c)p(\mathbf{x}|c, \boldsymbol{\theta}_c)$ , with  $p(\mathbf{x}|k, \boldsymbol{\theta}_c) = \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$ , and  $\boldsymbol{\theta} = \{P(c), \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c\}$  consists of mixing proportions  $P(c)$  as well as means  $\boldsymbol{\mu}_c$  and covariances  $\boldsymbol{\Sigma}_c$  of the Gaussian components. The model is estimated using the generalizable Gaussian mixture algorithm [15] and the subspace dimension  $K$  and number of components  $C$  are selecting using the BIC criterion, Eq. (2).

### Novelty detection

A general treatment of change detection is presented in e.g., [1, 8]. Here we suggest to deploy the novelty detection method proposed in [15, 3], which makes it possible to evaluate whether new examples conform to the trained model  $p(\mathbf{x}, \hat{\boldsymbol{\theta}})$ . A test sample  $\mathbf{x}$  conforming with the trained model will have high log-likelihood whereas a sample from another condition will have low log-likelihood value. In order to perform a formal comparison, we consider the cumulative density of the log-likelihood values.

$$Q(t) = \text{Prob}(\log p(\mathbf{x}|\hat{\boldsymbol{\theta}}) < t) \quad (7)$$

$Q(t)$  can be interpreted as the empirical estimate of the probability that the example  $\mathbf{x}$  (with log likelihood  $t$ ) belongs to the model given by the parameters  $\hat{\boldsymbol{\theta}}$ , i.e. the model that generated the training set.<sup>2</sup>Using a threshold, e.g.,  $t_{min} = 5\%$ , new examples where  $Q(t) < t_{min}$  are rejected under  $\mathcal{H}_0$  at a 5% significance level. See further figure 1.

<sup>2</sup>For a Gaussian density  $Q(t)$  is  $\chi^2$  distributed. In general, we can only compute this from samples, e.g., by generating an arbitrarily large sample from the generative model  $p(\mathbf{x}|\hat{\boldsymbol{\theta}})$ .

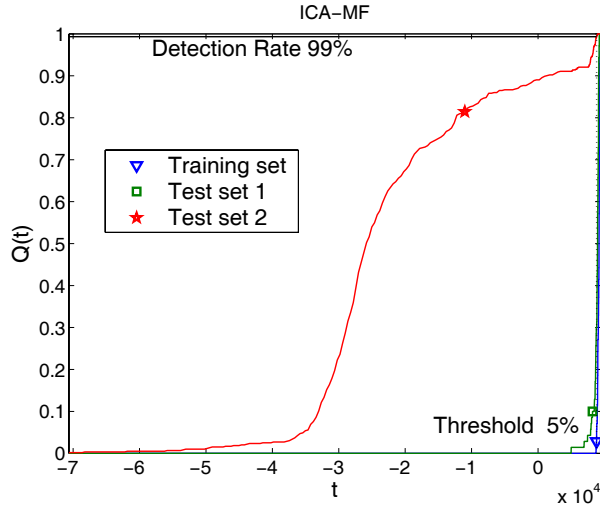


Figure 1: Cumulative log-likelihood density,  $Q(t)$ , from experiment 1 using ICA-MF with 6 components for *training set*, *test set 1* and *test set 2*. *training set* and *test set 1* are very close which means that if we use a 5% threshold on *training set* curve only very few example will be falsely detected. On the other hand, 99% of the examples in *test set 2* will be detected as novel, i.e., as a new condition.

## EXPERIMENTAL RESULTS

We consider data from three experiments described in the following section, however we only show results from the first experiment, in which the lubricating oil is shut off. The other experiments gives similar performance results besides from changes in the optimal number of components.

The  $Q(t)$  function of the trained models are computed from the training set. Choosing a specific threshold  $t_{\min}$  then the false alarm rate can be estimated as the fraction of examples in *test set 1* (belonging to  $\mathcal{H}_0$ ) for which log-likelihood  $\log p(\mathbf{x}|\hat{\boldsymbol{\theta}}) < t_{\min}$ . Similarly the true detection probability is estimated as the fraction of examples on *test set 2* (belonging to  $\mathcal{H}_1$ ) for which log-likelihood smaller than  $t_{\min}$ . By varying  $t_{\min}$  the so-called receiver operation characteristics (ROC) curves are formed, which is shown in figure 3. Larger area under the ROC curve implies higher true detection for a given false alarm. Clearly ICA-MF shows best true performance. In order to interpret the nature of the changed condition we can evaluate the difference between a test feature vector  $\mathbf{x}$  and its estimate under the model. For ICA-MF we first estimate the source  $\hat{\mathbf{s}}$  and then compute the estimate under the model  $\hat{\mathbf{x}} = \hat{\mathbf{A}}\hat{\mathbf{s}}$ . The interpretation is shown in figure 5.

