On the Implementation of Downsampling Permutation Entropy variants in the detection of Bearing Faults in Rotatory Machines

Antonio DÁVALOS-TREVIÑO

Meryem JABLOUN

Philippe RAVIER

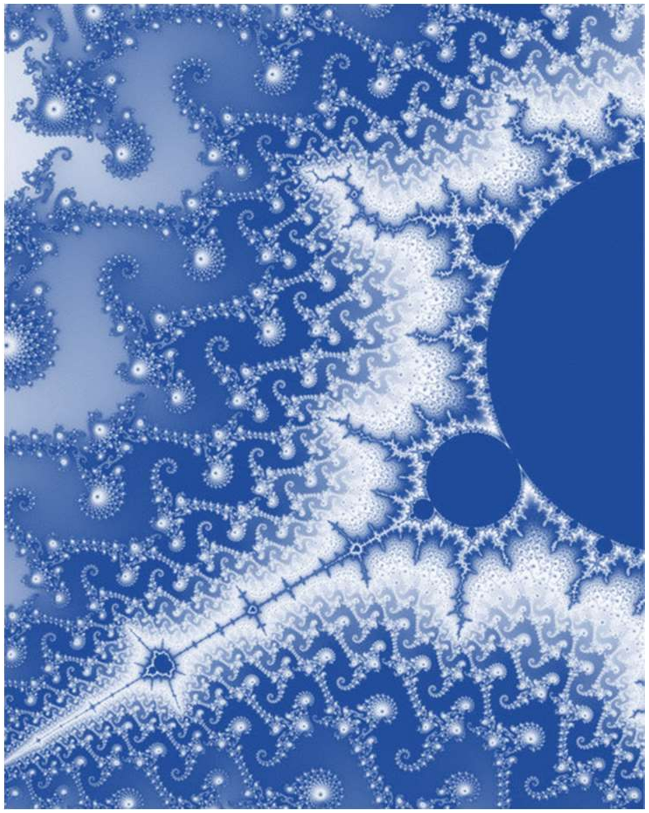
Olivier BUTTELLI

Mathématiques, Informatique, Physique, Théorique et Ingénierie des Systèmes [MIPTIS] Doctoral School

PRISME Laboratory,
University of Orléans, Fr.

MPE - Background

- Multiscale Permutation Entropy (MPE) of dimension d analyzes the information contained in the ordinal patterns of a signal at different time scales. (Aziz & Arif., 2005).
- Refinements to this method, like rcMPE (Humeau-Heutier 2015), greatly improve the precision of MPE, particularly for short time signals.
- Our contribution: First, we propose an alternative multiscaling method, using composite downsampling instead of composite coarse-graining, which further improves the method's precision.
- Motivation: Composite multiscaling creates artifact correlations within the signal, which increases the MPE variance. By downsampling, we avoid the unwanted effects of preprocessing in the final MPE value.



Contents

1 **Theoretical Background**

Permutation Entropy and Multiscaling techniques

2 **Refined Composite Downsampling Permutation Entropy**

Composite Downsampling and Statistics comparison

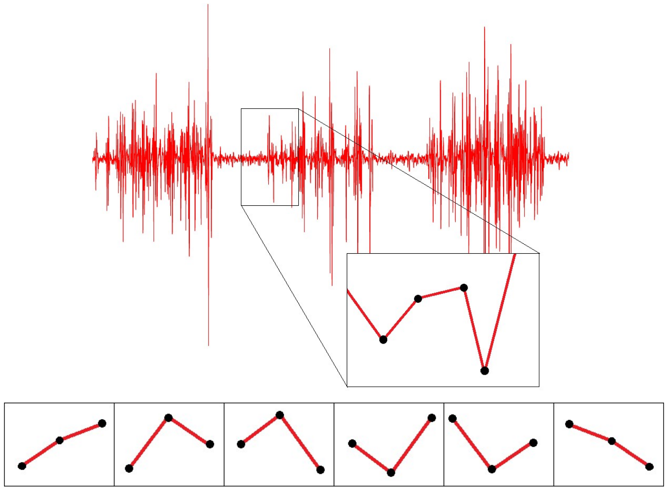
3 **Results**

Experimental Setup, Results and Discussion

Theoretical Background

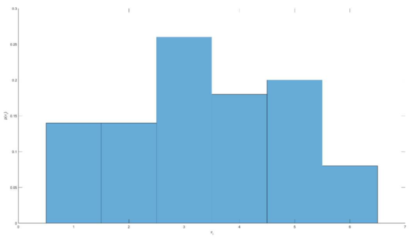
Multiscale Permutation Entropy

Permutation Entropy



All $d!$ possible ordinal patterns (for $d = 3$)

