Anomaly detection and noise reduction of urn-by-turn BPM signal

F. Bugiotti, A. Guelfane, Y. Zuo, I. Cherkaoui, J. KPODJADAN
(CentraleSupelec and Paris Saclay University)

Q. Bruant, B. Dalena, V. Gautard (CEA irfuand Paris-Saclay University)

Thanks to: J. Keintzel, M. Gael, Y. Ohnishi

Outline

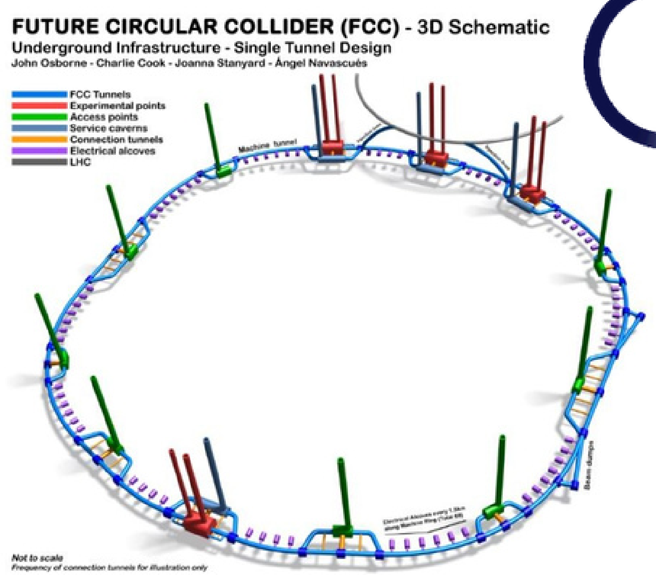
- Motivation and Context
- Turn by turn BPMs analysis pipeline
- Anomaly detection and noise reduction
- Result
- Conclusion

Motivation

- Application of optimization cost and/or performance of future high energy colliders to the case of FCC/SuperKEKB collider
- Correction of linear and non-linear imperfections
- Complex simulations studies that aims to predict interplay of different physics effects
- Study limitations from existing accelerators

➢ Exploit AI techniques to reduce computing resources and increase accuracy of measurements

FCC and SuperKEKB



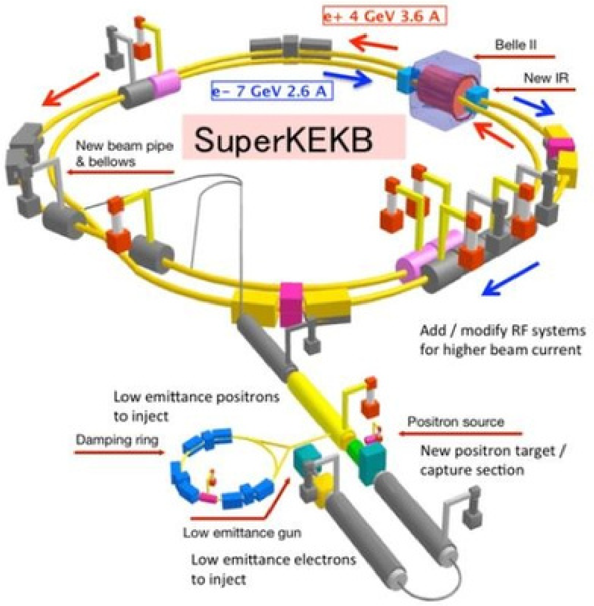
FUTURE CIRCULAR COLLIDER

International FCC project (CERN as host lab)

Continue to study experimental high energy particle physics
Accelerator options: FCC-hh, FCC-ee, FCC-he

CDRspublié dans **EuropeanPhysical Journal C (Vol 1) and ST (Vol 2-4)**

http://fcc-cdr.web.cern.ch/

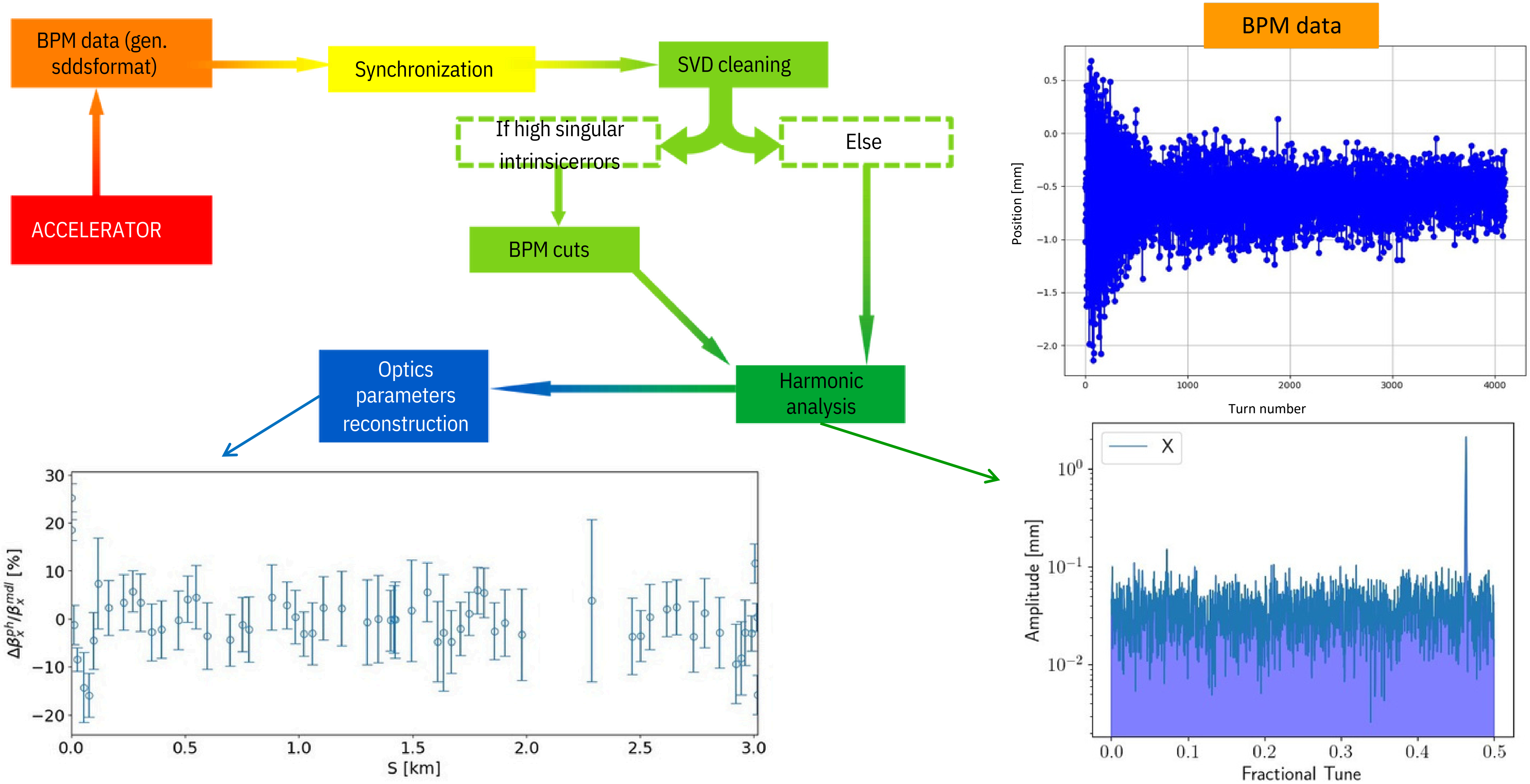


KEK

Existing e+e- collider:
small size FCC-ee
proofs of principle of several design choices

