The ET Science and Mock Data Challenge Generation Package

(Dated: February 15, 2012)

The purpose of the ET-SaMDC Generation Package is to produce simulated data to prepare the data analysis and science with the future Einstein Telescope. The end product are the time series of the 3 V-shaped detector outputs E1, E2 and E3 Fig. 1, arising both from the instrumental noise and from the GW signal. We propose to follow the procedure developed for the first ET-SaMDC [1], extending it to a large range of sources and models. In this document, we first describe quickly the structure of the ET-SaMDC Generation Package, then we give a first list of the different tasks, and people who already volunteered to participate.

I. BACKGROUND AND INTRODUCTION

Encouraged by the success of the first ET mock data challege [1], we would like to work on expanding the scope of the challenge so as to include a greater variety of sources as well as longer data train to gain a better understanding of the difficulties involved with analyzing that contain extremely long duration signals and occasional transients. First of all, we would like these to be called Mock SCIENCE Challenges rather than Mock DATA Challenges. Our goal should be to

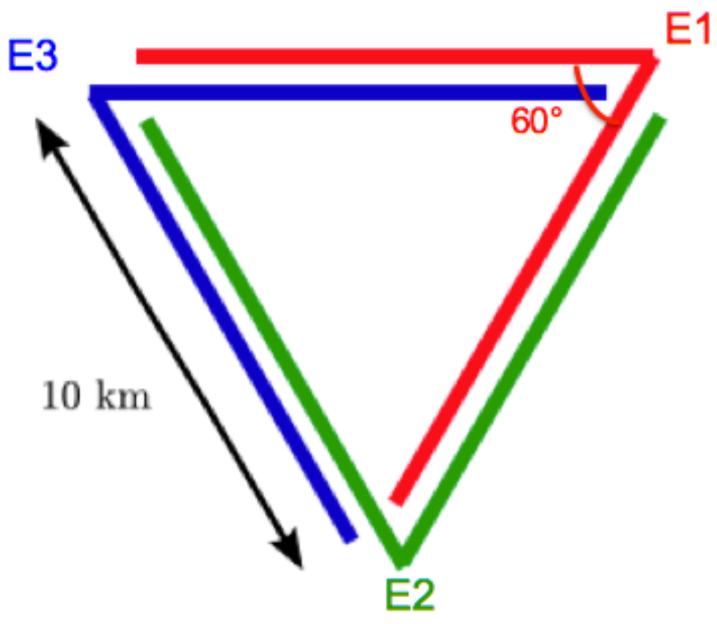

FIG. 1: The 3 ET detectors, in the triangle configuration

extract the best possible science out of the mock exercise and show the extent to which we can reconstruct the various "input" parameters. Input should be interpreted in the broadest possible term. It could represent ET PSD, waveforms, source parameters, source rate, cosmology, general relativity or any assumption we make in producing the mock data. Each of our next set of challenges could focus on one or more aspects of astrophysics, fundamental physics or cosmology. Here is what we suggest for the next round.

A. What will the Data contain

1. *Nature of the background noise:* Gaussian noise coloured with ET noise PSD. The first MDC used ET-B. In the second we should try and use ET-D, as it is the most recent design curve I. However, ET-D, being the quadrature sum of two different sensitivity curves, is somewhat complicated. We should try and develop a fit for each of the components and use an analytical formula that is the quadrature sum of two different sensitivity curves.

   Action: *We need someone to take on the task of producing an analytical fit to ET-D. Is Stefan Hild able to do this?*

2. *Lower frequency cutoff:* We should definitely aim for 5Hz, 3Hz might be possible, 1 Hz could be challenging. Assess the requirement after a careful study of required computational infrastructure.

   Action: *We need a quantitative evaluation of which lower frequency would be doable for this mock data.*

3. *Length of data:* First MDC used a month-long data set. To test certain claims about what is possible with ET (e.g. measuring the binary coalescence rate as a function of red-shift, usefulness of compact binaries for cosmography, etc.) it would be necessary to have a longer duration data. 1 year should be ample for most purposes. However, there might be merit in also producing a month-long data set sampled at a higher rate (see below).

4. *Sampling rate:* It has been suggested that we should generate two different data sets, one sampled at 4096 Hz, which could be used for most of the analysis, and a second sampled 8192 Hz, for injecting neutron star normal modes. 4 kHz is thought to be a bit too low in this case.

