

OSB Division 10 Data Analysis Platform

Elena Cuoco, Tania Regimbau, Chris Van Den
Broeck, John Veitch

May 10th 2023, ET Symposium

Division 10 organisation

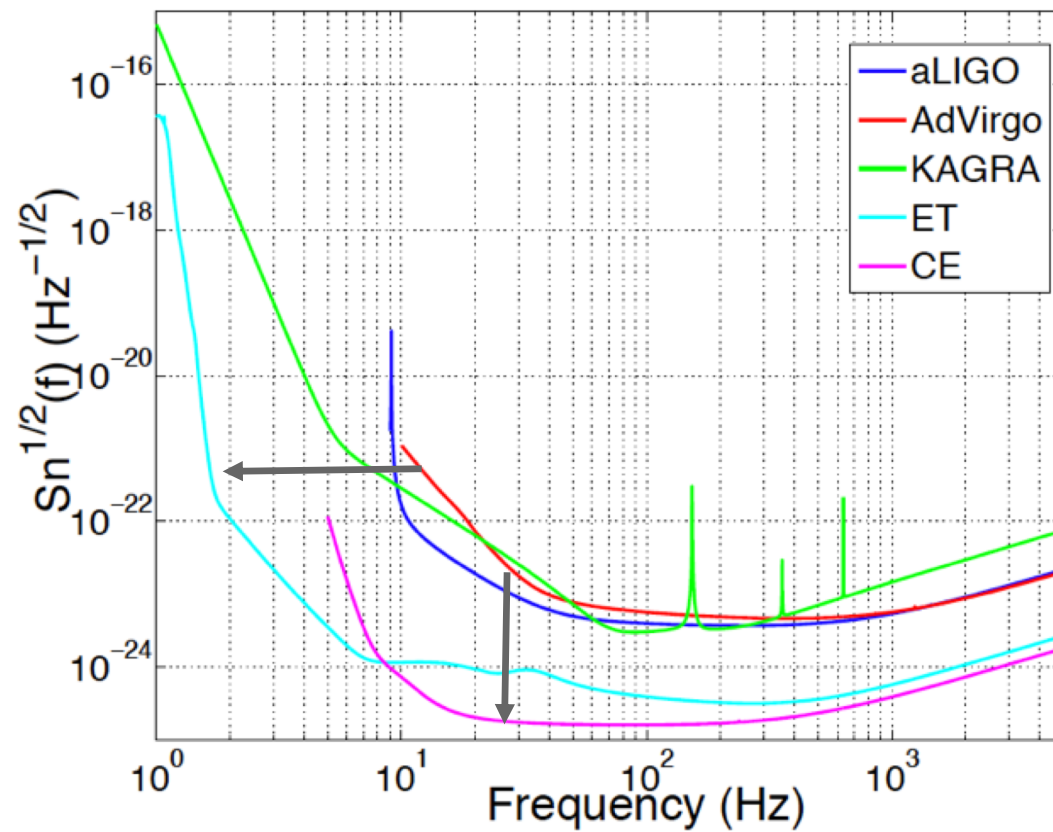
- 4 chairs: Elena Cuoco, Tania Regimbau, Chris Van den Broeck, John Veitch
- 122 subscriptions to the mailing list (<http://mail.et-gw.it/mailman/listinfo/et-osb-DA>)
- Monthly meetings
- Wiki page: <https://wiki.et-gw.eu/OSB/DataAnalysisPlatform>
- Liaison persons with other divisions and EIB

Goal of Division 10

The goal of Division 10 is to be ready for the analysis and parameter estimation with 3G data, in terms of methods and software, in order to exploit the full potential of 3G detectors

- new challenges due to a better sensitivity (very high SNR, long waveforms, overlapping events, new sources)

