FACULTY OF SCIENCES

A NOVEL APPROACH TO DISCOVER UNMODELLED FEATURES IN GRAVITATIONAL WAVE SIGNALS

Matteo Vereertbrugghen

GHENT UNIVERSITY

Promotor: Archisman Ghosh
Mentor: Freija Beirnaert

Recap

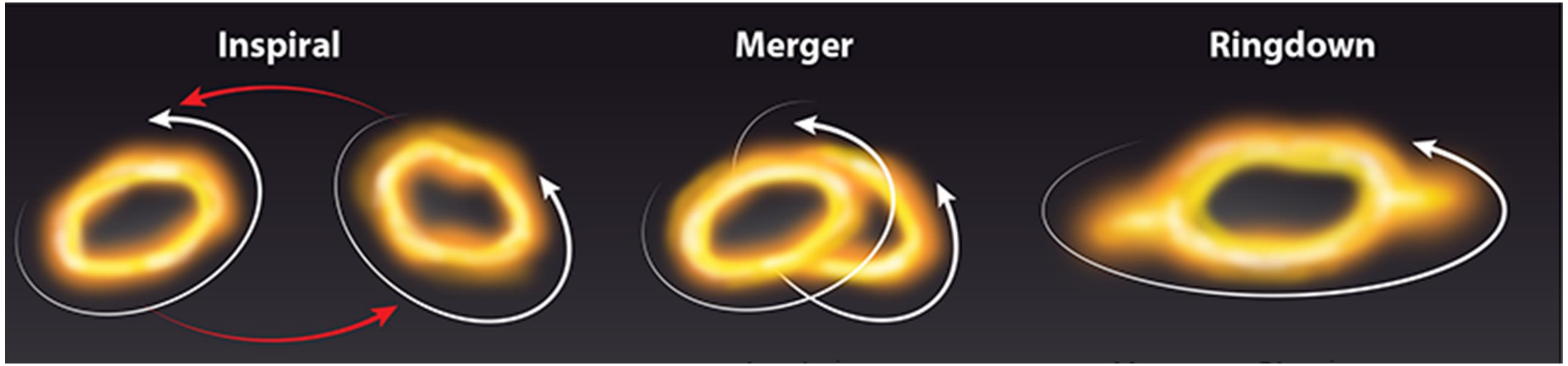
Machine learning algorithm

Results

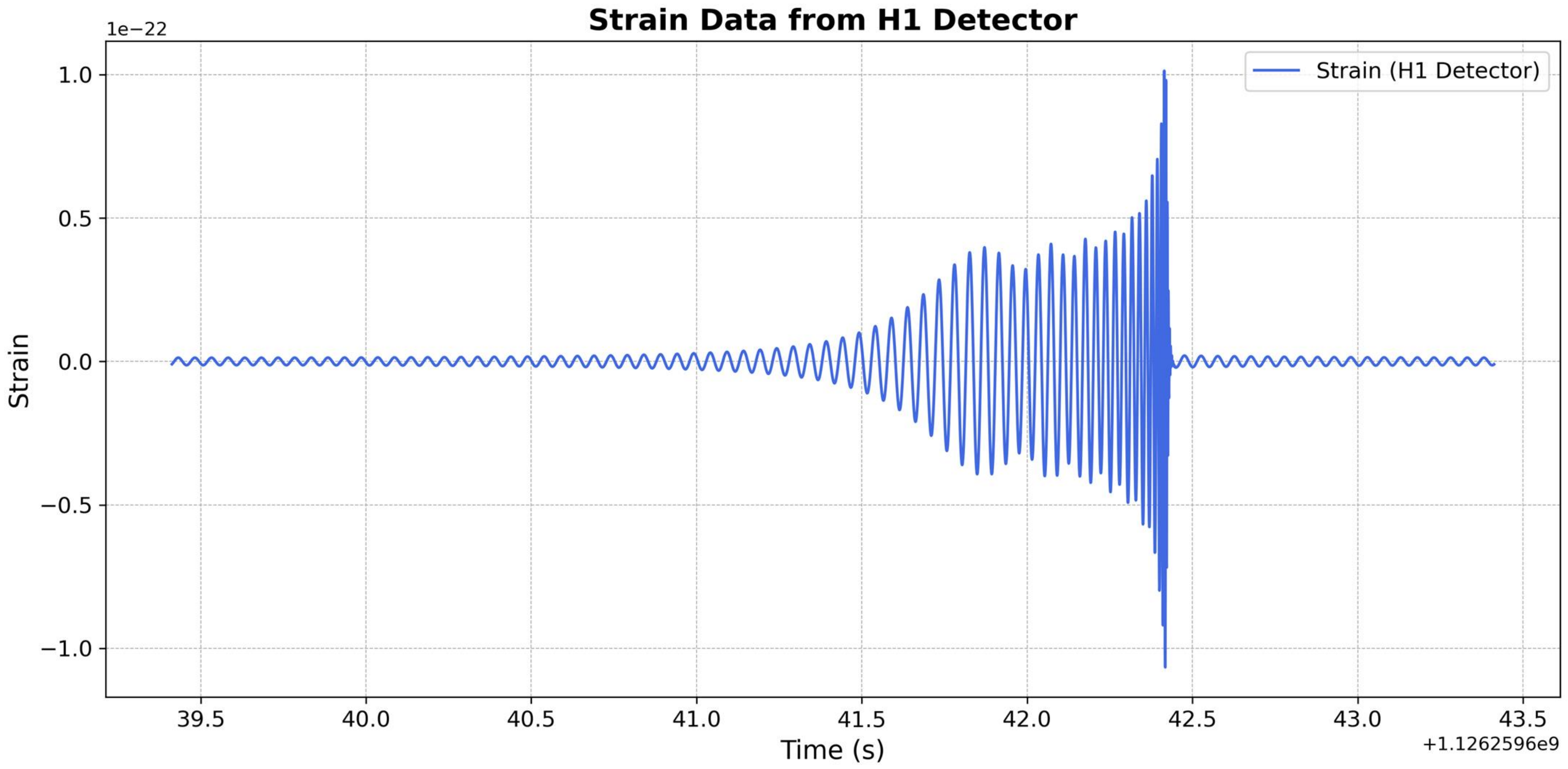
Future plans

GHENT UNIVERSITY

GRAVITATIONAL WAVES FROM CBC

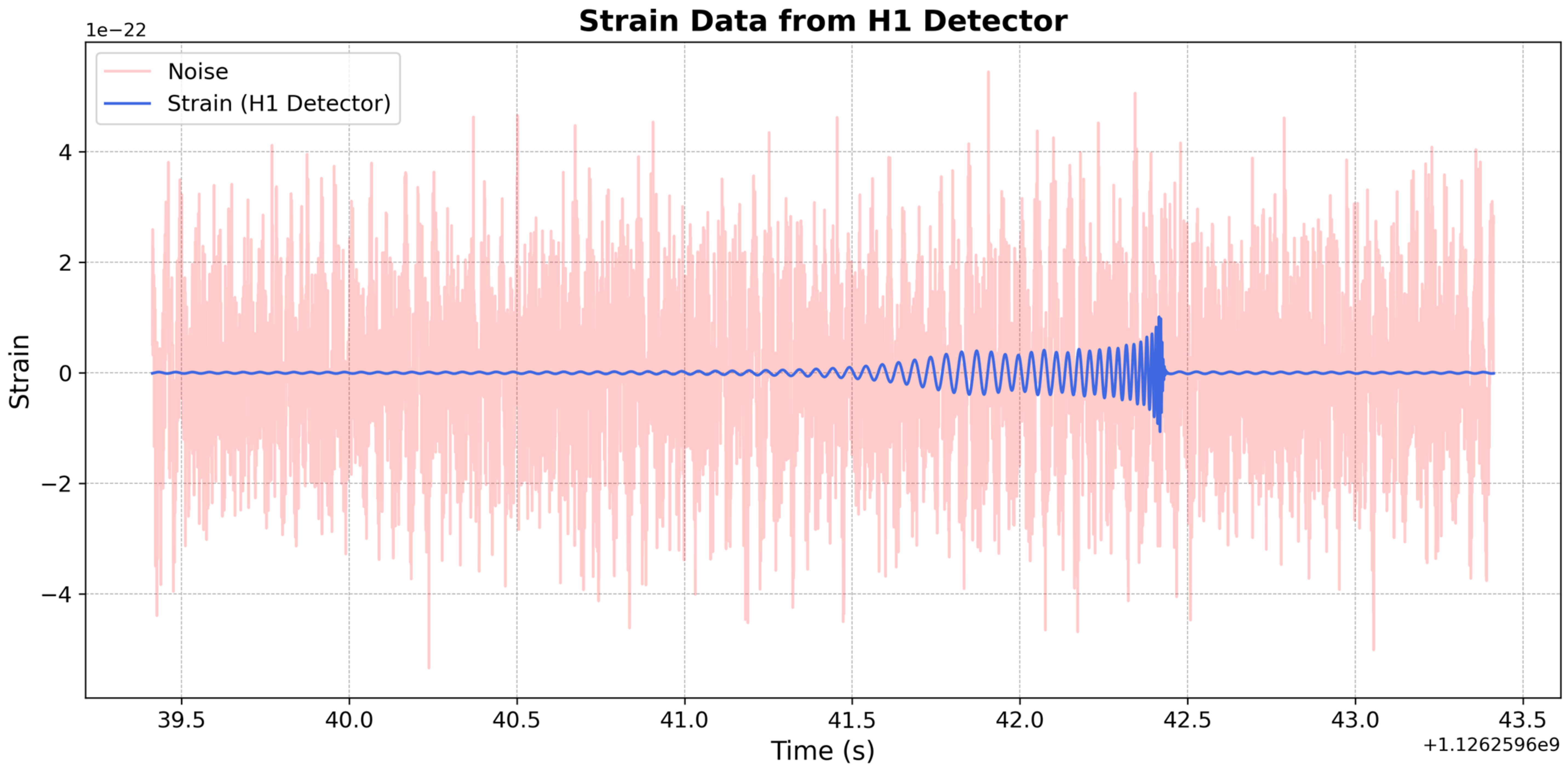


