Preparing for the future gravitational wave burst searches with machine learning techniques

TAUP conference 2021

Sophie Bini, Ph.D. student Università di Trento

Gabriele Vedovato, Marco Drago, Odysse Halim, Sergei Klimenko, Claudia Lazzaro, Andrea Miani, Edoardo Milotti, Giovanni Prodi, Francesco Salemi, Mishra Tanmaya, Shubhanshu Tiwari, Andrea Virtuoso

Preparing for the future gravitational wave burst searches with machine learning techniques

TAUP conference 2021

Sophie Bini, Ph.D. student Università di Trento

Gabriele Vedovato, Marco Drago, Odysse Halim, Sergei Klimenko, Claudia Lazzaro, Andrea Miani, Edoardo Milotti, Giovanni Prodi, Francesco Salemi, Mishra Tanmaya, Shubhanshu Tiwari, Andrea Virtuoso

Layout

I. All–sky short duration searches and Coherent WaveBurst (cWB) pipeline

II. New tools for the next observing run:

- pinpoint specific noise morphology
- upgrading of selection criteria and ranking procedures applied

III. Preliminary results

All sky short duration searches

Search for generic short-duration gravitational wave transients, without any assumptions on incoming signal direction, polarization or morphology.

Possible sources are: core-collapse supernovae, neutron star excitations, non-linear memory effects, cosmic strings

[Abbott, R., et al. "All-sky search for short gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run." arXiv (2021)]

cWB pipeline

Algorithm based on the maximum-likelihood-ratio statistic over all sky directions, applied to excesses of signal power in the time–frequency domain representation from the network of detectors

coherent WaveBurst
an open source software for
gravitational-wave data analysis

[S. Klimenko, G. Vedovato, M. Drago, F. Salemi, V. Tiwari, G. A. Prodi, C. Lazzaro, S. Tiwari, F. Da Silva, and G. Mitselmakher, Phys. Rev. D 93, 042004 (2015)]

Layout

I. All–sky short duration searches and cWB pipeline

II. New tools for the next observing run:

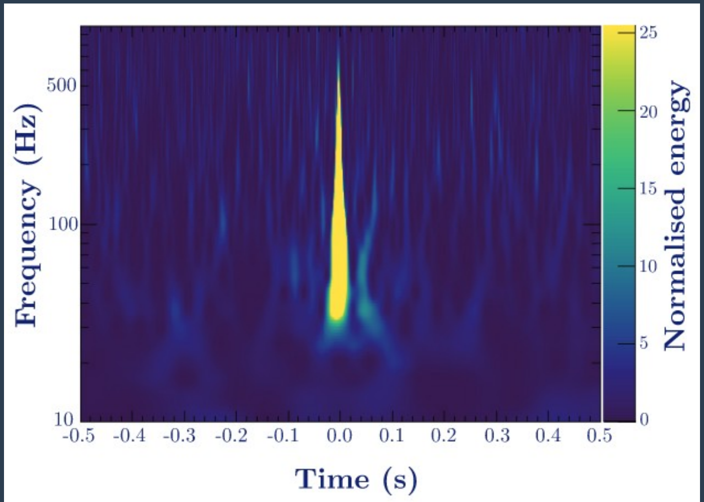
- pinpoint specific noise morphology

- upgrading of selection criteria and ranking procedures applied

III. Preliminary results

New tool → Autoencoder

There are short duration noise family particularly concerning, among which blip glitches[1,2]



Up to now, in cWB they were adressed by 2 parameters (Qveto[3]):

1) ratio between main peak amplitude and nearby peaks

2) model the signal as a CosGaussian and evaluate its Qfactor

[1. *Cabero, Miriam, et al. "Blip glitches in Advanced LIGO data." Classical and Quantum Gravity 36.15 (2019),*

2. Nitz, Alexander H. "Distinguishing short duration noise transients in LIGO data to improve the PyCBC search for gravitational waves from high mass binary black hole mergers." Classical and Quantum Gravity 35.3 (2018)

