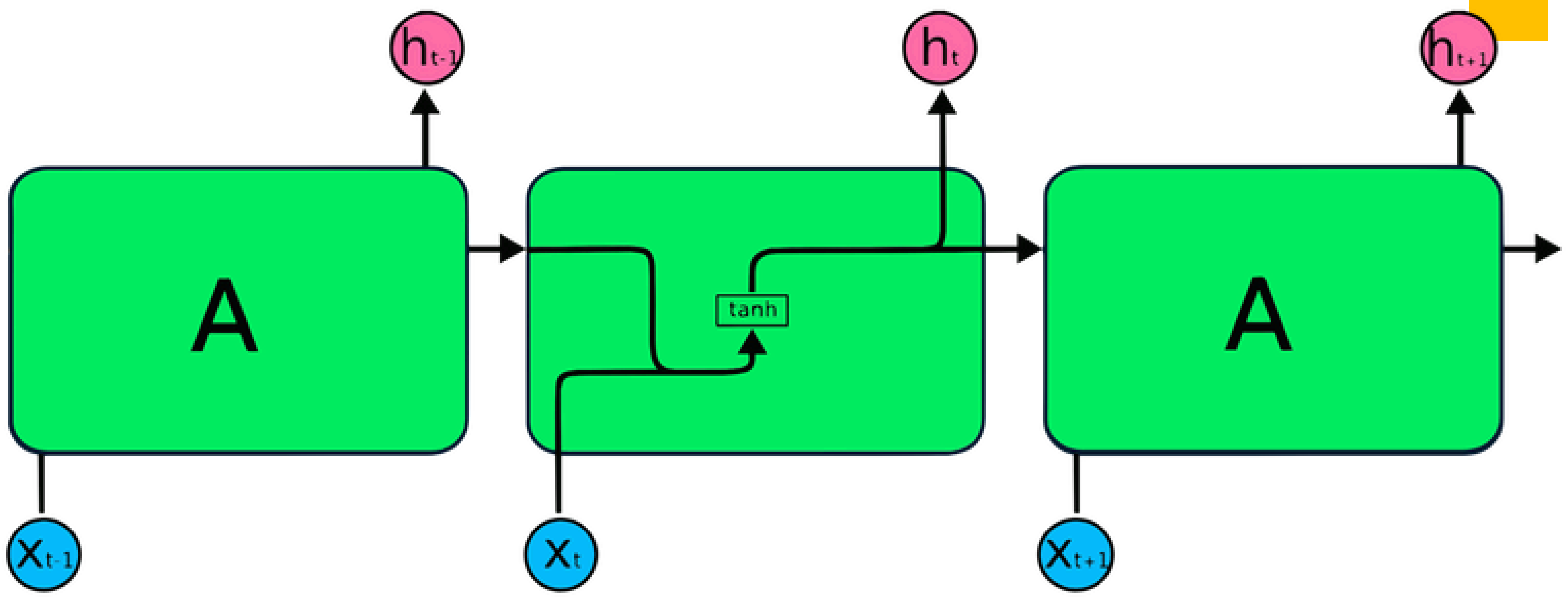
Long Short Term Memory

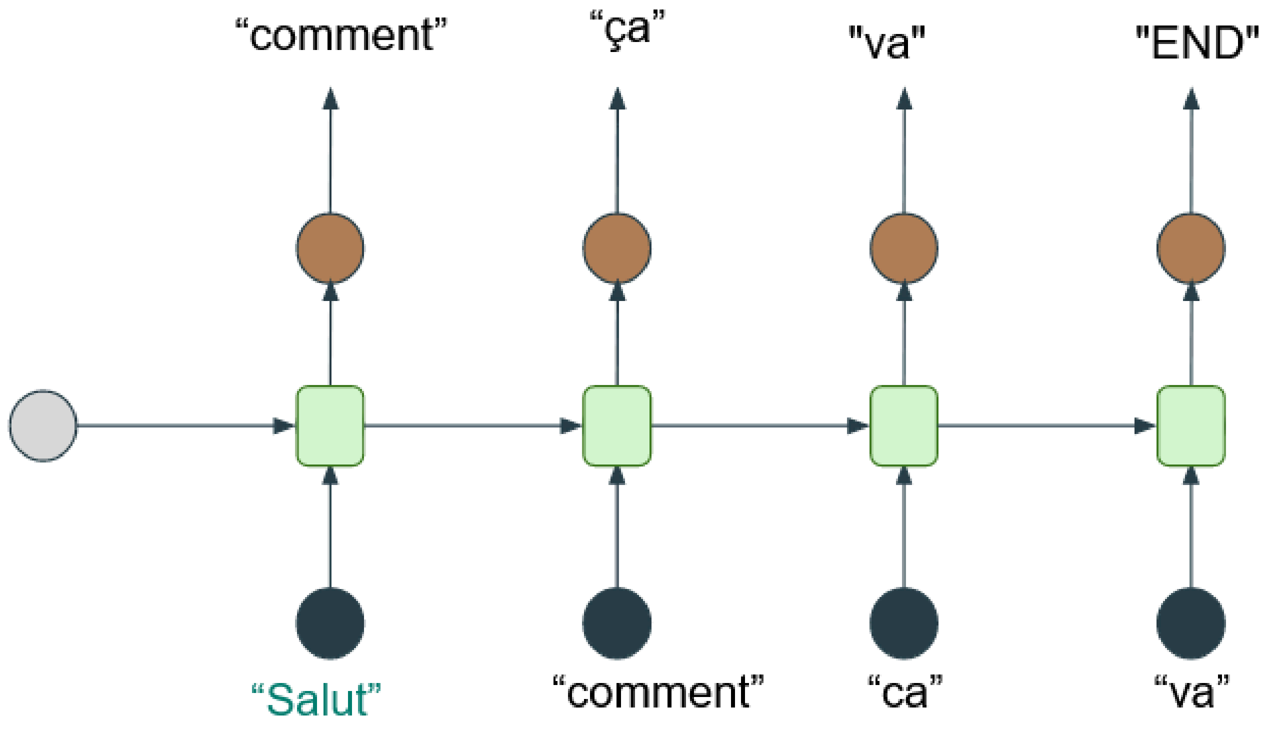
LSTM Networks for gravitational waves detection

Joubert
Gaspard

Recurent Neural Networks

- RNN remembers past inputs due to an internal memory
- Unidirectional recurrent neural networks cannot account for future events in their predictions
- Vanishing Gradient Problem





https://plainenglish.io/community/vanishing-gradient-problem-in-rnns-9d8e14
https://penseeartificielle.fr/comprendre-lstm-gru-fonctionnement-schema/gif-rnn-fonctionnement-cellule-rnn/
https://www.ibm.com/topics/recurrent-neural-networks?mhsrc=ibmsearch_a&mhq=recurent%20neural%20network

Long Short Term Memory

Forget Gate

Determines which information from the previous cell state should be discarded or forgotten.

