Detection of anomalies amongst LIGO's glitch populations with autoencoders

Melissa Lopez

m.lopez@uu.nl

ArXiv: 2310.03453

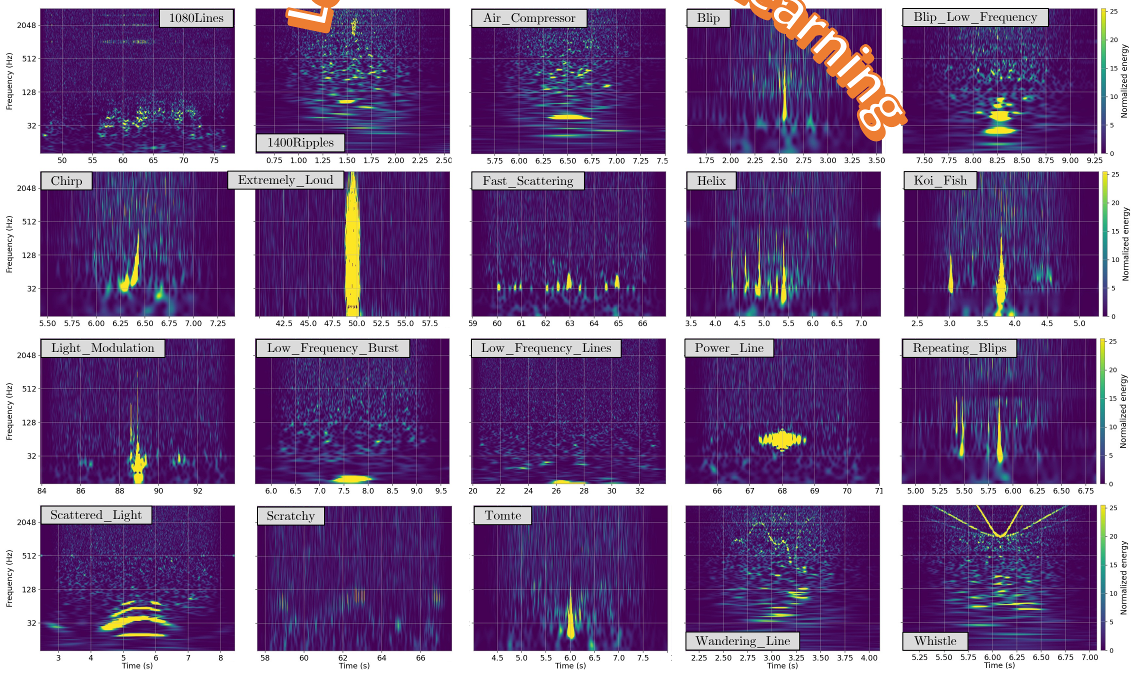
Transient noise a.k.a. glitches in LIGO

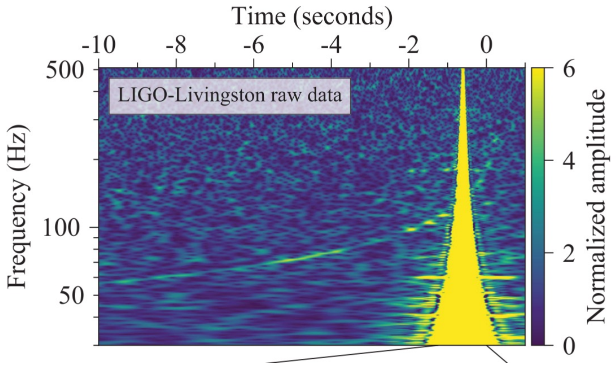
- Caused by instruments or environment (known or unknown)
- Diminish scientific data available
- Hinder GW detection (mask and/or mimic)
- Present in LIGO, Virgo and probably Einstein Telescope!

Idea: we need to mitigate them, so let's indetify them first

But… too many glitches! ~ 1 min^{-1} during O2

Let's use Machine Learning

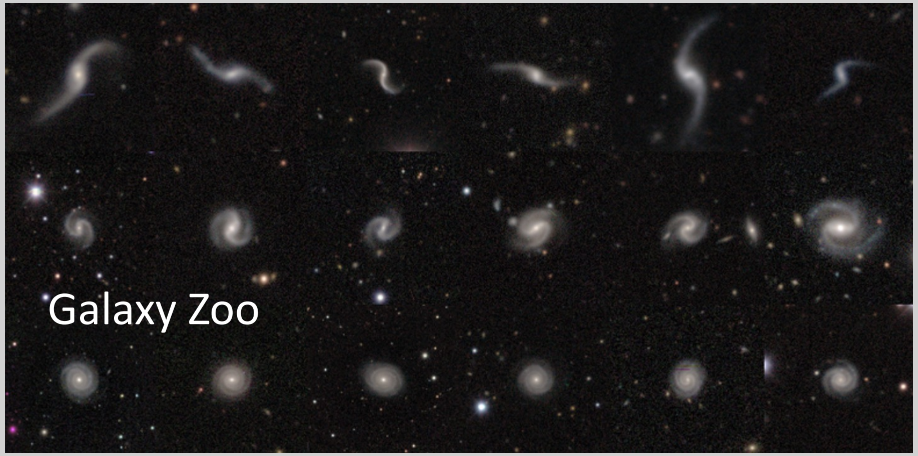


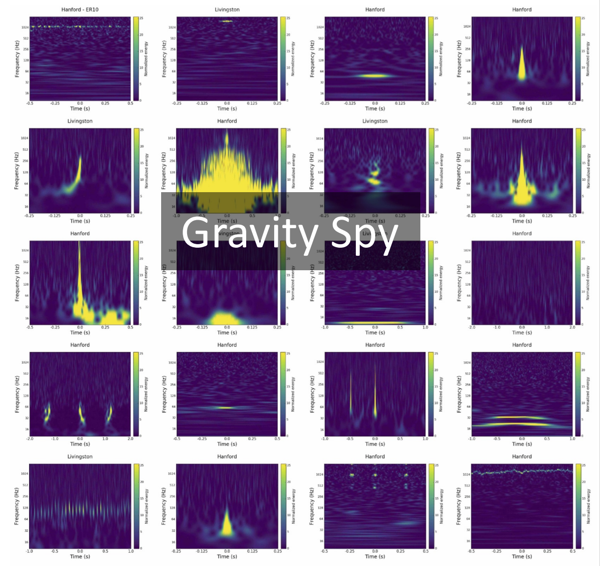


GW masked by glitch (GW170817)

Machine learning for glitch identification

Supervised learning: classification





Challenges

- Representation in the main strain of the detector
- Classes are rigid and labels expensive
- The detector evolves over time

Idea 1: we can use information from the detector itself, ie. *auxiliary channels*? → ~10^6 channels to process!