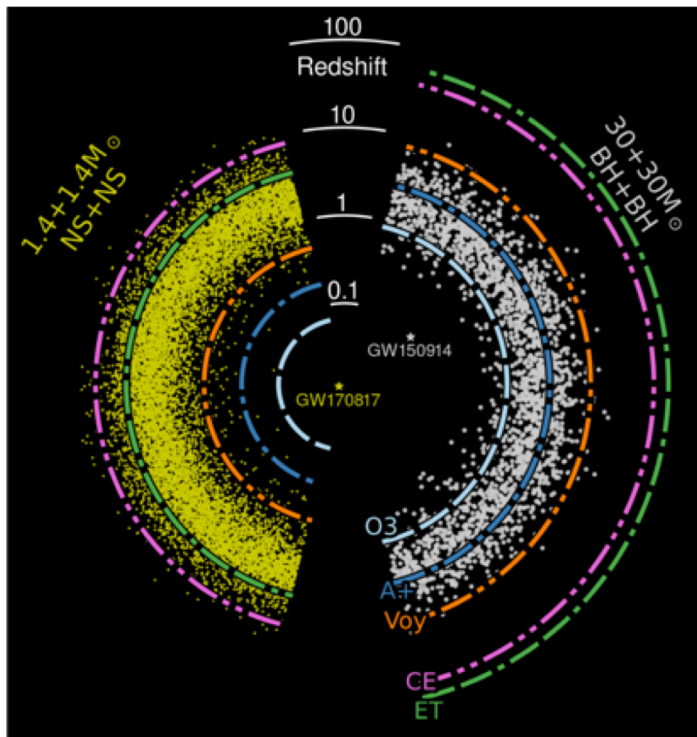
From 2G to 3G



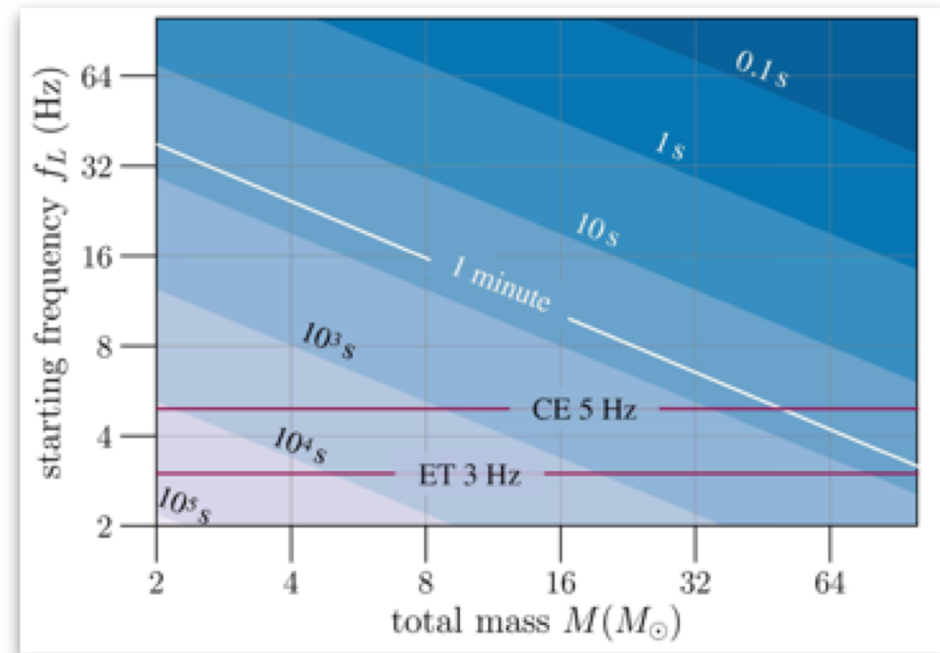
Lower frequency bound (10Hz \rightarrow 1-5Hz)

Sensitivity improvement by about x10

New challenges



May 10th 2023, ET Symposium



Credit Cosmic Explorer

Goal of Division 10

The goal of Division 10 is to be ready for the analysis and parameter estimation with 3G data, in terms of methods and software, in order to exploit the full potential of 3G detectors

- new challenges due to a better sensitivity (very high SNR, long waveforms, overlapping events, strong foreground, new sources)
- new geometry (null stream, correlated noise), new network of detectors (+CE)

Goal of Division 10

The goal of Division 10 is to be ready for the analysis and parameter estimation with 3G data, in terms of methods and software, in order to exploit the full potential of 3G detectors