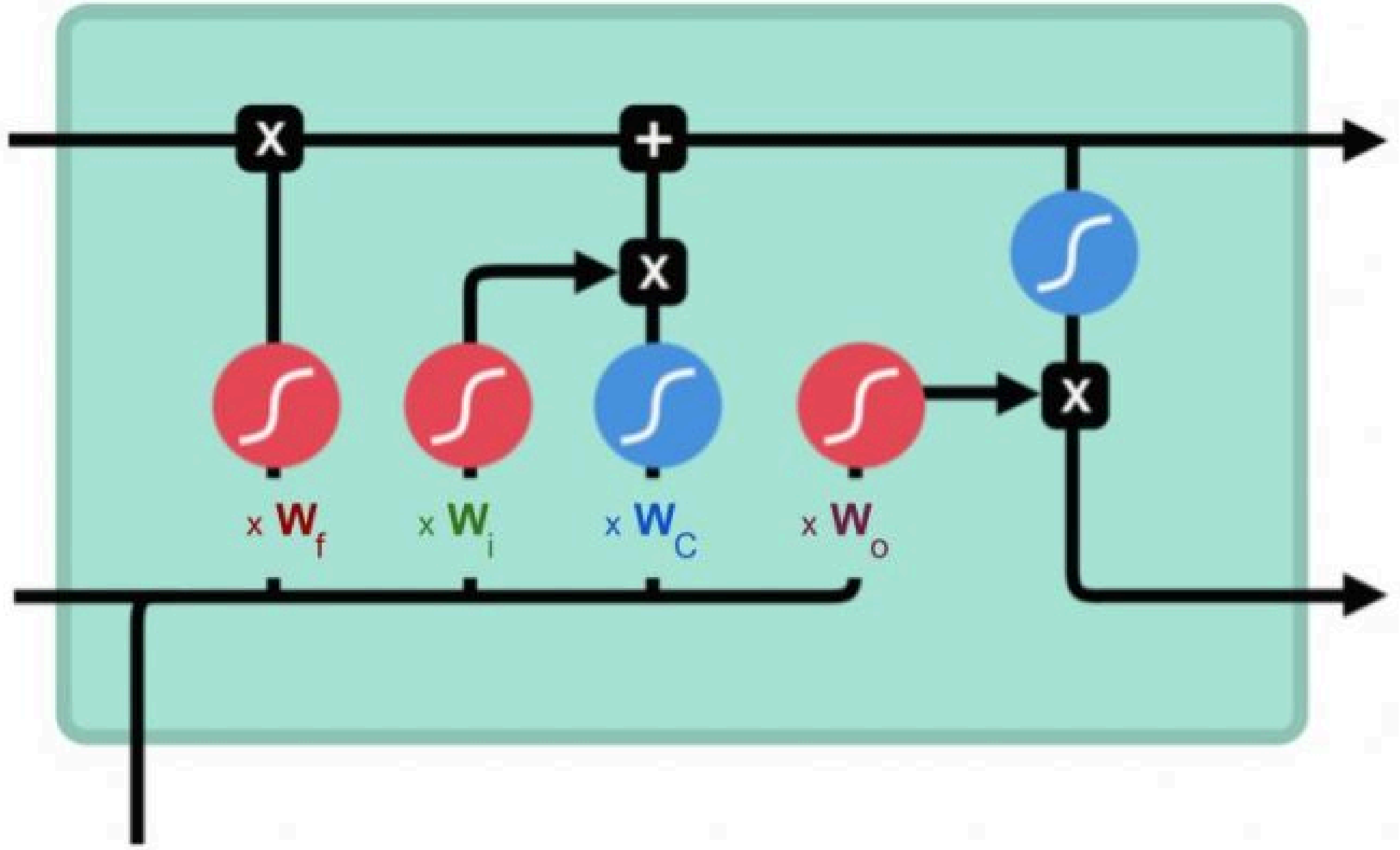
Input Gate

Determines which new information should be stored in the cell state.

Cell State

Output Gate

Determines what part of the cell state should be exposed as the output.

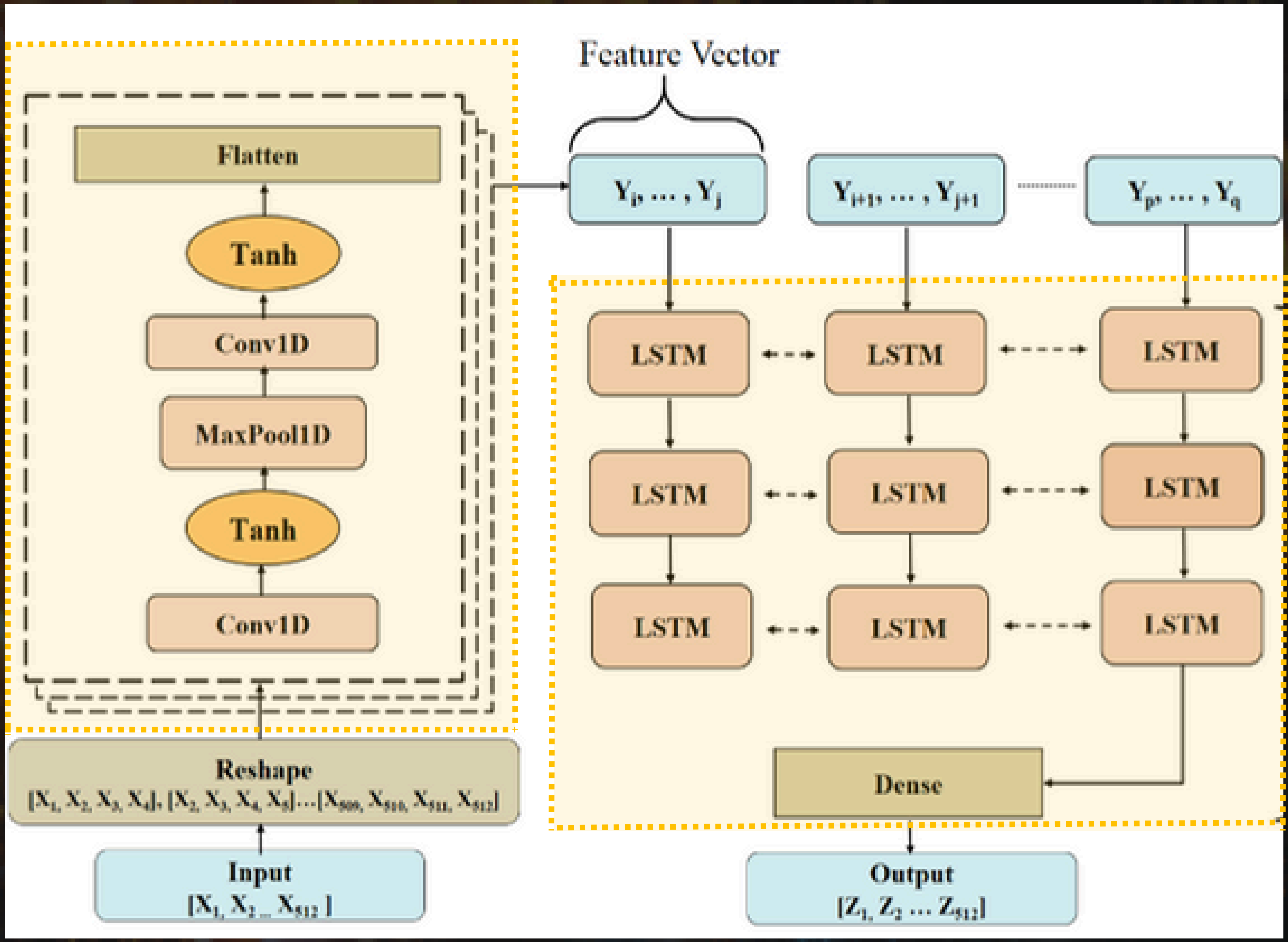


LSTM Cell

CNN-LSTM

Extraction of Binary Black Hole Gravitational Wave Signals from Detector Data Using Deep Learning