3. https://gwburst.gitlab.io/

Autoencoder - introduction

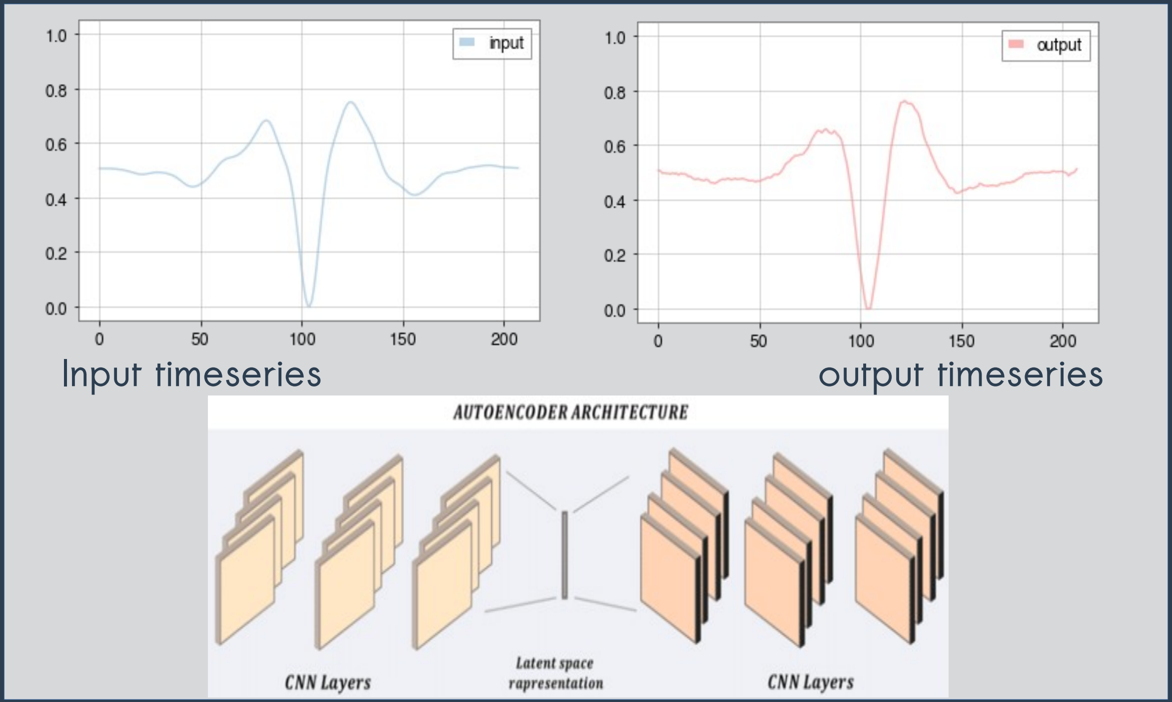
We investigate a more flexible algorithm, that can pinpoint also other specific morphologies

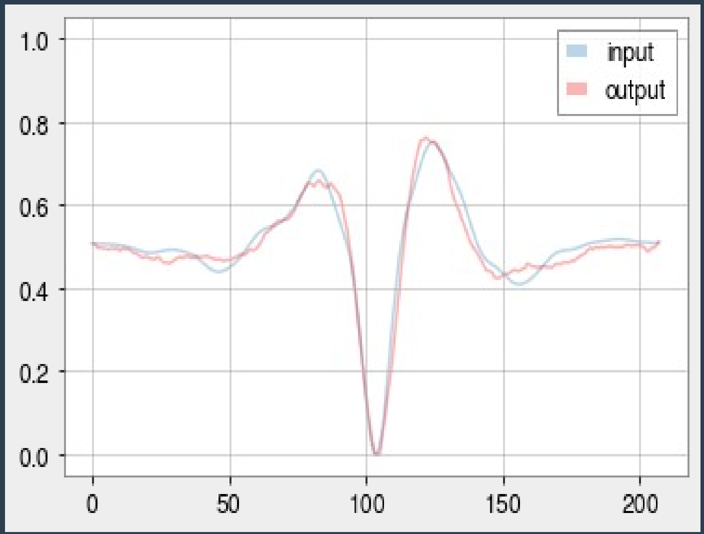
↓

An autoencoder neural network learns a specific morphology and evaluates how much a trigger is similar to the signals in the training set

Autoencoder - training

Trained on a single signal family (Blip glitches) using GravitySpy [1] labelled glitches