Idea 2: Let's the data speak for itself → unsupervised learning

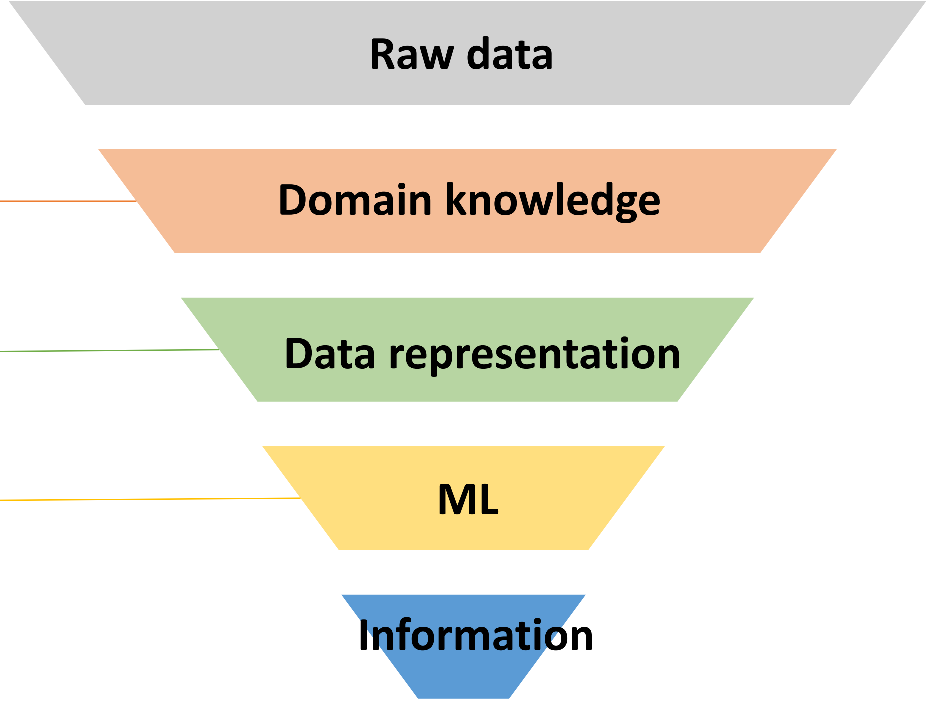
Encode, encode, encode

How can we reduce a 10^6 auxiliary channels (ac)?

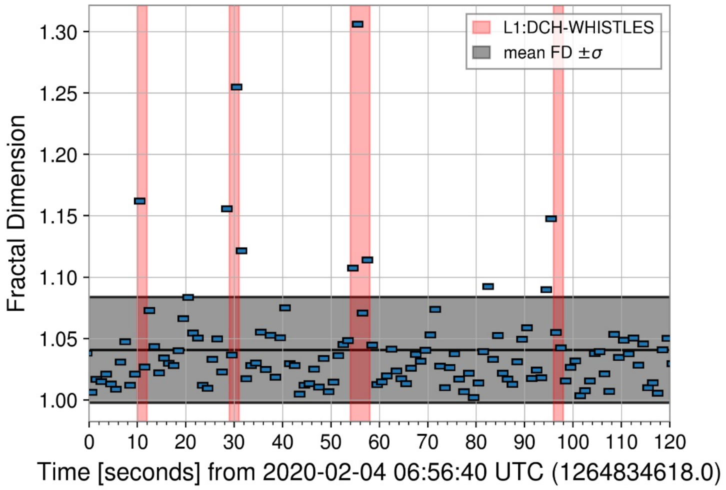
Select safe channels, i.e.not affected by GW (350 ac)

Encode with fractal dimension, i.e. measure complexity of the data

Use convolutional autoencoders



- M. Cavaglia 2022 → 1h of data encoded in 1h
- Our work → 1h of data encoded in 11s