(Top) Kip Thorne; (Bottom) B. P. Abbott *et al.* [8]; adapted by APS/Carin Cain



GHENT UNIVERSITY

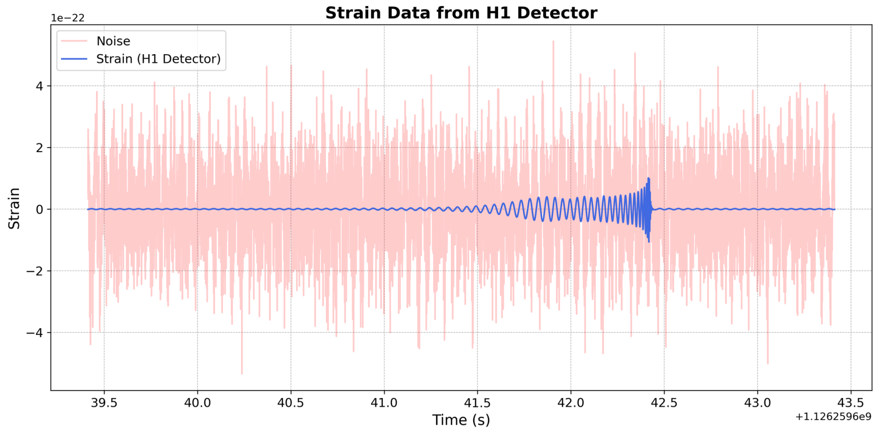
THE NOISE

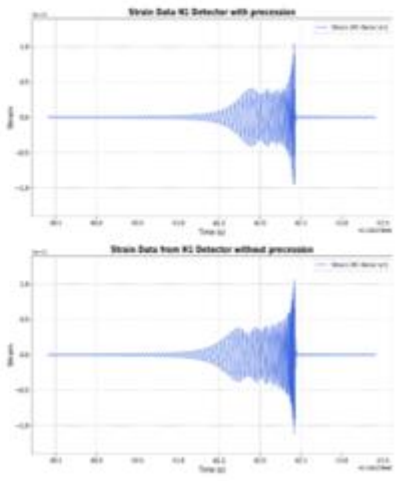


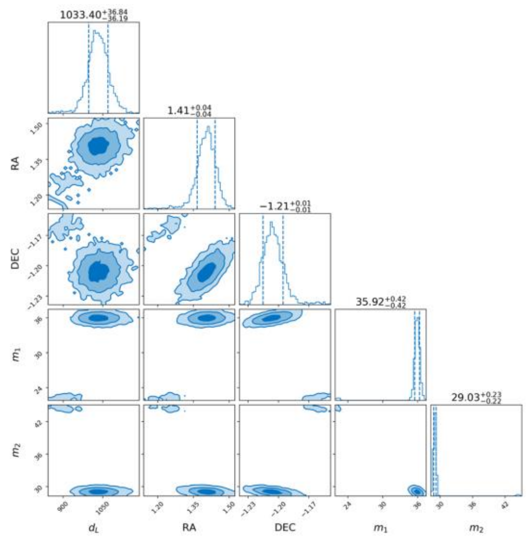
GHENT UNIVERSITY

EXTRACTING PARAMETERS

Injecting signal and noise into bilby → Choose model with parameters → Monte Carlo Markov chain → Returns Posterior distribution