Probability of each pattern

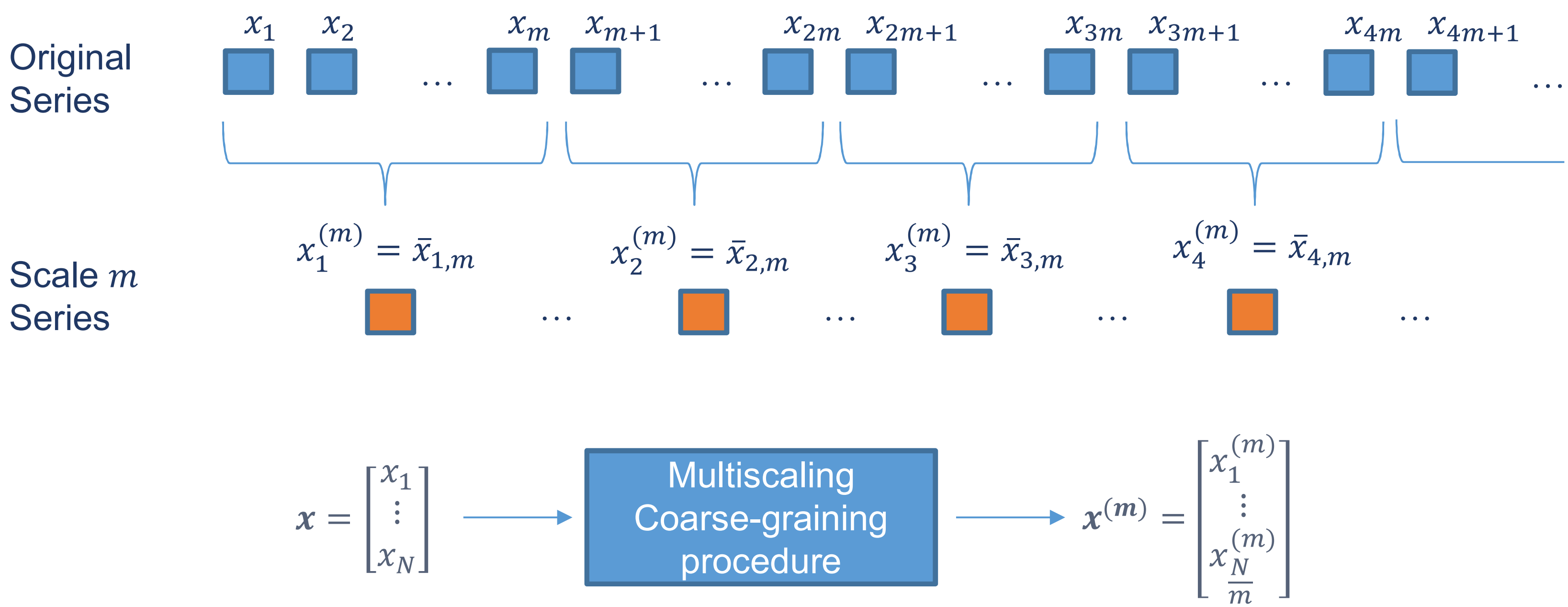


Permutation Entropy

$$\widehat{\mathcal{H}} = \frac{-1}{\ln(d!)} \sum_{i=1}^{d!} \hat{p}_i \ln(\hat{p}_i)$$

