

Data Analysis Platform

Elena Cuoco, Tania Regimbau, Chris van den Broeck and John Veitch

ET 2nd annual meeting

Goal

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 the next generation of detectors.

New challenges due to a better sensitivity (very loud sources, long waveforms, overlapping events, new sources)

New possibilities (network configurations, computing)

Key questions of the blue book

- What are challenges posed by low-frequency, long-signal, and high-SNR observations?
- What are challenges posed by overlapping signals for detection, and for parameter estimation?
- What are the computational challenges posed by large numbers of signals? What are promising technologies for solving these problems?
- What performance metrics can we provide to guide instrument design
- What are the unique advantages of a null stream? How will real instrumental noise impact the ideal case?

Contents

1	Introduction	2
2	Goals	2
3	Simulations and MDC	2
3.1	Description of the dataset (Tommaso)	3
3.2	Tutorials (John)	3
3.3	Metrics	3
3.4	Null stream	3
3.5	Overlapping signals	3
3.6	Correlated noise	3
3.7	Source subtraction for CBCs	4
4	Innovative new methods for data analysis	4
4.1	Machine Learning Methods (Elena)	4
4.1.1	Waveform simulation	4
4.1.2	Signal detection and classification	4
4.1.3	Parameter estimation	4
5	Computational Requirements (John, ...)	4
6	New paradigm for MMA	4
7	Organization	4
7.1	WP1 Identify and frame key challenges & opportunities	4
7.2	WP2 Integration with other Divisions?	4
7.3	WP3 Production of Mock Data Challenge	5
7.4	WP4 MDC Data Analysis and Results	5
7.5	WP5 Implications and Requirements for ET	5

Mock Data Challenges

- Training simulated data
- Find out the limitations of current methods
- Encourage the community to develop new tools
- Provide a common dataset for comparison of analysis methods
- Assess the science potential with ET or XG
- Assess the requirements for computing infrastructure

First ET/CE Mock Data Challenge

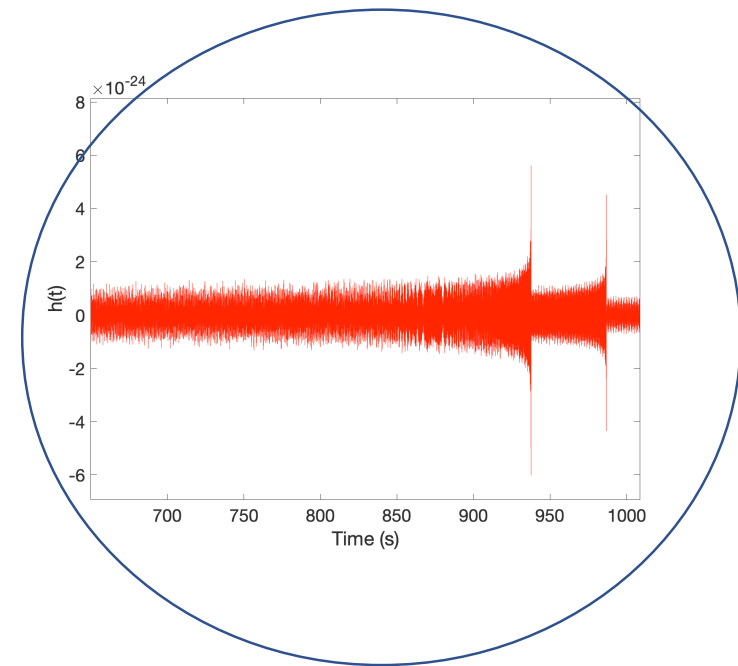
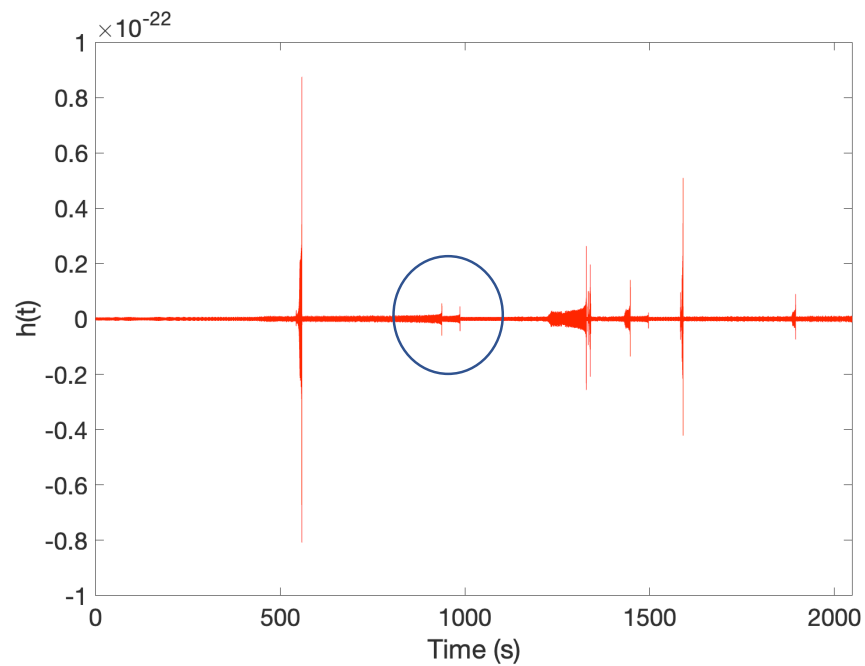
- 1 month duration
- Gaussian colored noise (ET-D 10 km, $f_{\text{Min}}=5\text{Hz}$), triangle
- Population of all type of binaries with isotropic distribution in the sky.
- Waveforms with higher modes for BBHs and tidal effects for BNS