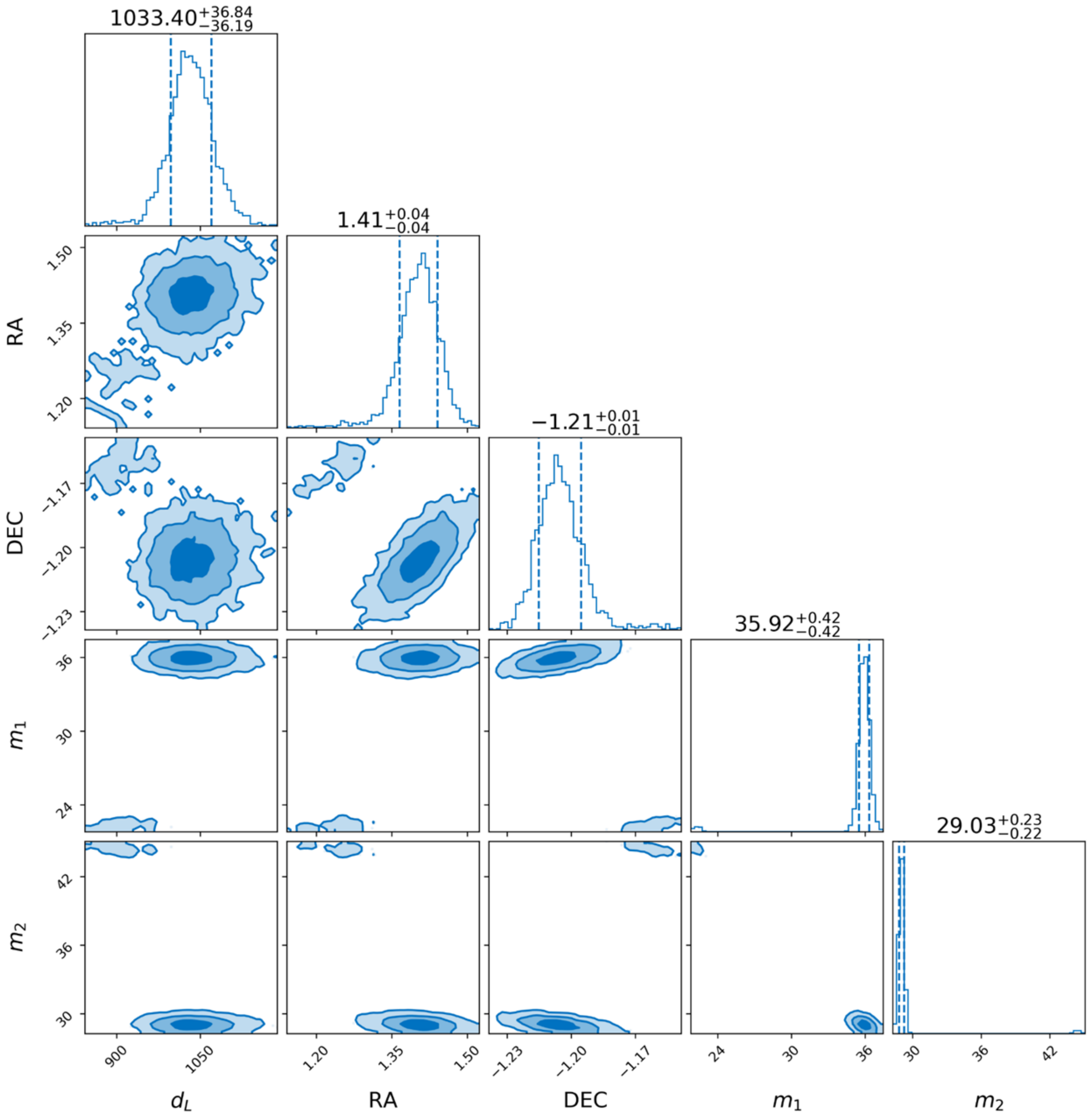






GHENT UNIVERSITY

CALCULATE PARAMETERS WITH BILBY

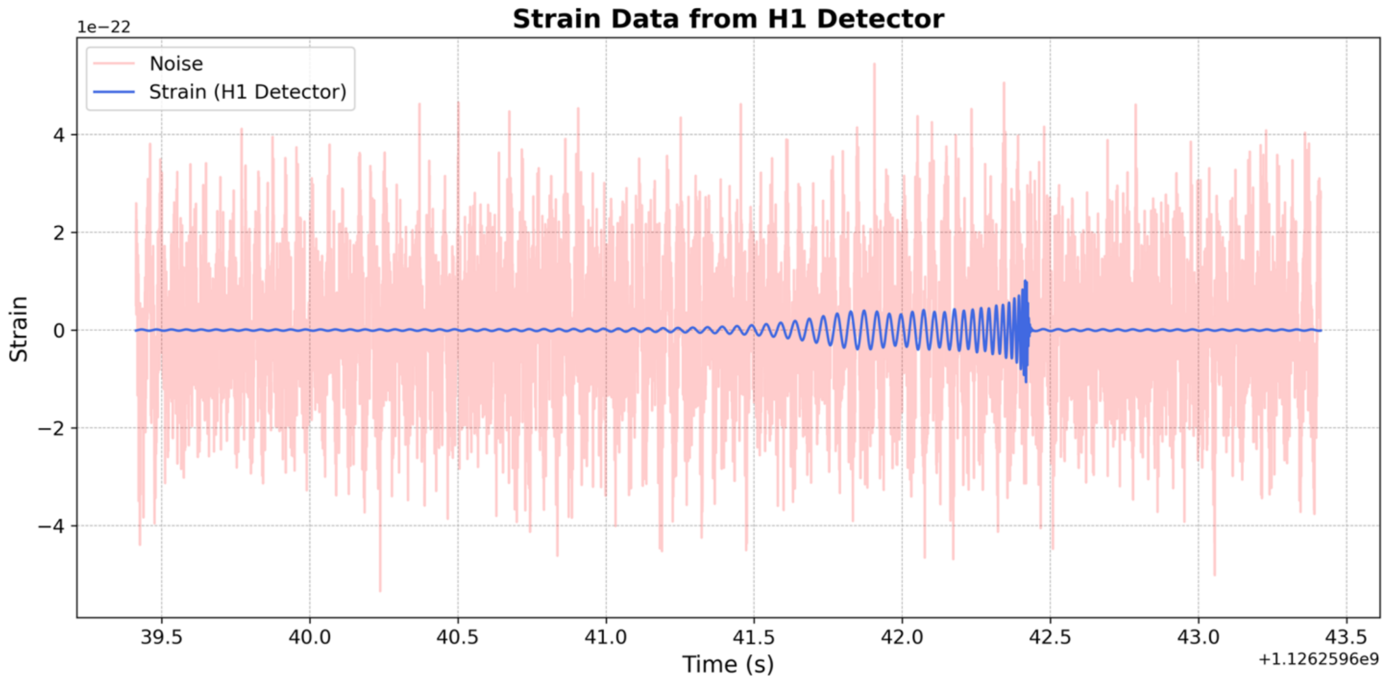


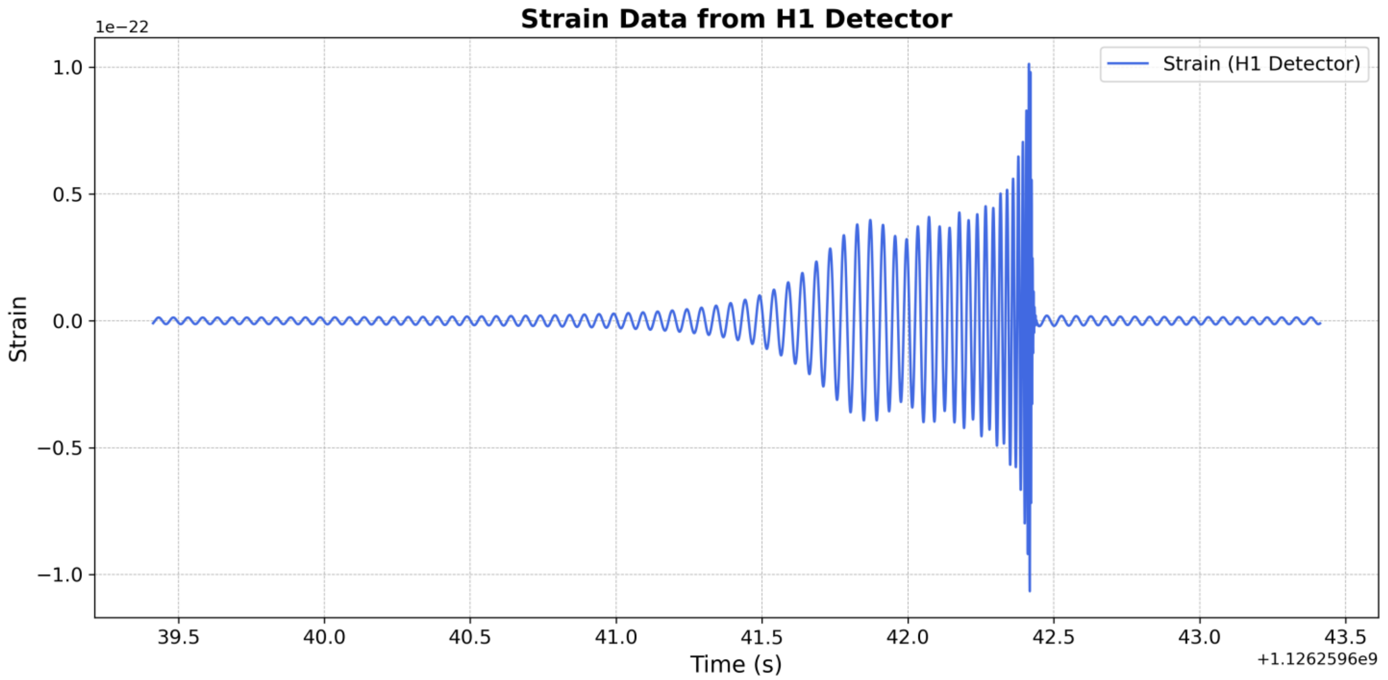
$$m_1 = \cdots$$
$$m_2 = \cdots$$
$$DEC = \cdots$$
$$RA = \cdots$$