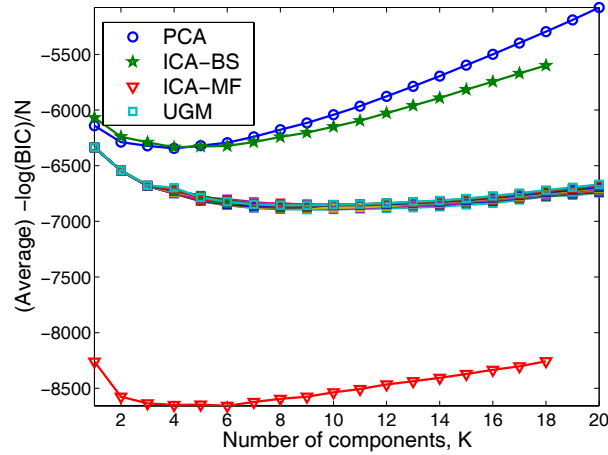


Figure 2: Model selection using the BIC criterion Eq. (2). The curves display  $-\log p(\tilde{\mathbf{X}}|K)/N$ . ICA-MF averaged over 4 runs and UGM over 20 runs. Further UGM curves for 2-16 clusters are plotted. PCA and ICA-BS achieved minimum BIC value for  $K = 4$ , ICA-MF for  $K = 6$ , and UGM for the combination  $K = 9$  and  $C = 2$ . The significantly lower BIC of the ICA-MF model indicates this model is preferable, possibly due to a more advanced noise model.

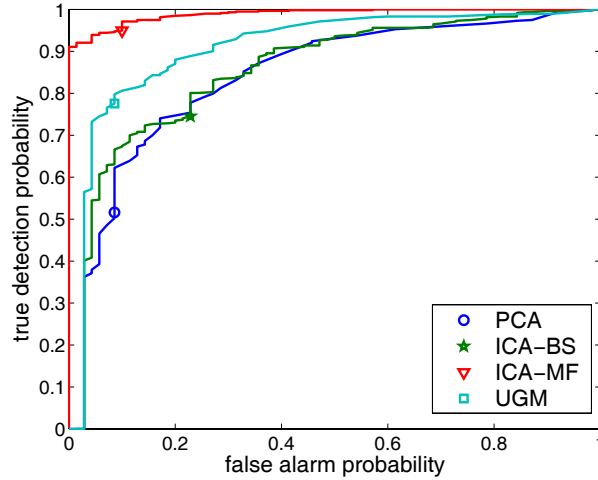


Figure 3: ROC curves from experiment 1 (shutting off lubrication) shows probability of false alarm versus true detection. Clearly the noisy ICA model (ICA-MF) provides best performance.

## EXPERIMENTAL SETUP

The data set consists of acoustical emission (AE) root-mean squared (RMS) signals acquired with four AE sensors. In this work we will use a single sensor placed on the liner (cylinder casing). Data was recorded for 10 seconds followed by a pause of 60 seconds as a simple compression scheme. Data

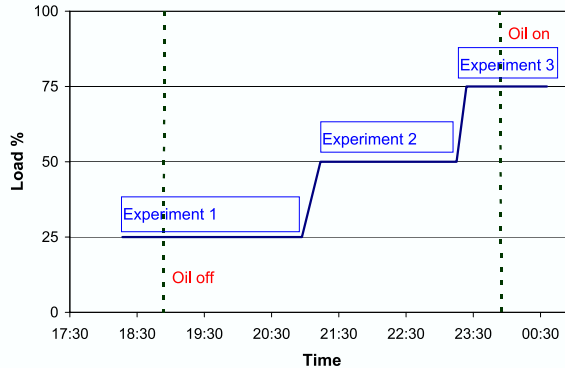


Figure 4: Time line of experiment. The curve shows the increasing load as function of time. The numbered boxes refer to the three experiments described in the experimental setup section. The two vertical lines indicate when lubricating was turned off and on.

