# Einstein Telescope
# Mock Data and Science Challenge

Tania Regimbau on the behalf of ET OSB Division 10

# What/Why ET MDC

- Produce realistic simulated data containing instrumental noise + GW signal from population of sources at the output of E1, E2 and E3

- Test, develop, optimize data analysis pipelines and parameter estimation

- Stress-testing the current computational infrastructure

# What's in the ET MDC data

- series of MDCs with increasing complexity

- first MDC with Gaussian colored noise and a cosmological population of CBCs (more details in the next slide)

- Next MDCs with all type of sources, glitches, correlated noise (need to be implemented in the simulation code)

# First ET MDC

- 1 month duration
- Gaussian colored noise (ET-D 10 km, fMin=5Hz), triangle configuration (+CE)
- Population of BBH (10%), BNS (87%) and NSBH (3%) from M. Mapelli's group (see CoBA) with isotropic distribution in the sky
- Waveforms: IMRPhenomXPHM for BBHs and BHNSs, and SimIMRPhenomP with tidal effects NRTidalv2_V for BNSs.
- SNR>8: 156902 BNSs , 73693 BBHs, 7563 BHNSs
  SNR>12: 57576 BNSs , 63711 BBHs, 3574 BHNSs

# The Data

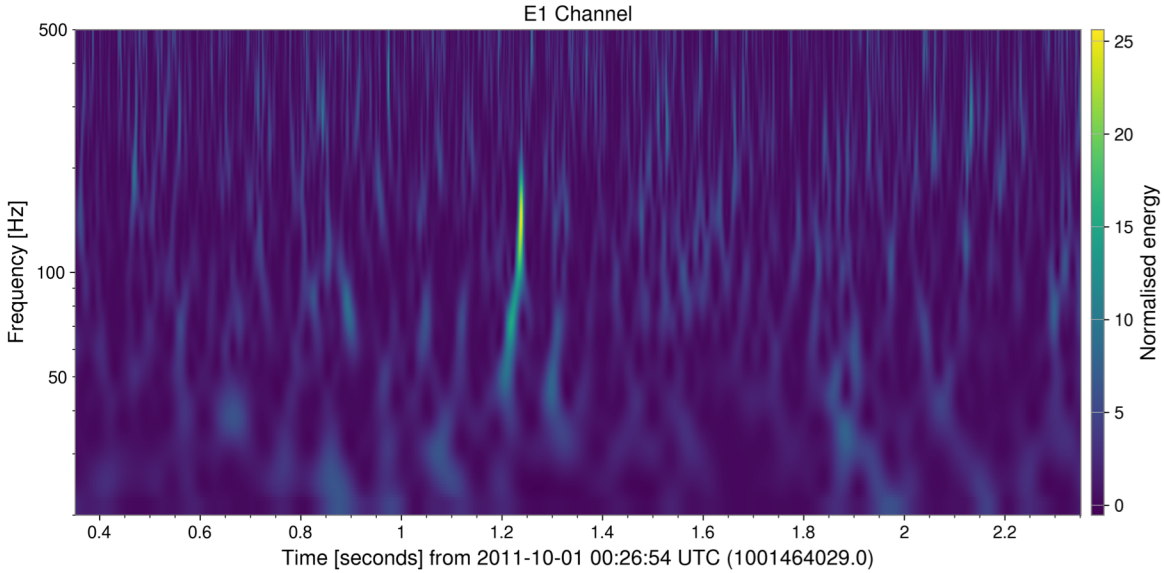
- /cvmfs/et-gw.osgstorage.org/et-gw/PUBLIC/MDC1
- Instructions here: https://wiki.et-gw.eu/EIB/SoftwareFrameworks/WebHome?validation_key=e2698d03b6eff5856cfab4654d3fbfe5