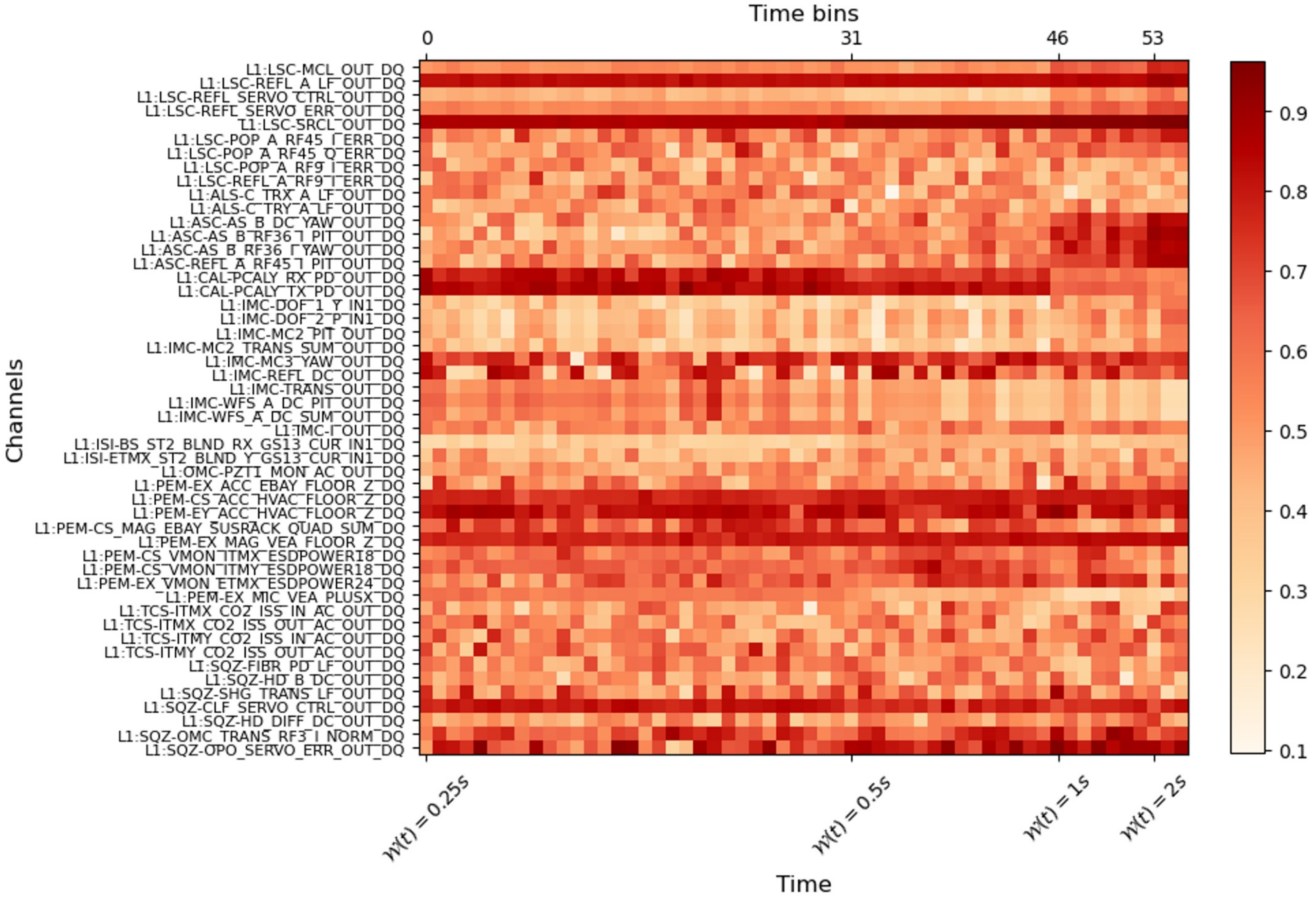


Robin van der Laag (UU)
Expert in high performance computing

Paloma Laguarta (UM)
Expert in ML
PhD at LHCb

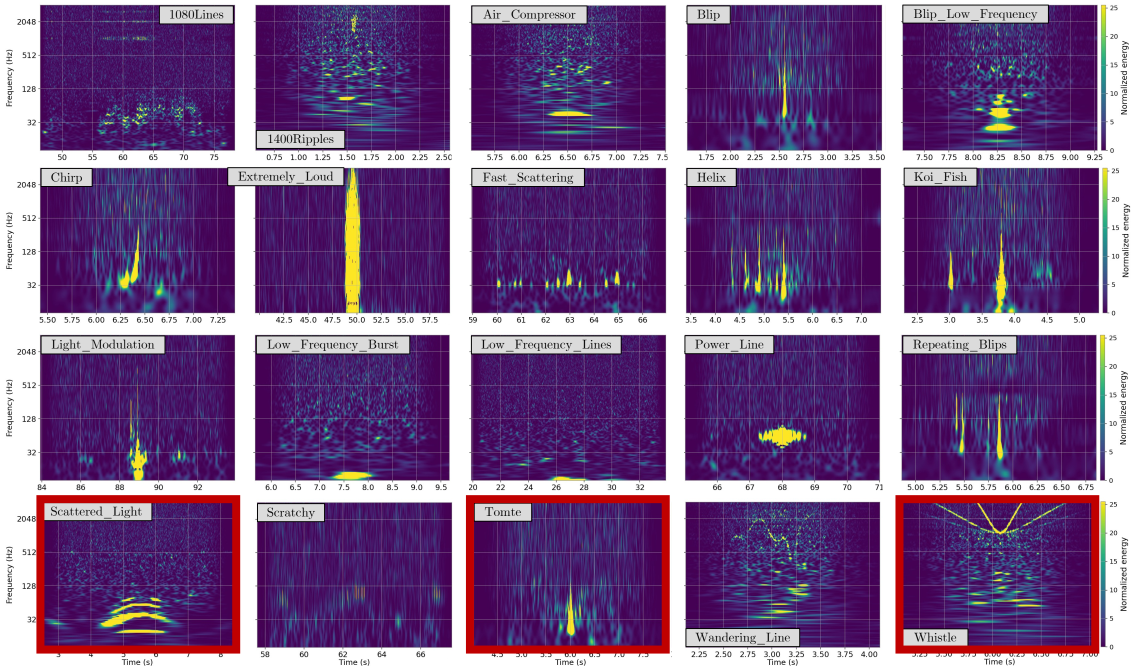
The input data

We preprocess 350 auxiliary channels → encode 50 auxiliary channels