https://doi.org/10.1093/ptep/pts083

TbT BPM data



BPM data (gen. sddsformat)
ACCELERATOR
Synchronization
SVD cleaning
If high singular intrinsicerrors
Else
BPM cuts
Harmonic analysis
Optics parameters reconstruction
BPM data
Position [mm]
Turn number
Amplitude [mm]
Fractional Tune
X
$\Delta\beta_x^{ph}/\beta_x^{mdl}$ [%]
S [km]

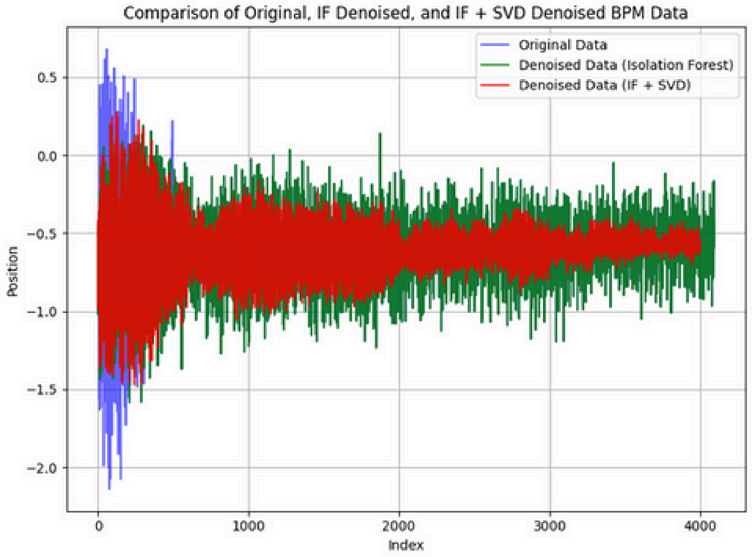
Isolation Forest + SVD denoising and anomaly detection

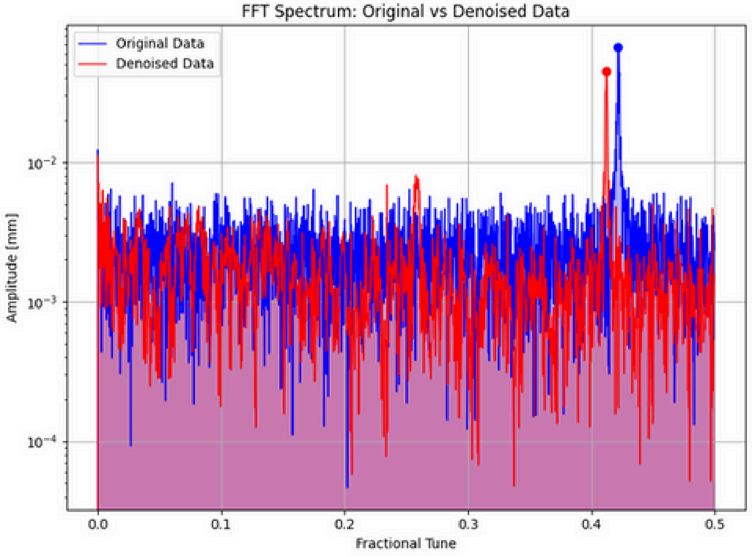
Isolation Forest (IF): identify and remove outliers
SVD : retains dominant singular values ⍰ Improved S/N in FFT of denoised data by IF+SVD

3 features are extracted to identify anomaly with IF:
The betatron tune (the main peak in the FFT spectrum), its amplitude and the signal to noise ratio

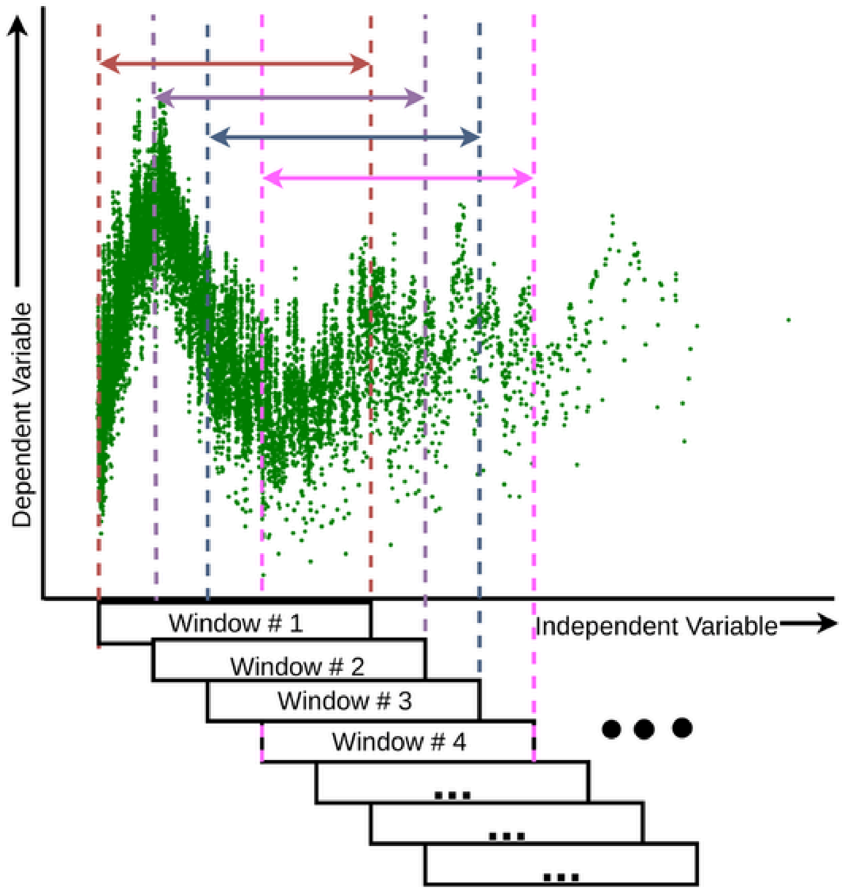
Table 1: Suspected Faulty or Faulty BPMs Detected

BPM Name	Classification	IF Anomaly Count
MQD3E18	suspected faulty	2
MQEAE23	suspected faulty	1
MQEAE35	faulty	5
MQLA2RE	suspected faulty	2
MQLC7RE	suspected faulty	1