5. *Detectors:* We will use the standard ET topology of three IFO outputs, ET1, ET2, ET3 - all at same position, that of Virgo detector. It has been suggested that we could produce six outputs, three for the high-frequency detector and three for the low-frequency detector. However, after some discussion it was concluded that combining data from a xylophone while conceptually straight forward is not something about which we have any experience. Perhaps this could be investigated outside of the MDC activity and the results could benefit future ET MDCs.

6. *Modulation of the antenna pattern:* The model used to modulate the detector antenna pattern should be made public.

   Action: *Tania Regimbau should write up a short document and a code describing the model that this used in producing the modulation. This will be needed by analysts who might want to "subtract" the effect of the modulation while drawing astrophysical inferences from the observed signals.*

7. *Misc:* Do we add artificial glitches? Should the single-ifo glitches be independent or should we consider having correlated noise? How could we model that?

   Action: *Decide if this is useful and identify people who can work on the required infrastructure.*

8. CBC Injections

   (a) *Binary neutron stars:* BNS signals should be injected with masses in a definite range but unknown to participants. The rate and the model for its evolution will not be known to everyone. The injection will use a specific distribution for injected intrinsic masses but these will not be known to analysts. Neutron star spins will be zero.

   Action: *Which waveform family should we use? TaylorT4?*

   (b) *BH-NS signals:* As above but BH will have a non-zero spin, chosen randomly (a flat distribution) in the range 0 to 1.

   Comment: *Is it really useful to have spins at this stage. We don't have a good analysis pipeline to deal with spins. The presence of systems with significant spins could seriously bias some of the conclusions. This bias might not be real in the sense that with a mature pipeline that can handle spinning systems it might not lead to such a bias. We should discuss this and drop spin from our consideration if we are not comfortable.*

   Action: *Which waveform family should we use? SpinTaylor or some other?*

(c) BBH signals: As above but both black holes could have spins (but see cautionary remark above).

Action: *Which waveform family should we use? EOBv2?*

(d) IMBBH signals: We should definitely have IMBBH signals. ET could map the IMBBH content of the Unvierse out to redshift of 5 to 6.

(e) *EMRI signals:*

Action: *Decide if we are going to include these.*

9. *A rare supernova signal:* ET is expected to observe a supernova once every two to five years. It would be have one SN signal over the entire duration of the data.

10. *A stochastic primordial background:* nearly flat power-law (cosmic string or inflation). The interest would be in disentangling this from the astrophysical (CBC) stochastic BG. Simply injecting at one amplitude is may not be very flexible. Could allow analysts to vary the amplitude and spectral index and do their own simulations.

11. *CW signals:* Isolated pulsars could be very useful but is there any expertise and interest in doing this within the MDC team?

    Action: *Should we consider injecting isolated pulsars? Is the analysis/science fundamentally different from CW searches in (a)LIGO data? Who might be intrested in producing and analysing CW signals?*

12. *Cosmology:* A flat Universe with $H_0 = 70\,\mathrm{km\,s^{-1}Mpc^{-1}}$, $\omega_\Lambda = 0.73$, $\omega_\mathrm{M} = 0.27$ and $w = -1$.

B. What are the Science Challenges for this run?

This is currently highly incomplete. Please improve it by providing your own ideas.

1. Determine the rate of BNS, NS-BH and BBH as a function of luminosity distance

2. Determine the cosmological parameters to the extent possible

3. Determine the mass function of neutron stars and black holes

4. Determine the presence/absence and parameters of a primordial power-law stochastic background, or determine the best sensitivity of a search for primordial BG in the presence of astrophysical BG.

5. . CW-related challenge?

C. What publications do we envisage for ourselves and for other groups?

Once the data are validated, release them and allow external groups to apply new techniques and write their own publications At least one general paper presenting the data as a whole.

II. MDC CODE STRUCTURE

The schematic structure of ET-SaMDC Generation Package is presented in Fig. 2. It contains different modules which are shortly described below.

1. Monte-Carlo Simulations : The core of the program which simulates the time series, is shown inside the central red box. This module has been implemented for the first ET-SaMDC [1] and is not expected to change significantly between the first and the second ET-SaMDC. As described in [1], the sources will be realistically distributed in the parameter space but also in time, assuming a Poisson distribution of the arrival times, and redshift, accounting for the star formation history of the Universe (block Universe). The noise (colored gaussian noise) will be generated using models of the ET sensitivity (block Detector), and considering a lower bound between $\sim 1 - 5$ Hz.

