UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE 182

Institut Pluridisciplinaire Hubert Curien

THÈSE présentée par :

Vincent JUSTE

Soutenue le : **06 Octobre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Physique des particules élémentaires

Search for compact binary coalescence using a single gravitational waves detector with the MBTA data analysis pipeline

THÈSE dirigée par :

M. Benoît MOURS	Directeur de recherche, IPHC Strasbourg

RAPPORTEURS :

M. Damien DORNIC	Chargé de recherche, CPPM Marseille
M. Sébastien VIRET	Chargé de recherche, IP2I Lyon

AUTRES MEMBRES DU JURY :

M. Dominique AUBERT	Professeur des universités, ObAS de Strasbourg
Mme. Frédérique MARION	Directrice de recherche, LAPP Annecy

Remerciements

J'aimerais tout d'abord remercier Benoît qui a supervisé mes trois années de thèse. Merci d'avoir été aussi disponible et pédagogique, j'ai pris beaucoup de plaisir à effectuer ce travail à tes côtés.

Je voudrais également remercier les membres de l'équipe OGMA à l'IPHC qui m'ont accompagné. Merci à Thierry, Dimitri, Antoine, Thomas et Florian.

Merci aussi aux membres de l'équipe MBTA dispersés dans les différents laboratoires du LAPP à Annecy, de l'IP2I à Lyon et de l'INFN à Urbino avec qui j'ai eu le plaisir de travailler et de voyager pour des conférences tout au long de ma thèse.

Un très grand merci à toute ma famille, parents, petit frère, grand-parents, tantes, oncles, cousins et cousines qui m'ont toujours soutenu et encouragé, grâce à qui j'ai pu effectuer ces longues études.

Je remercie bien sur également tous mes amis. Merci aux amis de longue date restés dans le sud de m'avoir encouragé, particulièrement Dylan et Jean, j'espère vous voir plus souvent dès que possible ! Merci à tous les amis que j'ai rencontrés à Strasbourg, j'ai passé de très belles années ici et c'est en grande partie grâce à vous ! Un merci particulier à Amandine, Gaëtan, Lucas, Mario et Vincent avec qui j'ai passé le plus clair de ma vie étudiante à Strasbourg; à Aurélie, Kenza et Salomé pour leur bonne humeur, les quizzs et les soirées; aux doctorants de l'IPHC pour les pauses cafés et les parties de ping-pong; à Adrien, Lucas, Marwan avec qui j'ai fait mes débuts en parkour ainsi que tous les traceurs (bien trop nombreux pour les nommer) avec qui j'ai eu le plaisir de bouger ou de boire !

Enfin, un grand merci aux membres du jury. Merci à Dominique Aubert d'avoir pris le rôle de président du jury, merci à Frédérique Marion, Sébastien Viret et Damien Dornic d'avoir fait le déplacement jusqu'à Strasbourg pour ma soutenance. Merci à Sébastien et Damien d'avoir également endossé le rôle de rapporteur.

Bonne lecture !

Contents

1 Overview

The era of gravitational waves astronomy kicked off in 2015 with GW150914, the very first detection of gravitational waves validating once more the theory of general relativity developped hundred years earlier. The source was identified as a merger of two black holes. This detection is the result of decades of work to build sensitive-enough detectors such as the two LIGO observatories and Virgo, as well as refined data analysis techniques to confidently claim the detection. The field of gravitational waves has been on the rise ever since: many other detections have occured, confirming the existence of other type of sources such as neutron star-neutron star mergers (GW170817) and neutron star-black hole mergers (GW200105 and GW200115). Joint detection of gravitational waves with electromagnetic waves during GW170817 also opened a new window for multi-messenger astrophysics.

This PhD work started at the end of the third observation period of LIGO and Virgo called O3 and will reach its end a few months after the start of the fourth observing run called O4. It therefore takes place during a time of upgrades, using knowledge accumulated through O3 to prepare for O4 as best as possible.

This work was done within the MBTA team. MBTA is an analysis pipeline that searches for compact binary merger signals in the LIGO and Virgo data. It participated in the low-latency online analysis and offline analysis of O3. Up to O3 MBTA only claimed signals found in coicindence between at least two detectors. The primary goal of this work was to develop a so-called single detector triggers search, to claim a detection if only one detector of the observatory network reported a significant signal. This is especially interesting at times were only one of the detector of the network was online, which can happen for various reasons. Most of the work regarding single detector triggers is about controlling and estimating the background to discover astrophysical signals.

Chapter 2 of this document is dedicated to a description of the state of the art of gravitational waves astronomy. It starts with a short introduction to gravitational waves. A description of the type of gravitational waves sources that were detected and those that are being searched for will also be given, along with some properties and parameters relevant for gravitational waves data analysis. A review of motivations that drive the search for gravitational waves will eventually be done.

Chapter 3 gives an overview of the detection of gravitational waves with LIGO and Virgo, without going into the many and complex details of the detectors. A review of the benefits of having multiple detectors will then be done. It will end with a short review of the future detectors planned for gravitational waves observations.

The analysis of LIGO and Virgo data will be described in chapter 4, first in a general way by explaining the principle of the matched-filtering technique which is the analysis method underlying most of the gravitational waves analysis. Then it focuses on the MBTA pipeline with details on the application of the matched-filtering technique to gravitational waves data analysis. The process of searching for signal in a single detector, several detectors and asserting the significance of an event is explained after that. There will also be a part dedicated to the noise rejection tools used by MBTA to better discriminate against background triggers. All of the work presented in this document relies on data analyzed with MBTA. There will be at the end of this section a presentation of the source classification and a definition of what are called “EM bright” candidates, to which we will pay a particular attention throughout this document.

The work carried out during the PhD starts with chapter 5. It is at the heart of the single detector triggers analysis. There, selection criteria based on data quality-related quantities will be

described. They were derived from O3 data in order to better control the background and reject as many bad triggers as possible while leaving as many astrophysical signals as possible untouched.

Chapter 6 will present a way to reliably estimate the background of single detector triggers by taking advantage of MBTA's way of analysing data in order to compute a false alarm rate for loud events and quantify their significance.

Chapter 7 will describe some work on a noise rejection tool called reweighted SNR to improve its ability to discriminate between noise and astrophysical signals.

Chapter 8 will show the results of an investigation on fluctuations of the PSD which may be the source of some bad triggers.

Finally, in chapter 9, we will present some preliminary results of the single detector trigger search for the begining of O4 with some discussions on the effectiveness of the work presented in this document.

This document is written in English for the benefit of gravitational waves researchers. However, the Ecole Doctorale 182 requires at least 10% of the manuscript to be written in French. Therefore, chapters 2 and 3, which are quite generic, are written in French.

2 Ondes gravitationnelles

Contents

2.1 Introduction aux ondes gravitationnelles

Détectées pour la première fois en 2015, les ondes gravitationnelles étaient prédites par la theorie de la relativité générale depuis 1916. Une manière "simple" de les faire apparaître dans les équations de la relativité générale est en introduisant une perturbation dans la métrique d'un espace-temps plat. Des calculs et changements de coordonnées, qui ne seront pas détaillés dans ce document, permettent d'arriver à une équation d'onde dont la métrique de l'espace-temps est solution. Les ondes gravitationnelles sont à la relativité générale ce que les ondes électromagnétiques sont à l'électromagnétisme. Tout comme la lumière, elles se propagent à la vitesse de la lumière dans le vide. Cependant leurs natures diffèrent grandement. Dans le cas des ondes gravitationnelles c'est la courbure de l'espace-temps elle-même qui suit cette équation d'onde. Cela implique que les ondes gravitationnelles se propagent à travers la matière sans altération. Elles peuvent toutefois être modifiées par de forts champs gravitationnels créés par des objets très massifs.

Bien que se propageant au travers de l'espace-temps, elles n'affectent au premier ordre que les composantes spatiales de la métrique. Elles n'ont d'effets que dans le plan transverse à leur direction de propagation. Ainsi, deux particules ponctuelles séparées dans l'espace uniquement le long de la direction de propagation de l'onde gravitationnelle ne subiront aucun effet. En revanche des particules séparées le long de directions transverses verront leur distance propre osciller proportionnellement à leur distance par rapport à la source et selon l'amplitude et la fréquence de l'onde.

De la même manière que des charges électriques en accélération émettent de la lumière en électromagnétisme, des masses en accélération produisent des ondes gravitationnelles. L'amplitude des ondes est extrêmement faible et croît avec la masse de l'objet accéléré. Pour pouvoir être détecté sur Terre, nous nous attendons donc à ce que l'objet en question soit très massif.

2.2 Aperçu des sources astrophysiques

Cette section présente les principales sources d'ondes gravitationnelles qui sont recherchées par la collaboration LIGO-Virgo-KAGRA (LVK).

Les coalescences de paires d'objets compacts

La première détection d'ondes gravitationnelles a été identifiée comme provenant d'une coalescence de binaire compact (CBC) qui consiste en la coalescence de deux objets compacts. Dans le cas de cette première détection, appellée GW150914 en référence à la date de l'observation, la forme des signaux observés a permis de conclure qu'il s'agissait de deux trous noirs. On parle alors de coalescences de trous noirs (Binary Black Hole, BBH). D'autres types de coalescences de binaires

étaient attendues comme les coalescences de binaires d'étoiles à neutron (Binary Neutron Star, BNS) ou encore les coalescences de binaires de trou noir-étoile à neutron (Neutron Star-Black Hole, NSBH). Il fallut attendre 2017 et 2020 respectivement pour la première observation d'une coalescence BNS et NSBH.

Le processus de coalescence d'une CBC consiste en trois étapes:

- La phase spiralante ou "Inspiral": les deux objets sont en rotation l'un autour de l'autre à grande vitesse. Ils se rapprochent graduellement et leurs vitesses augmentent, atteignant des valeurs du même ordre de grandeur que la vitesse de la lumière dans le vide. Des ondes gravitationnelles sont émises tout au long de ce processus. La fréquence et l'amplitude des ondes augmentent à mesure que les objets se rapprochent et accélèrent.

- La phase de fusion ou " Merger": les deux objets se sont tellement rapprochés qu'ils ont atteint leur dernière orbite circulaire stable (Innermost Stable Circular Orbit, ISCO). Ils "plongent" l'un vers l'autre et fusionnent rapidement. C'est à ce moment que le signal gravitationnel atteint son amplitude et sa fréquence maximales. Le corps résultant de la fusion dépend du système initial. Dans le cas d'une coalescence BBH il s'agira d'un trou noir plus massif. Pour une coalescence NSBH le trou noir va disloquer l'étoile à neutrons et l'absorber, résultant à nouveau en un trou noir. Plusieurs scénarios sont possibles lorsque l'on considère une coalescence BNS. L'objet final peut être un trou noir ou une étoile à neutrons. S'il s'agit d'une étoile à neutrons elle pourrait par la suite s'effondrer et former un trou noir. Cet effondrement peut arriver en quelques secondes ou bien sur une durée plus longue. Ces scénarios dépendent des propriétés des deux étoiles initiales et de l'objet résultant de la fusion.

- La phase de désexcitation ou "Ringdown": Cette étape n'est pertinente que si l'objet résultant est un trou noir massif. Suite à la fusion, le trou noir est dans un état excité et peut être très asymétrique. Il va rayonner de l'énergie par émission gravitationnelle jusqu'à atteindre un état d'équilibre.

La figure 1 montre une représentation du processus complet.

Recherche d'évènements subissant un effet de lentille gravitationnelle

Une lentille gravitationnelle est créée par un objet massif qui distord l'espace-temps. L'effet des lentilles gravitationnelles est observé sur des ondes électromagnétiques. Ces dernières sont déviées et créent pour l'observateur de multiples images de la source qui peuvent être magnifiées ou distordues. Les ondes gravitationnelles sont également sujettes aux effets des lentilles gravitationnelles si la source, par exemple une CBC, est située derrière un objet massif. Pour des objets très massifs, tel qu'un amas de galaxies, on parle alors de "lentille massive" ou encore "lentille gravitationnelle forte". Pour des objets de masses plus faibles, comme une étoile ou un objet compact, on parle de "microlentille". On s'attend à ce que des lentilles fortes produisent de multiples signaux dans nos détecteurs que l'on verrait comme une forme d'onde qui se répète dans les données espacés de temps pouvant atteindre quelques mois. Les lentilles fortes ne devraient pas modifier l'évolution en fréquence des ondes gravitationnelles [2]. Grâce aux multiples images les évènements subissant un effet de lentille gravitationnelle devraient permettre une meilleure localisation de leur source que les CBC habituels. Les microlentilles pourraient quant à elles introduire un rythme de battement dans la forme d'onde [3]. Cet effet pourrait apparaître si les conditions suivantes sont satisfaites:

- le delai temporel entre les différents chemins (qui produisent différentes images) doit être du même ordre de grandeur que la période de l'onde gravitationnelle.

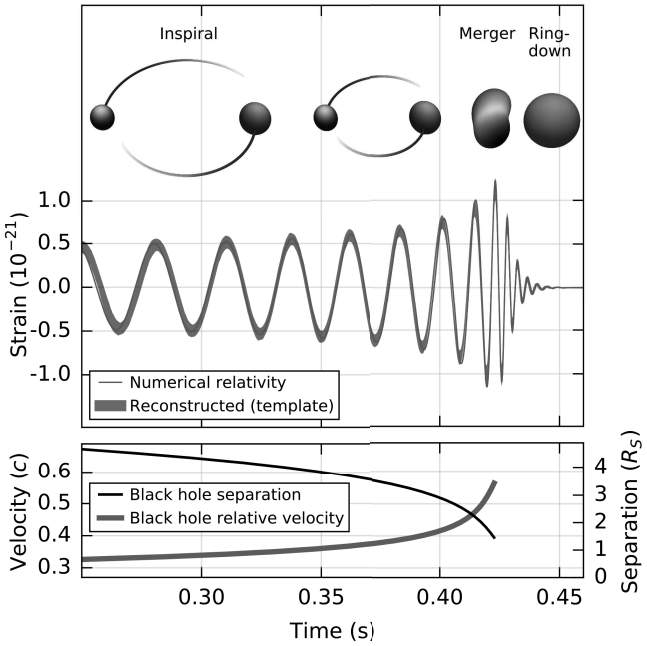


Figure 1: Figure issue de [1]. Panneau du haut: schéma de la phase d'inspiral, de fusion et de ringdown pour une coalescence de binaire de trous noirs, ainsi que la forme des ondes émises au fil du processus. Panneau du bas: évolution de la séparation entre les trous noirs et de leur vitesse relative au cours du temps.

- l'évènement doit être situé proche d'une caustique (ligne qui sépare les régions de différentes multiplicité d'images) afin d'être suffisamment magnifié pour permettre sa détection.

Recherche d'ondes continues

Les ondes continues sont des ondes gravitationnelles observables sur une échelle de temps bien plus grande que celles provenant des CBC. De possibles sources d'ondes continues sont les étoiles à neutrons en rotation avec une asymétrie ou une non-uniformité. Les ondes continues émises par un tel système auraient une fréquence proche de la fréquence de rotation de l'étoile (ou du moins reliée à celle-ci). Elles seraient donc quasiment monochromatiques avec des variations dûes à l'effet Doppler lié aux changements de la position de la source dans le ciel au cours du temps, ainsi qu'une évolution suivant le "spin-down" (ralentissement de la rotation) de l'étoile [4].

Les recherches d'ondes continues peuvent être

- ciblées, c'est à dire que les données analysée proviennent d'une région du ciel où la présence d'une étoile à neutrons en rotation est connue via l'observation d'un pulsar,
- dirigées, en analysant des données provenant d'une région spécifique du ciel dans laquelle il n'y a pas de source connue telle que le centre galactique,
- globale ou "all-sky", visant à chercher des sources dans toutes les directions.

Recherches de bursts d'ondes gravitationnelles

Les bursts sont des phénomènes très courts qui peuvent durer de quelques secondes à quelques millisecondes. Des source potentielles de burst d'ondes gravitationnelles sont les supernovae à effondrement de coeur, les excitations d'étoiles à neutrons, les cordes cosmiques [5]... Les recherches de bursts peuvent utiliser des modèles de formes d'ondes comme les recherches CBC ou bien être des recherches sans modèle pour identifier les excès de puissance corrélés dans les détecteurs.

Recherche d'un bruit de fond stochastique d'ondes gravitationnelles

Toutes les sources d'ondes gravitationnelles que nous ne détectons pas directement produisent tout de même des ondes qui peuvent nous atteindre. Ces sources non résolues sont distribuées dans tout le ciel. Les ondes qu'elles émettent devraient alors se superposer et créer un bruit de fond stochastique d'ondes gravitationnelles [6]. Ce bruit de fond devrait être isotrope au premier ordre mais pourrait avoir de faibles fluctuations le rendant anisotrope comme le fond diffus cosmologique [7]. Des recherches pour un bruit de fond isotrope et anisotrope existent [8, 9].

2.3 Propriétés des ondes gravitationnelles et des CBC

La forme d'onde des ondes gravitationnelles produites par une CBC est déterminée par les propriétés de la source, en particulier les masses et spins des deux objets du système binaire. On appelle par convention masse1, spin1 et on écrit m_1, s_1 la masse et le spin de l'objet le plus lourd. De même on appelle masse2, spin2 (m_2, s_2) la masse et le spin de l'objet le plus léger. Dans les faits, on ne considère en général que la composante de spin orthogonale au plan de rotation du système binaire $s_{1,2z}$, et on note alors $s_{1,2} = s_{1,2z}$. Les spins sont définis à partir du moment angulaire $J_{1,2}$ des deux objets et sont normalisés par les masses:

$$s_i = \frac{J_i}{m_i^2} \ , \ i \in \{1, 2\} \tag{1}$$

Ils prennent des valeurs entre -1 et 1.

Nous définissons cinq quantités supplémentaire qui sont courament utilisées. La masse totale

$$M_{\text{tot}} = m_1 + m_2 \tag{2}$$

la masse réduite

$$\mu = \frac{m_1 m_2}{M_{\text{tot}}} \tag{3}$$

la masse chirp

$$\mathcal{M} = m_{chirp} = \frac{(m_1 m_2)^{\frac{3}{5}}}{(m_1 + m_2)^{\frac{1}{5}}} = \mu^{3/5} M_{\text{tot}}^{2/5} \tag{4}$$

le rapport des masses

$$q = \frac{m_1}{m_2} \tag{5}$$

et le spin effectif

$$\chi_{\text{eff}} = \frac{s_1 m_1 \cos t_1 + s_2 m_2 \cos t_2}{m_1 + m_2} \overset{\text{(anti)aligned spins}}{=} \frac{s_1 m_1 + s_2 m_2}{m_1 + m_2} \tag{6}$$

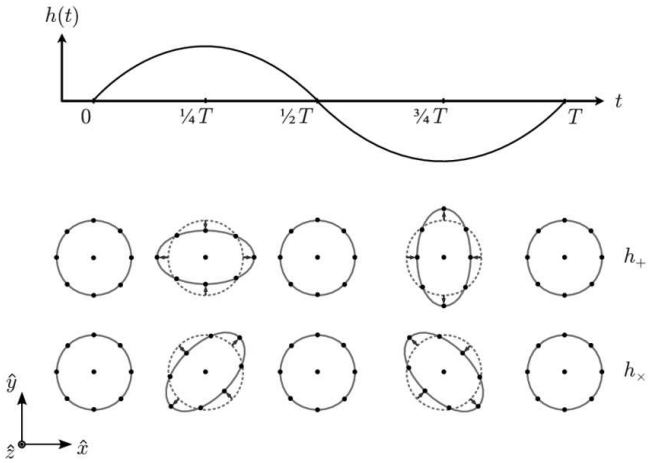


Figure 2: Figure issue de [14]. Polarisations plus et croix d'une onde gravitationnelle se propageant selon l'axe z et leur effet sur un ensemble de particules ponctuelles en chute libre.

avec t_1 et t_2 les angles d'inclinaisons de chaque objet par rapport au moment orbital angulaire du système binaire.

Les ondes gravitationnelles peuvent être écrites comme le mélange de deux polarisations formant un angle de 45° [10] appelées "polarisation plus", h_+, et "polarisation croix", $h_\times$, se rapportant aux directions dans lesquelles elles étendent et compressent l'espace. La figure 2 montre leurs effets sur des particules ponctuelles.

Ces polarizations peuvent être écrites, pour un système CBC en orbite circulaire, comme [11]

$$\begin{aligned} h_+ &= -\frac{2\mathcal{M}}{R}(2\pi\mathcal{M}f_b)^{2/3}(1+\cos^2 i)\cos(2\Phi_b(t)) \\ h_\times &= -\frac{2\mathcal{M}}{R}(2\pi\mathcal{M}f_b)^{2/3}(2\cos i)\sin(2\Phi_b(t)) \end{aligned} \tag{7}$$

Où i est l'inclinaison du système par rapport au plan du ciel, f_b est la fréquence orbitale du système binaire et $\Phi_b(t) = 2\pi\int^t f_b(t')dt'$ est la phase orbitale. R est l'échelle charactéristique de distance au système. Elle peut être prise comme la distance de luminosité

$$D_L(z) = (1+z)\int_0^z \frac{dz'}{H(z')} \tag{8}$$

où z est le redshift, H le facteur de Hubble tel que $H = a^{-2}da/d\eta$ with $a(\eta)/a(\eta_0) = (1+z)^{-1}$ et η est le temps conforme [12]. Ou encore la distance effective [13]

$$D_{\text{eff}} = D\left(F_+^2\left(\frac{1+\cos^2 i}{2}\right)^2 + F_\times^2\cos^2 i\right) \tag{9}$$

telle que $D_L \sim D_{\text{eff}}$ avec F_+ et $F_\times$ les réponses d'antennes de l'interféromètre aux polarisations plus et croix et D la distance à la source. Les réponses d'antennes sont rapidement décrites dans le chapitre 3.

Ce que l'on mesure sur Terre sont en fait les quantités redshiftées telles que la masse chirp redshiftée $\mathcal{M}_z = (1+z)\mathcal{M}$ et la fréquence redshiftée $f_z = f/(1+z)$.

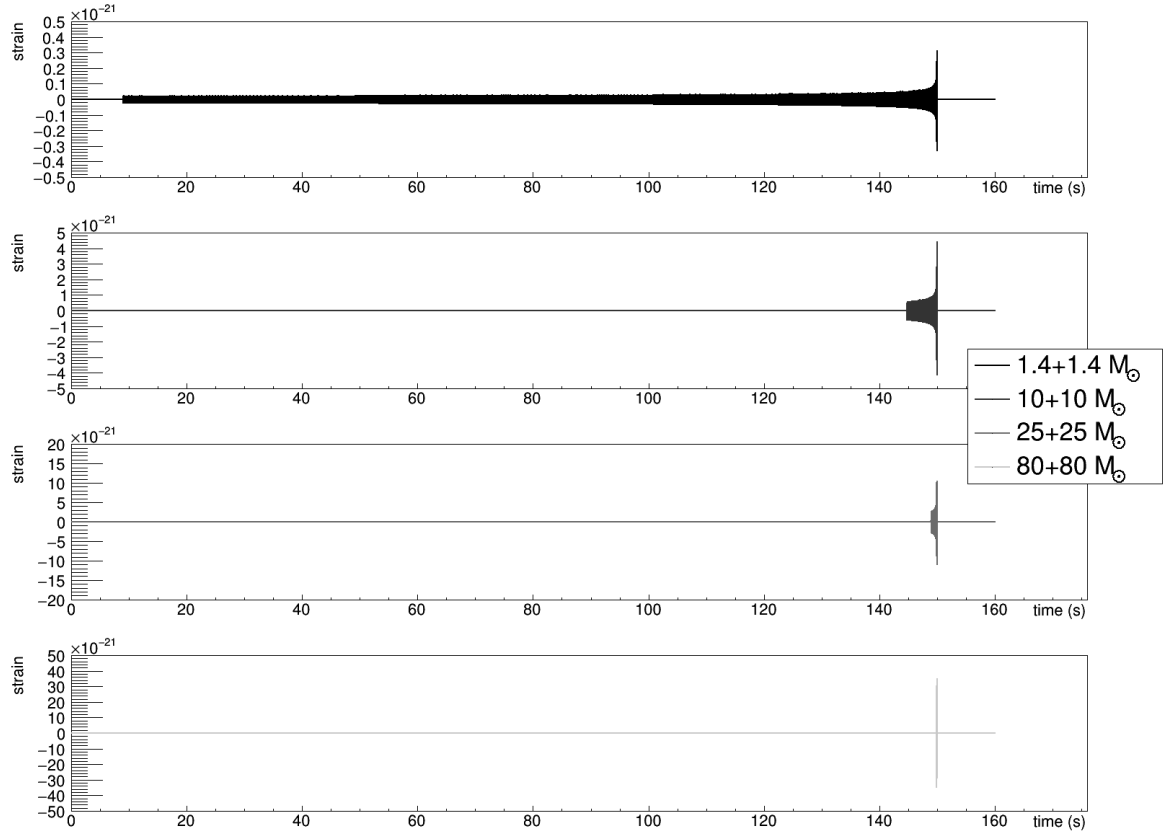


Figure 3: Forme d'onde pour des systèmes de différentes masses. Les autres paramètres sont spin1=spin2=0, fréquence de début de génération=21 Hz, distance effective=40 Mpc.

On voit en figure 3 que plus le système binaire est lourd, plus le signal émis dans une certaine bande de fréquence est court. On remarque également en figure 4 que des spins élevés sont associés à des formes d'ondes plus longues, et que des spins anti-alignés sont associés à des formes d'ondes plus courtes. Enfin, la figure 5 montre l'impact du choix de la fréquence de début de génération de la forme d'onde sur la durée de celle-ci.

2.4 Motivations pour la détection

En plus de la volonté de découvrir de nouvelles sources d'ondes gravitationnelles, de nombreuses motivations pour leur détection existent. Nous en présentons ici les principales.

Astronomie multi-messager

Le but de l'astronomie multi-messager est de mieux comprendre les populations d'objets astrophysiques, leur évolution et leur environnement. Le développement de l'astronomie multi-longueurs d'ondes a été motivé par le fait que différentes fréquences lumineuses fournissent différentes informations. De la même manière, l'ajout des informations obtenues par l'observation (ou l'absence d'observation) de rayons cosmiques, neutrinos de haute énergie ou d'ondes gravitationnelles a enormément de valeur. Un aperçu des jalons et découvertes multi-messagers est donné dans [15]. De nombreux efforts sont fournis pour faire des détections jointes de différent types de messagers. Une association entre onde électromagnétique et neutrinos de haute énergie a été rapportée par plusieurs

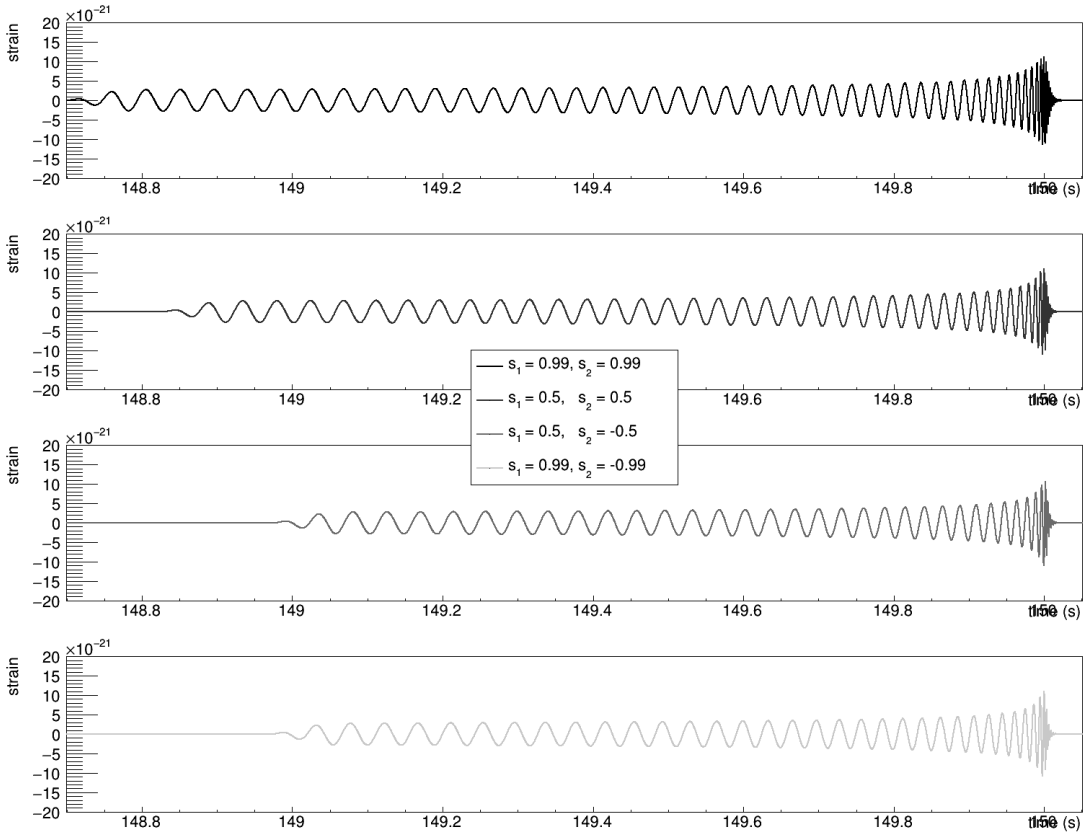


Figure 4: Forme d'onde pour des systèmes avec différents spins. Les autres paramètres sont mass1=mass2=25 $M_\odot$, fréquence de début de génération=21 Hz, distance effective=40 Mpc.

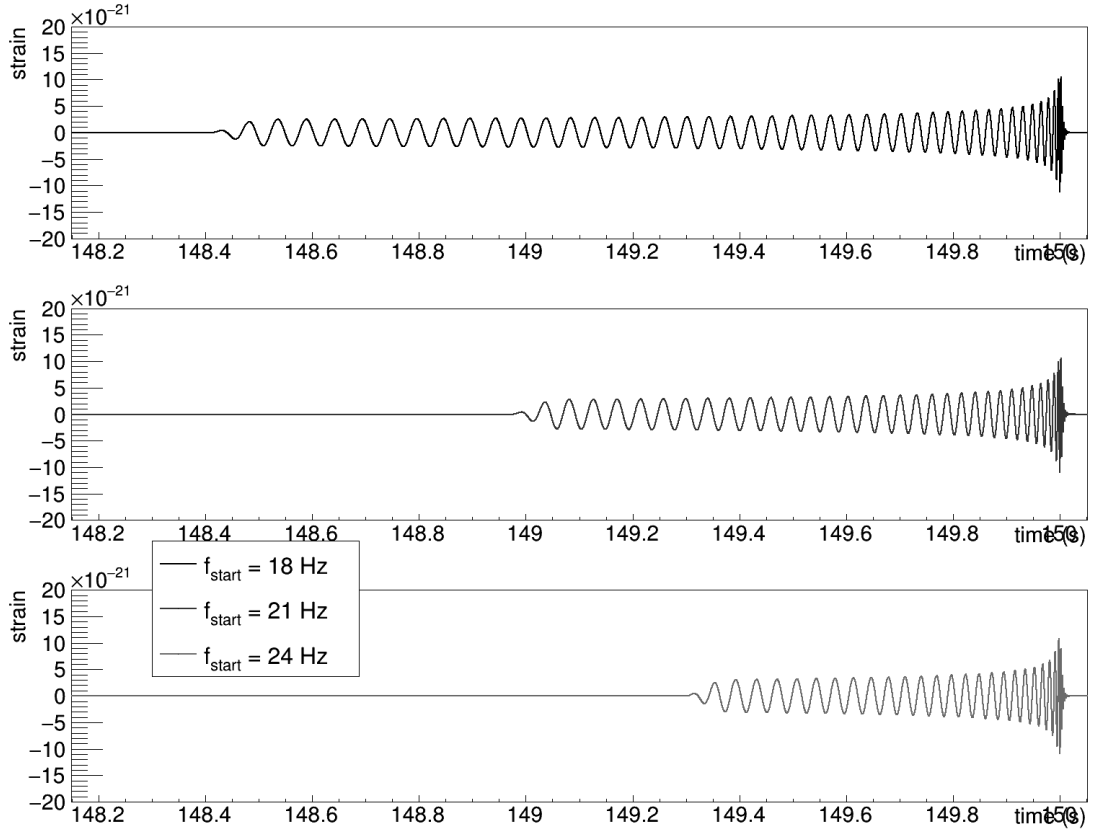


Figure 5: Forme d'onde représentée pour différentes valeurs de fréquence de début de génération pour un système de mass1=mass2=25 $M_\odot$, s1=s2=0 et distance effective=40 Mpc.

observatoires suite à la détection de neutrinos dont la position de la source était compatible avec un blazar de rayons γ connu [16]. La collaboration IceCube a également rapporté une preuve d'un excès de neutrinos compatibles avec la position d'une galaxie active proche [17]. Aucune détection jointe d'ondes gravitationnelles et neutrinos n'a été faite à ce jour [18, 19]. En revanche, une observation jointe d'ondes gravitationnelles et électromagnétiques s'est produite avec l'évènement GW170817 [20, 21], la toute première détection d'une coalescence BNS qui fut alors suivie d'un sursaut gamma. La détection d'onde gravitationnelle localisait la source dans une zone d'environ 31 degrés^2, à une distance de 40 Mpc (suivie par d'autres mesures de la position de la source). Le sursaut gamma a été détecté par plusieurs observatoires environ 1.7 s après la fusion avec une localisation de la source dans une zone de 1100 degrés^2. Des émissions d'ondes radios et de rayons X furent églament observées plusieurs jours après l'évènement.

Tests de la relativité générale

La liste de tests donnée ici n'est pas exhaustive et plus de détails peuvent être trouvés dans les différentes oeuvres citées.

Vitesse de propagation: La relativité générale prédit que les ondes gravitationnelles se propagent non dispersivement à la vitesse de la lumière dans le vide. Toute déviation serait donc un argument fort en faveur de théories alternatives de la gravité qui prédisent certaines variations de cette vitesse de propagation. Elles pourraient être causées, par exemple, par le couplage des ondes avec des champs gravitationnels résiduels ou bien si la gravitation est associée à un champ massif impliquant l'existence d'un graviton massif [22, 11, 23]. D'un point de vue observationnel, une approche directe pour tester la vitesse de propagation des ondes gravitationnelles est de comparer leurs temps d'arrivée à ceux d'ondes électromagnétiques pour un même évènement. Cela requiert toutefois une connaissance précise des temps d'émissions des deux. La détection de GW170817 et du sursaut gamma associé GRB170817a a permis de contraindre la différence entre les vitesse de propagations (Δv) des ondes gravitationelles et de la lumière dans le vide (c) [21]:

$$-3 \times 10^{-15} \leq \frac{\Delta v}{c} \leq +7 \times 10^{-16}$$

Ce resultat a éliminé un certain nombre de théories alternatives de la gravitation.

Polarisations: Un autre test de la relativité générale consiste à chercher de nouvelles polarisations pour les ondes gravitationnelles. En effet, la relativité générale ne permet que deux polarisation tensorielles tandis que d'autres théories en prédisent quatres additionnelles. Il s'agirait de deux polarisations scalaires et de deux polarisations vectorielles comme représenté en figure 6. Ces autres polarisations sont recherchées en même temps que les ondes gravitationnelles continues [24], le bruit de fond stochastique [10] ainsi que des évènements CBC. Des perspectives de recherches à l'aide de futurs détecteurs existent également [23, 25].

Déviations des prédictions de formes d'ondes: Les formes d'ondes sont prédites par la relativité générale et sont utilisées pour chercher du signal dans les données des détecteurs (voir section 4.3). Des déviations sont recherchées en regardant la puissance résiduelle dans le détecteur après avoir soustrait aux données la forme d'onde prédite d'un signal détecté. Un excès de puissance suffisament significatif serait un indicateur de déviation.

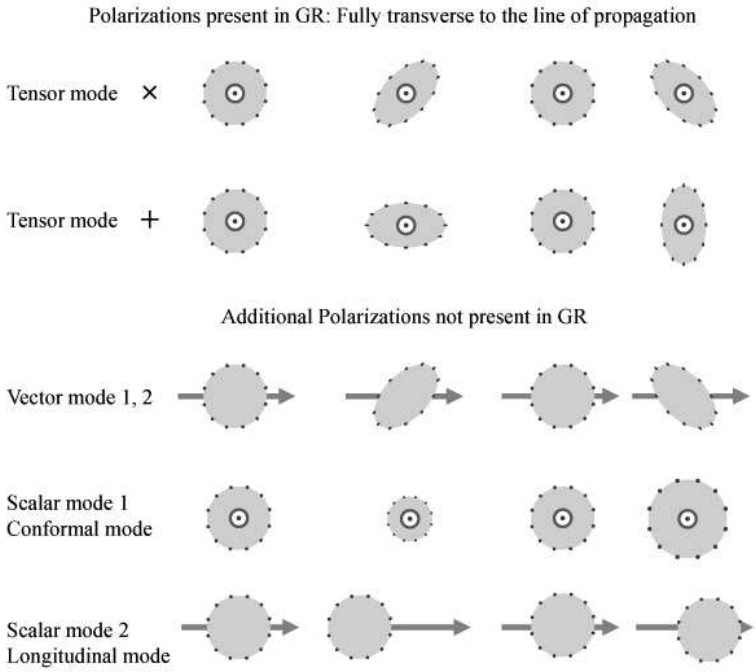


Figure 6: Figure issue de [26]. Polarisations "+" et "×" ainsi que les autres types de polarisations qui ne sont pas prédits par la relativité générale.

Mesure de la constante de Hubble

La valeur de la constante de Hubble est un sujet très actif depuis de nombreuses années à cause des différences entre mesures directes et indirectes qui sont en désaccords allant parfois jusqu'au niveau de $4-6\sigma$ level [27] comme montré en figure 7. Les mesures directes et indirectes correspondent à des valeurs de redshift différentes (univers "jeune" et "vieux"), et permettent donc de tester la cohérence des modèles d'évolution cosmologique.

La constante de Hubble, H_0, est mesurée en étudiant la relation en distance et redshift pour des populations d'objets astrophysiques. Des objets permettant une telle mesure sont appelés chandelles standards. Il est possible de mesurer la distance de luminosité (eq. 8) des sources CBC avec les ondes gravitationnelles mais il manque une mesure indépendante du redshift pour établir la relation de l'un en fonction de l'autre. GW170817 était une chandelle standard grâce à la contrepartie électromagnétiques. Cet évènement a permis de mesurer $H_0 = 70.0^{+12.0}_{-8.0}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ [28]. Pour les autres évènements on peut utiliser des catalogues de galaxies afin d'identifier la galaxie hôte ou, de manière plus réaliste, marginaliser sur de multiples potentielles galaxies hôtes pour obtenir une mesure indépendante du redshift. Une analyse raffinée utilisant cette méthode a donné pour résultat $H_0 = 68.7^{+17.0}_{-8.3}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ [29]. Une mesure utilisant uniquement l'information onde gravitationnelle et se reposant sur les particularités des distributions de masse des CBC est également possible. De futures mesures avec une meilleure précision vont permettre de contraindre encore plus la valeur de la constante de Hubble pour les modèles cosmologiques.

Contraintes sur l'équation d'état des étoiles à neutrons

L'équation d'état (EOS) des étoiles à neutrons relie les variables d'état telles que pression et densité entre elles et décrit la matière nucléaire à extrêmement haute densité à l'intérieur de l'étoile. Elle

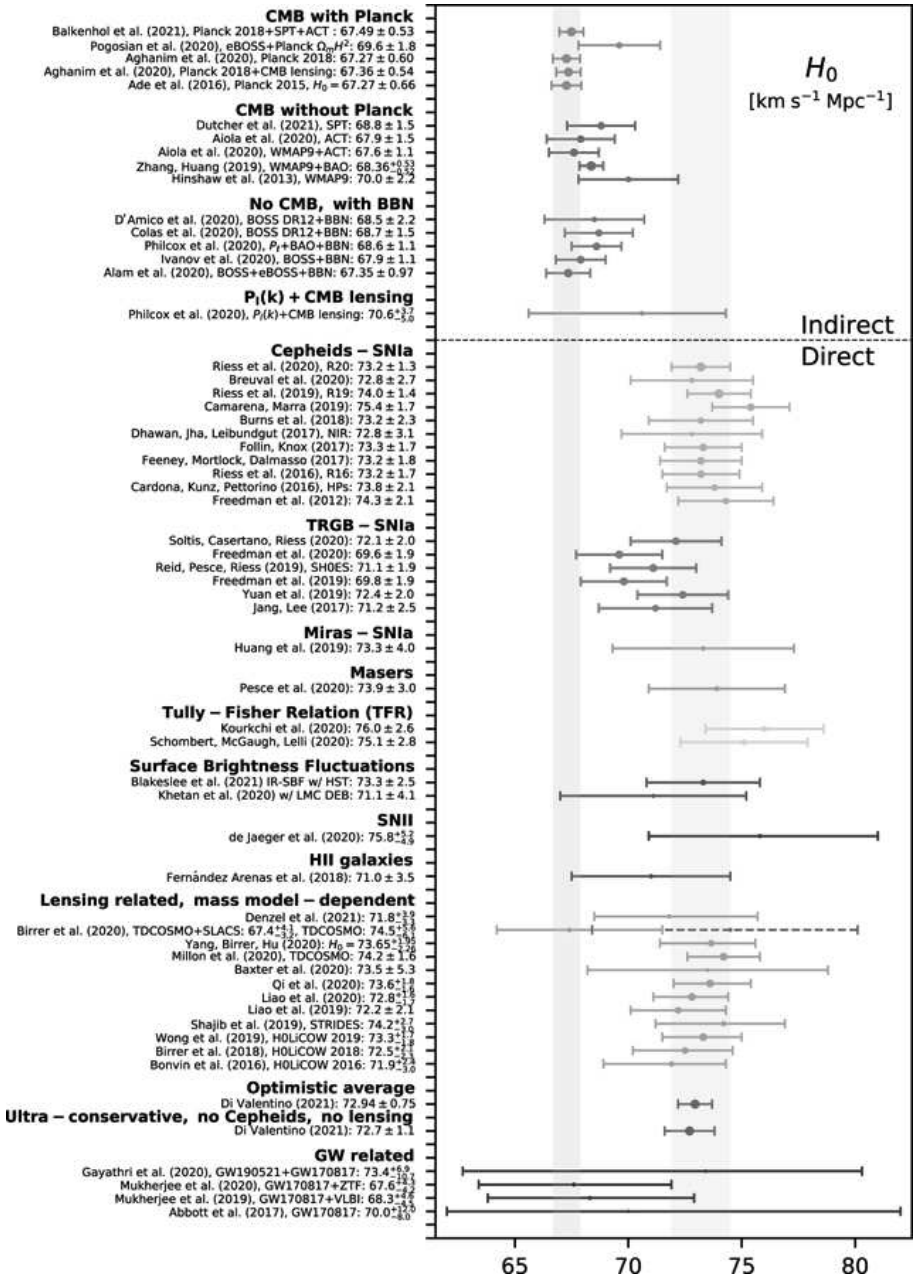


Figure 7: Figure issue de [27]. Mesures directes et indirectes de H_0 avec intervalles à 68% de niveau de confiance effectuées par diverses missions et collaborations. Traduit de [27]: "La bande verticale cyan correspond à la valeur de H0 donnée par SH0ES [30] (R20, $H0 = 73.2 \pm 1.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ at 68% CL) et la bande verticale rose correspond à la valeur de H0 rapportée par Planck 2018 [31] dans un scénario ΛCDM."

est utilisée pour dériver de nombreuses propriétés de l'étoile à neutrons telles que son rayon, sa masse, son moment d'inertie, sa déformabilité... L'EOS permet également, en se basant sur les masses et populations d'étoiles à neutrons, de contraindre les mécanismes de formation des étoiles à neutrons. De nombreux efforts ont été faits afin de développer des EOS permettant de décrire la matière d'une étoile à neutrons de la croûte jusqu'au coeur intérieur. Mais l'EOS est encore très peu contrainte et différentes approches donnent des équations différentes [32, 33]. Les expériences sur Terre qui visent à étudier la matière nucléaire à très faible température et forte densité ne peuvent pas reproduire les conditions de l'intérieur d'une étoile à neutron. Il est donc impératif d'avoir recours à des observations astrophysiques via (de manière non exhaustive) les pulsars radios, les binaires à rayons X et, plus intéressant pour ce document, via les ondes gravitationnelles émises par les fusions d'étoiles à neutrons [34, 35, 36]. La déformabilité causée par la force de marée de chacune des étoiles du système binaire sur l'autre est la quantité qui est la mieux mesurée pour contraindre l'EOS. Elle est utilisée pour calculer les rayons des deux étoiles du système binaire. En combinant ces mesures à celle des masses des deux étoiles à neutrons, il est possible de placer des contraintes sur l'équation d'état.

Taux et populations de CBC

Il est possible de considérer les évènements CBC non pas comme des évènements individuels mais comme une population de sources pour étudier leurs propriétés. Cette considération est d'autant plus intéressante avec le début de O4 en Mai 2023 durant lequel on s'attend à plus de détections. Les quantités que l'ont peut dériver de ces études sont par exemple le taux de merger des différents types de CBC et leur distribution en masse. En utilisant le catalog GWTC-3 [37], les taux de fusions ont été rapportés dans [38] entre

- $10\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ et $1700\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ pour des évènements BNS,
- $17.9\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ et $44\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ pour des évènements BBH à redshift $z \leq 0.2$,
- $7.8\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ et $140\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ pour des évènements NSBH.

Les distributions de masses rapportées sont

- uniforme entre $1.2^{+0.1}_{-0.2}\ \mathrm{M}_\odot$ et $2.0^{+0.3}_{-0.3}\ \mathrm{M}_\odot$ pour les étoiles à neutrons,
- loi de puissance avec pics à $\mathcal{M}_c = 8.3^{+0.3}_{-0.5}\ \mathrm{M}_\odot$ et $\mathcal{M}_c = 27.9^{+1.9}_{-1.8}\ \mathrm{M}_\odot$ pour les BBH.

Recherche de masses subsolaires

La découverte d'un objet compact avec une masse inférieure à $1\,\mathrm{M}_\odot$ serait la preuve de nouveaux mécanismes de formations d'objets compacts. En effet les modèles actuels d'évolution stellaire pour les étoiles massives ne permettent un effondrement qu'en étoile à neutrons ou trou noir plus lourd que $1\,\mathrm{M}_\odot$ [39, 40]. Ces objets de masse subsolaire pourraient être des trous noirs primordiaux [41] ou bien des trous noirs formés par effondrement de matière noire [42, 43, 44] et indiqueraient l'existence d'une nouvelle physique. Des recherches de masses subsolaires sont menées par certaines chaînes d'analyses CBC avec un espace des paramètres dédié. MBTA (chapitre 4) est l'une d'entre elles.

Recherche d'objets compacts exotiques

De nombreux objets compacts exotiques ont été théorisés, mais leur existence n'est toujours pas prouvée. La forme d'onde crée par la coalescence de certains de ces objets est attendue comme identique à celle d'un signal BBH jusqu'à la phase de merger ou post-merger. D'autres objets exotiques pourraient radier des ondes gravitationnelles par eux-même ou encore avoir un effet sur les ondes gravitationnelles comme du lensing [45, 46].

3 Détection des ondes gravitationnelles

Contents

3.1 Principe de détection

Nous nous concentrerons dans ce chapitre sur les détecteurs d'ondes gravitationnelles actuels qui reposent sur des techniques d'interférométrie laser. Ils sont basés sur la comparaison de deux distances orthogonales de l'espace. Ces distances sont délimitées par des masses libres qui sont les miroirs d'un interféromètre de Michelson, interféromètre qui nous permet de comparer ces distances.

Dans un interféromètre, un faisceau laser est scindé en deux et envoyé vers les miroirs. Les deux faisceaux sont réfléchis, se recombinent et le faisceau résultant est analysé par un photodétecteur. Un schéma simple est donné en figure 8. Pour optimiser la sensibilité, cette recombinaison des faisceaux est choisie de sorte à ce qu'ils interfèrent de manière quasi-destructive et qu'il n'y ait donc presque aucune lumière en sortie de l'interféromètre. On parle alors de configuration "franges sombres". Le phénomène d'interférence, et donc la quantité de lumière qui peut être transmise en sortie de l'interféromètre, dépend de la différence de phase des faisceaux recombinés. Cette différence de phase dépend à son tour directement de la difference des chemins optiques parcourus par chaque faisceau. C'est ce concept qui est exploité pour la détection d'ondes gravitationnelles. Pour un interféromètre dont la longueur des bras est fixée et dont les sources de bruits sont sous contrôle, la configuration de franges sombres est maintenue.

Considérons maintenant le passage d'une onde gravitationnelle au travers de cet interféromètre, avec une direction de propagation colinéaire à l'un des bras de l'interféromètre. Ce bras va être aveugle au passage de l'onde car cette dernière n'a des effets que le long des directions orthogonales à sa direction de propagation. En revanche, ce n'est pas le cas pour l'autre bras de l'interféromètre. Ainsi, la distance propre entre la séparatrice (Beam Splitter, BS) et le miroir de ce bras va osciller à cause du passage de l'onde. Le chemin optique parcouru par le laser va donc être modifié, ce qui va causer un décalage dans la différence de phase des deux faisceaux. Le passage de l'onde va ainsi modifier la configuration de franges sombres. Une photodiode sensible permet de détecter cette variation et donc d'identifier le passage de l'onde gravitationnelle.

3.2 Description des détecteurs

Les détecteurs LIGO et Virgo sont bien plus complexes qu'un interféromètre de Michelson usuel. La quantité mesurée est appelée "contrainte" ou "*detector strain*". Elle est calculée à partir de la

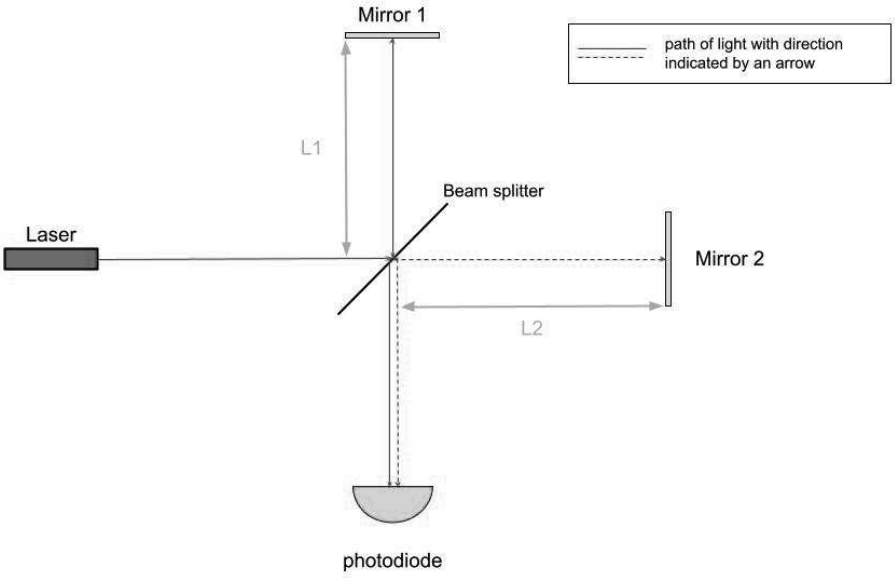


Figure 8: Schéma d'un interféromètre de Michelson

variation de longueur d'un bras par

$$h = \frac{\Delta L}{L_0} \tag{10}$$

Où $\Delta L = L_1 - L_2$ est la différence de longueur entre les deux bras de l'interféromètre et $L_0 = 3\,\mathrm{km}(4\,\mathrm{km})$ pour Virgo (LIGOs) est la longueur des bras au repos. Cette contrainte est sans dimension par construction et varie avec le temps. Elle est communément appellée "h de t" et notée $h(t)$.

Une schématisation de la disposition optique de Virgo pour la période d'observation O3 est présentée en figure 9. Les détails de l'interféromètre peuvent être trouvés dans [47, 48, 49] et les divers papiers référencés dans ce chapitre. Nous n'allons ici qu'en présenter les parties principales afin de décrire la figure 9. Le laser permet de délivrer un faisceau très stable. Des modulateurs électro-optiques (*Electro-Optic Modulator*, EOM) sont placés après celui-ci pour créer des fréquences supplémentaires utilisées pour assurer le fonctionnement du détecteur. Des cavités Fabry-Pérot sont présente dans chaque bras de l'interféromètre entre les "*West/North input mirrors*" (WI/NI) et "*West/North end mirrors*" (WE/NE). Le faisceau, réfléchi de nombreuses fois dans ces cavités, parcours alors un chemin optique bien plus grand que la longueur des bras. Les miroirs de 40 kg sont faits de silice fondue à faible absorption. Leurs surfaces sont recouvertes d'empilements de couche minces assurant leur propriété réfléchissante, ou anti-réfléchissante. Ils sont suspendus à une chaine de filtres mécaniques, fonctionnant dans les six degrés de libertés, assurant leur isolation des mouvements du sol pour des fréquences supérieures à quelques Hz. Un miroir de recyclage de puissance (*Power Recycling mirror*, PR) est utilisé pour réinjecter dans l'interféromètre la puissance qu'il réfléchit et augmenter de plus d'un ordre de grandeur la puissance des faisceaux circulant dans l'interféromètre. Cela permet de réduire l'incertitude liée au bruit quantique en sortie du détecteur.

Un miroir de recyclage du signal (*Signal Recycling mirror*, SR) permet également d'atteindre une meilleure sensibilité [50, 51]. Des fractions des faisceaux sont prélevées à différents endroits de l'interféromètre pour fournir des signaux permettant d'assurer le contrôle continu du positionnement des différents éléments de l'interféromètre. Des nettoyeurs de modes sont placés en entrée et sortie du détecteur et sont notés IMC/OMC pour *Input/Output Mode Cleaner*. Ils assurent la qualité du faisceau laser en entrée et sortie de l'interféromètre. La photodiode B1 est placée à la sortie du détecteur pour collecter le signal contenant l'information onde gravitationelle. L'utilisation d'une source de "*squeezed light*" [52], injectée dans l'interféromètre, permet de réduire le bruit quantique lié au comptage des photons. La plupart de ces composants sont placés dans le vide.