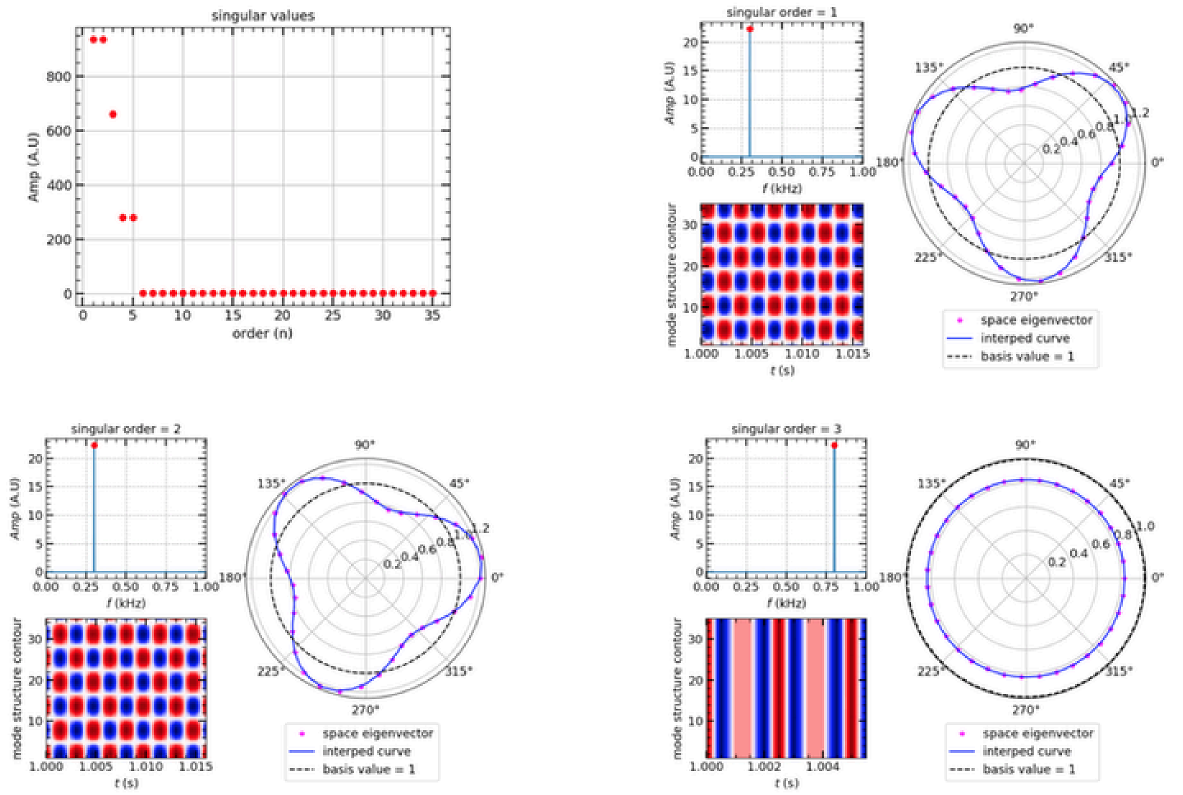




Denoising - SVD (on raw data)



Split the 1D data into overlapping windows



SVD: Retain dominant singular values and reconstruct signal

Denoising - IF

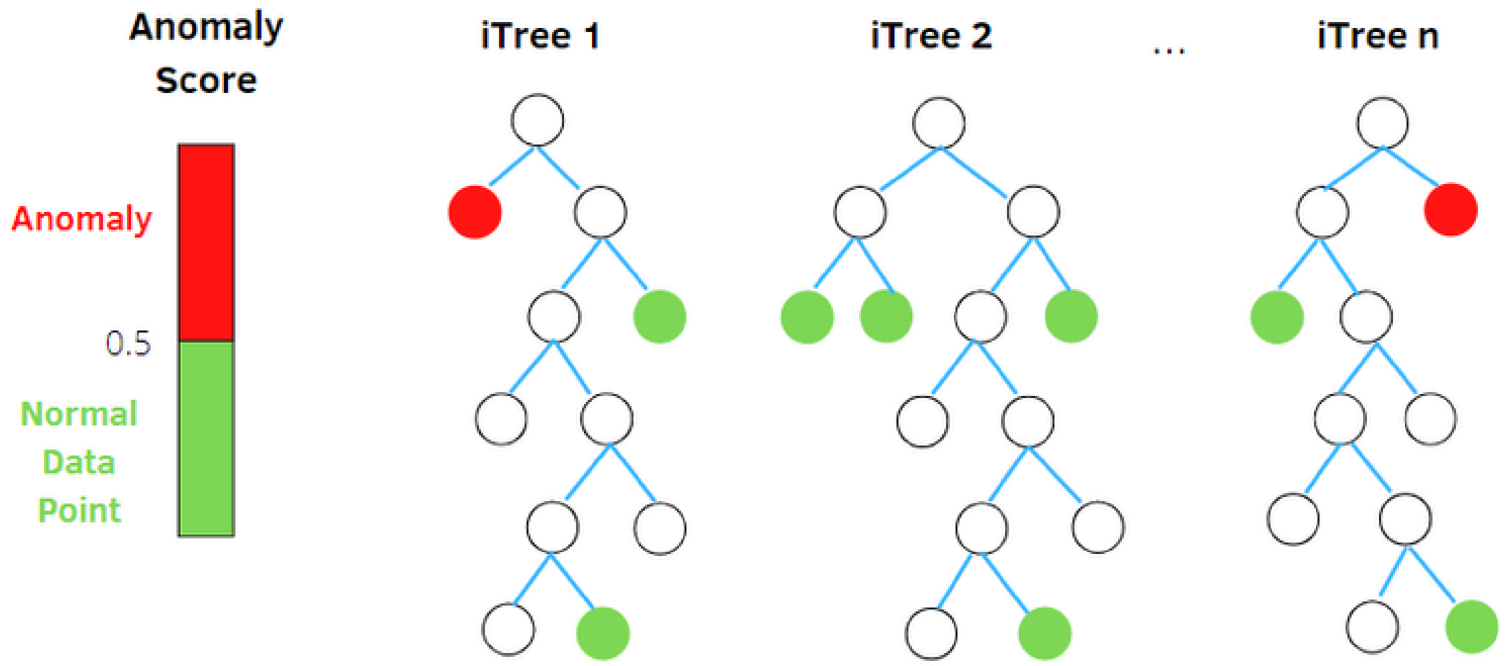


Figure 6: How Isolated forest works

Denoising - IF + SVD (on raw data)