More details in the presentation tomorrow

Example Training Data Set

28 signals with $\text{SNR} > 6$, largest at $\text{SNR} = 85$

BNSs merging at 937s and 986s, that are both long duration and overlapping

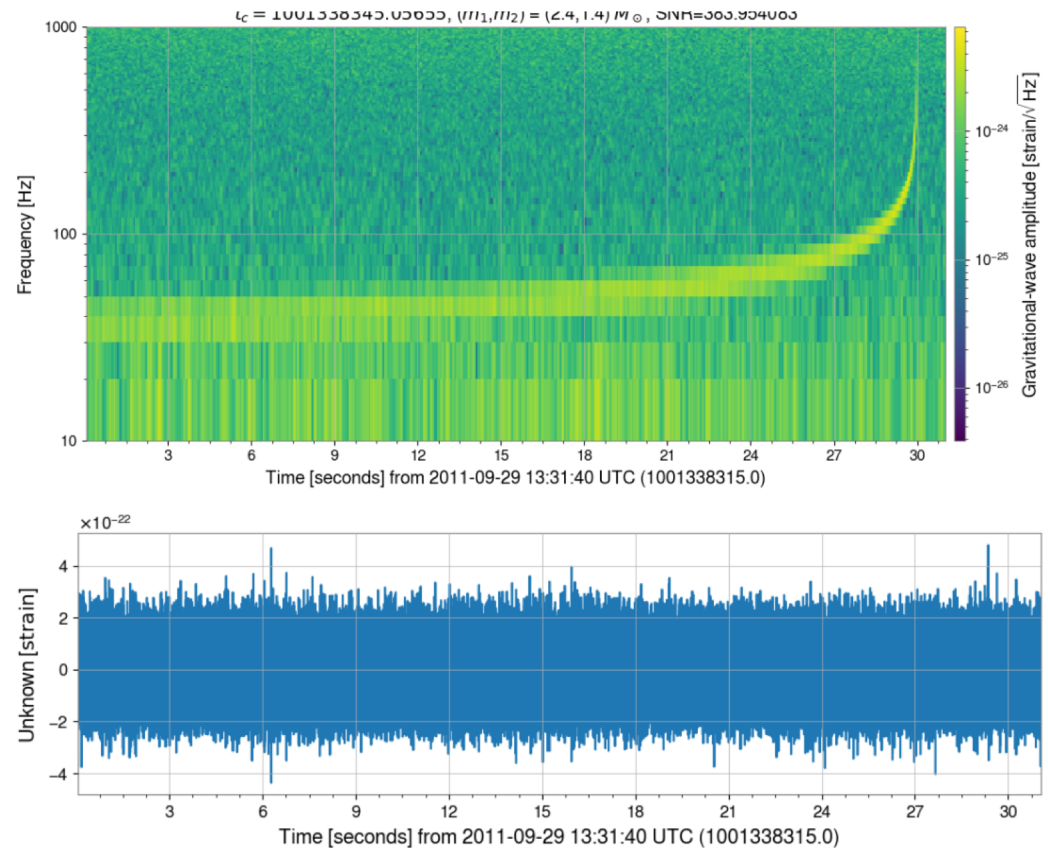


Tutorials

<https://gitlab.et-gw.eu/osb/div10/mdc-tutorial>

We have set up a tutorial for beginners to access the data and run simple scripts on the data.

We have selected 6 ultra high SNR signals with $\text{SNR} > 300$.



Innovative new methods

- 3G analysis will require the development of new techniques that go beyond what has been used in GW astronomy to date.
- These developments may be evolutions of existing methods, or introduction of entirely new techniques.
- Machine learning (ML) and Deep Learning (DL) techniques, which in recent years have changed the way data analyzed (waveform classification, signal detection and classification, parameter estimation)

Blue book agenda

- **WP1 Identify and frame key challenges & opportunities**

Deadline Sept 2021 (1st workshop)

- **WP2 Integration with other Divisions**

Identify overlap and build connections

- **WP3 Production of Mock Data Challenge**

Design (with other divisions), implement required software, release data

- **WP4 MDC Data Analysis and Results**

Using existing and prototype pipelines.

- **WP5 Implications and Requirements for ET**

In context of MDC, what are the required future developments for ET?

What is mature enough already? What approaches are most promising?