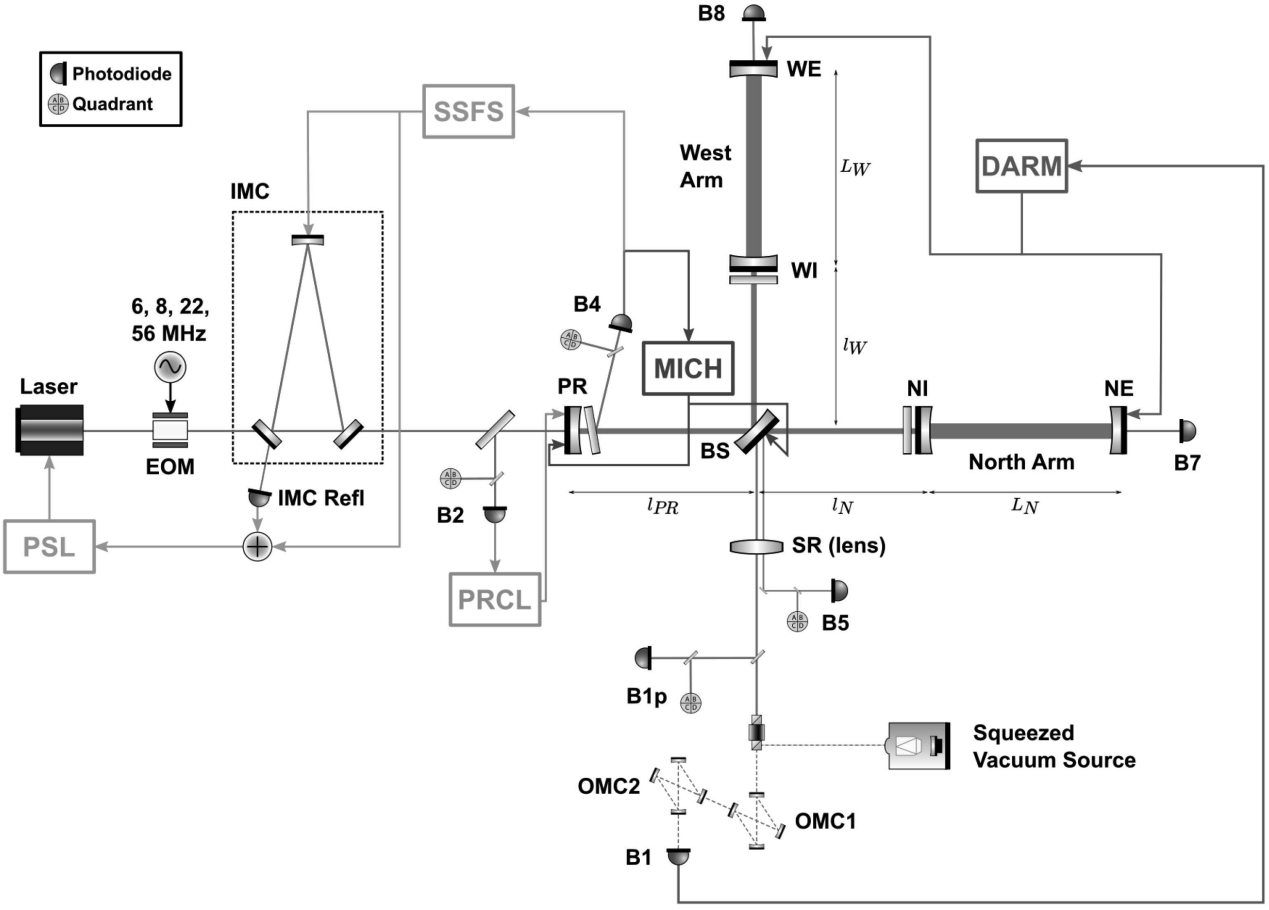


Figure 9: Dispositif optique de Virgo durant O3 provenant de [53] décrit en partie dans le texte. MICH, PRCL, CARMN et DARM sont les degrés de libertés longitudinaux charactéristiques du détecteur. Plus de détails peuvent être trouvés dans la référence.

Des techniques de calibration ont été développées pour obtenir une mesure précise de $h(t)$:

- Le calibrateur photonique [54] (PCal) utilise la modulation de la pression de radiation d'un laser pour induire un déplacement des miroirs de fin. Il s'agissait de la méthode de calibration de référence durant O3.
- Le calibrateur newtonien [55, 56] (NCal) est une nouvelle technique développée pour le détecteur Virgo à Annecy puis Strasbourg. Le principe de cette méthode est d'induire un

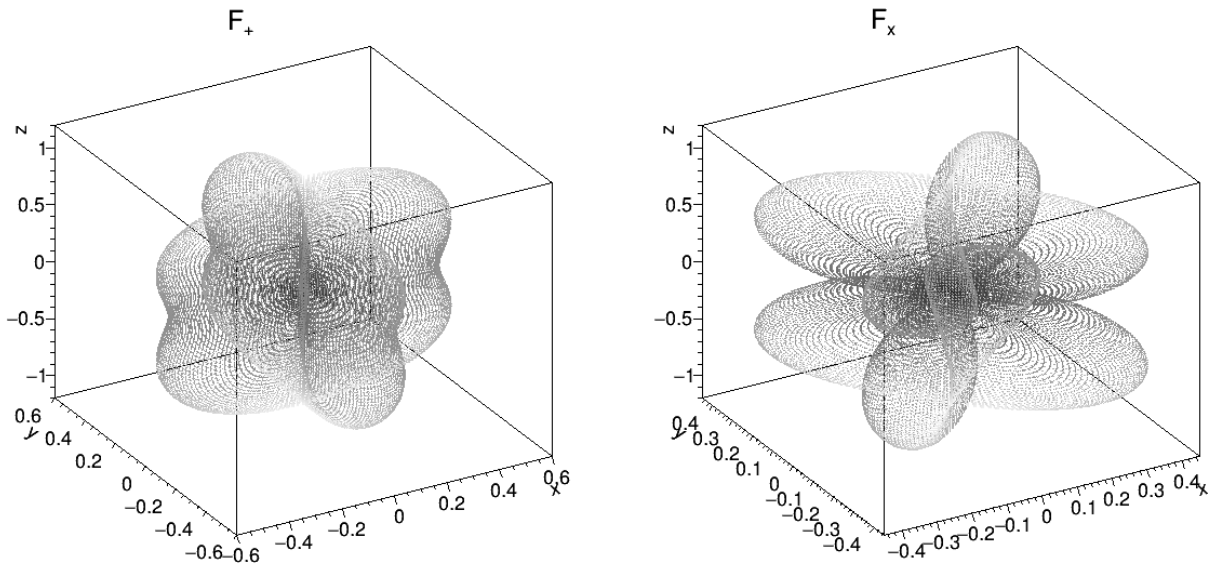


Figure 10: Diagramme d'antenne d'un détecteur aux polarisations plus et croix.

déplacement du miroir de fin à l'aide d'un champ gravitationnel variable, produit par des masses en rotation. Cette technique a, entre autre, l'avantage de ne pas nécessiter de fenêtre d'accès au miroir puisque la force gravitationnelle agit à travers la matière.

La référence [57] donne une vue d'ensemble de la calibration et de la reconstruction de $h(t)$ pour Virgo durant O3.

3.2.1 Réponse d'un détecteur d'ondes gravitationnelles

La réponse d'un détecteur à une onde gravitationnelle peut être modélisée par son diagramme d'antenne. Cette réponse est une fonction de la direction de propagation de l'onde par rapport à l'orientation du détecteur telle que

$$h = F_+ h_+ + F_\times h_\times \tag{11}$$

où [58]

$$F_+ = -\frac{1}{2}(1+\cos^2\theta)\cos 2\phi\cos 2\psi - \cos\theta\sin 2\phi\sin 2\psi \tag{12}$$

$$F_\times = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\sin 2\psi - \cos\theta\sin 2\phi\cos 2\psi \tag{13}$$

sont les réponses d'antennes du détecteur aux polarisations "plus" et "croix" respectivement. θ et ϕ sont les angles de position dans le ciel de la source des ondes et ψ est l'ange de polarisation. La figure 10 montre ces réponses pour un angle de polarisation de 0. On remarque que le cas optimal pour une détection est lorsque l'onde se propage selon une direction orthogonale aux deux bras de l'interféromètre.

La forte sensibilité des détecteurs fait qu'ils sont facilement pollués par de nombreuses sources de bruit. La sensibilité d'un détecteur est caractérisée par la densité spectrale de puissance (*Power Spectral Density*, PSD) de son bruit. Les figures 11 et 12 montrent les contributions de différentes

sources de bruit à l'amplitude de densité spectrale (*Amplitude Spectral Density*, ASD = $\sqrt{\text{PSD}}$) des détecteurs afin d'en illustrer la complexité. La figure 13 montre la comparaison des ASD des détecteurs LIGO et Virgo, pour les périodes d'observations O2 et O3.

3.2.2 Horizon et *range* d'un détecteur d'ondes gravitationnelles

Différentes quantités sont utilisées pour quantifier la sensibilité des détecteurs avec une unique valeur [13, 59, 60].

La distance d'horizon est définie pour un type de source et un seuil de ratio signal-sur-bruit ρ (*Signal-to-Noise Ratio*, SNR détaillé en section 4). Il s'agit de la distance effective (eq. 9) pour laquelle ce type de source serait détecté avec un SNR attendu de ρ pour l'orientation optimale. En d'autres termes l'horizon est la distance maximale à laquelle il est possible de détecter cette source avec le seuil donné. L'horizon est donné par [13] (Appendix D)

$$D_{\text{horizon}} = \frac{1}{\rho}\frac{(GM_c)^{5/6}}{\pi^{2/3}c^{3/2}}\frac{\text{M}_\odot{}^{5/3}}{1\,\text{Mpc}}\sqrt{\left(\frac{5}{6}\right)\int_0^\infty \frac{f^{-7/3}}{S_n(f)}df} \tag{14}$$

où $S_n(f)$ est la PSD du bruit du détecteur. En pratique l'intégration est faite en commençant à f_{low}, la fréquence à laquelle la bande sensible du détecteur commence, et f_{ISCO} la fréquence du système binaire à sa dernière orbite circulaire (dite ISCO ou encore LSO). Cette fréquence est définie à l'ordre newtonien par

$$f_{\text{ISCO}} = \frac{c^3}{6\sqrt{6}\pi G M_{\text{tot}}} \tag{15}$$

L'horizon est typiquement calculé pour un système BNS de $1.4 + 1.4$ $\text{M}_\odot$et un seuil de SNR $\rho = 8$.

Le volume comobile redshifté V_z est exprimé par unité de temps [59] et est défini par:

$$V_z = \frac{\int_{D<D_{\text{horizon}}} \frac{D^2}{1+z(D)} dD\, d\Omega\, \sin i\, di\, d\psi}{\int \sin i\, di\, d\psi} \tag{16}$$

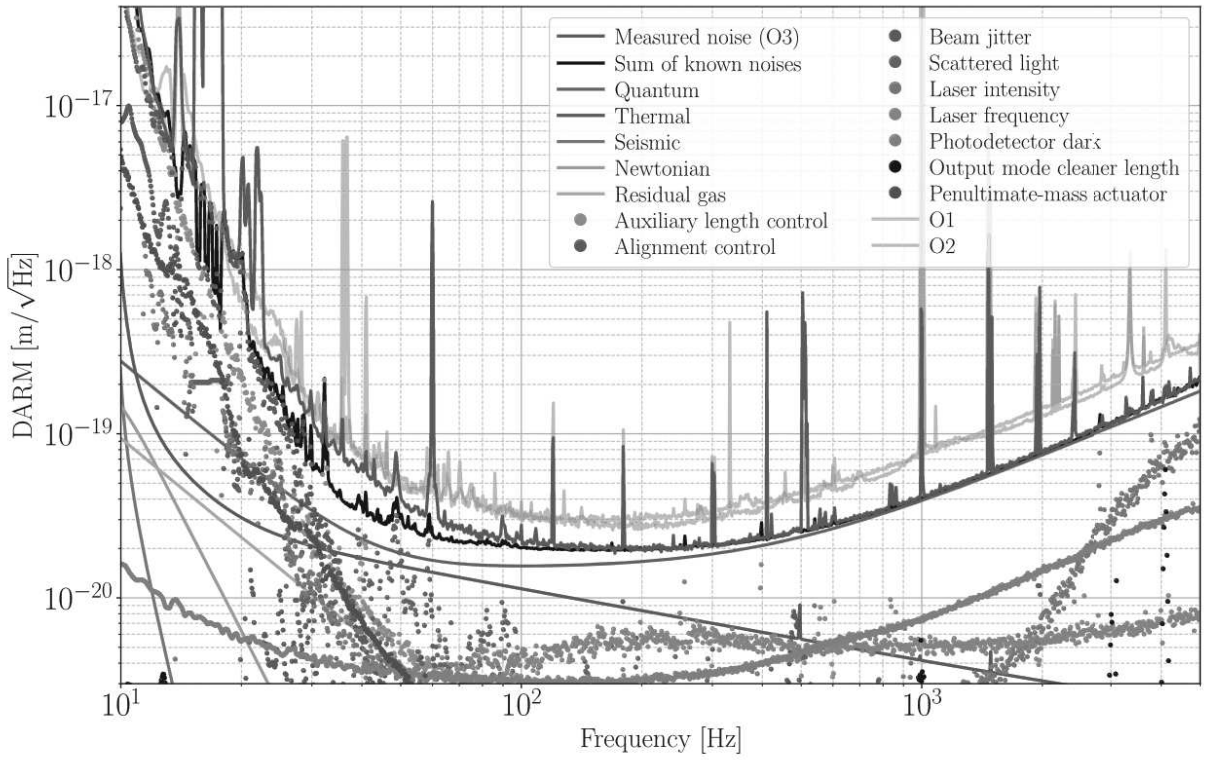
où D est la distance comobile à la source, Ω est l'angle solide dans le ciel, i est l'inclinaison du système binaire et ψ est l'orientation de la source.

Le volume V_z multiplié par le temps effectif d'observation est appelé "volume-temps" (VT). Pour des évènements de type CBC, les sources des ondes gravitationnelles sont situées à de très grandes distances, bien au-delà de l'échelle de notre galaxie. Ainsi, la distribution de ces sources peut être considérée comme isotrope. Pour une population de N_{source} sources distribuées isotropement dans une sphère de rayon D_{horizon}, seul un petit nombre N_{det} d'entre elles sera détectable car elles vérifieront $D_{\text{eff}} < D_{\text{horizon}}$. Un volume plus grand augmente donc le nombre de sources détectables. Le nombre de détections est proportionnel au VT.

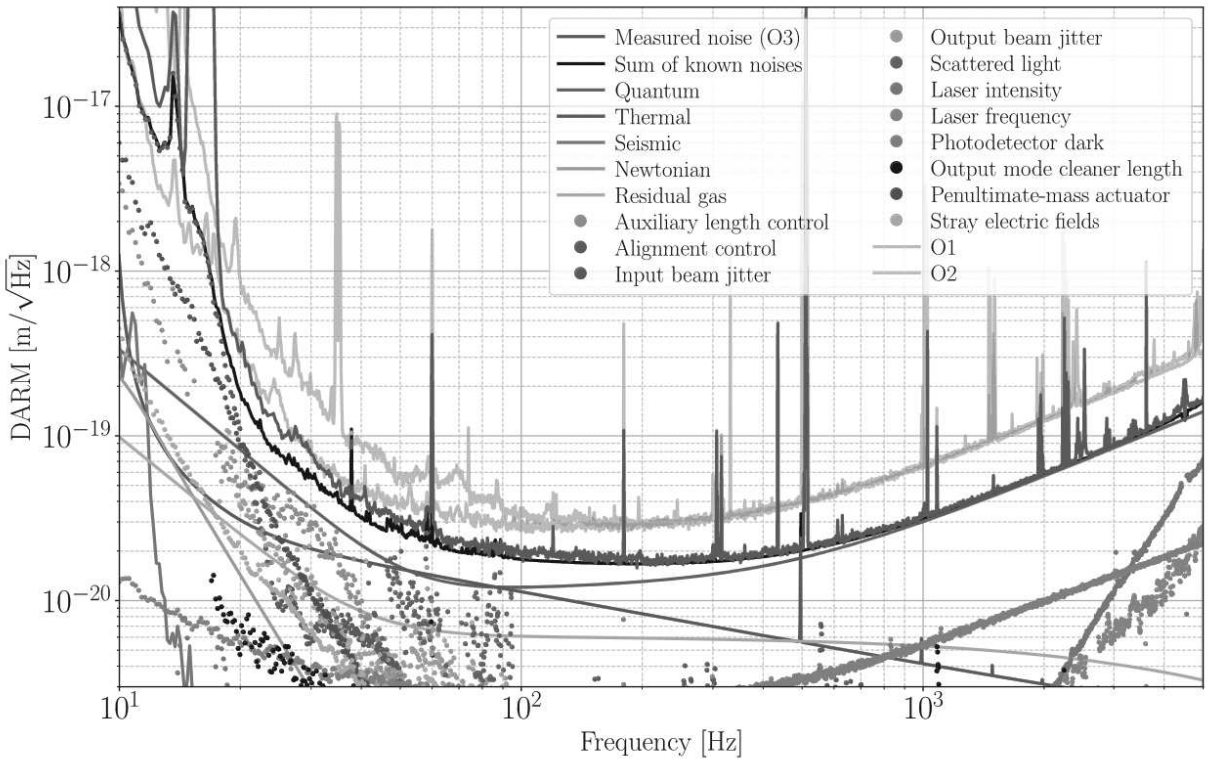
On utilise plus communément la portée du détecteur ou "*range*" pour se référer à la sensibilité des détecteurs. La portée est définie comme le rayon d'une sphère de volume V_z:

$$\frac{4}{3}\pi \times \text{range}^3 = V_z \tag{17}$$

L'horizon étant calculé pour un type de source et un seuil de SNR donné, c'est également le cas pour la portée du détecteur. Comme pour V_z, il est typiquement calculé pour un système BNS de $1.4 + 1.4$ $\text{M}_\odot$et un seuil de SNR de 8. On parle alors de "portée BNS" ou "*BNS range*". Une comparaison des sensibilités observées et anticipées pour les futures prises de données des détecteurs LIGO, Virgo et KAGRA est montrée en figure 14.



(a) LHO



(b) LLO

Figure 11: Contributions de différentes sources de bruit aux ASD des détecteurs LIGO Livingston et Hanford durant O3. Figure issue de [61].

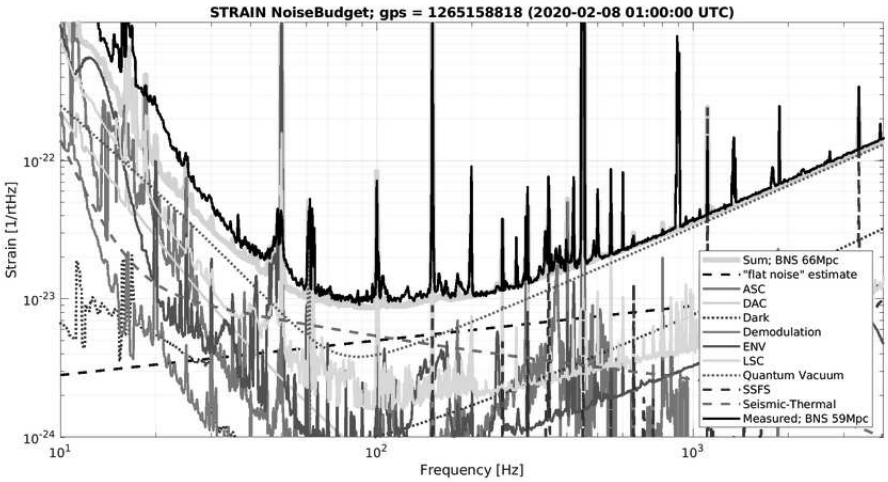


Figure 12: Contributions de différentes sources de bruit à l'ASD du détecteur Virgo pendant O3. Figure issue de [62].

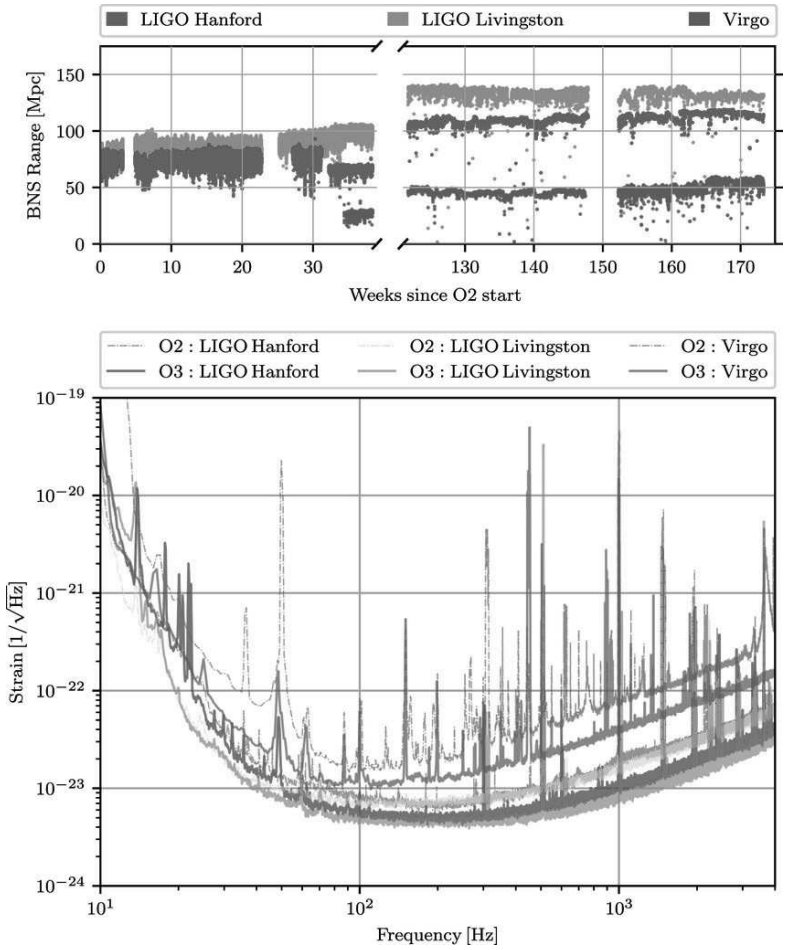


Figure 13: Comparaison des sensibilités de O3 et O2 pour les détecteurs LIGO et Virgo [63].

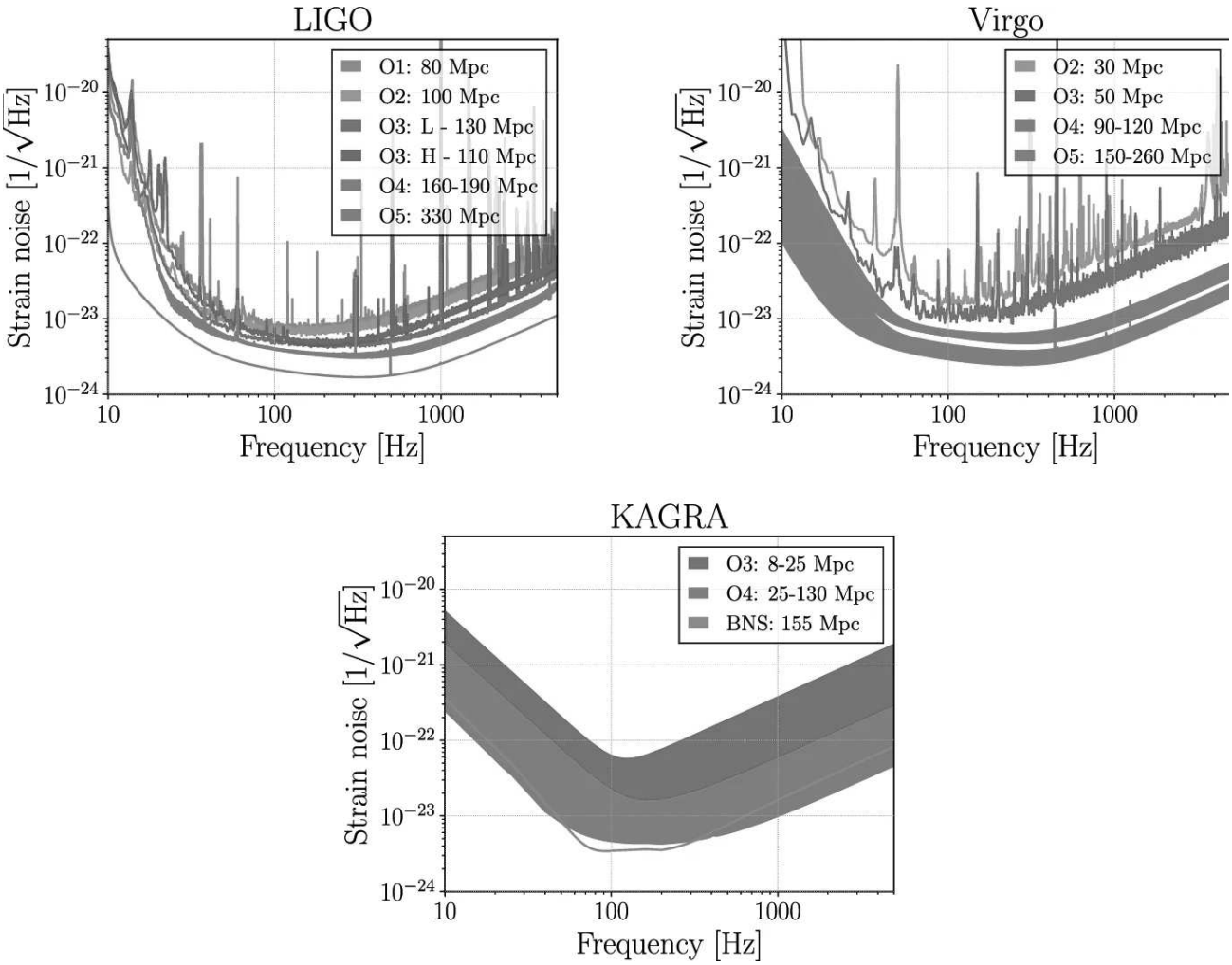


Figure 14: Evolution des sensibilités pour les détecteurs LIGO, Virgo et KAGRA [64].

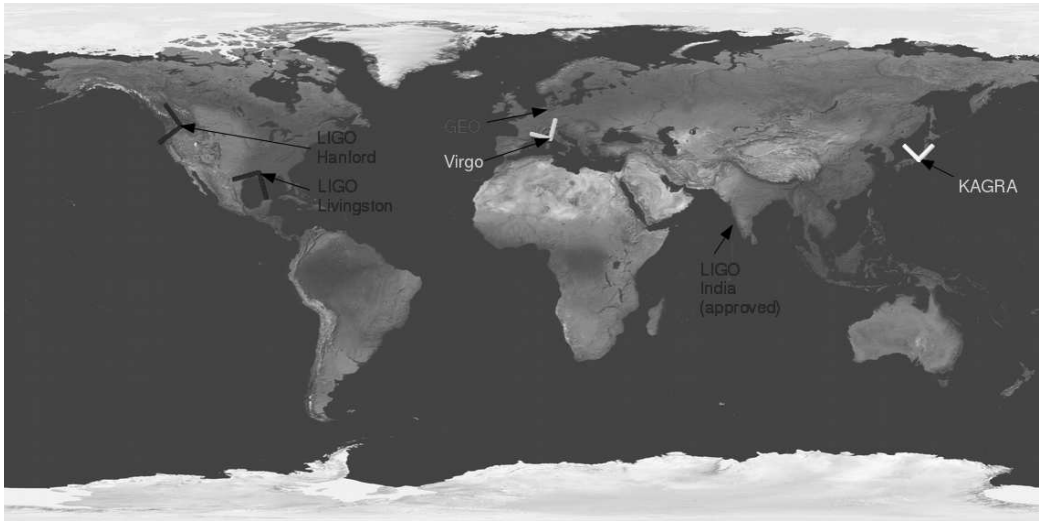


Figure 15: Carte du réseau international d'observatoires d'ondes gravitationnelles (IGWN) indiquant également l'orientation des détecteurs. Issue de site web Virgo.

3.3 Réseau de détecteurs et localisation des sources

Pendant O3, trois détecteurs d'ondes gravitationnelles ont été utilisé pour faire des détections: LIGO Hanford (H1, USA), LIGO Livingston (L1, USA) et Virgo (V1, Italie). Le détecteur GEO600 (Allemagne) observe le ciel depuis plusieurs années. Il est principalement utilisé pour faire de la recherche et du développement d'équipements dédiés aux autres détecteurs. Le détecteur KAGRA (Japon) a effectué une courte période d'observation à la fin de O3. Un troisième détecteur LIGO est également prévu en Inde. La figure 15 montre ce réseau d'interféromètre sur un planisphère.

Il y a de multiples bénéfices à avoir plusieurs détecteurs. Tout d'abord, en les plaçant à des positions différentes sur Terre avec des orientations également différentes, les diagrammes d'antennes des détecteurs vont se compléter. Cela permet d'avoir une bien meilleure couverture du ciel et d'augmenter les chances de détection.

Un autre avantage est la possibilité de rechercher des signaux en coïncidence dans plusieurs détecteurs. Dans le cas de la recherche de CBC, nous pouvons en effet supposer que le bruit des différents détecteurs n'est pas corrélé car il ne dépend que de facteurs locaux (nous ne considérons aucune source de bruit astrophysique) et les détecteurs sont largement séparés. Ainsi, un signal fort causé par du bruit dans un détecteur a très peu de chances d'être associé à un signal fort causé par du bruit dans un autre détecteur au même moment. Par contre, dans le cas d'évènements astrophysiques, les ondes traversent la Terre et tous les détecteurs sont susceptibles de les détecter. Dans ce cas on s'attend à avoir des signaux significatifs similaires dans plusieurs détecteurs à la fois (en prenant en compte le temps de vol des ondes entre les détecteurs). C'est un critère puissant qui permet de discriminer signaux astrophysiques et bruit de fond.

Un troisième bénéfice, qui est en fait le produit des deux considérations précédentes, est qu'un plus grand nombre de détecteurs permet d'avoir une meilleure localisation de la source des ondes. Pour une amplitude relative et un delai de détection entre des détecteurs, il est possible de trianguler la position la plus probable pour la source des ondes gravitationnelles. La non-détection d'un signal par l'un des détecteurs est également une information intéressante car cela indique que la source peut être localisée dans une direction à laquelle ce détecteur n'est pas sensible. Le processus de localisation est illustré en figure 16 en prenant l'exemple de GW170817. De gauche à droite

et de haut en bas: les réponses d'antennes de H1 et L1 indiquent les directions auxquelles ils sont sensibles. L'information temporelle permet de définir une ligne sur le ciel sur laquelle il est probable que la source soit positionnée. En prenant en compte les rapports d'amplitude dans H1 et L1 il est possible de dériver un contour pour la position de la source autour de cette ligne. Si on ajoute maintenant V1, qui n'a pas détecté l'évènement, on peut raffiner la recherche de la position. Comparer les SNR mesurés aux SNR attendus en fonction de la position de la source permet également d'inférer un contour plus restreint pour la position de la source.

3.4 Publication d'alertes en ligne

Dans le cas de GW170817, une information supplémentaire était disponible pour localiser la source. En effet, une contrepartie électromagnétique fut détectée. La figure 17 résume les informations données par les divers observatoires. Cette détection conjointe est rendue possible grâce au systeme d'alertes publiques émisent par la collaboration LVK et utilisée par de nombreux observatoires dans le monde.

Les chaînes d'analyse utilisent un seuil de taux de fausse alarme (FAR), typiquement de l'ordre de un par deux heures, pour envoyer leurs candidats à la base de donnée commune GraceDB. Le FAR est le taux d'évènements de bruit de fond que l'on attend avec une statistique de classement supérieure à un seuil donné. Les candidats enregistrés par GraceDB sont rangés dans des structures appelées "G-event" avec certaines propriétés rapportées par la chaîne d'analyse. Si plusieurs chaînes ont détecté un même évènement, un "Superevent" est créé et rassemble les différents G-events associés à la détection. Les propriétés listées dans le superevent sont celles du G-event dit "préféré", choisi comme celui ayant le plus grand SNR. Les notices contiennent de nombreux paramètres comme le type et le temps de l'alerte, la chaîne d'analyse qui a rapporté la détection, les détecteurs impliqués, la carte du ciel et les diverses probabilités cité plus bas dans cette section.

Le FAR des candidats est utilisé pour définir leur significativité. Lorsque le FAR d'un candidat rapporté par au moins l'une des chaînes d'analyse dépasse un certain seuil, fixé par la collaboration, le candidat est rendu public (au-delà de la collaboration LVK) et des circulaires et notices sont envoyées aux observatoires qui en ont fait la demande. Plusieurs chaînes d'analyse CBC (MBTA, PyCBC, GstLAL, SPIIR, cWB_BBH, RAVEN) peuvent participer aux détections. Pour prendre en compte cette multiplicité, le bruit qu'elles détectent est supposé indépendant et donc les seuils par chaîne d'analyse valent le seuil global divisé par le nombre de chaînes d'analyses CBC. On parle de "*trial factor*".

Durant O3 le seuil de FAR global était défini comme un tous les deux mois pour les candidats CBC. Comme il y avait cinq chaînes d'analyse CBC, le seuil par chaîne d'analyse était donc de un tous les dix mois.

Pour O4 deux seuils globaux sont considérés:

- FAR<1 par mois pour les candidats significatifs,
- 1 par mois<FAR<2 par jour pour les candidats faiblement significatifs.

Cela donne respectivement, pour les FAR calculés par les chaines d'analyse, des seuils à 1 par 6 mois et 1 par trois jours après application du *trial factor*. Le seuil pour les recherches de bursts est quant à lui de 1 par an avec un *trial factor* de 1/3 pour prendre en compte les trois chaînes d'analyse. Les candidats significatifs sont validés par les experts des chaînes d'analyse et de la caractérisation des détecteurs ayant participés à la détection sous la forme d'une circulaire. Cela

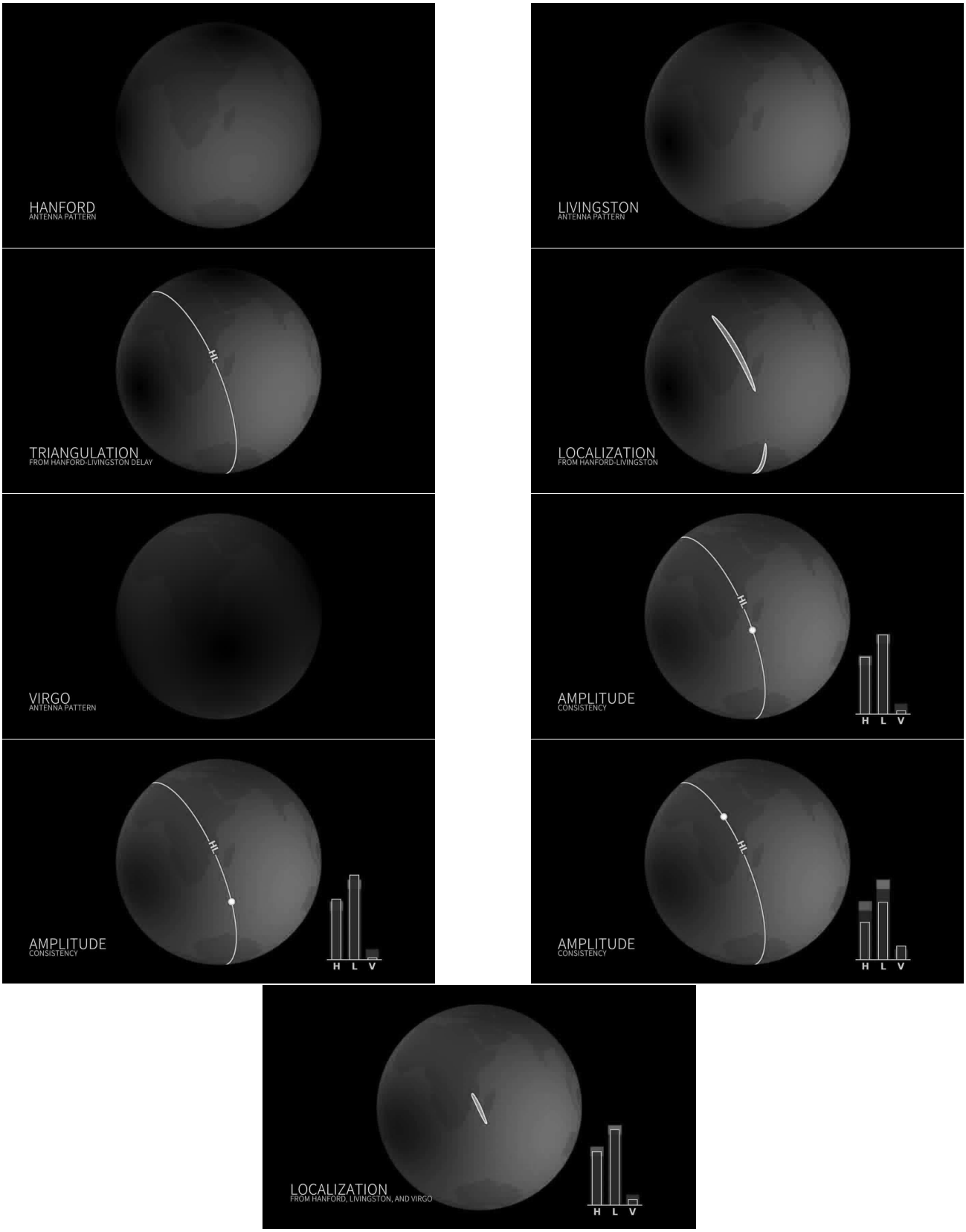


Figure 16: Illustration du processus de localisation d'une source d'ondes gravitationnelles, appliquée à GW170817. Voir le texte pour une description. Crédit: Leo Singer (format original disponible à `https://dcc.ligo.org/LIGO-G1702012/public`).

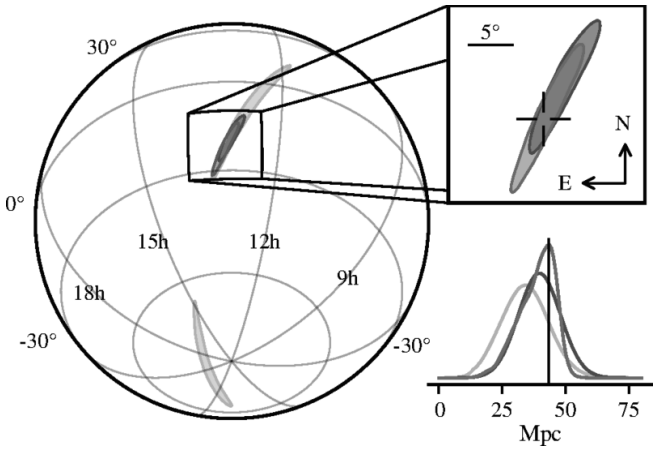


Figure 17: Reconstruction de la position de la source dans le ciel pour GW170817 issue de [65]. L'aire bleu-clair (190 $\deg^2$) est calculée par un algorithme de réponse rapide en n'utilisant que H1 et L1. La bleu-foncé (31 $\deg^2$) inclut V1. Le contour vert (28 $\deg^2$) inclut les trois détecteurs, il est donné par un algorithme avec une latence plus élevée que le précédent. Le réticule dans le panneau en haut à droite indique la position de la galaxie hôte identifiée. Le panneau en bas à droite donne la distribution postérieure pour la distance de luminosité.

permet de confirmer le potentiel de l'évènement ou bien de retirer l'alerte si un expert juge que l'évènement n'est pas d'origine astrophysique.

Pour O4 une catégorie supplémentaire d'évènements existe: les candidats dits de "haut profil" auxquels une attention particulière est portée. Il s'agit de candidats significatifs qui satisfont l'une des conditions suivantes:

- Une contrepartie multi-messager a été détectée,
- Le G-event avec le plus faible FAR a été téléchargé par une chaîne d'analyse de burst,
- Le G-event préféré a une probabilité d'origine astrophysique (voir section 4.12 pour une définition de ces quantités) supérieure à 0.5 et:
  - soit la probabilité qu'il s'agisse d'une coalescence BNS est supérieure à 10%,
  - soit la probabilité qu'il s'agisse d'une coalescence NSBH est supérieure à 10%,
  - soit la probabilité qu'il y ait une masse résiduelle après la fusion est supérieure à 10%,
  - ou bien le contour à 90% de niveau de confiance pour la position de la source est plus petit que 100 $\deg^2$.

Plus d'informations sur le système d'alertes peuvent être trouvées dans le guide d'utilisateur fourni par la collaboration LVK: `https://emfollow.docs.ligo.org/userguide/`.

3.5 Autres détecteurs d'ondes gravitationnelles

D'autres détecteurs d'ondes gravitationnelles sont proposés avec des stades de progression plus ou moins avancés. Leur modèles peuvent donc encore être amenés à changer.

Détecteurs de troisième génération

Le "Einstein Telescope" (ET) [66] est une proposition d'interféromètre souterrain dix fois plus sensible que les détecteurs de seconde génération LIGOs et Virgo. Deux propositions ont été formulées concernant sa forme: triangulaire ou en forme de L (à l'instar des LIGOs/Virgo) avec des bras bien plus grand que les détecteurs acutels. Deux sites sont pour le moment considérés: la Sardaigne et l'Euroregio Meuse-Rhin.

Cosmic Explorer (CE) est une proposition de détecteur aux USA. Il est prévu en forme de L avec des bras de 40 km et une sensibilité dix fois plus grande que les détecteurs Advanced LIGO.

ET et CE permettraient des mesures extrêmement précises des paramètres des sources, ce qui contraindrait de nombreux modèles. Il y a également des attentes concernant la première détection directe de bruit de fond stochastique d'ondes gravitationnelles dans le domaine de fréquence de ces interféromètres (quelques Hertz à quelques kiloHertz) et un raffinement des mesures de LIGO et Virgo.

LISA

Le "Laser Interferometer Space Antenna" (LISA) [67] est une proposition d'observatoire basé dans l'espace. Son but serait d'étudier les basses fréquences entre $0.1\,\mathrm{mHz}-0.1\,\mathrm{Hz}$. Son design consiste en trois satellites formant un triangle. Chacun abrite deux masses libres et deux lasers qu'ils pointent vers leurs voisins. Les satellites seraient séparés d'environ ~ 2.5 million km. Les variations relatives de distance entre les masses libres situées dans les satellites sont mesurées par interférométrie laser hétérodyne.

Pulsar Timing Array

Le réseau de synchronisation de pulsar international, ou *International Pulsar Timing Array* (IPTA), est un détecteur d'ondes gravitationnelles "naturel" exploitant les émissions périodiques de pulsars connus [68]. Le passage d'une onde gravitationnelle causera un delai dans le temps de vol des impulsions émises par les pulsars, corrélées avec la position dans le ciel de ces derniers permettant la détection. L'éloignement des pulsars étant bien plus grand que les bras des détecteurs terrestres, IPTA pourra sonder des fréquences très faibles allant jusqu'au nanoHertz.

La collaboration NANOGrav, membre du consortium IPTA a récemment publié des premières observations de l'existence d'un bruit de fond stochastique d'ondes gravitationnelles [69]. Une corrélation entre 67 pulsars a été observée en utilisant 15 ans de données collectées par NANOGrav. Cette corrélation suit le modèle de Hellings-Downs, attendu pour un signal stochastique d'ondes gravitationnelles. Le spectre et l'amplitude observés correspondent aux attentes pour une population de binaires de trous noirs supermassifs.

4 Searching for CBC with MBTA

Contents

The Multi-Band Template Analysis (MBTA) [70] is one of the search pipelines running on the LIGO and Virgo data streams. It uses the matched filtering technique to search for CBC signals, both online at low-latency and offline. This technique, detailled in section 4.3, consists in computing the cross-correlation of the data with templates which are predictions of the signal being searched for. The result of this operation, called also filtering, is the Signal-to-Noise Rartio (SNR) time series, which becomes large at the time of an event. Since we don't know the parameters of the source, this operation is performed on multiple templates, the template bank. The goal is to find the template which maximizes the SNR in order the find the source parameters and get a candidate, or trigger, if the SNR is above a given threshold.

MBTA filters the data at a low computational cost by splitting the filtering in two separate frequency bands. This splitting has two benefits. First, it allows a tuning of the sampling frequency differently and template duration for each band, effectively reducing the time needed to compute the Fast Fourier Transforms (FFT) needed for the filtering. Second, it significantly reduces the number of templates needed to cover the parameter space. The result of the matched-filtering in

each band is then coherently combined yielding the time series for the full frequency band. But a search pipeline is not just the matched filtering. There are preprocessing steps as well as post processing, to clean the data and evaluate the quality of the candidates. This chapter describes all the steps of the MBTA pipeline.

4.1 Overview of the MBTA pipeline

A schematic view of the pipeline is given in figure 18. We describe here the main search of the pipeline dedicated to BNS, BBH and NSBH detection. The online version for more specific searches (sub-solar mass search and early-warning) and offline version follow the same logic, with a few changes for some parameters like the sampling rate of analysis.

MBTA takes the data of the LIGO and Virgo detectors to search for CBC signals. There are two pre-processing steps before MBTA analyzes the data: the data coming from LIGO and Virgo are resampled by MBTA from 16 384 Hz to 4096 Hz. Then MBTA applies what is called the gating. The gating is a noise rejection tool which removes loud glitches (transient noises) from the time series. The gating is described in more details in section 4.2.

The data is then broadcasted to the machines that will do the filtering. The filtering is done in parallel with 43 jobs (job 90 is specific to very short templates, see section 4.7.2) plus an additional one without gating (named region 4 or R4, see section 4.2). Each of these jobs analyzes the data with a fraction of the template bank (and therefore a fraction of the parameter space) sorted by template duration. This sorting allows to share them among jobs in such a way that most real templates are used by only one job. Jobs filtering using long templates are given less templates to balance their computing usage.

The filtering is done using MBTA's own template banks. The templates used for the matched-filtering are derived from accurate models of general relativity. The template banks are built in order to have a loss of at most few percent in SNR (with respect to the optimal SNR) when detecting a signal. They will be described in section 4.7.

If at some point the SNR is greater than a pre-defined threshold we say that we have a single detector trigger and the event will be saved. During O3, the time of the trigger was taken as the time of the end of the template. For O4 it is taken as the time of the maximum amplitude, the difference being the ringdown time.

The MBTA pipeline also uses data quality monitoring and a variety of noise rejection tools to either downgrade the SNR of triggers or straight off reject them to better discriminate astrophysical signals from background triggers. Some are applied during the filtering and others at the post-filtering step.

The pipeline then searches for coincident triggers in different detectors, meaning observed with the same template and matching time accounting for the time of flight of gravitational waves between each pair of detectors.

An astrophysical signal or loud noise may match with several templates thus creating several triggers. To avoid this redundancy, triggers that are close in time are clustered (more on this in sections 4.8 and 4.10).

The pipeline also computes a probability of astrophysical origin, a source classification and a False Alarm Rate (FAR) (see sections 4.11 and 4.12). Significant triggers are astrophysical candidates. If their FAR is below a given threshold, they are uploaded to the Gravitational-Wave Candidate Event Database (GraCEDb) and online public alerts may be issued.

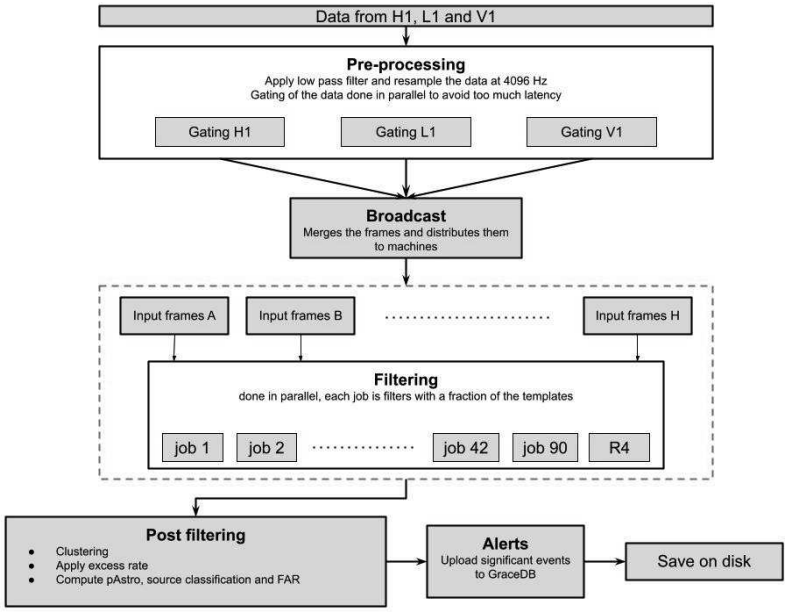


Figure 18: Schematic of the processes and workflow of the online MBTA pipeline.

4.2 Gating

Strain data sent to the pipeline are accompanied by a state vector. This vector indicates whether the data can be safely analyzed. Offline, this role is filled by the category 1 (CAT1) flags [71, 72] which indicate times at which major issues took place. These times are not analyzed.

Yet, not all noisy periods are excluded by such vetoes. For this reason, pipelines have to use a variety of noise rejection tools in order to control their background. The gating is one of them. We present here the gating as it was used during O3. More details on the development of the gating can be found in [73]. The gating is a tool used to reject loud glitches. It works by applying a gate function, hence the name, to set the $h(t)$ time series to 0. Simply ignoring the data at the time where the glicth was identified is not enough because due to the finite size of the FFT, loud glicthes could spoil the full time series of one FFT. It could therefore contaminate the data even after the glitch, it is thus mandatory to nullify the data.

Figure 19 shows the application of gating around a glitch in L1. The gating process takes place before MBTA starts analyzing the data. The criteria to apply the gating is based on the BNS range of the detector which follows the variations of the PSD of the detector. If at some point the BNS range drops below 60% of the median BNS range over the last 10 seconds, the gate function is applied. This threshold is updated every second. The range is computed for a BNS signal from 10 Hz to 2048 Hz. The PSD used to compute the range is actually the maximum of the value of the current PSD and the median PSD. This biases the range and a 1.2 empirical factor has to be applied to correct it. To smooth the 1 to 0 and 0 to 1 transitions, the chosen gate function is a Tukey window which goes from 1 to 0 in 0.315 s. The range is computed at 32 Hz to follow precisely the rapid fluctuations of the PSD due to glitches. Figure 20 shows the evolution of the range over time with some gated times.

For O4 some changes were applied to the gating to resolve some noise issues encountered with the early warning search. Excess of triggers were caused by noise at low frequency to which the standard gating is not sensitive. This was solved by computing the range for the gating of the early

warning search from 10 Hz to 37 Hz and lowering the threshold to trigger the gating from 60% to 50%. Other tuning of the gating parameters were done for all searches like the enlargement of the tapering window time to 0.5 second.

Another issue occured during times were the range experienced frequent drops. This caused the threshold for the gating to fluctuate a lot and was characterized by excess of triggers. To keep track of the range reductions, a small fraction of the PSD computed during gated times are kept. If over 1 s, more than 50% of time is gated, the 4 s that follow are also gated.

To avoid cutting out very short and loud astrophysical signals that could trigger the gating, a search without gating is also run with a much higher SNR threshold and a reduced parameter space. This search is called region 4. During O3 the region 4 was defined as all region 3 templates (see section 4.7.1) with duration smaller than 2.7 s (starting at 21 Hz). A ranking statistic threshold of 12 was set for triggers from this region. The ranking statistics quantifies and ranks the loudness of events, it is defined in section 4.8.3) The region 4 for O4 is defined as any template with a duration shorter than 6 s (see section 4.7.2), that trigger the gating with an SNR of at most 25 computed with the expected O4 L1 "high-sensitivity" PSD.

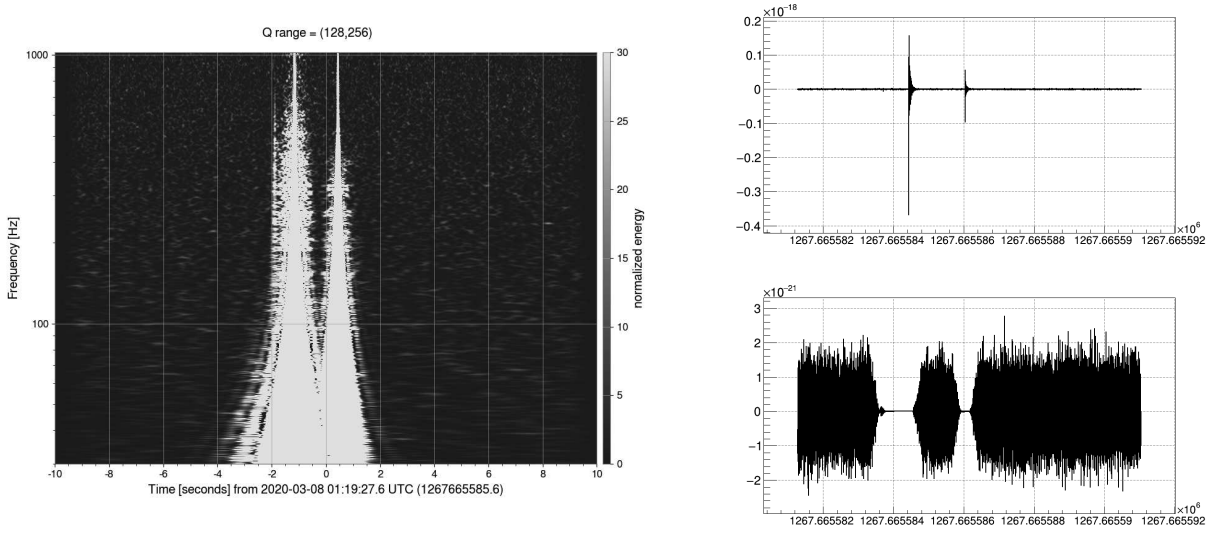


Figure 19: Left: Spectrogram around two glitches in L1. Right: strain before (top) and after (bottom) gating around these glitch.

4.3 Matched Filtering

The matched filtering technique is the optimal detection strategy in the presence of stationary noise [74]. The performance of the filter is caracterized by the signal-to-noise ratio at its output.
Let $n(t)$ be the stationary Gaussian noise for a detector.
The signal at the output of the detector is

$$h(t) = n(t) + s(t) \tag{18}$$

where $s(t)$ is the possible astrophysical signal.

The principle of the matched-filtering technique is to compute the correlation of the detector output $h(t)$ with expected GW waveforms called templates, weighted by the PSD. If a signal

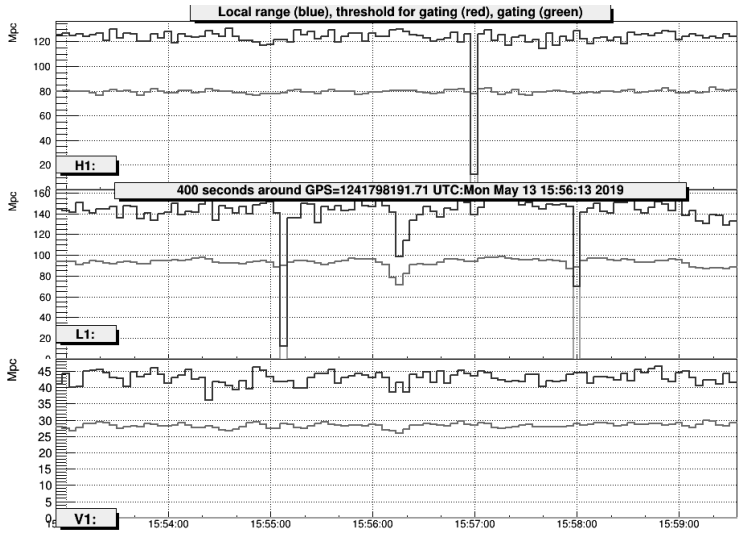


Figure 20: Example of the range (blue) and threshold (red) for gating over time for a segment of O3. The green line indicates gated times.

matching the template is mixed with the detector noise, then the correlation will be high. To increase the chances of detection, many templates are prepared in advance for the filtering in a way that covers a large parameter space.
From [75], we can write the filter output for the correlation between $h(t)$ and a template $s_{\text{template}}(t)$ in the presence of signal as

$$M(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{h}(\omega)\tilde{s}^*_{\text{template}}(\omega)e^{i\omega t}d\omega \tag{19}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{n}(\omega)\tilde{s}^*_{\text{template}}(\omega)e^{i\omega t}d\omega + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{s}(\omega)\tilde{s}^*_{\text{template}}(\omega)e^{i\omega t}d\omega \tag{20}$$

$$= \nu(t) + \eta(t) \tag{21}$$

$$\tag{22}$$

where $\nu(t)$ is the filter output for the noise and $\eta(t)$ the output for the signal we are looking for. The noise being a random process it follows from Wiener-Khinchin theorem [75] that

$$\overline{n^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_n(\omega)e^{i\omega t}d\omega \tag{23}$$

where $\overline{n^2}$ is the average noise power and $S_n(f)$ the PSD of the noise. Therefore the noise power at the output of the filter will be

$$\overline{\nu^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|\tilde{s}^*_{\text{template}}(\omega)\right|^2 S_n(\omega)e^{i\omega t}d\omega \tag{24}$$

We define the Signal-to-Noise Ratio (SNR) time series for a signal $h(t) = n(t) + s(t)$ (with $n(t)$

stationnary and Gaussian) and a template $s_{\text{template}}(t)$ as

$$\rho(t) = \frac{\eta^2(t)}{\overline{\nu^2}} \tag{25}$$

$$= \frac{1}{2\pi} \frac{\left| \int_{-\infty}^{+\infty} e^{i\omega t} \tilde{s}^*_{\text{template}}(\omega) \tilde{s}(\omega) d\omega \right|^2}{\int_{-\infty}^{+\infty} \left| \tilde{s}^*_{\text{template}}(\omega) \right|^2 S_n(\omega) e^{i\omega t} d\omega} \tag{26}$$

Our goal is to have maximal SNR, to this end we search the template which achieves this. Using Cauchy-Schwartz inequality for square-integrable complex-valued functions, we can write

$$\left| \int_{-\infty}^{+\infty} e^{i\omega t} \tilde{s}^*_{\text{template}}(\omega) \tilde{s}(\omega) d\omega \right|^2 \leq \int_{-\infty}^{+\infty} e^{i\omega t} \left| \tilde{s}^*_{\text{template}}(\omega) \right|^2 d\omega \int_{-\infty}^{+\infty} e^{i\omega t} \left| \tilde{s}(\omega) \right|^2 d\omega \tag{27}$$

the maximal SNR value is therefore reached when we attain equality, that is when the template is proportional to the signal: $\tilde{s}_{\text{template}} = \text{constant} \times \tilde{s}(\omega)$. In practice the noise of the detector is neither Gaussian nor stationnary and the template does not reproduce perfectly the signal but we can achieve high precision with fine enough grid of templates (see section 4.7).