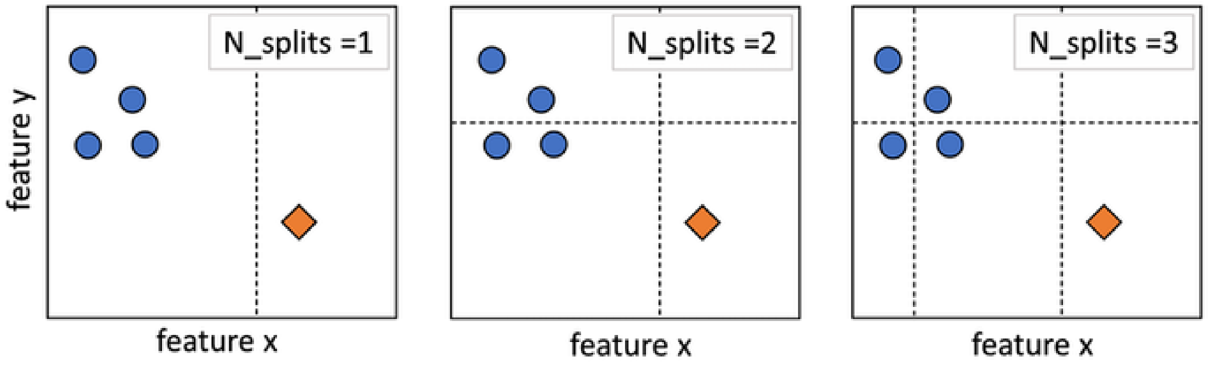
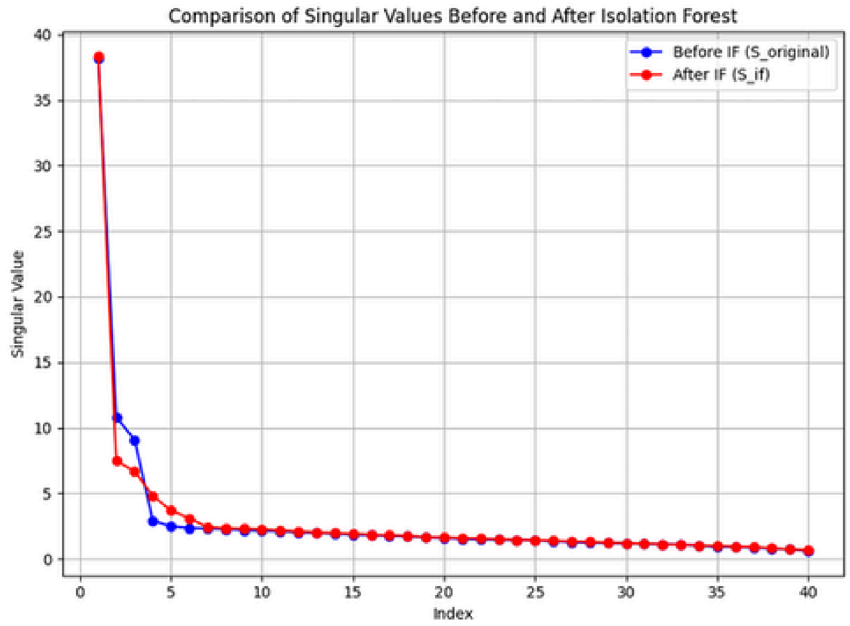
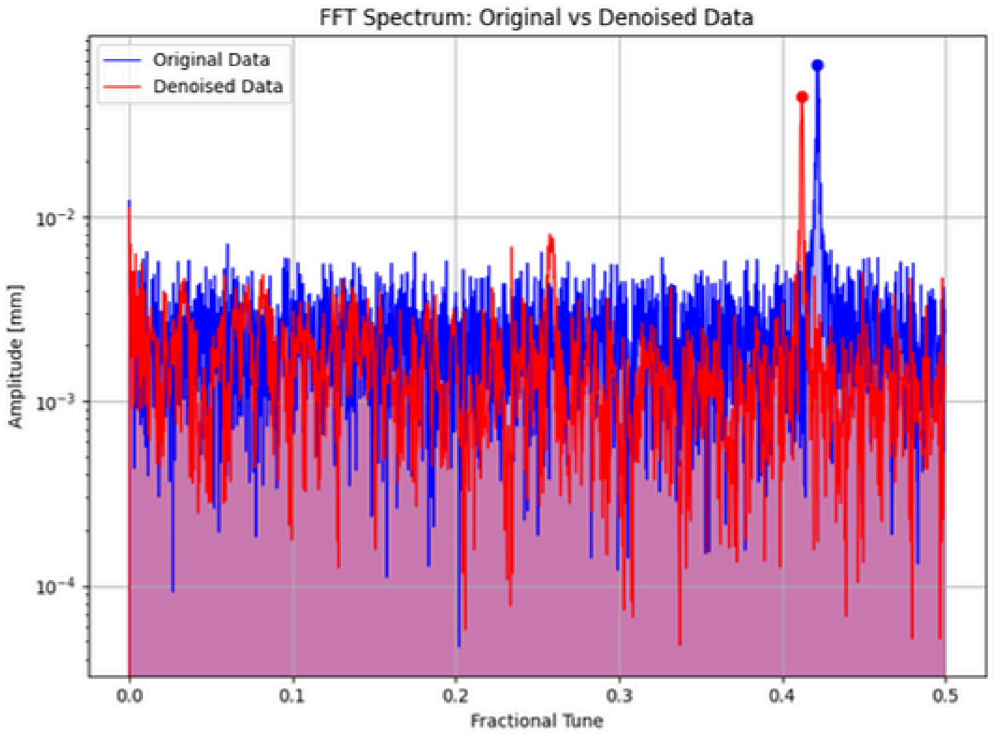


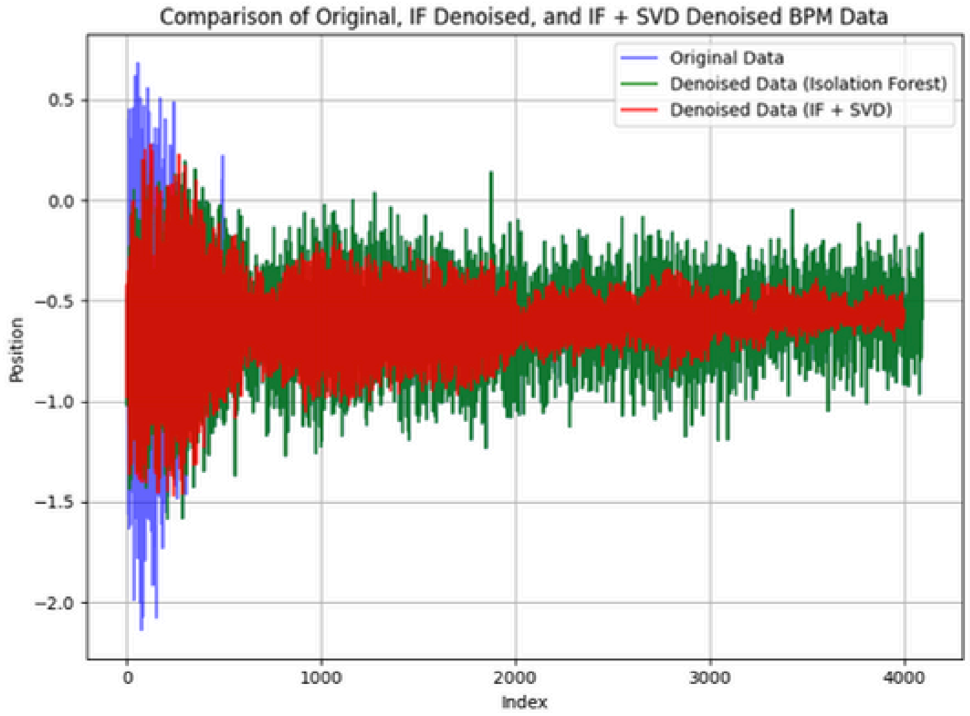
Figure 2: A conceptual illustration of the IF algorithm.