Chatterjee, Chayan and Wen, Linqing and Diakogiannis, Foivos and Vinsen, Kevin



CNN encoder

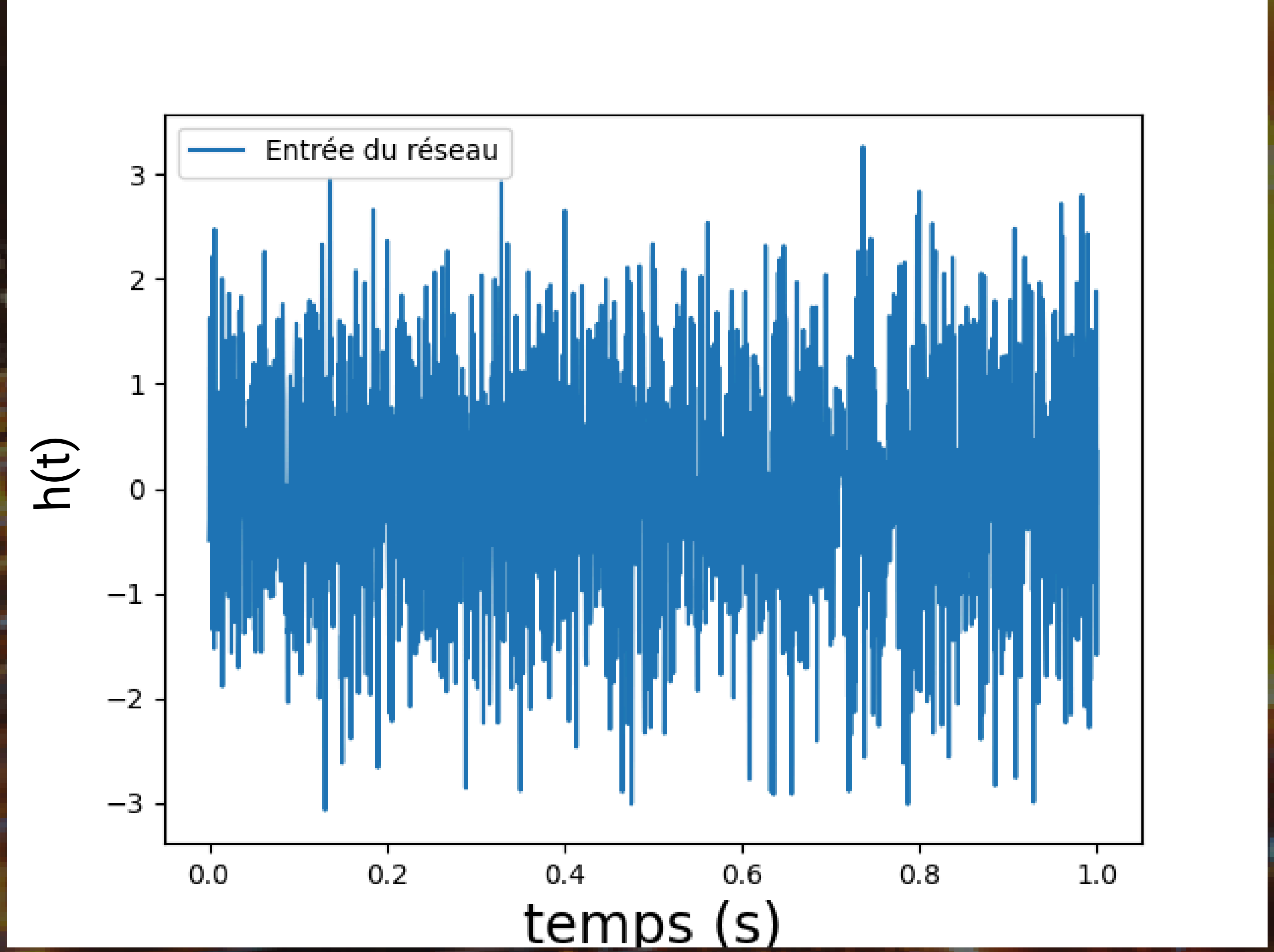
Extract the main features of th noisy signal

Bi-LSTM decoder

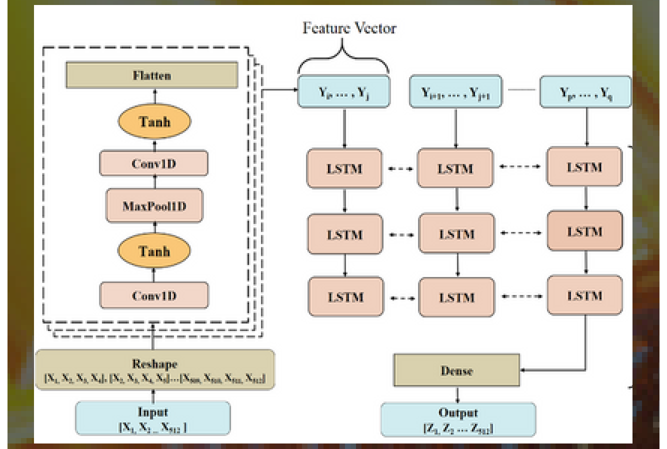
Reconstructs the denoised signal from the features

http://dx.doi.org/10.1103/PhysRevD.104.064046

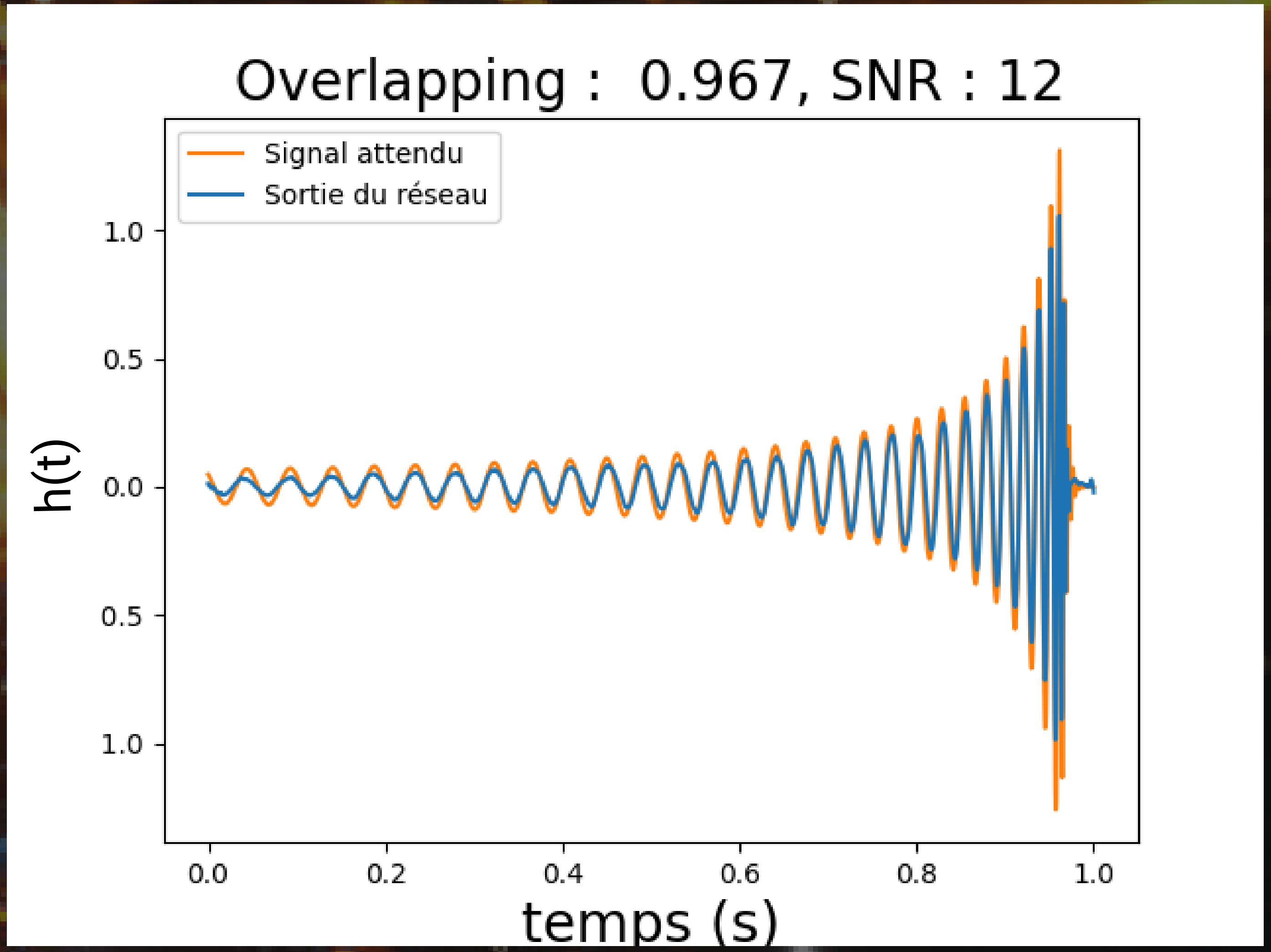
Principle



Entrée du réseau
h(t)
temps (s)



