Overlapping signal detection using cWB pipeline

GRAvitational-wave Science & technology Symposium, 6-7th June 2022, Padova, Italy

Andrea Virtuoso [1 2], Philip Relton, Sophie Bini, Marco Drago, Ian Harry, Claudia Lazzaro, Andrea Miani, Shubhanshu Tiwari, Vivien Raymond

[1]Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy

[2]INFN, Sezione di Trieste, I-34127 Trieste, Italy

Contact: andrea.virtuoso@ts.infn.it

Why do we have to take care about overlapping signals

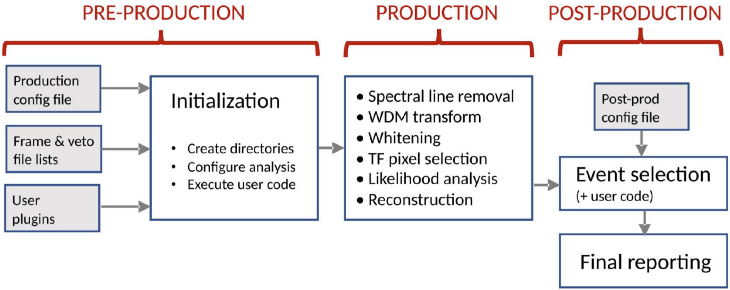
- in the past observing runs the possibility in which the interferometers detect two different signals at the same time (i.e. overlapping signals) wasn't relevant at all
- also for O4 the expected detection rate gives a very unlikely chance to have overlapping signals ($9.0^{+5.6}_{-3.2} \times 10^{-6}$)[1]
- instead, with the expected rate for the next generation of detectors (e.g. Einstein Telescope) the overlap between different signals will be very common

[1]Relton, Philip, and Raymond, Vivien. "Parameter estimation bias from overlapping binary black hole events in second generation interferometers." *Physical Review D* 104.8 (2021): 084039.

Why do we have to take care about overlapping signals

- $\Rightarrow$ we need to understand how the current pipelines behave in analyzing data with (at least) two different signals
- how do overlapping signals affect detection efficiency? Which kind of biases in parameter estimation?
- a study is going on both modeled (PyCBC) and unmodeled (coherent WaveBurst) pipelines; in this presentation we'll focus on coherent WaveBurst analysis

cWB in a nutshell



In coherent Wave-Burst (cWB) data are decomposed into wavelets in time-frequency (TF) maps made of pixels, each of these labeled with the index $[i]$. Whitening is carried out pixel by pixel using the averaged pixel mean squared fluctuation, detector by detector

$$\boldsymbol{w}[i] = \mathbf{f}_{+}(\theta, \phi)[i] h_{+}[i] + \mathbf{f}_{\times}(\theta, \phi)[i] h_{\times}[i] + \boldsymbol{n}[i] \tag{1}$$

cWB in a nutshell

The main step in the analysis is the maximization of a log-likelihood ratio ("referred as "the likelihood")

$$\ln \frac{p(\boldsymbol{w}|S)}{p(\boldsymbol{w}|B)} = \boldsymbol{w}^2 - (\boldsymbol{w} - \mathbf{f}_+ h^r_+ - \mathbf{f}_\times h^r_\times)^2 \equiv \boldsymbol{w}^2 - (\boldsymbol{w} - \boldsymbol{s})^2 \qquad (2)$$

The likelihood is maximized finding the correct source location (θ, ϕ)

$$\mathcal{L} = (\mathcal{P}(\theta, \phi)\boldsymbol{w})^2 = \boldsymbol{w}^\dagger \mathcal{P}(\theta, \phi)\boldsymbol{w} \qquad (3)$$

where the projector has been defined as

$$\mathcal{P}_{nm}(\theta, \phi) = e_{n+}(\theta, \phi)e_{m+}(\theta, \phi) + e_{n\times}(\theta, \phi)e_{m\times}(\theta, \phi) \qquad (4)$$

with $\boldsymbol{e}_{+,\times}(\theta, \phi) = \mathbf{f}_{+,\times}(\theta, \phi)/\left|\mathbf{f}_{+,\times}(\theta, \phi)\right|$

$\Rightarrow$ **take home message:** the coherence is strictly related to the finding of the correct antenna patterns for the reconstructed detector response, i.e to the finding of the correct source location

cWB in a nutshell

Starting from the maximum likelihood let's define the following statistics

- the coherent energy

$$E_c = \sum_{n \neq m} w_n^* \mathcal{P}_{nm} w_m \tag{5}$$

- the residual noise energy

$$E_n = (\boldsymbol{w} - \mathcal{P}\boldsymbol{w})^2 = \boldsymbol{w}^\dagger (I - \mathcal{P}) \boldsymbol{w} \tag{6}$$

- the correlation coefficient

$$cc = \frac{|E_c|}{|E_c| + E_n} \tag{7}$$

- the penalty

$$\chi^2 = E_n / N_{df} \tag{8}$$

  where N_{df} is the number of degrees of freedom, i.e. the number of pixels used for describing the detected event.