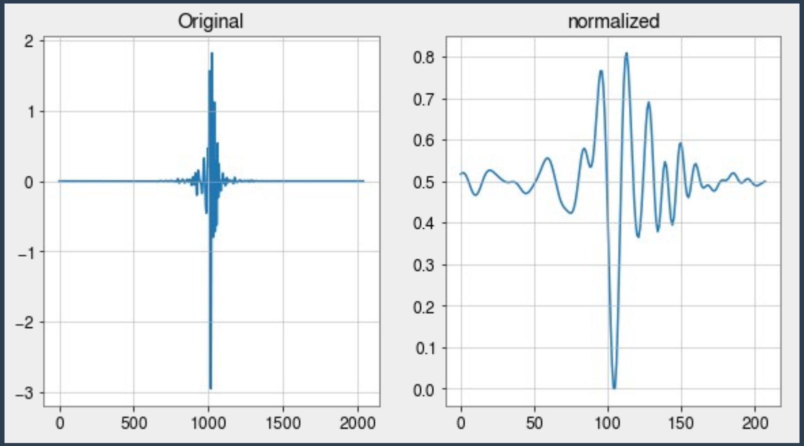




[1. Zevin, Michael, et al. "Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science." Classical and quantum gravity 34.6 (2017)]

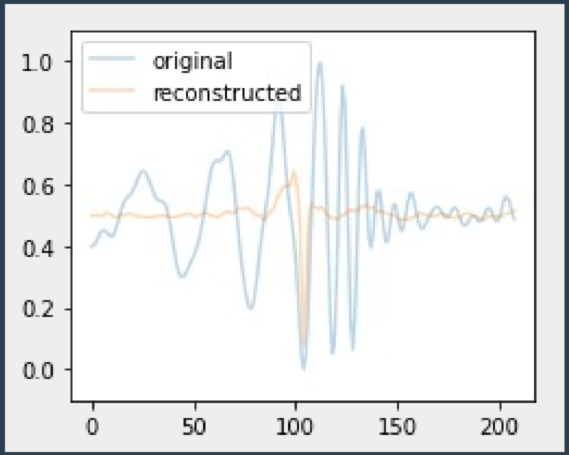
Autoencoder – input and fit

Input → cWB reconstructed waveforms
(normalized and in a fixed window
centered around the peak)



A Blip timeseries before and after the pre-processing

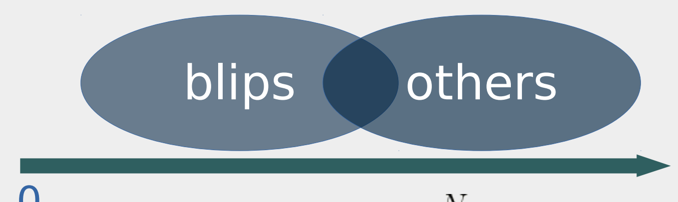
Fit → the timeseries reconstructed
by the autoencoder



Example on a BH – BH coalescence injection:
the network struggles to reconstruct a timeseries different
from the one in the training dataset

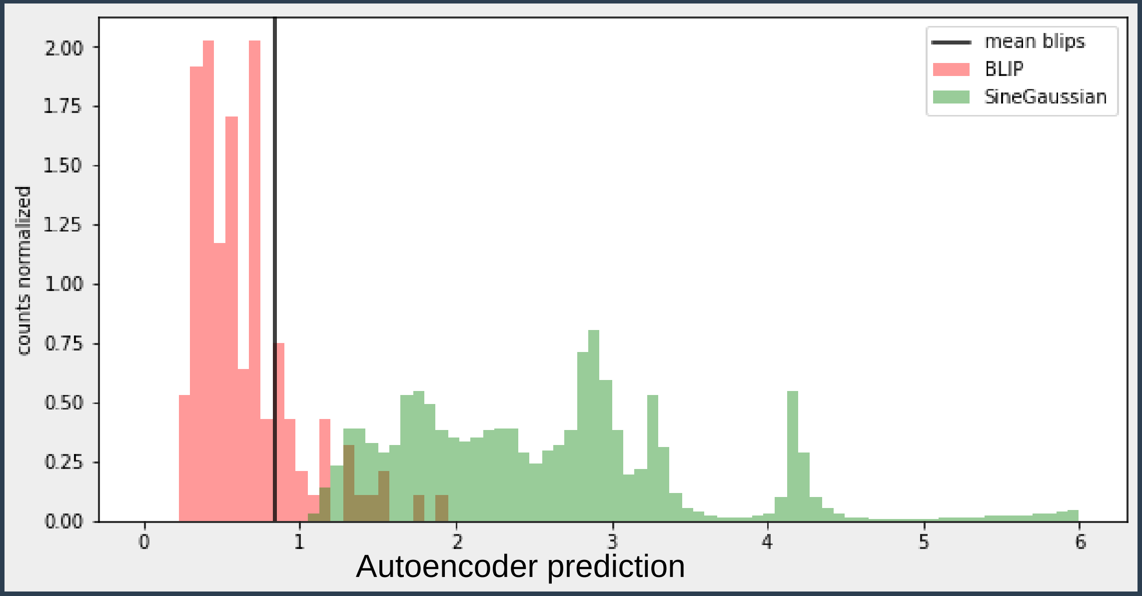
Autoencoder - prediction

Prediction → the mean square error between the input and the reconstructed timeseries