- new challenges due to a better sensitivity (very high SNR, long waveforms, overlapping events, strong foreground, new sources)
- new geometry (null stream, correlated noise), new network of detectors (+CE)
- new computing technologies (GPU/TPU/FPGA, cloud computing) and algorithms (classical development, IA, Quantum algorithm)

Mock Data Challenges

Produce realistic simulated data at the output of ET (and CE), in order to:

- Train newbies in data analysis (tutorials/simple data sets)
- Stress-test the current computational infrastructure
- Develop/test adapted data analysis and parameter estimation methods
- Prepare the interpretation of the results in term of astrophysics, cosmology and fundamental physics

Previous MDSC (2012–2015)

MD&SC 1	MD&SC 2
1 month	1 year
Gaussian coloured noise (ET-B)	Gaussian coloured noise (ET-D)
Low frequency at 10 Hz	Low frequency at 5 Hz
E1, E2 and E3 collocated	Relative position of the vorticed accounted *
NS-NS (Gaussian mass distribution)	<ul style="list-style-type: none"> • NS-NS, NS-BH, BH-BH (StarTrack) • IMBHs + IMRIs (J. Gair & I. Mandel) • 1 f-modes & 2 supernova (J. Clark)
1 data set: ET MDSC1	3 data sets : <ul style="list-style-type: none"> • stellar CBC only ET MDSC2a • undetected stellar CBC ET MDSC2b • all type of sources ET MDSC2c

Results previous MSDC

First ET MSDC

- Demonstrated null stream
- Showed source overlap is not a problem at 25 Hz
- Measured chirp mass (<0.5%) and total mass (<a few percent)
- Measured non Gaussian astrophysical background (5% accuracy)

Results previous MSDC

Second ET MSDC

- CBC analysis down to 5 Hz with GSTLAL - Measured chirp mass ($<0.05\%$)
- Recovery of overlapping sources.
- Measured residual astrophysical background (non isotropic bias corrected)
- PE for IMBHB and IMRIs
- Burst analysis

Some previous work

- Detection+signature in the null stream

First MDC: <https://arxiv.org/pdf/1201.3563>

- Residual background from CBC (BBH, BHNS, BNS)

Second MDC: <https://arxiv.org/pdf/1404.1134.pdf>

- Overlapping signals (analysis with GSTLAL down to 5 Hz)