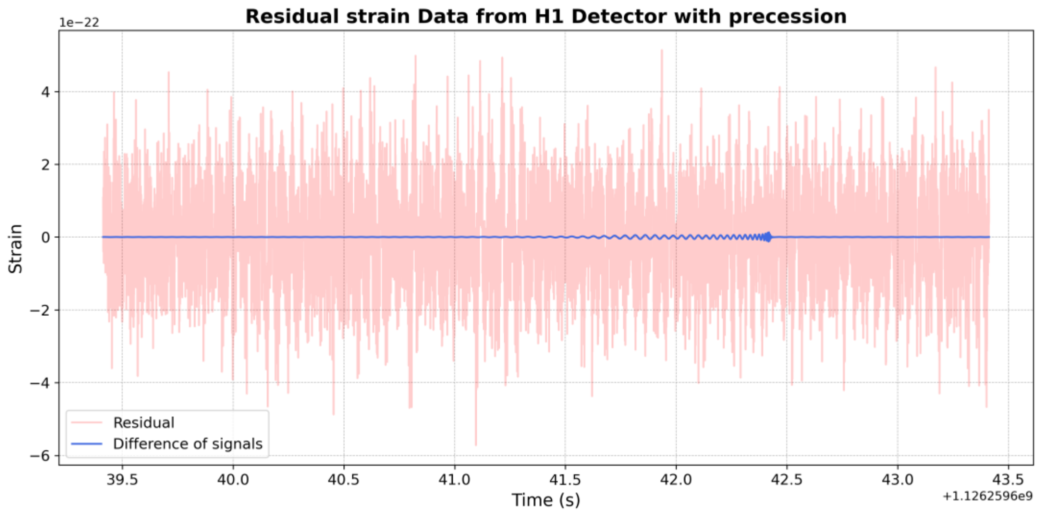
Very slow

GHENT UNIVERSITY

RESIDUAL ANALYSIS







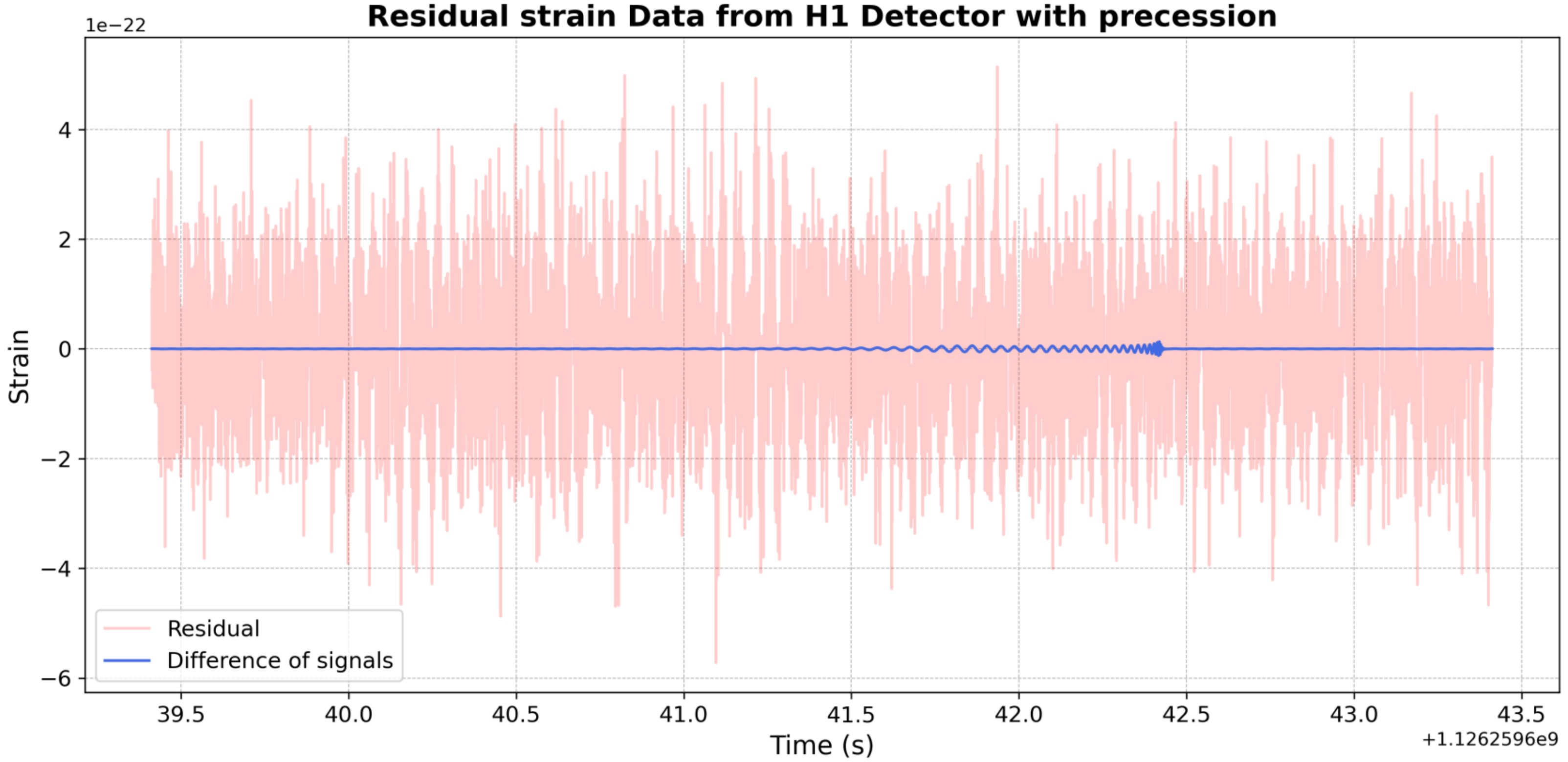
WHICH WAVEFORM TO USE?

Waveform	Precession	Higher harmonics
IMRPhenomD	✗	✗
IMRPhenomHM	✗	✓
IMRPhenomPv2	✓	✗
IMRPhenomXPHM	✓	✓