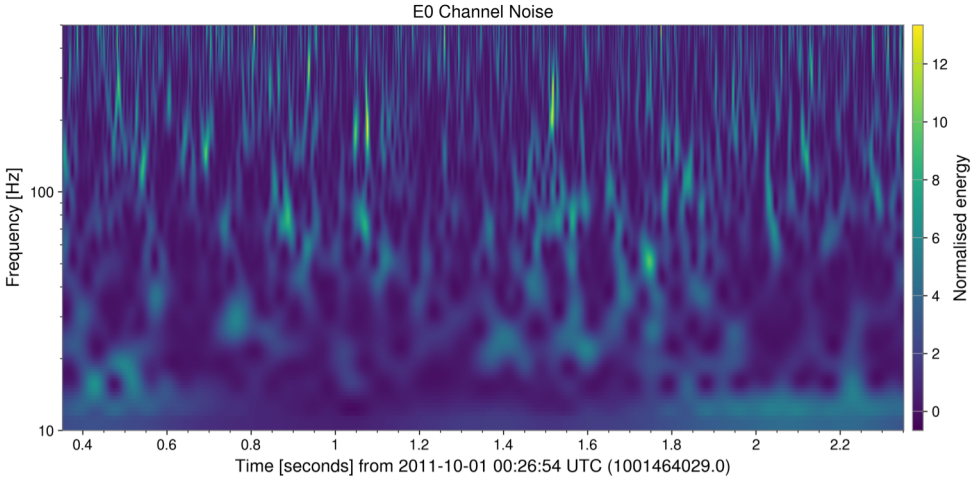
- Frame files for E1, E2, E3 and E0 (set 0: noise only, set 1: noise+GWs)

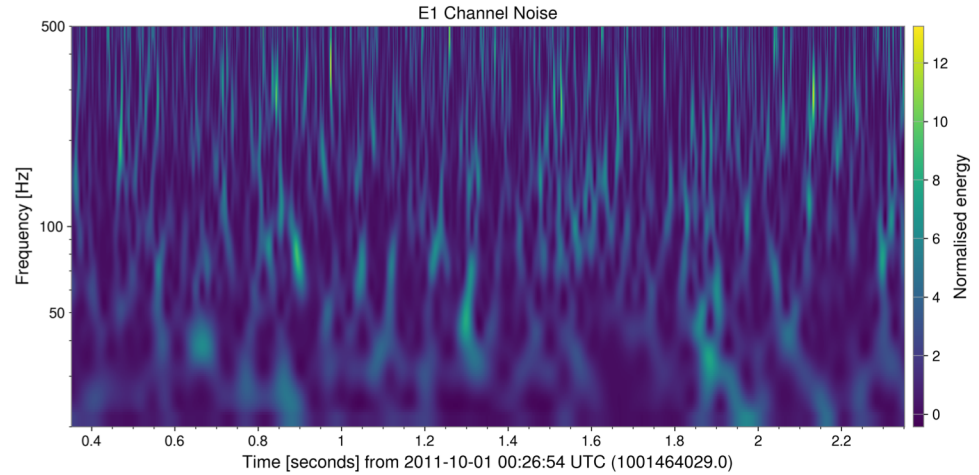
1300 frames per detector of length 2048s and sampling rate 8192 Hz - In total 1.3 TB (20 GB memory requested for some jobs)

+ text files with lists with the source parameters and expected SNR
Soon data for CE