Second MDC: <https://arxiv.org/pdf/1511.01592>

- Foreground subtraction

<https://arxiv.org/pdf/2002.05365.pdf>, <https://arxiv.org/pdf/gr-qc/0511092>

- Studies on the Null stream

<https://arxiv.org/pdf/2204.08533>, <https://arxiv.org/pdf/2205.00416>

New 3G Mock Data and Science Challenges

Series of MDCs with increasing complexity containing noise + GW signals

Noise

- Different design sensitivity curves
- Correlations between IFO channels

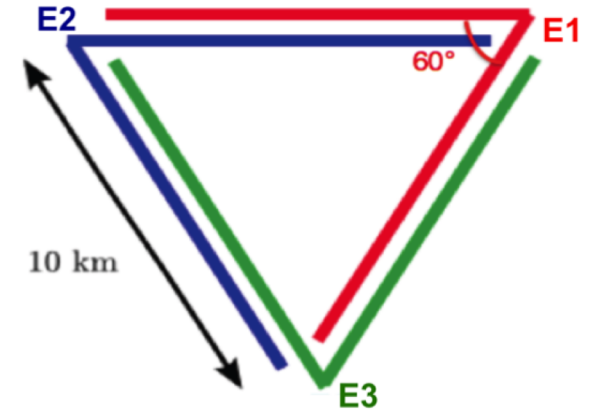
Janssens 2022:magnetic, seismic&NN

- New families of glitches
- Calibration uncertainties

Signals

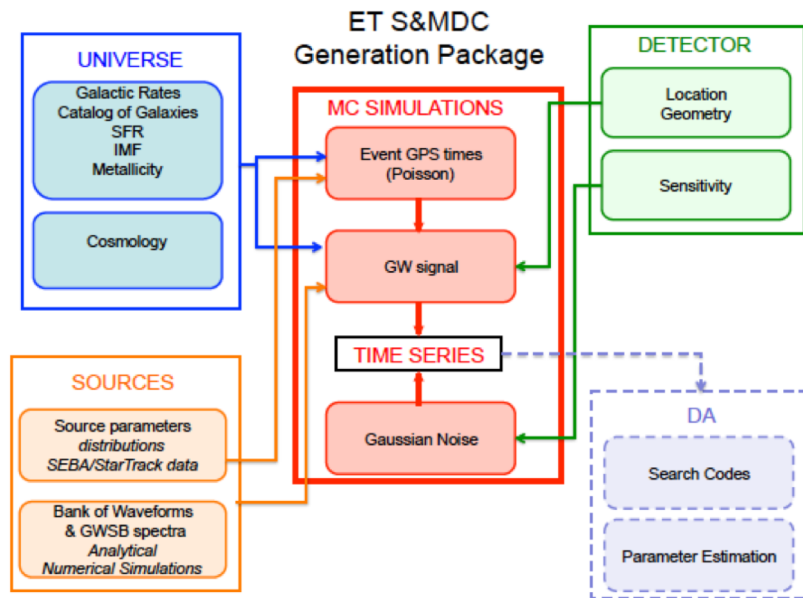
- CBCs /CWs / SNae / Stochastic
- Exotica
- Input from other divisions for specifications

First MDSC



- One month duration
- Gaussian colored noise in the triangle configuration, no glitches (+ CE)
- Cosmological population of BBHs, BNSs and NSBHs from Mapelli's group catalogs used for CoBA
- Sources with isotropic distribution in the sky/orientation

MDC_Generation



<https://dcc.ligo.org/LIGO-T1400401>

May 10th 2023, ET Symposium

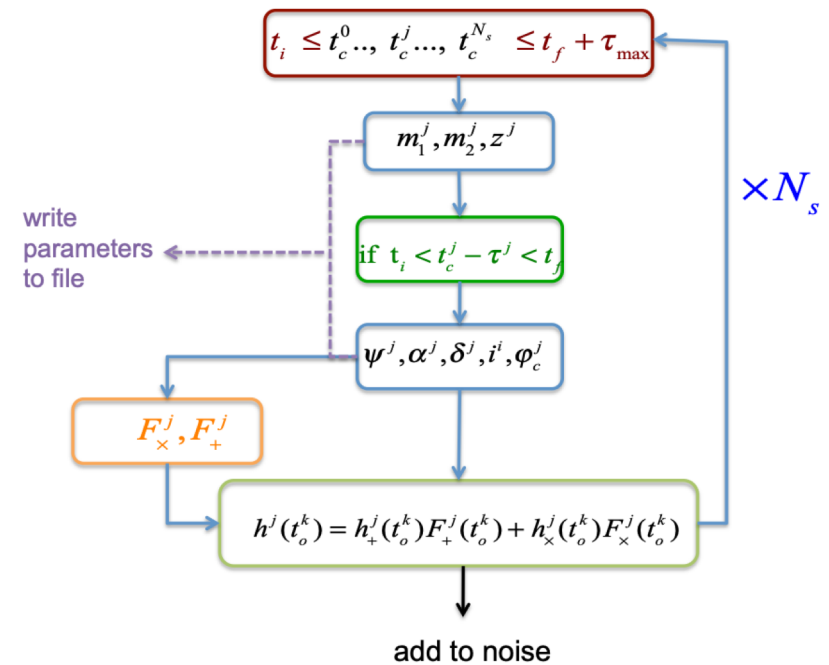


FIG. 3: Flow diagram of the Monte Carlo simulation code

MDSC1: data storage

Where is the data available?

/cvmfs/et-gw.osgstorage.org/et-gw/PUBLIC/MDC1

Instructions can be found here: https://wiki.et-gw.eu/EIB/SoftwareFrameworks/WebHome?validation_key=e2698d03b6eff5856cfab4654d3fbfe5

MDSC1: data files

What format is the data in?

In Frames for E1, E2, E3 and the null stream E0. It represents 1.3 TB in total. Future releases will have hdf5.

How long is the data set?

The data is one-month long split in 1270 segments of 2048 s, sampled at 8192 Hz. The GW signal and noise are continuous from one segment to the other.

MDSC1: noise and signal

What does the data contain?