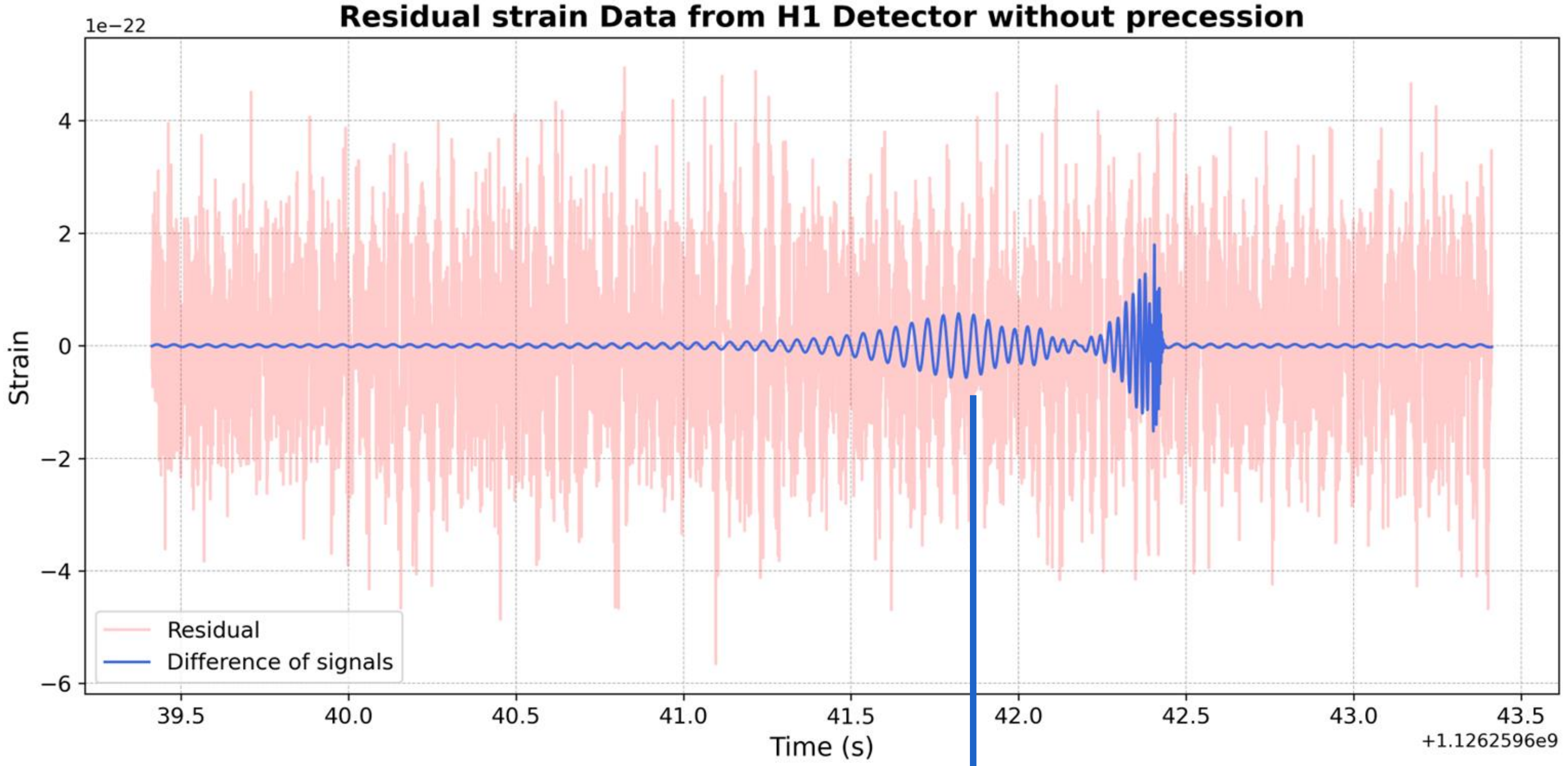
GHENT UNIVERSITY

DIFFERENT RESIDUALS

Correct Waveform



Wrong waveform



Real GW event: No blue part can be plotted

GHENT UNIVERSITY

GOAL OF THESIS

- Advancing detectors → More data
- Better waveform → Curse of dimensionality

  → long runtimes
- Less advanced waveforms → less parameters
- 2 goals:

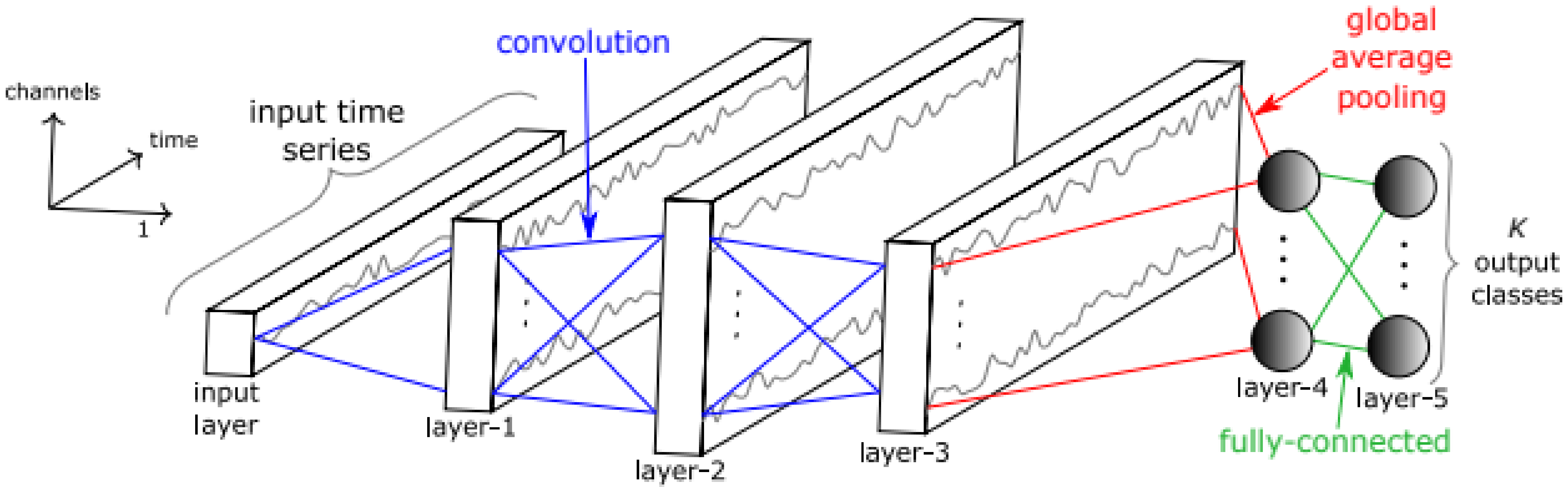
  Dirty run →Which features are we missing?

  Search for unmodeled features

CONVOLUTIONAL NEURAL NETWORK

GENERAL STRUCTURE

- Convolutional layers
- Pooling layers
- Dropout, fully connected layers

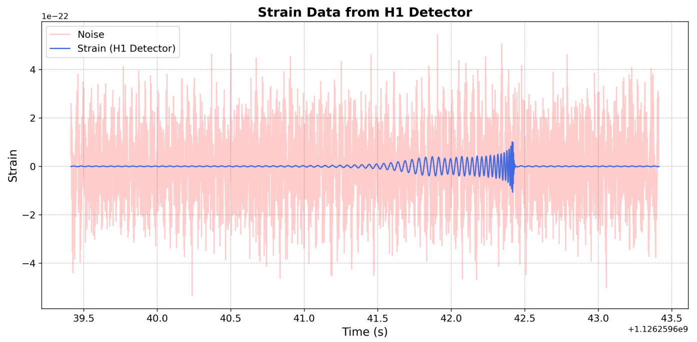


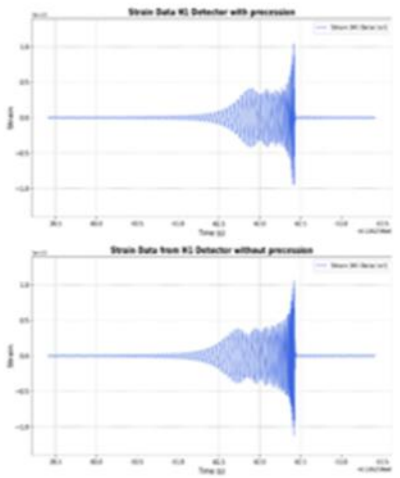
Taherkhani, A., Cosma, G. & McGinnity, T.M. A Deep Convolutional Neural Network for Time Series Classification with Intermediate Targets