was originally sampled at 2.5 MHz using the RMS time constant  $50 \mu\text{s}$ . The signal is resampled into crank angle domain using a crank encoder. This partially compensates for variations in rotation speed and establish the relation between AE signal expression and mechanical events during the combustion cycle. Define the  $d = 2048$  dimensional feature vector

$$\mathbf{x}_n = [x_n(1), \dots, x_n(i), \dots, x_n(2048)]^\top, \quad (8)$$

where  $x_n(i)$  is the RMS AE signal for cycle  $n$  at angle  $(i - 1) \cdot 360/2048^\circ$ .  $0^\circ$  corresponds to the top position as indicated by top pulse signal. 21 other channels (including top and crank-pulses) were acquired from the cylinder at MAN B&W Diesel's Research Engine<sup>3</sup> in Copenhagen. These additional signals can be used to interpret the results of AE signal analysis.

During the experiment, the engine load was changed from 25% to 75%. In the middle of the 25% load period the cylinder lubrication was turned off, and in the middle of the 75% load period lubrication was re-established. Figure 4 shows the actual timing of these events.

From the entire data set we have selected periods where the engine displays non-trivial abrupt condition changes. Thus we are not interested in detecting that the load changes but e.g., that lubrication is turned on or off.

Knowledge about condition changes is obtained from manual annotations by MAN B&W and from additional 21 sensor channels. This information is not directly passed to the algorithms, but is used in order to design relevant data periods and for performance evaluation.

We consider three experiments indicated in Figure 4.

**Experiment 1: Shutting Off Lubrication** After turning on the engine, the load stabilized at 25% on the propeller curve. After a while the

<sup>3</sup>Test bed, 4 cylinders, 500 mm bore, 10.000 BHP.

lubrication to the cylinder is turned off. The objective is to detect this operation condition change.

**Experiment 2: Unstable Revolution Speed** The engine is running at 50% load with the lubrication system turned off. Inspection of the revolution speed obtained from timing signal indicates that the engine condition undergoes sudden changes in the middle of this period, which is probably caused by engine load fluctuations. We aim to detect the start and end of this period.

**Experiment 3: Re-establishment of lubrication** The engine is running at 75% load without lubrication. After 30 minutes lubrication is re-established, possibly lowering the wear rate. We aim to detect this change.

In order to validate the performance of the detection we consider a null-hypothesis  $\mathcal{H}_0$ , the current normal condition, and a new condition,  $\mathcal{H}_1$ . The data from each experiment are divided into:

*Training set* contains examples from the current engine operation condition  $\mathcal{H}_0$ .

*Test set 1* contains examples from current condition,  $\mathcal{H}_0$ , and is used for model validation.

*Test set 2* contains examples that we based on annotations believe come from the the new condition,  $\mathcal{H}_1$ .

## CONCLUSION

This paper presented a novel probabilistic change detection framework based on independent component analysis (ICA) modeling. The method was successfully applied to unsupervised condition change detection in large diesel engines using acoustical emission sensors. The overdetermined noisy ICA model using mean-field Bayesian learning showed best performance.

**Acknowledgment.** The work is supported by EU Competitive and Sustainable Growth Programme GRD2-2001-50014 through the AE-WATT project <http://isp.imm.dtu.dk/aewatt>. Torben Fog Man B&W is acknowledged for fruitful discussions and data support. Further we would like to thank Kaare B. Petersen and Lars Kai Hansen for critical comments.

## REFERENCES

- [1] M. Basseville and I. V. Nikiforov, **Detection of Abrupt Changes: Theory and Application**, Prentice-Hall, 1993.
- [2] A. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," **Neural Computation**, vol. 7, pp. 1129–1159, 1995.