Two sets of data, one with GW signal+noise in MDC1/E0,E1,E2,E3 and one with only noise in MDC1/Noise/E0,E1,E2,E3.

- Colored Gaussian noise with no noise artifacts and noise PSD mimicking the expected PSDs of ET 10 km
- GW signal from BNS (59540, 87%), BBH (6578, 10%) and BHNS (1977, 3%)

MDSC1: waveforms

Which waveforms were used?

- **BBH and NSBH:** IMRPhenomXPHM
- **BNS:** IMRPhenomP

with tidal effects (NRTidalv2_V) (*EOS provided by M. Oertel and T. Hinderger*)

Neutron stars have no spins.

MDSC1: informations on source parameters

Are the source parameters available?

starting_time, final_time, time@Maximum, mass1, mass2, spin1, spin1x, spin1y, spin1z, spin2, spin2x, spin2y, spin2z, lambda1, lambda2, z, dist(Mpc), ra, declination, polarization, inclination, initial phase, SNR, type

The lists of sources are available in the folder lists for each segment (cbc_\$segment.dat) and for the full data set (list_cbc.txt).

MDSC1: cosmology

What cosmological model was used?

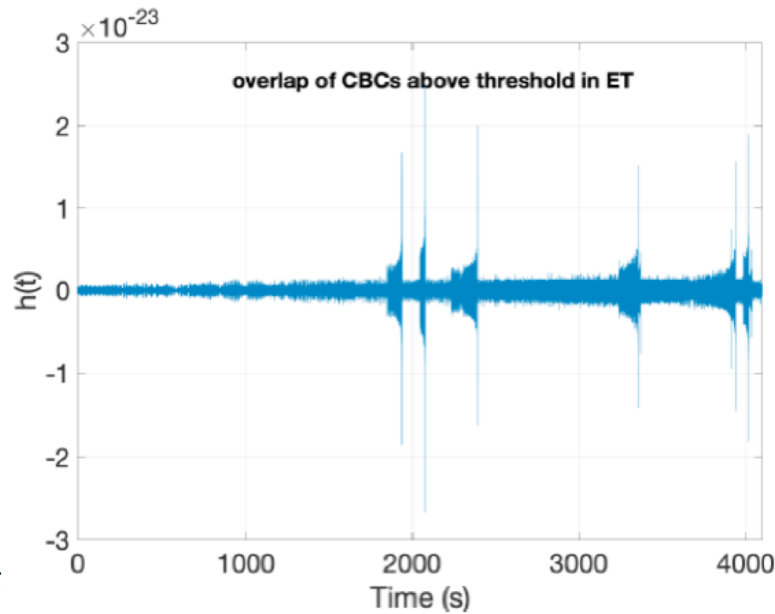
Lambda-CDM model with a total $\Omega=1$ (Planck 2018)
 $\omega_m=0.3111$, $\omega_\lambda=0.6889$ and $h_0=0.6766$