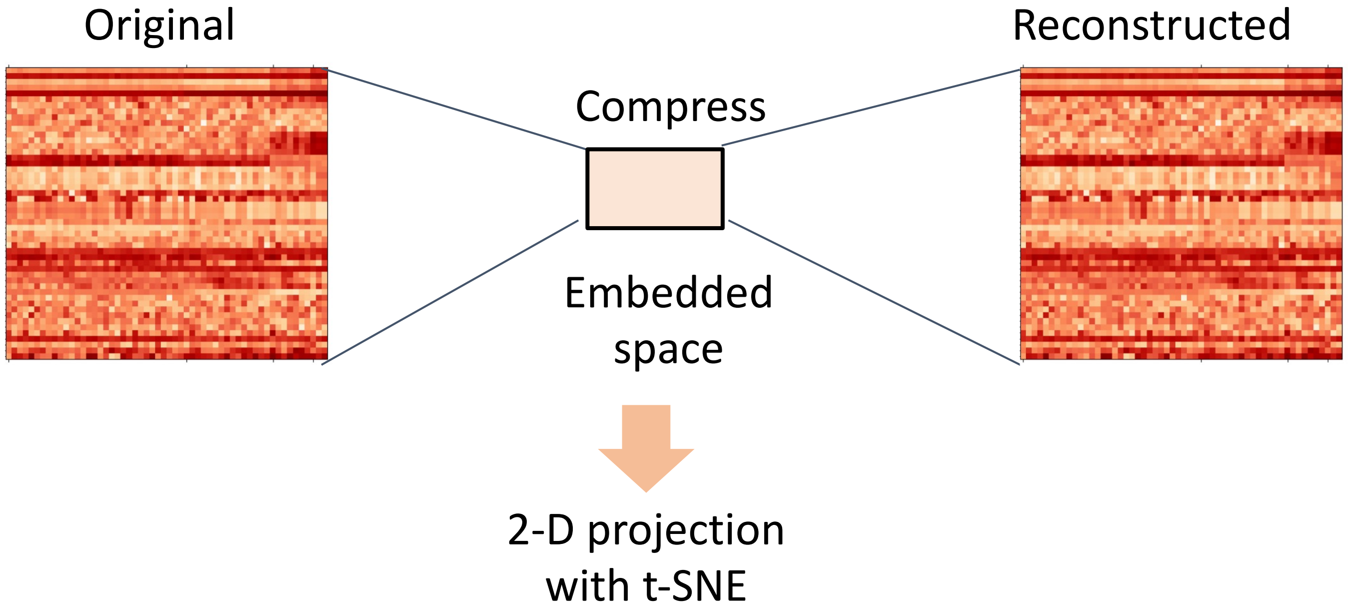
SCAN ME

ArXiv: 2310.03453



Classified with supervised learning

Results of compressed data



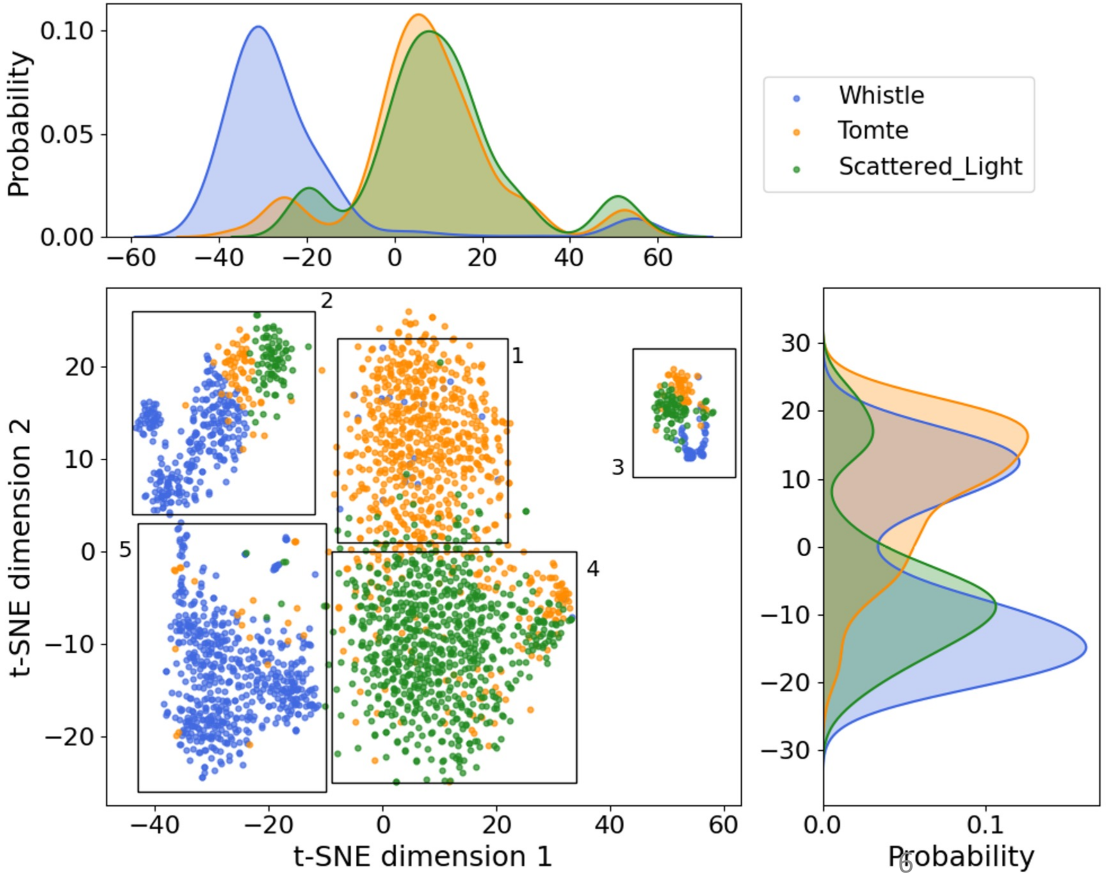
Benchmarking against supervised learning:
Clusters consistent with Gravity Spy, but