A comparison of the singular values before and after applying the IF

Result :Denoising - IF + SVD (on raw data)

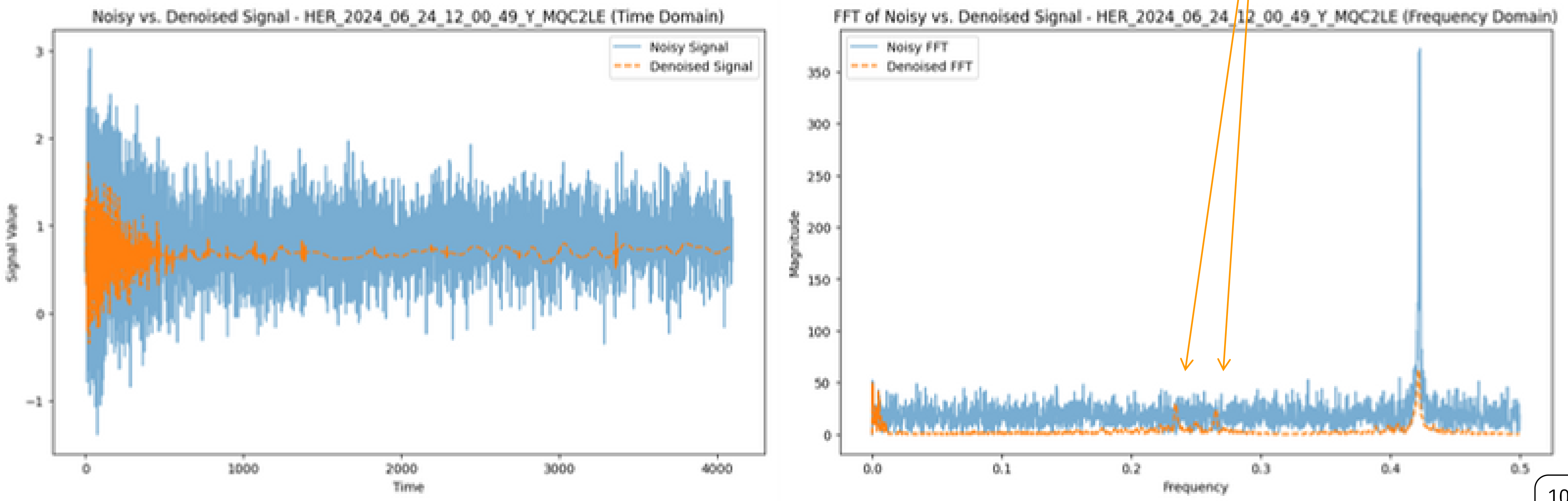




Wavelets and LSTM denoising

1. Data Normalization
2. Wavelet decomposition: Daubechies wavelet (db4)
3. Noise estimation (MAD)
4. Adaptive soft thresholding: $\lambda = \sigma\sqrt{2\log n}$
5. Wavelet reconstruction: inverse wavelet transform
6. Post-processing with a LSTM Neural Network

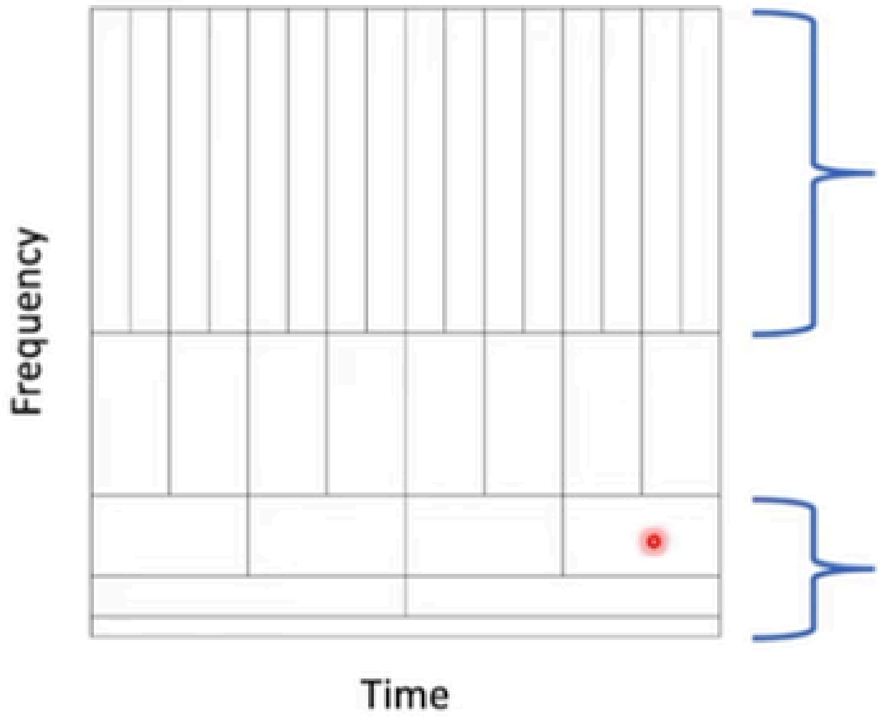
- Denoised signal shows interesting peaks



Denoising - Wavelet Transform

The Wavelet Transform

The Wavelet Transform results in analyzing a signal into different frequencies at different resolutions, known as multiresolution analysis.