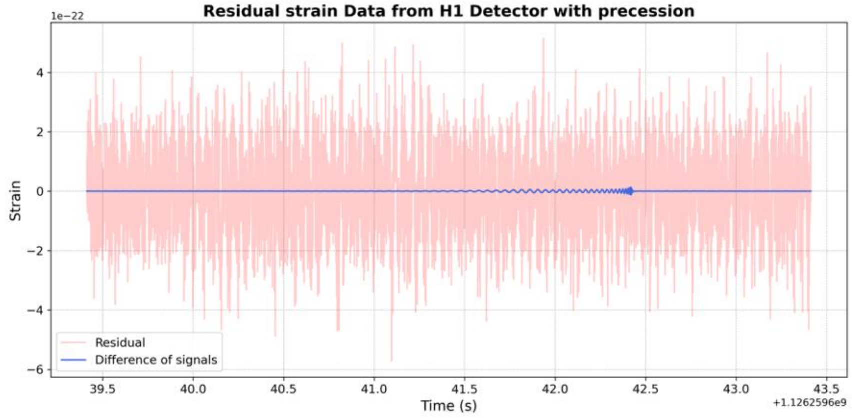
GHENT UNIVERSITY

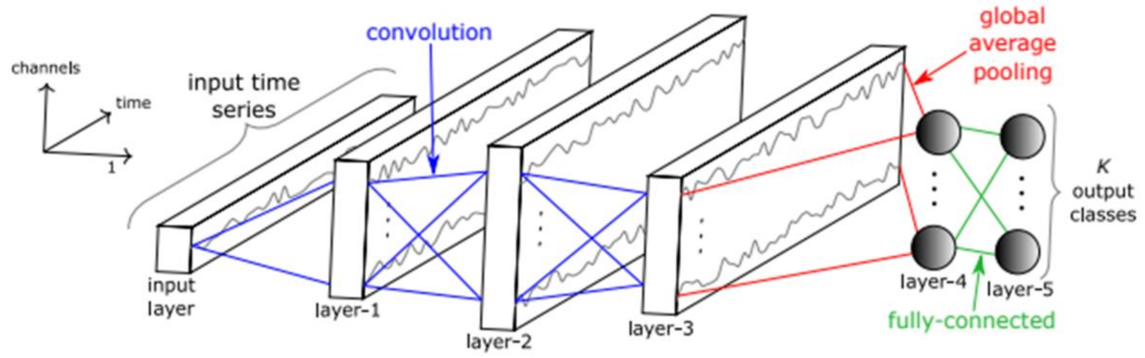
PIPELINE

Inject simulated signal → Choose waveform for reconstruction → Create residual → Train ML model → Test classifier with simulated signals not in the training set









RESULTS

GHENT UNIVERSITY

DATA GENERATION

100k residuals + Bilby = Computational expensive

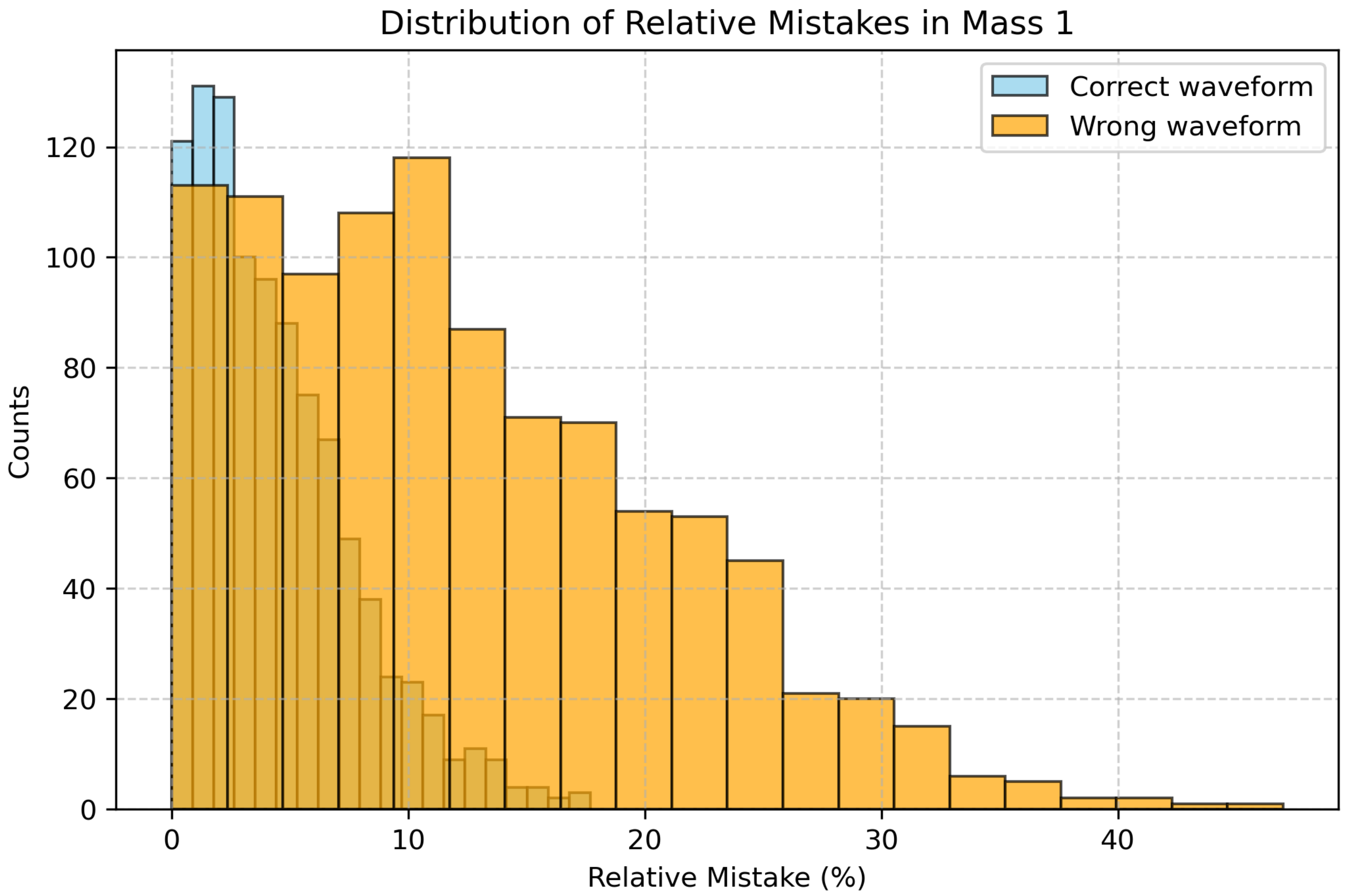
→Solution : SciPy optimize

- Injection signal → randomly generated parameters
- Injection waveform → IMRPhenomXPHM
- Recovered signal → SciPy parameters
- Recovered waveform → IMRPhenomXPHM or IMRPhenomPv2