Since the phase of the gravitational wave at merger time Φ_0 is not known we want to maximize the SNR over the phase. To this end the SNR time series is decomposed into in-phase h_P and in-quadrature h_Q components

$$\rho(t) = \rho_P(t) \cos \Phi_0 + \rho_Q(t) \sin \Phi_0 \tag{28}$$

such that $\rho_P = \rho(\Phi_0 = 0)$, $\rho_Q = \rho(\Phi_0 = \pi/2)$. This decomposition is related to the polarization of the wave as at merger time ρ_P and ρ_Q are equal to the cross and plus polarizations respectively. In order to maximize over the phase we have to solve $\rho'(\Phi_{0,best}) = 0$. The solution to this equation is

$$\Phi_{0,best} = \tan^{-1} \frac{\rho_Q}{\rho_P} = arg(\rho_P + i\rho_Q) \tag{29}$$

Putting this back into the expression of $\rho(t)$ yields

$$\rho^2(\Phi_{0,best}) = \rho_P^2 + \rho_Q^2 \tag{30}$$

This leads us to define the complex SNR time series

$$Z = \rho_P + i\rho_Q \tag{31}$$

and the SNR is taken as

$$\text{SNR}^2 = |Z|^2 = \rho_P^2 + \rho_Q^2 = \rho^2(\Phi_{0,best}) \tag{32}$$

We call matched-filtering output (MFO) the SNR time series. In practice we naturally deal with discrete series instead of integrals. Figure 21 shows the matched filtering output for a BNS astrophysical signal.

4.4 PSD computation

The PSD used in the matched filtering should be accurately estimated to have a proper estimation of the background and therefore of the significance of any trigger. Since high frequency (HF) and low frequency (LF) templates have different duration (the low frequency part of the signal is much

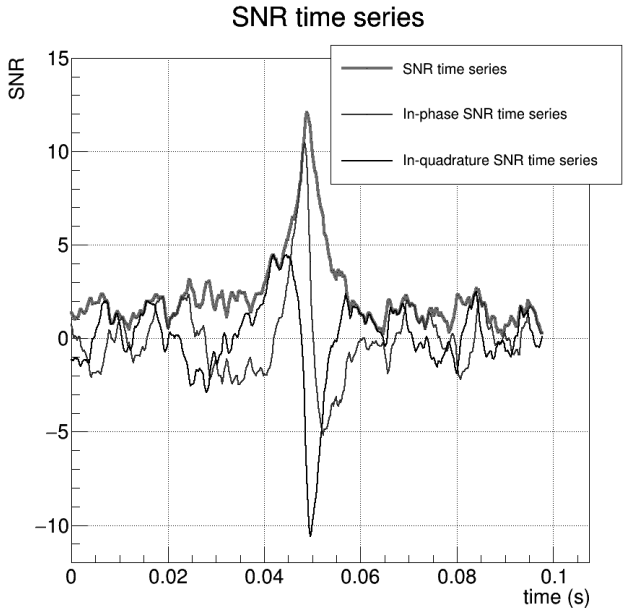


Figure 21: Matched filtering output for a BNS event (GW190425) with SNR 12.07 in L1

longer), the PSD is computed separately for each frequency band. The computation in MBTA is done using the FFTW3 [76] implementation of the "real to half-complex" FFT. To limit the bias due to possible noise fluctuations, the PSD used for the matched filtering is the median PSD computed over a user-defined number of FFT. When running MBTA the PSD is first initialized either by loading the Amplitude Spectral Density (ASD) from a previous MBTA run or by loading a detector strain data file and computing the PSD using this file. Note that some additional steps can be done when computing the PSD:

- smoothing of the ASD by averaging some frequency bins together,
- canceling of some frequency bands if some are known to be noise-polluted: we know for instance that right after the lock of an interferometer, the suspension fibers' transversal mechanical modes of vibration (so-called "violin modes") can cause a lot of triggers. Canceling the frequency band (like $512\,\mathrm{Hz} \pm 7.5\,\mathrm{Hz}$ in L1) associated to those modes avoids bad triggers at those times.
- application of a high-pass shaping filter, a common one being a Butterworth filter of order 4 and cut-off frequency 40 Hz for instance. A high-pass filter reduces the signal dynamic and limits numerical issues which would pollute the FFT. The filter is applied in time domain to the strain data, the inverse filter is applied in frequency domain.

Estimations of the SNR^2 sharing between the bands and of the low/high frequency cutoffs to have close to 100% of the SNR are also done at initialisation of the PSD.

Since the noise in the LIGO and Virgo detectors is not sationnary, the PSD computed at initialization will not properly caracterize the noise at a later time and it should therefore be updated over time, as shown by figure 22. As a typical example, during the test performed to prepare the O4 online configuration, FFTs were computed every 4 seconds over 87.5 s for the LF band and over 8 s for the HF band. The median PSD was computed over 4000 s for the LF band

and 1000 s for the HF band. This means roughly over 90 and 250 FFT respectively (acounting for an overlap of 50% between consecutive FFTs). The computation of the median PSD can be very time consuming and was therefore updated every 18 (50) FFT for the LF (HF) band. Meaning that we compute a new median each time 1/5 of the FFTs were updated. Since the PSD does not fluctuate so fast this choice allows to have a good accuracy in accounting the noise fluctuations while not being too computationally expensive.

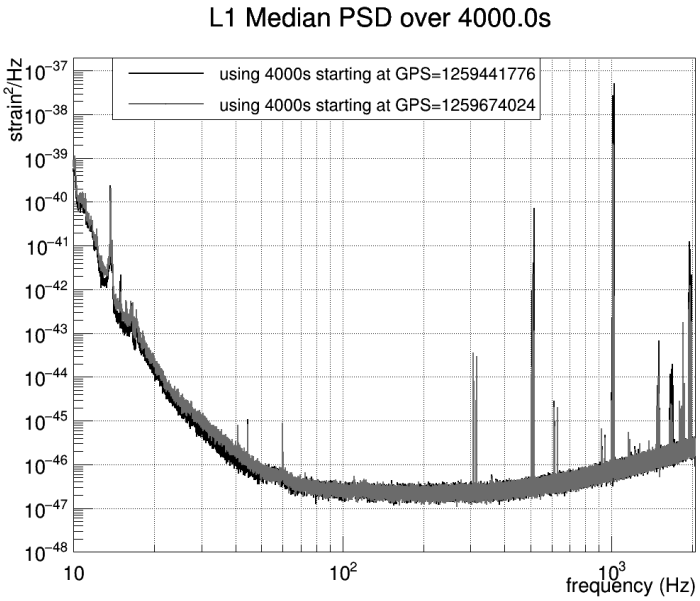


Figure 22: Difference in the PSD of LIGO Livingston computed over 4000 s few days apart.

4.5 Frequency range of the search

To detect signal in the $h(t)$ data stream we first need to decide the frequency range we want to probe. The larger the frequency band of the search, the higher the chance to find a signal because the SNR integrated over the band will be higher (for an actual astrophysical event). But a wider band also requires more computing time. Furthermore, most of the SNR is collected in a limited frequency band.

To choose the start and end frequency of the search, we compute the SNR loss for several values of them. The computation depends on the PSD of the detector and the mass of the system considered since the maximum frequency of the signal decreases when the mass of the binary increases. Figure 23 shows the SNR loss as a function of the starting and end frequency of the search, in H1 and L1, for a $1.4\,M_\odot$+$1.4\,M_\odot$ BNS system. Figure 24 shows the same for several masses computed at the same O4 time. The plots were produced using O3 strain data from GPS = 1268100004 to GPS = 1268100800 (March 13 2020) and O4 strain data from GPS = 1370500020 to GPS = 1370500800 (June 11 2023).

Usually, the starting frequency is chosen at 24 Hz. We see that going below would only allow to recover less than 1% of the SNR. It would however result in higher computing cost and is therefore not interesting. This may change for future runs with the expected detector improvements at low frequency.

Figure 25 shows that the frequency at the last stable orbit for low mass systems can go up to a few kilohertz. Although figure 23 shows that stopping at 1024 Hz would allow to recover the quasi-totality of the SNR, the end frequency of the search is typically chosen at 2048 Hz. This an empiric choice motivated by the fact that extending the search to a higher frequency yielded better results in terms of timing resolution.

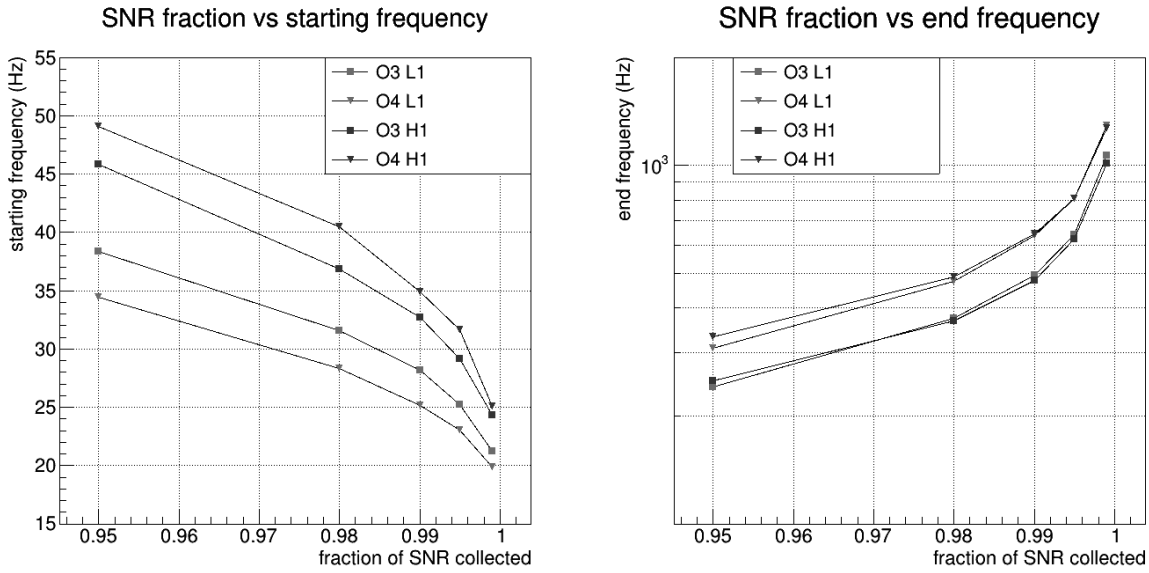


Figure 23: Fraction of SNR collected as a function of the starting frequency (left) and end frequency (right) of the search, respectively assuming an end frequency of 2048 Hz and start frequency of 10 Hz.

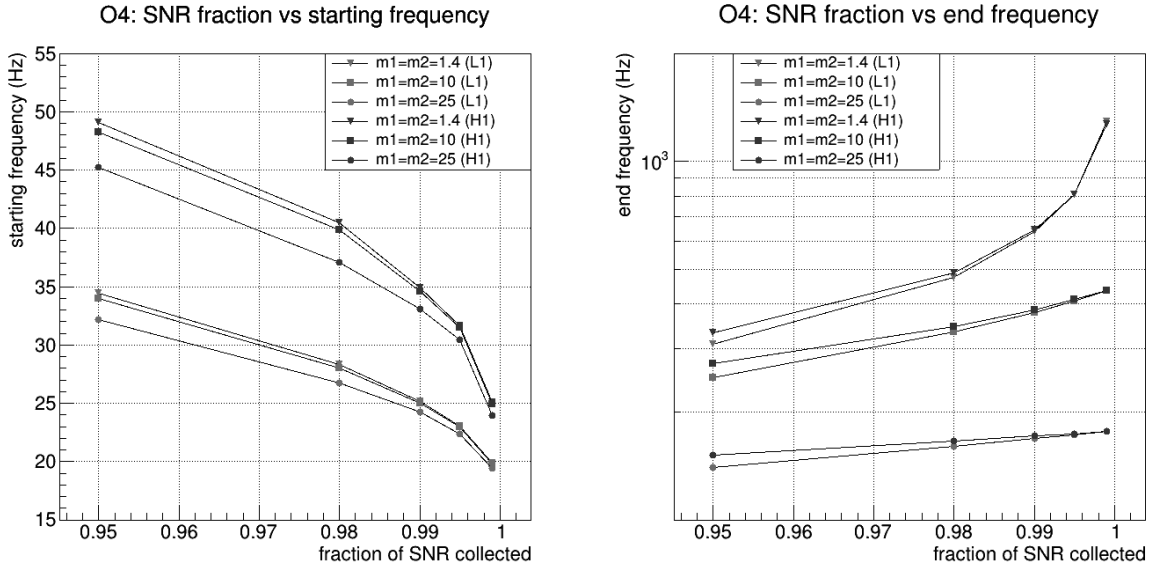


Figure 24: Fraction of SNR collected using an O4 PSD for several masses as a function of the starting frequency (left) and end frequency (right) of the search, respectively assuming an end frequency of 2048 Hz and start frequency of 10 Hz.

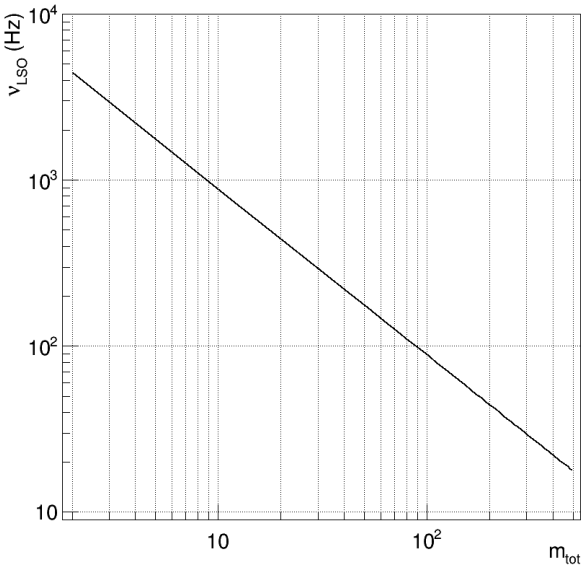


Figure 25: Signal frequency at the last stable orbit as a function of the total mass (see eq. 15).

4.6 Multi-band analysis

As mentionned previously, the filtering is done in two separate frequency bands. The splitting frequency can be chosen such that the SNR squared is roughly equally shared between the two bands. This is typically achieved by splitting frequencies around 100 Hz for a BNS system. In practice, an empiric choice is to generally have slightly more SNR in the high frequency band with a splitting frequency at 80 Hz. This is because the search is run with the same splitting frequency for a large part of the parameter space. Having a bit more SNR in the HF band when computed for BNS, provides more high mass templates with appropriate signal in the HF band. Another consideration is the impact of the different shape of the detector sensitivity, which means different SNR per frequency band. Since the spliting frequency must be the same for all detectors in the current MBTA implementation, this splitting is done on the most sensitive detector, L1 for O3 and O4. For instance during the first weeks of O4 the expected SNR^2 sharing was 46% for the HF band and 54% for the LF band in L1, and 30% for the HF band and 70% for the LF band in H1. Figures 26 to 28 show the SNR sharing for singles detector triggers obtained during an analysis of data with injected simulated signals (see section 5.3), for O3 single detector triggers respectively and for triggers obtained on Gaussian noise. For O3 single detector triggers, the SNR seems larger in the low frequency band.

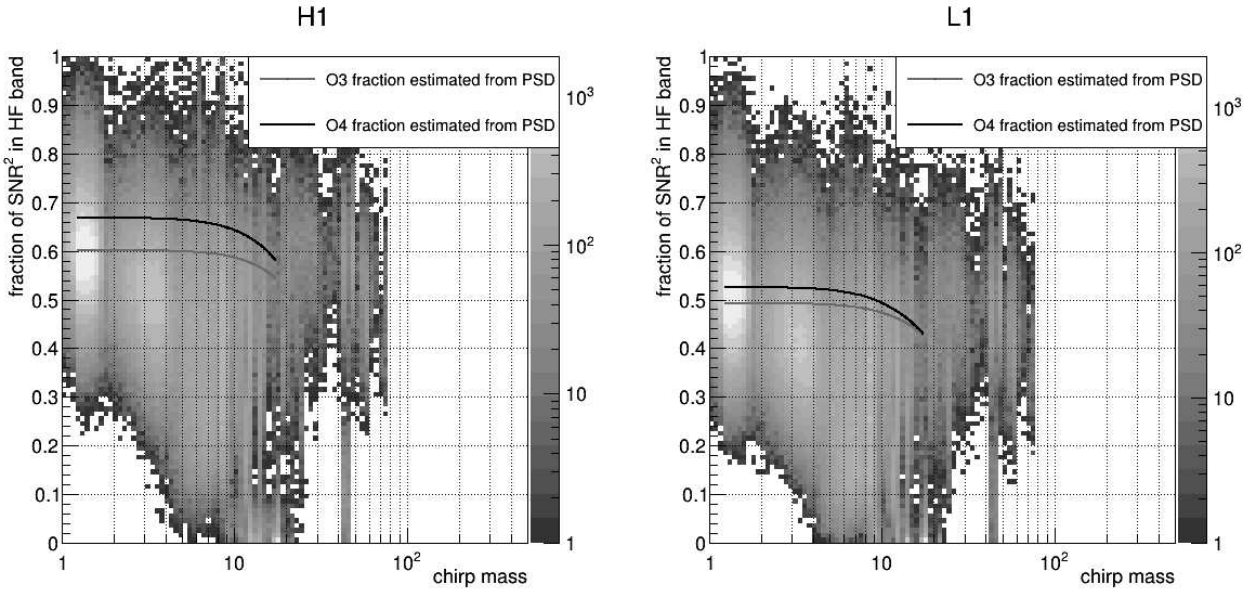


Figure 26: Fraction of SNR^2 in the HF band as a function of the chirp mass for O3 common injections on top of O3 data (section 5.3) analyzed with the pipeline O3 configuration. The expected fraction is also plotted for O3 and O4.

Summarizing the discussion on the starting, end and split frequencies, the bands are chosen as:

- a low frequency band ranging from 24 Hz to 80 Hz,
- a high frequency band ranging from 80 Hz to 2048 Hz.

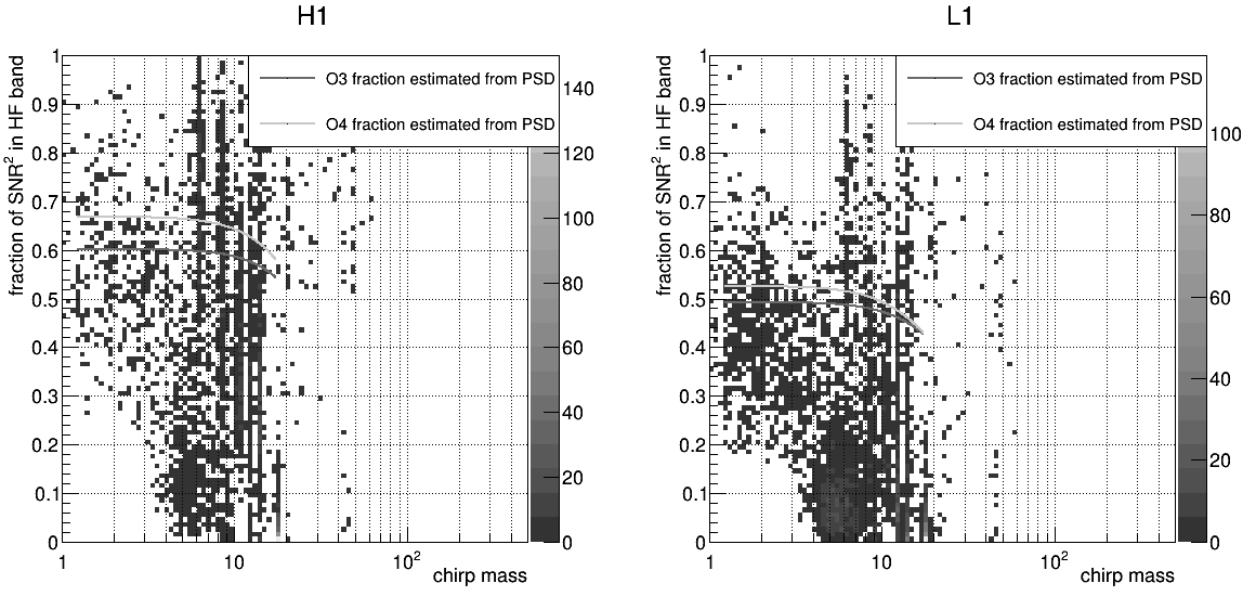


Figure 27: Fraction of SNR^2 in the HF band as a function of the chirp mass for O3 single detector triggers analyzed with the pipeline O3 configuration having a rwSNR larger than 8. The rwSNR, defnined in section 4.8.2, is a modified SNR. The expected fraction is also plotted for O3 and O4.

4.6.1 Combining bands

The result of the matched filtering in the two frequency bands is coherently summed following

$$\mathrm{SNR}(t, \mathrm{template}) = \int_{\omega_{min}}^{\omega_{max}} \tilde{s}(\omega)\tilde{h}^*_{\mathrm{template}}(\omega)e^{i\omega t}d\omega \tag{33}$$

$$= \int_{\omega_{min}}^{\omega_{split}} \tilde{s}(\omega)\tilde{h}^*_{\mathrm{template}}(\omega)e^{i\omega t}d\omega + \int_{\omega_{split}}^{\omega_{max}} \tilde{s}(\omega)\tilde{h}^*_{\mathrm{template}}(\omega)e^{i\omega t}d\omega \tag{34}$$

to get the SNR time series.

An interpolation of the low-frequency band MFO is needed before combining with the high-frequency band MFO due to the difference in sampling rate between the two bands. Accounting for the time and phase shift the individual bands MFOs are recombined following

$$\rho_P(t) = \cos(\Delta\phi)\rho_{LF,P}(t) - \sin(\Delta\phi)\rho_{LF,Q}(t) + \cos(\Delta\phi)\rho_{HF,P}(t+\Delta t) - \sin(\Delta\phi)\rho_{HF,Q}(t+\Delta t) \tag{35}$$

$$\rho_Q(t) = \sin(\Delta\phi)\rho_{LF,P}(t) + \cos(\Delta\phi)\rho_{LF,Q}(t) + \sin(\Delta\phi)\rho_{HF,P}(t+\Delta t) + \cos(\Delta\phi)\rho_{HF,Q}(t+\Delta t) \tag{36}$$

Δt is the time it takes for the signal to go from the low-frequency band to the high-frequency band (i.e. the time "spent" in the low-frequency band), $\Delta\phi$ is the phase offset of the signal between the two bands. Both typically depend on the templates for each band, as well as the full band template we want to emulate and are computed at the initialization of the pipeline. Figure 29 shows the low frequency, high frequency and combined MFOs for an astrophysical event.

Analyzing the detector strain in two frequency bands implies that we have templates for each frequency band. The templates used for the filtering are called real templates (RT). MBTA has

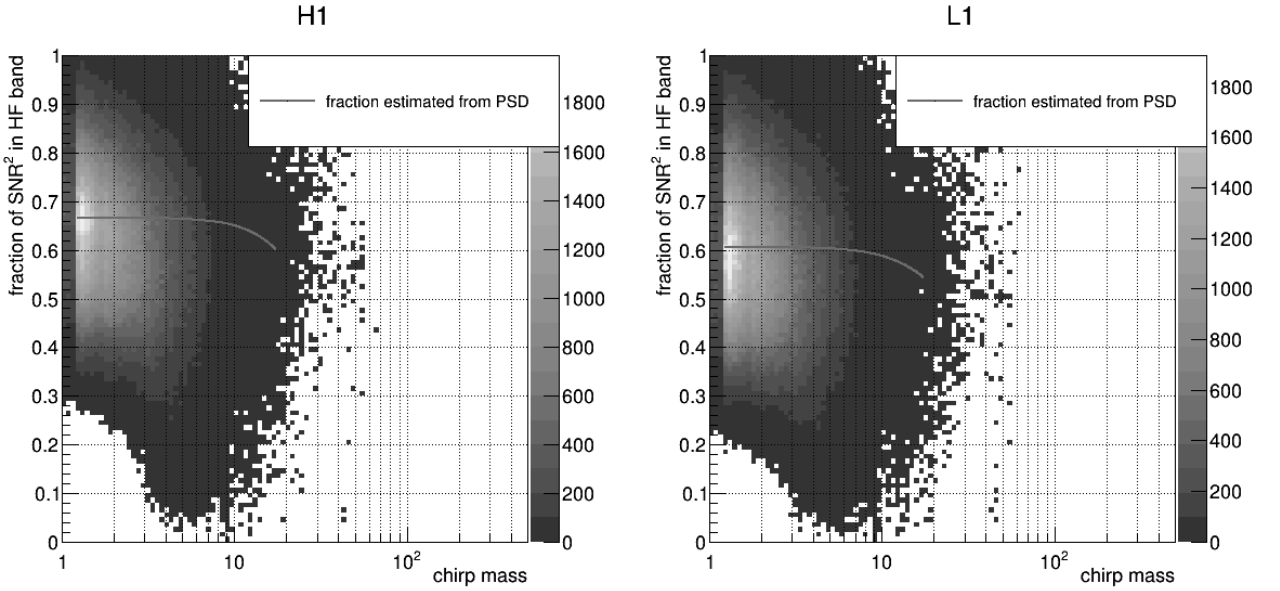


Figure 28: Fraction of SNR^2 in the HF band as a function of the chirp mass for single detector triggers obtained on simulated Gaussian strain (see section 5.2) analyzed with the pipeline O3 configuration. The expected fraction is also plotted for the PSD used to generate the strain.

therefore two RT banks, one for each band. To describe the full waveform after combination of the bands MBTA uses an additional bank of virtual templates (VT). When claiming a detection, the parameters announced for the source are those of the VT with the best match.

We do not have the same number of templates for each frequency bands, meaning that the templates do not share exactly the same parameters either. To properly combine the result of the matched filtering of each bands, we need to be able to tell which templates of the LF bands are compatible with which templates of the HF band to form a VT. This is done using a modified version of the PyCBC "banksim" algorithm [77] by injecting each template of the VT bank as a signal in the data and finding the RT which best match each VT. This allows us to have for each VT the combination of RTs that match it best. A same RT can be combined to form several VTs, leading to a reduction of usually more than an order of magnitude in the number of RT compared to the numbers of VT.

4.7 MBTA parameter space and template bank generation

We need to build templates to filter the $h(t)$ data. For CBC searches parameter space, we consider the masses and spins of the objects of the binary. The parameters are the observed ones and therefore in the detector frame (i.e. redshifted masses). Only the cases of either aligned or anti-aligned spins are considered in all searches. It was shown that considering misaligned spins does not improve significantly the detection rate [1, 78].

During O3, MBTA's parameter space covered masses up to $195\,M_\odot$. In preparation for O4, the mass range of the parameter space was extended up to $500\,M_\odot$. The spins consideration were kept the same. A new algorithm was used by the Urbino group to generate the template bank.

A template bank for a sub-solar mass (SSM) search was also created for the O3 offline analysis and adjusted for the O4 online analysis. The SSM search, lead by the IP2I in Lyon, considers

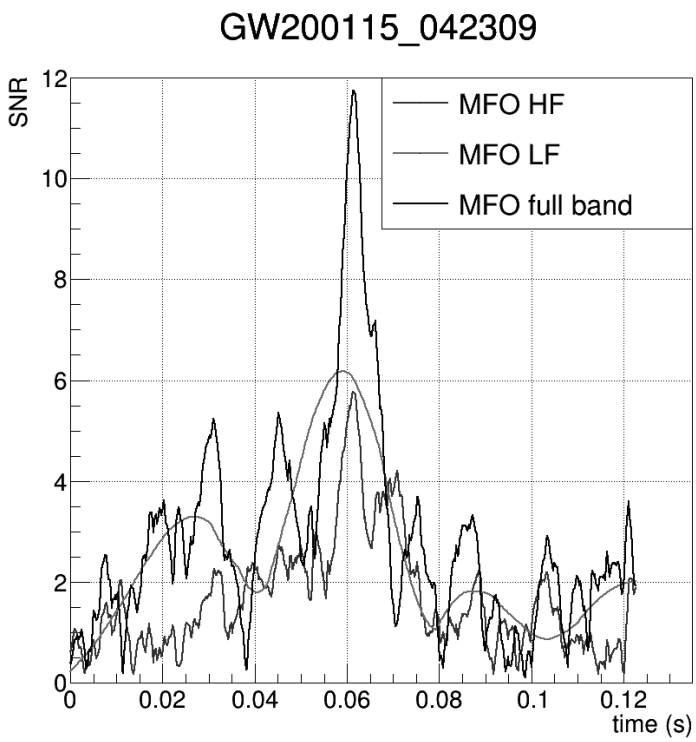


Figure 29: Matched Filtering Outputs (MFOs) for the LF, HF and full frequency band for astrophysical event GW200115_042309 detected on O3-replay data using MBTA O4 configuration.

CBC signals with at least one component of the binary lighter than 1 $M_\odot$. Considered masses are $m_1 \in [0.2, 10]M_\odot$, $m_2 \in [0.2, 1]M_\odot$with $m_1/m_2 \in [1, 10]M_\odot$. Spins are limited to $[-0.1, 0.1]$ for masses below $0.5M_\odot$ and $[-0.9, 0.9]$ otherwise.

4.7.1 MBTA's O3 main template bank

During O3, MBTA ran the main search in parallel on three independent parts of the parameter space. They were named region 1, 2 and 3 and corresponded to what we can expect roughly in terms of parameters for BNS, NSBH and BBH respectively. This includes region between 2 and 5 $M_\odot$where the separation between NS and BH is unclear. The bank of template for region 1 had low masses and small spins:

$$1M_\odot \leq m_{1,2} \leq 2M_\odot$$

$$|s_{1,2}| < 0.05$$

For region 2 the parameter space was defined as

$$1M_\odot \leq m_1 \leq 2M_\odot, |s_1| < 0.05 \text{ for what could be a neutron star,}$$

$$2M_\odot \leq m_2 \leq 99M_\odot, |s_2| < 0.997 \text{ for what could be a black hole.}$$

Region 3 included templates with

$$2M_\odot \leq m_{1,2} \leq 195M_\odot,$$

$$m_1 + m_2 \leq 200M_\odot,$$

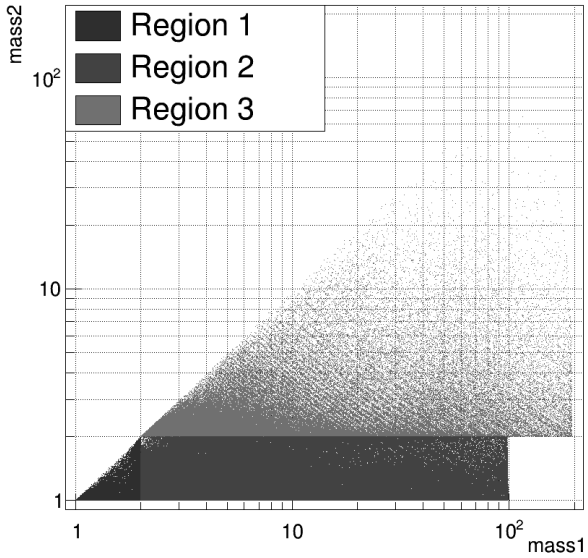


Figure 30: Template distribution of the template bank used by MBTA during O3

$$|s_{1,2}| < 0.997 \ .$$

The limit of $2\,M_\odot$ is fairly small for expected BNS, but it allows for objects with high spin as light as $2\,M_\odot$. Region 1 was generated using the TaylorF2 approximant [79] for the bank generation and SpinTaylorT4 [80] for the analysis. Region 2 and 3 were both generated using SEOBNRv4_ROM and SEOBNRv4 [81] for the bank generation and analysis respectively.

The templates of the O3 bank were computed starting at a frequency of 25 Hz, 23 Hz and 21 Hz for regions 1, 2 and 3 respectively.

The distribution of the templates in the parameter space depends on the generation algorithm and the parameters used. One of the most important parameters is the minimal match. The minimal match is the lowest value allowed for the match between any given waveform and the template that matches the best [82]. In other words, if we construct a bank with a minimal match of 0.97, as for the O3 template bank, for any astrophysical signal with parameters inside our parameter space reaches the detector, the match is expected to be larger than 0.97. But there could be few cases with slighty smaller match.

The two main types of algorithms for template bank generation are the following:

- Geometric algorithms which map the parameter space with lattices and place templates from neighbour to neighbour such that neighboring templates have a match close to the minimal match [83].
- Stochastic algorithms place templates at random within the parameter space and then removes the one that are not necessary, i.e. those that can be removed without dropping the minimal below the requested value [84].

The template bank used by MBTA during O3 was a geometric template bank. Figure 30 shows the distribution of templates in the 3 regions.

4.7.2 MBTA's O4 template bank

For O4 MBTA changed completely its template bank. The parameter space was extended and made of a single region. Furthermore, a novel template placement algorithm was used. This algorithm is called hybrid, it is a mix of geometric and stochastic placement. It starts by initializing a set of randomly placed templates that will act as seeds, one of them is taken to be the starting point for the building of the bank. From this chosen point templates are placed following a geometric procedure. Some of them that are "far-away" from the seed will then be considered as seed one after the other to repeat the procedure until no new seed can be found. The procedure is explained in more details in [85, 86].

MBTA's O4 bank is created in several steps:

- First a BNS seed template bank is generated with a geometric placement algorithm and a minimal match of 0.98 . This BNS seed contains templates with $1\mathrm{M}_\odot \leq m_{1,2} \leq 3\mathrm{M}_\odot$ and $|s_{1,2}| \leq 0.05$ for masses below $2\mathrm{M}_\odot$, $|s_{1,2}| \leq 0.997$ for others.
- Then a BBH seed is created using the hybrid algorithm, with a minimal match of 0.98, $5\mathrm{M}_\odot \leq m_{1,2} \leq 500\mathrm{M}_\odot$, a mass ratio limit $m_1/m_2 \leq 3$ and $|s_{1,2}| \leq 0.997$.
- The two seeds are eventually used to run the hybrid algorithm and complete the bank with a minimal match of 0.965 . The final bank has parameters $m_{1,2} \in [1, 500]\mathrm{M}_\odot$, $m_{\mathrm{tot}} \in [1, 500]\mathrm{M}_\odot$, $q \in [1, 50]$. Spin are limited to 0.05 in magnitude for masses below $2\mathrm{M}_\odot$, 0.997 otherwise.

The minimal match is thus higher in the most interesting regions with the largest source populations. Figure 31 shows the repartition of the templates in the mass1-mass2 plane and $\mathrm{m}_{\mathrm{chirp}}$-$\chi_{\mathrm{eff}}$ planes.

Templates of the BBH seed and full bank with duration larger than 200 ms are discarded. Among the templates generated for the full bank we consider a sub-category: those that have a merging frequency lower than the frequency band separation frequency. Since in this case the quasi-totality of the signal is in the low frequency band, there is a high chance that the LF RT associated to the VT will likely be combined with a random HF RT. Any VT with peak frequency smaller than the separation frequency or associated to a LF RT with peak frequency lower than the separation frequency, or to a HF RT with duration below 20 ms are therefore run only on one frequency band. These templates are filtered by the so-called "job 90".

The waveform approximant used for the BNS seed are TaylorF2 [79] for the bank generation and SpinTaylorT4 [80] for the analysis. The BBH seed and full bank are both generated using SEOBNRv4_ROM SEOBNRv4_opt [81] for the bank and analysis respectively.

The templates of the O4 bank are generated starting at a frequency of 25 Hz for the BNS seed and 18 Hz for the BBH seed and the full bank.

The real template banks are generated in the same way as the VT full bank, with the same parameters but no requirement on the template duration. RTs are generated for a separation frequency of the bands at 80 Hz.

4.8 Single detector triggers search

The search for astrophysical signals starts with independent searches in each of the interferometer of the network. As mentioned previously the matched-filetring of each interferometer data is carried out in two frequency bands.

The matched-filtering being done in the two frequency bands, triggers may arise (when the maximum of the SNR time series goes above a given threshold) in one or the other. In this case the

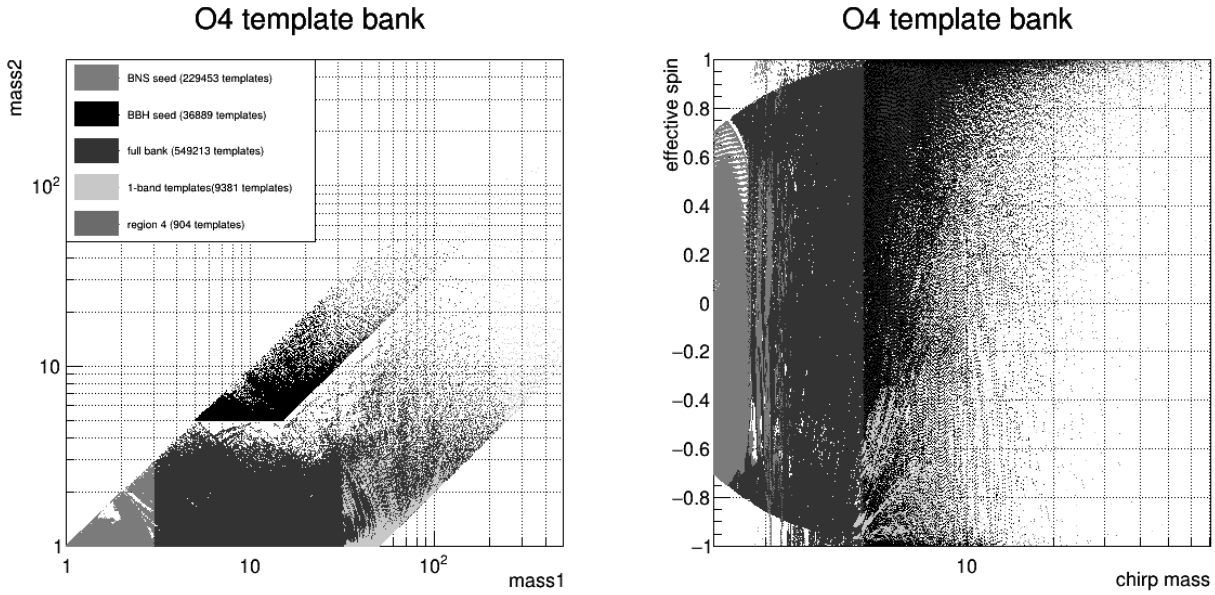


Figure 31: MBTA O4 template bank. Left: mass1 vs mass2. Right: chirp mass vs effective spin.

combination process (eq. 36) is done using the RTs of the other frequency band for every VT that is associated to the RT of the trigger within a certain time window. This window is centered on the time of the trigger and its duration depends on the width of the maximum SNR peak of the MFO. The duration of the window can't exceed 25 ms. If the combined SNR is in turn above a specified threshold we have a **single detector trigger**. The two real templates are then recombined on a longer time window to enable parameter consistency tests.

Some triggers produced by noise can sometimes have quite large values of SNR. In order to identify and reject them some tools were developped. We present here three of them.

4.8.1 Signal consistency test: χ^2

Gravitational waves signals have a very specific power distribution in the time-frequency domain. The role of the χ^2 test presented in this section is to check that the distribution of the signal as a function of the frequency is consistent with what is expected for a GW signal. The χ^2 was mostly used during O2. It is a poor version of the autoχ^2 (see next section) by which it was surclassed during O3.

A cut on this χ^2 quantity is still applied to triggers during MBTA analysis. The principle of the χ^2 is to compare the difference in measured SNR vs theoretical SNR, in-phase and in-quadrature, in each frequency band:

$$\chi^2 = \sum_{i=0}^{N_{\text{band}}} \Big[\left(\text{SNR}_{\text{P,meas},i} - \text{SNR}_{\text{P,th},i} \times \alpha_i \times \cos\left(\Delta\Phi_{i-1,i}\right)^2\right) + \left(\text{SNR}_{\text{Q,meas},i} - \text{SNR}_{\text{Q,th},i} \times \alpha_i \times \cos\left(\Delta\Phi_{i-1,i}\right)^2\right) \Big] \tag{37}$$

where $\text{SNR}_{\text{P/Q, meas/th, }i}$ is the measured (meas) or theoretical (th) in-phase (P) or in-quadrature (Q) SNR for each frequency band, α_i is the fraction of the SNR in band i and $\Delta\Phi_{i-1,i}$ the phase difference between the frequency bands (0 for $i = 0$). The expected value for the ideal case of an

astrophysical signal without noise is, by construction, $\chi^2 = 0$ as the measured and theoretical SNR would be equal. In practice a trigger is rejected if its χ^2 is larger than a threshold defined as

$$\chi^2_{cut} = A(2 + B \times \mathrm{SNR}^2); \tag{38}$$

with $A = 3$ and $B = 0.025$.

This signal consistency test is not used anymore for O4.

4.8.2 autoχ^2 and rwSNR

Similarly to the way χ^2 checks for consistent repartition of the SNR in frequency, we can test the consistency of the time evolution of the SNR around its maximum value. This is done thanks to a quantity call autoχ^2 .

The autoχ^2 quantifies the mismatch between the measured matched-filtering output (MFO = SNR time series) and the theoretical MFO that is expected from the autocorrelation of the template. It is computed following:

$$\mathrm{auto}\chi^2 = \frac{1}{2\Delta t}\int_{t_0-\Delta t/2}^{t_0+\Delta t/2} \left|\left| \begin{pmatrix}\rho_P \\ \rho_Q\end{pmatrix} - \rho_{max} \begin{pmatrix}\cos\Delta\Phi & -\sin\Delta\Phi \\ \sin\Delta\Phi & \cos\Delta\Phi\end{pmatrix} \begin{pmatrix}\rho_{P,th} \\ \rho_{Q,th}\end{pmatrix} \right|\right| \tag{39}$$

where ρ is the SNR time series, ρ_{th} is the theoretical SNR time series, t_0 the time of the maximum of the SNR time series ρ_{max} (SNR of the event), $\Delta\Phi$ the phase difference between measured and expected SNR time series and Δt a time interval centered on t_0.

In practice however we deal with discrete series and MFOs, the autoχ^2 is then computed on 400 points (usually at 4096Hz) following:

$$\mathrm{auto}\chi^2 = \frac{1}{2\mathrm{N}}\sum_{i=0}^{\mathrm{N\text{-}1}} \left[(\mathrm{MFO}_{P,meas}[i] - \mathrm{Rot}_{P,exp}[i])^2 + (\mathrm{MFO}_{Q,meas}[i] - \mathrm{Rot}_{Q,exp}[i])^2\right] \tag{40}$$

where

$$\mathrm{Rot}_{P,exp}[i] = \cos\Delta\Phi \times \mathrm{MFO}_{P,exp}[i] - \sin\Delta\Phi \times \mathrm{MFO}_{Q,exp}[i] \tag{41}$$
$$\mathrm{Rot}_{Q,exp}[i] = \sin\Delta\Phi \times \mathrm{MFO}_{P,exp}[i] + \cos\Delta\Phi \times \mathrm{MFO}_{Q,exp}[i] \tag{42}$$

The autoχ^2 is then used to reweight the SNR. The reweighted SNR (rwSNR) was introduced as a mean to downgrade the significance of triggers that were not consistent with the templated they were matched with. It is a powerful tool to discriminate background from astrophysical candidates. The reweighting of the SNR is then done following [73]

$$\mathrm{rwSNR} = \begin{cases} \mathrm{SNR}, & \text{if auto}\chi^2 \ \leq 1. \\ \mathrm{SNR} \times \left(\frac{\mathrm{A}+\left[\mathrm{auto}\chi^2\right]^{\alpha}}{A+1}\right)^{-\frac{1}{\beta}}, & \text{otherwise.} \end{cases} \tag{43}$$

for O3 the parameters took the values A=10, α=5 and β=8. Figure 32 shows the measured and expected MFOs for an astrophysical event and a loud noise trigger with the associated autoχ^2 values and rwSNR.

Further considerations on the autoχ^2 are given in chapter 7.

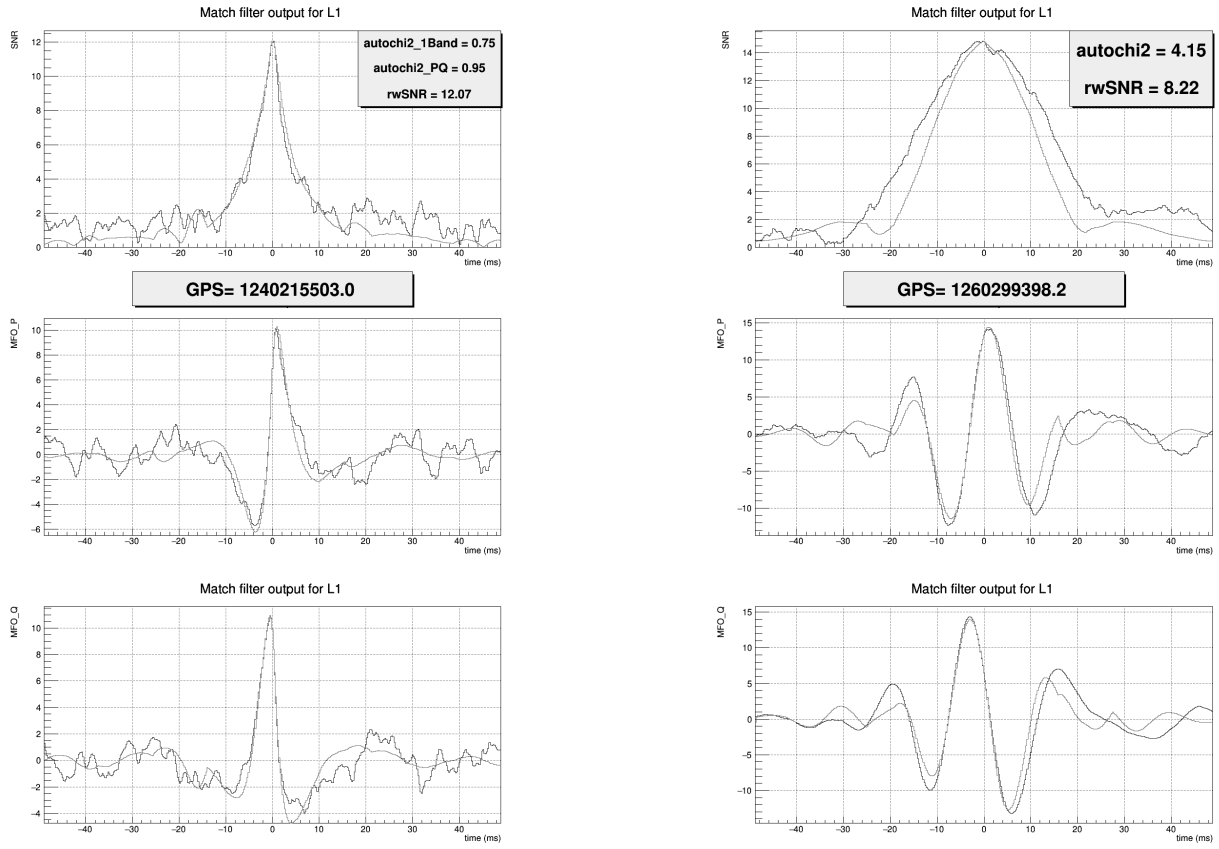


Figure 32: Left: Measured (blue) and expected (red) MFOs for astrophysical event GW190425 in L1. Right: Measured (blue) and expected (red) MFOs for a loud noise event.

4.8.3 Excess rate and ranking statistics

The excess rate is computed from the trigger rates before and after applying the rwSNR: let r_{raw} be the raw rate of triggers (i.e. above the SNR threshold with only the gating applied) and r_{sel} the rate of triggers above the rwSNR threshold, the excess rate at time t_0 is

$$\mathrm{ER}(t_0) = \underset{t_0 - 10s \leq t \leq t_0}{\mathrm{median}} \left(\frac{r_{raw}(t) - r_{sel}(t)}{r_{raw}(t)} \right) \tag{44}$$

It is then used to reweight again the SNR **after the coincidence step** (see section 4.9) and yield the ranking statistics (RS) [73]:

$$\text{rank stat} = \begin{cases} \text{rwSNR}, & \text{if ER} \leq 0.3. \\ \text{rwSNR} \times \left[1 - A(\text{ER} - 0.3)^{\alpha}\right], & \text{otherwise.} \end{cases} \tag{45}$$

with A = 1, $\alpha = 2$.

4.9 Coincidence search

We expect astrophysical events to produce correlated signals across detectors in terms of arrival time, phase and amplitude. It is much more likely for an astrophysical signal to be found in several

detectors than for a glitch or noise fluctuations due to non-stationarity. Considering coincidences therefore excludes many noise triggers from the analysis. Thus, to ensure more significant and reliable detections as well as to better probe low SNR events, coincidences between the different detectors are searched for. Events found in coincidences between detectors also have the benefit to provide a much better estimation of the source location as explained in section 3.3. Finally a threshold is applied on the ranking statstic of coincidences to keep only the most significant events.

Triggers found in different detectors must satisfy two conditions to be considered as coincident:

- the time delay between them must be smaller than a given threshold to take into account the time of flight of the GW from one detector to the other, typically 15 ms for H1-L1 and 35 ms for H1-V1 and L1-V1;
- their template parameters must match exactly (i.e. same template).

A coincidence between two (three) detectors is called a double (triple) coincidence. They are denoted by the initial letters of the interferometers that participated in the coincidence. For example a coincidence between LIGO Livingston and Virgo is an LV coincidence, a triple is written HLV. Coincidences are first constructed as doubles and if an HL and HV share a common H1 trigger they are upgraded to an HLV coincidence. During triple detector time, a double coincidence can be significant enough to be uploaded to GraceDB. In this case we also search for a subthreshold trigger in the third detector. If one is found the event is upgraded to a pseudo-triple coincidence and, if for instance the third detector is Virgo, it is named HL-Von. . Pseudo-triple coincidences uploaded to GraceDB have the combined SNR and skymap computed using the information of the three detectors but their significance is kept as the one of the double. This is done in order to provide a better sky localization of the source.

The SNR^2 of a coincidence is taken as the quadratic sum of the SNR of the single detector triggers, for example

$$\rho_{HL} = \sqrt{\rho_{H1}^2 + \rho_{L1}^2} \tag{46}$$

Contrary to the combined SNR of the triple coincidence, its combined ranking Statistic (cRS) is not simply the quadratic sum of the single detector triggers ranking statistics. This is motivated by the fact that for an astrophysical source, the signal observed across the different detectors is expected to have some correlation from one detector to another. The quantities expected to show some correlation are namely the arrival time of the signal, the phase of the signal and its amplitude. For any two detectors a, b the combined ranking statistic is given by

$$\rho_{RS,ab}^2 = \rho_{RS,a}^2 + \rho_{RS,b}^2 + 2\ln\left(P_{\Delta t_{ab}} P_{\Delta\Phi_{ab}} P_{\Delta RA_{ab}}\right) \tag{47}$$

with the probabilities $P_{\Delta t_{ab}}$, $P_{\Delta\Phi_{ab}}$, $P_{\Delta RA_{ab}}$ respectively for the time of flight, phase difference and relative amplitude and $\rho_{RS,a}^2$ as given by eq. 45.

The combined ranking statistic for a triple coincidence (with detectors labeled a, b, c) is then derived as

$$\rho_{RS,HLV}^2 = \rho_{RS,HL}^2 + \rho_{RS,HV}^2 - \rho_{RS,H}^2 \tag{48}$$

where we avoid counting twice the RS of H1 since triple coincidences are built from HL and HV triggers.

The probabilities in eq. 47 are called parameter consistency tests. They are derived from distributions of simulated BNS with chirp mass $1.2\,\mathrm{M}_\odot$, no spin and uniformely distributed in a sphere of radius 300 Mpc. Since these parameters are derived from the source location, we expect

the same behaviour for all type of sources at first order (the timing resolution for heavy BBH compared to BNS is neglected). Not all of these injections can actually be detected because some are too distant. An injection is detectable by a detector if its effective distance is within the horizon of this detector. We consider here horizons of 100 Mpc, 140 Mpc and 50 Mpc for H1, L1 and V1 respectively. The distribution of the parameters for the injections detectable as double coincidences (triples included) are shown in figure 33. Note that we are only selecting on the generated parameters of the injections, no MBTA analysis is done here.

The similarity in sensitivity and antenna pattern between the two LIGO detectors make the HL distribution rather simple. The phase difference between H1 and L1 follows a distribution centered on π. The mean distance ratio between the two is slightly less than one, as expected due to their range difference. Large time differences are deprecated because they would require a source located on the line passing by the two detectors which constitutes a small part of the sky. It is therefore less likely to have such detections.

Regarding the two LIGOs and Virgo, the overlap between their different antenna patterns make things more complicated. Their antenna pattern are rotated by $\sim 45^\circ$ causing more detection with phase difference closer to $\pi/2$ and $3\pi/2$. The mean distance ratio is larger than one in favour of H1 and L1, owing to the differences in ranges. The time of flight distributions are roughly uniform and also depend on the orientation of the detectors.

The correction applied to the cRS are derived from these distribution after including a model of the detector resolution for these parameters. More details can be found in [73]. We show in figure 34 the corrections applied to the cRS^2 as a function of the various parameters for each pair of detector.

Finally we show in figure 35 the distributions of the same parameters for O3 common injections recovered by MBTA. We can see that their shape is very close to the ones described previously.

4.10 Clustering

A loud astrophysical signal or glitch can match with more than one templates, producing several triggers. To avoid counting such events multiple times the concept of clustering was introduced in the search. Clustering consists in merging every trigger within a certain time gap of each other into what is called a clustered event or simply cluster. This means that triggers apart from more than the time gap can still be part of the same cluster as long as they have a common neighbor within said time gap, as shown by figure 36. The time gap is typically of ~ 10 ms. The clustered event takes all the parameters (SNR, masses, spins...) of the most significant trigger in the cluster also called "cluster head". MFO vectors of the head are also saved, along a few vectors containing the most relevant parameters of the other triggers of the cluster. The size of the cluster is characterized by the number of trigger it contains as well as its t_{before} and t_{after}, the time differences between the cluster head with the earliest and latest trigger in the cluster respectively.

4.11 FAR computation during O3

4.11.1 FAR for a single search

In the following paragraphs we will refer to a given region of the O3 parameter space and a given type of coincidence as an individual search.

The false alarm rate (FAR) indicates how likely it is for a background trigger to have a given cRS. It is computed as the rate of background triggers expected above a given cRS threshold.