- Gravity Spy → spectrograms of h(t)
- Our work → fractal dimension with auxiliary channels

OK, now let's represent anomalies in spectrograms of h(t)

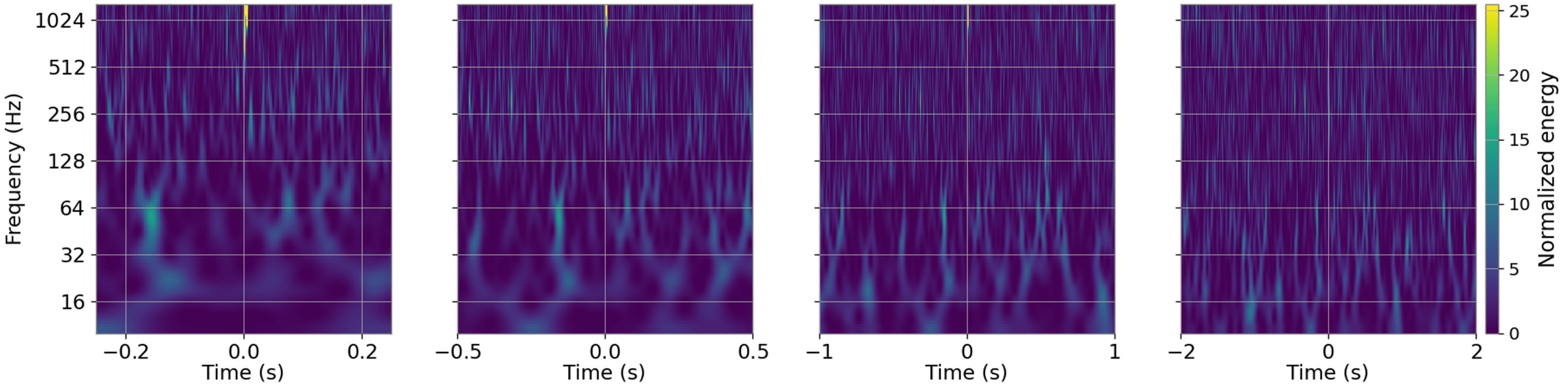
SCAN ME

ArXiv: 2310.03453

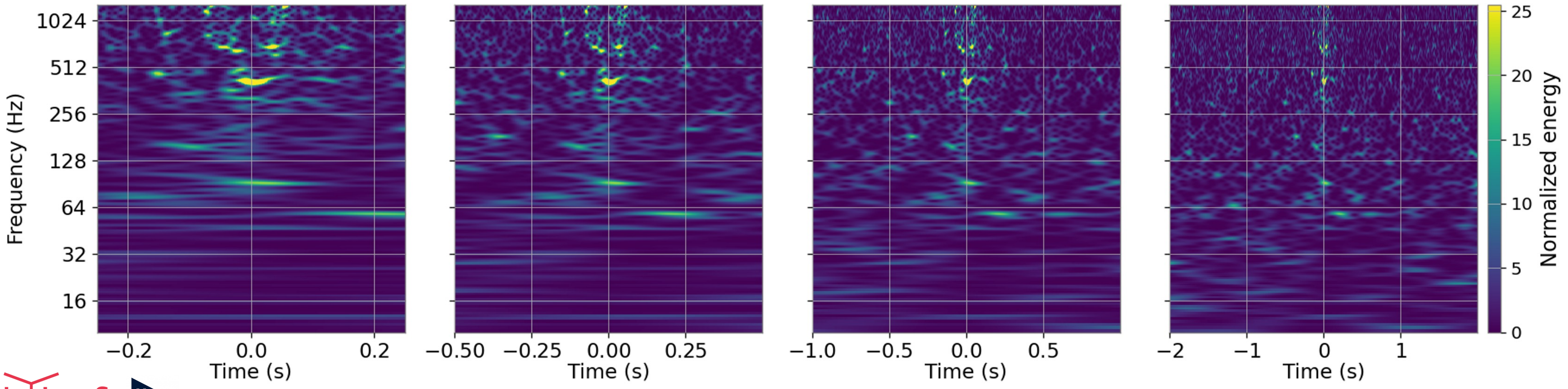


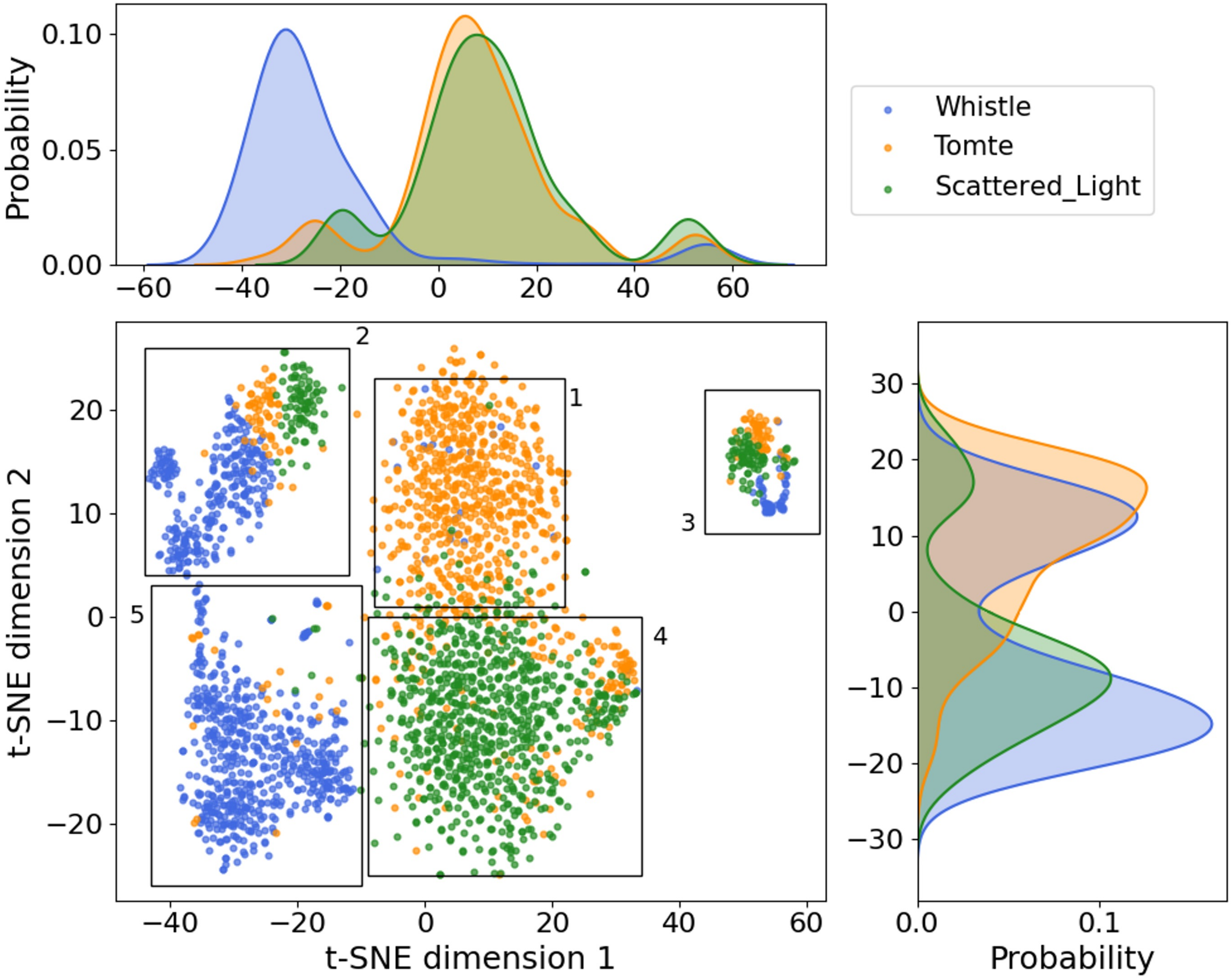
Section 1

“Standard” Whistle



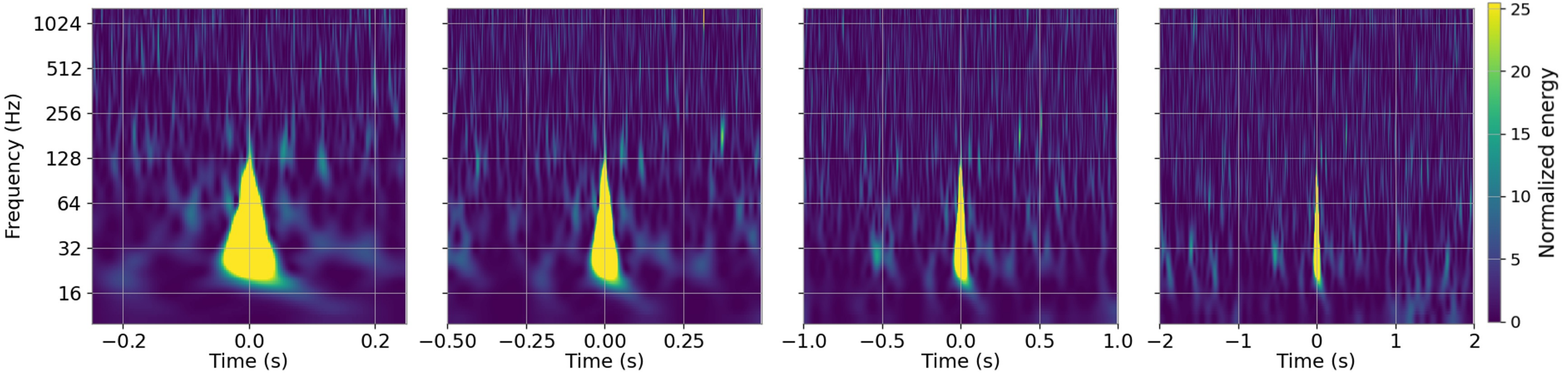
Labelled as Whistle but anomalous Whistle morphology



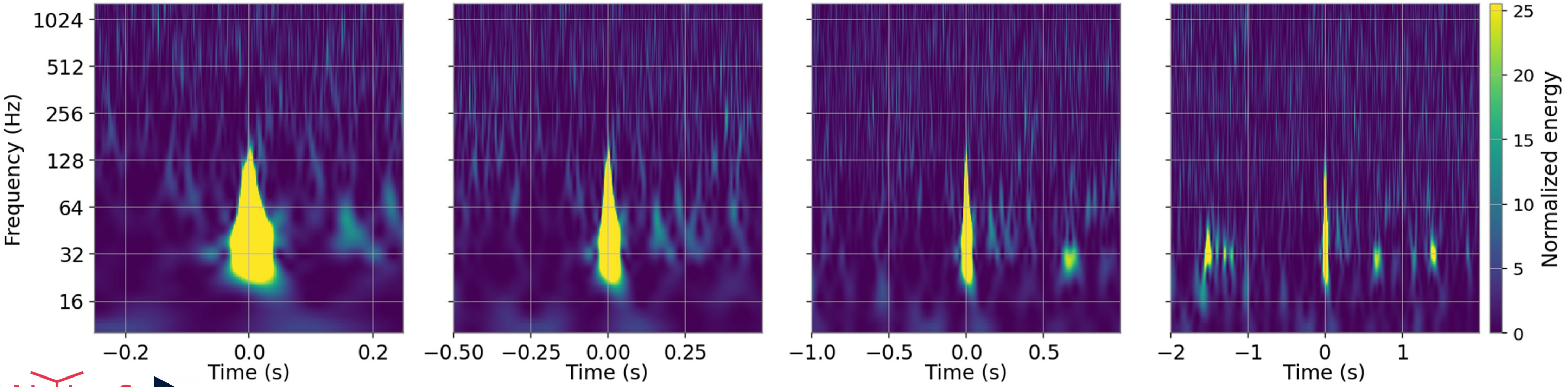

Probability
0.10
0.05
0.00
−60
−40
−20
0
20
40
60
Whistle
Tomte
Scattered_Light
1
2
3
4
5
t-SNE dimension 2
t-SNE dimension 1
Probability
0.0
0.1