Feature Vector
Flatten
Tanh
Conv1D
MaxPool1D
Tanh
Conv1D
Reshape
Input
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
LSTM
Dense
Output

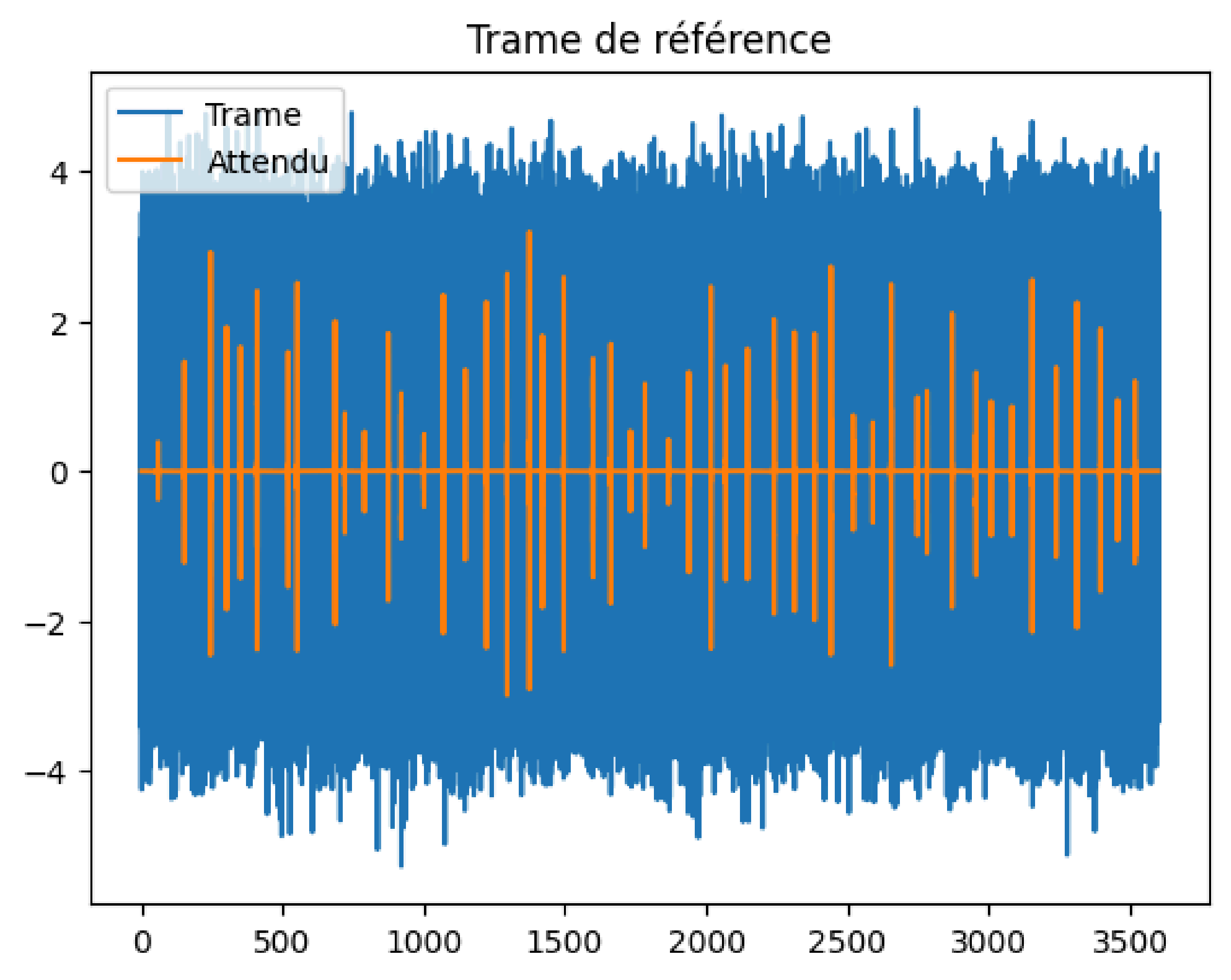


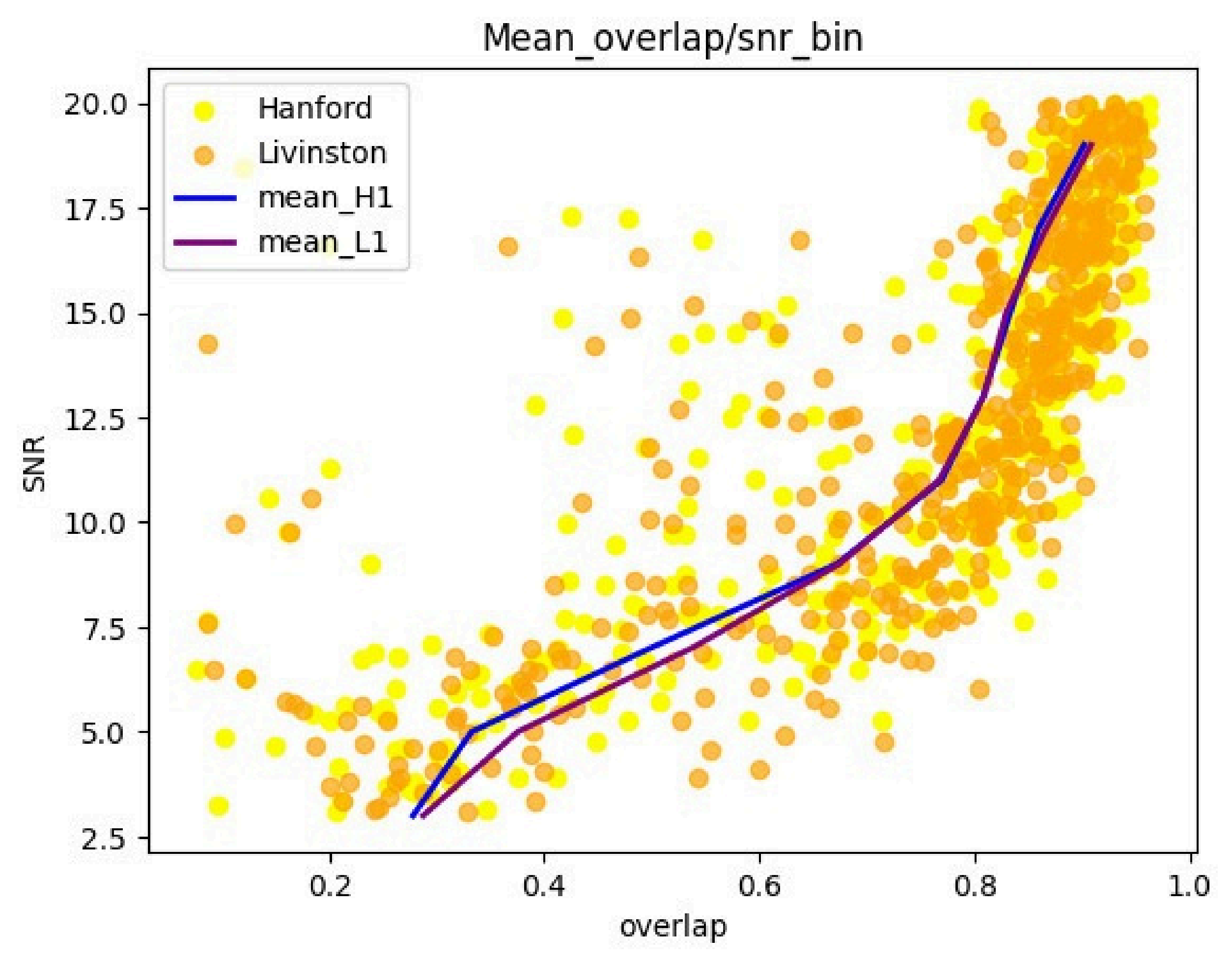
Overlapping : 0.967, SNR : 12
Signal attendu
Sortie du réseau
h(t)
temps (s)