# The Data



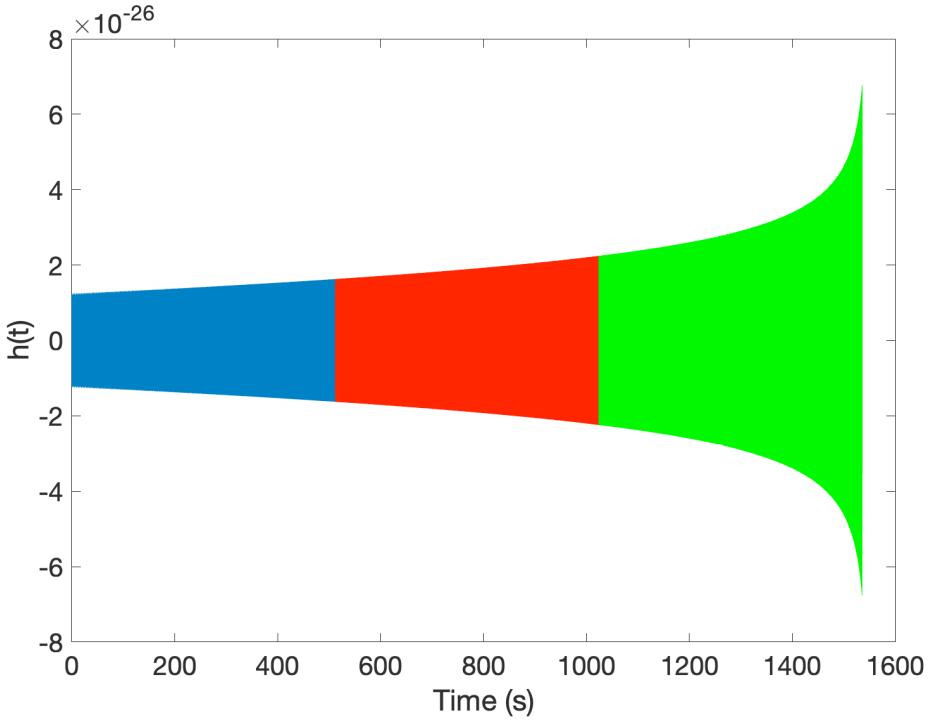



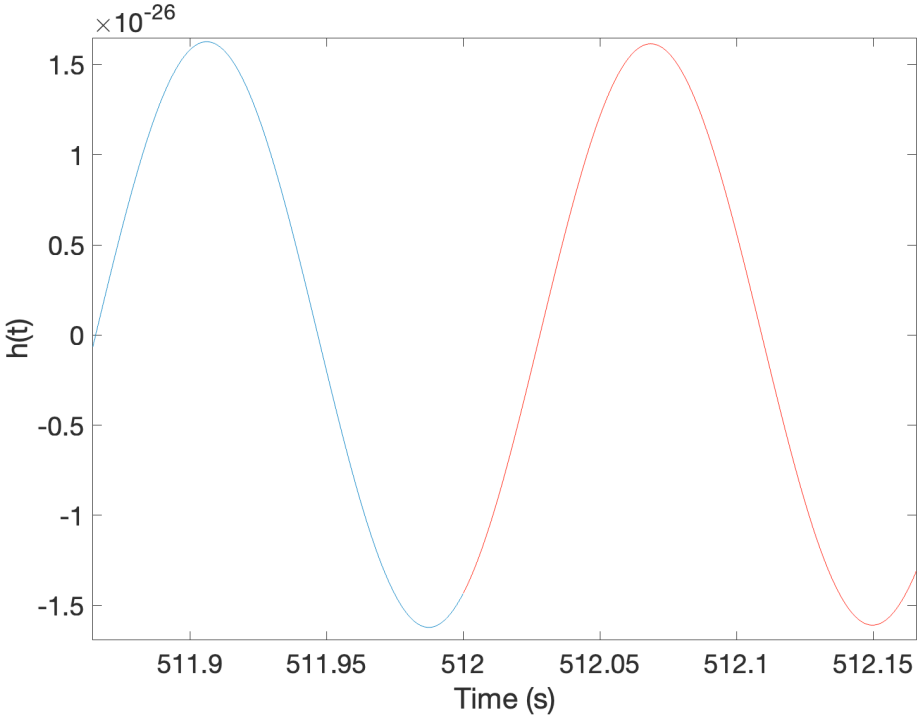




Plots produced by Jishnu Suresh

# Generation code not optimized

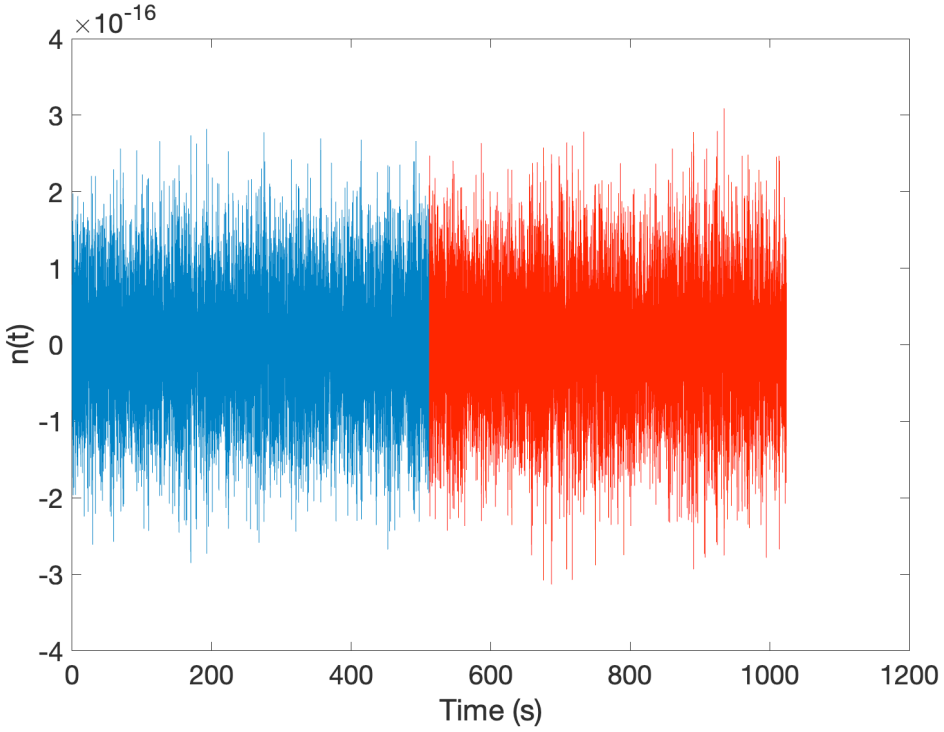
Data are produced in parallel using Condor for each segment. In order to keep the continuity of the GW signal, the ET waveforms are produced several times. Also true for the noise.

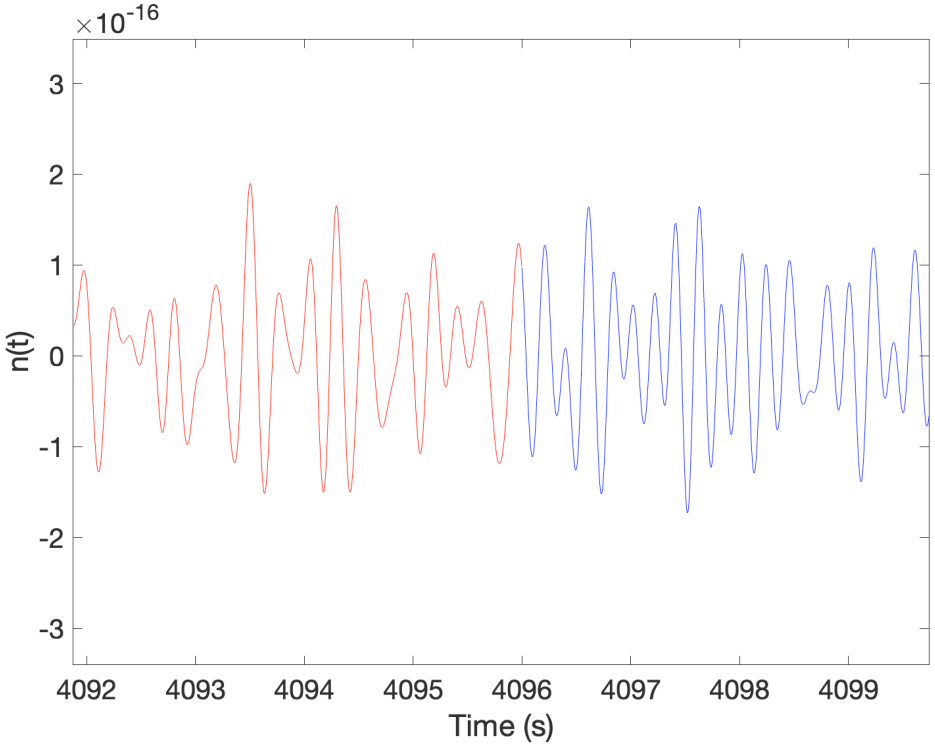


# Generation code not optimized

Data are produced in parallel for each segments. In order to keep the continuity of the GW signal, the log ET waveforms are produced several times. Also true for the noise.



# Challenges

- Stress-testing the current computational infrastructure
- CVMFS, Open Science Grid
- Computational cost of the current pipelines for high rate/large template bank

- Detection efficiency
- Parameter inference in presence of overlapping signals

- Infer the rate of each population, the mass/spin distributions, the NS EOS, the cosmological model.