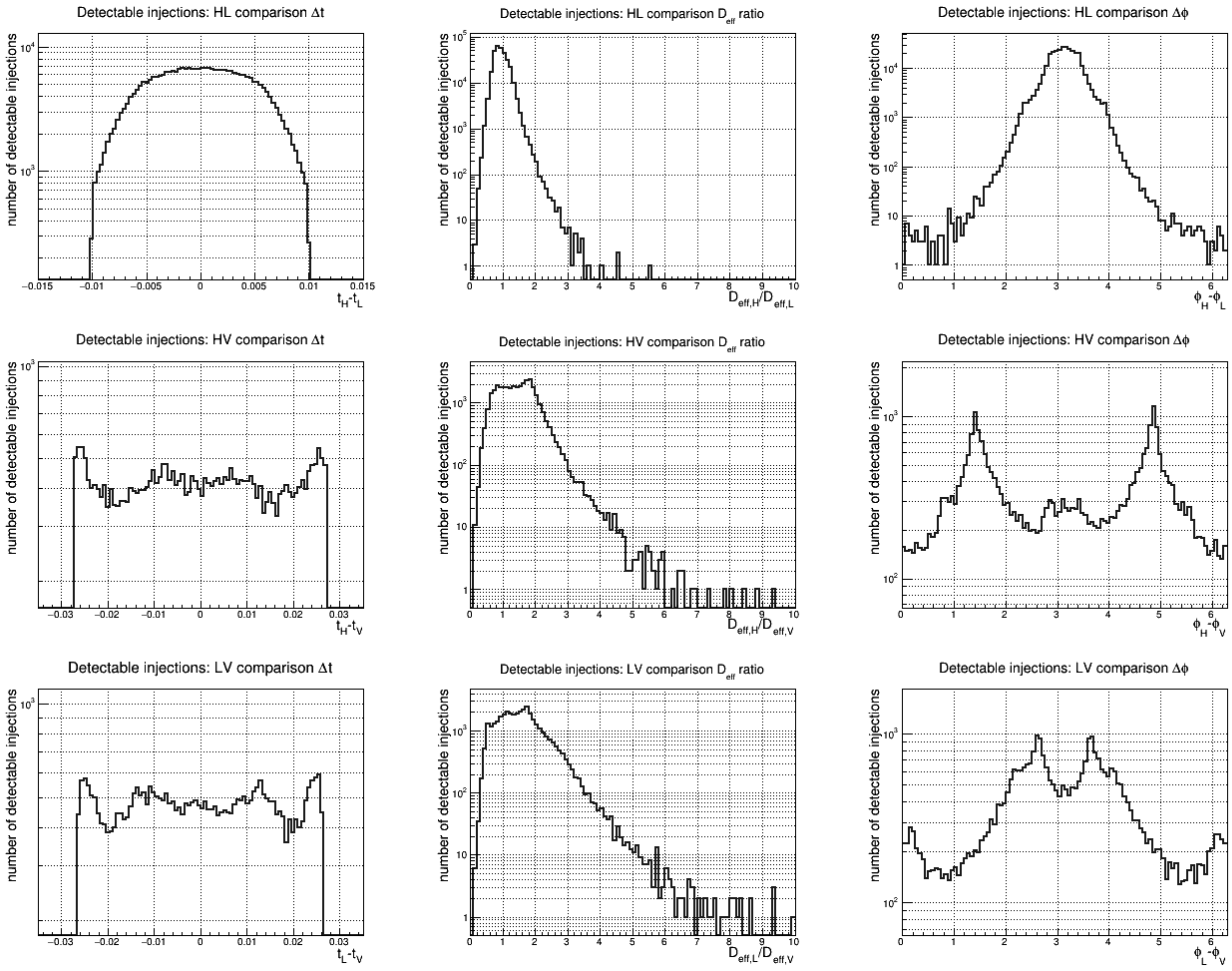


Figure 33: Comparison of the parameters of the injections detectable by at least two detectors.

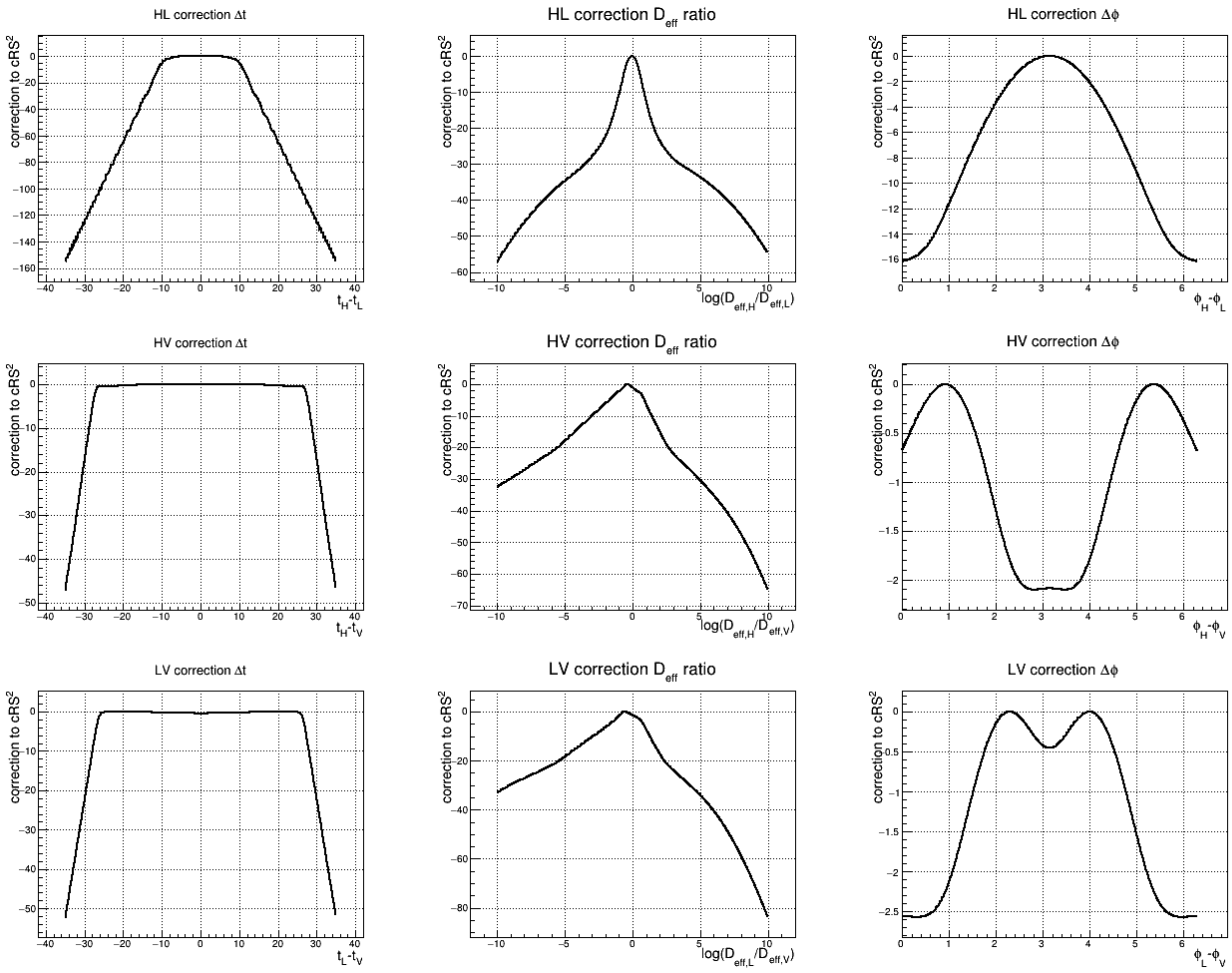


Figure 34: Correction applied to the cRS^2 for each parameter and pair of detector.

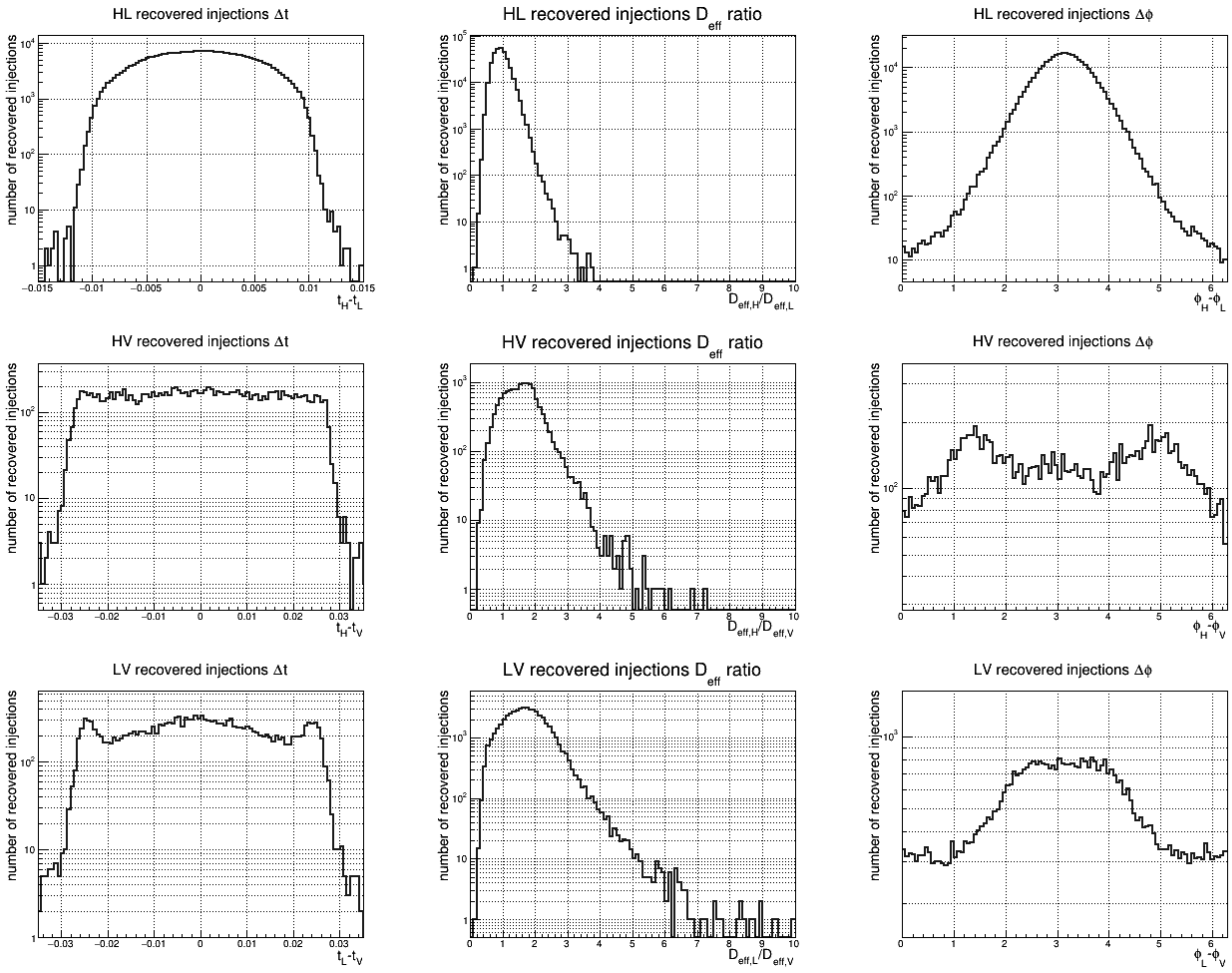


Figure 35: Comparison of the parameters of recovered O3 common injections.

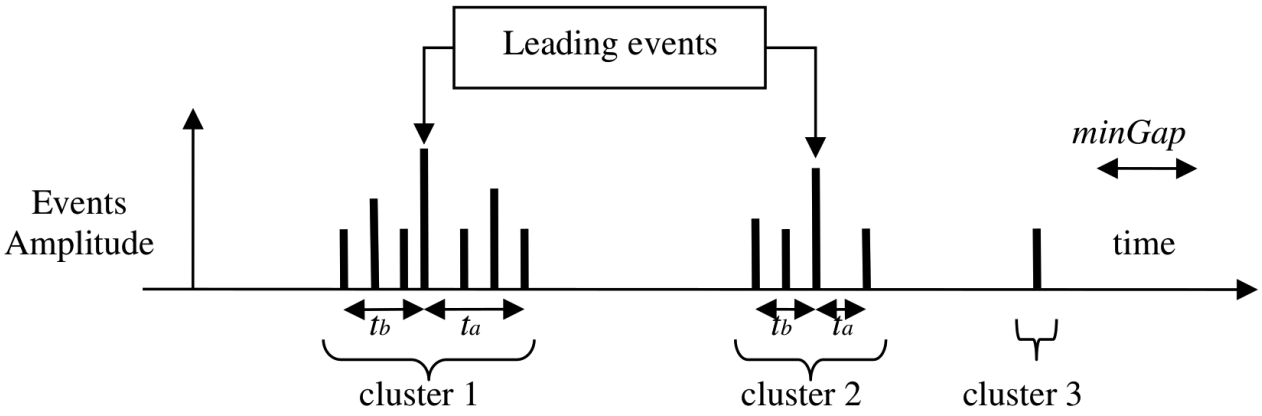


Figure 36: Scheme for the clustering process taken from MBTA's documentation. Vertical bars represent events and are proportional to their SNR, t_b and t_a correspond to t_{before} and t_{after} respectively. In the case of cluster 3, which contains only one event, $t_{\text{before}} = t_{\text{after}} = 0$.

Upon detection of a candidate with combined ranking statistics $\rho_{RS,\text{candidate}}$ we need to know what this expected rate is. To this end we compute a cRS distribution for background triggers which straightforwardly gives a FAR vs cRS distribution.

The background distribution used in the computation of the FAR for an individual search is built by making random coincidences between the two relevant detector's single detector triggers. Times where astrophysical signals have been identified are excluded. Due to the very high number of low SNR events they are down-sampled to reduce the computational cost. During O3 the online analysis collected single detector triggers in such a way over the last 24 hours for each search region, while the offline analysis used ~ 6 days. The background distribution for the coincidences is then built by combining any single detector triggers with matching templates, independently of their arrival time. The large number of fake coincidences is expected to smooth large noise fluctuations or correlations that might have occured in the detectors. The cRS of the fake coincidences is computed using the single detector triggers parameters for the phase and amplitude and a random time of flight value (within the range permitted for the two considered detectors).

A FAR is first computed for each individual search and then a final FAR is computed by taking into account the different searches (mass regions and coincidence types) and the time covered by at least two properly working detectors. This is motivated for instance by the fact that a double coincidence during triple detector time is not the same as a double coincidence during double detector time, we have less information for the latter.

The FAR for a given CRS threshold is then given by

$$\text{FAR}_{ab}(\rho_{RS,ab}) = N_{ab}(\rho_{RS,ab})\frac{\omega_{ab}}{T_a T_b} \tag{49}$$

where $N_{ab}(\rho_{RS,ab})$ is the number of fake coincidences with CRS equal or larger than the threshold $\rho_{RS,ab}$, ω_{ab} is the coincidence time window and T_a and T_b are the analyzed times for detector a and b.

In the case of a triple coincidence, since they are built for HL and HV coincidences, we can compute their FAR from the FAR of the doubles. This is done by integrating the product of the FAR of the HL and HV coincidence while being careful not to count twice the triggers in H1 . We then need to renormalize it by the triple detector time T_{HLV} and by the number of triggers in H1

$N_H(\rho_{RS,H})$:

$$\mathrm{FAR}_{HLV} = \int\int \frac{\mathrm{FAR}_{HL}(\rho_{RS,HL})\mathrm{FAR}_{HV}(\rho_{RS,HV})}{N_H(\rho_{RS,H})/T_{HLV}} d\rho_{RS,HL} d\rho_{RS,HV} \tag{50}$$

with $\rho^2_{RS,H} = \rho^2_{RS,HL} + \rho^2_{RS,HV} - \rho^2_{RS,HLV}$ (eq. 48). All of this allows to compute the FAR for the $3 \times 4 = 12$ individual searches. Figure 37 shows a typical FAR vs cRS distribution obtained by making fake coincidences with single detector triggers.

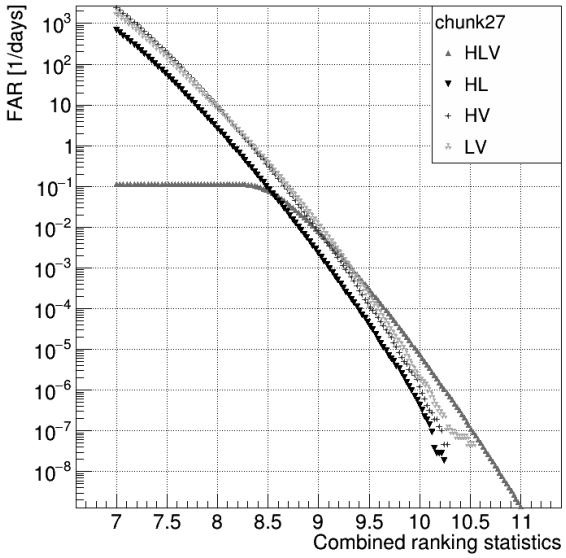


Figure 37: Double and triple coincidences FAR vs cRS distribution obtained by combining single detector triggers of the O3 BNS (region 1) search from December 3 2019 to December 11 2019. Clustering and trial factors are not applied yet.

The background is computed using unclustered triggers, but the clustering reduces the overall number of triggers which would lead to inconsistent IFAR cumulative distributions. To take this effect into account the FAR is scaled by a factor k_{cluster} which is the average ratio of the number of clustered events versus the number of events before clustering. For the O3 online analysis we had $k_{\mathrm{cluster}} = 0.59$ for the BNS region and $k_{\mathrm{cluster}} = 0.44$ for the BBH and NSBH regions. The Inverse False Alarm Rate (IFAR) of a coincidence from a given individual search is given by

$$\mathrm{IFAR} = \frac{1}{k_{\mathrm{cluster}}\mathrm{FAR}(\rho_{RS})} \tag{51}$$

4.11.2 Global search FAR

By construction each of the background distributions computed for the individual searches yields the same rate of noise coincidences with FAR below a given threshold. This means that we could take a straightforward approach and simply weight each of the 12 distribution before combining them into a single one. This was done by putting the same weight $k_{\mathrm{region}} = 1/3$ on the three search regions, which favours the BNS region because of its astrophysical interest despite its reduced number of templates. For the type of coincidences, however, some astrophysical priors were included to improve the search: First, not all coincidences are equivalent in the sense that they are not as

likely to seize astrophysical signals. The distributions are scaled by a trial factor k_{coinc} to take into account the relative difference in search volume of the different type of coincidences. This search volume is estimated through an injection run (see later section 5.3) by counting the number of injections recovered by the different type of coincidences. During O3 triple detector times we had $k_{\mathrm{coinc}} = 0.909$, 0.002, 0.007 and 0.083 for HL, HV, LV and HLV respectively. For double detector time it was first set to 1 for the online analysis and later set to the fraction of double-detector volume versus triple-detector volume for the offline analysis. Figure 38 shows the IFAR cumulative plot for double and triple coincidences during a few days (same period as figure 37) of O3 where no astrophysical signals were detected.

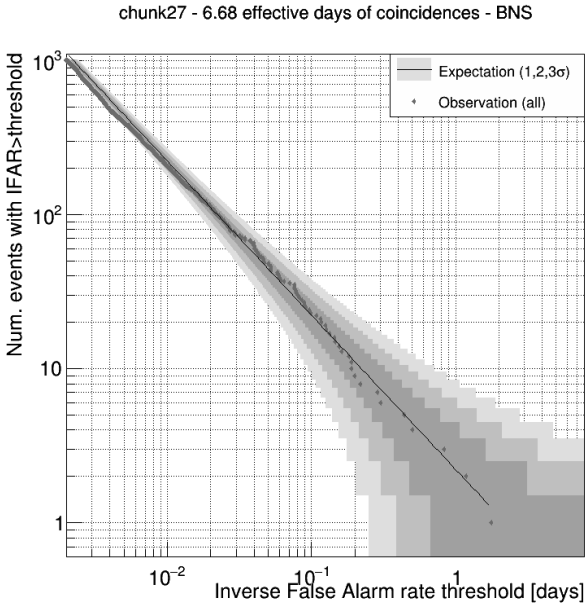


Figure 38: IFAR cumulative plot for double and triple coincidences during 6.68 effective days for the O3 BNS (region 1) search. The observation is consistent with expectations for background only.

4.12 Probability of astrophysical origin and source classification

The increasing number of GW detections allows to have more information on the source populations. The knowledge of these populations can be used to compute a probability of astrophysical origin when a candidate is detected. This probability is complementary to the FAR. It will help detecting more astrophysical signals in regions of the parameter space that are denser in sources.

In addition to the probability of astrophysical origin, source populations can be used to infer a source classification of the candidates. Being able to classify the sources allows to compute merger rates for the different type of sources.

Both the probability of astrophysical origin and the source classification are helpful to astronomer during low-latency searches to decide whether to follow-up on the alert. The methods to compute them for MBTA were developped during the offline analysis of O3.

More recently, methods to compute the probability of having a neutron star and a probability of having a remnant mass after merger were developped for MBTA in preparation for O4.

4.12.1 $\mathbf{p_{astro}}$, $\mathbf{p_{source}}$

With the insight gained on the source populations, thanks to all the detections made, it is possible to infer a probability of astrophysical origin as well as developp a source classification for the candidates we detect. The method used by MBTA, fully described in [87], is the following. The probability of astrophysical origin is computed as

$$p_{\text{astro}} = \frac{\text{astrophysical foreground rate}}{\text{astrophysical foreground rate} + \text{background rate}} \tag{52}$$

Assuming a normalized foreground and background distribution in cRS^2, $f(\text{cRS}^2)$ and $b(\text{cRS}^2)$ respectively, for a given observing time where we expect N_f foreground triggers and N_b background triggers we have

$$p_{\text{astro}} = \frac{N_f f(\text{cRS}^2)}{N_f f(\text{cRS}^2) + N_b b(\text{cRS}^2)} \tag{53}$$

The source (BNS, BBH, NSBH) classification is derived from this expression as

$$p_{\text{source}} = \frac{N_{\text{source}} f_{\text{source}}(\text{cRS}^2)}{N_f f(\text{cRS}^2) + N_b b(\text{cRS}^2)} \tag{54}$$

with

$$p_{\text{BNS}} + p_{\text{BBH}} + p_{\text{NSBH}} = p_{\text{astro}} \tag{55}$$

In practice the expected rates of astrophysical and background candidates are not a priori known and need to be estimated from the data. This is done by assuming the foreground and background as independent Poisson processes and using Bayes' theorem to compute the posterior distribution of counts assuming a distribution of cRS^2. Since we do not expect the same type and rate of astrophysical events in all regions of the parameter space, the latter is divided in bins of chirp mass and mass ratio for a total of 165 bins during O3. Each bin has its own foreground and background distribution. One last thing that needs to be taken into account is that the different types of coincidences (doubles, doubles in triple detector time, triples) will have different astrophysical foreground rates because of the detector sensitivities and orientations. This is accounted for by weighting the foreground for each type of coincidence, the weight being computed using injections.

4.12.2 hasNS, hasRemnant

As was previously discussed, one of the motivation for gravitational waves detection is to participate in multi-messenger searches. Since we can now provide a source classification, it would be interesting to be able to tell if a signal is likely to have an electromagnetic counterpart. We call such events "EM bright". A definition of the EM bright population for the pipeline MBTA will be given in section 5.4. For that we define two quantities: hasNS and hasRemnant. The first one is the probability that the binary system had a neutron star, the latter is the probability that there is some matter remaining after the merger (with mass above a given threshold). The definition for hasNS is rather straightforward:

$$\text{hasNS} = \frac{p_{\text{BNS}} + p_{\text{NSBH}}}{p_{\text{astro}}} \tag{56}$$

For hasRemnant it gets more complicated because not all NSBH mergers have a remnant mass. We define hasRemnant as

$$\text{hasRemnant} = \frac{p_{\text{BNS}} + p_{\text{NSBH-bright}}}{p_{\text{astro}}} \tag{57}$$

To define what is a bright NSBH we make two assumptions:

- BNS mergers have a remnant mass, meaning $p_{\mathrm{BNS}} = 1 \Rightarrow \mathrm{hasNS} = \mathrm{hasRemnant} = 1$,
- BBH mergers have no remnant mass, meaning $p_{\mathrm{BBH}} = 1 \Rightarrow \mathrm{hasNS} = \mathrm{hasRemnant} = 0$.

We need to tell whether a NSBH will have a remnant mass. Foucart et al. [88] have given a parametrization for the remnant mass after the merger of a neutron star and black hole with masses M_{BH} and M_{NS} respectively:

$$M^{\mathrm{rem}} = M^{b}_{\mathrm{NS}} \left[\mathrm{Max} \left(\alpha \frac{1 - 2C_{\mathrm{NS}}}{\eta^{1/3}} - \beta \frac{R_{\mathrm{ISCO}}}{M_{\mathrm{BH}}} \frac{C_{\mathrm{NS}}}{\eta} + \gamma, 0 \right) \right]^{\delta} \tag{58}$$

where M^{b}_{NS} is the baryonic mass of the neutron star, C_{NS} its compactness (neutron star mass over neutron star radius in natural units), η the symmetric mass ratio $q/(1+q)^2$ and R_{ISCO} the radius of the innermost stable circular orbit. The baryon mass is related to the neutron star mass by the binding energy

$$\mathrm{BE} = M^{b}_{\mathrm{NS}} - M_{\mathrm{NS}} \tag{59}$$

which is in turn related to the compactness of the star [89] by

$$\frac{\mathrm{BE}}{M_{\mathrm{NS}}} = d_1 C_{\mathrm{NS}} + d_2 C^2_{\mathrm{NS}} \tag{60}$$

Since the compactness directly depends on the neutron star equation of state, and since this latter is not too constrained (as presented in section 2.4), a marginalization over the mass-radius EOS posterior given in [90] is carried out. A NSBH is considered bright if the predicted remnant mass is larger than $1 \times 10^{-3}\,\mathrm{M}_{\odot}$.

4.12.3 Evolution for O4: FAR($\mathrm{p_{astro}}$)

When making catalogs of detected events during O3, some issues were raised regarding the consistency between FAR and $\mathrm{p_{astro}}$. Indeed, for a same value of $\mathrm{p_{astro}}$, there were some large differences in FAR. This can be explained by the fact that, during O3, the FAR was computed on the 3 regions of the parameter space while $\mathrm{p_{astro}}$ was computed on 165 bins: the astrophysical foreground can vary greatly (several orders of magnitude) from one bin to the other but the background changes by at most an order of magnitude.

To solve these issues, it was decided to use another method for O4. A FAR based on the cRS tells how likely it is for background to produce a loud event. This is definitely good and was especially useful when we had little knowledge on the signals we detect. We now have more knowledge and we use it to tell how likely it is for a trigger to be an astrophysical signal through the computation of $\mathrm{p_{astro}}$. The question that we ask now is: how likely is it for a background event to have a given $\mathrm{p_{astro}}$ value. To answer this question we decide to compute a FAR of $\mathrm{p_{astro}}$, meaning that $\mathrm{p_{astro}}$ will play the role previously fullfilled by the cRS^2. We have on one side $\mathrm{FAR}(\mathrm{cRS}^2)$ and on the other $\mathrm{p_{astro}}(\mathrm{cRS}^2)$ which can be inverted and combined to give a $\mathrm{FAR}(\mathrm{cRS}^2(\mathrm{p_{astro}}))$ which is computed for each type of coincidence, each type of network configuration (number of detector online) and each $\mathrm{p_{astro}}$ bin. This FAR is then summed over all bins and integrated over the observing times corresponding to the network configurations available for each type of coincidence (i.e. over single, double and triple detector time for single detector triggers, double and triple detector time for double coincidences and triple detector time for triple coincidences).

5 Selecting single detector triggers

Contents

5.1 Introduction

The core of this PhD project is the developpment of the analysis of MBTA's single detector triggers, not just as part of concidences, but as standalone candidate events. The goal is to make the background low enough so we can retrieve as many astrophysical candidates as possible and compute the FAR of the candidates.

It is difficult to predict what the background will be like during O4. We can base our investigation on the O3 background as it is the best we have and the sources of noise will not be fundamentally different. Figure 39 shows the ranking statistics background distribution for single detector triggers during double or triple detector time and the background for HL coincidences. By comparing them we can see that high ranking statistics values are present in the single detector triggers distributions and not in the coincidence distribution. These triggers are therefore due to noise. We need to get rid of them if we want to claim and release significant detections for single detector triggers.

We ask the following question: Can we use selection criteria based on pipeline-related quantities to discriminate against triggers of "poor" quality and effectively reduce the background? Considered quantities were the autoχ^2 (eq. 40), the excess rate (eq. 44) and the gating (section 4.2). The criteria defined in this section do not impact the filtering of the data and can be applied as a post-processing step.

Since reducing the background is difficult we may also wonder whether all kind of triggers are equally interesting. Multi-messenger considerations have a lot of weight, especially from the point of view of online analyses. Although the sky localization of a single detector trigger is bound to be poor, especially during single detector time, the timing information still has a high scientific value in case another observatory detects some signal at the same time. BNS and NSBH events are also rarer than BBH events so every additional detection has a lot of value. This advocates for an emphasis on EM bright candidates that will be presented in section 5.4. Another consideration in

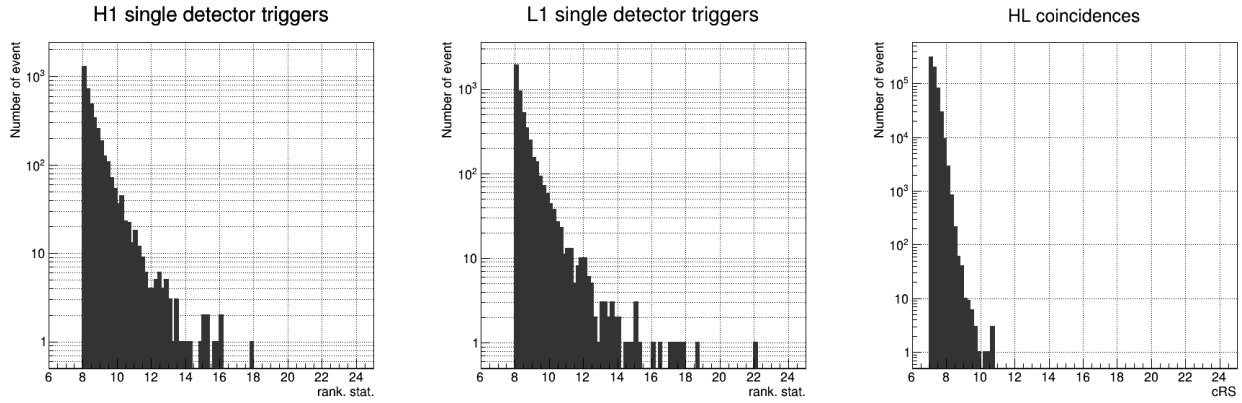


Figure 39: Left and middle: single detector triggers ranking statistics distribution in H and L during O3 double and triple detector time. Right: combined ranking statistics distribution of HL triggers during the same times.

favor of EM bright candidates is that they are associated to longer signals due to their lower mass (see section 5.5), and are therefore easier to identify. But it does not mean that EM-dark triggers should be excluded by default. We will also study them in parallel to the EM bright population, to prepare a possible extension of the single detector triggers analysis to all type of sources.

This chapter starts by giving some details on the Gaussian noise mentioned previously. This Gaussian noise will be used throughout the chapter. We then define the EM-bright population for the MBTA single detector triggers search. We will also develop some considerations on the importance of the duration of the templates. This will be followed by a study on a selection using the autoχ^2 and excess rate. There will then be further considerations which lead to consider the gating and to be more strict on the presence of excess rate around triggers to exclude times when one interferometer misbehaves.

5.2 Ideal case of noise

We want, through the application of selection criteria, to reduce the background as much as possible. A reference for "good" or "clean" background is the ideal case where the detector strain is stationary Gaussian noise. Background obtained by running MBTA on such noise is the limit of what we can hope to achieve.

To simulate the strain observed by a perfect GW detector, we start by simulating stationary white Gaussian noise, i.e. time-independent noise with mean 0, standard deviation 1 and constant PSD. It is then colored using the detectors O3b PSD. An example of stationary gaussian noise colored with L1 PSD is shown in figure 40. The Gaussian noise used throughout this section is computed for an effective time equivalent to the effective observing time of the first month of O3. The effective time is generally smaller than the observed time as it only takes into account segments that were analyzed by MBTA.

This Gaussian noise was analyzed with MBTA using the pipeline's O3b configuration. In order to compare the background obtained by MBTA on this Gaussian noise to the background observed during O3, the number of triggers was scaled using the effective time ratio $T_{eff}(\text{O3})/T_{eff}$ (Gaussian noise), with $T_{eff}(\text{O3}) \leq T_{obs}(\text{O3})$ and $T_{eff}(\text{Gaussian noise}) \leq T_{simu}(\text{Gaussian noise})$. Histograms of the single detector triggers produced by analyzing this simulated Gaussian noise with MBTA are shown in figure 41. Unlike in figure 39, the Gaussian noise barely reaches rwSNR values of

8 while the O3 background has much higher values. This confirms that an excess of significant noise triggers is present in MBTA's single detector triggers. This Gaussian noise analysis will be a reference for the effectiveness of the selection criteria defined in this chapter.

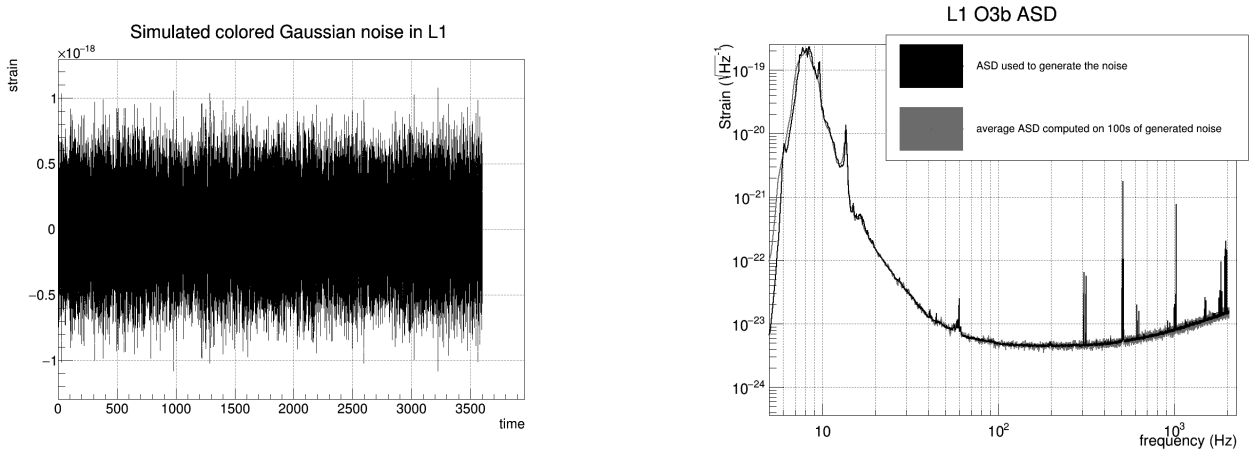


Figure 40: Left: One hour of simulated stationary colored Gaussian noise. Right: L1 O3b Amplitude Spectral Density (ASD = $\sqrt{\text{PSD}}$) used to color the noise and average ASD computed on 100s of the simulated noise.

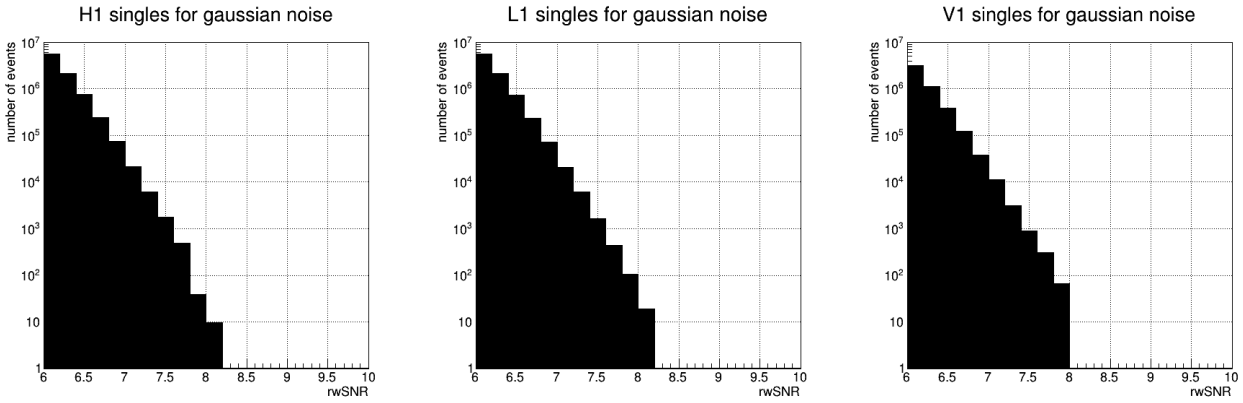


Figure 41: Single detector triggers for the BNS search region of a simulated Gaussian noise computed on about a month of simulated strain, scaled to the O3 effective observing time. Scaling factor for H1, L1 and V1 are 9.36, 9.13 and 9.06 respectively.

5.3 Monte Carlo simulations of astrophysical signals

To test the efficiency of the pipeline or of some tools on a large number of astrophysical events, we rely on Monte Carlo simulations. The simulated waveforms cover a large parameter space. These signals are then added to the real data of the detector and are called injections.

We can then analyze those data with injections using MBTA to see which injections were found and missed. An injection is considered as found, or recovered, if it passes the SNR (or cRS) threshold of the search and it is found within ± 100 ms of the injected time. Note that the PSD used for the filtering and the FAR vs cRS distributions are computed on the data without injections.

During the online analysis of O3, MBTA generated its own injections to validate the proper behaviour of the pipeline.

To compare pipelines' performances with each other, a set of common injections was made by the rates and populations LVK group. They were added on top of the full O3 data and processed by MBTA with the regular search configuration.

The BBH common injections were generated as follow:

- $2 < m_{1,2} < 100\,\mathrm{M}_\odot$, proba$(m_1) \propto m_1^{-2.35}$, proba$(m_2) \propto m_2$
- $|s_{1,2}| < 0.998$, uniform in spin magnitude, isotropic in spin orientation
- redshift $z < 1.9$, proba$(z) = d(\text{comoving volume})/dz \times d(t_{\text{source}})/dt \times (1+z)$
- frequency$\in [10, 1500]$ Hz

For the BNS common injections:

- $1 < m_{1,2} < 2.5\,\mathrm{M}_\odot$, uniformly distributed
- $|s_{1,2}| < 0.4$, uniform in spin magnitude, isotropic in spin orientation
- redshift $z < 0.15$, uniform in comoving volume-time
- frequency$\in [15, 1500]$ Hz

Finally for the NSBH common injections:

- $2.5 < m_1 < 60\,\mathrm{M}_\odot$, following a Salpeter power law with slope -2.35
- $1 < m_2 < 2.5\,\mathrm{M}_\odot$, uniformly distributed
- $|s_1| < 0.998$, $|s_2| < 0.4$, uniform in spin magnitude and isotropic in spin orientation for both
- redshift $z < 0.25$, uniform in comoving volume-time
- frequency$\in [15, 1500]$ Hz

BNS injections were generated using the SpinTaylorT4threePointFivePN waveform, NSBH and BBH were generated using SEOBNRv4pseudoFourPN. Figure 43 shows the masses and distance distribution of the O3 common injections.

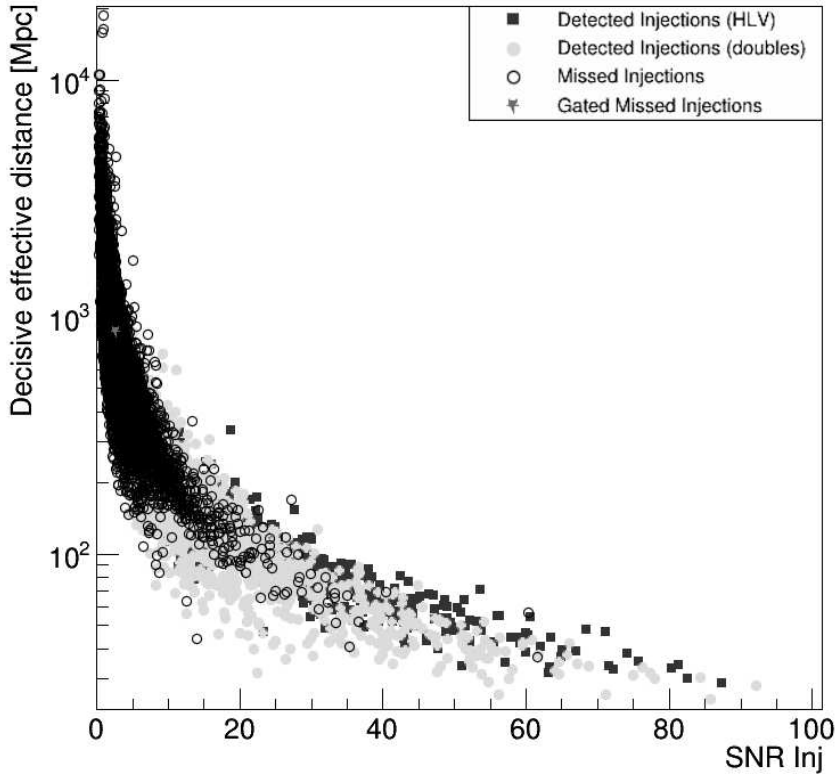


Figure 42: Missed and found injections. Plots of the MBTA production monitoring for BNS injections on around 6 days.

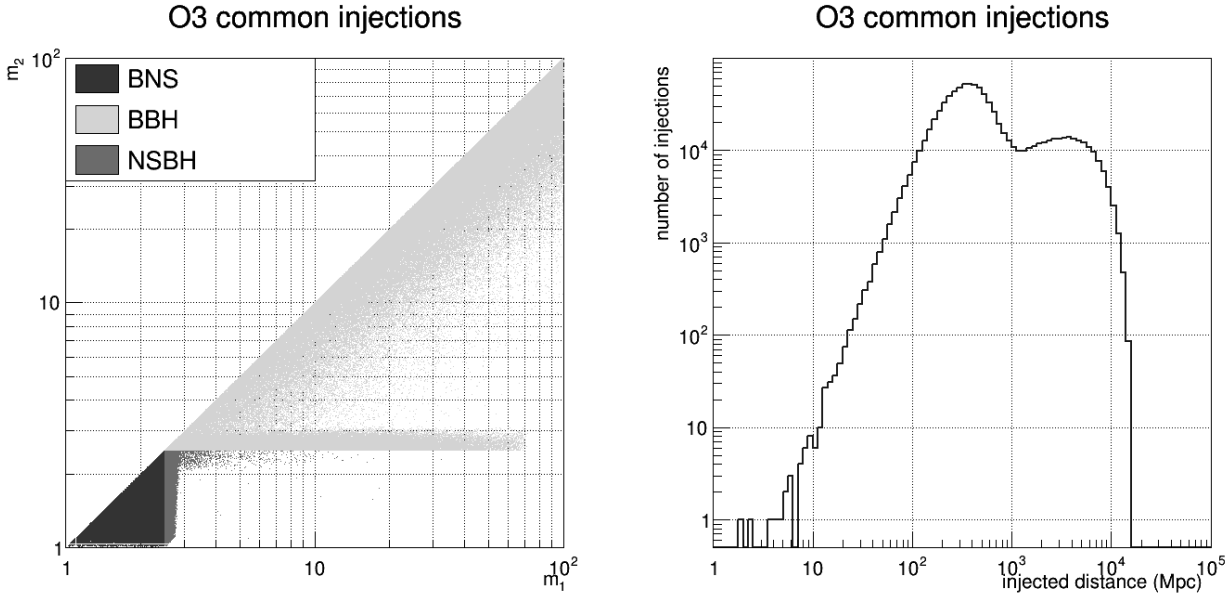


Figure 43: O3 common injections. Left: distribution of the masses for the injections. Right: distance distribution of the injections.

5.4 Defining the EM bright population for single detector triggers

We want to define a population of interest for the single detector trigger search. This population, which we call EM bright, should contain most of the candidates susceptible of having an electromagnetic counterpart. The most common process that is expected for the emission of EM waves is through the presence of a remnant mass [91]. Thus it becomes natural to define this EM bright population according to the hasRemnant quantity computed by the pipeline (see section 4.12). To be safe we decide to consider as EM bright any candidate that has hasRemnant$> 0.1\%$. However the hasRemnant quantity was not computed for O3 injections as it was added after O3 catalogs were produced. Therefore in this PhD work, we translate this definition in terms of just masses because the recovery of the spins is usually poor. Figures 44 and 45 show the hasRemnant value as a function of the individual masses and chirp mass for an injection run on 40 days of O3 replay data.

Based on these plots, we choose to define the EM bright population as follows:

$$1\,\mathrm{M}_\odot \leq m_1 \leq 50\,\mathrm{M}_\odot, \tag{61}$$

$$1\,\mathrm{M}_\odot \leq m_{chirp} \leq 5\,\mathrm{M}_\odot \tag{62}$$

which includes all recovered injections with hasRemnant$> 0.1\%$. We believe that this definition is conservative enough and should encompass more than just the EM bright candidates. We use those constraints on the detected parameters as given above without concern for the effect of redshift since we do not expect high redshift for such sources.

Anything that is not part of this EM bright population is called EM dark.

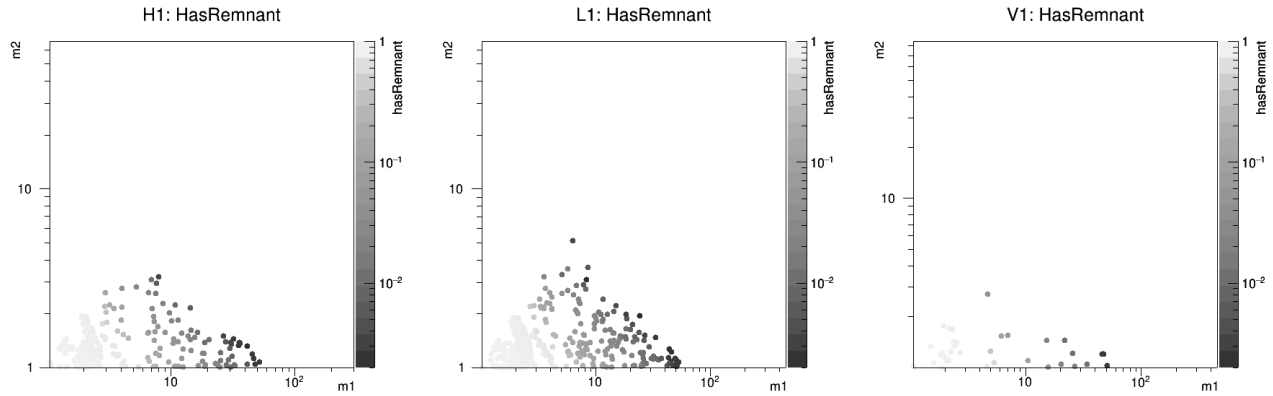


Figure 44: hasRemnant $> 1\%$ versus m_1 and m_2 for single detector triggers during an injection run on O3 replay data.

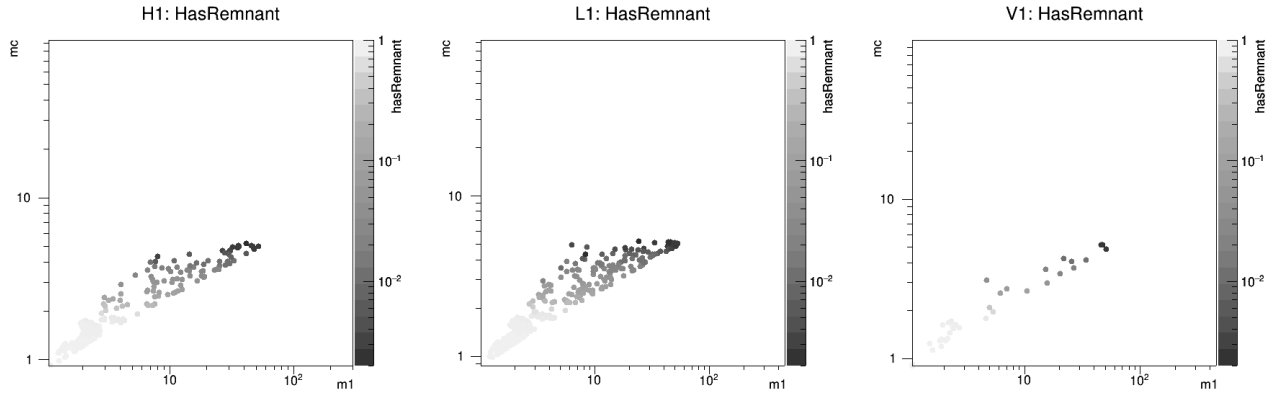


Figure 45: hasRemnant $> 1\%$ versus m_1 and m_{chirp} for single detector triggers during an injection run on O3 replay data.

5.5 Restraining the parameter space

We want to focus on the EM bright triggers, as defined in section 5.4 for the single detector triggers analysis. We show as a reference the background distribution for O3 singles detector triggers saved at all detector times in figure 46 (unlike figure 39 which was for double and triple detector time only). We easily verify that the EM bright population produces less noise triggers than the EM dark one by looking at the rwSNR distributions in figures 47 and 48.

We know that we want to keep the EM bright candidates but we wonder whether we could also keep a part of the EM dark ones. Since, as mentionned previously, the duration of the template may have an importance on the discrimination between noise and signal, we want to investigate this point. Figure 49 shows, for each frequency bin, the number of single detector triggers saved during O3 divided by the number of template (O3 bank, all regions mixed). The density of template as a function of the template duration is also shown. We see that short templates produce way more triggers than the long ones. Excluding such noisy templates from the single detector trigger search should therefore allow to reduce significantly the background. We can go further with figures 50 and 51. They show respectively the distribution of EM bright and EM dark single detector triggers in the rwSNR versus template duration plane and confirm that EM bright templates are generally longer. We have therefore no reason to split the EM bright population based on the template duration. EM dark templates however exhibit many triggers with short duration and large rwSNR. What appear as vertical lines in figure 51 are very noisy templates which produce many triggers. To exclude these triggers we decide to split the EM dark population in a "long EM dark" population with template duration $\geq 0.8\,\mathrm{s}$ and a "short EM dark" population with template duration $< 0.8\,\mathrm{s}$.

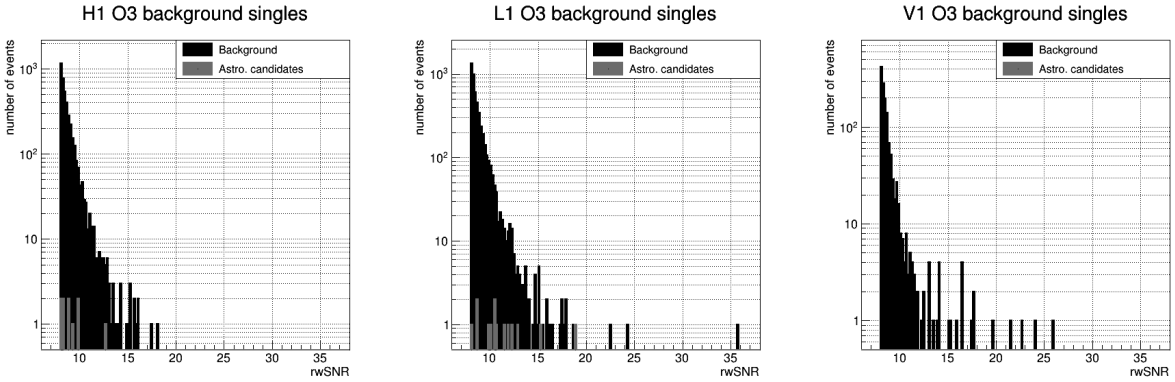


Figure 46: MBTA O3 single detector triggers from the offline analysis.

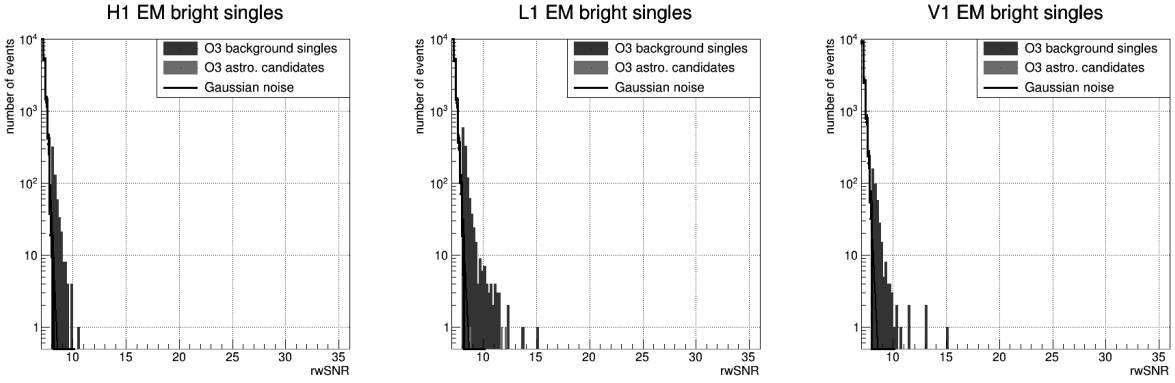


Figure 47: rwSNR distribution for O3 EM bright single detector triggers before any specific single detector trigger selection.

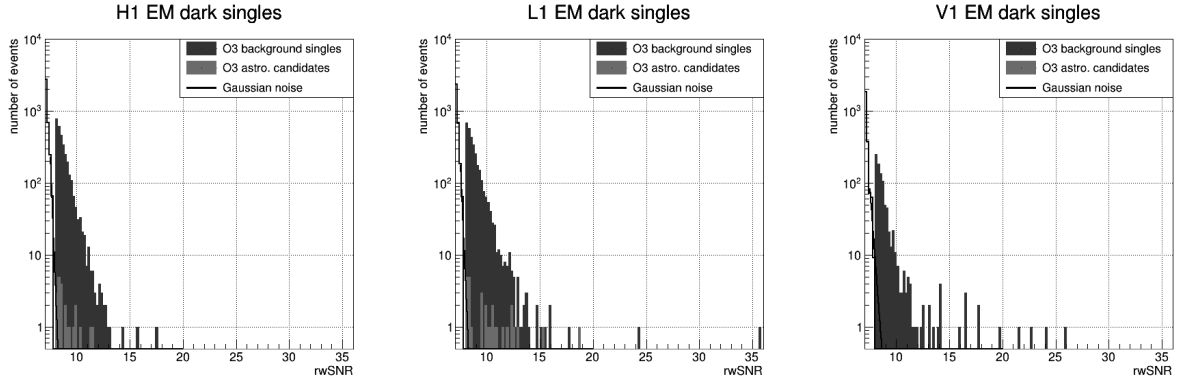


Figure 48: rwSNR distribution for O3 EM dark single detector triggers before any specific single detector trigger selection.

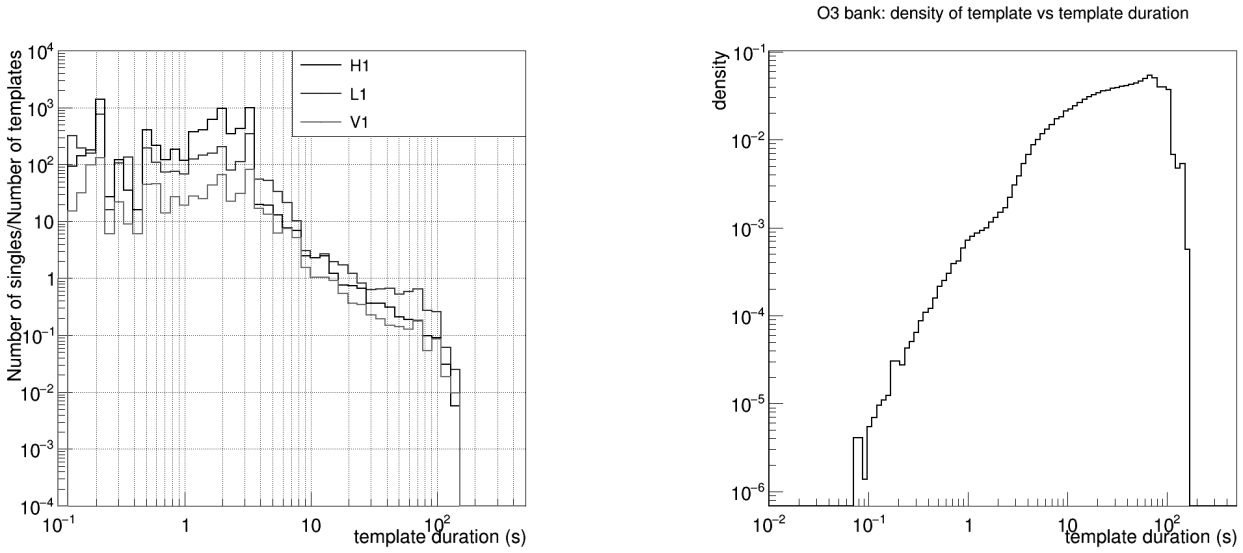


Figure 49: Left: Number of MBTA O3 single detector triggers per template as a function of the template duration. Right: Density of templates (number of templates per bin relative to the total number of templates) as a function of the template duration for MBTA's O3 template bank.