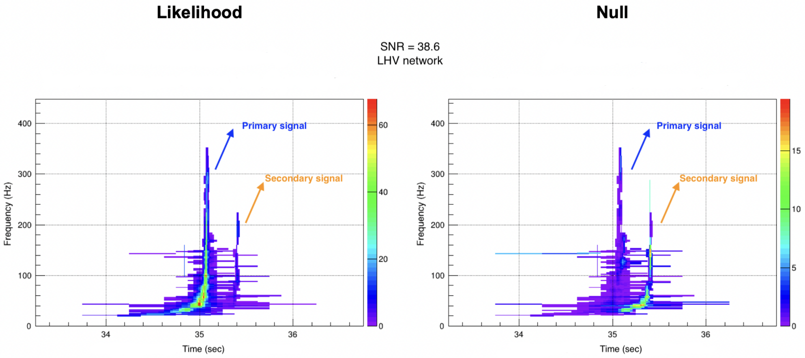
cWB post-production cuts

- the latter statistics are used in cWB post-production analysis to recover triggers which could be given by noise transients rather than GWs
- indeed, non Gaussian noise transients (usually referred as glitches) are not coherent between different detectors, so a lot of their energy goes in the null (E_n): this leads to a lower value of correlation coefficient and an higher value of penalty
- in post production analysis some cuts are applied: e.g. during O3 triggers were selected if $cc > 0.7$ and $\log_{10}(\chi^2) < 0.2$

Coherent statistics behaviour with overlapping signals

- let's suppose that cWB is triggered by two overlapping CBC signals
- cWB then maximizes the likelihood (i.e. find the correct source location) with respect to the "dominant" (let's call it "primary") signal
- in that way the "not dominant" (let's call it "secondary") one is not necessarily seen as coherent between different detectors because the likelihood hasn't been maximized with respect to his own source location
- so in some cases a fraction of the energy of the secondary signal could go to the null, lowering the correlation coefficient and increasing the penalty
- this could lead to a penalisation of the trigger in post-production analysis
  $\Rightarrow$ how much is this effect relevant in cWB analysis? (spoiler: this affects the detection efficiency, but not in a drastic way)

Likelihood and null energy time-frequency representation



As above mentioned, the secondary signal is not coherent with respect to the sky position found maximizing the likelihood for the primary signal, so the null energy for the secondary signal is really high.

Analysis setup

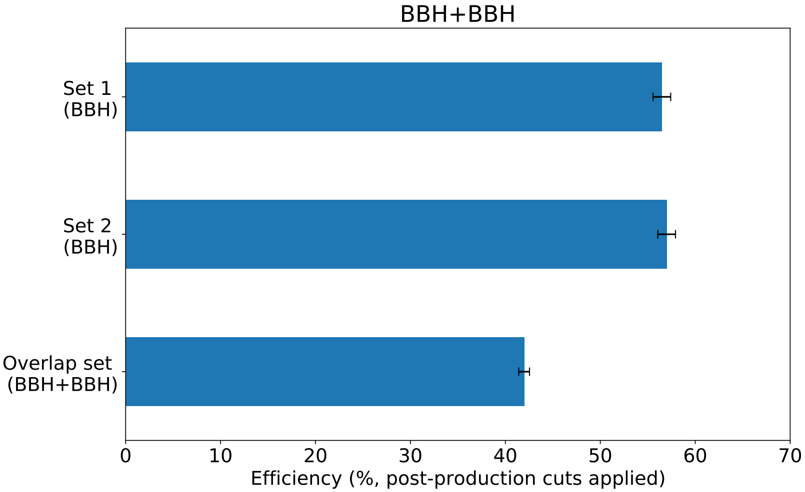
To better understand the behaviour of cWB with overlapping signals we perform an analysis run as follows

- we consider two sets of BBH waveforms ("Set 1" and "Set 2") uniformly distributed in the sky with a wide range of SNRs (from 1 to 100)
- the two sets are separately injected on the same three detectors network data stream made by a Gaussian noise generated using the O5 power spectral density
- these two data streams are separately analyzed with cWB, then the above mentioned post-production cuts are applied ($cc > 0.7$ and $\log_{10}(\chi^2) < 0.2$) and the detection efficiency is found, i.e. the number of detected triggers over the number of injections