Bandt & Pompe, 2002

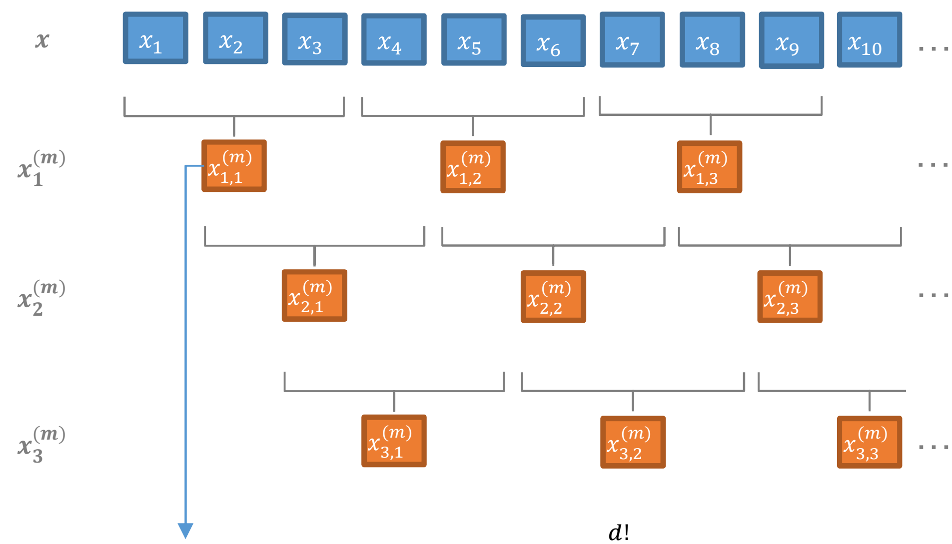
Multiscale Permutation Entropy



Aziz & Arif, 2005. Costa *et al.* 2002

Refined Composite Multiscale Permutation Entropy

- For $\mathrm{k} = 1, \ldots, m$, we can construct m different coarse signals by starting the classical coarse-graining procedure at element k. (Humeau-Heurtier *et al.* 2015)
- Improved Precision
- Artifact Cross-Correlation
    - $x_{k=1,1}^{(m=3)} = \frac{1}{m}(x_1 + x_2 + x_3)$
    - $x_{k=2,1}^{(m=3)} = \frac{1}{m}(x_2 + x_3 + x_4)$



$$x_{1,1}^{(m)} = \frac{x_1 + x_2 + x_3}{m}$$

$$H_{rc}(\hat{\boldsymbol{p}}^{(m)}) = \sum_{i=1}^{d!} \hat{\bar{p}}_i^{(m)} \, log(\hat{\bar{p}}_i^{(m)})$$

Multiscale Permutation Entropy Characteristics

Ordinal
Patterns are invariant to signal's amplitude.

Robust
Robust to signal's noise and artifacts.

Length Constraint
The precision of the pattern probability estimations decrease for short signals.

Artifact Cross-Correlation
For composite approaches, the use of redundant terms leads to MPE underestimation.

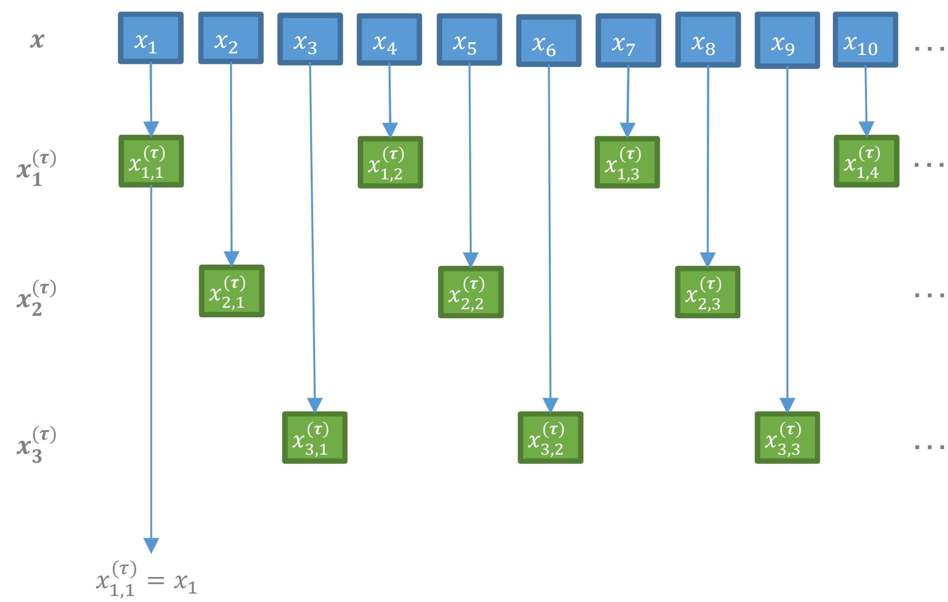
Refined Composite Downsampling Permutation Entropy