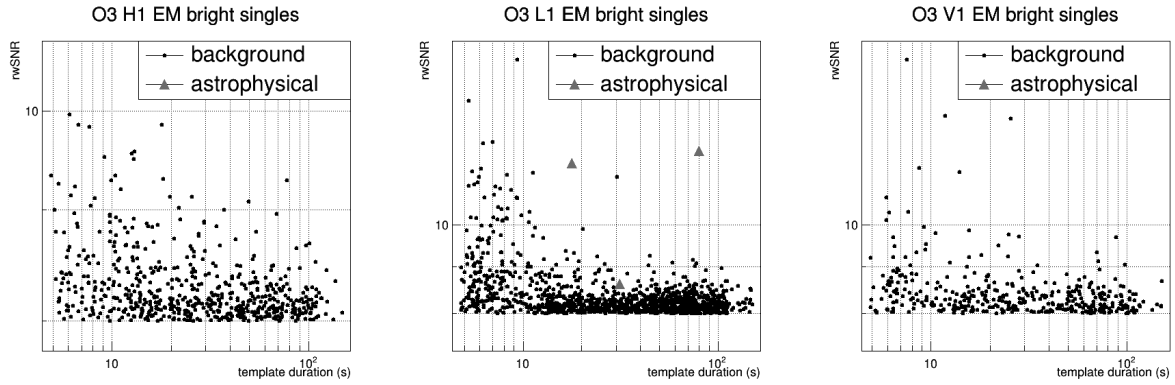


Figure 50: Distribution of the O3 EM bright single detector triggers in the template duration vs rwSNR plane.

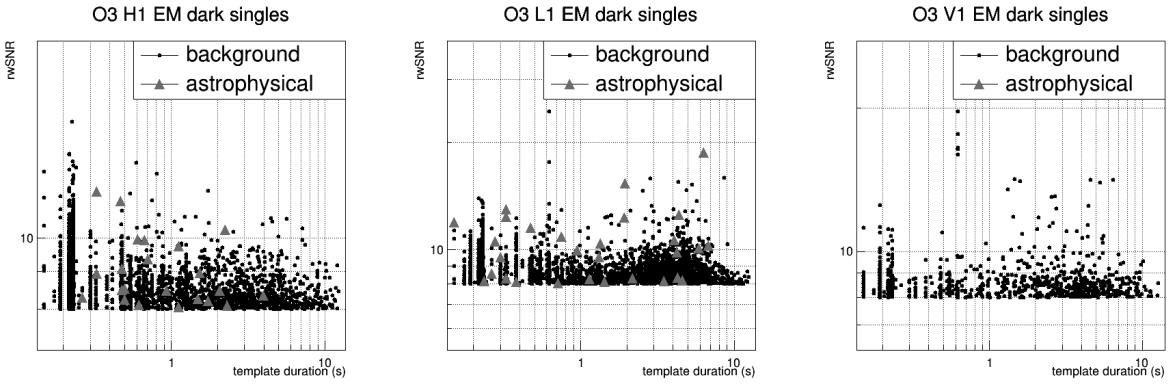


Figure 51: Distribution of the O3 EM dark single detector triggers in the template duration vs rwSNR plane. A few short templates are triggered very often, showing as vertical lines.

5.6 Selection using the autoχ^2

The first quantity we consider to define specific single detector trigger selection criteria is the autoχ^2 . It is already used to compute the reweighted SNR but does not remove all noise triggers and relies on coincidences to clean what is left. Here we want something more powerful even at the cost of some loss of duty cycle. We therefore decide to investigate a cut on the autoχ^2 .

There are no physical motivations to choose a specific cut value, but we know that the autoχ^2 has a tendency to grow with the SNR. The cut should therefore take this into account to avoid rejecting loud astrophysical signals. We use a set of injections to choose a cut that allows to reject background triggers while only removing at most a few percent of the injections. The tentative cut is the following:

$$\text{auto}\chi^2 \leq 2 + 0.005 \times \text{SNR}^2 \tag{63}$$

The quadratic dependence in the SNR is motivated by the formula of the autoχ^2 . More details about this SNR dependence of the autoχ^2 are given in section 7.1. This dependence should not change from one O3 region of the parameter space to the other since the banks are made with the same minimal match parameter. This cut is therefore tested for the three regions all together.

The corresponding threshold is shown in figure 52, drawn over the distribution of recovered BBH injections in the plane autoχ^2 vs SNR. The fractions of astrophysical injections surviving the cut in the 3 regions are given in table 1. When taking this cut and applying it to O3 EM

	H1	L1	V1
BNS	0.999	0.995	0.995
BBH	0.980	0.954	0.997
NSBH	0.982	0.974	0.982

Table 1: Fraction of injections surviving the autoχ^2 cut for each search region and detector.

bright singles it turns out that none of them falls outside of the acceptable (SNR, autoχ^2) values as shown by figure 53. EM dark single detector triggers are barely affected and the tail of their rwSNR distribution is unchanged. As can be seen on figure 52, a more aggressive cut would start removing astrophysical signals. This cut is therefore not considered further.

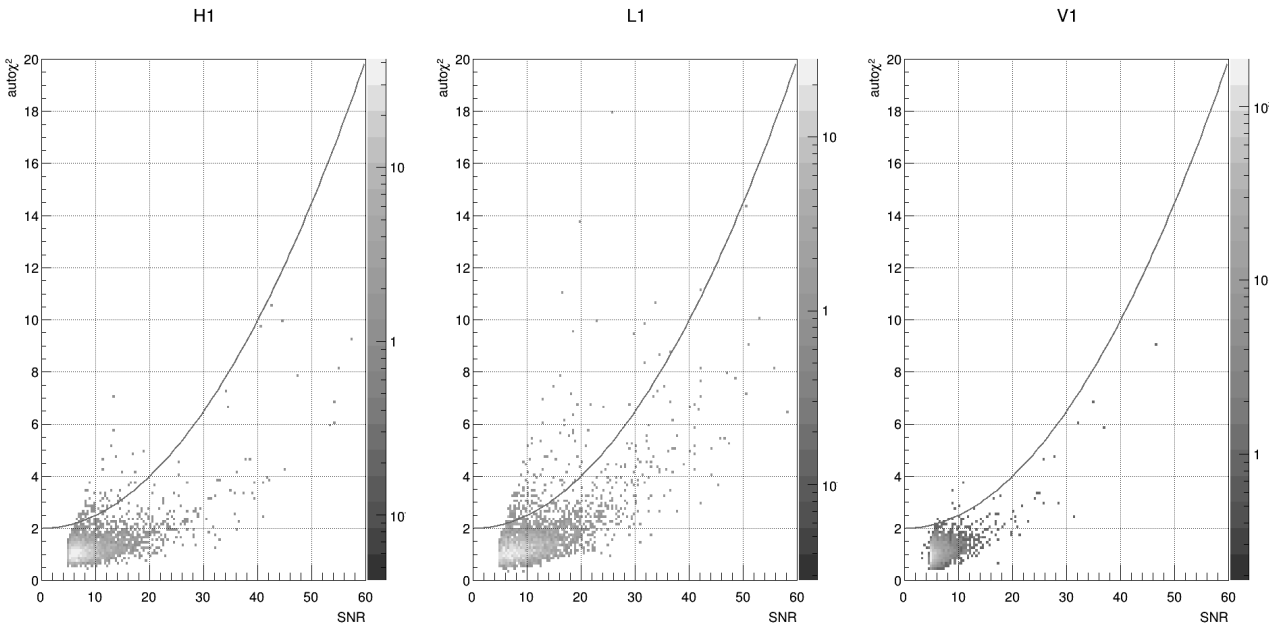


Figure 52: BBH injections distribution in the autoχ^2 vs SNR plane. The tentative threshold chosen for the cut is drawn in red: anything above the line is rejected.

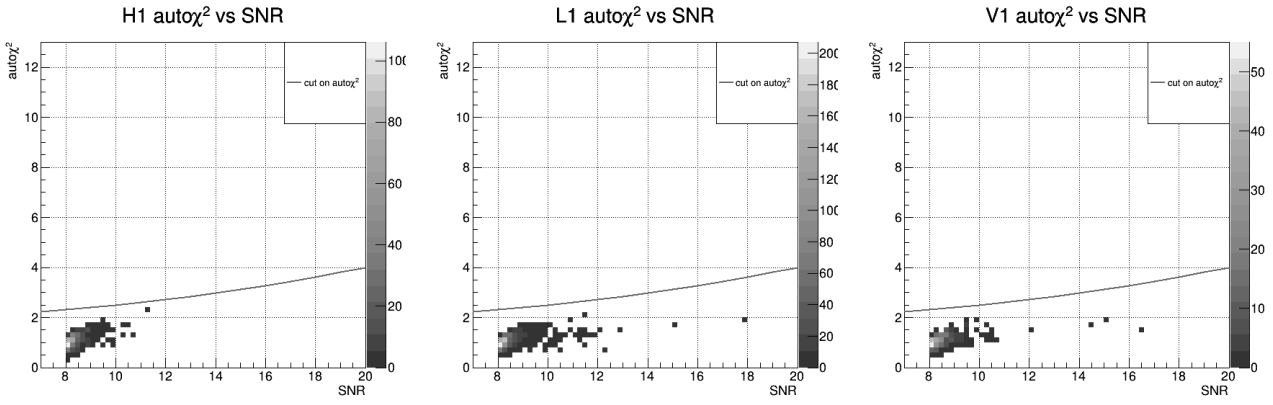


Figure 53: O3 EM bright single detector triggers distribution in the autoχ^2 vs SNR plane. The tentative threshold chosen for the cut is drawn in red: anything above the line is rejected.

5.7 Selection using the excess rate

The second quantity we consider to define selection criteria is the excess rate (ER). We want to reject triggers that were detected during bad data quality times and this is exactly what the ER is telling us. Our approach is straightforward: we want to define a threshold on the ER above which candidates will be rejected. Reweighting the SNR as for the coincidences could work but very loud noise triggers could still be quite significant after reweighting. We therefore choose to apply a cut to be on the safe side.

5.7.1 Which excess rate should we use?

There were 3 types of ER during O3, one for each search region (BNS, BBH, NSBH) but this will not be the case for O4 anymore. We can show (figure 54) that the different ER are correlated. Since we want to find a cut value for these excess rates that would allow to reduce the tail of the background SNR distribution, this correlation is a good argument to be even more restrictive by cutting on the maximum of the three excess rates.

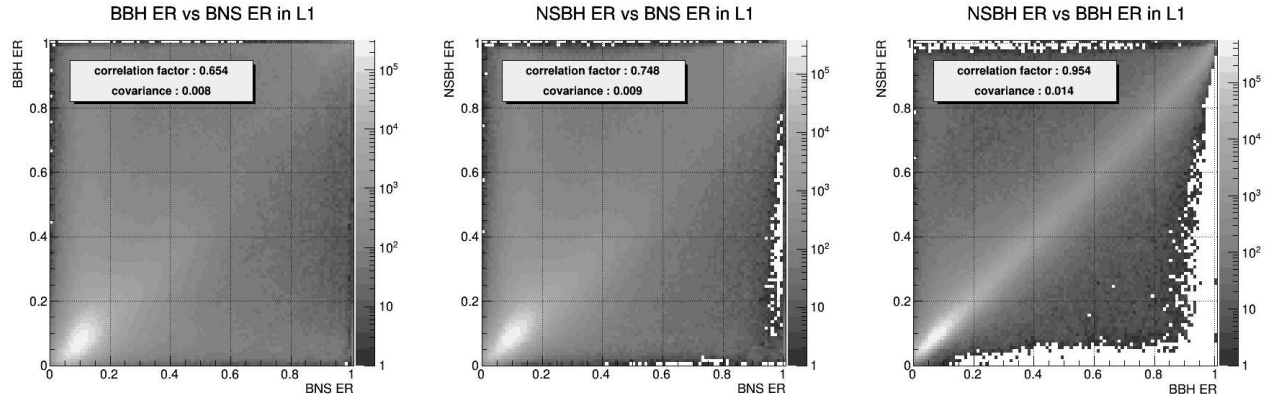


Figure 54: Pairwise comparison of the 3 types of excess rates observed during O3 in L1. Correlations factor greater than 0.5 indicate strong correlation between the different ER. The correlation factor is the covariance of the excess rates divided by the product of their standard deviation.

5.7.2 Excess rate on EM bright single detector triggers

Figure 55 shows the effect of a cut rejecting all background trigger at times with ER > 0.3 on the EM bright SNR distribution for the 3 excess rates and their maximum. All confirmed candidates published in the GWTC-2.1 [92] and GWTC-3 [37] catalogues were removed. The maximum ER gives by construction the best result with only few SNR values above 9 and does not reject much more duty cycle than the others, which comforts us in choosing it to apply a cut.

The value of ER $= 0.3$ for the cut is equal to the threshold used to reweight the SNR with the excess rate for coincidences. Softer values of cut did not clean as well the SNR distribution. Using a harsher cut would reduce even more the duty cycle while not removing much more background triggers as shown by figure 56. Thus we settle for a cut at ER $= 0.3$. But this will have to be checked again with O4 noise.

Figure 57 shows the effect of the cut on the EM bright background along the astrophysical candidates given in table 2 (with parameter within the EM bright region, coming from the GWTC2.1

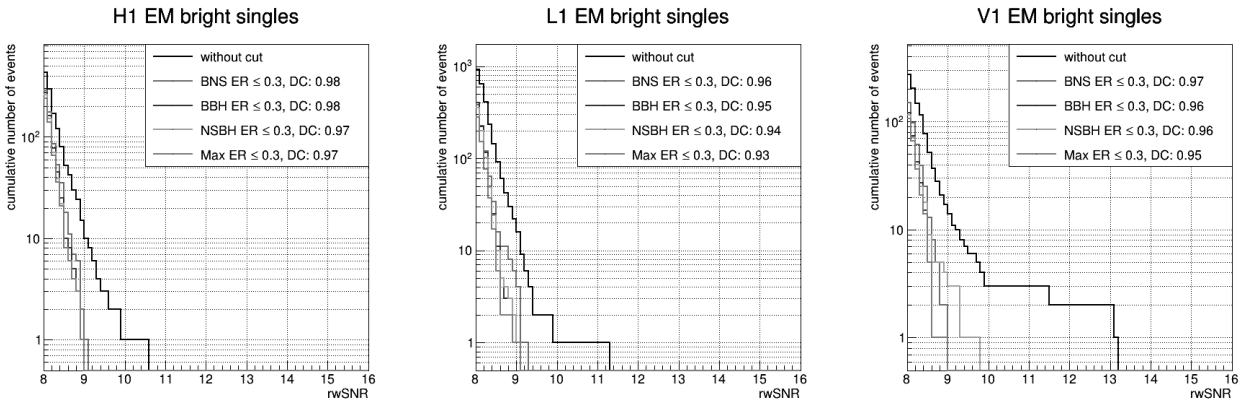


Figure 55: Cumulative rwSNR distribution of O3 background EM bright single detector triggers for a cut on the different excess rates. The legend also indicates in each case the remaining duty cycle (DC).

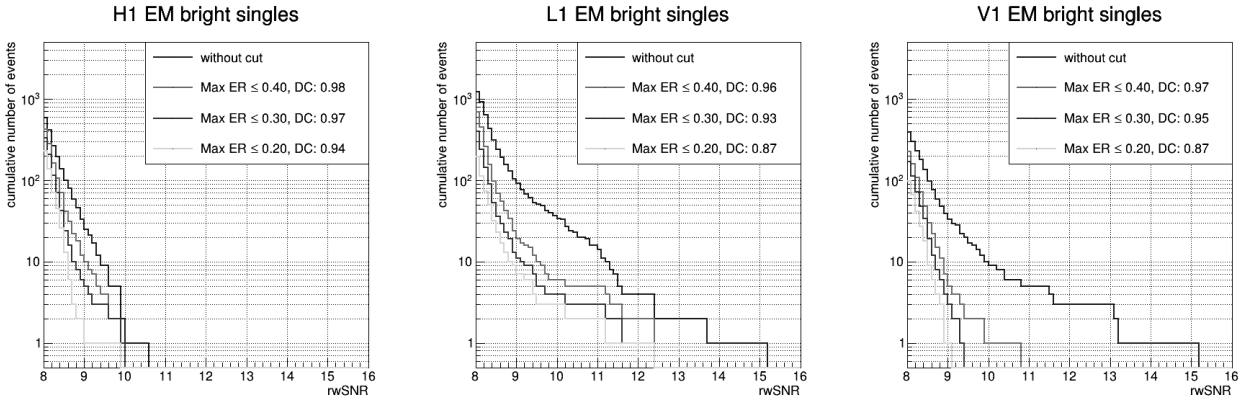


Figure 56: Cumulative rwSNR distribution of O3 background EM bright single detector triggers for different excess rate cuts and remaining duty cycle (DC) for each.

[92] and GWTC3 catalogs [37]). Plots are only shown for L1 because it is the most sensitive out of the three detectors and the only one in which we have single detector triggers from EM bright astrophysical signals. We see that we have rejected many triggers at high rwSNR but the astrophysical signals, although not impacted by the cut, still do not stand out of the distribution because of a few significant background triggers remaining.

Name	GPS time	rwSNR	ER
GW190425	1240215503.027	12.07	0
GW200105_162426	1262276684.066	11.69	0
GW200115_042309	1263097407.744	8.63	0.066

Table 2: MBTA's EM bright single detector triggers associated to GWTC-2.1 and GWTC-3 astrophysical events.

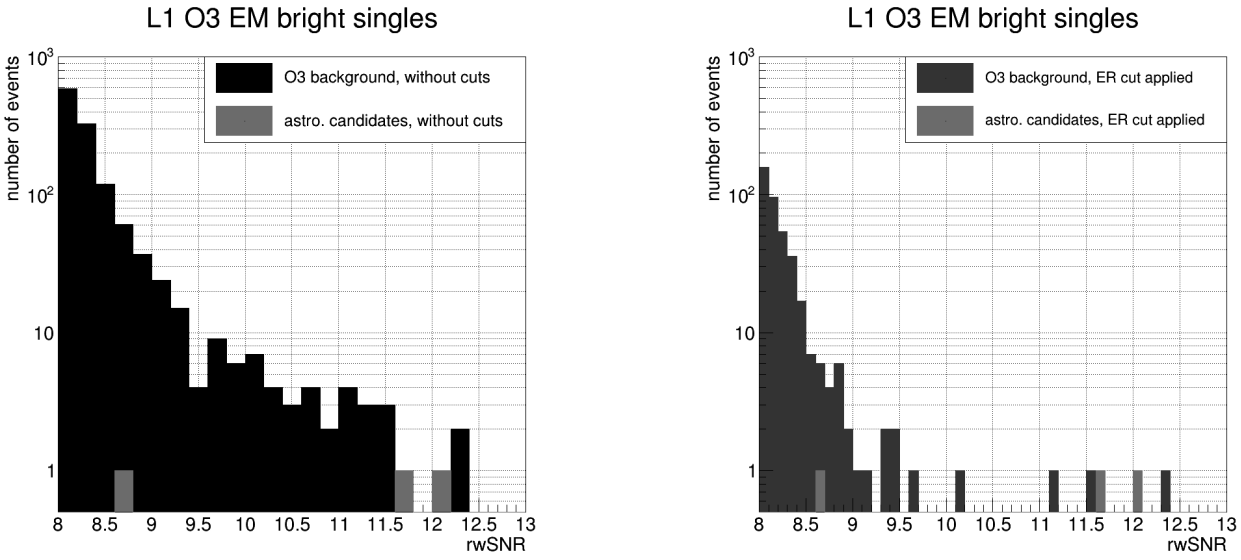


Figure 57: rwSNR distribution of O3 background EM bright single detector triggers and astrophysical candidates before (left) and after (right) applying the excess rate cut.

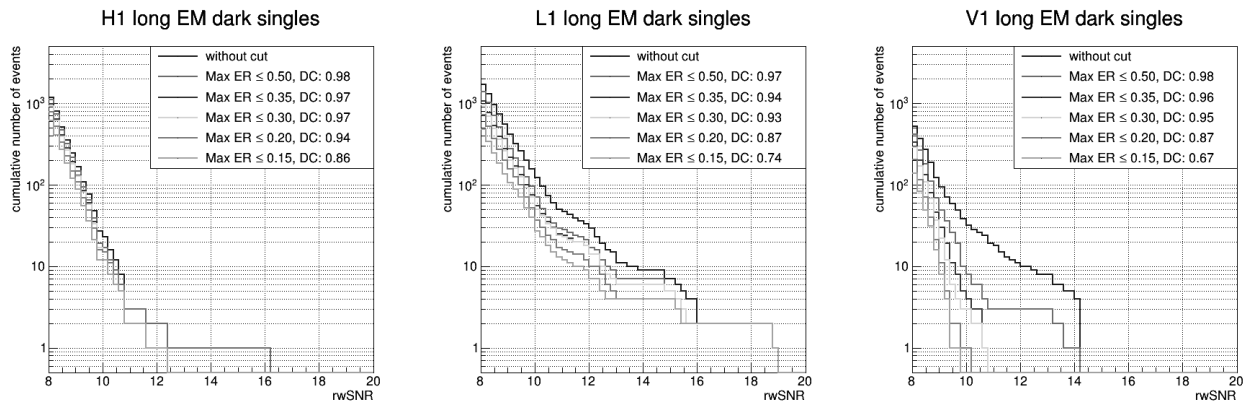


Figure 58: Different excess rate cuts on O3 long EM dark single detector triggers and remaining duty cycle (DC) for each.

5.7.3 Excess rate on EM dark single detector triggers

Since for the EM dark region we have more single detector triggers ($> 10^3$ per detector) and we are looking at higher masses, we wonder whether we should be stricter on the excess rate cut. To this end we will also investigate cuts below ER = 0.3. Figure 58 shows the SNR distribution of the EM dark population for different values of cut on the max ER. We see that cutting on the max ER allows for some background rejection but the value of the cut has little impact unlike for EM bright triggers. Since being more restrictive does not make things significantly better and aggressive cuts can be dangerous (risk of missing astrophysical events due to lower duty cycle), we settle for a cut on the excess rate at 0.3. The comparison with the astrophysical candidates is shown in figure 59.

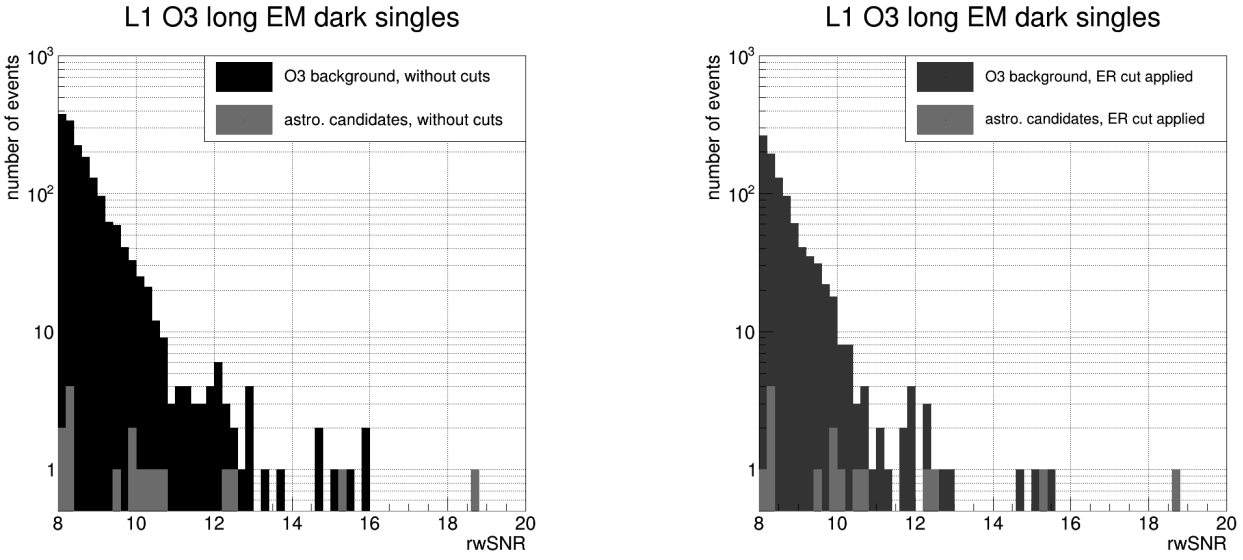


Figure 59: rwSNR distribution of O3 background EM dark single detector triggers and astrophysical candidates before (left) and after (right) applying the excess rate cut.

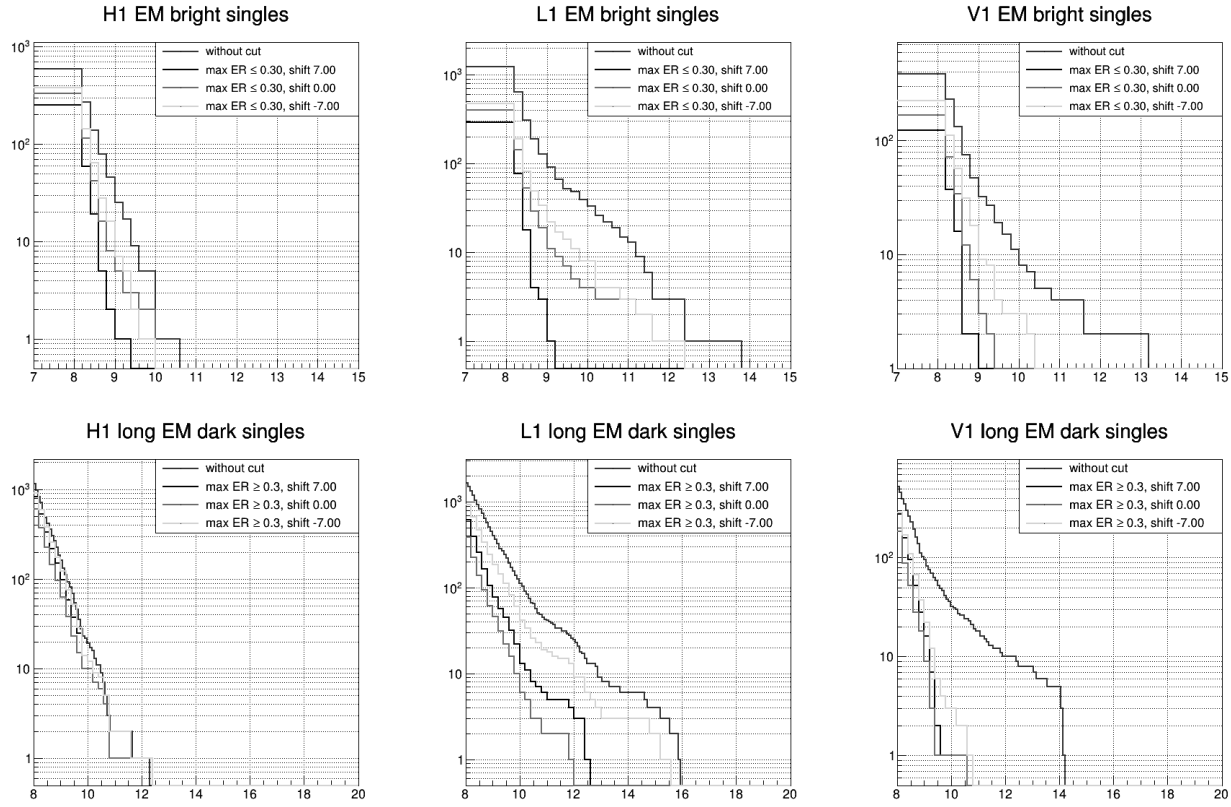


Figure 60: rwSNR distribution with different time shifts on the excess rate. EM bright (top) and long EM dark (bottom) single detector triggers.

5.8 Need of a larger cut on the excess rate and gating

We wondered at some point whether the time at which we take the excess rate for the single detector triggers was of high importance for our cut, or if a shift of a few seconds would not change much. The reason for this thought is that the ER is computed as a median value over 10 s so a shift in time may bring some differences. To investigate this question, we apply the selection with different time offset, as presented in figure 60. We see that the delay does not matter much in H1 and V1 but there are some more significant differences in L1 for both EM bright and long EM dark single detector triggers. Since this effect appears only in L1 it is not clear at this point wether it is actually due to the delay introduced for the excess rate. We can further investigate the matter by looking at events with large rwSNR that were not rejected previously by the ER cut. Figure 61 shows the excess rate vs time with indication of gated times. We see that there is clearly an issue with those single detector triggers: the presence of gating close to the trigger indicates a bad period of time and the spike in the excess rate right after the trigger explains why shifting the time at which we take the excess rate had an impact on the SNR distribution. Further investigation showed that these single detector triggers were not isolated cases and many other displayed such behaviour.

We therefore decide to reject any single detector trigger for which a gated time is found within $[-20, +5]$ s and we require that the excess rate should not exceed 0.3 in the time window $[-7; 7]$ s centered on the trigger. The result is shown in figures 63, 64 for the EM bright single detector triggers (to be compared to figure 62) and 66, 67 for the EM dark single detector triggers (to

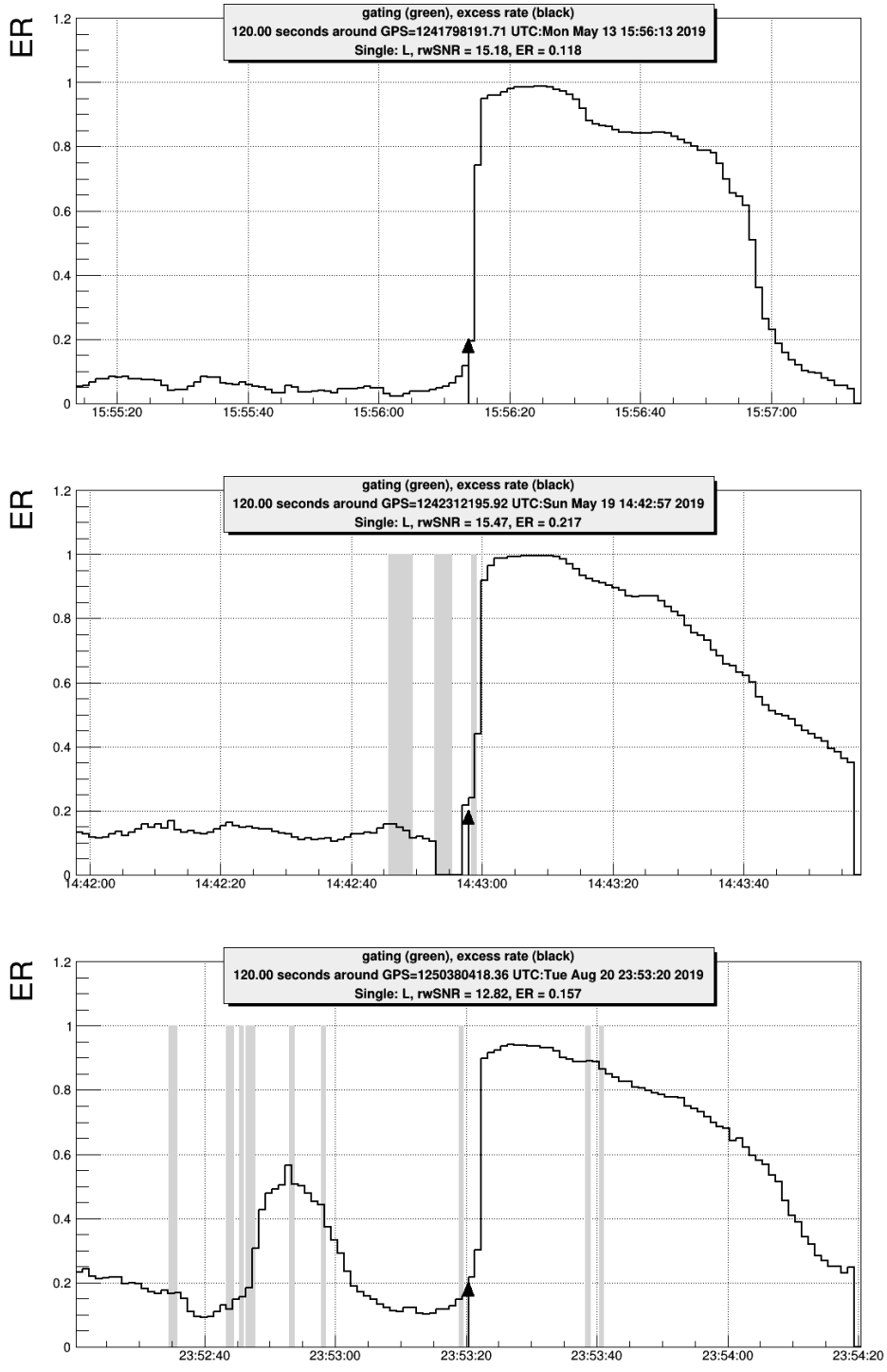


Figure 61: Gating and excess rate around some single detector triggers in L1. The time of the trigger is indicated by an arrow each time.

be compared with figure 65). As can be seen the tail of the distribution is cleaner and the total number of events has decreased. The total loss of duty cycle for O3 due to the application of those cuts is 9.1% in H1, 15.4% in L1 and 9.8% in V1.

As a final plot for the selection we show the comparison of the EM bright and long dark triggers distribution with the Gaussian noise after the selection in figures 68 and 69. We also show the template duration versus rwSNR distribution of EM dark single detector triggers in figure 70 to confirm that the separation between long and short templates still holds. The separation based on the duration is still meaningful after applying the selection as we can still see some noisy templates with very short duration.

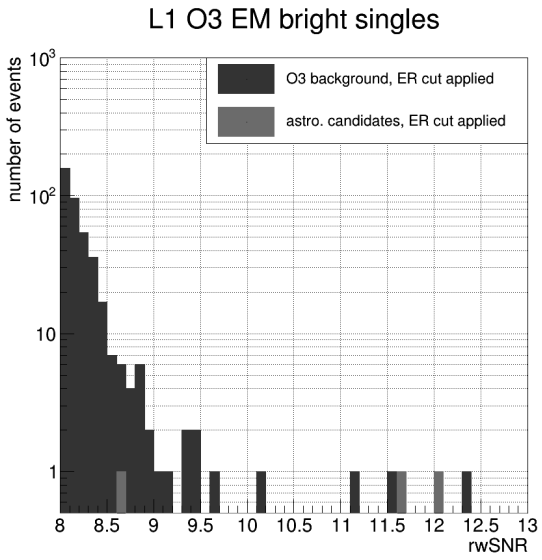


Figure 62: rwSNR distribution for EM bright single detector triggers with the simple excess rate cut described in section 5.7.3.

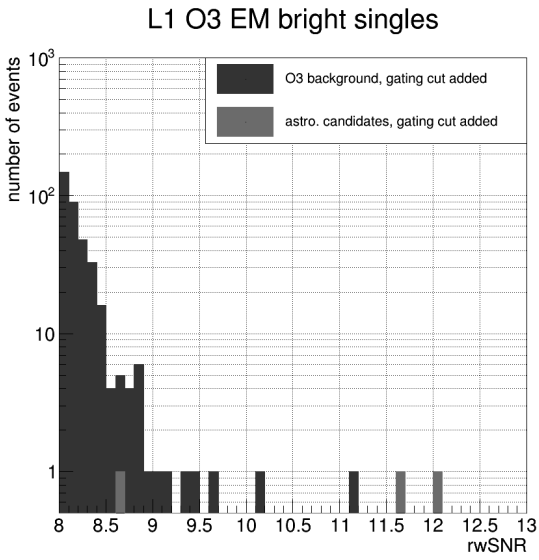


Figure 63: rwSNR distribution for EM bright single detector triggers after applying the cut on the gating.

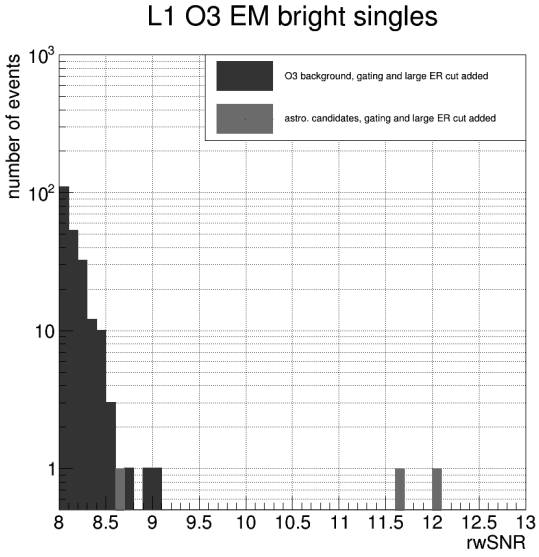


Figure 64: rwSNR distribution for EM bright single detector triggers after applying the larger cut on the excess rate in addition to the cut on the gating.

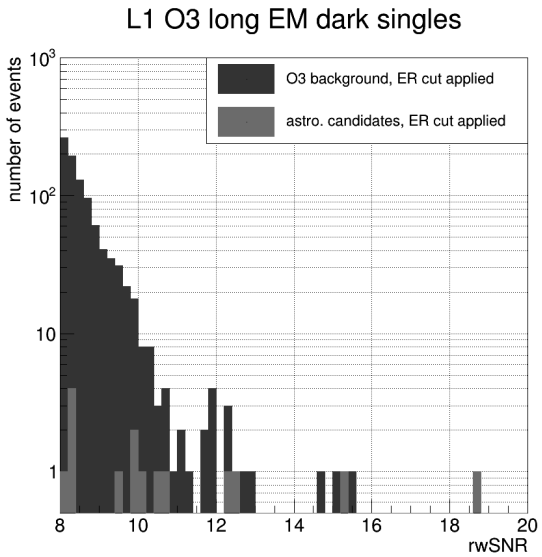


Figure 65: O3 long EM dark single detector triggers: SNR distribution with the simple excess rate cut described in section 5.7.3.

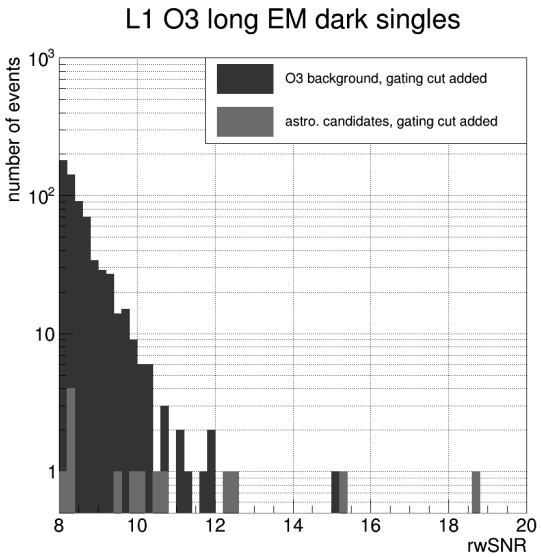


Figure 66: O3 long EM dark single detector triggers: SNR distribution after applying the cut on the gating.

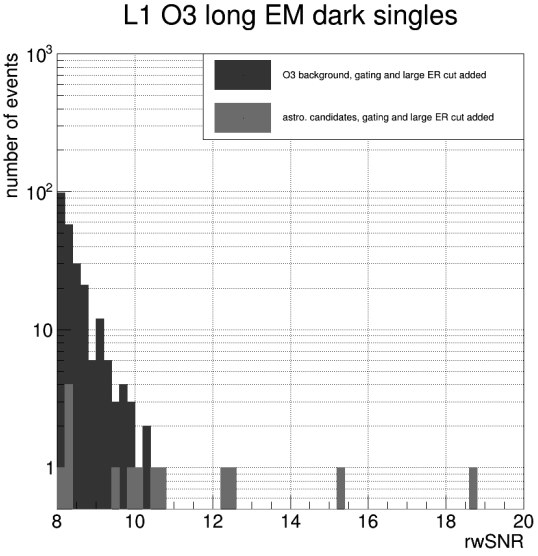


Figure 67: O3 long EM dark single detector triggers: SNR distribution after applying the larger cut on the excess rate in addition to the cut on the gating.

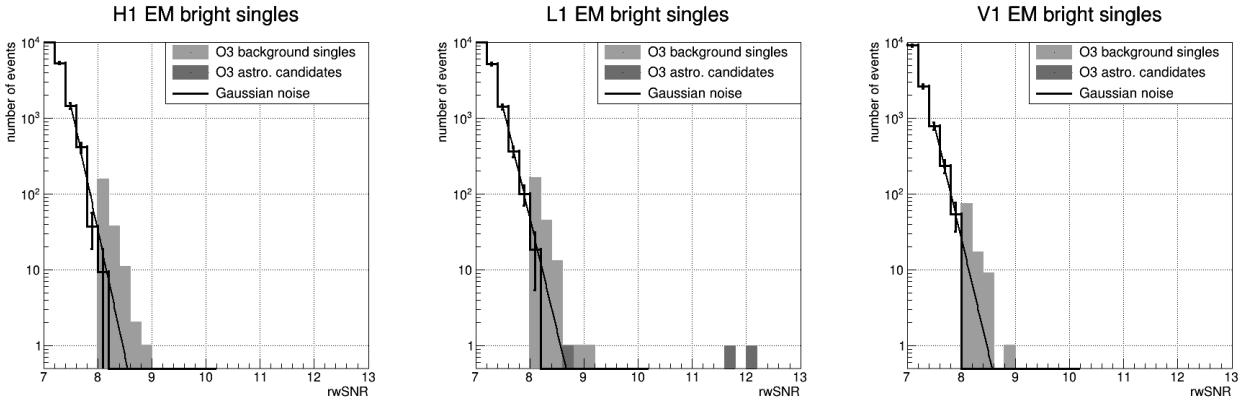


Figure 68: rwSNR distribution: O3 EM bright single detector triggers after selection (gating and wide ER cut) vs Gaussian noise

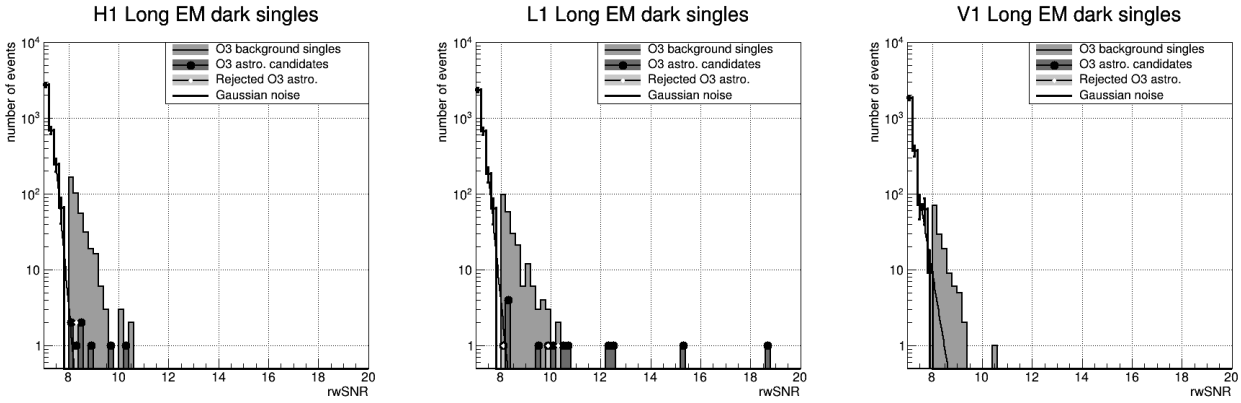


Figure 69: rwSNR distribution: O3 Long EM dark single detector triggers after selection (gating and wide ER cut) vs Gaussian noise

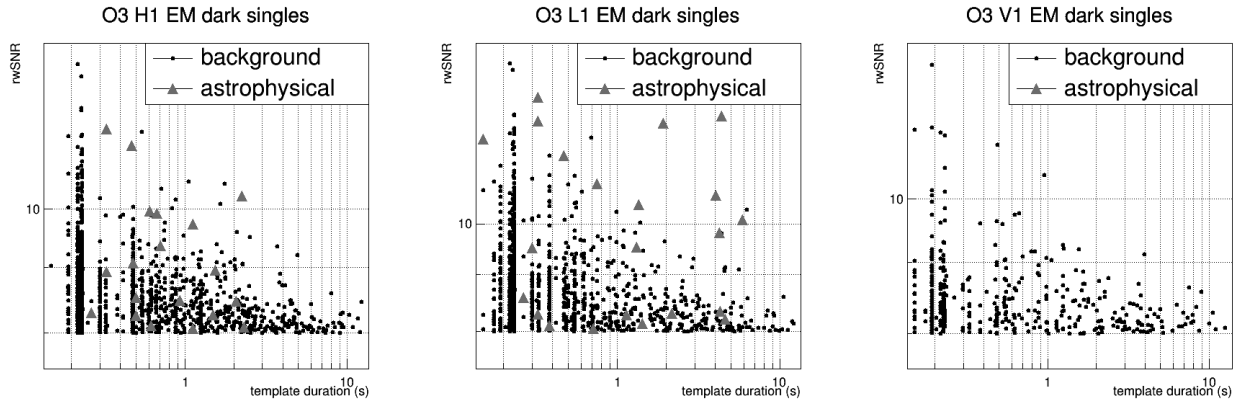


Figure 70: Distribution of the EM dark single detector triggers in the template duration vs rwSNR plane. The cuts on the gating and excess rate are applied. A few short templates are still triggered very often, showing as vertical lines.

5.9 Search sensitivity improvement

We have seen that we can reject background by cutting on the excess rate and gating at the cost of some duty cycle. We want to know how much was gained in terms of efficiency of the search.

One way to answer is by looking at the Volume-Time (VT) (see section 3.2). This VT is roughly proportional to the number of detections. The volume scales with the distance cubed, which in turn scales with the inverse of the SNR threshold (eq. 14) (or rwSNR in the present case). Having lower background means that astrophysical signals with smaller rwSNRs can pass the threshold for detection. This means an increase in volume. Hence, we have on one side a loss in observing time and on the other a gain in volume. We therefore need to estimate both to anwser the initial question.

As mentioned previously, the effective observing time loss caused by the application of the selection is 9.1% in H1, 15.4% in L1 and 9.8% in V1. Note that during these times with excess rate, the search is less efficient because of the reweighting and there is already some VT loss. Therefore when cutting these times, the VT reduction is less than the sole observing time loss.

Looking at EM bright single detector triggers in L1 (figure 71) and considering a FAR thresold of 2 per year, the ranking statitstics threshold using the simple reweighting would be ~ 11.2. When applying the selection criteria this threshold becomes ~ 8.8. Thus, in first approximation, the gain in volume is $(11.2/8.8)^3 - 1 = 106.2\%$. Following the same reasoning, the improvement in H1 and V1 is $(9.5/8.6)^3 - 1 = 34.8\%$ and $(9.4/8.5)^3 - 1 = 35.2\%$ respectively. This gain is much larger than the loss in duty cycle, validating the use of the selection criteria. The same conclusion can be derived from figure 72 for long EM dark single detector triggers.

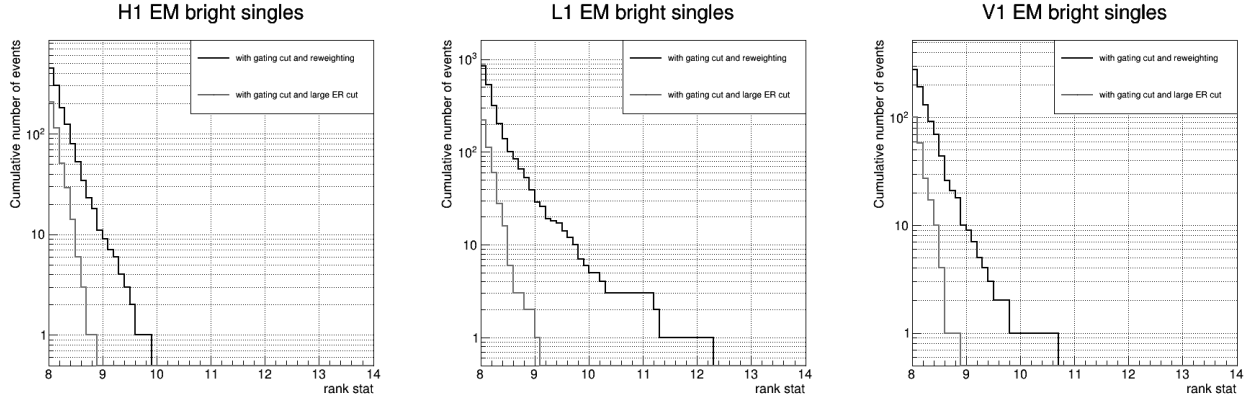


Figure 71: Reweighting vs cutting on the excess rate for O3 EM bright single detector triggers : gating and wide ER cuts.

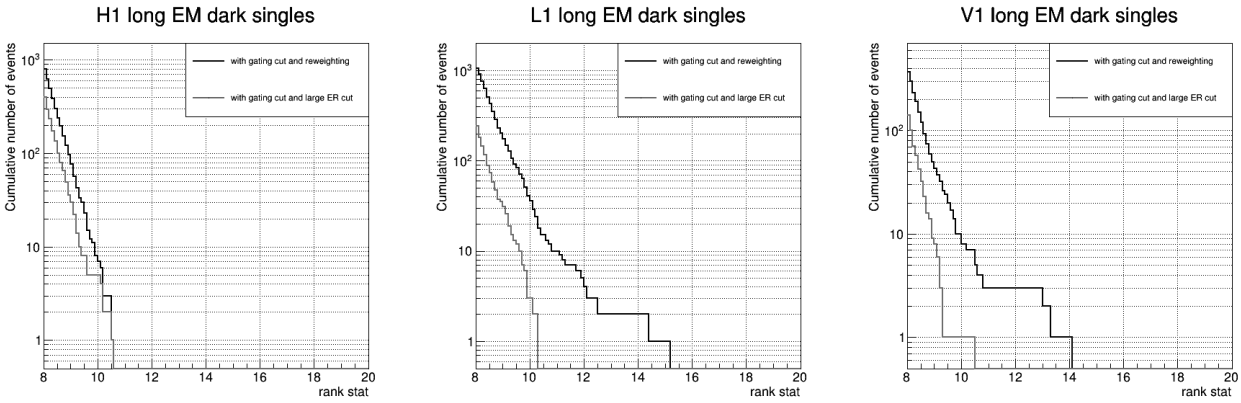


Figure 72: Reweighting vs cutting on the excess rate for O3 long EM dark single detector triggers : gating and wide ER cuts.

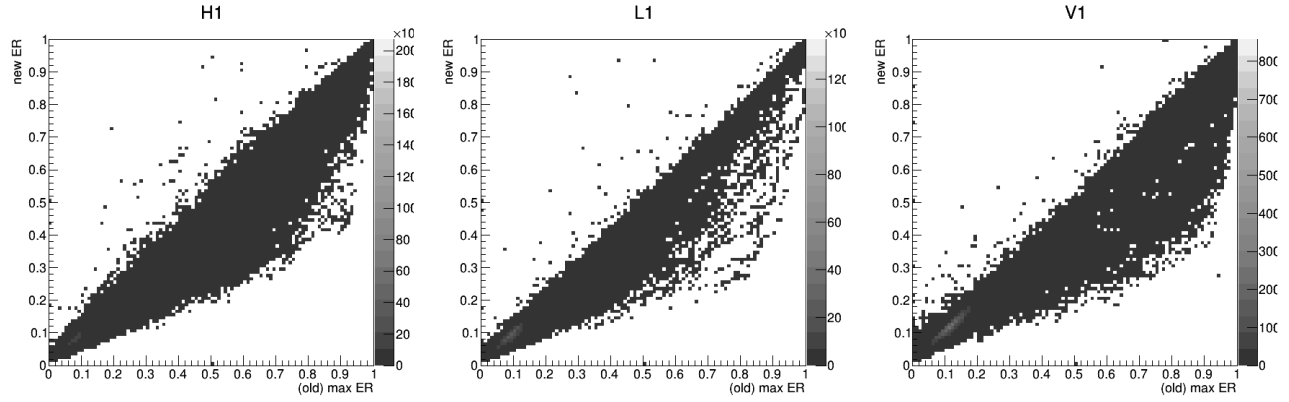


Figure 73: Comparison between the old max ER and the new ER.

5.9.1 Computation of a new ER in anticipation for O4 bank

We have shown what can be achieved by applying cuts on the excess rate and the gating. We have until now used the maximum of the 3 search regions excess rates as explained in section 5.7.1, but for O4 we will have only one region. In this section we explore how the result will change with an excess rate computed on only one region.

The excess rate is computed from the trigger rates before and after data quality checks following equation (44). For O3 it is computed for each search region by taking independently the rates for each of them. To achieve our goal we can simply add up those rates to have one total rate before and after quality checks. We can then compute a single excess rate following equation 44. Figure 73 shows the comparison between the maximum of the three search regions' excess rates and the newly computed excess rate. As can be seen the new one has a tendency to be lower than the maximum ER. This is expected since in the case where one of the 3 ER is high while the other two have smaller values, the maximum will be high while our new ER will be closer to a mean value of the 3.

We can now apply the same cuts as before with a substitution of the max ER by the new ER to compare the effect of both. The result is shown in figures 74 and 75 for EM bright and long EM dark single detector triggers respectively. As can be see, there is little difference between the 2 excess rate. A few more events survive the cuts when using the new ER since it is generally lower but it does not significantly affects the tail of the distributions. We can also note that using the new ER, we retrieve one more astrophysical event among the short EM dark single detector triggers in L1. Figures 76 and 77 show the comparison with Gaussian noise after the full selection. The total loss of duty cycle when applying the cut on the gating and new ER is of 6%, 14.1% and 9% for H1, L1 and V1 respectively. For comparison, using the old (max) ER, the loss was of 9.1%, 15.4% and 9.8% for H1, L1 and V1.

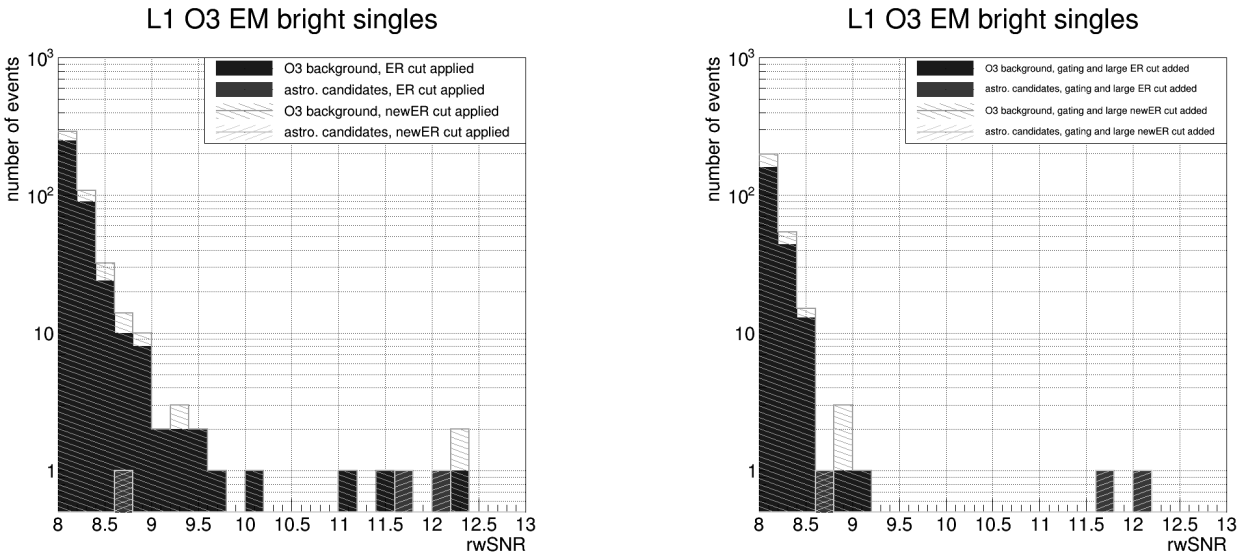


Figure 74: rwSNR distributions for L1 O3 EM bright single detector triggers with ER selection (left) and ER+gating selection (right) using the old (max) and new ER (common region).