- [3] C. M. Bishop, "Novelty Detection and Neural Network Validation," **IEEE Proceedings - Vision Image and Signal Processing**, vol. 141, no. 4, pp. 217–222, 1994.
- [4] G. Chandroth, A. J. C. Sharkey and N. E. Sharkey, "Cylinder Pressures and Vibration in Internal Combustion Engine Condition Monitoring," in **Comadem 99**, July 1999.
- [5] G. Chandroth, A. J. C. Sharkey and N. E. Sharkey, "Vibration signatures, wavelets and principal components analysis in diesel engine diagnostics," in **Marine Technology ODRA 99**, October 1999.
- [6] T. Fog, **Condition Monitoring And Fault Diagnosis in Marine Diesel Engines**, Ph.D. thesis, Technical University of Denmark, 1998, IMM-PHD-1998-52.
- [7] T. L. Fog, L. K. Hansen, J. Larsen, H. S. Hansen, L. B. Madsen, P. Srensen, E. R. Hansen and P. S. Pedersen, "On Condition Monitoring of Exhaust Valves in Marine Diesel Engines," in Y. H. Hu, J. Larsen, E. Wilson and S. Douglas (eds.), **Proceedings of the IEEE Workshop on Neural Networks for Signal Processing IX**, IEEE, Piscataway, New Jersey, 1999, pp. 225–234.
- [8] F. Gustafsson, **Adaptive Filtering and Change Detection**, John Wiley and Sons LTD, 2000, ISBN: 0 471 49287 6.
- [9] L. Hansen and J. Larsen, "Unsupervised Learning and Generalization," in **Proceedings of the IEEE Int. Conf. on Neural Networks 1996**, 1996, vol. 1, pp. 25–30.
- [10] L. K. Hansen, J. Larsen and T. Kolenda, "Blind Detection of Independent Dynamic Components," In **proc. IEEE ICASSP'2001**, vol. 5, pp. 3197–3200, 2001.
- [11] L. K. Hansen, S. Sigurdsson, T. Kolenda, F. . Nielsen, U. Kjems and J. Larsen, "Modeling Text with Generalizable Gaussian Mixtures," in **IEEE ICASSP'2000, Istanbul, Turkey**, June 2000, vol. VI, pp. 3494–3497.
- [12] P. Højen-Sørensen, O. Winther and L. K. Hansen, "Mean Field Approaches to Independent Component Analysis," **Neural Computation**, vol. 14, pp. 889–918, 2002.
- [13] M. Kaveh, A. H. Tewfik, K. M. Buckley et al., **Integrated Predictive Diagnostic Tools**, MultiUniversity Center for Integrated Diagnostics, chap. 4.b: Signal processing for the detection and classification of acoustic emissions in integrated diagnostics, 2000.
- [14] T. Kolenda, S. Sigurdsson, O. Winther, L. K. Hansen and J. Larsen, "DTU:Toolbox," Internet, 2002, <http://isp.imm.dtu.dk/toolbox/>.
- [15] J. Larsen, L. K. Hansen, A. Szymkowiak, T. Christiansen and T. Kolenda, "Webmining: Learning from the World Wide Web," **Computational Statistics and Data Analysis**, vol. 38, pp. 517–532, 2002.
- [16] T.-W. Lee, **Independent Component Analysis: Theory and Applications**, Kluwer Academic Publishers, September 1998, ISBN: 0 7923 8261 7.
- [17] D. J. C. MacKay, "Bayesian Model Comparison and Backprop Nets," in **Advanced of Neural Information Processing Systems 4**, 1992, pp. 839–846.
- [18] T. P. Minka, "Automatic choice of dimensionality for PCA," in T. Leen, T. Dietterich and V. Tresp (eds.), **Advances in Neural Information Processing Systems 13**, MIT Press, 2001, pp. 598–604.
- [19] G. D. Neill, S. Benzie, J. D. Gill, P. M. Sandford, E. R. Brown, J. A. Steel and R. L. Reuben, "The relative merits of acoustic emission and acceleration monitoring for detection of bearing faults," in **COMADEM**, December 1998,

pp. 651–661.

- [20] R. L. Reuben, “The Role of Acoustic Emission in Industrial Condition Monitoring,” **International Journal of COMADEM**, vol. 1, no. 4, pp. 35–46, Oct. 1998, ISSN 1363-7681.
- [21] L. M. Rogers, “Use of acoustic emission methods for crack growth detection in offshore and other structures,” **Trans IMarE**, vol. 110, no. part 3, pp. 171–180, 1998.
- [22] A. Ypma, **Learning methods for machine vibration analysis and health monitoring**, Ph.D. thesis, Dept. of Applied Physics, Delft University of Technology, Nov. 2001.