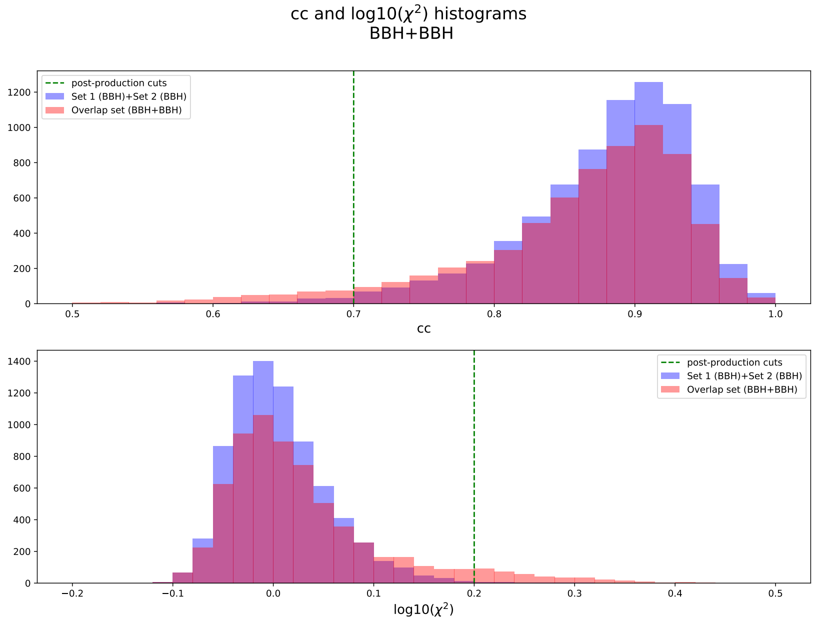
Analysis setup

- then the sets are injected on the same data stream so that each signal of "Set 1" overlaps with a signal of "Set 2"
- this "Overlap set" is analysed by cWB, applying also in this case the post-production cuts, and then the detection efficiency is evaluated
- the same analysis is performed injecting two sets of overlapping BNSs and a set of BBHs overlapping with BNSs