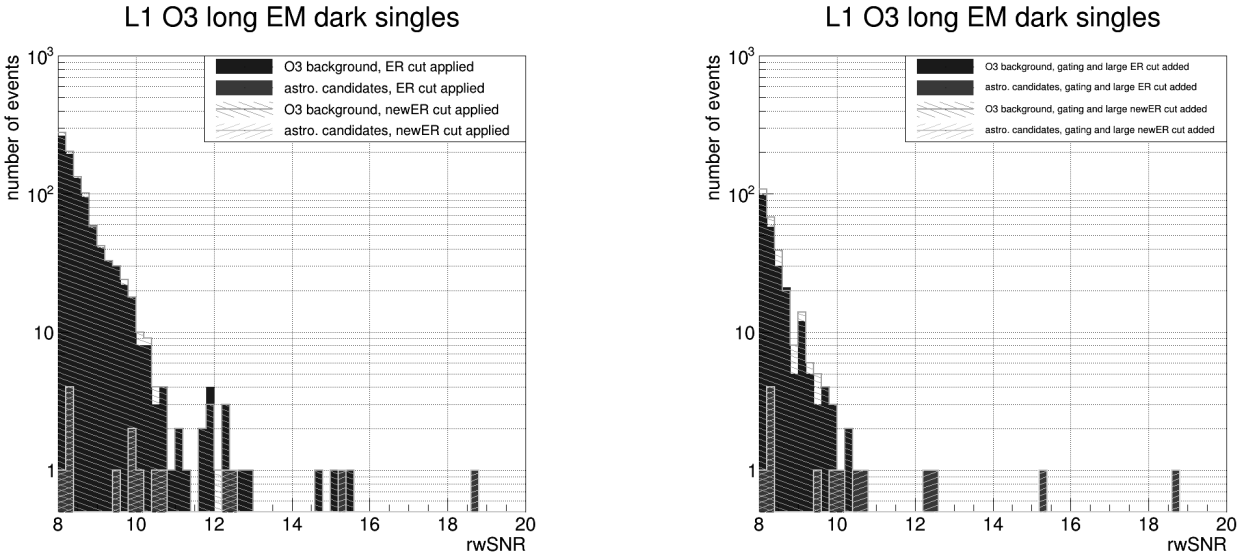


Figure 75: rwSNR distributions for L1 O3 long EM dark single detector triggers with ER selection (left) and ER+gating selection (right) using the old (max) and new ER (common region).

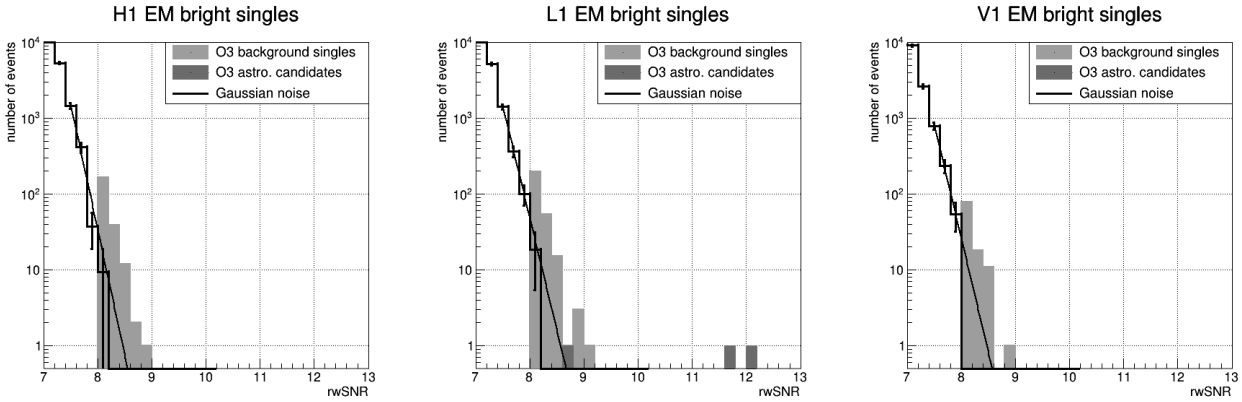


Figure 76: rwSNR distribution: O3 EM bright single detector triggers after selection (gating and wide new ER cuts) vs Gaussian noise

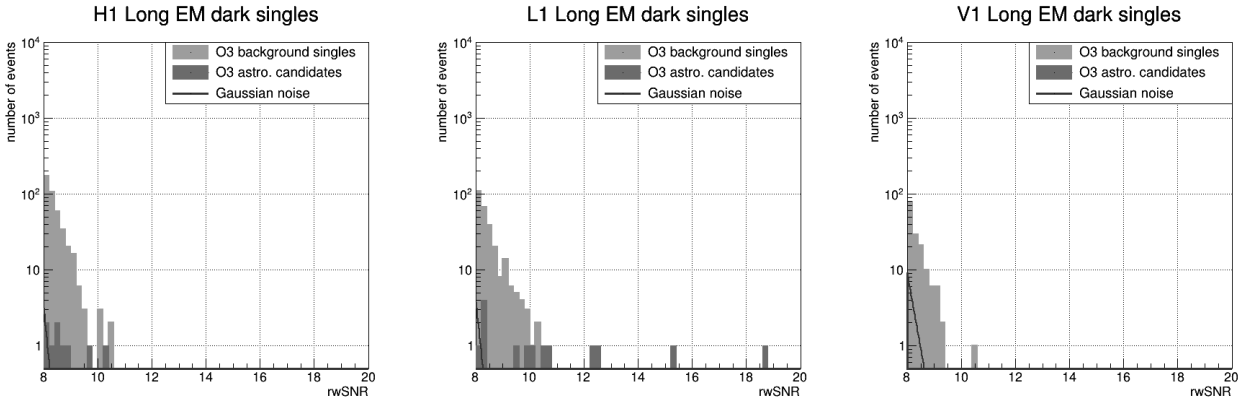


Figure 77: rwSNR distribution: O3 Long EM dark single detector triggers after selection (gating and wide new ER cuts) vs Gaussian noise.

6 Assessing the significance of single detector triggers

Contents

Now that we have cleaned our single detector triggers SNR distribution we want to assess the significance of the triggers, especially for those in the tail of the distribution since this is where we expect to find astrophysical events. The ranking statistics tells us how loud an event was. It is only relevant when comparing MBTA triggers with each other and does not tell whether a trigger is likely to be of astrophysical origin or not. This selection, for sending public alerts, is instead done using the False Alarm Rate (FAR) described in 4.11. This chapter describes several option for the single detector triggers FAR computation.

6.1 Extrapolating the observed background

To compute a FAR for the single detector triggers, a first possibility is to extrapolate the observed O3 background. In chapter 5 we have shown that we can significantly reduce the background for EM bright single detector triggers. This allows to remove the tails of the distribution. The remaining background rwSNR distribution is then close to an exponential distribution. We can therefore fit the distribution using an exponential function to reach lower rates for background events which can be used to assign lower FAR values to high rwSNR candidates. Figure 78 shows such a fit on O3 EM bright single detector triggers after aplying the selection criteria. The fitting function is $\text{FAR}_{8.6} \times \exp(\text{slope} \times (\text{rwSNR} - 8.6))$, such that "$\text{FAR}_{8.6}$" is the cumulative background rate (equal to the FAR) at rwSNR= 8.6. This choice is made in order to reduce a bit the correlation between the parameters of the fit. Two fits are performed. One on the tail of the distribution only ($8.4 < \text{rwSNR} < 9.1$) and the other on the full distribution. Results of the fits are reported in table 3. The errors on the cumulative rate were adjusted to have a χ^2/NDF close to 1 in order to have a reasonable estimate of the errors on the fit.

Detector	fit range	$\text{FAR}_{8.6}$ (days^{-1})	slope	IFAR_8 (days)	IFAR_9 (years)	IFAR_{10} (centuries)
H1	[8.0, 9.1]	0.019 ± 0.002	-6.9 ± 0.2	0.8 ± 0.03	2.3 ± 0.3	23.1 ± 7.2
L1	[8.0, 9.1]	0.022 ± 0.002	-6.8 ± 0.2	0.8 ± 0.04	1.9 ± 0.4	16.0 ± 6.4
H1	[8.4, 9.1]	0.017 ± 0.001	-7.9 ± 0.6	0.5 ± 0.2	3.8 ± 1.2	107 ± 100
L1	[8.4, 9.1]	0.023 ± 0.003	-5.2 ± 1.3	1.9 ± 1.4	1.0 ± 0.5	1.7 ± 3.1

Table 3: Fit parameters obtained on O3 EM bright single detector triggers after selection criteria.

We can also test this method on MDC data, i.e. O3 data analyzed with the O4 configuration of MBTA (updated code, new bank...). Figure 79 shows the rwSNR distributions of EM bright single detector triggers that pass the selection criteria in H1 and L1. The distributions are fitted with an exponential function from rwSNR=7.3 onwards. The two top plots are given for an effective time of 7.2 days for H1 and 7.9 days for L1. The two middle plots at the bottom are given for 7.8 days in H1 and 7.9 days in L1, separated from the previous ones by around four days. The last two plots are given for an effective time of 5.7 days in H1 and 6.0 days in L1, again separated by a

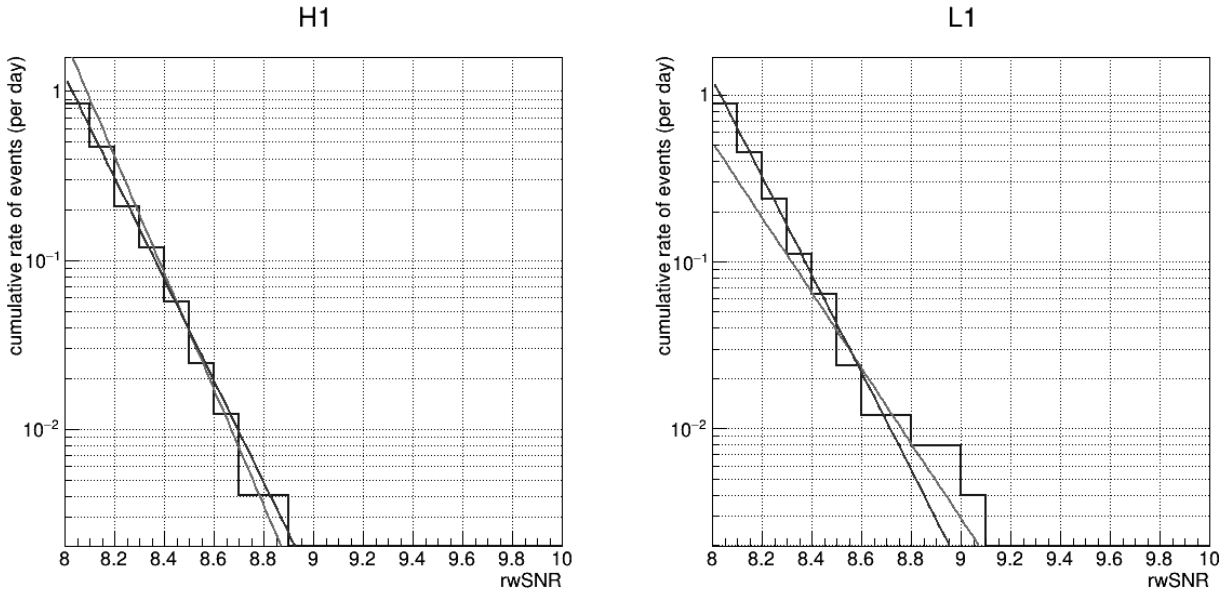


Figure 78: rwSNR distribution of O3 EM bright single detector triggers after selection with an exponential fit. The red curves are fitted on [8.4, 9.1]. The blue curves ares fitted on the full distribution to show in table 3 the fit parameters when smoothing the tail of the distribution.

few days from the previous ones. Table 4 gives the parameters of the fits performed. We see that the background is rather stable over time with fit parameters that are similar from one period of time to the other. The FAR at a rwSNR of 10 obtained using each parametrization is also given to explicit what kind of rate we can reach. The bottom line of these fits is that the value of IFAR_8 is around unity with a slope slightly smaller than -6.

Detector	data duration (days)	$\mathrm{FAR}_{8.6}$ (days^{-1})	slope	IFAR_8 (days)	IFAR_9 (years)	IFAR_{10} (centuries)
H1	5.7	0.028 ± 0.003	-6.7 ± 0.1	0.62 ± 0.03	1.4 ± 0.2	11.8 ± 2.9
H1	7.2	0.039 ± 0.005	-6.4 ± 0.1	0.57 ± 0.03	0.9 ± 0.1	5.2 ± 1.4
H1	7.8	0.021 ± 0.003	-6.8 ± 0.1	0.80 ± 0.06	2.0 ± 0.4	18.7 ± 6.0
L1	6.0	0.030 ± 0.004	-6.6 ± 0.1	0.61 ± 0.04	1.3 ± 0.2	9.9 ± 2.6
L1	7.9	0.030 ± 0.004	-6.6 ± 0.1	0.63 ± 0.04	1.3 ± 0.2	9.4 ± 2.7
L1	7.9	0.024 ± 0.003	-6.8 ± 0.1	0.69 ± 0.05	1.7 ± 0.3	15.4 ± 4.8

Table 4: Results of the fits performed on the MDC observed background for different detectors and duration. Fit function is $\mathrm{FAR}_{8.6} \times \exp(\mathrm{slope} \times (\mathrm{rwSNR} - 8.6))$.

Let's take the L1 trigger of GW190425 with rwSNR=12.07 as an example. Using the parameters of the last L1 fit in table 4 for the parametrization, one could give this trigger a FAR close to $1.4 \times 10^{-12}\,\mathrm{days}^{-1}$. In practice we would not trust the fit over so many orders of magnitude and we would therefore put an upper limit on the FAR. One way to do that would be to choose the highest $\mathrm{FAR}_{8.6}$ and mildest slope (in table 4 for L1: 0.039 and -6.4 respectively). In this case the FAR would be $8.8 \times 10^{-12}\,\mathrm{days}^{-1}$.

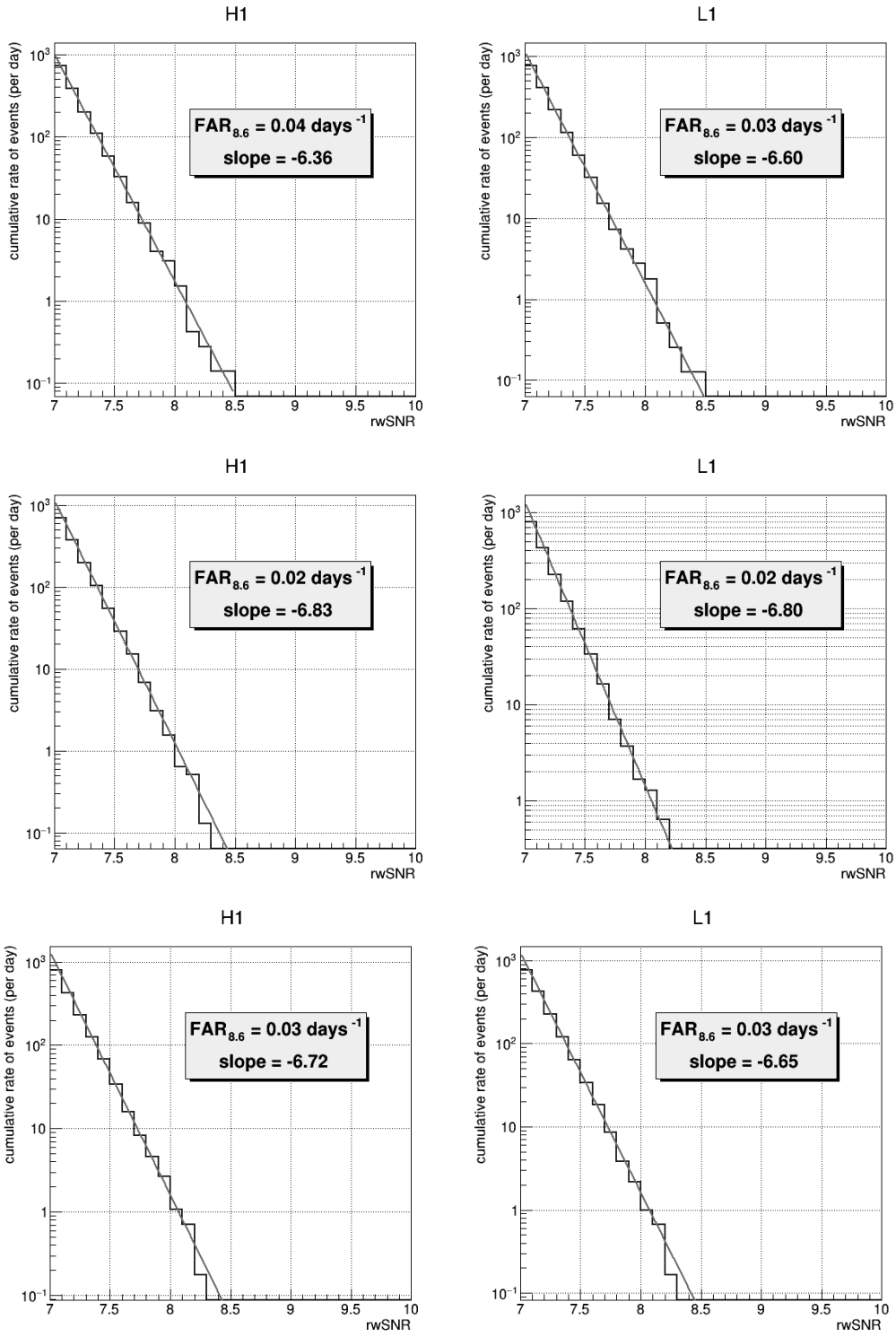


Figure 79: rwSNR distribution for EM bright single detector triggers which pass the selection criteria on O3 data analyzed using MBTA O4 configuration. The duty cycles from top to bottom are 7.2, 7.8, 5.7 days for H1 and 7.9, 7.9, 6.0 days for L1. Fit with $\mathrm{FAR}_{8.6} \times \exp(\mathrm{slope} \times (\mathrm{rwSNR} - 8.6))$.

6.2 Combining individual band triggers with random noise

Fitting the observed background works well when the background does not change much from one day to another. If, however, the background distribution that we fit has some tail because there was a lot of noise some day, the parametrization will certainly not be reliable. Furthermore, we are limited to a few days of observation to perform the parametrization if we want for instance to update the background distribution every week or so. We would therefore like to investigate a method that would estimate the background for FAR below one per few days..

In the case of coincidences, a background distribution is generated by running a coincidence search with time-shifted detector data to ensure any coincidence found is fortuitous. The FAR corresponding to a given cRS_0 is then computed as the rate of background triggers (obtained via these fake coincidences) with cRS larger than cRS_0. We obviously cannot generate a background distribution using time shifted data of several detectors in the case of single detector triggers. But MBTA allows us to perform a trick to work in a similar way. The matched-filtering is performed in two separate frequency bands. In the case of an astrophysical signal we expect the data in the two bands to be correlated at adjacent times (figure 80). We can therefore make an analogy between our two frequency bands and two detectors of the interferometer network. Taking this analogy as the starting point, we can start working in a similar way as for the coincidence search. We follow the logic of the single detector trigger constructions which starts by detecting a trigger in one band and combining its MFOs with the available signal in the other band to see if it passes the selection threshold.

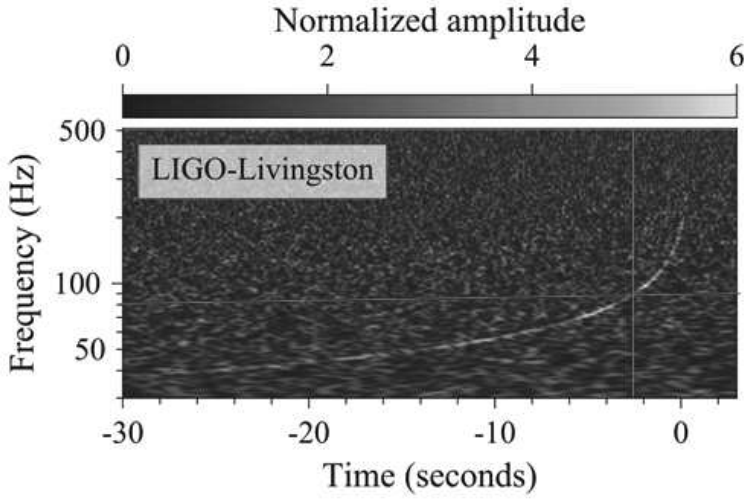


Figure 80: GW170817 in the time-frequency plane, extracted and adapted from [65]. The horizontal line shows the limit between the 2 frequency bands defined in MBTA, the vertical line shows the separation in time of the low frequency and high frequency components of the signal.

Therefore, to emulate this process we need to store single band triggers and single band random noise (random MFOs). This option is available in the MBTA version prepared for O4 and works as follow:

- Individual band triggers of each band are saved with an SNR threshold of 5. Since there is a very high number of low SNR triggers, they are down-sampled to reduce the computational cost and disk space usage. The down-sampling is done between SNR=5 and SNR=9. A

weight is computed as $w_{ds} = 10^{0.3\times(9-\mathrm{SNR})}$. A random number r is then generated following a uniform distribution, if $r \times w_{ds} \leq 1$ the trigger is saved.

- MBTA periodically saves random noise in each band for every Real Template (RT, individual band templates), in general every 2000 s+ϵ with $\epsilon = 0.11$ s. This choice is motivated by the fact that FFTs are computed with an integer interval and we could have some edge effects at the time the FFT is computed (such as excess of triggers) although it is unlikely to occur. by using an ϵ we can select our random noise at different time offsets with respect to the time of the computation of the FFT.

We show in figure 81 the SNR distribution of the individual band triggers and random noises over two different periods of time (for EM bright templates and times that pass the single detector triggers selection criteria). We see that we have high rwSNR triggers in the HF band but they constitute at most $\sim 0.1\%$ of the total number of triggers. Also they will not produce loud noise triggers as we will see, because the autoχ^2 and rwSNR step will downgrade them.

We then proceed to make the fake coincidences:

- The MFOs of the triggers of one band are combined with all random noises of the other band that are associated with a compatible RT. The combination is done following

$$\begin{cases} \mathrm{MFO}_{\mathrm{combi,P}}(t) & = \mathrm{MFO}_{\mathrm{LF,P}}(t_{LF}) + \cos(\phi)\mathrm{MFO}_{\mathrm{HF,P}}(t_{HF}) - \sin(\phi)\mathrm{MFO}_{\mathrm{HF,Q}}(t_{HF}) \\ \mathrm{MFO}_{\mathrm{combi,Q}}(t) & = \mathrm{MFO}_{\mathrm{LF,Q}}(t_{LF}) + \sin(\phi)\mathrm{MFO}_{\mathrm{HF,P}}(t_{HF}) + \cos(\phi)\mathrm{MFO}_{\mathrm{HF,Q}}(t_{HF}) \end{cases} \tag{64}$$

  which is the same as for the band MFO combination in MBTA (eq. 36) with the LF part at time t_{LF} and the high frequency part at t_{HF}. The MFO of the low frequency part of the combination (either the trigger or the random noise) first has to be interpolated to account for the difference in sampling rate between the two frequency bands. Since we have no information about the phase ϕ, the combination is done for two orthogonal values: $\phi = 0$ and $\phi = \pi/2$.

- The autoχ^2 and rwSNR are computed.

- Repeat with the triggers of the second band and random noise of the first one. Note that a small fraction of noise events may come from a trigger in each frequency band. In the way we proceed here, this event will be counted as two different events because we have two triggers. This means that we have a slight over-estimation and we are therefore being conservative.

- The effective time for the computed background is

$$2 \times \frac{(\text{observed effective time})^2}{\text{random noise saving period}} \tag{65}$$

  The factor 2 stands for the two phase values used in the combination process.

An example of combination of a high frequency trigger with a low frequency random noise is shown in figures 82 and 83. The rwSNR distribution of the computed background is shown in figure 84 using 3.6 days of data in H1 and 6.4 days of data in L1. The effective time for the computed background is 3.1 years in H1 and 9.7 years in L1. It is compared to the background observed during the time for which we saved the triggers and random noises, Gaussian noise and O3 background. We see that we can compute a distribution which follows the observed background, especially at

Detector	data duration (days)	$FAR_{8.6}$ ($days^{-1}$)	slope	$IFAR_8$ (days)	$IFAR_9$ (years)	$IFAR_{10}$ (centuries)
H1	3.6	0.019 ± 0.001	-8.1 ± 0.1	0.34 ± 0.01	3.7 ± 0.4	125 ± 23
L1	6.4	0.042 ± 0.002	-7.40 ± 0.04	0.279 ± 0.005	1.26 ± 0.07	20 ± 2

Table 5: Parameters of the fits using trigger-random noise coincidences

low rwSNR (around 7), without any arbitrairy scaling factor. However, for larger rwSNR, the computed background is usually below the observed one.

Figure 85 shows the cumulative rate for the computed and observed background. The computed background is fitted using $FAR_{8.6} \times \exp(\text{slope} \times (\text{rwSNR} - 8.6))$. Fit results are given in table 5. The slope in this case is steeper than what we had previously for the observed background in table 4.

Figure 86 shows the rwSNR distribution of the computed background split in two: pseudo-events made with a HF trigger and pseudo-events made with a LF trigger. We see that in H1 the HF band had a greater contribution to the final background but in the case of L1 the contribution of each band is almost equal.

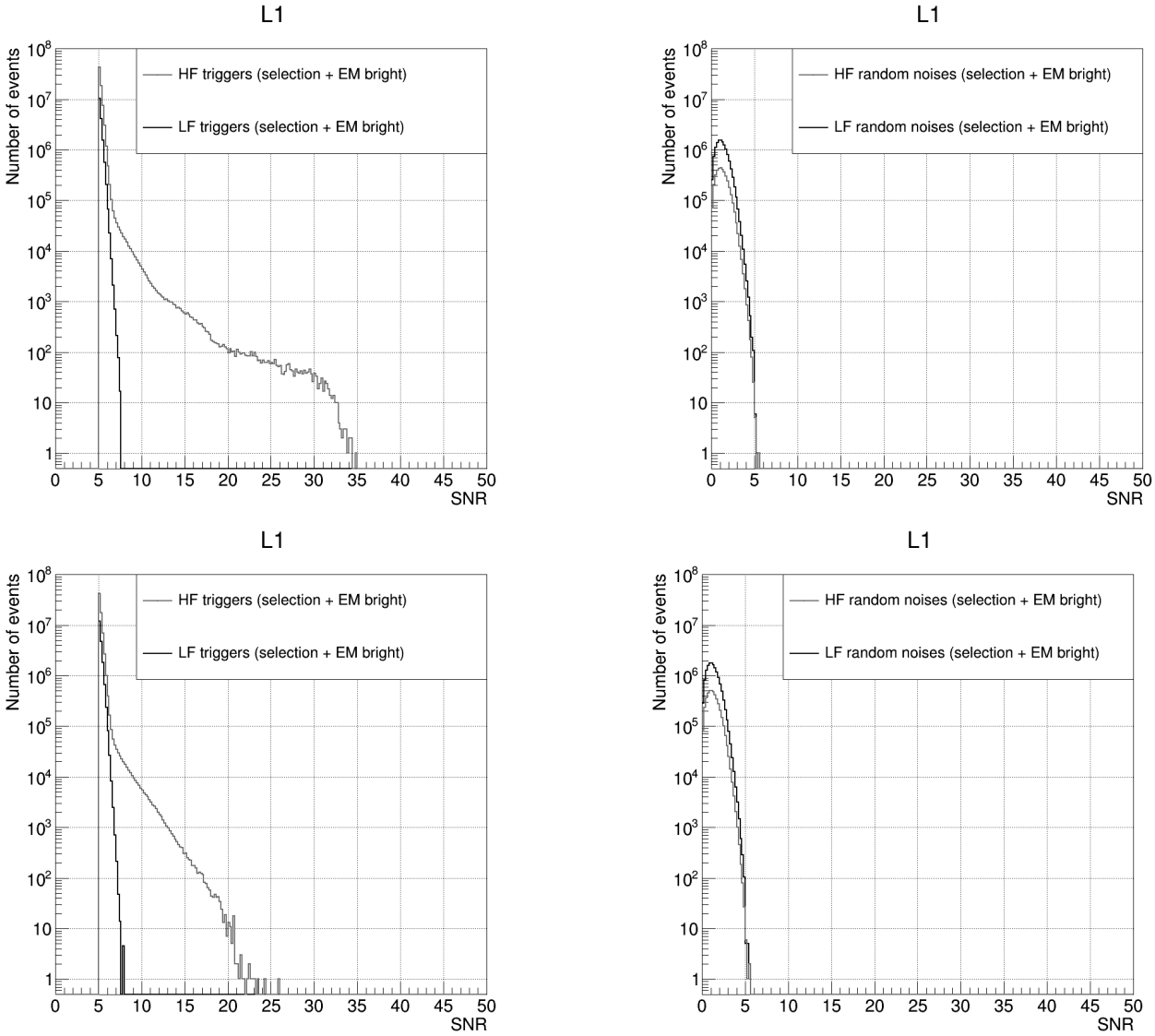


Figure 81: Example of triggers (left) and random noise (right) SNR distribution over effective ~ 6.4 days (top, from February 12 2023 to February 24 2023) and ~ 6.34 days (bottom, from February 28 2023 to March 10 2023) in L1, only for EM bright with single selection criteria applied. The downsampling weights of the triggers are taken into account. The 6.4 days are the same ones as those used for the computation of the background in L1 presented in this chapter.

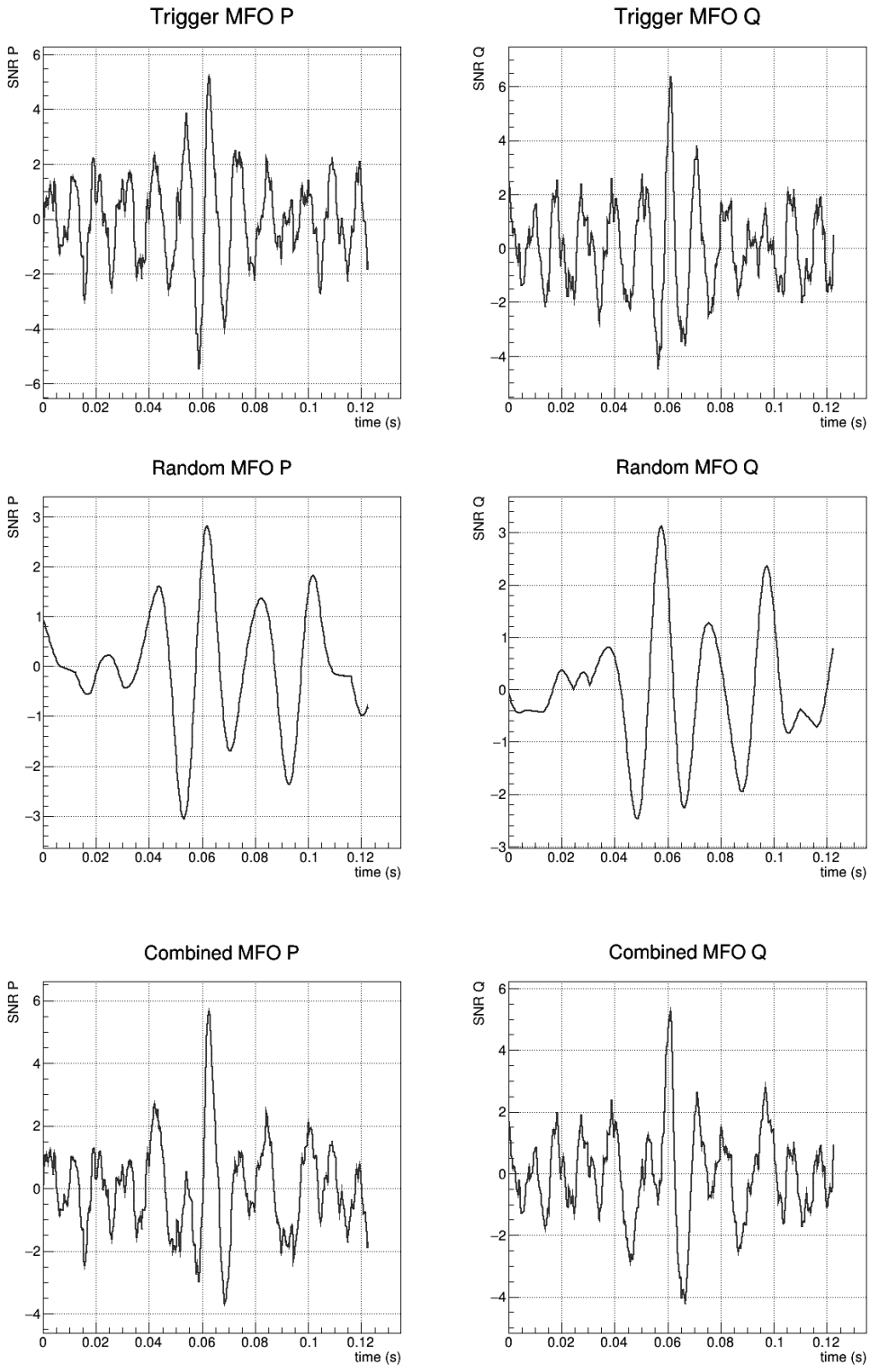


Figure 82: Top: P and Q MFOs for a HF trigger in L1. Middle: P and Q MFOs for a LF random noise in L1. Bottom: In-phase and in-quadrature SNR time series for the event resulting from the combination of the two. Mind the differences in Y-axis range.

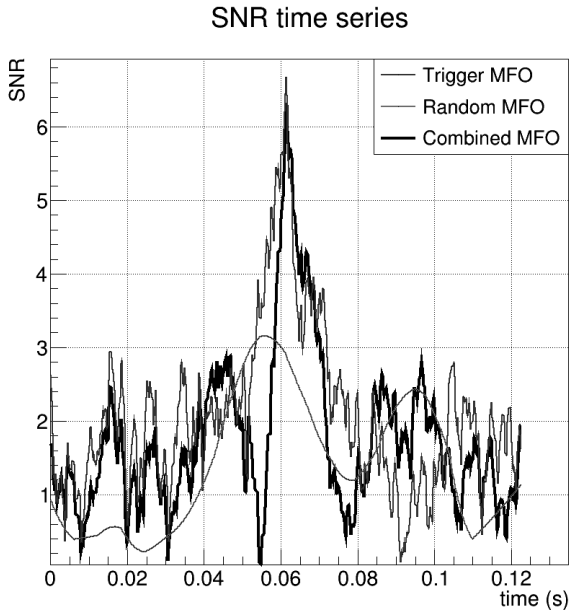


Figure 83: SNR time series for the trigger, random noise and pseudo-event resulting from the combination of the trigger and random MFO show in figure 82

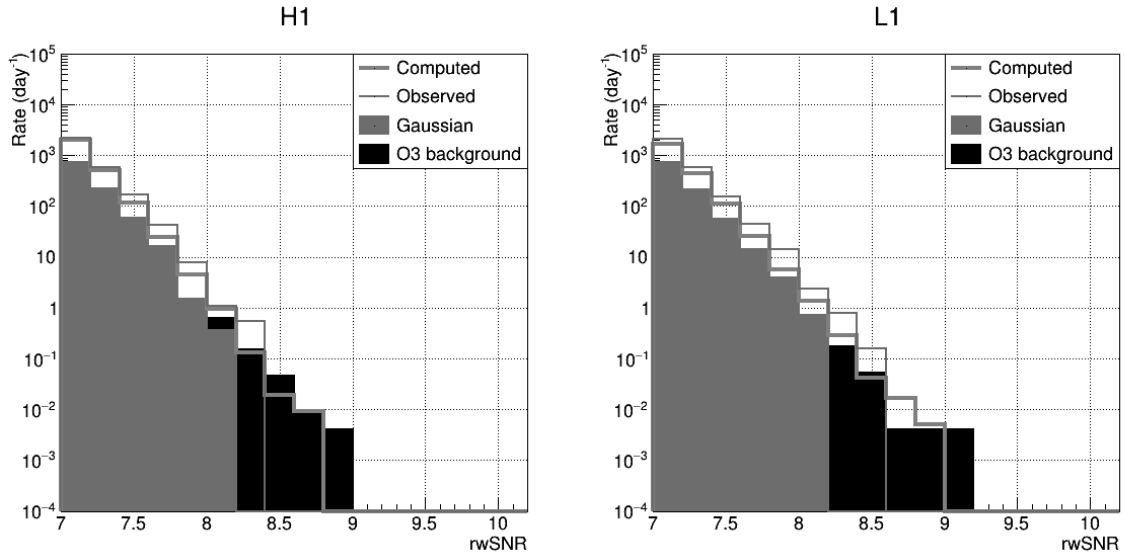


Figure 84: Comparison of the non-cumulative rate of computed background (blue, trigger-random noise coincidences) with the background observed during the time used for the computation (red), simulated gaussian noise (green, for around one month of data) and all-O3 background (black). Left is for 3.6 days of data in H1. Right is for 6.4 days of data in L1. The "observed background" is O3 data analyzed with MBTA O4 configuration as opposed to O3 background which was analyzed with the O3 configuration. We only consider triggers from the EM bright region which passed the single selection.

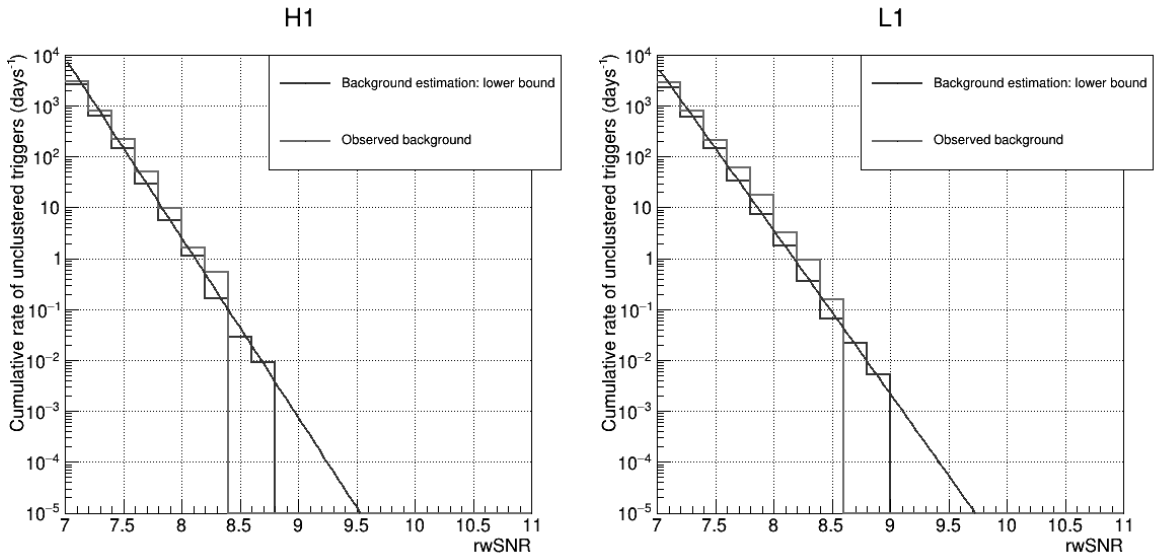


Figure 85: Cumulative rate for the computed background (blue) and observed background (red) shown in figure 84. Left is for 3.6 days of data in H1. Right is for 6.4 days of data in L1.

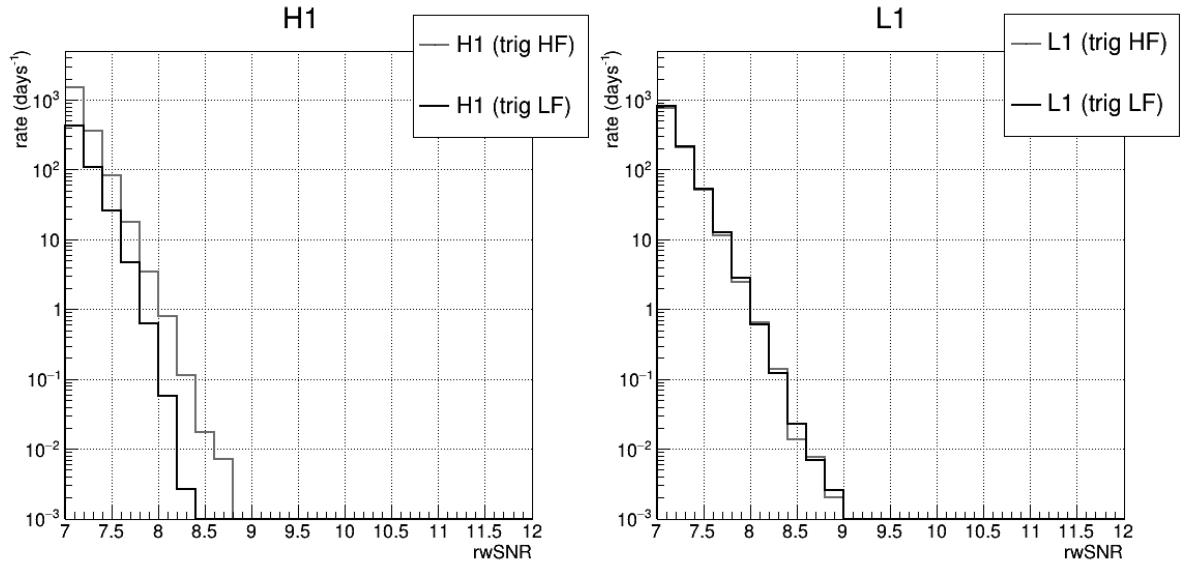


Figure 86: Computed background rwSNR distribution for 3.6 days in H1 (left) and 6.4 days in L1 (right) using trigger-random noise coincidences separated based on the frequency band of the trigger. Red is for pseudo-events made with a HF trigger. Black is for those made with a LF trigger.

6.3 Combining individual bands triggers with other triggers

We have until now only considered triggers-random noises associations. This means that we assumed our frequency bands to be completely uncorrelated. But what if, for example, a glitch occurs at the time the signal changes frequency bands, and therefore could pollute both MBTA's frequency bands? Then our single detector trigger will likely be formed from two single band triggers, implying some correlation between our bands. We will therefore also consider trigger-trigger concidences when esimating our background. This way we can have anything between fully correlated frequency bands (100% trigger-trigger coincidences) and completely uncorrelated bands (100% trigger-random noise coincidences).

We will proceed in the same way as for the coincidences with random noise but we have to be careful about two things:

- triggers were not saved periodically as random noises were, meaning that we do not have a fix number of them for each template. In order to easily scale the computed distribution to a rate comparable to the observed background we decide to do a fixed number of coincidences for each template. To account for the down-sampling we have to use the down-sampling weight of the second trigger (the one which takes the place of the random noise in the previous section) when counting the number of coincidences we make.

- triggers were down-sampled unlike random noise, we have to take this into account when building the distributions (rwSNR distribution for instance) for the pseudo-events. Thus, a pseudo-event made from a trigger-trigger combination will have count equal to the product of the downsampling weights of the two triggers.

For example combining a trigger from the first band with two triggers from the other band, one with a weight of 3.4 and the other of 1, will count as a total of 4.4 coincidences. Then when building distributions we will count the coincidences as the product of the down-sampling weight of the first and second trigger. So in the previous case if the first trigger had a weight of, say 2.1, our first fake single detector trigger would count as $2.1 \times 3.4 = 7.14$ and the second one as $2.1 \times 1 = 2.1$. Regarding the number of coincidences that we want to make, as explained in the first point, it is quite arbitrary although constrained by the quantity of data available or that we want to analyze: the more coincidences the better (computing time aside) but we have to have enough triggers for that. We want to at least reach the threshold for public alerts. This means that the computed background should be given for an effective time of at least one year. Since we typically consider close to a week of data to make the fake coincidences, this can be achieved by doing around 100 coincidences per trigger. But our goal is also to give high rwSNR trigger a FAR that reflects properly their significance. Using 1000 coincidences allows to reach FARs of around 1 per 33 years when using 6 days of effective data before even extrapolating. Thus we choose to do 1000 coincidences per triggers.

If we have enough triggers in each band to make 1000 coincidences per trigger, then we can use any of the two to pick primary triggers and the other to pick secondary triggers (playing the role of the random noise in the previous section). In case not enough coincidences could be made for a given real template, we decide to use the adjacent ones (RT index-1 and RT index+1) to compensate. This has also prompted us to use LF triggers as secondary triggers and the HF triggers as primary triggers, since LF RTs should be very similar for adjacent templates.

To summarize:

- Consider one HF trigger with down-sampling weight ω_{HF}.

- The real template of this HF trigger is compatible with one or several LF RTs.
- Let's take the first LF RT which has index k and all associated triggers.
- Let N be the number of such LF triggers, each of them has a down-sampling weight ω_i with $i \in [|1; N|]$.
- We want to make 1000 fake coincidences, this means that if $N > 1000$ we will in fact only use the first $n < N$ LF triggers such that

$$\sum_{i=1}^{n-1} \omega_i < 1000 \leq \sum_{i=1}^{n} \omega_i \tag{66}$$

- If $N < 1000$ the relation still holds, simply some of the triggers will have index $k-1$ and/or $k+1$ (although we cannot exclude that for a very small number of templates it may not be possible to gather enough triggers).
- When counting coincidences, each of those coincidences will count as $\omega_{\text{HF}} \times \omega_i$.
- Since we have effectively done (close to) 1000 coincidences, we can scale the computed distribution with

$$T_{\text{eff}} = T_{\text{eff,ini}} \times 2 \times 1000 \tag{67}$$

  where $T_{\text{eff,ini}}$ is the initial effective time of the data we use to make the fake coincidences and the factor 2 comes from the two values we use for the phase like with the random noises.

Proceeding in this way leads us to a very large over-estimation of the background since it assumes that all triggers in one band are connected to a trigger in the other band. In order to have a background estimation that can act as an upper bound we choose to scale the distribution computed with trigger-trigger coincidences to force an equal rate with the observed background for a rwSNR of 7.5, the scaling factor is in this case around 10^{-4}. We chose to scale at rwSNR=7.5 as it is a bin with many pseudo-events, meaning reduced errors and also because at lower rwSNR we observe a threshold effect due to the selection of triggers starting at SNR=5. While completely ad hoc this allows us to have an estimation of the background rate close to the observed one and a less steep slope. Results are shown in figure 87. Note that for H1, due to the smaller observing time used for the computation, only 100 coincidences per trigger were done. This upper bound estimation is safer in terms of FAR computation because it will lead us to over-estimate the FAR (whereas the lower bound would lead to an under-estimation) and therefore we are less likely to claim false detections since the lower the FAR the more significant the event. Like for trigger-random noise coincidences, the computed background using trigger-trigger combinations follows an exponential shape. We can thus go further and extrapolate the computed distribution (actually the cumulative distribution of the computed background) with an exponential function to reach even lower values of FAR as shown in figure 88. Results of the fit for several background estimations are given in table 6. These background parametrization can be done in terms of cRS^2 in order to be used by the p_{astro} code to compute in the end a $\text{FAR}(\text{p}_{\text{astro}})$.

In all this chapter we have computed this background using different methods. It should be higlighted that in all cases the background distribution could be fitted with an exponential function with the same slope over the full rwSNR range. This gives us confidence in the background extrapolation we can do to reach lower FAR.

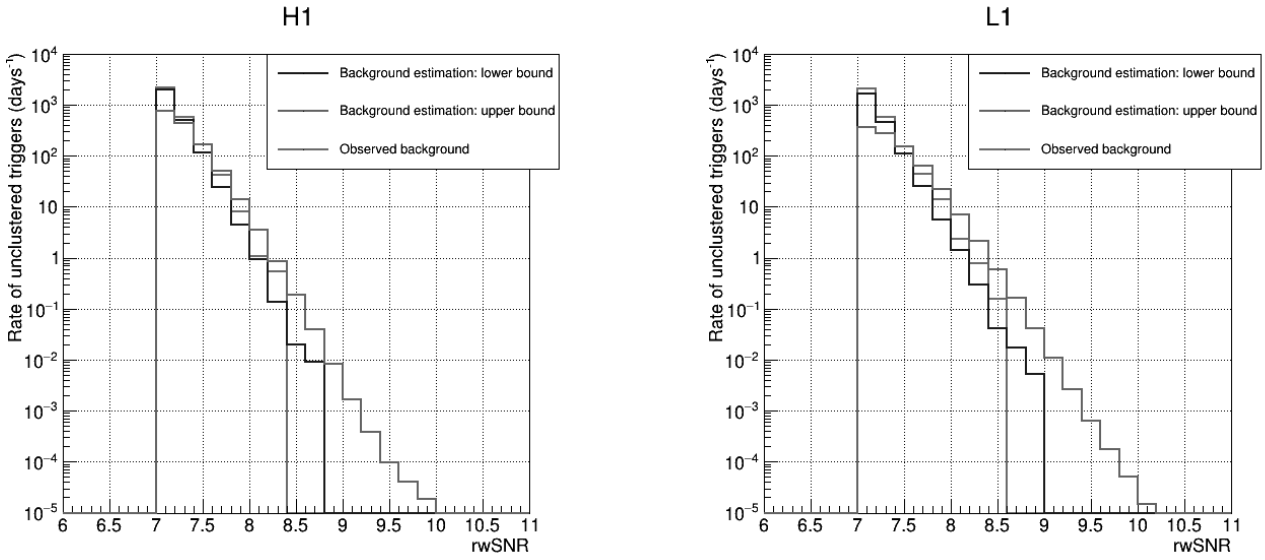


Figure 87: Comparison of the non-cumulative rate for the computed lower bound, upper bound and observed background for H1 (left) and L1 (right). Computation was done over ~ 3.6 and ~ 6.4 days for H1 and L1 respectively.

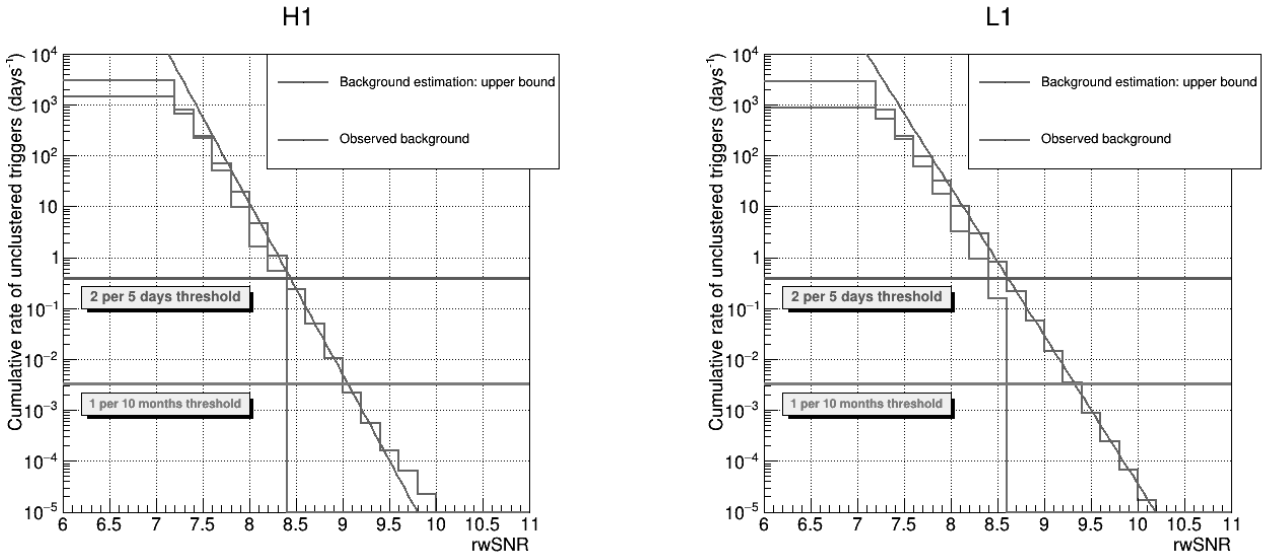


Figure 88: Comparison of the cumulative distributions of the observed background and computed upper bound for the background in H1 (left, for ~ 3.6 days of data) and L1 (right, for ~ 6.4 days of data) fitted with an exponential function. Fit results are given in table 6. Horizontal lines show the FAR thresholds used to send low-significance (2 per 5 days) and significant (1 per 10 months) public alerts at the begining of O4.

Detector	data duration (days)	$FAR_{8.6}$ ($days^{-1}$)	slope	$IFAR_8$ (days)	$IFAR_9$ (years)	$IFAR_{10}$ (centuries)
H1	3.6	0.109 ± 0.001	-7.8 ± 0.1	0.087 ± 0.003	0.56 ± 0.02	13.1 ± 1.2
H1	3.7	0.132 ± 0.002	-7.8 ± 0.1	0.069 ± 0.004	0.48 ± 0.02	11.9 ± 1.8
L1	6.4	0.418 ± 0.003	-6.70 ± 0.04	0.043 ± 0.001	0.096 ± 0.002	0.8 ± 0.1
L1	3.5	0.243 ± 0.001	-7.24 ± 0.04	0.053 ± 0.001	0.205 ± 0.004	2.9 ± 0.2
L1	3.6	0.122 ± 0.002	-8.1 ± 0.1	0.065 ± 0.005	0.56 ± 0.03	17.8 ± 3.2
L1	3.8	0.149 ± 0.003	-8.0 ± 0.1	0.055 ± 0.004	0.45 ± 0.03	13.6 ± 2.4

Table 6: Results of the fits performed on the computed over-estimation of the background for different detectors and times. Fit function is $FAR_{8.6} \times \exp(\text{slope}(\text{rwSNR} - 8.6))$.

7 Improving the rwSNR

Contents

We have shown in section 5 that we can reject background triggers using selection criteria based on the excess rate and gating. In the same section we also considered a cut on the autoχ^2 to reduce the background. This cut turned out to unsafe. This is explained by the fact that for injections, and therefore astrophysical signals, the autoχ^2 grows with the SNR. Therefore, removing high autoχ^2 events at high SNR will remove both noise and astrophysical triggers.

In this chapter we will study this dependency of the autoχ^2. The goal would be to find a way to improve the autoχ^2 or the rwSNR such that very loud astrophysical signal and injections are not downgraded by them. This study is not conclusive, therefore the purpose of this section is to document it in order to highlight the challenges and what might be done in the future.

The rwSNR (eq. 43) was introduced to downrank background candidates relative to astrophysical events. It is based on the computation of the autoχ^2 (eq. 40) which computes, upon detection of a candidate, the excess of power due to mismatch between the measured signal and the theoretical one, expected from the template that triggered. The rwSNR proved to be very useful in background removal. As already reminded, there is however a downside to the autoχ^2 : its value grows as the SNR grows. This is especially visible when looking at simulated events (injections). This is shown on figure 89 where we can see that almost no BBH injection with SNR greater than 30 keeps more than 80% of its SNR after reweighting. It also shows the rwSNR vs SNR distribution for those injections. We clearly see that many loud injections end up with a much lower rwSNR. This means that some loud injections and astrophysical events can be missed because their SNR is too heavily reweighted.

The origin of of higher autoχ^2 at high SNR is the discreet nature of the template bank. It results in a small mismatch between the source and template parameters and thus in waveforms. These small differences become significant for high SNR events.