MDSC1: expectations

Detected sources

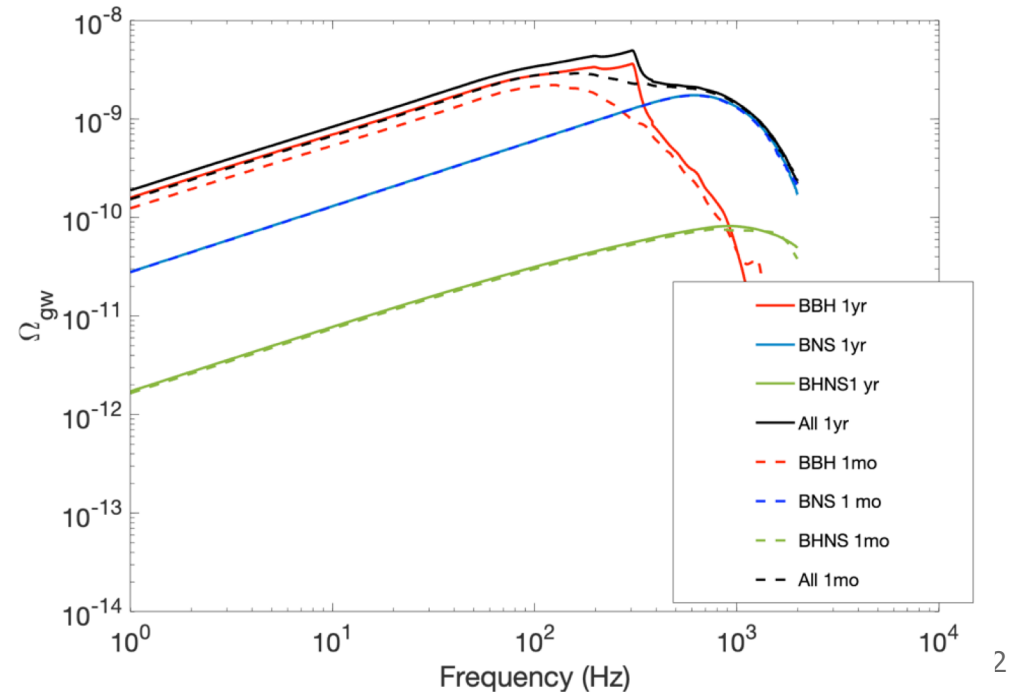
SNR>8: 13,000 BNSs , 6,100 BBHs

SNR>12: 4,800 BNSs , 5,300 BBHs