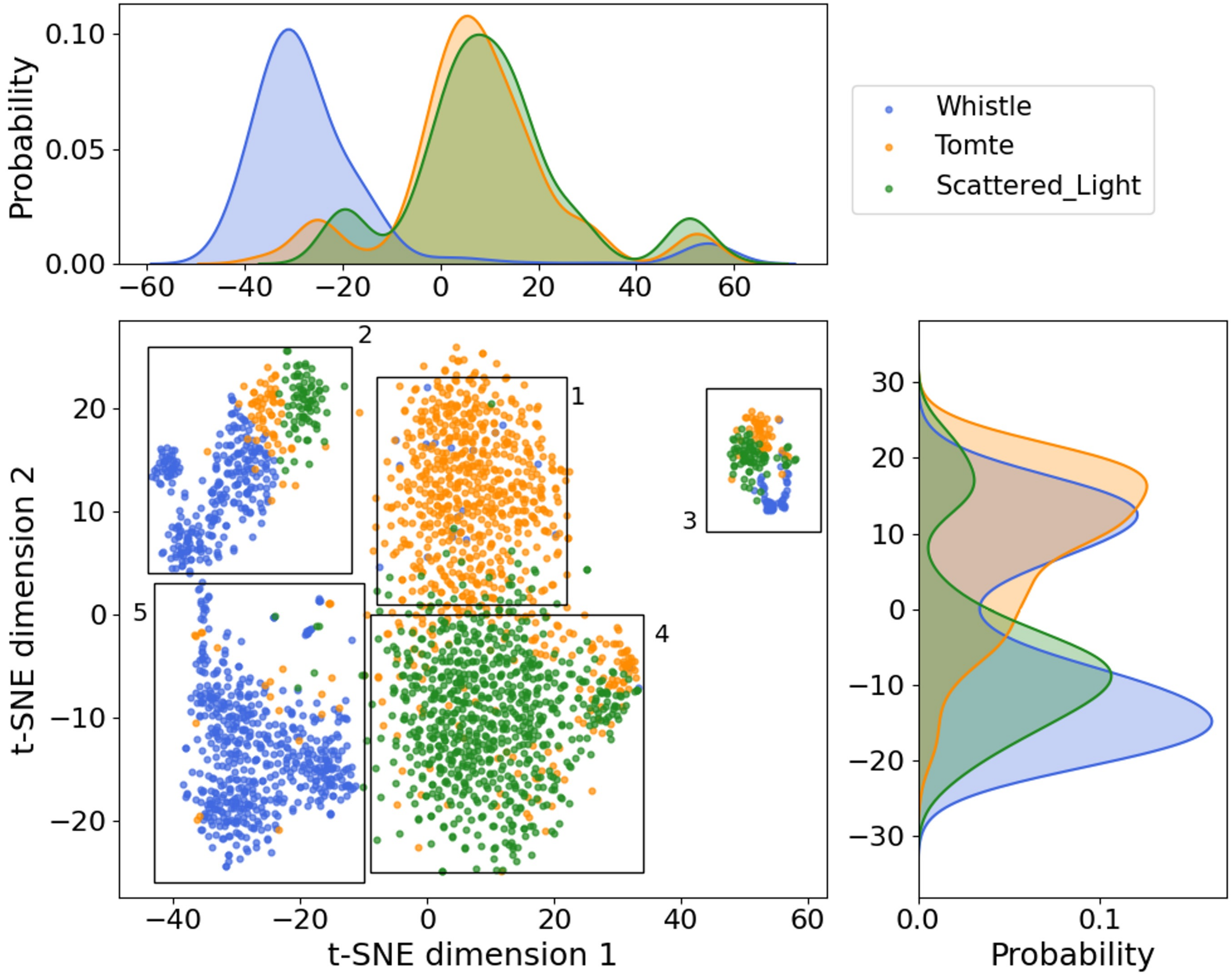
Section 2

“Standard” Tomte



Label Tomte, but Tomte overlap with Scratchy

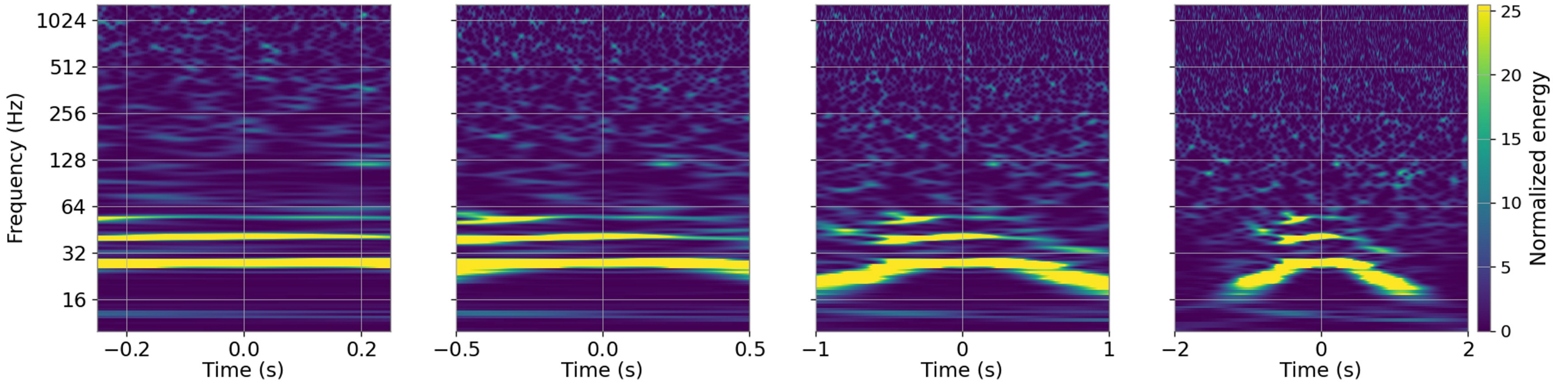




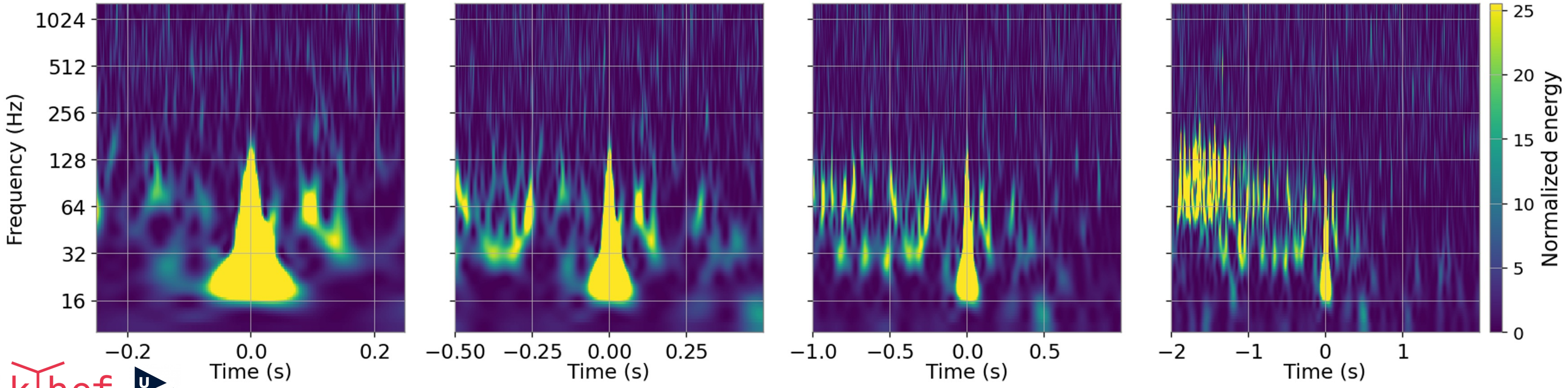
Probability
0.10
0.05
0.00
−60
−40
−20
0
20
40
60
Whistle
Tomte
Scattered_Light
t-SNE dimension 2
t-SNE dimension 1
1
2
3
4
5
20
10
0
−10
−20
30
−30
0.0
0.1
Probability