Composite Downsampling and Statistics comparison

Composite Downsampling

- We combine the classical downsampling procedure with the multiscaling.
- Build downsampled signals with different starting point.
- Retains improved precision from Composite Coarse-graining.
- Avoids Artifact Cross-correlation.

$$H_{rcd}(\hat{p}^{(\tau)}) = \sum_{i=1}^{d!} \hat{p}_i^{(\tau)} \, log(\hat{p}_i^{(\tau)})$$



Davalos *et al.*, 2020.

MPE Statistics

MPE Exp. Value

$$E[H_c(\hat{\boldsymbol{p}}^{(m)})] \approx H(\boldsymbol{p}^{(m)}) - \frac{1}{2}\ (d! - 1)\left(\frac{m}{N}\right)$$

rcMPE Exp. Value

$$E[H_{rc}(\hat{\boldsymbol{p}}^{(m)})] \approx H(\boldsymbol{p}^{(m)}) - \frac{1}{2}\ (d! - 1)\left(\frac{1}{N}\right)$$

rcDPE Exp. Value

$$E[H_{rcd}(\hat{\boldsymbol{p}}^{(\tau)})] \approx H(\boldsymbol{p}^{(\tau)}) - \frac{1}{2}\ (d! - 1)\left(\frac{1}{N}\right)$$

MPE Variance

$$var\left(H_c(\hat{\boldsymbol{p}}^{(m)})\right) \approx \left(\frac{1}{N}\right) m\, \boldsymbol{l}^{(m)\prime}\, \boldsymbol{\Sigma}_p^{\{(m)\}}\, \boldsymbol{l}^{(m)} + \left(\frac{1}{N}\right)^2 m^2 \left(\boldsymbol{1}'\, \boldsymbol{l}^{(m)} + d!\, H(\boldsymbol{p}^{(m)}) + \frac{1}{2}\ (d! - 1)\right)$$

rcMPE Variance

$$var\left(H_{rc}(\hat{\boldsymbol{p}}^{(m)})\right) \approx \left(\frac{1}{N}\right) \boldsymbol{l}^{(m)\prime}\, \boldsymbol{\Sigma}_p^{\{(m)\}}\, \boldsymbol{l}^{(m)} + \left(\frac{1}{N}\right)^2 \left(\boldsymbol{1}'\, \boldsymbol{l}^{(m)} + d!\, H(\boldsymbol{p}^{(m)}) + \frac{1}{2}\ (d! - 1)\right)$$

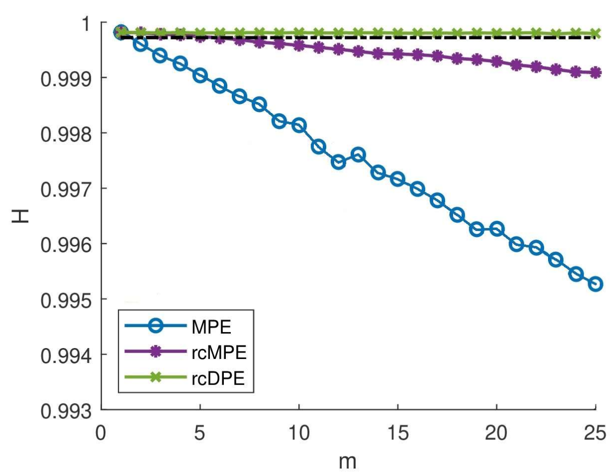
rcDPE Variance

$$var\left(H_{rcd}(\hat{\boldsymbol{p}}^{(\tau)})\right) \approx \left(\frac{1}{N}\right) \boldsymbol{l}^{(\tau)\prime}\, \boldsymbol{\Sigma}_p^{\{(\tau)\}}\, \boldsymbol{l}^{(\tau)} + \left(\frac{1}{N}\right)^2 \left(\boldsymbol{1}'\, \boldsymbol{l}^{(\tau)} + d!\, H(\boldsymbol{p}^{(\tau)}) + \frac{1}{2}\ (d! - 1)\right)$$