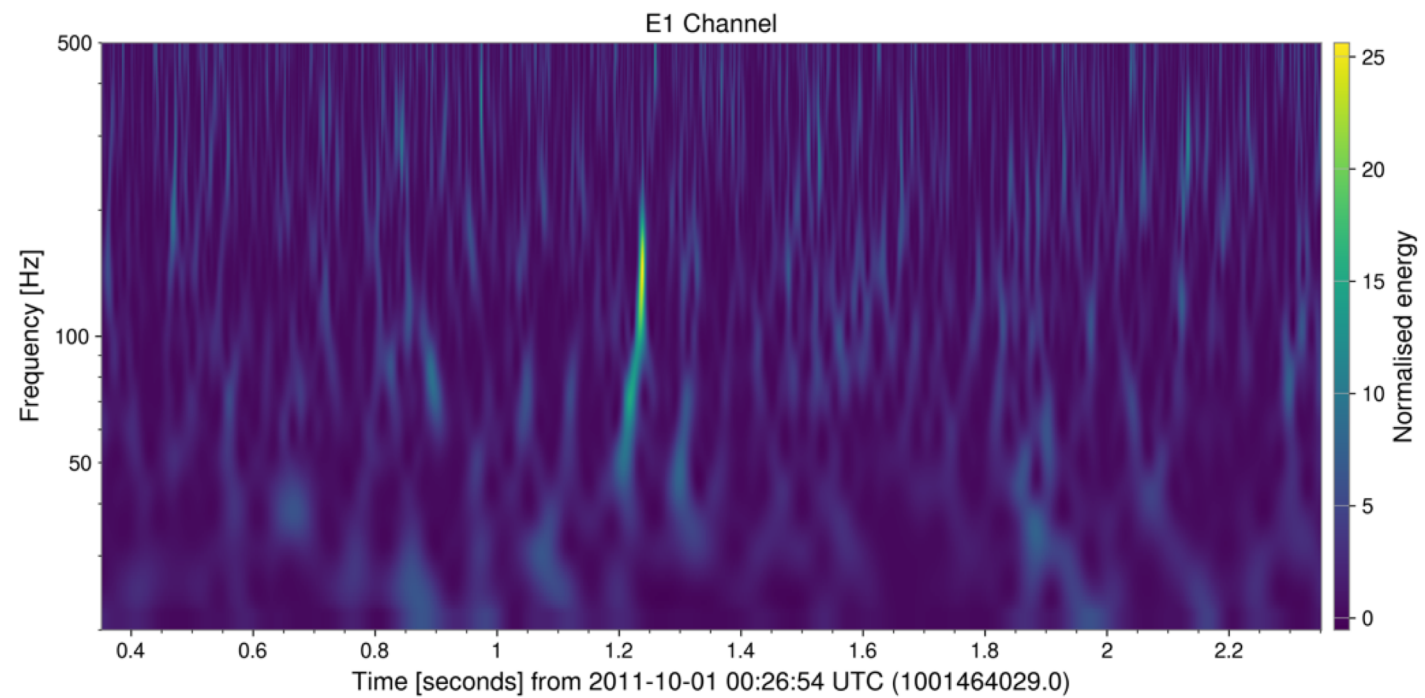
May

Background



2

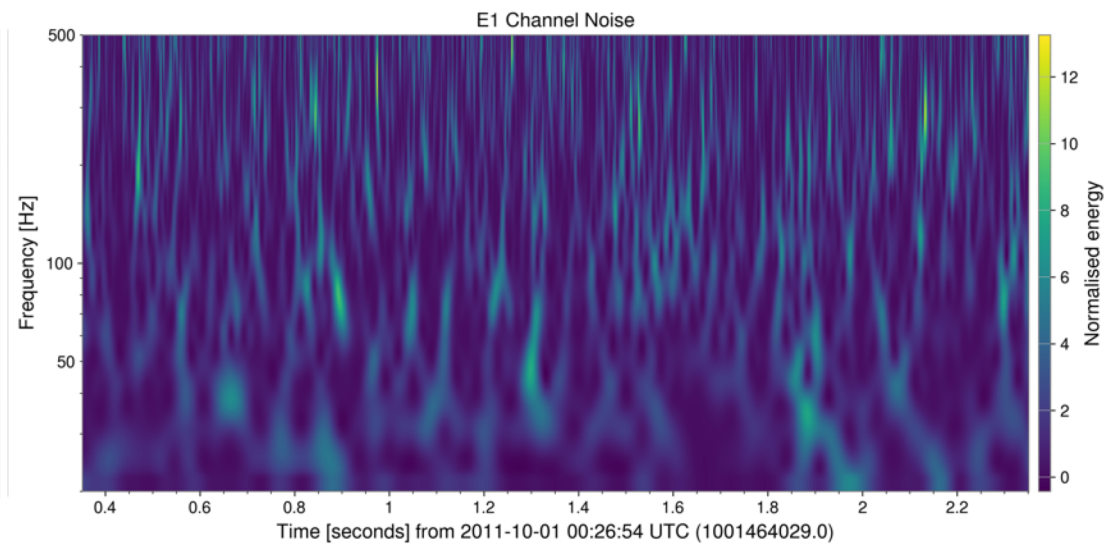
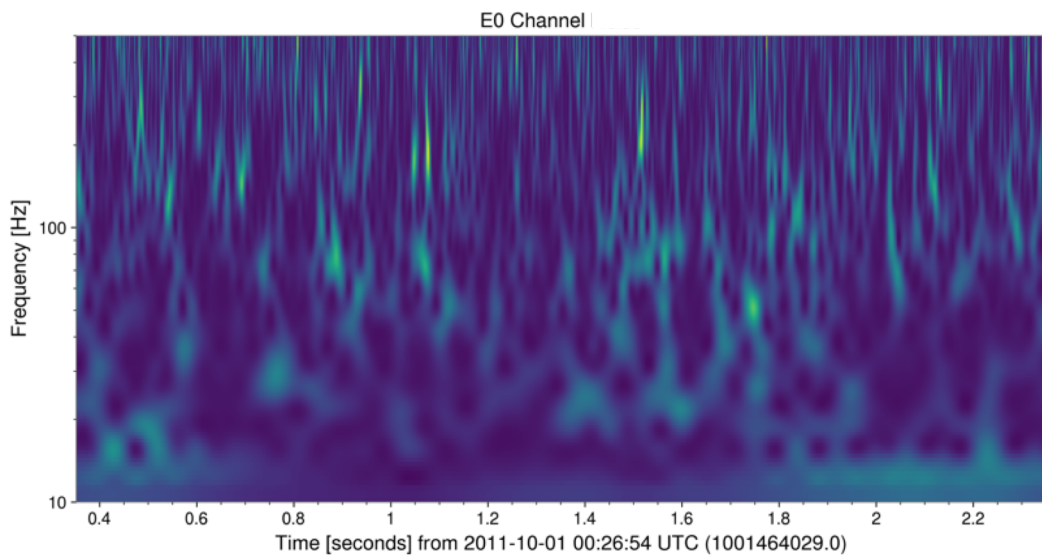
MDSC1: data