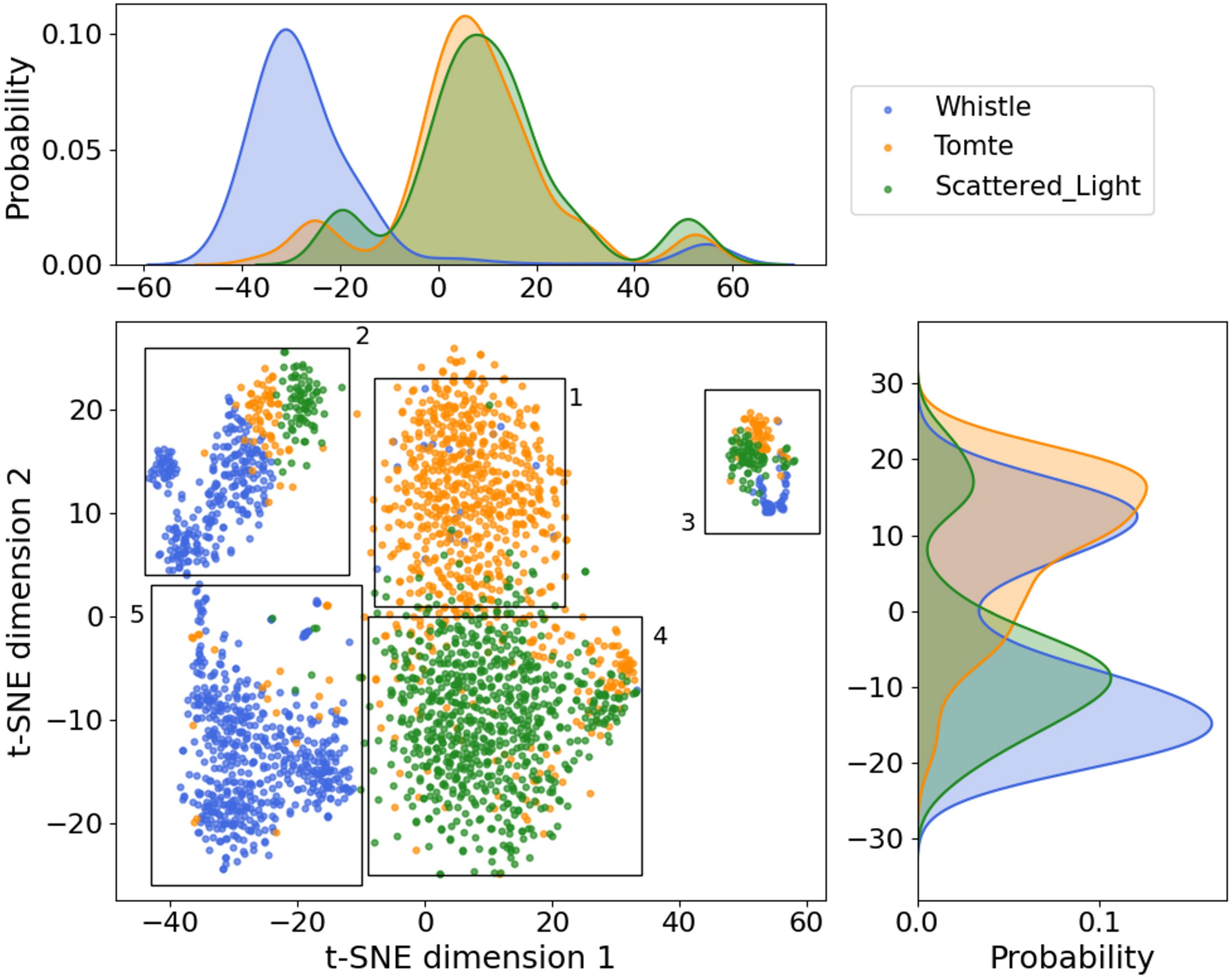
Section 3

"Standard" Scattered Light



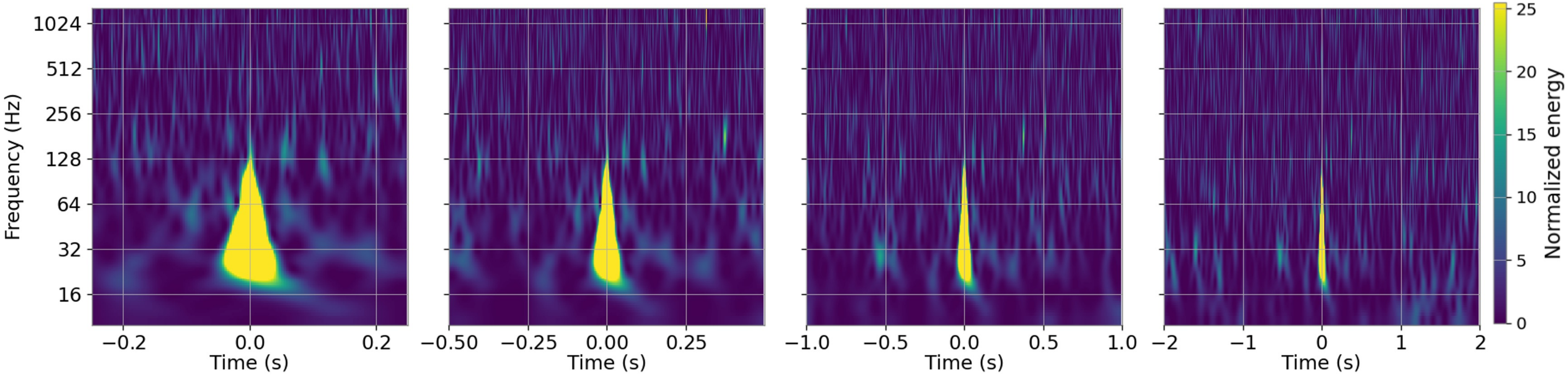
Labelled as a Scattered Light, but overlap between Tomte and Scratchy



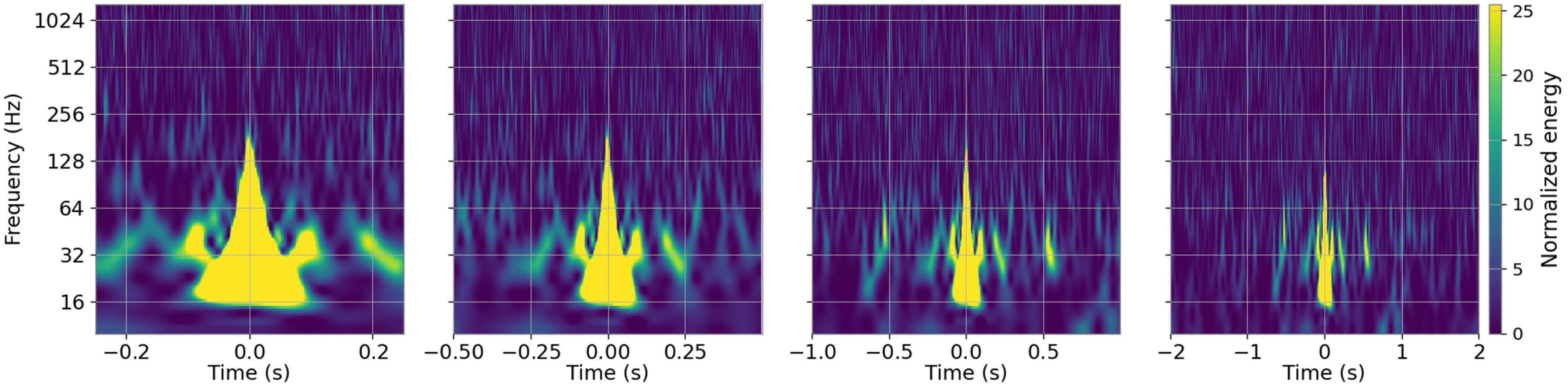

Probability
0.10
0.05
0.00
−60
−40
−20
0
20
40
60
Whistle
Tomte
Scattered_Light
1
2
3
4
5
20
10
0
−10
−20
t-SNE dimension 2
−40
−20
0
20
40
60
t-SNE dimension 1
30
20
10
0
−10
−20
−30
0.0
0.1
Probability