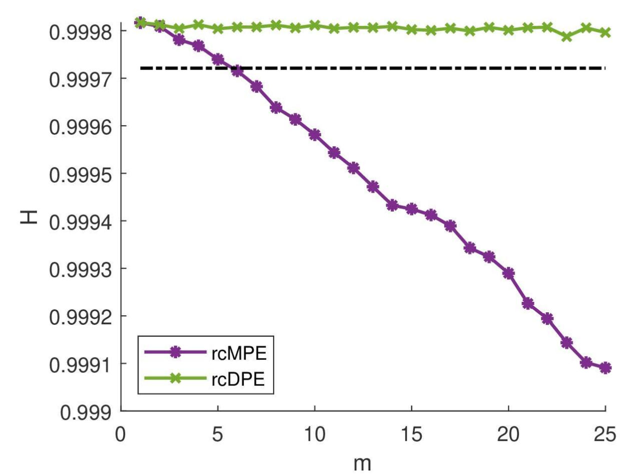
Embedding dimension d, time scale m, signal length N, and probability p.

MPE Expected Value – White Noise

All methods

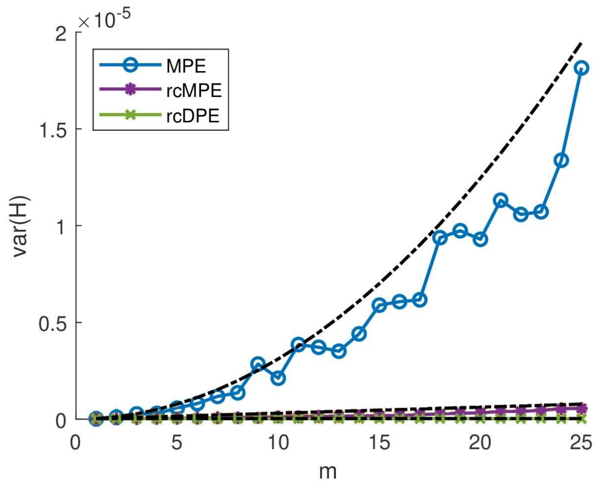


Refined methods

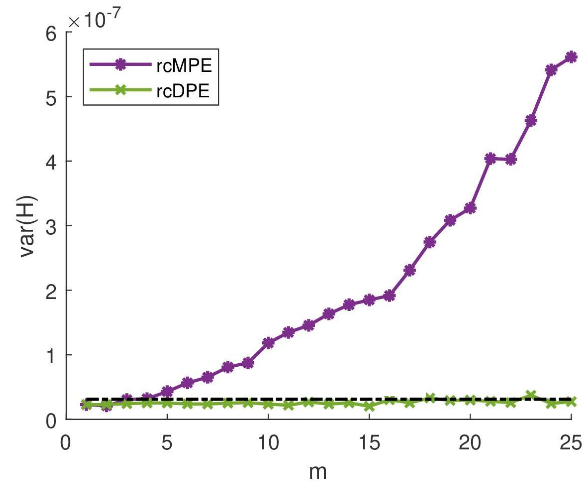


MPE Variance – White Noise

All methods



Refined methods



Main Findings

Artifact Cross-correlation
Limits the refined composite precision

Downsampling
Avoids redundancy by using composite downsampling instead of coarse-graining

Variance
rcDPE presents the lowest variance of all the methods discussed so far. It outperforms classic MPE by two orders of magnitude on white noise.

Scale-invariant
rcDPE is the only method here discussed with no effects related to scale.

Results

Experimental Setup, Results,
and Discussion.