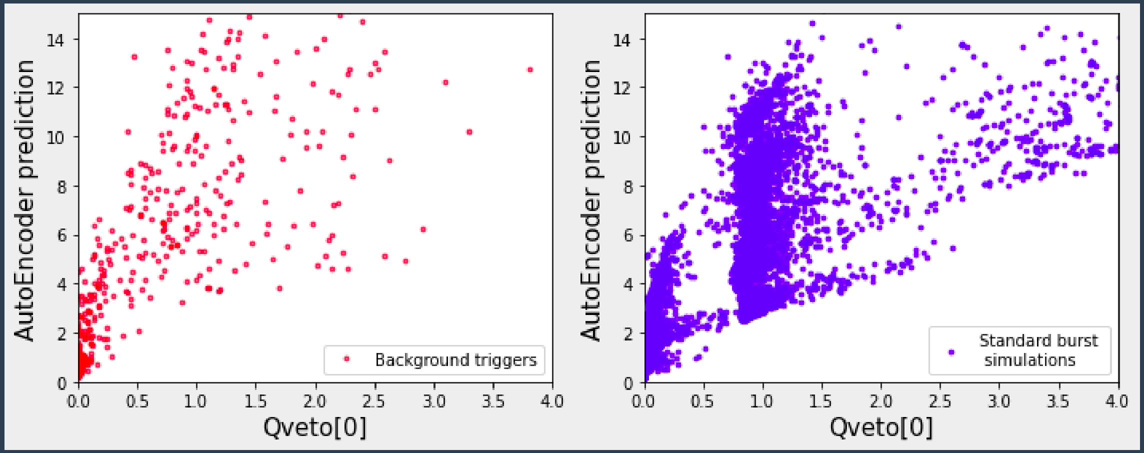
$$\text{Prediction} = \sum_{i}^{N} (x_i - y_i)^2$$

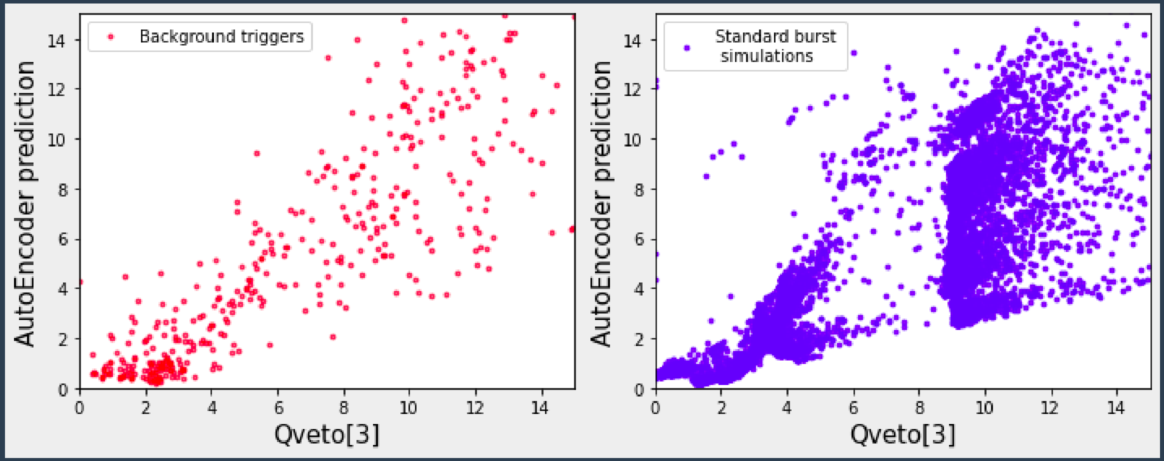
Example with Blips and SineGaussian injections



Autoencoder - comparison with the existing solution

Autoencoder prediction is highly correlated with the Qveto parameters





Layout

I. All–sky short duration searches and cWB pipeline

II. New tools for the next observing run:

- pinpoint specific noise morphology
- upgrading of selection criteria and ranking procedures applied

III. Preliminary results

New tool → XGBoost

Up to now, selection criteria and ranking procedures based on hard threshold on different parameters

A decision tree learning algorithm is being tested to improve signal-noise classification in all sky short searches.