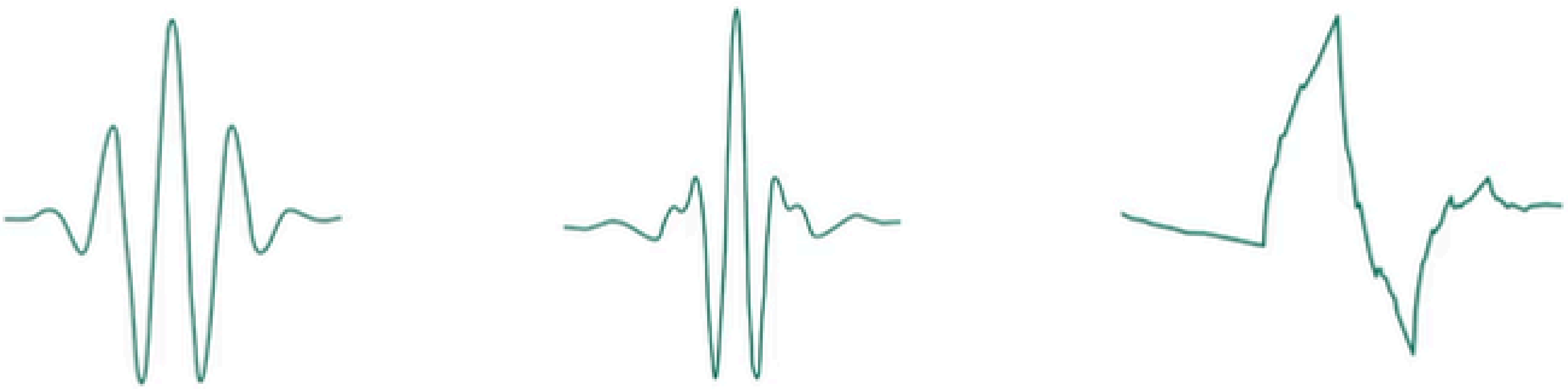
Good time resolution and poor frequency resolution at high frequencies.

Good frequency resolution and poor time resolution at low frequencies.

Denoising - Wavelet Transform

The Wavelet

The wavelet is a small wave such as these...



The Wavelet is our new basis function, and acts as a window function.

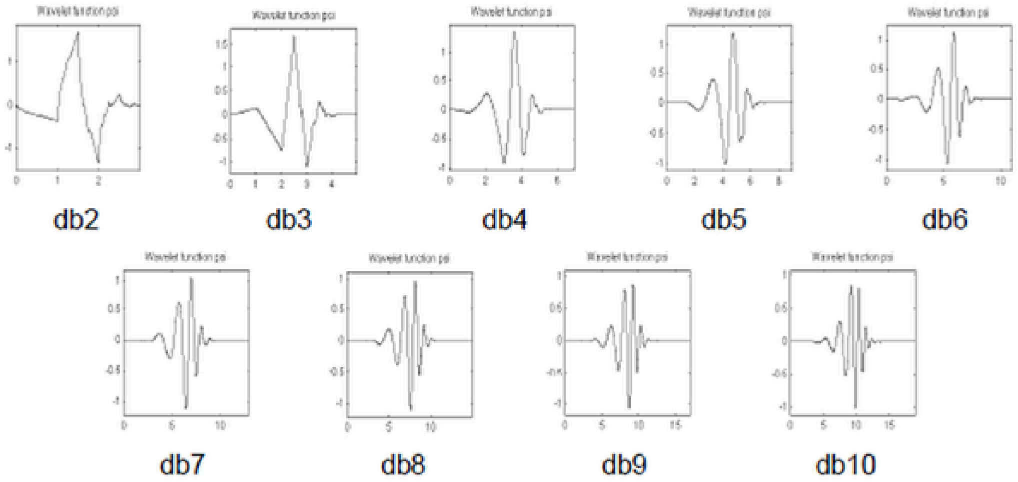
Denoising - Wavelet Transform

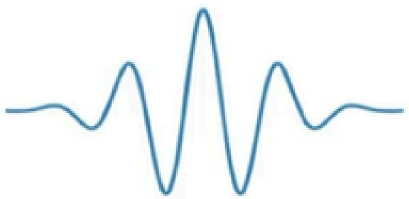
We can change the width of the wavelet and its central frequency as we move it across our signal by changing s. This is called scaling.

Obtained from the function $\psi(t)$, known as the wavelet.

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t)\, \psi^*\left(\frac{t-\tau}{s}\right) dt$$

Scale Parameter (1/frequency)





Expanded wavelet is better at resolving low frequency components of the signal with bad time resolution. (Large values of s)

Shrunken wavelet is better at resolving high frequency components of the signal with good time resolution. (Small values of s)

Wavelets on irregular point sets. Ingrid Daubechies, Igor Guskov, Peter Schröder, Wim Sweldens. Phil.Trans.Roy.Sec.A 357,2397–2413 (1999).

Denoising - Wavelet Transform + AI (on raw data)