Experimental Setup

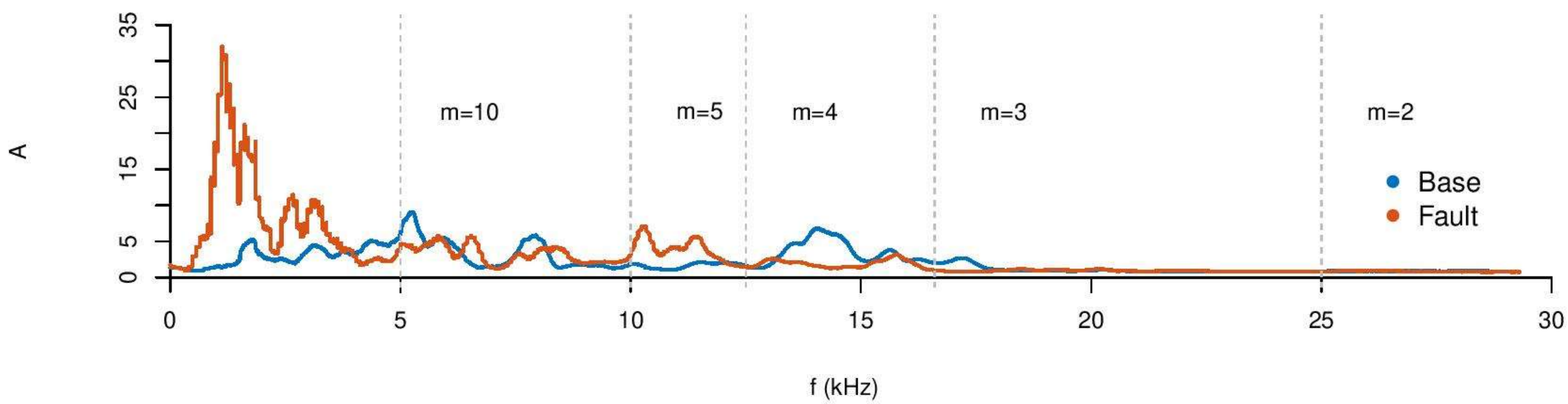
Dataset

- Bearing fault signal dataset (Bechhoeffer 2013).
- Sampling frequency 100kHz for 6 seconds.
- Bearing test rig with 270 lb load.
- Rotation frequency 25 Hz.
- 3 signals with labels:
  - Base: no defects.
  - Fault: with defects

Methods

- 3-way ANOVA test.
- Factors:
  - Type: Base and Fault
  - Methods: MPE, rcMPE, rcDPE, and rcDPE (filtered)
  - $d = 3, \ldots, 6$.
- $m = 1, \ldots, 20$, with $\tau = m$.

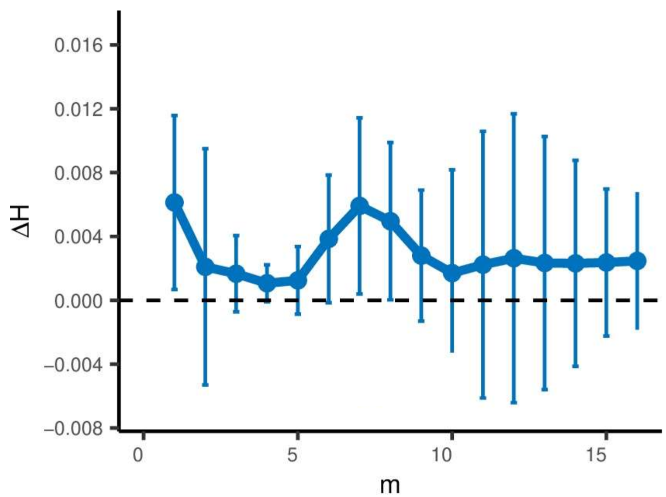
Sample Signal Spectrum



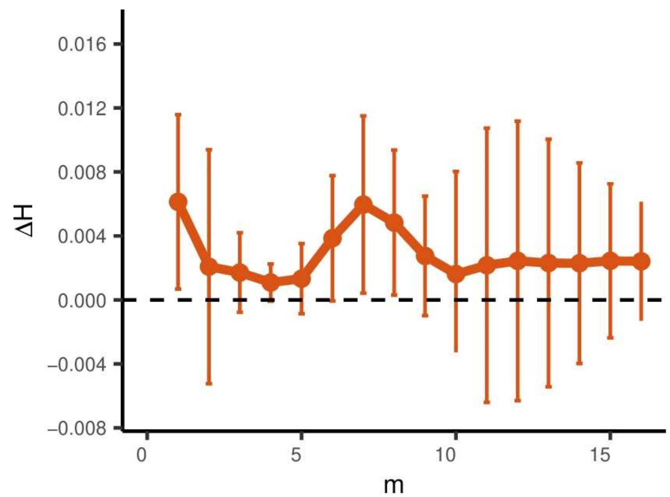
* Dotted lines showcase the Nyquist frequency at each scale. For m>3, we experience aliasing.