Credit Jishnu Surech

May 10th 2023, ET Symposium

MDSC1: null channel and noise



Credit Jishnu Surech

May 10th 2023, ET Symposium

MDSC1: challenges

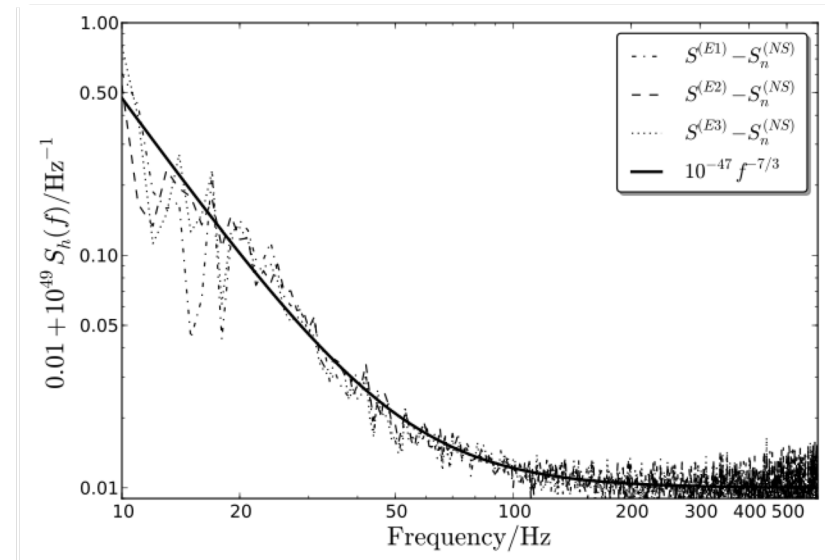
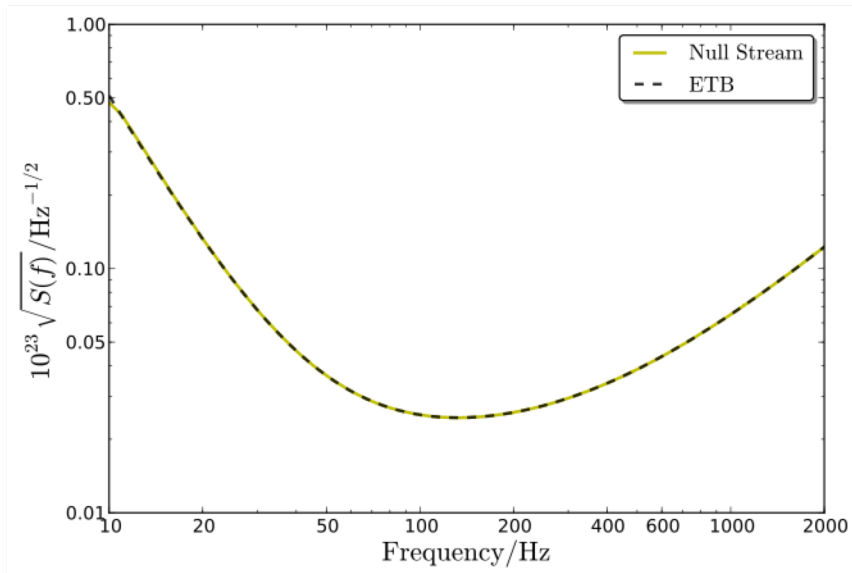
Computing

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

Data Analysis and Science

- Compute the detection efficiency
- Obtain the mass/spin distributions
- Determine the rate of each population
- Measure the neutron star equation of state
- Measure H_0
- Measure energy density of the background

Background signature



First ET MDC: <https://arxiv.org/pdf/1201.3563>

May 10th 2023, ET Symposium

Planned Tutorials

- MDC_Generation for data production
- PyCBC, BayeStar (others?) for CBCs
- PyGWB and PyStoch for stochastic
- ...

Dates will be announced through div 10 mailing list

Summary

- A lot to do to be ready for the analysis of 3G data
- Series of MDCs, first round starting now
- Series of tutorials (announced by email and in the wiki)

Mailing list: <http://mail.ego-gw.it/mailman/listinfo/et-osb-DA>

Wiki page: <https://wiki.et-gw.eu/OSB/DataAnalysisPlatform>