2. Universe : This module will contain information on the Universe, local or at cosmological distances : cosmological parameters, star formation history, initial mass function, metallicity, catalog of galaxies or galactic rates. All these informations are needed to calculate the source rates (and then the average time interval between events), the distribution of the redshift, but also the distance for high redshift sources. Galaxy catalogs may not be used for the second ET-SaMDC, but we should consider including it in the future.

3. Sources : This module is the one that needs the most to be developed. It will contain distributions of the source parameters (masses, spins, delays...), needed to simulate realistic populations of sources. These distributions can be analytical models or ascii files produced by sophisticated star evolution codes such as StarTrack [2] or SEBA [3]. It will also contain banks of waveforms that will be used to generate the GW signal. These waveforms could be taken from analytical calculations or numerical simulations. Also we plan to build a catalog of

spectra for GW stochastic backgrounds [4] which will be generated in the frequency domain, and Fourier transformed to the time domain.

4. Detector : This module contains all the informations related to the detector, geometry, location, response tensors needed to calculate the beaming factor, and various models for the sensitivity and lower frequency bound.

The simulated time series will be later analyzed to produce lists of detected events (block analysis block, search codes). The results of this analysis combined with parameter estimation and informations on the universe and the sources, could then be interpreted in term of astrophysics, cosmology and fundamental physics.

III. TASK LIST

This is a list of the different tasks needed to develop the four modules described above, with the people who have volunteered to work on them.

A. UNIVERSE

We already have a set of SFR and cosmological models. The Rome Group will provide a star formation history derived from a cosmological hydrodynamical simulation, which well reproduces the observational data at $z < 6 - 7$.

B. SOURCES

1. Stellar compact binaries :

a. 1. parameter distributions: At the moment we have simple analytical distributions for the masses and the delay between formation of the progenitor and merger. Different groups propose to provide data files from evolutionary codes:

1. StarTrack : The Warsaw group (Belczynsky, Bulik, Dominik, Kowalska) has developed a state of the art model of binary evolution and DCO formation, utilizing the most recent results in simulations of supernovae explosions, realistic treatments of common envelope phases. A grid of eleven metallicity values ranging from 0.0001 to 0.03 (with solar being 0.02) will be calculated for this model in order to cover chemical properties of Pop II and III