Entropy Difference $\Delta\mathrm{H} = \overline{H}_{\mathbf{Base}} - \overline{H}_{\boldsymbol{Fault}}$

MPE



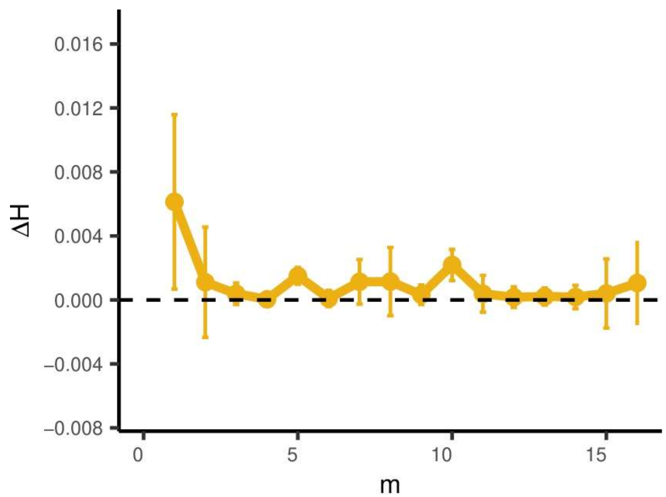
rcMPE



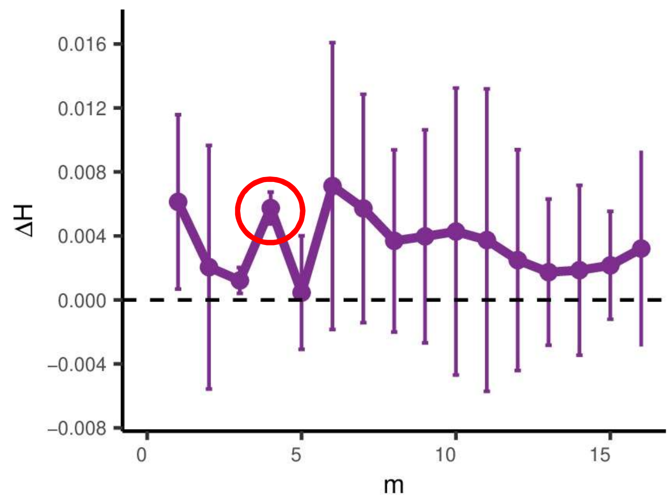
Error bars computed using $S_{\Delta H} = \pm 1{,}96\sqrt{S_{H,Base}^2 + S_{H,Fault}^2}$ with significance $\alpha = 0{,}05$

Entropy Difference $\Delta H = \overline{H}_{\mathrm{Base}} - \overline{H}_{Fault}$

rcDPE



rcDPE – Aliasing Filter



Error bars computed using $S_{\Delta H} = \pm 1{,}96\sqrt{S_{H,Base}^2 + S_{H,Fault}^2}$ with significance $\alpha = 0{,}05$

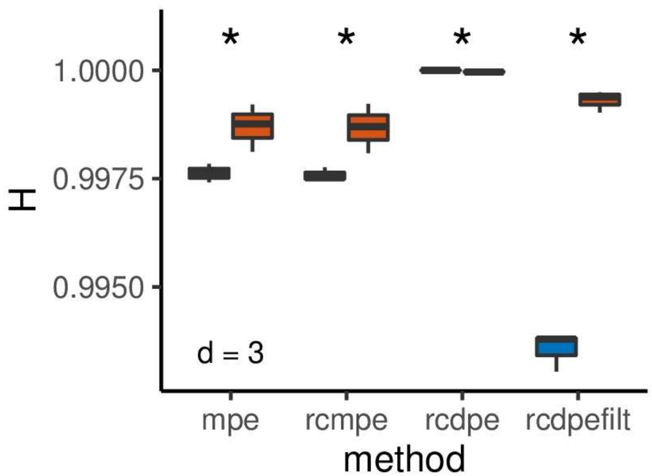
	Df	Sum Sq.	Mean Sq.	F-value	P-Value	Significant
Type	1	0.0000358	0.0000358	12.346	**0.000730**	***
Method	3	0.0007515	0.0002505	86.392	**< 2e-16**	***
Dimension	1	0.0004970	0.0004970	171.397	**< 2e-16**	***
Type & Method	3	0.0003514	0.0001171	40.398	**5.37e-16**	***
Type & Dimension	1	0.0000410	0.0000410	14.134	**0.000322**	***
Method & Dimension	3	0.0000955	0.0000318	10.981	**4.09e-06**	***
Type & Method & Dimension	3	0.0000219	0.0000073	2.514	0.064282	
Residuals	80	0.0002320	0.0000029			