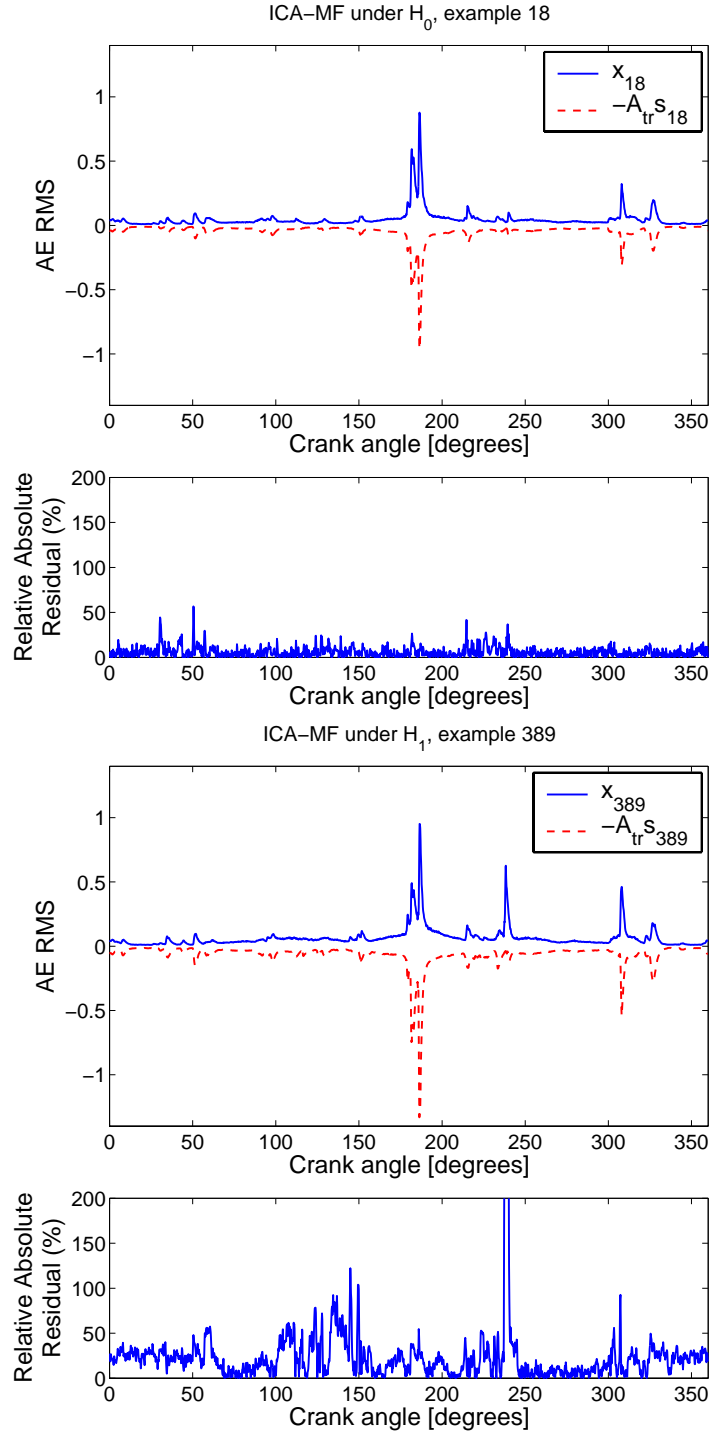


Figure 5: Interpretation of examples which ICA-MF detected as belonging to  $\mathcal{H}_0$  or  $\mathcal{H}_1$ . The upper panel shows  $\mathbf{x}$  and the (negative) estimate  $\hat{\mathbf{x}} = \mathbf{A}\mathbf{s}$ , and the lower panel the relative error  $100\% \cdot |(\mathbf{x} - \hat{\mathbf{x}})/\hat{\mathbf{x}}|$ . Under  $\mathcal{H}_0$  the relative error typically is around 10% while the example under  $\mathcal{H}_1$  possesses very high error around 1500% for crank angle position close to  $240^\circ$ . Knowledge about the engine combustion cycle at crank angle position  $240^\circ$ , can then be used to identify the nature and impact of detected condition change.