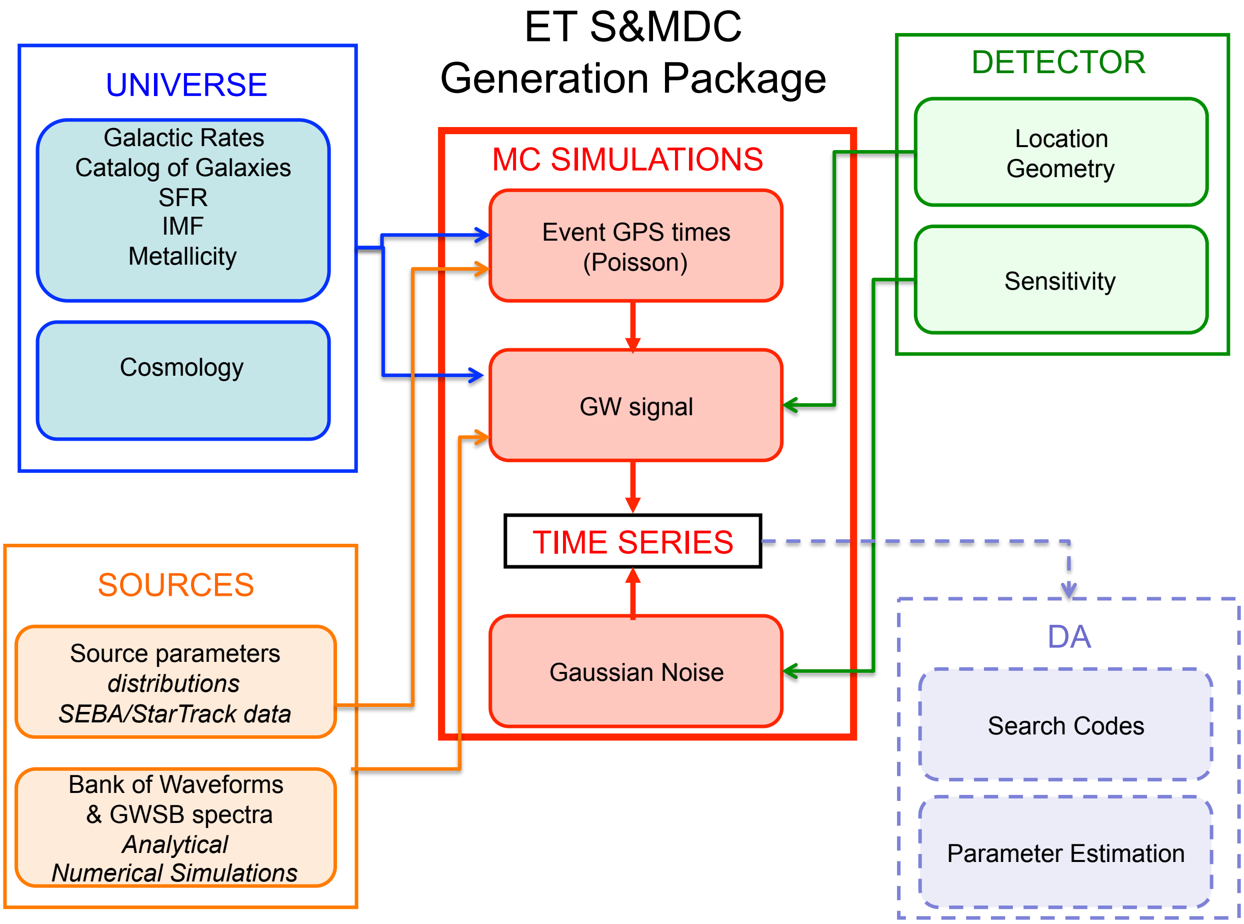

FIG. 2: Structure of the ET Generation package. The different modules are independent and can be developed separately. To illustrate, we also included the module Data Analysis, which is not part of the generation package but uses the simulated time series.

stars in the Universe. The star formation rate, metallicity and galaxy mass as functions of redshift will be provided, allowing for calculations of distance-specific Galactic merger rates of DCOs in the Universe. These results translated into detection rates will meet ET's potential to probe distant regions of space. Compared with previous studies, which investigated only properties of DCOs in the "local" Universe, utilized simplified cosmology or used unphysical models of binary evolution, this adds new quality to the subject.

2. SeBa (Rome group): The Rome group (Corvino, Ferrari, Marassi, Maselli, Schneider) can provide a set of compact binary populations generated through the SeBa population

synthesis code. These have been produced exploring a wide range of physical parameters, such as the core mass threshold for BH formation, the kick velocity distribution and common-envelope parameters. The Rome group has already explored how these substantially modify the predicted (BH-BH) merger time distribution and coalescence rate in http://adsabs.harvard.edu/abs/2011PhRvD..84l4037M (see section VI-B). These or additional models could be injected in the ET data stream using the Monte Carlo simulation code.

b. 2. waveforms: The Rome group will provide analytical waveforms for BNS and BH-NS systems up to the merger, based on TaylorTn approximants for aligned spinning bodies, including finite size effects at the leading and next to leading order [5]. These waveforms can be provided for any NS equation of state available in the literature. In the short term the group will work on how to improve the contribution of the NS tidal deformation to the orbital motion, going beyond the adiabatic approximation, and model the Love number as a suitable function of the orbital distance and will be able to provide the corresponding waveforms ($\sim$ 2 months). At present the tidal corrections to the orbital phase are known to 1PN order. The group plans to calculate next order corrections (NNL) and release the corresponding waveforms ($\sim$ 4 months)

Ilya and Stas also suggested they could provide models of BBH including spin.

2. Intermediate mass BH coalescences

Jon, Stas, Sofiane and Ilya will provide waveforms and parameters

3. Pulsar Glitches and magnetars

James will provide waveforms

4. Supernova

Siong will provide waveforms

5. Sochastic Background

Stefanos and Tania will work on simulations of Gaussian primordial GWSB in the frequency domain, or in the time domain for GWSB from cosmic strigs and astrophysical origin.

C. DETECTOR

Stefan will provide analytical fits of the ET D sensitivity.

(to be completed...)

---

[1] *A Mock Data Challenge for the Einstein Gravitational-Wave Telescope* Regimbau T., Dent T. and ET Science and MDC team, http://arxiv.org/abs/1201.3563

[2] *A Comprehen- sive Study of Binary Compact Objects as Gravitational Wave Sources: Evolutionary Channels*, K. Belczynski, V. Kalogera, and T. Bulik, , Rates, and Phys- ical Properties, Astrop. J. 572, 407 (2002);

[3] http://www.sns.ias.edu/∼starlab/

[4] http://homepages.spa.umn.edu/∼cwu/

[5] J. Vines, E.E. Flanagan, T. Hinderer, Phys. Rev. **D83**, 084051 (2011)