In this chapter we only study L1 as it is the detector with the highest sensitivity and therefore the highest SNRs. It makes a good case of the issues at hand.

7.1 SNR dependency of the autoχ^2

We want to parametrize the evolution of the autoχ^2 as a function of the SNR. This would allow to take it into account in our reweighting to try to improve the rwSNR. As mentionned earlier, it is caused by the discretization of the template bank. The signal that we measure is in fact the theoretical signal to which we have to add the background and also the mismatch due to our discretized template bank. The theoretical MFO is

$$\mathrm{MFO}_{P/Q,exp}[i] = \mathrm{SNR} \times \mathrm{MFO}_{P/Q,exp,\mathrm{normalized}}[i] \tag{68}$$

The mismatch between the template and measured signal is therefore going to be proportional to the SNR. If we go back to the expression of the autoχ^2 (eq. 40) and we write

$$\mathrm{MFO}_{P/Q,meas}[i] = \mathrm{MFO}_{P/Q,exp}[i] + \epsilon\mathrm{SNR} + \mathrm{bkg} \tag{69}$$

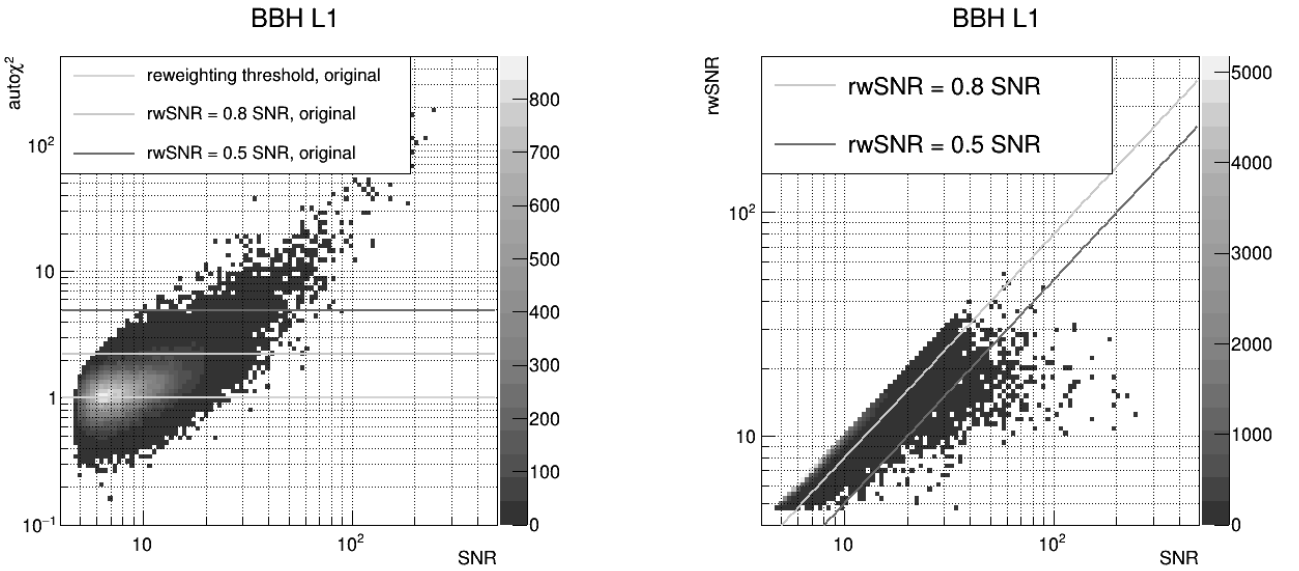


Figure 89: O3 BBH recovered injections: autoχ^2 vs SNR (left) and rwSNR vs SNR (right) for single detector triggers associated to coincidences. Various rwSNR threshold are plotted using equation 43.

where ϵSNR is a term relative to the mismatch with the template proportional to the SNR and bkg is a background term of order 1, then we can write the autoχ^2 as follows:

$$\text{auto}\chi^2 = \frac{1}{2\text{N}} \sum_{i=0}^{\text{N-1}} \left[2(\epsilon\text{SNR} + bkg)^2\right] = \epsilon^2\text{SNR}^2 + 2\epsilon\text{SNR}\times\text{bkg} + \text{bkg}^2 \tag{70}$$

$$\text{auto}\chi^2 \xrightarrow[\text{large SNR}]{} \epsilon^2\text{SNR}^2 \tag{71}$$

Hence, we expect the autoχ^2 to grow with the SNR^2 at large SNRs. We can now use the formula 70 to parameterize the autoχ^2 as a function of the SNR. Figure 90 shows the profile histograms of the autoχ^2 vs SNR graphs (like figure 89 for BBH) for the BBH, BNS and NSBH injections. The profiles are fitted with

$$\text{auto}\chi^2 = (p_0 + p_1\text{SNR})^2 \tag{72}$$

We retrieve a value of order 1 for $p_0 = C$ as expected for the background term, and close to 3% for $p_1 = \epsilon$ which is of the order of magnitude for the minimal match of the template bank.

7.2 Mitigating the SNR dependence in the reweighting

We would like to mitigate the SNR dependence of the reweighting induced by the SNR dependence of the autoχ^2. We tried to modify the reweighting formula 43 as follows:

$$\text{rwSNR} = \begin{cases} \text{SNR} & , \text{ if auto}\chi^2 \ \leq (C + \epsilon\ \text{SNR})^2\,. \\ \text{SNR}\left(\dfrac{A + \left[\dfrac{\text{auto}\chi^2}{(C+\epsilon\ \text{SNR})^2}\right]^{\alpha}}{A+1}\right)^{-\frac{1}{\beta}} & , \text{ otherwise.} \end{cases} \tag{73}$$

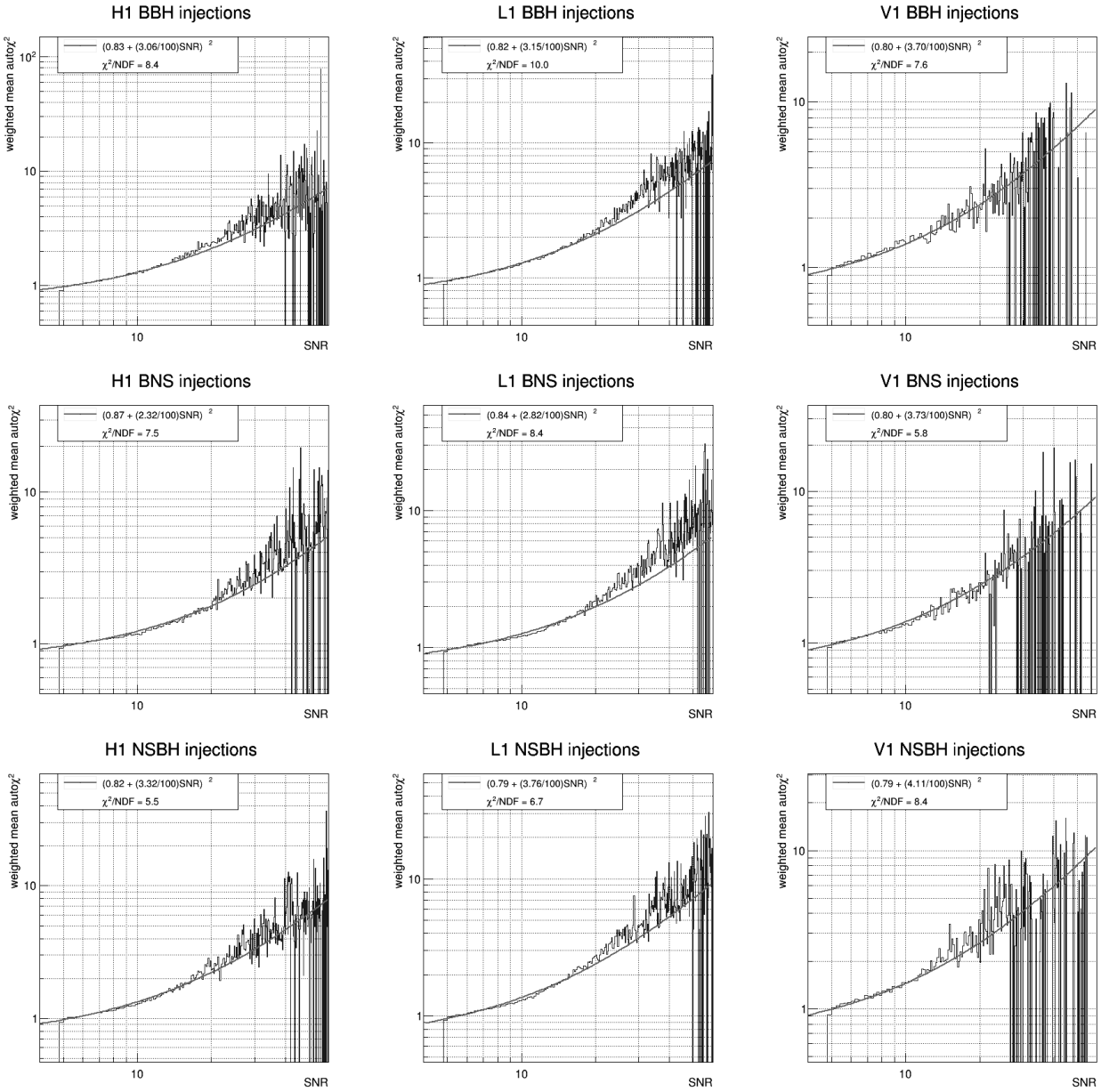


Figure 90: Injections autoχ^2 vs SNR profile histograms. From top to bottom: BBH injections, BNS injections and NSBH injections.

The idea here is to start reweighting at higher autoχ^2 for higher SNRs, the modified numerator ensures the continuity of the weight applied at 1. This is actually identical to defining a new autoχ^2 such that

$$\text{auto}\chi^{2\prime} = \frac{\text{auto}\chi^2}{(C + \epsilon\ \text{SNR})^2} \tag{74}$$

and we would have the following reweighting formula:

$$\text{rwSNR} = \begin{cases} \text{SNR}, & \text{if auto}\chi^{2\ \prime} \leq 1. \\ \text{SNR} \times \left(\frac{A + \left[\text{auto}\chi^{2\prime}\right]^{\alpha}}{A+1} \right)^{-\frac{1}{\beta}}, & \text{otherwise.} \end{cases} \tag{75}$$

which is identical to equation 40 when switching auto$\chi^2 \leftrightarrow$ auto$\chi^{2\prime}$.

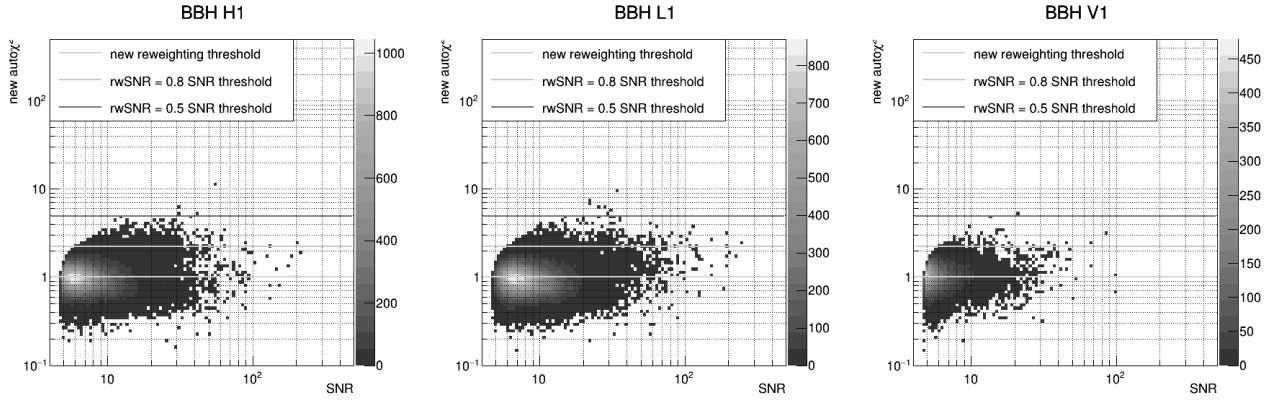


Figure 91: O3 BBH recovered injections: auto$\chi^{2\prime}$ (or new autoχ^2) vs SNR for single detector triggers associated to coincidences.

As the parameter values of the fits presented in the different plots in figure 90 seem to be uncorrelated with the detector and the type of source (as they should be), we decide to take the mean value over the detectors and sources: we therefore settle for $C = 0.82$ and $\epsilon = 0.033$. Figure 91 shows the distribution of O3 BBH injections in the auto$\chi^{2\prime}$ versus SNR plane. The auto$\chi^{2\prime}$ works as intended with a much flatter distribution in SNR. Figure 92 shows the new thresholds for rwSNR = 100%/80%SNR using the formula with the auto$\chi^{2\prime}$ (eq. 75). We were able to make an SNR-dependent reweighting that is less harsh at high SNR but we are now barely reweighting the background: almost all triggers keep more than 80% of their SNR.

The next step of the study was to explore different values of the parameters with the auto$\chi^{2\prime}$. The figure of merit of this analysis is the number of recovered injections at high SNR versus the number of background triggers. But we also have to keep in mind that the efficiency of the search is computed using the global number of recovered injections, meaning that we do not want to lose (too many) injections at low SNR. In an attempt to tune further the formula we explore higher values of the parameter α to give more weight to the term containing the auto$\chi^{2\prime}$. After several attempts we choose a value of $\alpha = 12$ instead of 5. Figure 93 shows the same thresholds as figure 92 but with $\alpha = 12$ for injections and background. Detection thresholds are also plotted.

We show also in figures 94 and 95, for injections and background respectively for the full O3 run, the rwSNR distribution obtained using the formula 73 with $C = 0.82$, $\epsilon = 0.033$ and $\alpha = 12$

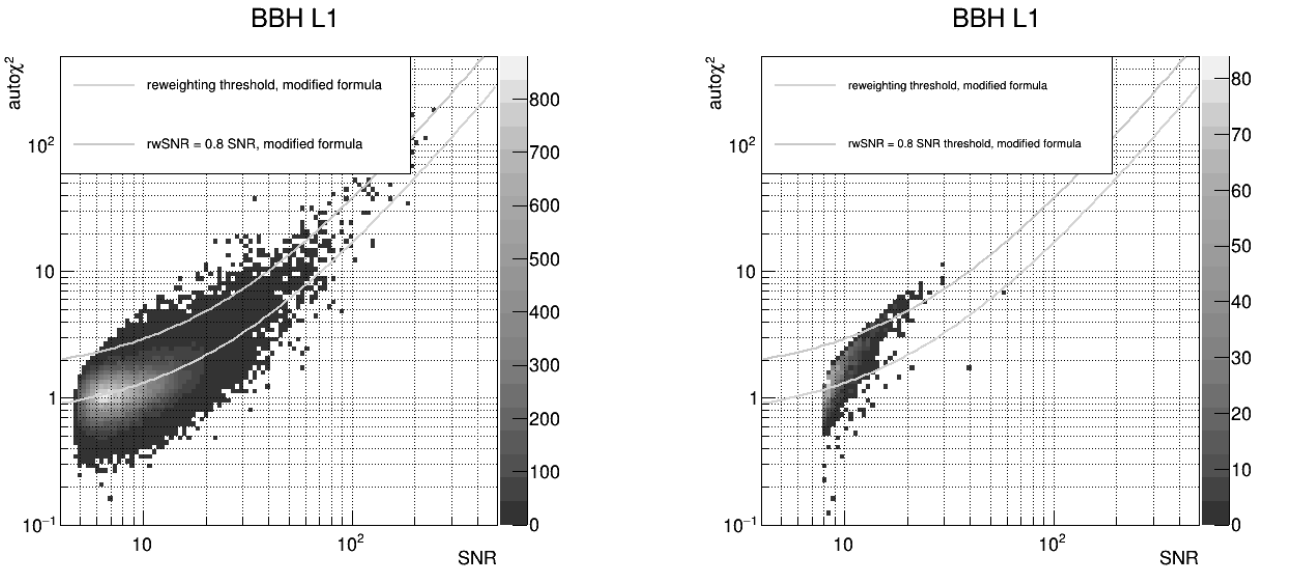


Figure 92: autoχ^2 vs SNR. Left: single detector triggers associated to coincidences for O3 BBH recovered injections. Right: O3 BBH background single detector triggers. The reweighting threshold and rwSNR = 80% SNR obtained using the modified formula 73 with $\alpha = 5$ are also plotted.

(mind the rwSNR threshold value that is different for background and injections in these plots). As intended we now recover more injections at high SNR and the background remains globally unchanged but on the downside we lose some injections at low rwSNR and we actually lose more of them than we gain at high rwSNR, meaning that the efficiency of the search is actually lowered. Other modifications of the reweighting formula were considered to try to improve the rwSNR. They did not allow for any significant improvement of the autoχ^2 but they were documented with their pros and cons [93].

In figure 95 in L1, we note the presence of 3 triggers with a modified rwSNR larger than 30. All 3 of those are caused either by scattered light in the detector or glitches and are associated to the exact same short (template duration=0.62), very asymmetric (mass1=193.011, mass2=2.02908) template. A scan centered on the time of one of them is shown in figure 96. It is therefore safe to consider that the presence of these 3 events at high modified rwSNR does not contradict the work presented here. This issue of "noisy templates" was investigated in section 5.

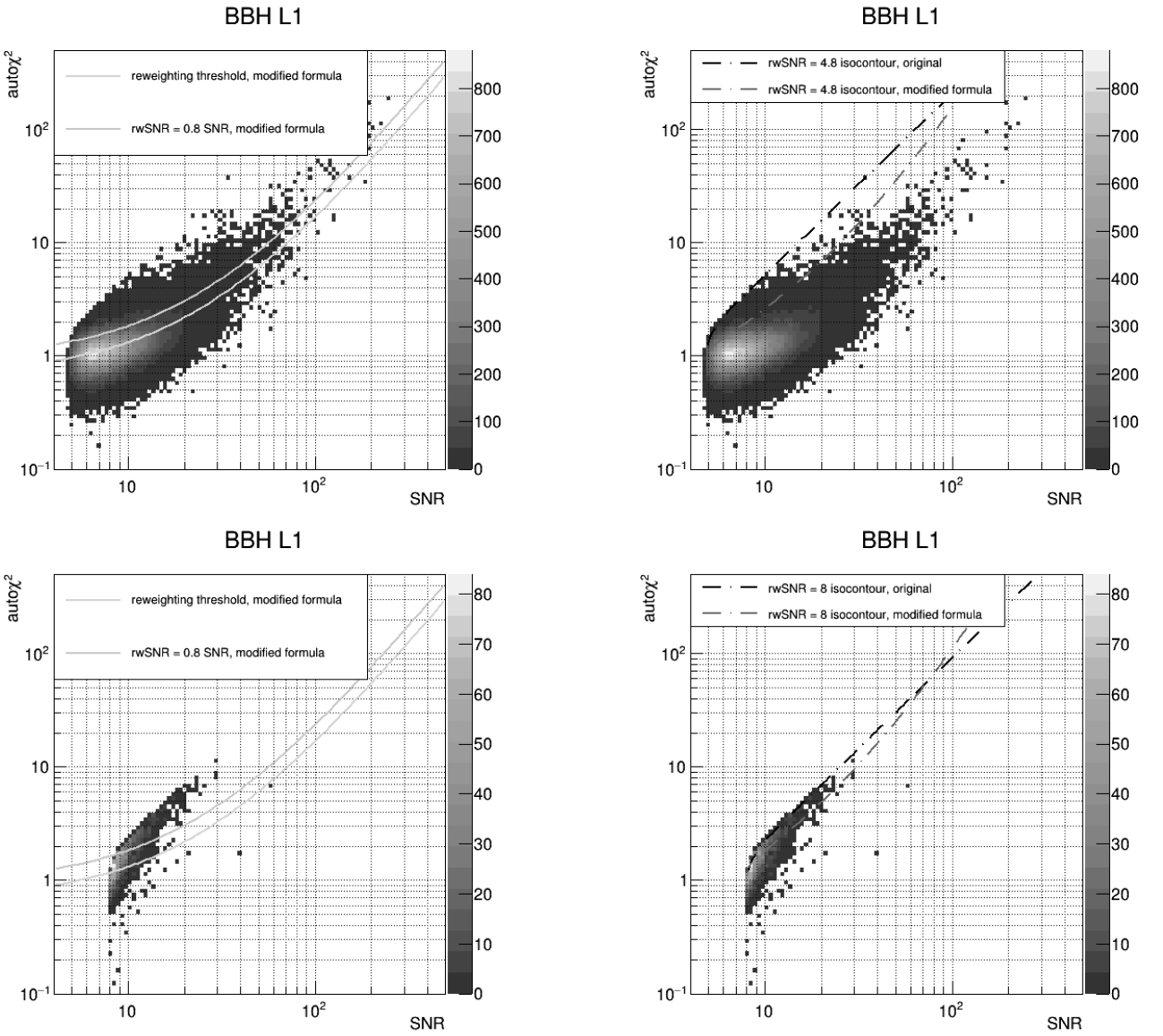


Figure 93: Top plots: O3 BBH injections autoχ^2 vs SNR dsitribution with rwSNR thresholds as a fraction of the SNR (left) and detection threshold (right) using $\alpha = 12$ in formula 73. Bottom plots: same plots for O3 BBH background single detector triggers.

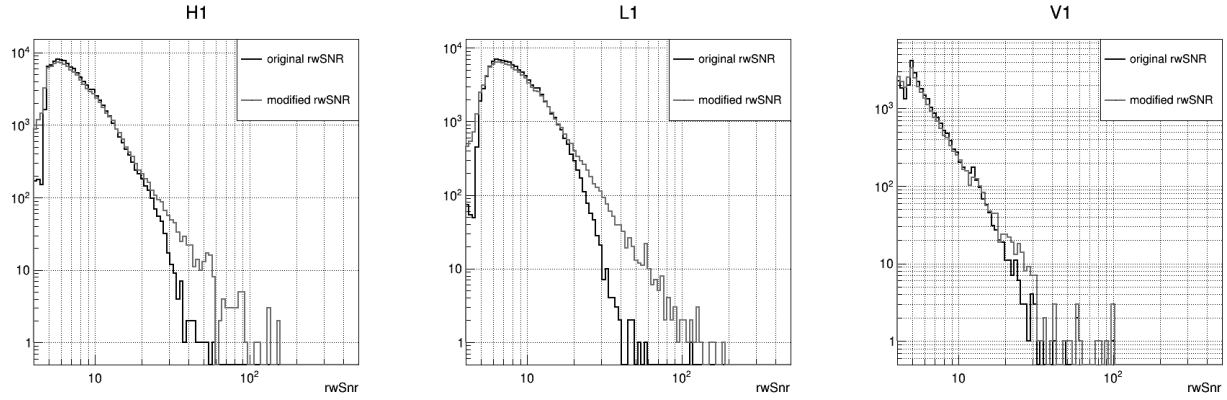


Figure 94: O3 BBH injections: comparison of the rwSNR distribution using the modified reweighting formula versus the original one.

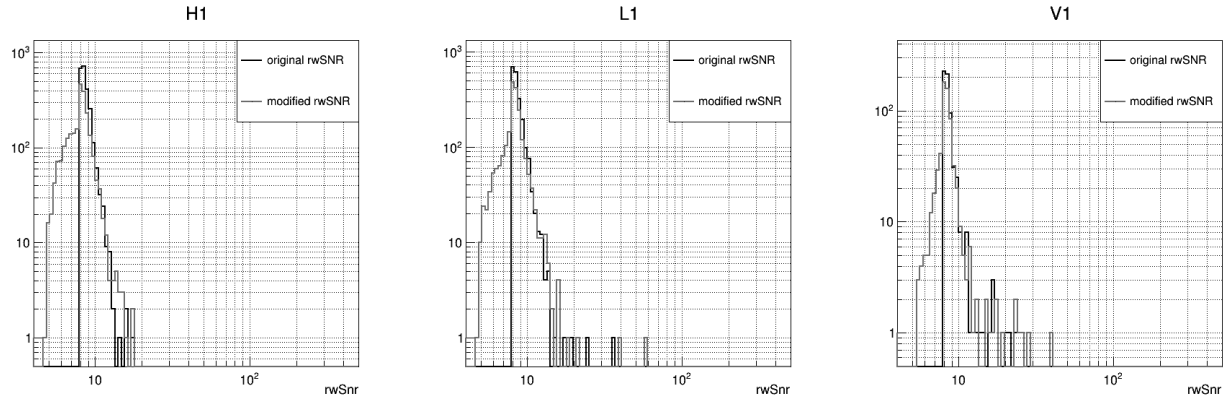


Figure 95: O3 BBH background single detector triggers: comparison of the rwSNR distribution using the autoχ^2 2' and $\alpha = 12$ versus the original one.

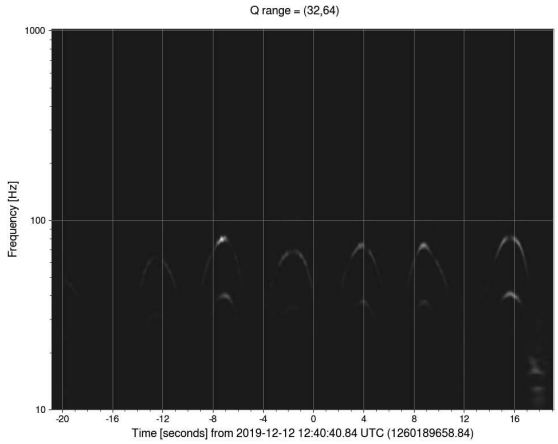


Figure 96: Trigger in L1 at GPS=1260189658.84 with modified rwSNR=59.28 .

8 Trigger rate and PSD fluctuations

Contents

Following the study on the rejection of noise triggers using criteria presented in chapter 5 we will, in this chapter, further investigate why we have more single detector triggers on real data than on Gaussian noise. Understanding why O3 single detector triggers have higher SNR could help in reducing their number or at least take actions to treat them adequately. We will investigate one hypothesis in the present chapter: there may be excesses of noise on short time scales which cause fluctuations of the detector sensitivity that are not properly accounted for.

8.1 Triggers boosted by fluctuations of the PSD

The detector sensitivity is directly related to its PSD and can be monitored using the local BNS range computed by the gating process (see section 3). This BNS range is computed over 0.25 s at 32 Hz, it has therefore fast variations. Rather than this BNS range we will use the median range used for the gating, to monitor the sensitivity changes. This median range is computed each second on the previous 10 s. We will call it range_{10s}.

The reason why we decide to monitor the range fluctuations rather than the PSD itself is because the range allows to summarize the information given by the PSD in a single value. Also the range is computed with the proper weight within the frequency band in which we expect to find astrophysical signals. Noise at frequencies outside of this bandwidth will thus not make it fluctuate.

When doing the matched-filtering MBTA uses the median PSD over the last 1000 s for the high frequency band and the last 4000 s for the low frequency band. We want to investigate how the range_{10s} behaves close to the single detector triggers times and compare it to the median range computed over 1000 s (range_{1000s}) or 4000 s (range_{4000s}). This may highlight some effect that could explain, at least in part, why we have more single detector triggers on real data than Gaussian noise.

8.1.1 Range drop around triggers

Figure 97 shows the distribution of the range_{10s} at times where MBTA saved a background single detector trigger (i.e. excluding GWTC-2/3 triggers). Only times which pass the single detector triggers criteria are considered. We see that this distribution is shifted towards lower values when compared to the range_{10s} distribution all over O3. If this is due to extra noise at the time of the trigger, and not just due to the extra power coming from the event, we would be under-estimating the PSD and thus over-estimating the SNR of the triggers.

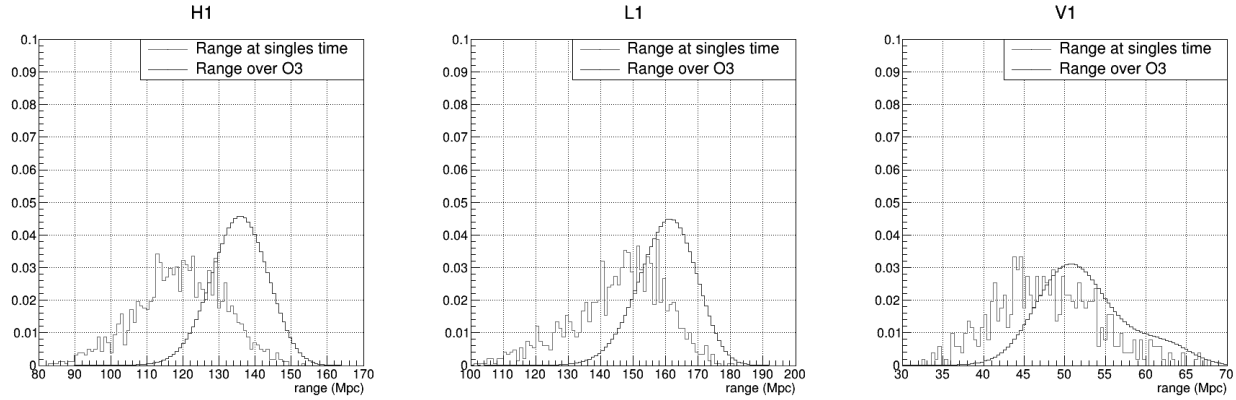


Figure 97: range_{10s} distribution over O3 (times which do not fit single selection criteria were removed).

Figure 98 shows the distribution, for O3 background single detector triggers, of the range_{10s} at the time of the single detector triggers compared to the range_{1000s} and range_{4000s}. We see that the range_{10s} at the time of single detector trigger is indeed smaller compared to the median range computed on a longer time by more than 10%.

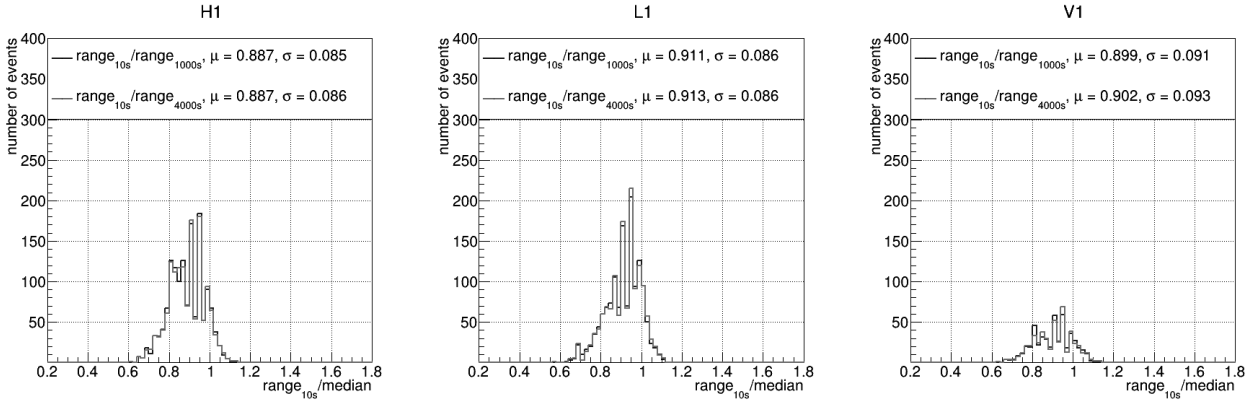


Figure 98: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for all of O3 singles (see text for the definition of the ratio).

We wonder what the effect is on EM bright single detector triggers, as we have seen that we have less noise triggers for this population. We see on figure 99 that the effect is still present, although smaller ($\sim$ 1.7-2.9%$\pm \sim$ 0.4%) than when looking at all single detector triggers.

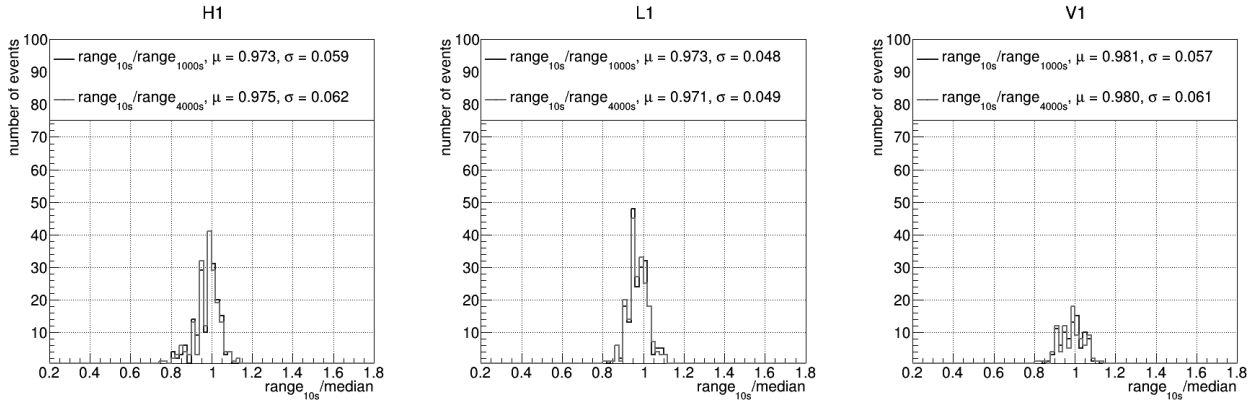


Figure 99: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for O3 EM bright singles with selection criteria applied. Their rwSNR distribution is shown in figure 68. Considering ~ 200 triggers in H1 or L1, the error on the mean value is $\sim 0.4\%$ in H1, $\sim 0.3\%$ in L1. For ~ 100 triggers in V1 the error on the mean value is $\sim 0.6\%$.

Looking at O3 long EM dark single detector triggers after selection on ER and gating in figure 100, we see that there is still a large offset of more than 10%. This is expected because although we called them "long EM dark" many of these triggers have a duration of only a few seconds. An excess of power in the detector during a small duration will make the range drop on a short time.

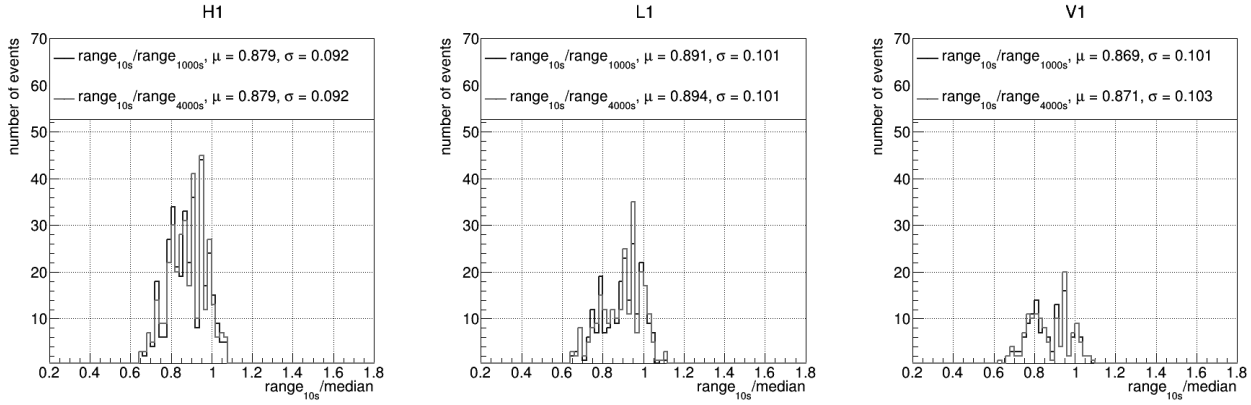


Figure 100: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for O3 long EM dark singles with selection criteria applied.

We can also check whether this effect occurs for Gaussian noise, as it consists of low SNR triggers. The range computed on Gaussian noise should not fluctuate a lot. Thus, we expect the effect to be much smaller, if there is an effect. Figure 101 shows the distribution of the range ratio for single detector triggers obtained on Gaussian noise. We see that we have here an effect smaller than 1%.

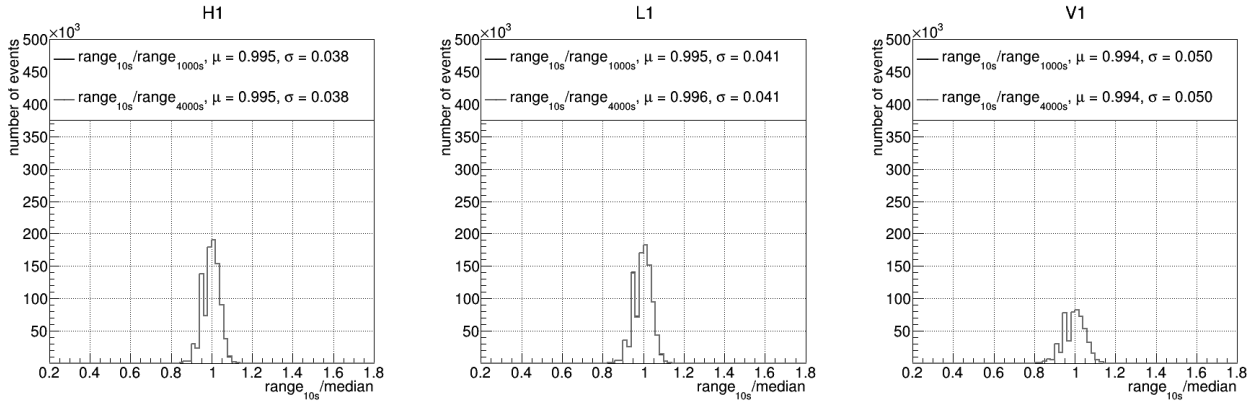


Figure 101: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for all (EM bright or not) single detector triggers obtained from a Gaussian noise analysis. The distribution is rescaled to O3 observing time.

The EM bright single detector triggers which passed the selection criteria defined in section 5 were compared to triggers obtained on Gaussian noise in figure 68. Taking the mean value of the histograms in figures 99, we can shift this Gaussian noise by 2.6% in H1 and 2.9% in L1. Figure 102 shows the comparison of the shifted Gaussian noise distribution with the single detector triggers. We see that we can explain part of the difference between the two with the observed effect. However, this shift of the Gaussian noise assumes that all the range reduction observed at the time of single detector triggers is coming from extra noise. We have to see if this is a background-only effect. One way to investigate this, is to have a look at astrophysical signals and randomly selected times.

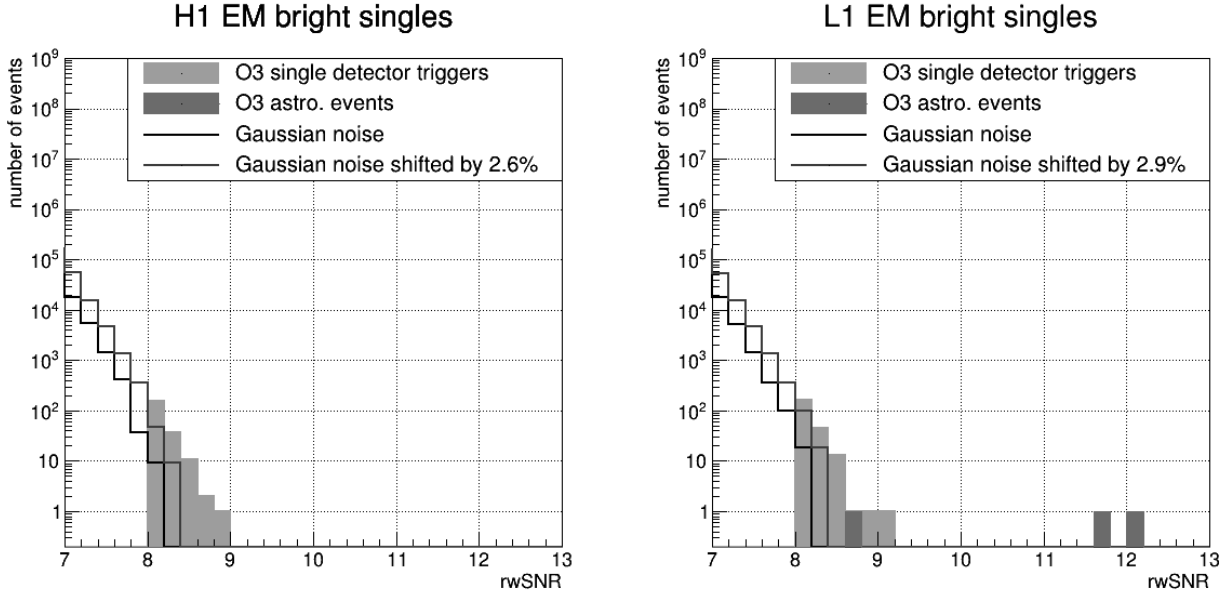


Figure 102: Comparison of O3 EM bright single detector triggers which passed the selection to shifted Gaussian noise distributions using the observations described in this section.

8.1.2 Range drop due to astrophysical signals

Astrophysical signals could also reduce the "local" range because they are adding signal on top of the detector noise. Therefore they can increase the PSD of the detector. In this section, we are investigating this effect, to see if it explains all or only part of the observed range change described in the previous section.

Figure 103 shows the same distributions as figure 98 for single detector triggers associated to O3 astrophysical events (all type mixed, selection criteria applied). We see that the effect is indeed present but not as strong as in figure 98, however there is only a small number of detection which prevents us to draw any solid conclusion.

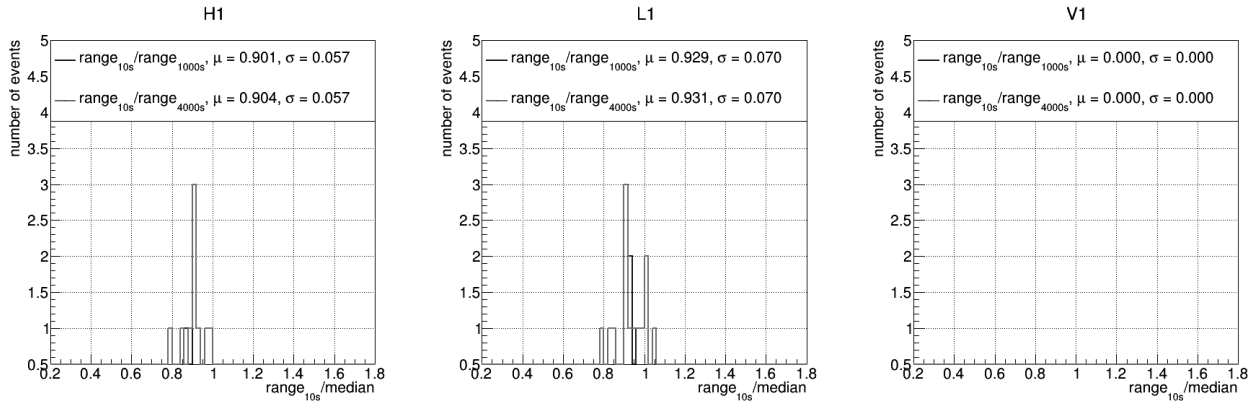


Figure 103: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for all O3 astrophysical triggers.

To overcome the low statistics issue, we can also see this effect for simulated astrophysical signal as shown in figure 104 for injections on O3 data, analyzed with the O4 configuration of the pipeline. They are split in EM bright and EM dark populations. The effect is smaller for the EM bright injections, with a shift smaller than 1% in the distribution. This is for the same reason as O3 single detector triggers. BBH injections are loud and short, causing a drop in the range of the detector. On the other hand, for EM dark injections, the effect is of the same order of magnitude as for real astrophysical signals: $\sim$ 7.1-9.3%. Comparing the offset for O3 EM bright single detector triggers ($\sim$ 1.7-2.9%$\pm \sim$ 0.4%) with the offset for EM bright injections ($\sim$ 0.5-0.8%$\pm \sim$ 0.1%), the effet is statistically significant.

On the contrary, we do not expect to see such an effect if we look at the variation of the range_{10s} at random times. This is confirmed by figure 105 which shows a deviation smaller than 0.3%.

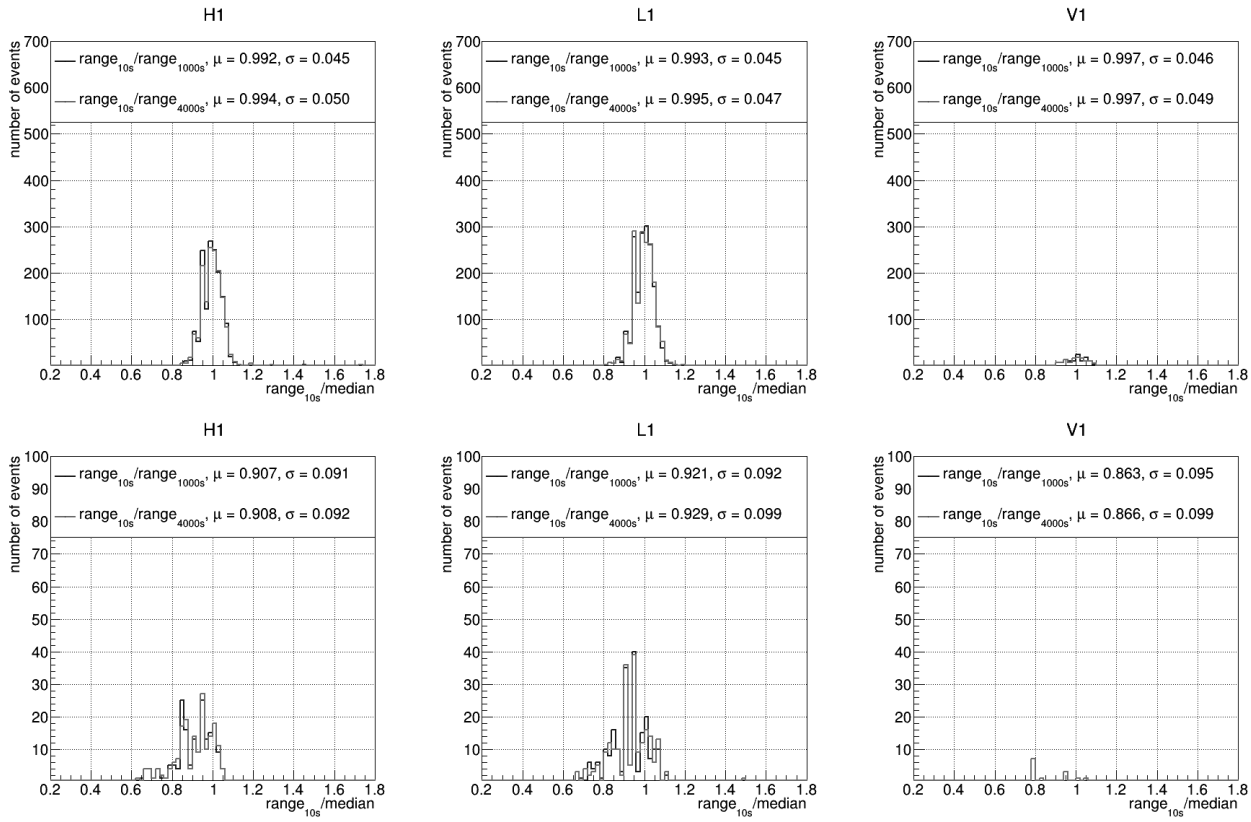


Figure 104: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for recovered injections added to O3 data, analyzed using MBTA O4 configuration. Top: EM bright population. Bottom: EM dark population. For EM bright injections, considering ~ 1500 recovered injections in H1 and ~ 1750 in L1 the error on the mean is $\sim 0.1\%$ for both. It is $\sim 0.4\%$ in V1 due to fewer recovered injections.

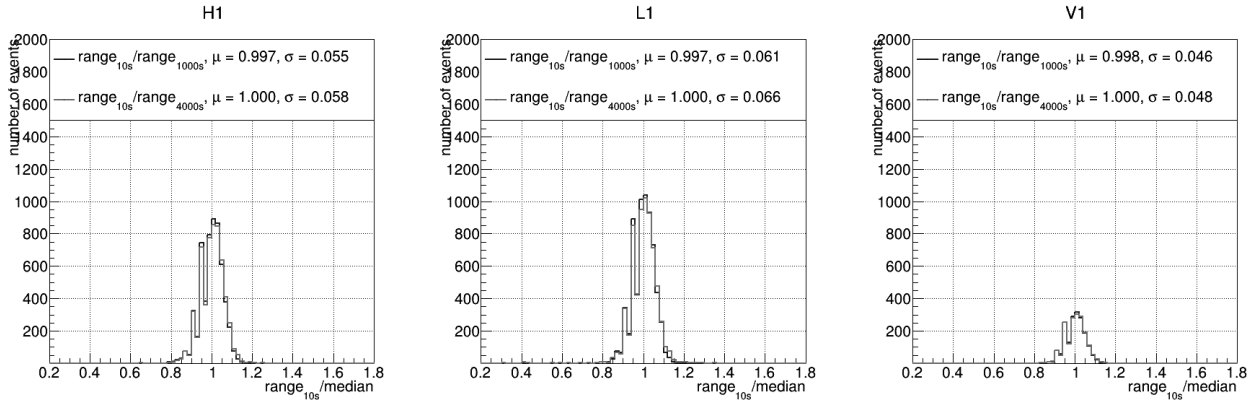


Figure 105: $\text{range}_{10s}/\text{range}_{1000(4000)s}$ for random times.

This is strong evidence in favour of our initial hypothesis: there are PSD fluctuations that are not properly accounted for when performing the matched-filtering, causing the PSD to be under-estimated at trigger times leading to over-estimate their SNR values. This results in a larger number of single detector triggers above a given SNR threshold.

8.1.3 Time scale of the drop in range

We want to know on what time scale the drop in sensitivity happens. Starting with all background single detector triggers, we compute for all of them the ratio of the range at each time with the range_{10s} at the time of the trigger. Then, at each time, we take the median value of the ratio over all EM bright single detector triggers. The black curve in figure 106 shows the evolution over time of the median of the ratio over all EM bright triggers. We see that the time scale and magnitude of the effect depends on the detector, with L1 being the one in which it is the most visible. It also appears that the minimum for the ratio is reached just a few seconds before the trigger.

We have seen that there is no offset for random times. But we can create one by selecting cleverly our random times (which are therefore not so random anymore). We can search, for each random time, the time which has the minimum range_{10s} value within for example $\pm$ 40 s. If we then compute the range ratio around this time, we should see a similar effect to the one observed for single detector triggers. This is shown by the red curve in figure 106. We see that the drop in the ratio is much larger than what is observed for single detector triggers. But unlike the effect observed in L1, the drop happens only in a matter of seconds corresponding to the time window on which the local range_{10s} is computed.

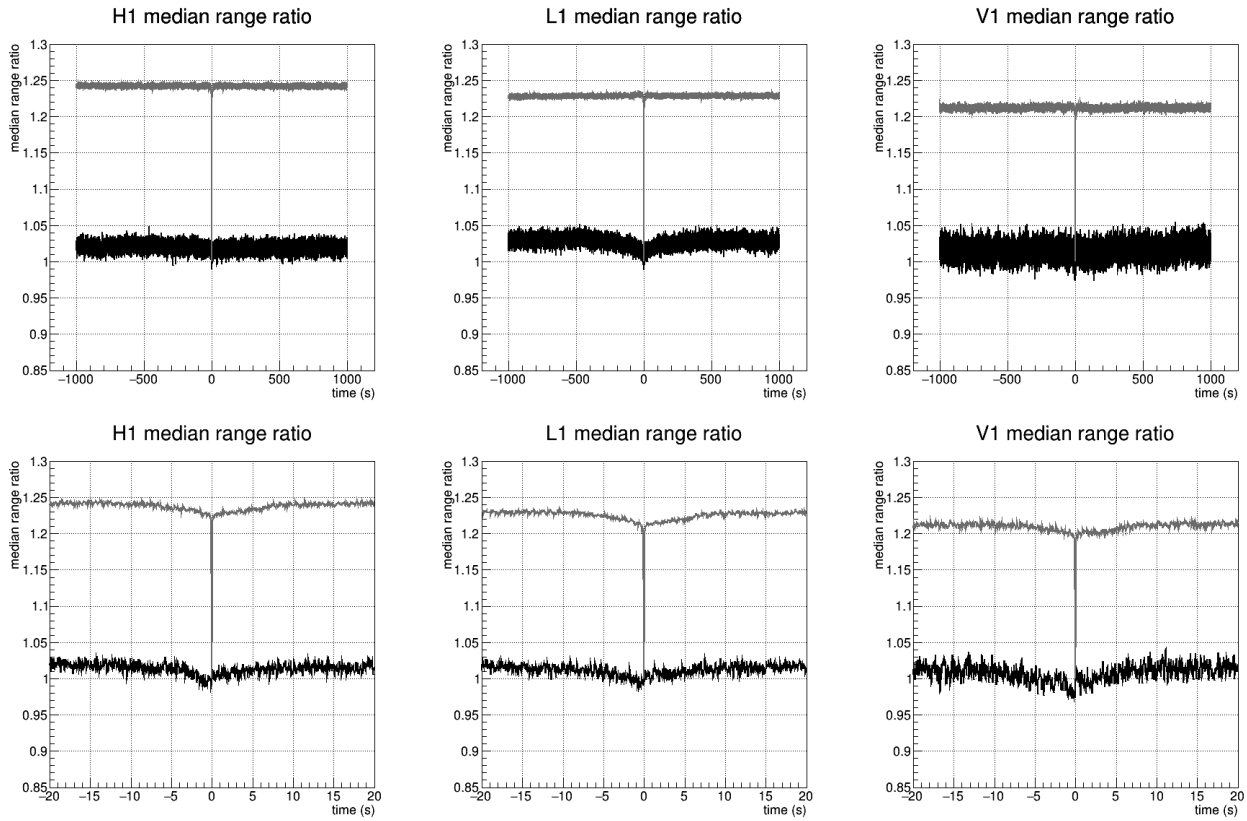


Figure 106: range(t)/range(t_0=0) (t_0 being the time of the single detector trigger or minimum range for random times). The figure is showing in black for each bin the median value over all EM bright single detector triggers. The red line shows the same for the lowest value of random segments (see text). Top: looking at $[-1000; +1000]$ s. Bottom: zoom on $[-20; +20]$ s

8.2 Correcting the SNR for range fluctuations

Solving the issue highlighted in the previous section is no easy task. The PSD can't be measured with precision on short durations with high frequency resolution. Instead an idea was to apply a correction to the SNR by taking into account the range fluctuations around the time of the trigger. The weight applied would be the ratio between the range_{10s} at the time of the single detector trigger and the median range computed over a "much longer" duration before the trigger. It would also mean reducing the SNR of astrophysical signals. For instance, the 3 EM bright astrophysical events of O3, GW190425, GW200105_162426 and GW200115_042309 have a range ratio of 0.99, 0.97 and 1.00. But we can still have a look at the effect of the correction for background triggers.

Figure 107 shows the comparison of the uncorrected distribution with the distribution corrected using the range ratios $\text{range}_{10s}/\text{range}_{1000s}$. Two types of corrections are shown: considering all range ratios (meaning that the SNR can be increased by the correction) and considering only ratios smaller than 1. We see that, considering all ratios, even if we have less single detector triggers above rwSNR=8 overall, the result is not satisfactory as we now have more triggers with SNR around 9. Using only ratios smaller than 1, however, allow to reduce nicely the background. We will therefore only consider corrections when the range ratio is smaller than 1 in the following.

Accounting for the time scale of the range drop as shown in figure 106, we decide to investigate a correction using a median value of the range_{10s} instead of the instantaneous value at the time of the trigger. This should avoid to be sensible to the rapid fluctuations of the range ratio. The result is shown in figure 108. The correction here is done using the median value of the range_{10s} over the 4 s preceding the triggers (chosen arbitrarily), divided by range_{1000s}. It is compared to the correction using the range_{10s} at the time of the event. It makes things slightly better at high rwSNR.

Additionnally, we may wonder how important the choice of 4 s was for the computation of the median. Figure 109 shows that the duration chosen for the computation indeed has an impact on the result, but it is not clear what the best choice is. Note that we have only considered range ratios smaller than 1 here. The range_{1000s} was chosen for this figure.