Training efficiency

Testing on a long data frame

To test the effectiveness of the network training and show its strengths/weaknesses, we build a 10h reference frame with 500 signal injections
(1st hour display here)

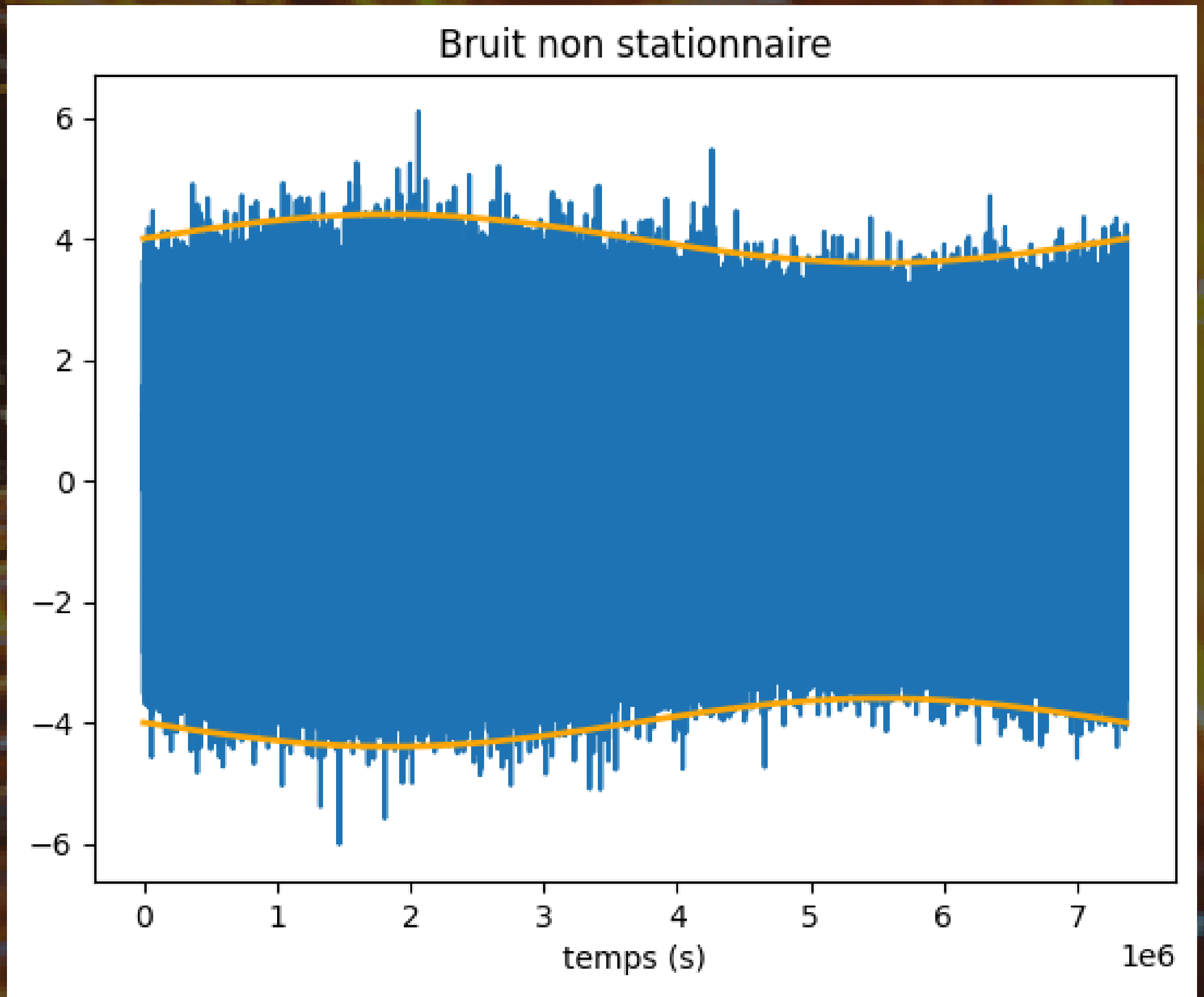


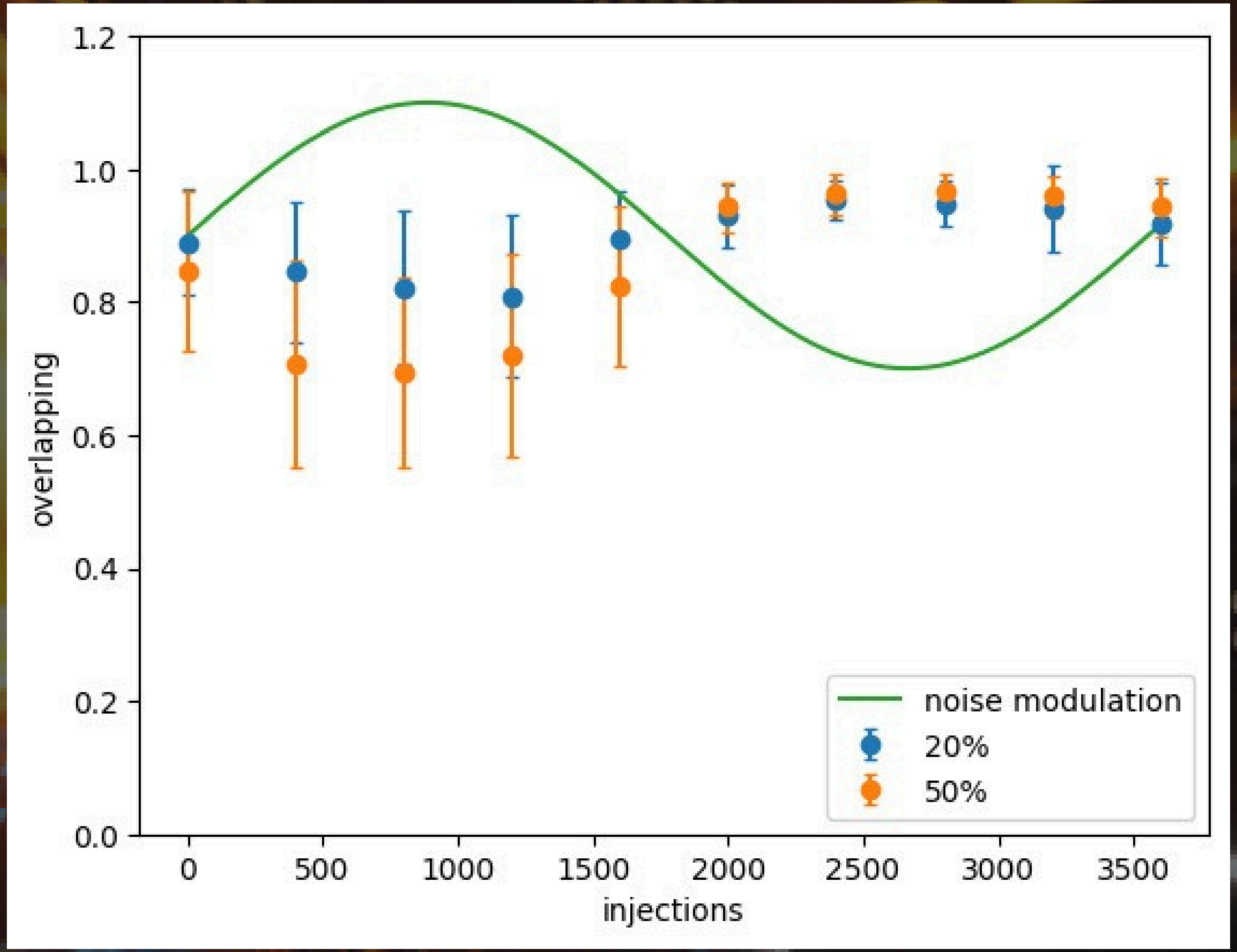


For each signal detected by the network, we plot its SNR as a function of its overlapping with the injection.