It builds the model using a subset of summary statistics generated from the cWB reconstructed events

[Mishra, Tanmaya, et al. "Optimization of model independent gravitational wave search using machine learning." arXiv (2021). Enhancing detection of gravitational waves with machine learning, link]

Layout

I. All–sky short duration searches and cWB pipeline

II. New tools for the next observing run:

- pinpoint specific noise morphology
- upgrading of selection criteria and ranking procedures applied

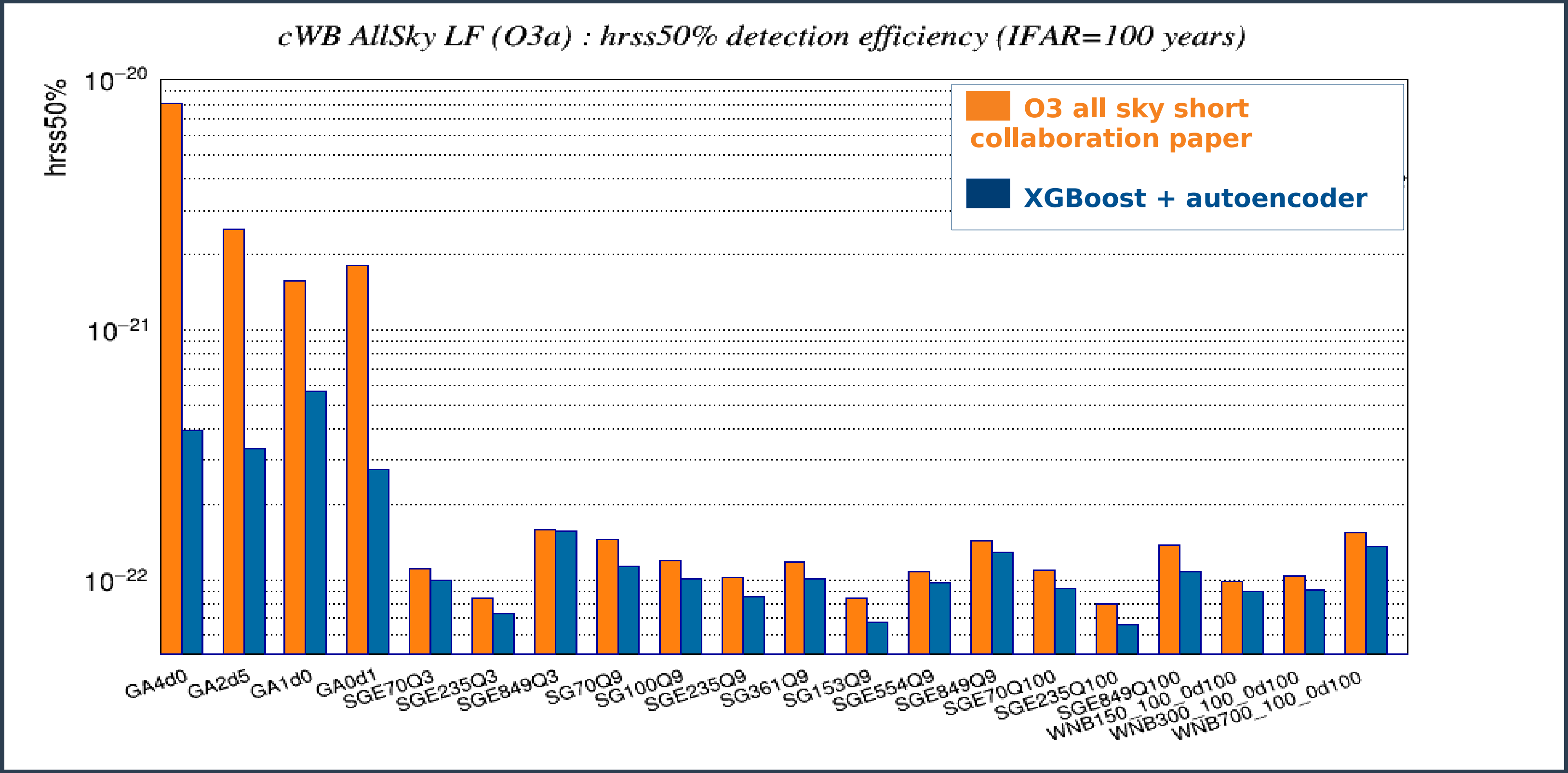
III. Preliminary results

Preliminary results standard burst injections

Autoencoder trained on Blip glitches

XGBoost trained on standard burst injections*

*** a set of ad hoc waveforms (Gaussian Pulse (GA), sine-Gaussian wavelets(SG), White Noise Burst (WNB)) used to estimate the search sensitivity**