DATA

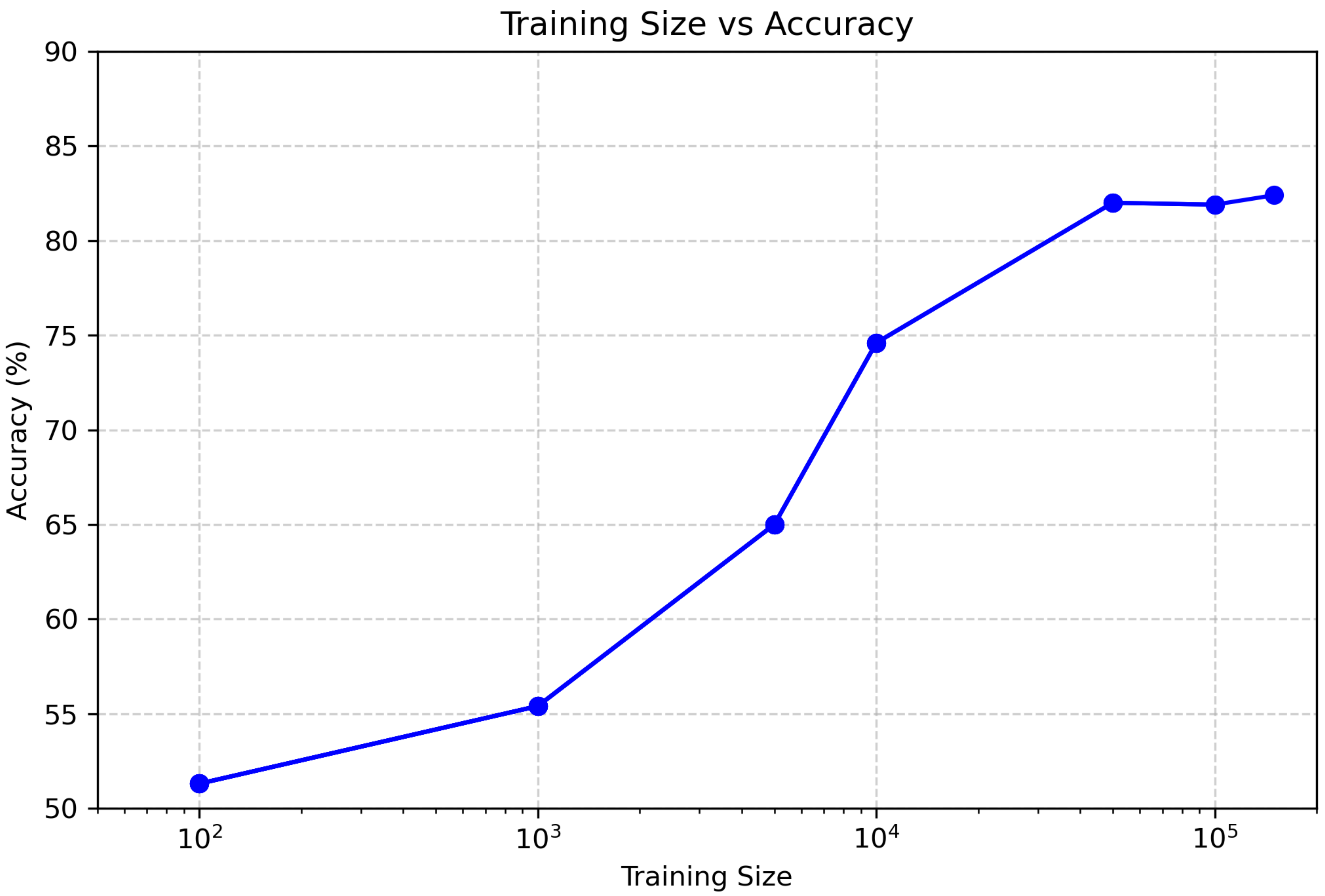
- Class labels 1 or 0:
  - 1: injection waveform = recovered waveform
  - 0: injection waveform ≠ recovered waveform
- If 0 → recovered parameters error larger
  - Check this: Relative mistake, 2000 runs