Significance $\alpha = 0{,}05$

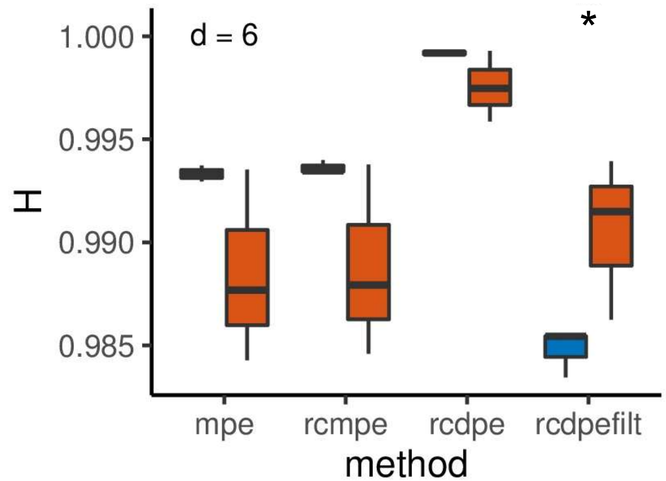
ALL FACTORS AND INTERACTIONS ARE SIGNIFICANT

Base-Fault Classification

d=3



d=6



Significant with $\alpha = 0{,}05$, $\mathrm{m} = \tau = 4$

Main Findings

Optimal Settings

rcDPE with aliasing filter
Time scale m=4
Low dimension (d=3)

MPE behavior

Unfiltered rcDPE has the lowest variation, but erases the difference between faulty and non-faulty signals.

MPE dimension

For this particular dataset, higher dimensions actually perform worse than lower dimensions. Filtered rcDPE maintains significant differences.

Aliasing

There is a trade-off between statistical precision and aliasing effects.

Conclusions

Main Insights

Conclusion

- The refined Composite Downsampling Permutation Entropy improves the precision of the classical rcMPE by avoiding the artifact cross-correlations, product of the signal's preprocessing.
- rcDPE enhances the classification between faulty and non-faulty bearings in rotatory machines.
- In order to avoid aliasing effects, we applied an anti-aliasing filter matching the Nyquist frequency for each time scale. This allowed a better classification, even with increased variance compared to the unfiltered rcDPE method.
- For this dataset, lower dimensions allowed a better classification of the faulty components. Nonetheless, the filtered rcDPE method successfully detected the difference, even for higher dimensions.

References

- W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in 2005 Pakistan Section Multitopic Conference, 2005, pp.1–6.
- A. Humeau-Heurtier, C.-W. Wu, and S.-D. Wu, "Refined composite multiscale permutation entropy to overcome multiscale permutation entropy length dependence" *IEEE Signal Processing Letters*, vol. 22, pp. 2364{2367, 12 2015.
- C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," Physical Review Letters, vol. 88, no. 17, p. 174102, 2002.
- M. Costa, A. Goldberg, and C. Peng, "Multiscale entropy analysis of biological signals," Phys. Rev. E 71, 021906, 2005.
- A. Davalos, M. Jabloun, P. Ravier, and O. Buttelli, "Improvement of Statistical Performance of Ordinal Multiscale Entropy Techniques Using Refined Composite Downsampling Permutation Entropy," *Entropy* **2021**, *23*(1), 30, 2020.
- Bechhoefer, E. Fault Data Sets. Society For Machinery Failure Prevention Technology. Vibration Institute. 518 2013. https://www.mfpt.org/fault-data-sets/, 2013.

Thank you!

Q&A Session