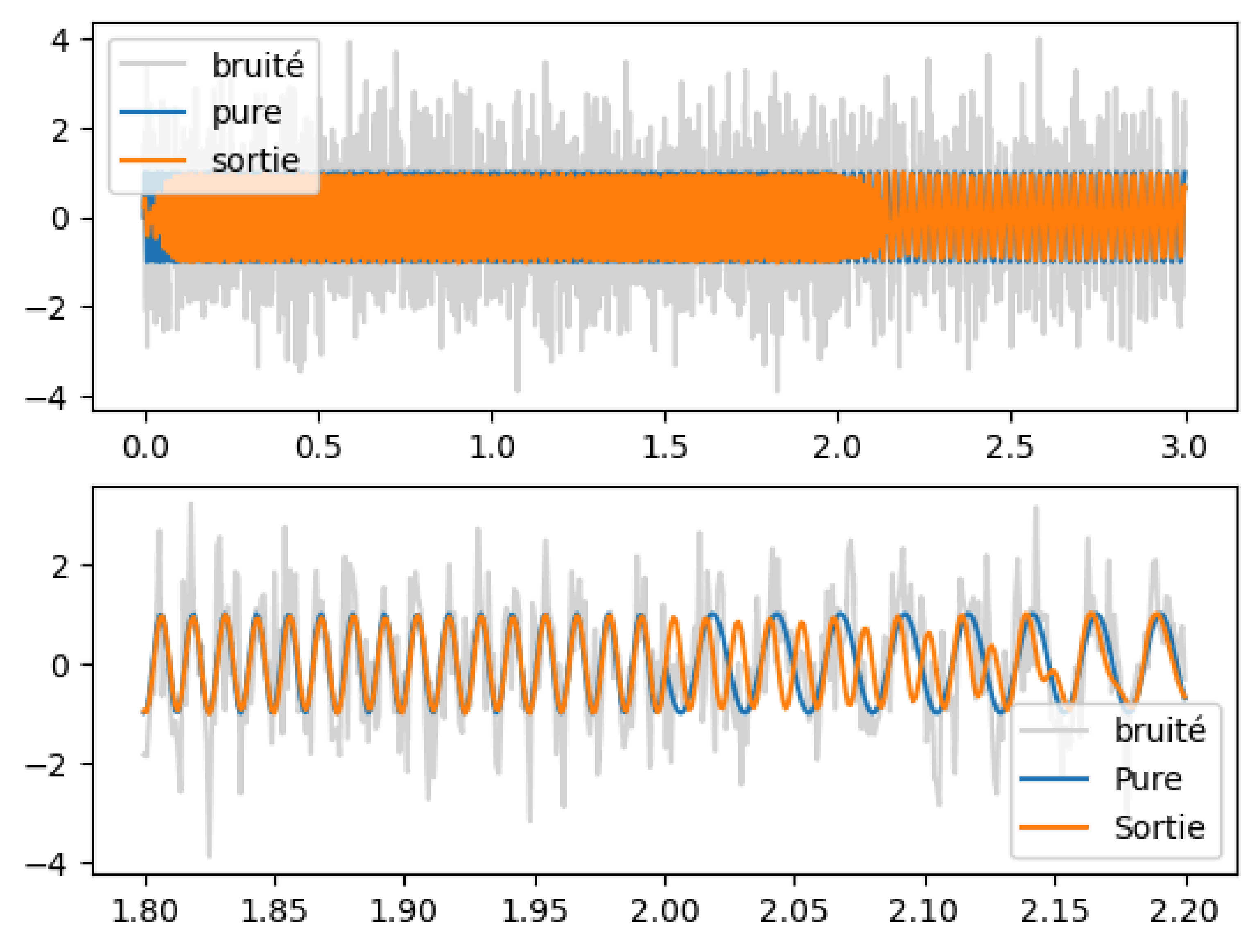
We look at the average overlapping for different SNR values to get an idea of the effectiveness of the training.