DIFFERENCE RELATIVE ERROR



GHENT UNIVERSITY

TEST ON TRAINING SIZE



GHENT UNIVERSITY

BEST RESULTS

100k training set

+

50k test set

+

IMRPhenomPv2
and
IMRPhenomXPHM

=

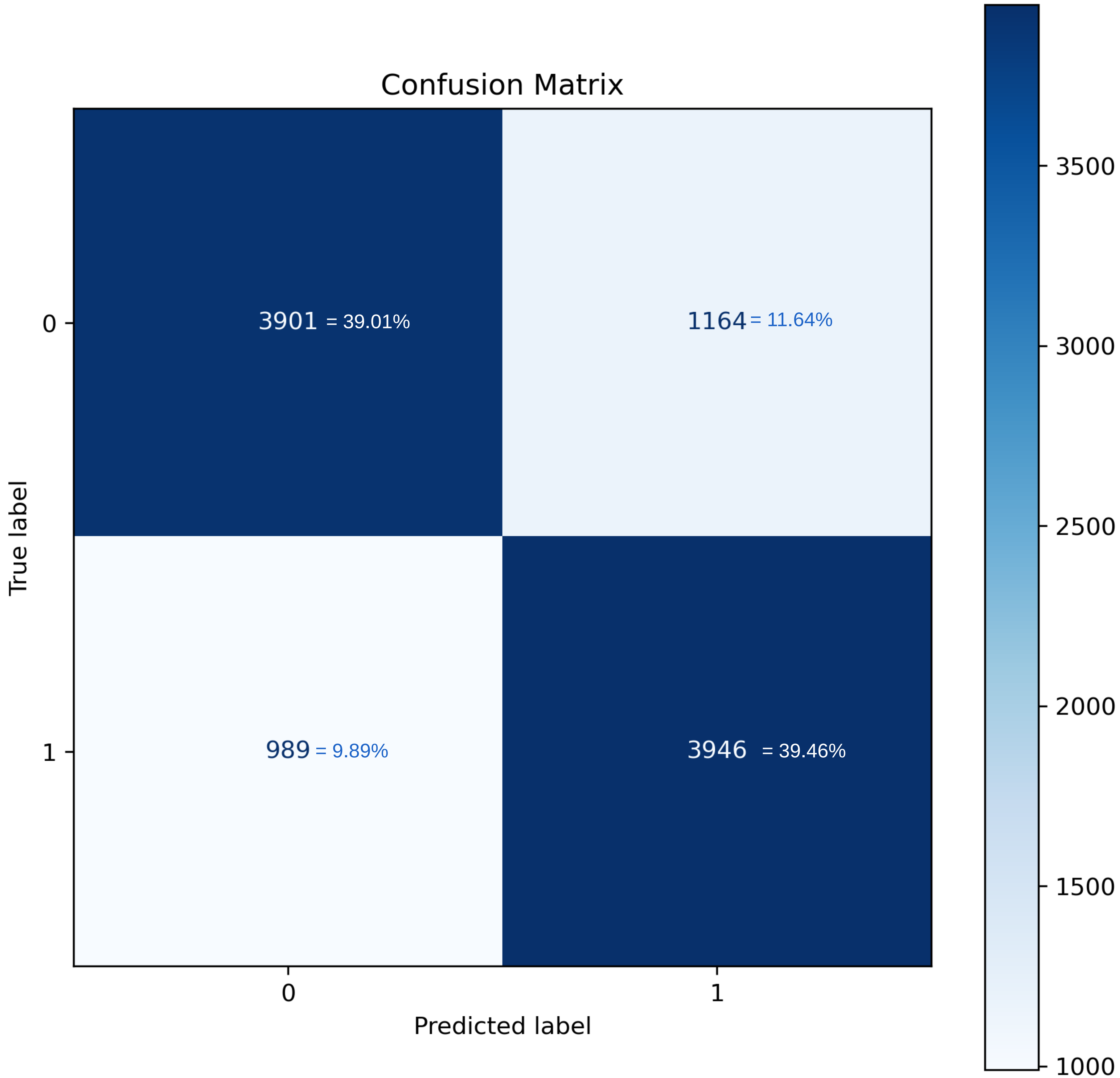
Result:
82.9%
accuracy

GHENT
UNIVERSITY

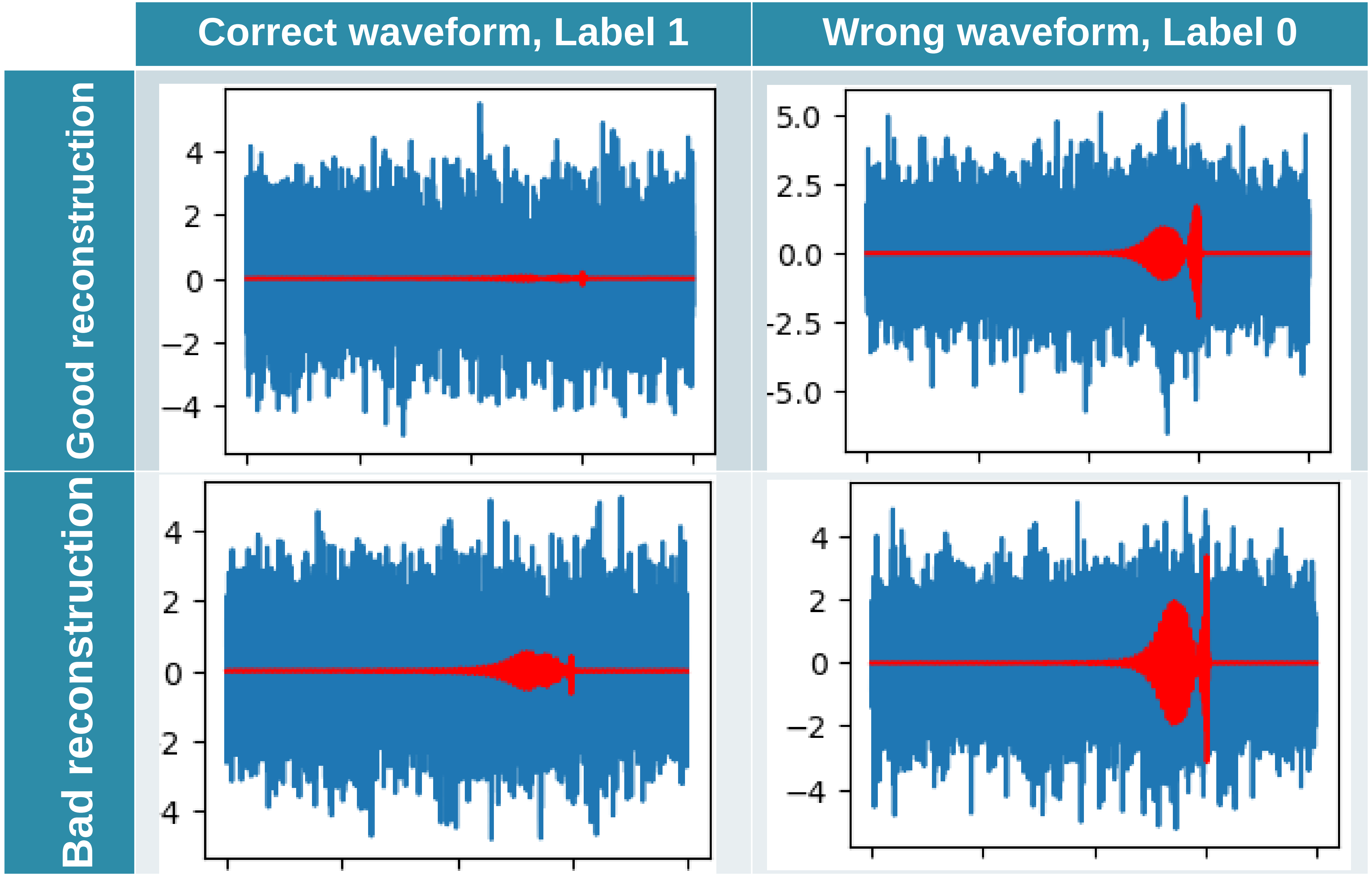
CONFUSION MATRIX

Correct waveform =1

Wrong waveform =0



POSSIBLE REASON FOR 80% ACCURACY



FUTURE PLANS

GHENT UNIVERSITY

OPTIMIZATION OF ALGORITHM

ADDING MORE WAVEFORMS

TESTING ON REAL GRAVITATIONAL WAVES

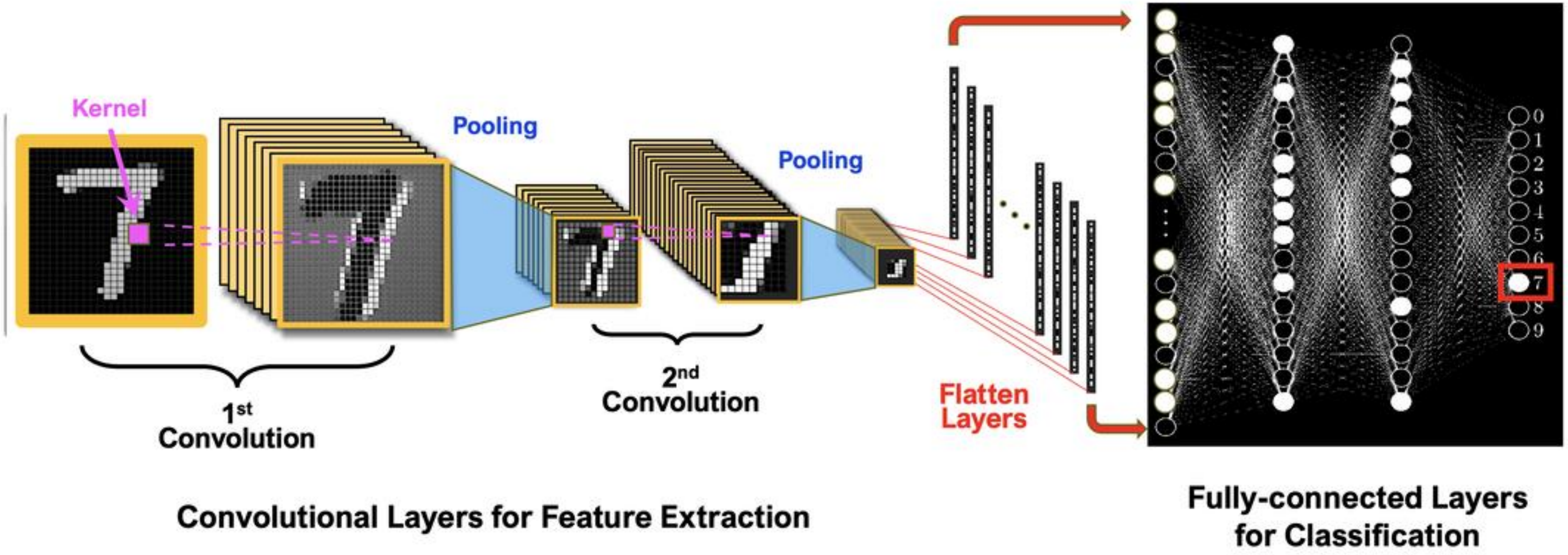
SEARCH UNMODELLED FEATURES

QUESTIONS?

GHENT UNIVERSITY

EXTRA CNN STRUCTURE

```
model = nn.Sequential(
    nn.Conv1d(1, 16, kernel_size=5, padding=2),
    nn.ReLU(),
    nn.MaxPool1d(2),
    nn.Dropout(0.4),  # Dropout added

    nn.Conv1d(16, 32, kernel_size=5, padding=2),
    nn.ReLU(),
    nn.MaxPool1d(2),
    nn.Dropout(0.4),

    nn.Conv1d(32, 128, kernel_size=5, padding=2),
    nn.ReLU(),
    nn.MaxPool1d(2),
    nn.Dropout(0.4),

    nn.Flatten(),
    nn.Linear(128 * (num_timesteps // 8), 512),  # Increase neurons
    nn.ReLU(),
    nn.Dropout(0.5),

    nn.Linear(512, 256),  # Add extra hidden layer
    nn.ReLU(),
    nn.Dropout(0.5),

    nn.Linear(256, num_classes)  # Output layer
)
```


Kernel
Pooling
Pooling
1st
Convolution
2nd
Convolution
Flatten
Layers
0
1
2
3
4
5
6
7
8
9
Convolutional Layers for Feature Extraction
Fully-connected Layers
for Classification