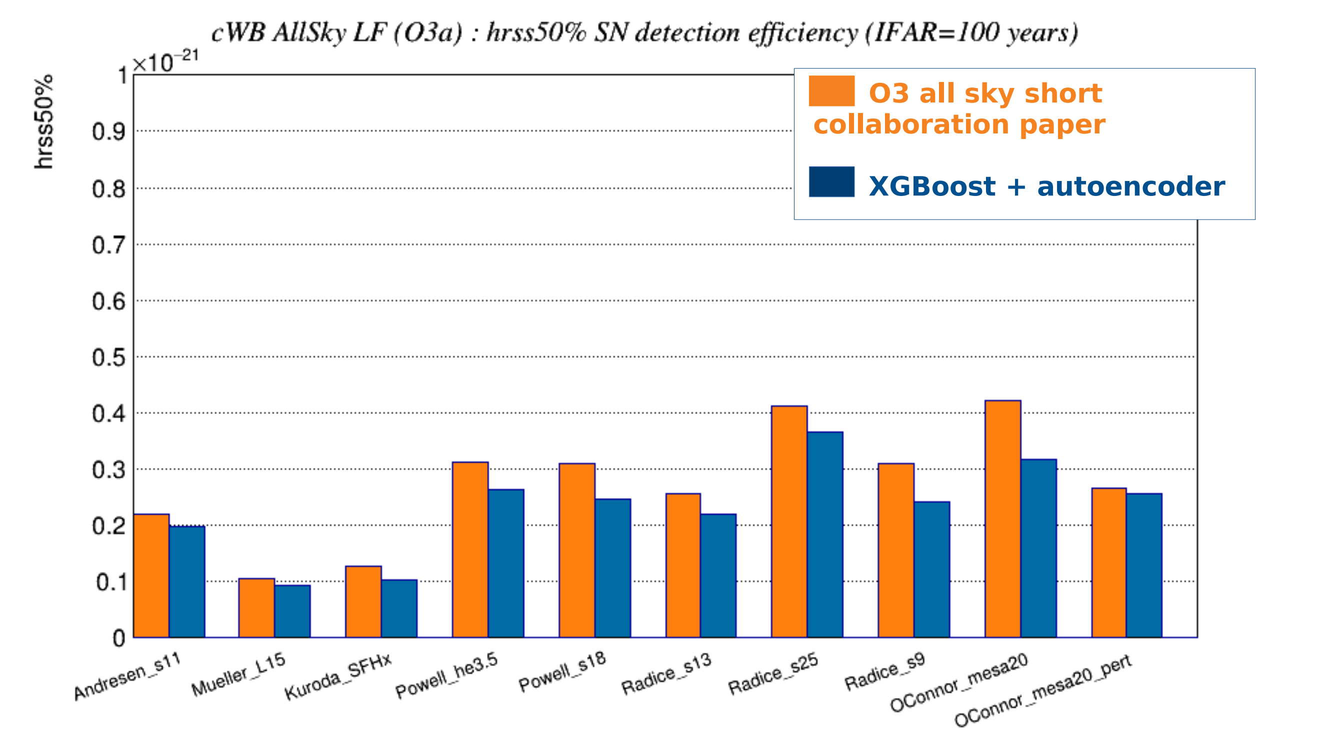


Preliminary results supernovae injections

Autoencoder trained on Blip glitches

XGBoost trained on standard burst injections*

* a set of ad hoc waveforms (Gaussian Pulse (GA), sine-Gaussian wavelets(SG), White Noise Burst (WNB)) used to estimate the search sensitivity



Robustness tests

We want that both these procedures are robust and do not limit the general character of the search.

Several tests are ongoing:

- Tests on injections not included in the training data set

- Check of performance degrations after the removal of one or more input parameters

Conclusions

We are testing two new tools in cWB pipeline to pinpoint specific noise morphology and to avoid hard threshold in the post-production stage.

Appropriate machine learning tools could be promising also for all-sky unmodelled gravitational wave searches

The results seems promising and more tests will be performed

(short author list paper regarding XGBoost for all sky searches is in preparation)