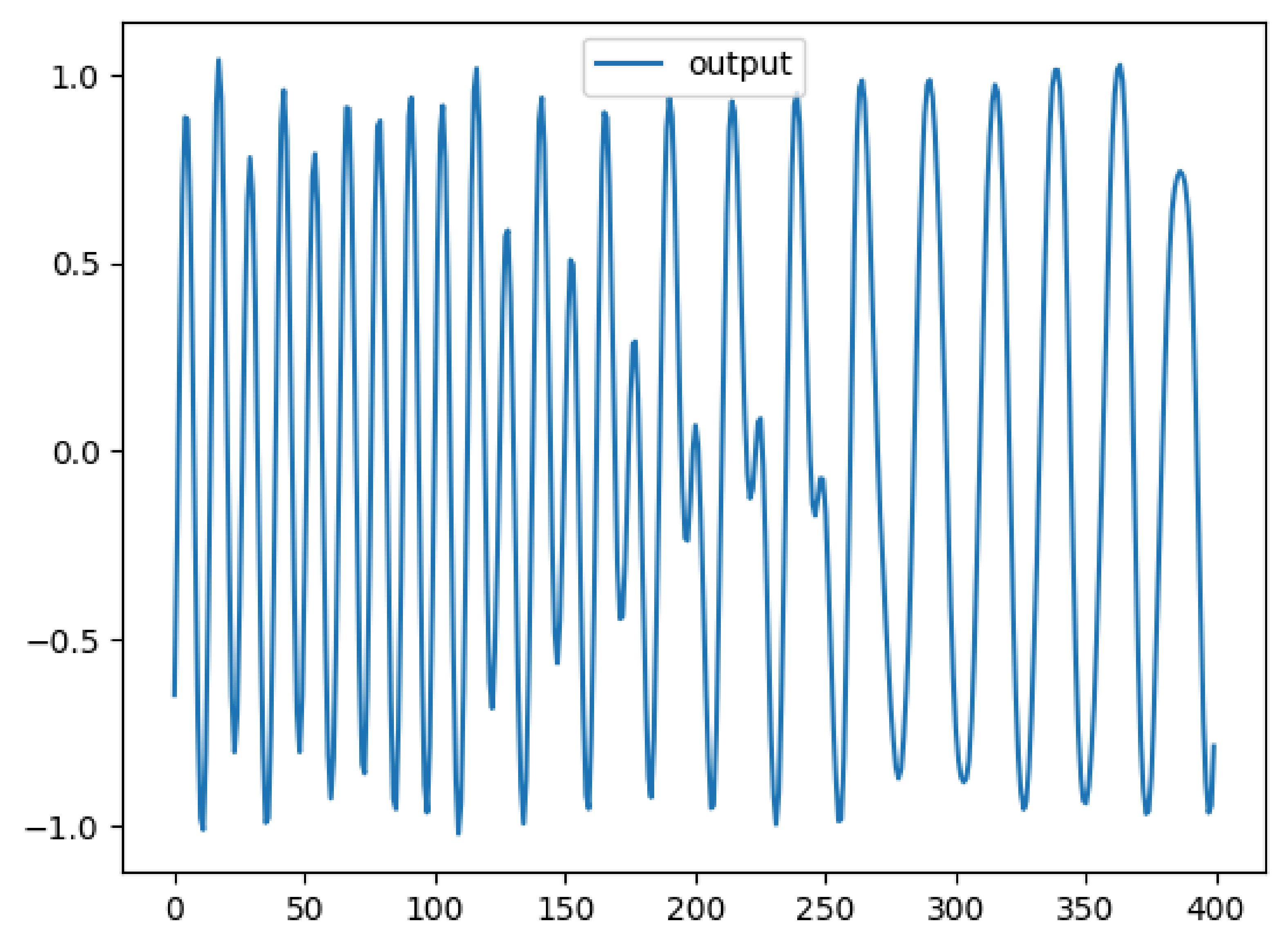
Frame modulation





Two-frequency problem





Tanimoto distance

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

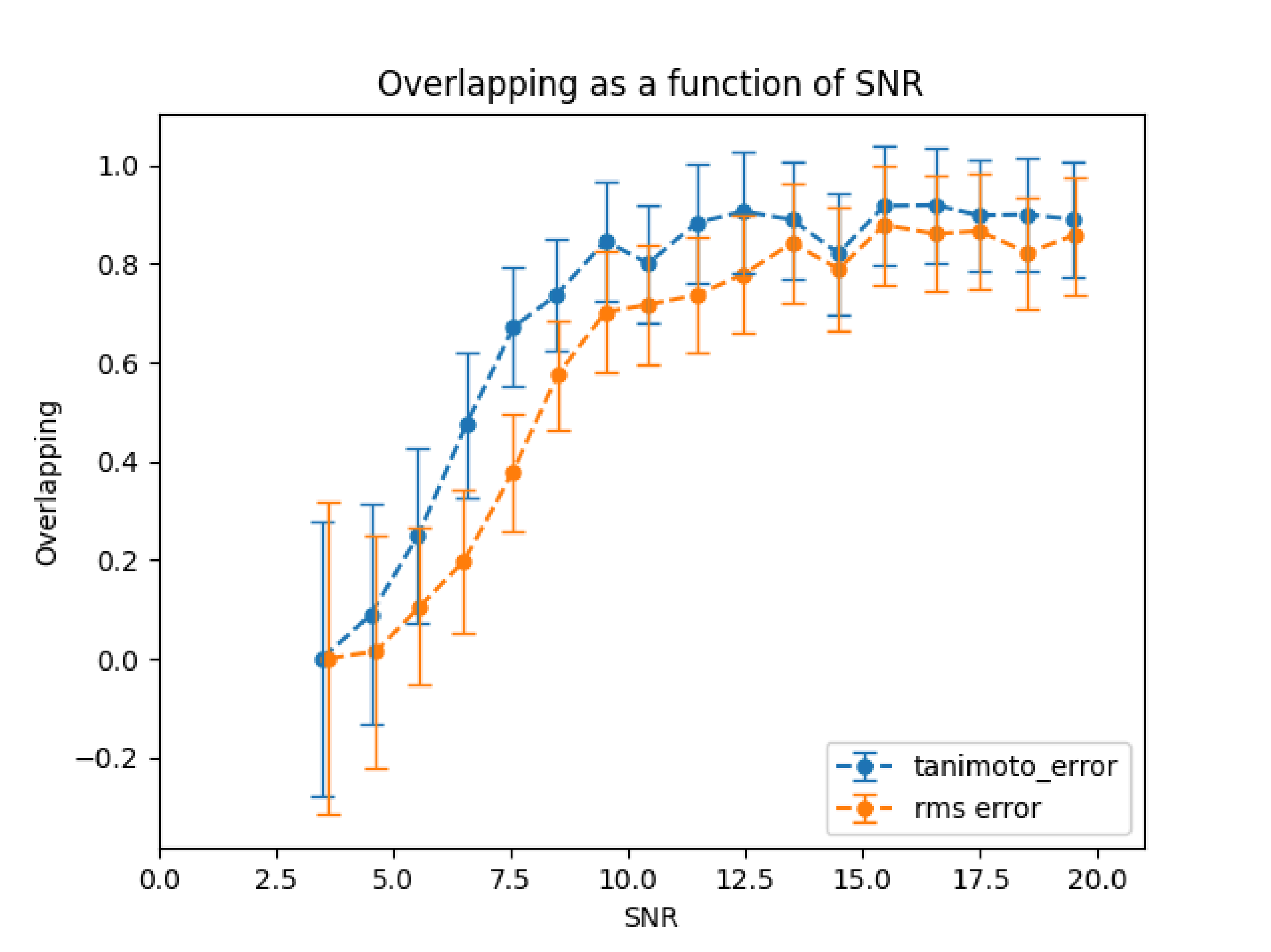
$\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are predicted values

$y_1, y_2, \ldots, y_n$ are observed values

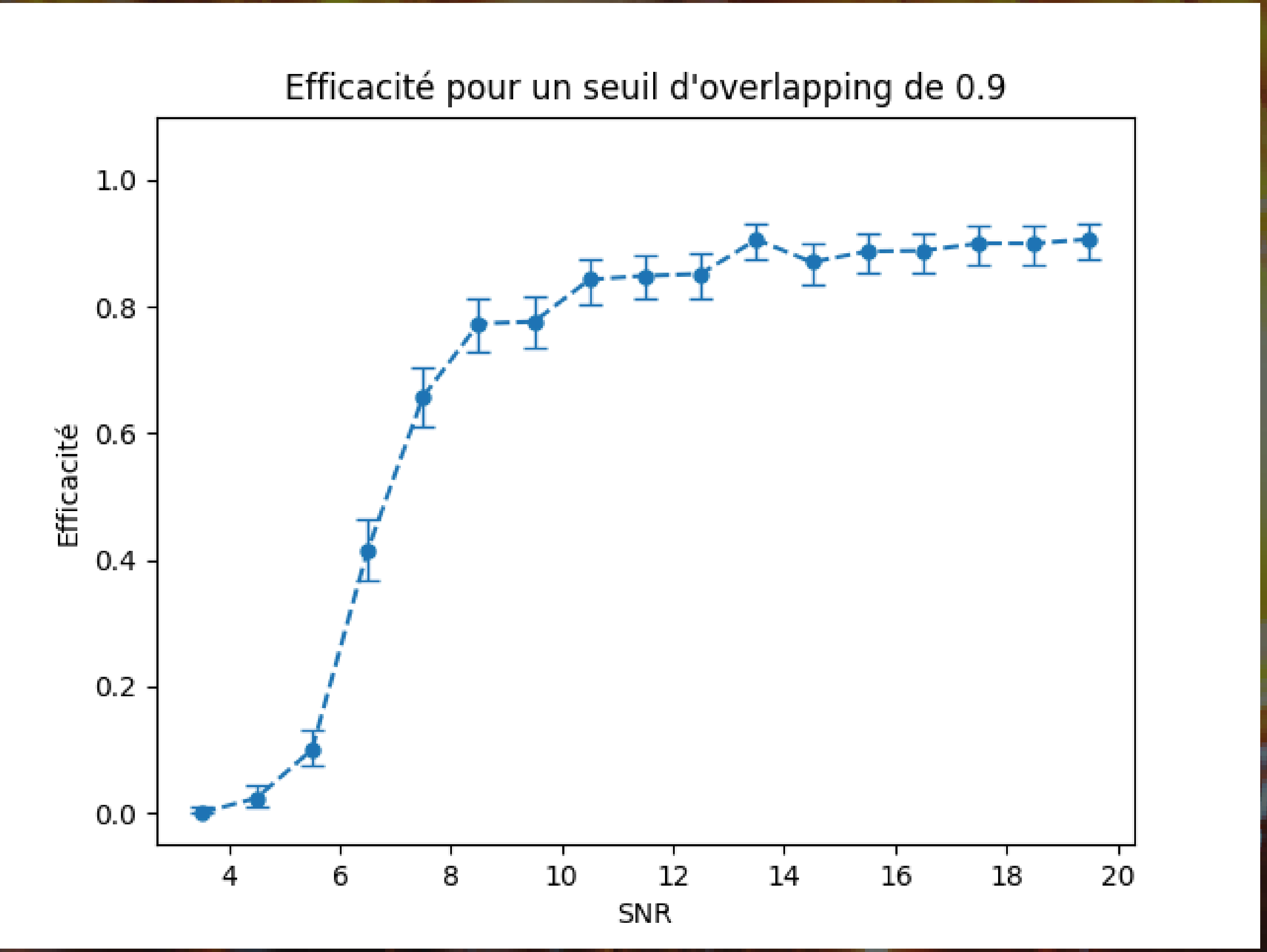
$$L_{z,x} = \frac{(\mathbf{x}_i - \mathbf{z}_i)^2}{n} - r^d_{w,z,x},$$