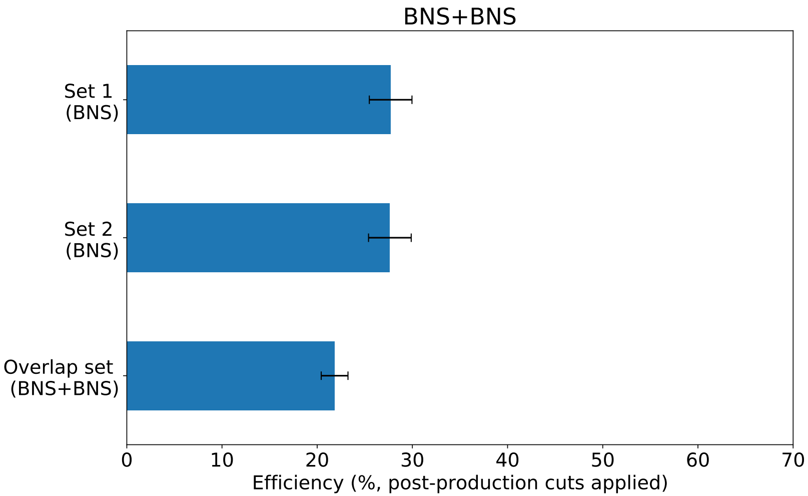
BBH-BBH analysis results



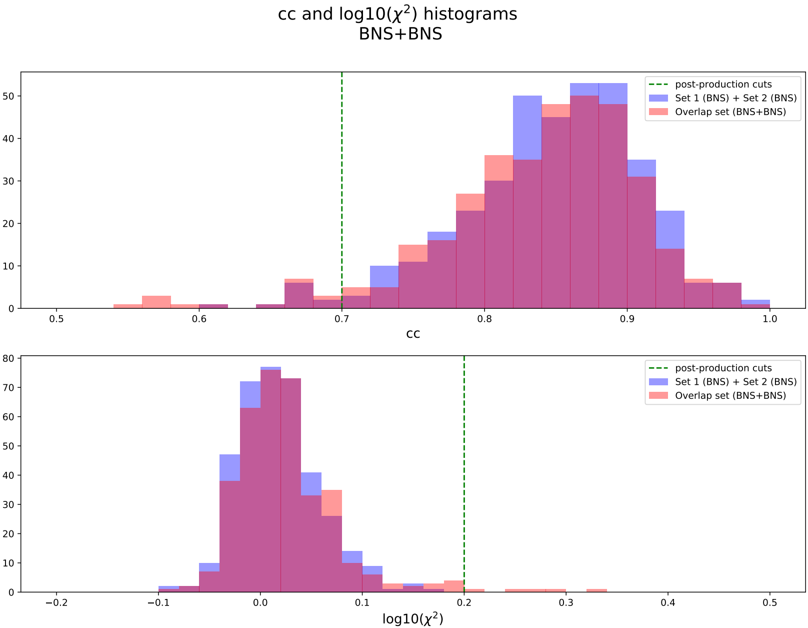
BBH-BBH analysis results



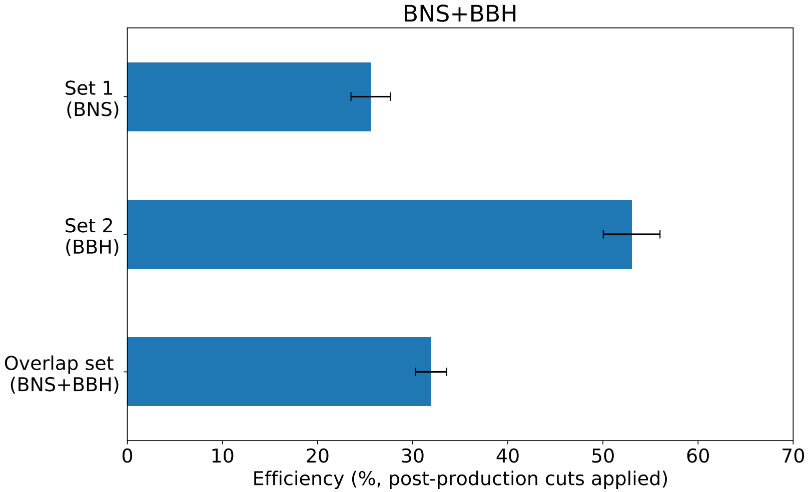
BNS-BNS analysis results



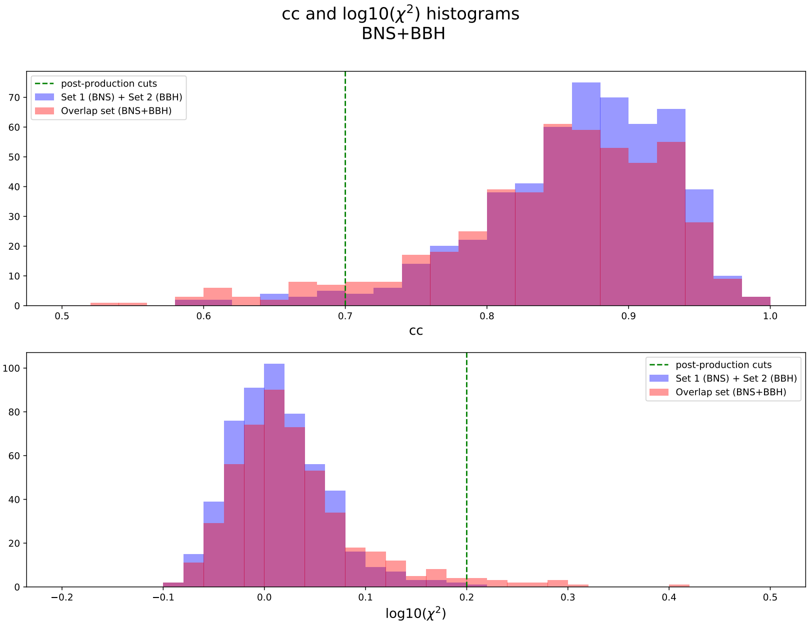
BNS-BNS analysis results



BNS-BBH analysis results



BNS-BBH analysis results



Conclusions

- comparing "Set 1" and "Set 2" with the "Overlap set", cWB seems to have a not negligible loss in detection efficiency, even if it is still able to analyze a relevant fraction of signals
- further studies will test the pipeline with a wider range of boundary conditions, resulting in overlapping signals which could interfere from a negligible to a significant way (e.g. varying the merger time separation, the SNR of a binary with respect to the other one, the chirp mass etc.)
- a paper comparing cWB and PyCBC search results is in preparation
- thanks to these studies we can start to think how to set the pipeline for being able to analyze more than one signal at the same time (e.g. signals are often clearly disentangled in TF representation for most of their duration)