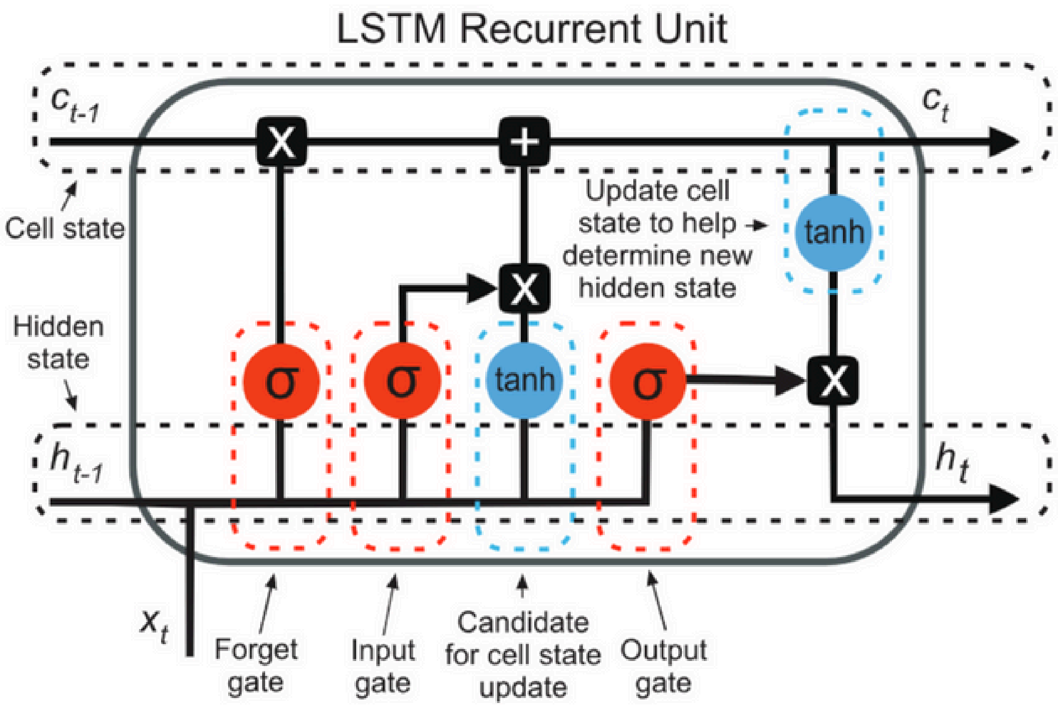


Figure 7: LSTM Model

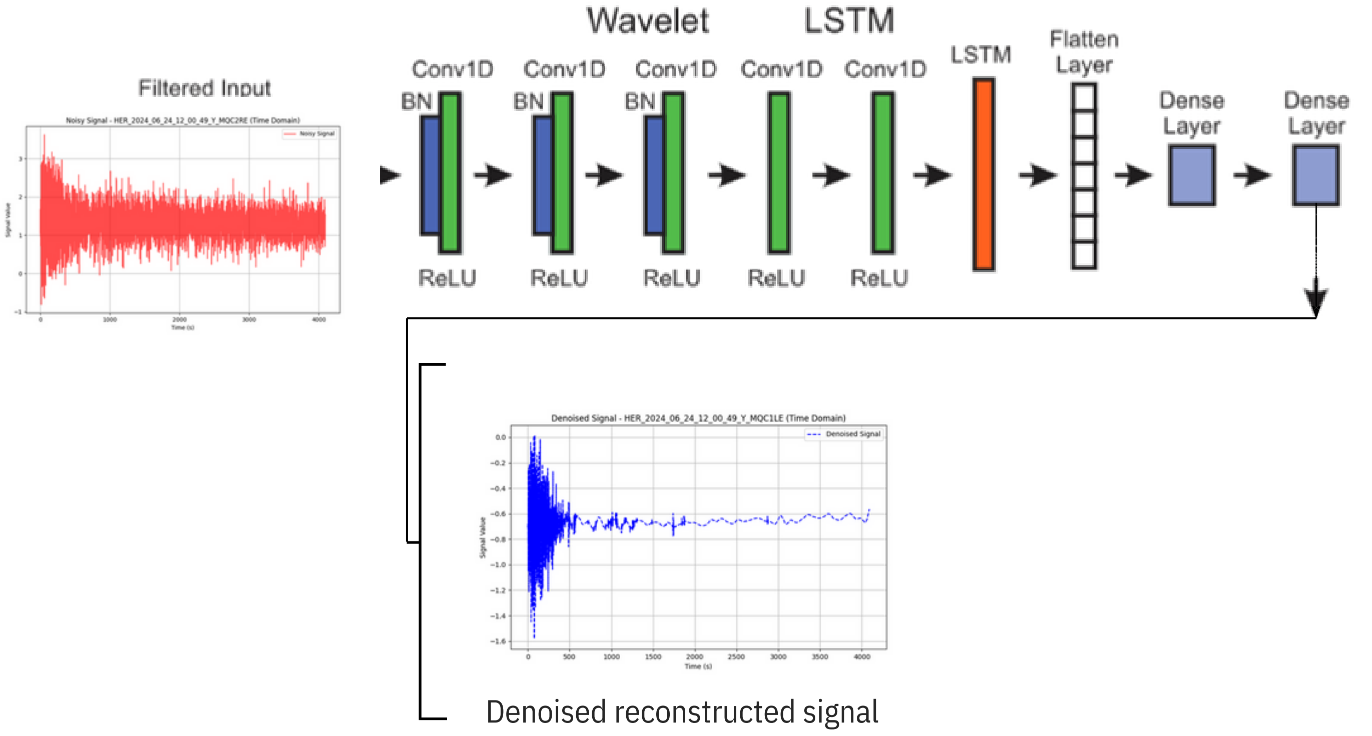
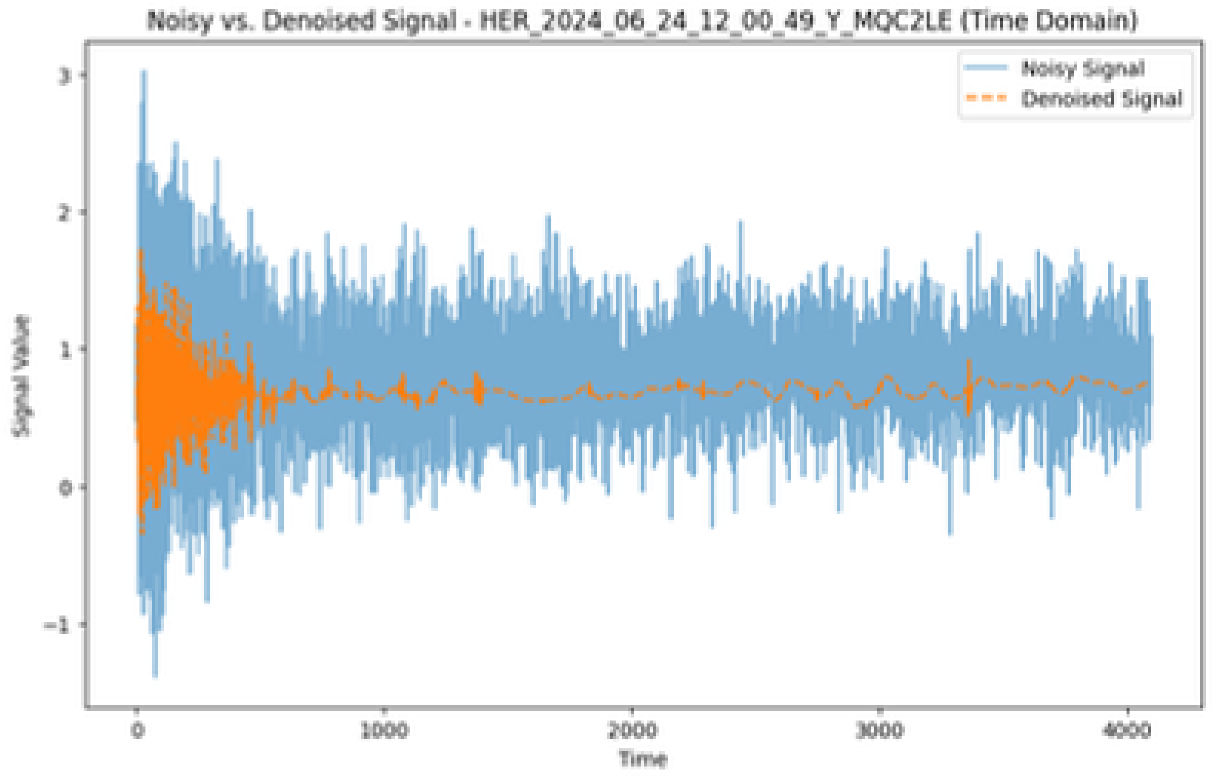
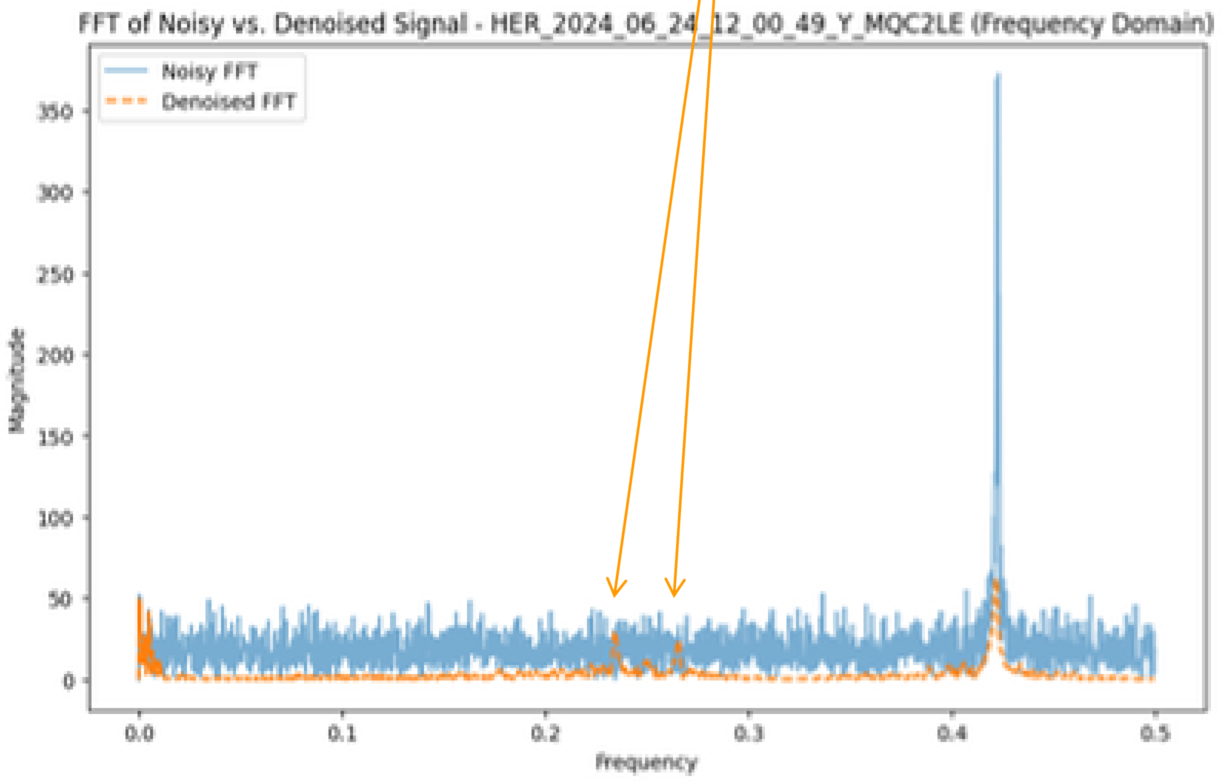