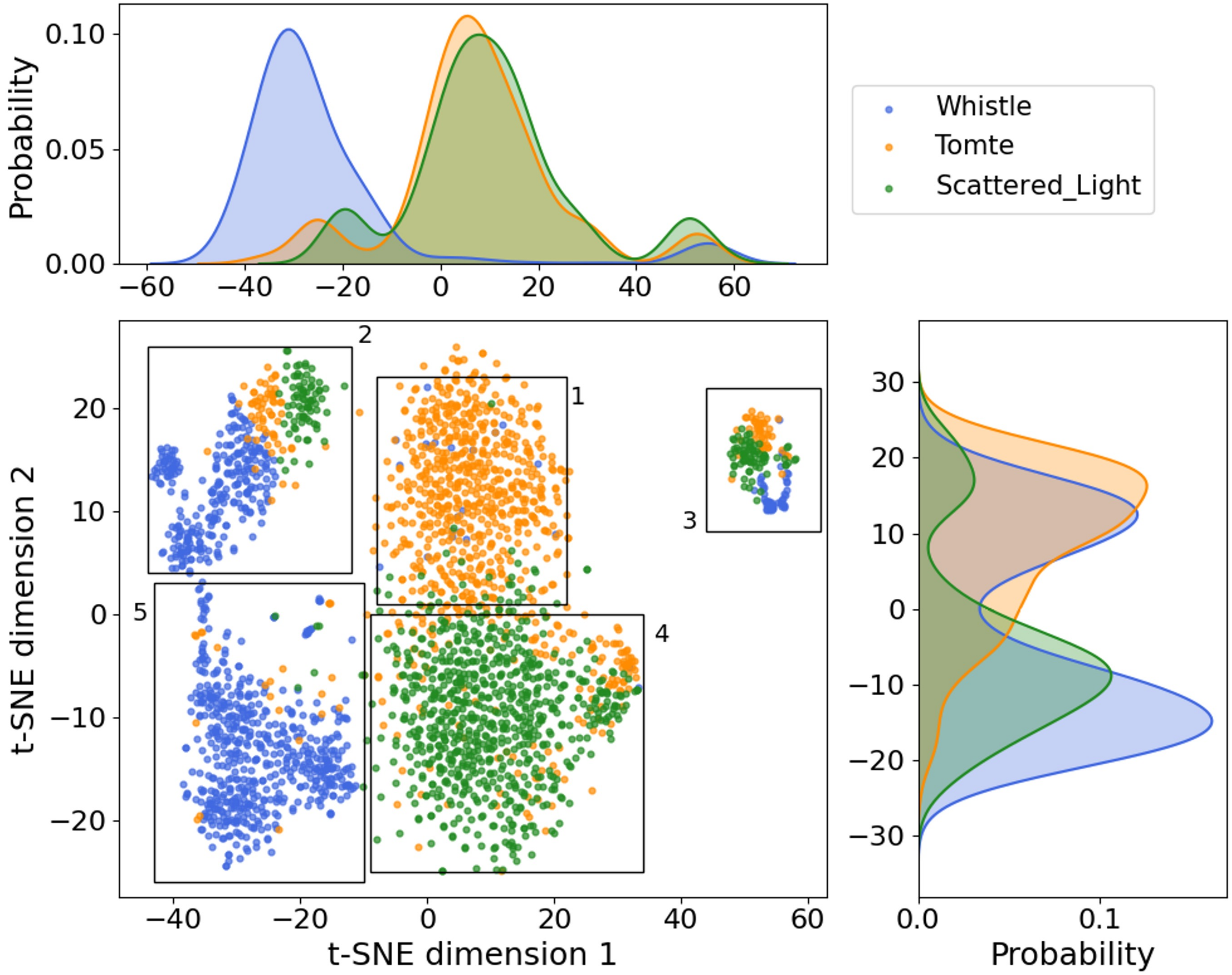
Section 4

"Standard" Tomte



Labelled as Tomte but it is a Koi Fish

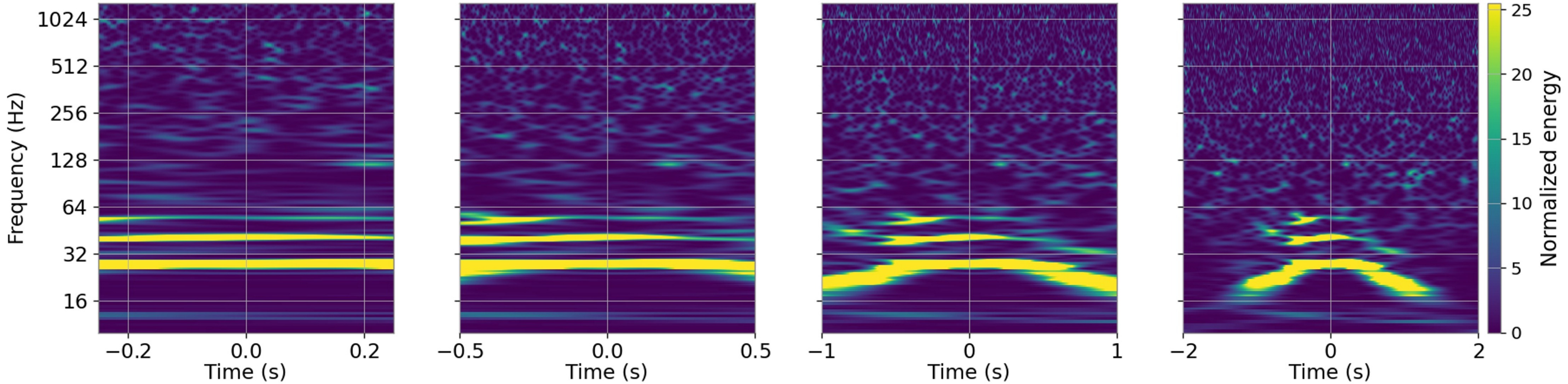




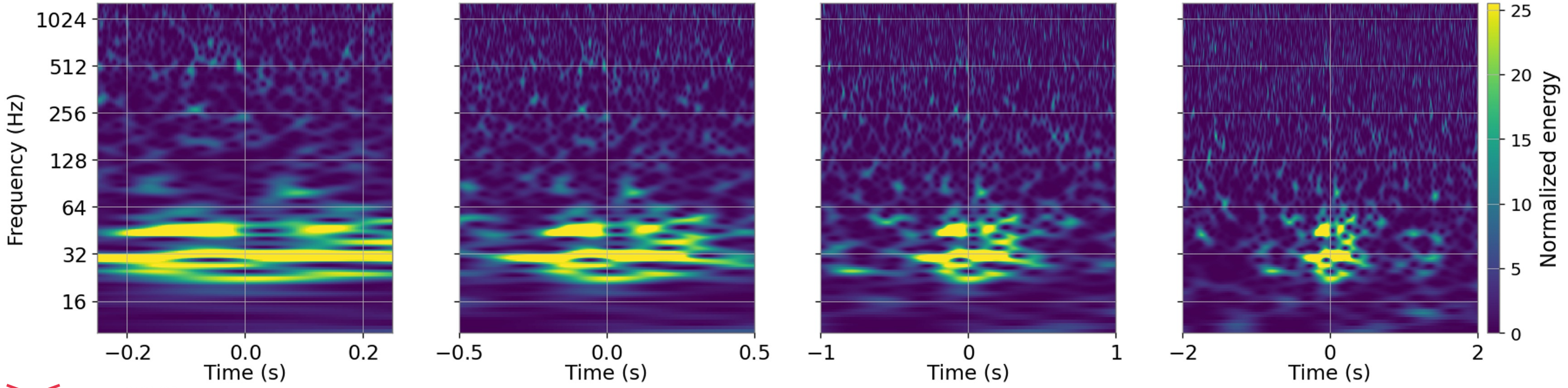
Probability
0.10
0.05
0.00
−60
−40
−20
0
20
40
60
Whistle
Tomte
Scattered_Light
t-SNE dimension 2
t-SNE dimension 1
1
2
3
4
5
20
10
0
−10
−20
−40
−20
0
20
40
60
30
20
10
0
−10
−20
−30
0.0
0.1
Probability

Section 5

“Standard” Scattered Light



Labelled as Scattered Light but unknown morphology



Results of compressed data

In total 177 anomalies were found, which constitute 6,6% of the data.

- **Anomalous whistles (49):**
  - 45% unknown morphologies, 28% misclassifications, 27% overlaps.
- **Anomalous Tomtes (57):**
  - 32% unknown morphologies, 21% misclassifications, 47% overlaps.
- **Anomalous Scattered Lights (71):**
  - 28% unknown morphologies, 72% misclassifications, 1 overlap.

SCAN ME

ArXiv: 2310.03453

Conclusions and future work

- ✓ Fractal dimension representation is complementary to h(t)
- ✓ Unsupervised learning can reveal misclassifications of supervised learning, glitch overlaps and novel morphologies
- ➢ Extend to glitch populations of GW detectors
- ➢ Relate glitches to auxiliary channels via explainable ML

SCAN ME

ArXiv: 2310.03453

Thank you for listening! Questions?

m.lopez@uu.nl

SCAN ME

ArXiv: 2310.03453