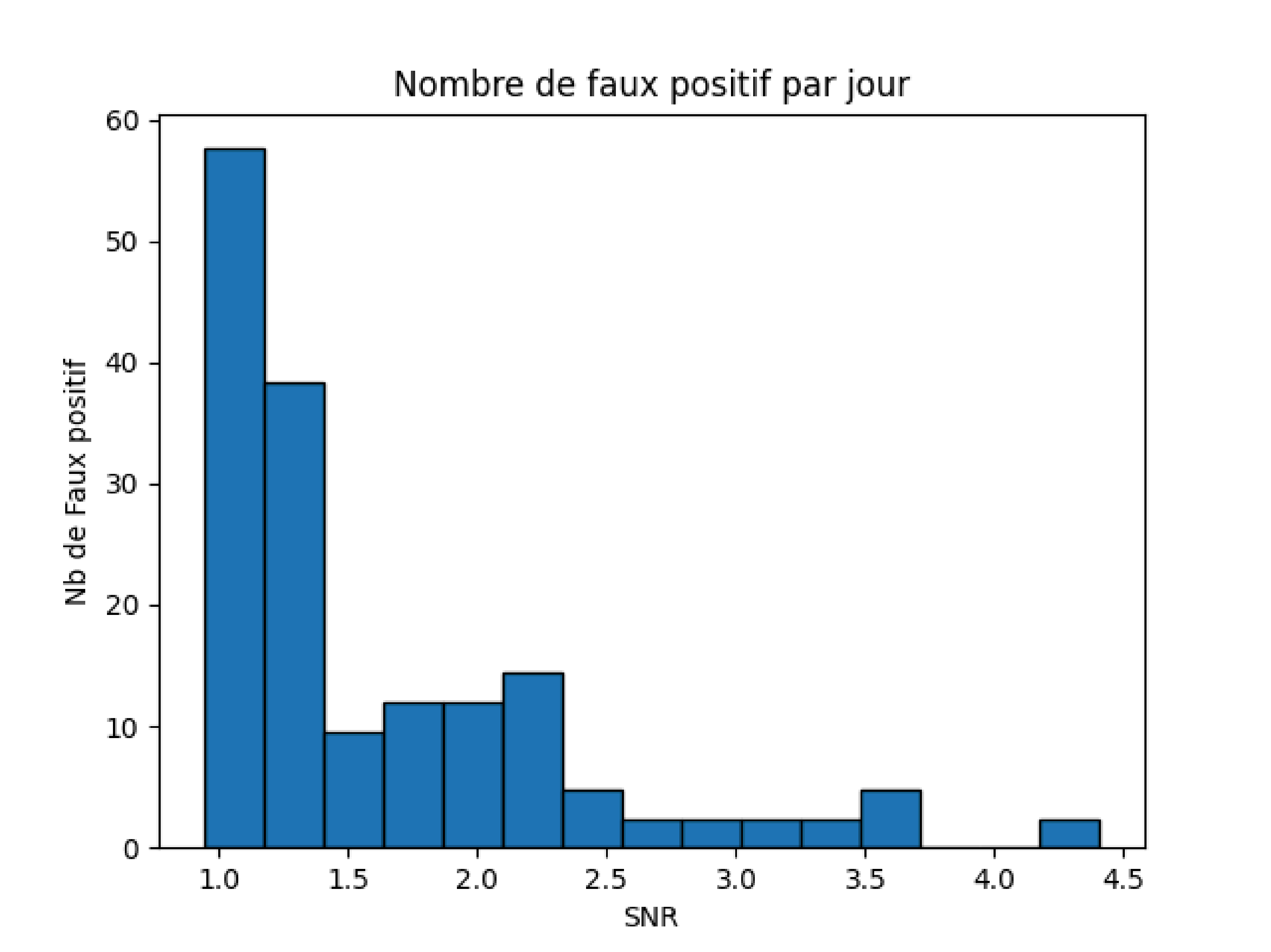
- **x :** Expected signal
- **z :** Network output
- **w :** Weight
- **d :** Hardness parameter

$$r^d_{w,z,x} = \frac{\sum_i^n w_i.\mathbf{z}_i.\mathbf{x}_i}{2^d \sum_i^n w_i.({\mathbf{z}_i}^2 + {\mathbf{x}_i}^2) - (2^{d+1} - 1)\sum_i^n w_i.\mathbf{z}_i.\mathbf{x}_i}$$

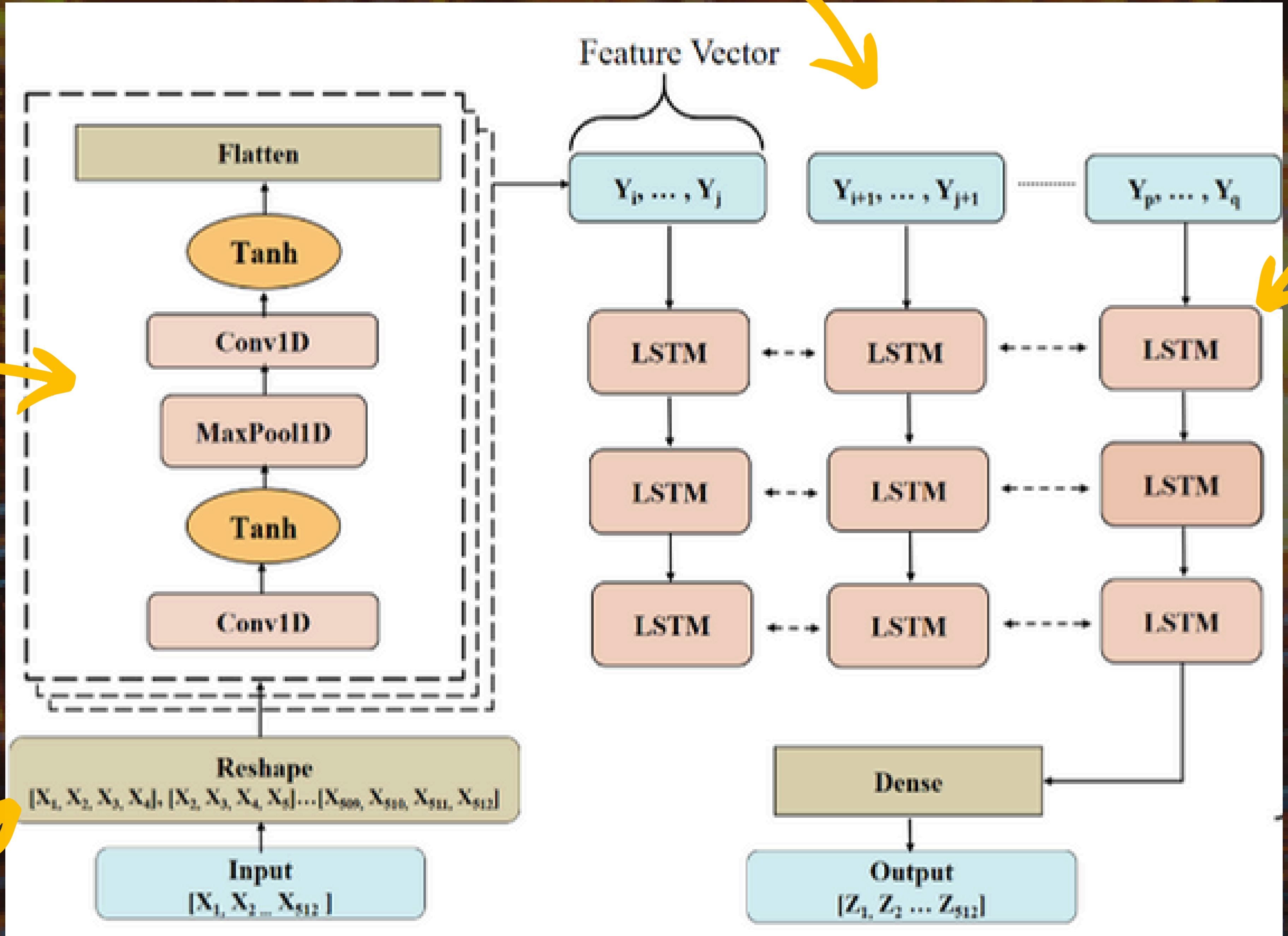


holesjoliplot





Hyperparameters



- Batch size
- Numbers of epochs
- Percentage of signal
- Learning rate
- Weights in tanimoto
- SNR during training

Thank You