Figure 8: Pipeline of our Model

Result - Wavelet Transform + AI

Denoised signal shows interesting peaks





Denoising - Wavelet Transform + AI : Model Validation

F(t) --> F(t) + White_noise(t)

The SNR[2] compares the power of the signal to the power of the noise. It is given by:

$$\text{SNR (dB)} = 10 \cdot \log_{10}\left(\frac{\|s_{\text{clean}}\|^2}{\|s_{\text{clean}} - s_{\text{denoised}}\|^2}\right)$$

Signal-to-Noise Ratio (SNR)

SNR = 8.61 dB

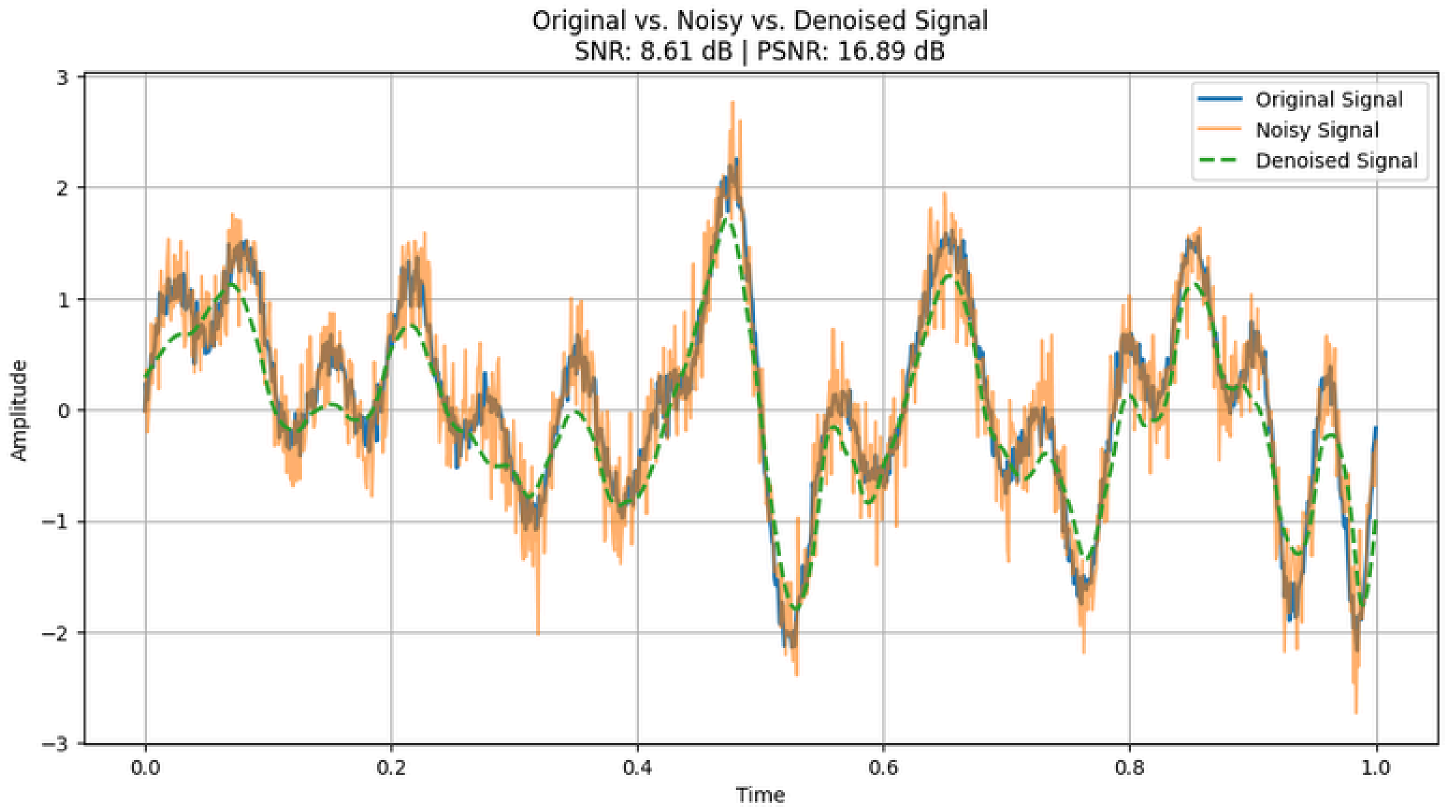


Figure 9: Model Validation

[2] **:** Source : Sherman, C., & Butler, J. (2007). Transducers and Arrays for Underwater Sound. Springer New York. ISBN: 9780387331393. Retrieved from https://books.google.fr/books?id=srREi-ScbFcC

Perspectives

- Further understanding and improvements of the algorithms:
  - Optics function reconstruction of denoised data
  - Comparison with OMC* FFT spectrum
- Investigate other AI techniques
- Understand the possible source (physics ?) of the two peaks in 0.2-0.3 frequency region

* Cern software using harpy library https://github.com/pylhc/omc3/tree/master/omc3

Thank you !