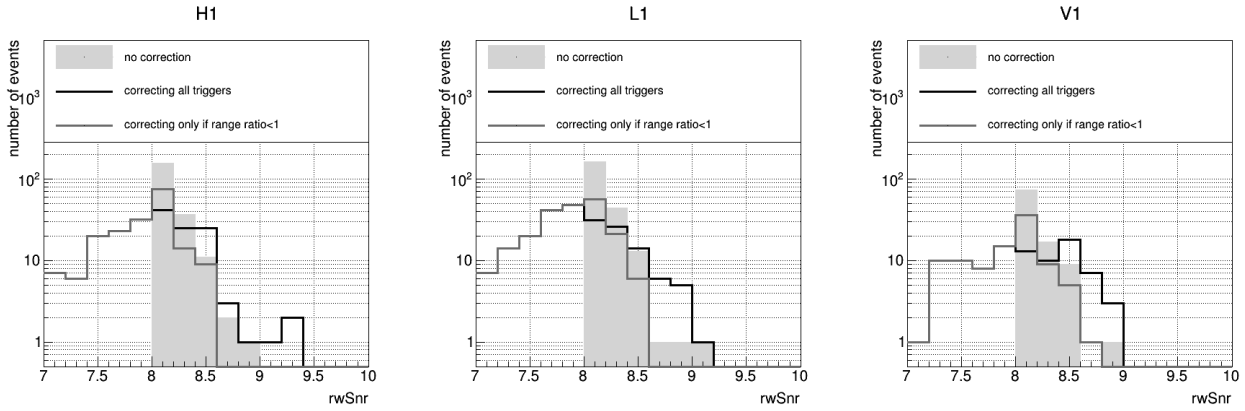


Figure 107: Corrected distribution compared to the original distribution for O3 EM bright singles (with selection criteria) using $\text{range}_{10s}/\text{range}_{1000s}$. Black: correcting all triggers. Red: correcting only if the ratio is smaller than one.

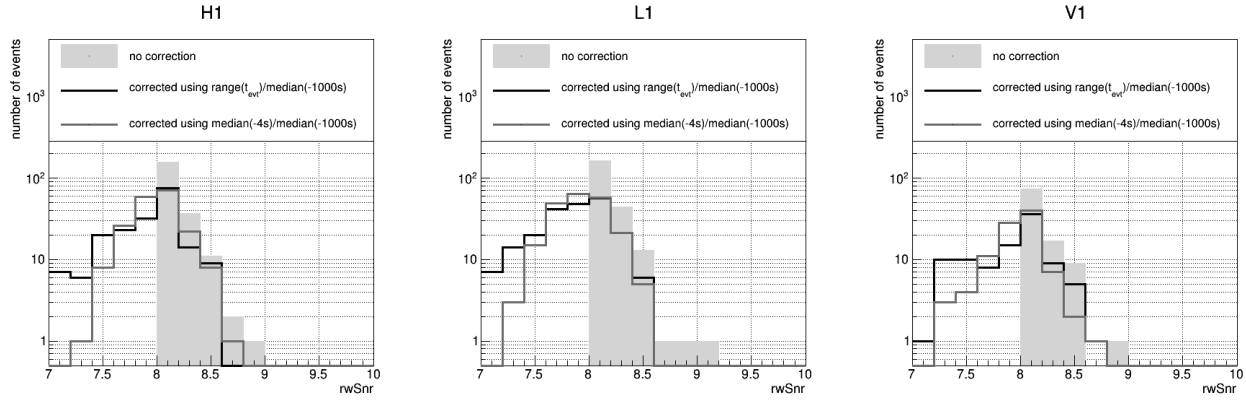


Figure 108: Corrected distribution for O3 EM bright singles (with selection criteria). Black: $\text{range}_{10s}/\text{range}_{1000s}$. Red: using the median over 4 s before the trigger of the range_{10s} instead of the median range. Corrections are only applied if the ratio is smaller than one.

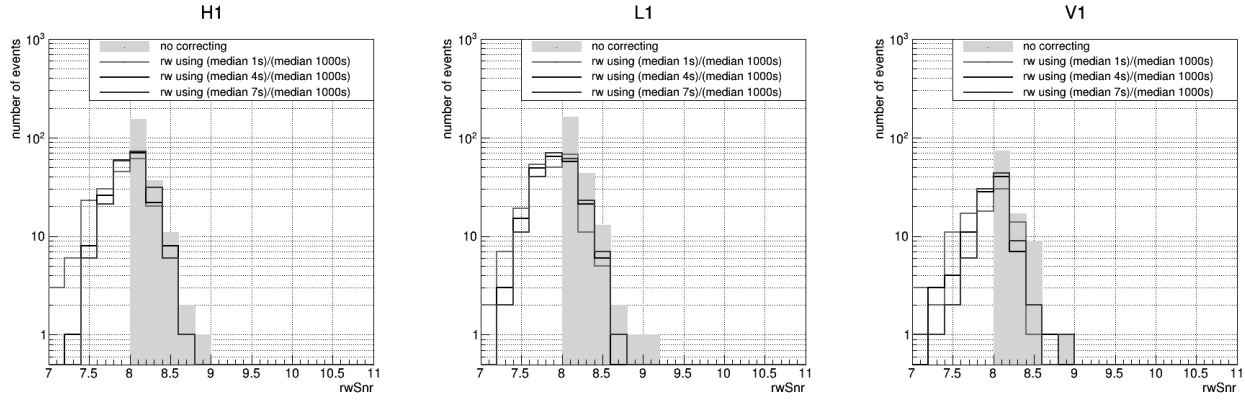


Figure 109: Corrected distribution compared to the original distribution for EM bright singles using various median values of the range_{10s} before the trigger. Corrections are only applied if the ratio is smaller than one.

We saw earlier that shifting the Gaussian noise rwSNR distribution by a few percent (according to the previous section) reduced the difference between O3 EM bright single detector triggers and the Gaussian noise distribution. We now try another approach and investigate if correcting the SNR using the range ratio affects the single detector triggers obtained on Gaussian noise. Figure 110 shows that the Gaussian noise is barely affected by this reweighting. Figure 111 shows the comparison of the corrected distribution for O3 EM bright single detector triggers with the corrected distribution for the Gaussian noise. The two distributions are now indeed closer to one another, although there is still room for improvement.

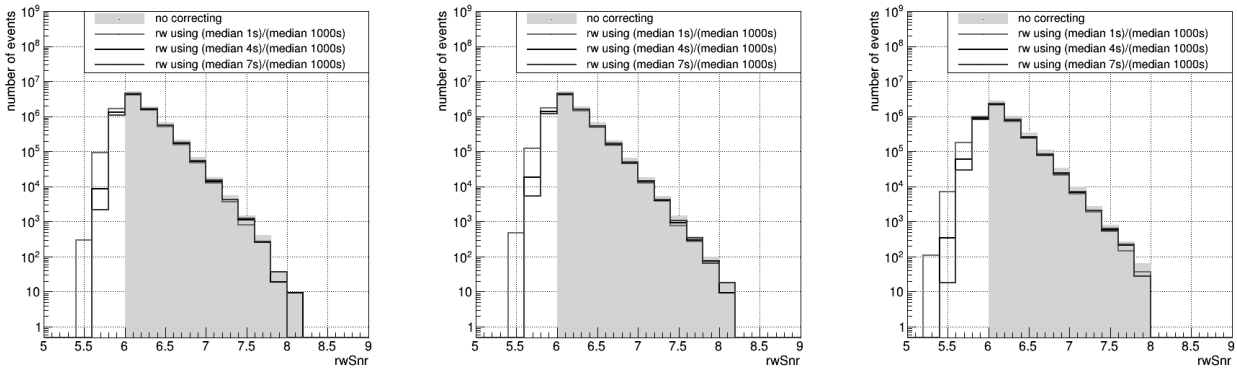


Figure 110: Corrected distribution for the EM bright single detector triggers obtained by analyzing the same gaussian noise as in section 5. Note that the Gaussian noise distribution is scaled to the observing time of O3. Corrections are only applied if the ratio is smaller than one.

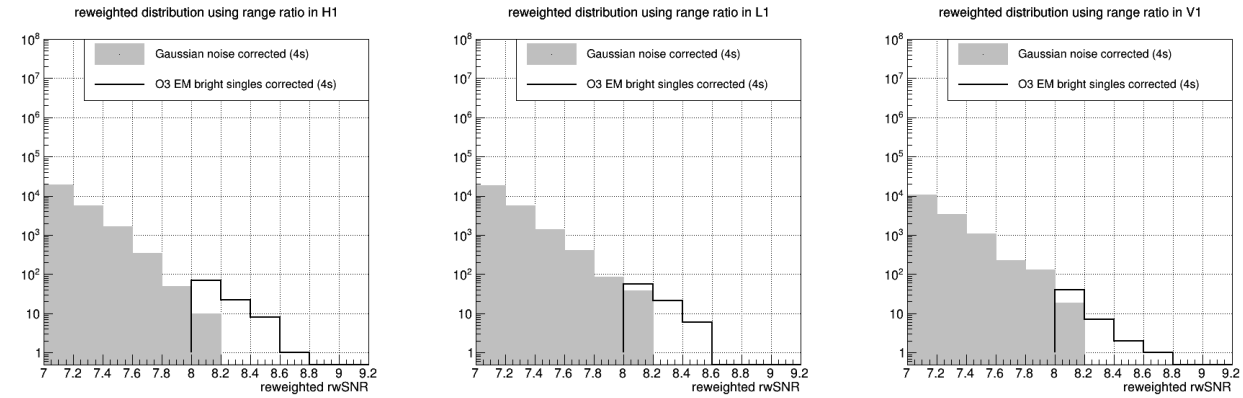


Figure 111: Corrected O3 EM bright single detector triggers distribution compared to the corrected gaussian noise distribution (also EM bright singles, scaled to O3 observing time). The correction is the ratio of the median computed over 4 s before the trigger with the range_{1000s}. Corrections are only applied if the ratio is smaller than one. O3 events with a corrected rwSNR smaller than 8 (initial selection threshold) were not included in this figure.

We can also investigate a correction for the long EM dark single detector triggers. We saw that for injections, the mean value of the range ratio distribution was smaller than 1, for instance 0.92 in L1. In order to correct the SNR of the single detector triggers without impacting too much the injections, we choose to correct the SNR if the range ratio is smaller than 0.92. Figure 112 shows the effect of the correction when considering all long EM dark single detector triggers and when considering only those with ratio smaller than 0.92. We want the correction we apply to start at 1 when the ratio is 0.92 and decrease for lower values. The correction applied is therefore the ratio divided by 0.92. Figure 113 show the effect of the correction (for range ratios smaller than 0.92) on O3 long EM dark single detector triggers when considering the median of the range_{10s} over the 4 s preceding the trigger in the ratio. We see that for long EM dark triggers, using the range_{10s} at the time of the trigger allows to reject more triggers while using the median over 4 s has almost no impact. This is explained by the fact that the drop in the range for long EM dark triggers happens

in a short time and is more important than for EM bright single detector triggers. Therefore taking the median value of the range_{10s} over the 4 s before the trigger makes the difference with the range_{1000s} before the trigger much smaller. Since we chose to correct triggers with a ratio smaller than 0.92, only a very small number of triggers are corrected in this case.

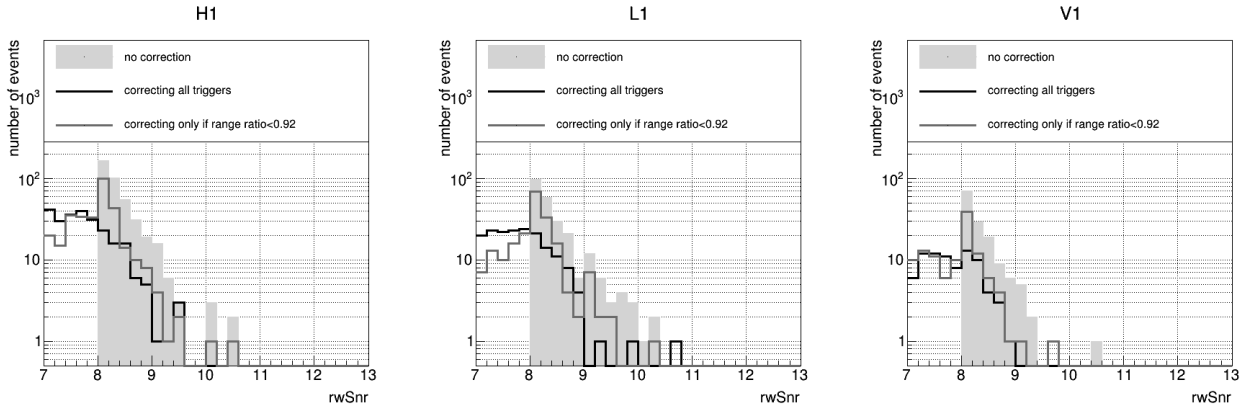


Figure 112: Corrected distribution compared to the original distribution for O3 long EM dark singles (with selection criteria) using $\text{range}_{10s}/\text{range}_{1000s}$. Black: correcting all triggers. Red: correcting only if the ratio is smaller than 0.92.

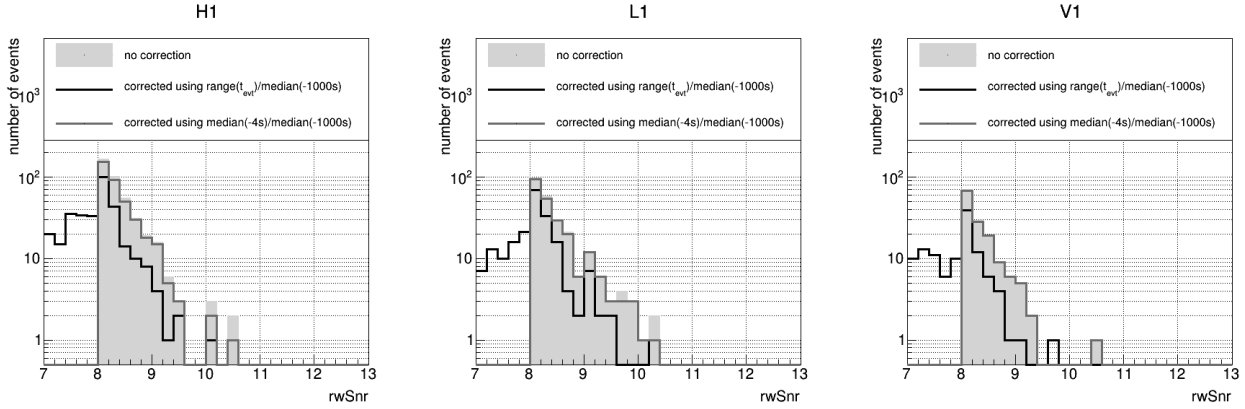


Figure 113: Corrected distribution for O3 long EM dark singles (with selection criteria). Black: using $\text{range}_{10s}/\text{range}_{1000s}$. Red: using the median over 4 s before the trigger of the range_{10s}. Corrections are only applied if the ratio is smaller than 0.92 and is equal to the ratio divided by 0.92.

It was decided to leave things as they were for O4. The understanding we acquired of this mechanism tells us that this effect is taken into account in the method we use to compute the background for the FAR of the single detector triggers, presented in chapter 6.

8.3 Detector range: real data vs Gaussian noise

We have seen that by studying the variations of the range for the gating, we were able to identify an issue that increases our background. We wonder if there is more that can be learned. We decide this time to have a look at the range itself (not the range_{10s}).

It was shown that MBTA single detector triggers are quite far from the Gaussian noise level, although we were able to reduce this difference in chapter 5. We decided to investigate the differences in the range computed on real data compared to the range computed on Gaussian noise. The goal is not to propose a solution to fix specific issues but rather to highlight possible discrepancies in the amplitude and time scale of the variations of the range which could be studied in the future.

We start by having a look at the BNS range, computed for a $1.4 + 1.4\text{M}_\odot$ system starting at a frequency of 10 Hz. The evolution of the range in time and frequency for real data and Gaussian noise are shown in figure 114 and 115 respectively. We show the range for several decimation factors because it allows to be more sensitive to low frequency fluctuations, especially in the time domain, since we expect differences to be at low frequencies. The first thing we notice is the presence of many drops in the range computed on real data due to glitches and the absence of such drops in the range computed on Gaussian noise. We also see that there are more fluctuations on large time scales in the range computed on real data, characterized by an excess in the (average) ASD for frequencies smaller than 0.2 Hz when compared to the ASD of the range computed on Gaussian noise. The ASD of the range computed on real data also decreases slightly from 0.2 Hz to 1 Hz while the ASD of the range computed on Gaussian noise is flatter. The ASDs are however rather similar for higher frequencies with a drop around 4 Hz, corresponding to the time window (0.25 s) of the FFT used to compute the instantanious range (reminder: this time window is shifted by a step of 1/32 s to compute a range at 32 Hz).

Figures 116 and 117 show the 1-Dimensional projection of the range computed on real data and Gaussian noise respectively. The mean value of the range is roughly the same for real data and generated Gaussian noise but the standard deviation is larger in the case of real data up to a factor 2 when using a decimation factor of 256.

We also show in figures 118 and 119 the real data range and gaussian noise range as a function of time and frequency with starting frequency of 50 Hz. Starting the computation of the range at 50 Hz reduces the difference between the two ranges at low frequency.

We now wonder whether a range computed on a BBH system of $30 + 30\text{M}_\odot$ ($\nu_{\text{ISCO}} \sim 73\,\text{Hz}$) will exhibit the same effect. Figures 120 and 121 show the time and frequency dependency of the range computed on real data and gaussian noise respectively. Here again the ASD of the range computed on real data is higher at low frequencies but the difference is smaller than for the BNS range. Looking at the 1-D projections in figures 122 and 123, the two ranges have close mean values but different standard deviations with increasing difference as the decimation factor grows. Similarly to the BNS range, we show figures 124 and 125 for the computation of the range starting at 50 Hz. As for the BNS range, this makes the difference between the two ASDs at low frequencies barely noticeable.

This did not lead to any changes for O4. The goal is only to highlight some behaviours that may be of more interest in the future.

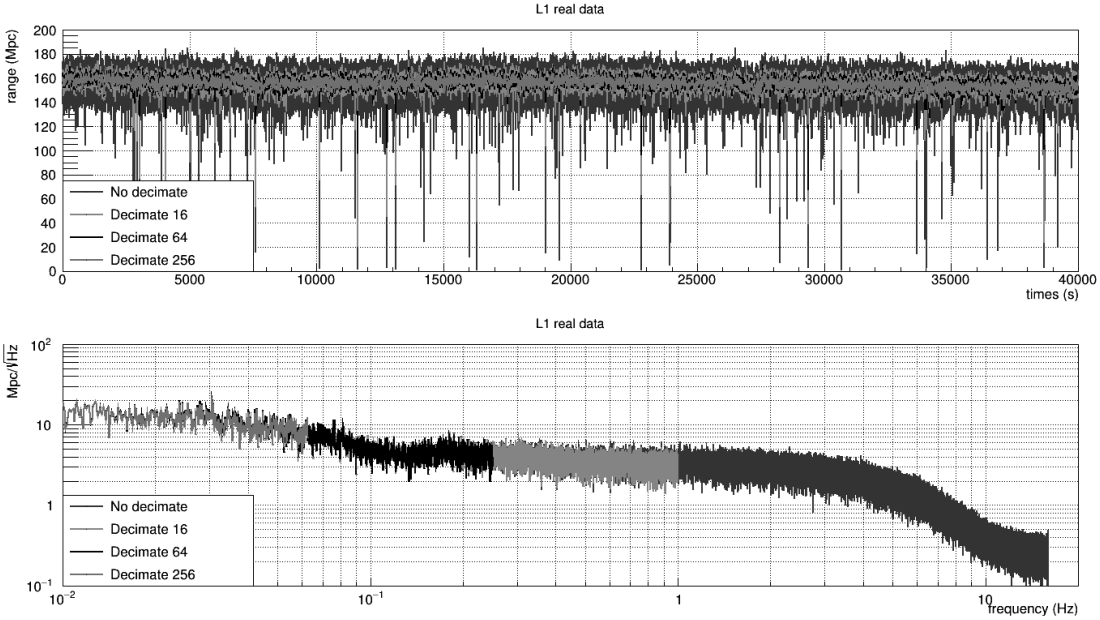


Figure 114: BNS range in the time and frequency domain, computed from 10Hz on real data

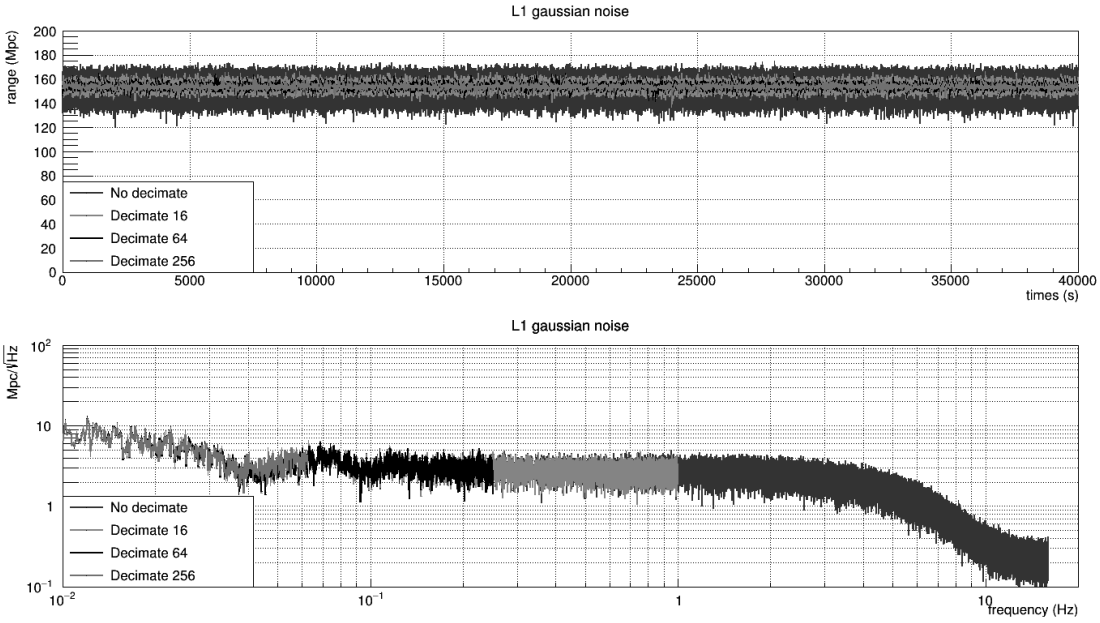


Figure 115: BNS range in the time and frequency domain, computed from 10Hz on Gaussian noise

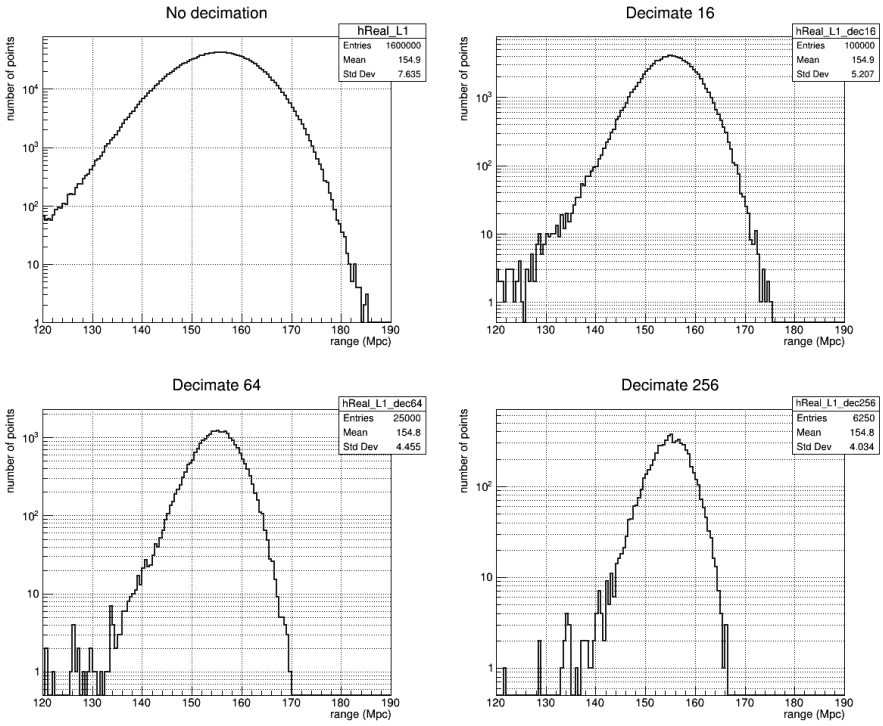


Figure 116: 1D projection of the BNS range in time domain computed from 10Hz on real data

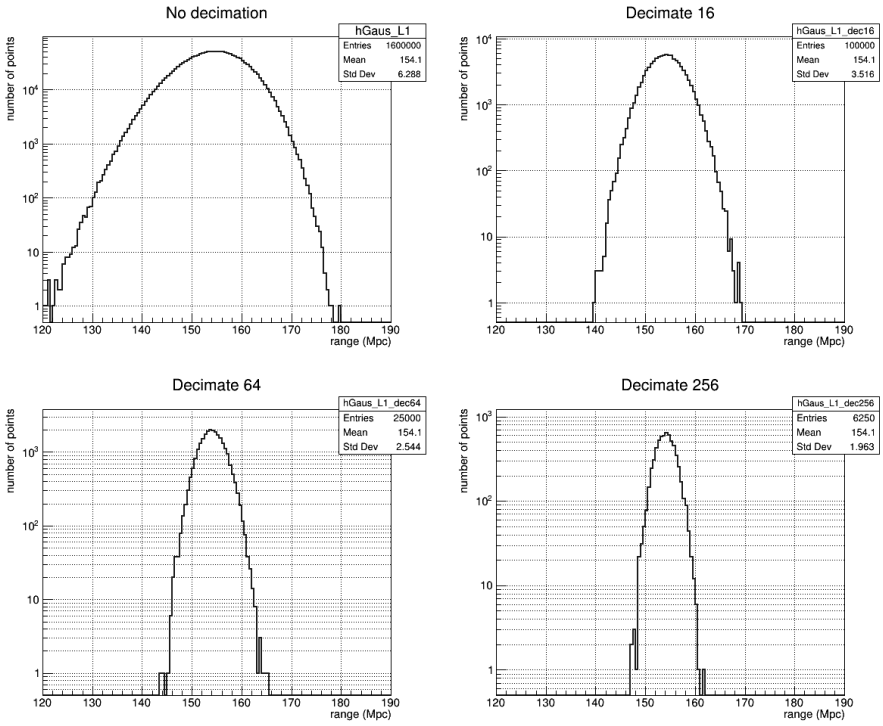


Figure 117: 1D projection of the BNS range in time domain computed from 10Hz on Gaussian noise

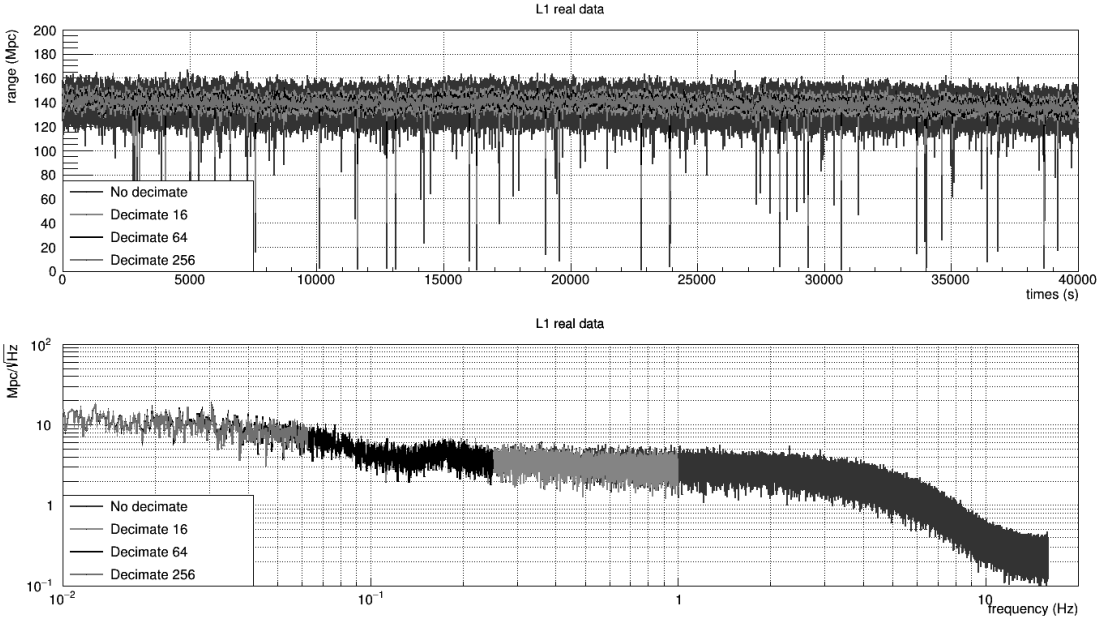


Figure 118: BNS range in the time and frequency domain, computed from 50Hz on real data

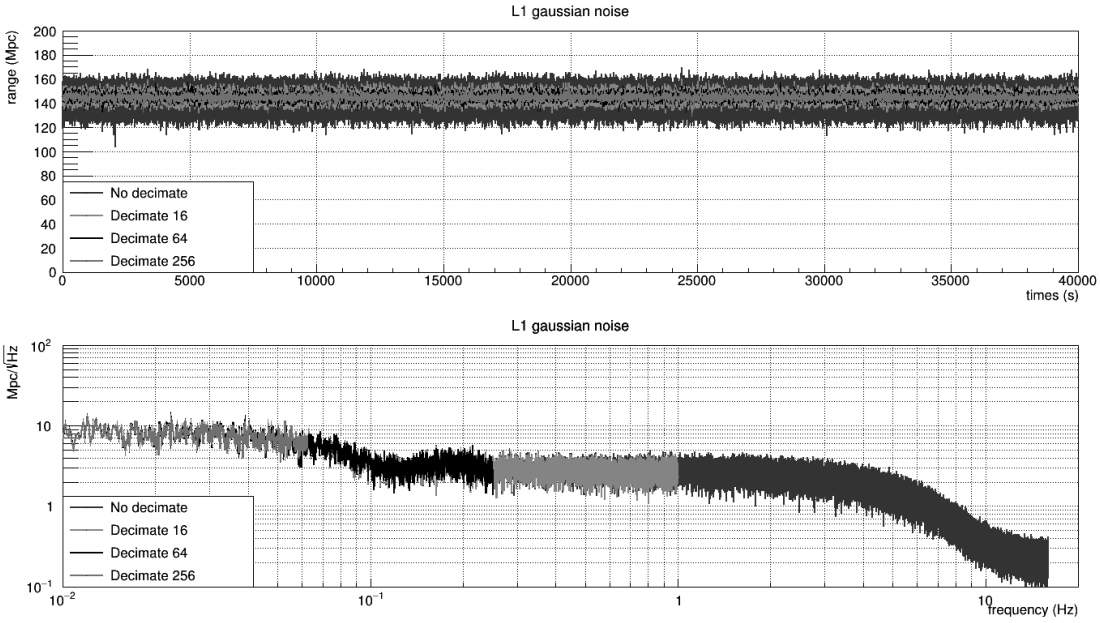


Figure 119: BNS range in the time and frequency domain, computed from 50Hz on Gaussian noise

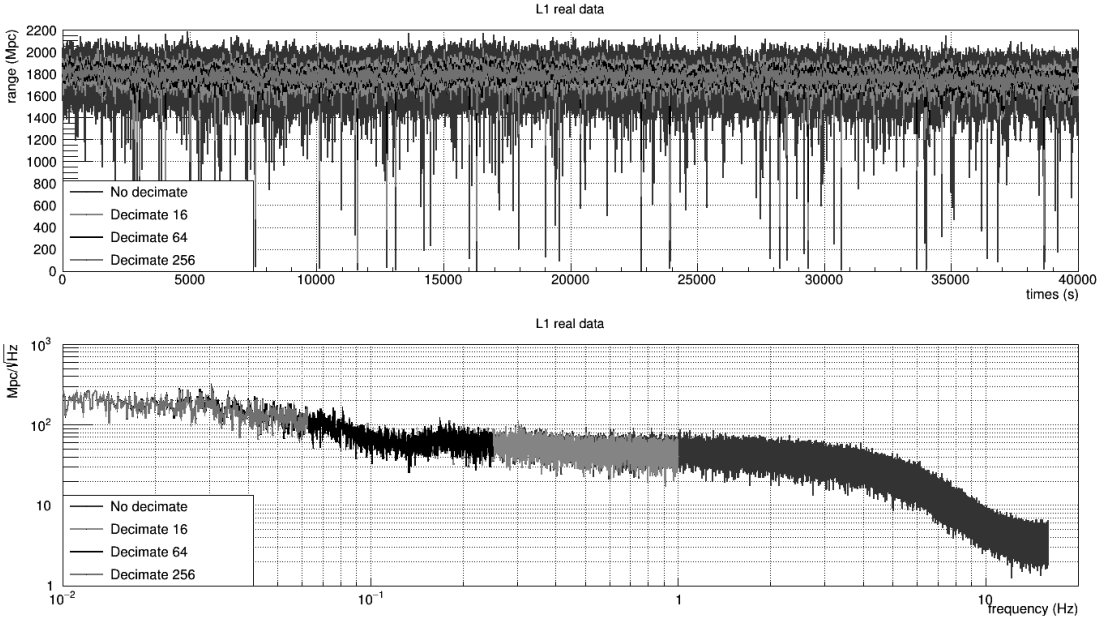


Figure 120: BBH range in the time and frequency domain, computed from 10Hz on real data

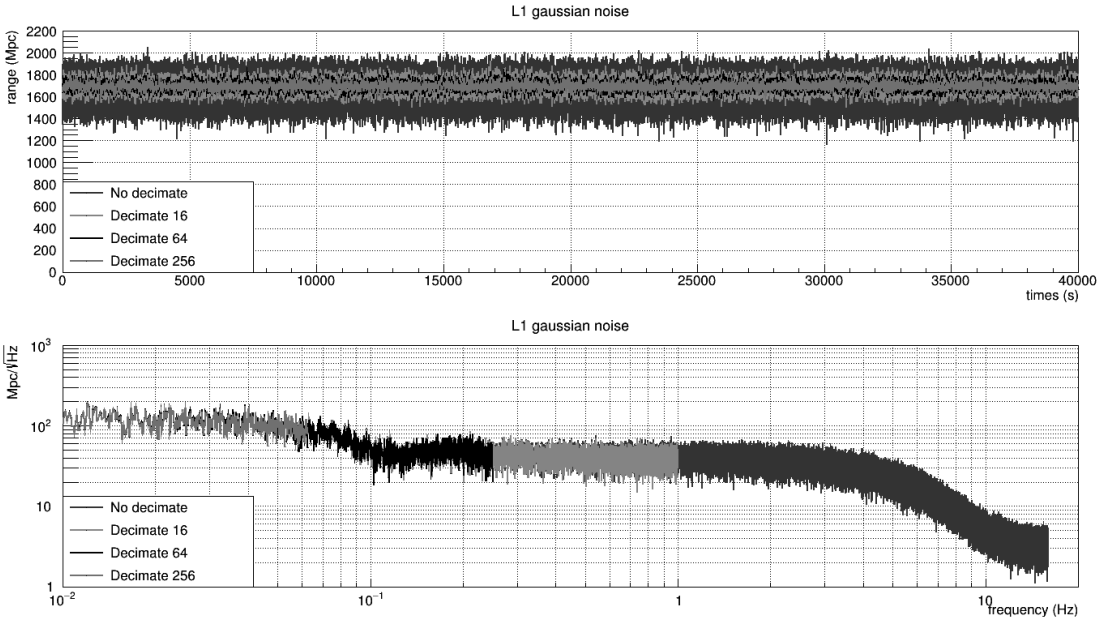


Figure 121: BBH range in the time and frequency domain, computed from 10Hz on Gaussian noise

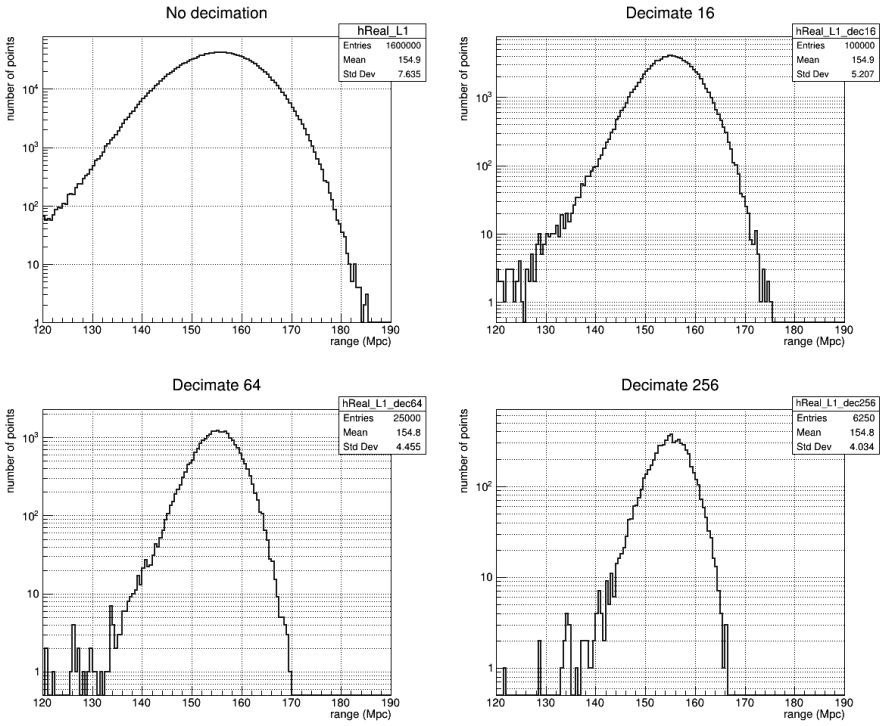


Figure 122: 1D projection of the BBH range in time domain computed from 10Hz on real data

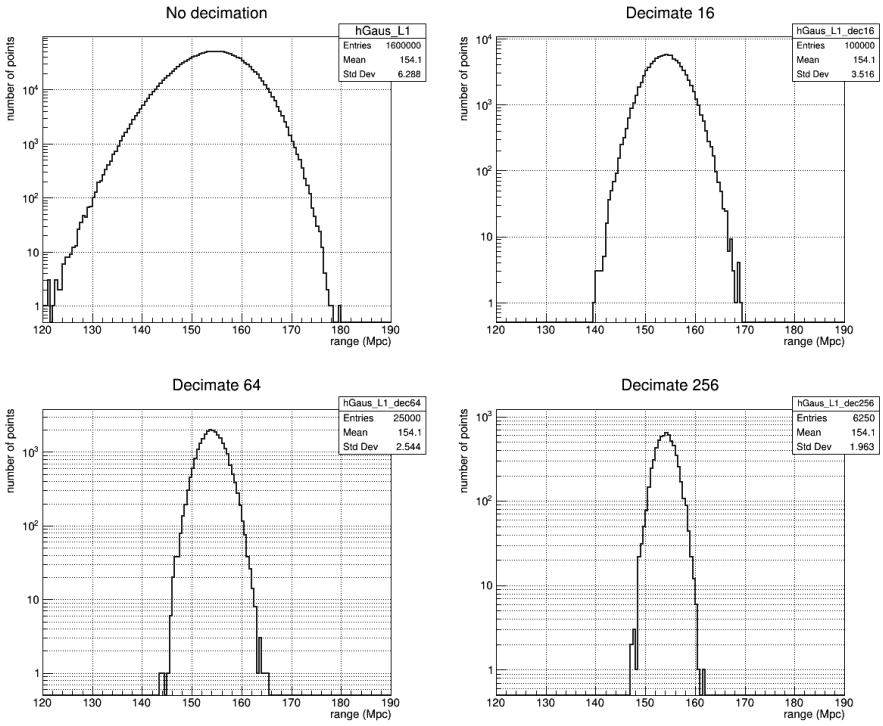


Figure 123: 1D projection of the BBH range in time domain computed from 10Hz on Gaussian noise

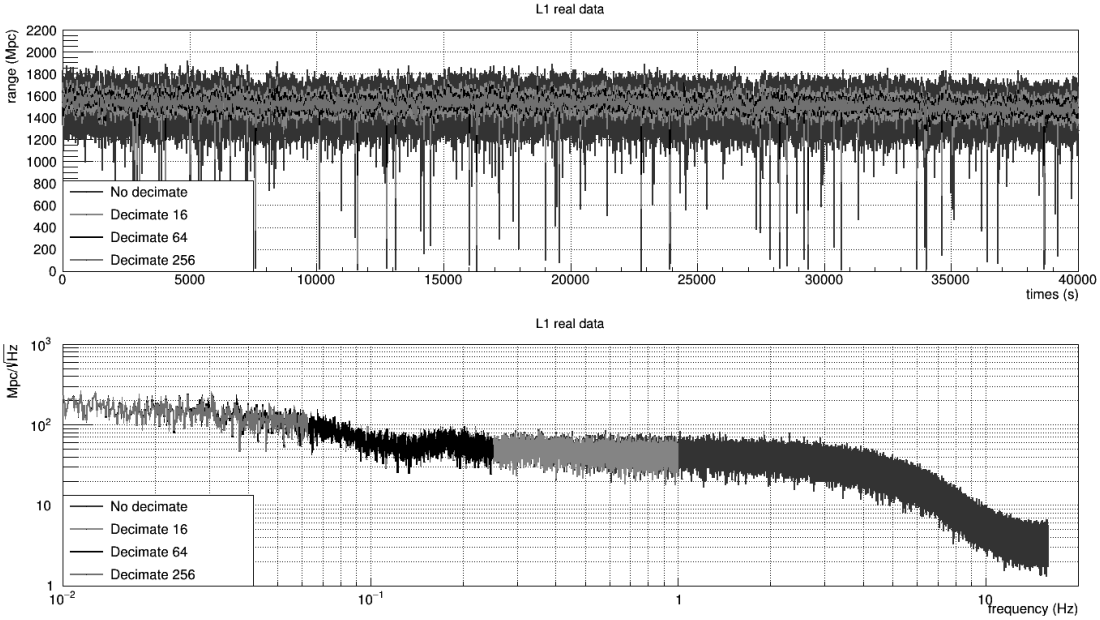


Figure 124: BBH range in the time and frequency domain, computed from 50Hz on real data

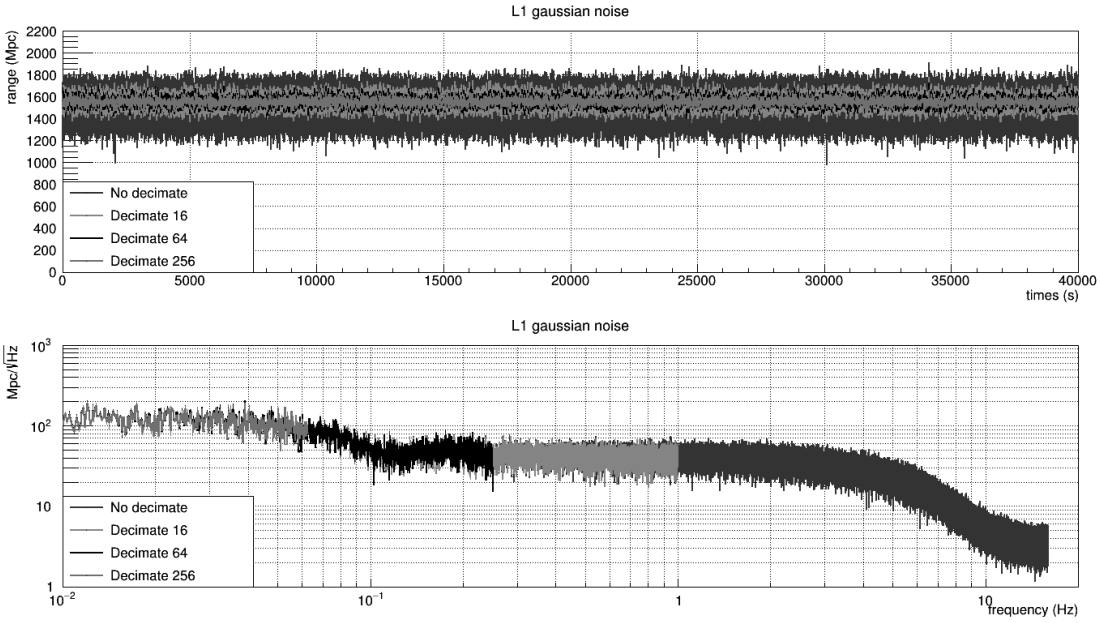


Figure 125: BBH range in the time and frequency domain, computed from 50Hz on Gaussian noise

9 Single detector search during the begining of O4

Contents

The fourth observing run started on the 24th May 2023 with only the two LIGO detectors taking data, Virgo needing some more time to reach the expected sensitivity. This means that there were only single and double detector time. Figure 126 shows the duty cycle of the detector network from the 24th May 2023 (GPS=1368933107) to the 9th July 2023 (GPS=1372947450) ($\sim$ 46.5 days). We give here an overview of the results of the single detector search for this period of time.

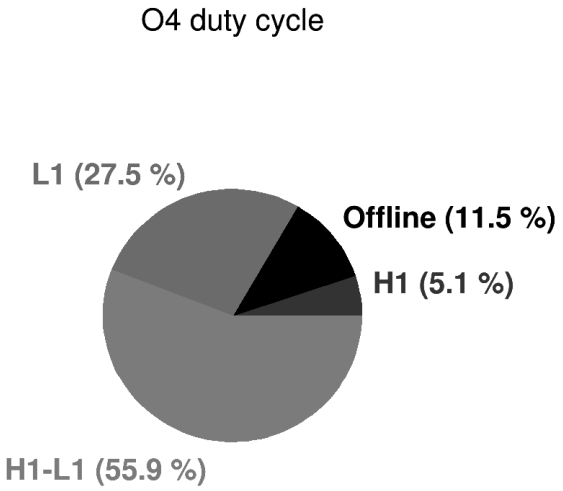


Figure 126: Detector network duty cycle for the first 46.5 days of O4.

9.1 Effectiveness of the selection criteria

We first want to check whether the single detector selection criteria defined previously worked as intended. In this chapter, we exclude all single detector triggers that are associated to a significant public alert. Figure 127 shows all single detector triggers for the same period of time as the duty cycle figure. Note that we do not distinguish here between background triggers and the few triggers that are likely to be of astrophysical origin.

Figure 128 shows the rwSNR vs template duration distribution (starting from 24 Hz, or 20 Hz for the high mass one band templates) of those single detector triggers. We can see that there are still some short templates which produce many triggers with sometimes high rwSNR.

For the single detector analysis done on O3 data, we defined the EM bright population to restrain the parameter space and reduce the background. The population was defined based on the chirp mass and mass1 of the systems in such a way that it would be similar to a selection on

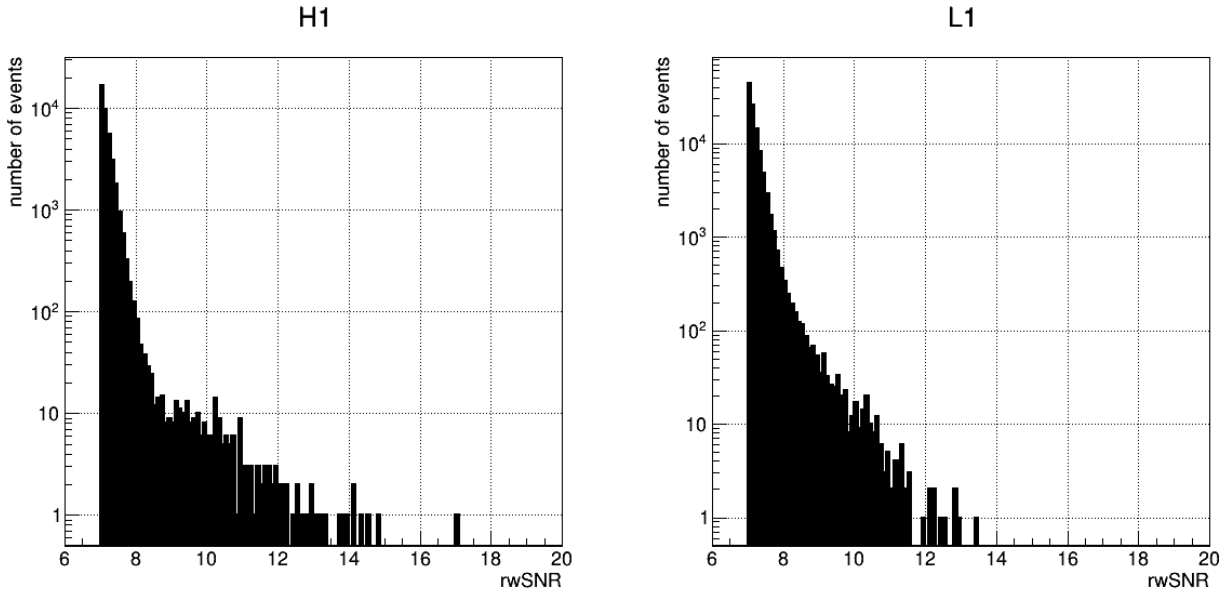


Figure 127: rwSNR distribution of all single detector triggers for the first 46.5 days of O4. Effective times in H1 and L1 are 28.4 days and 38.7 days respectively. Triggers associated to significant public alerts were removed.

hasRemnant $> 0.1\%$. This was because we did not compute hasRemnant during O3. For O4, the hasRemnant quantity is computed by MBTA. The EM bright population is therefore taken as hasRemnant $> 0.1\%$. This definition includes 87% of the templates of the O4 bank.

Figure 129 shows the rwSNR vs template duration for EM bright single detector triggers. We see that the noisy templates seen in figure 128 are not part of the EM bright population.

Looking at figure 130, which shows the rwSNR distribution of EM bright single detector triggers before and after the selection criteria defined in chapter 5. We see that the selection criteria reduce the overall number of triggers, especially in L1 where many high rwSNR trigger are rejected. With the selection criteria applied, the distribution of EM bright triggers follows an exponential shape with a few outliers in L1. Most of them have a very asymmetric template in terms of masses. Adding an extra selection on the mass ratio $m_1/m_2 < 30$ allows to remove these triggers as shown in figure 130. Furthermore, some changes to the gating that occured after the detection of these triggers allows to remove two of them when analyzing the data offline with the latest MBTA configuration.

Another thing to be checked is the criteria themselves. They were defined by looking at O3 background, which is different from O4 background. In addition the pipeline went through some changes between the runs, with in particular a new template bank and modification to the gating. The criteria were defined to be conservative and avoid sending false alerts at the begining of the run. They may be improved to be more performant and/or cause less duty cycle loss. Although we do not have a lot of data yet we can still have a first look.

Figures 131 and 132 show the rwSNR for EM bright single detector triggers with several cuts on the excess rate and gating respectively. As it can be seen the effect of the selection on the gating is marginal on this period of time. The selection on the excess rate on the other hand allowed to reject some triggers. We see that the time window in which we look for excess rate has a limited impact. The same can be said about EM dark single detector triggers, as shown by figures 133 and

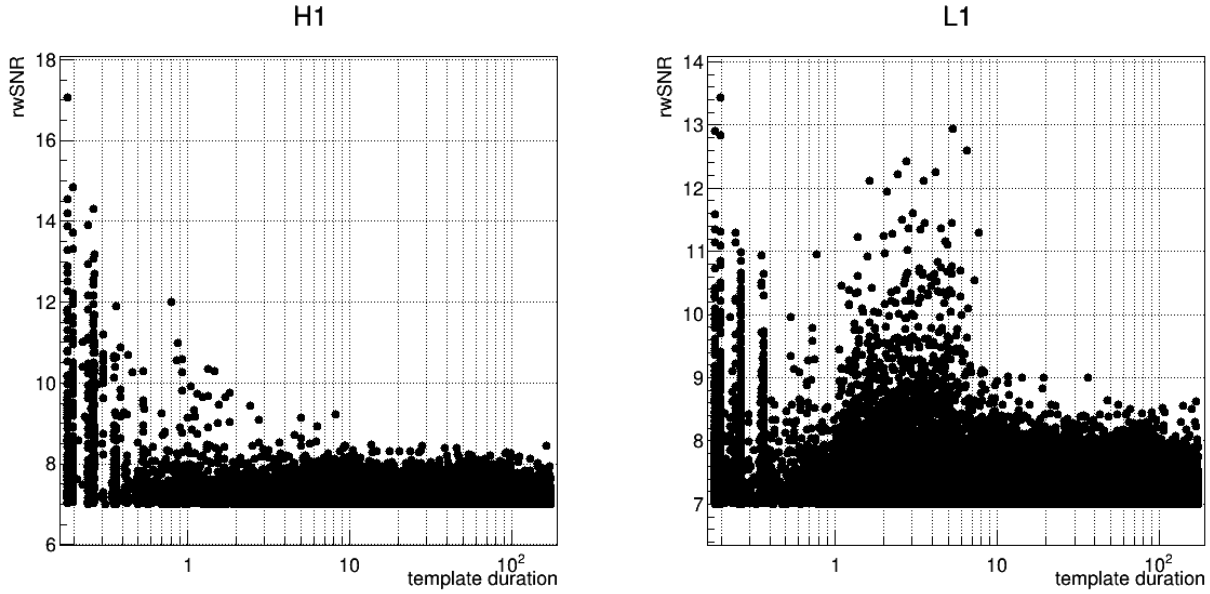


Figure 128: rwSNR vs template duration distribution of all single detector triggers for the first 46.5 days of O4. Triggers associated to significant public alerts were removed.

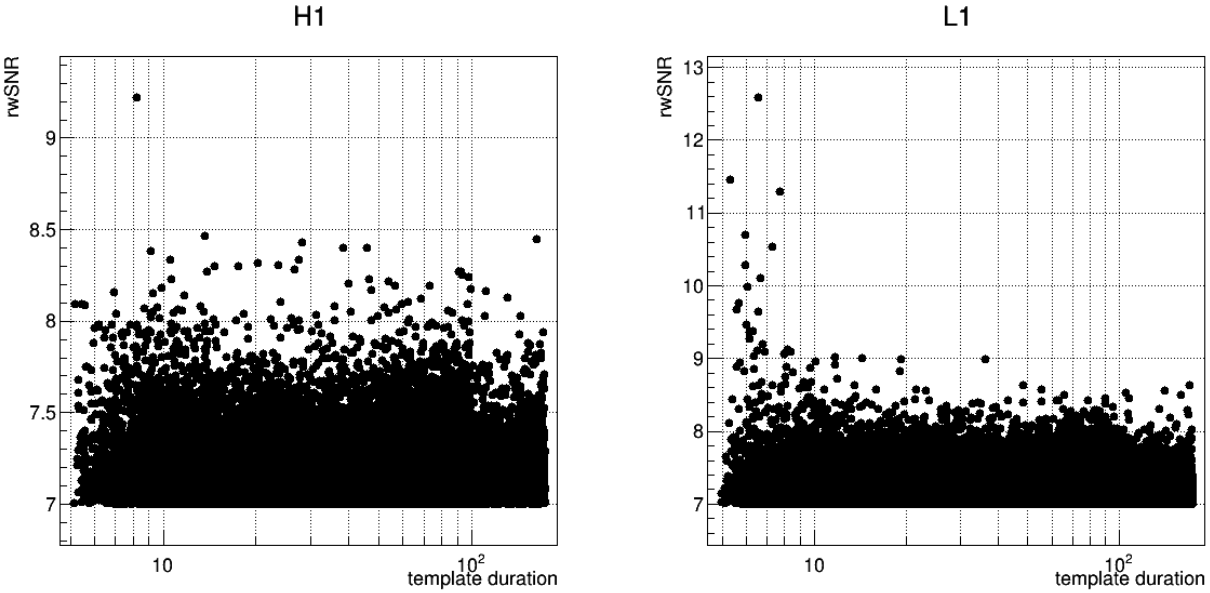


Figure 129: rwSNR vs template duration distribution of EM bright single detector triggers for the first 46.5 days of O4. Triggers associated to significant public alerts were removed.

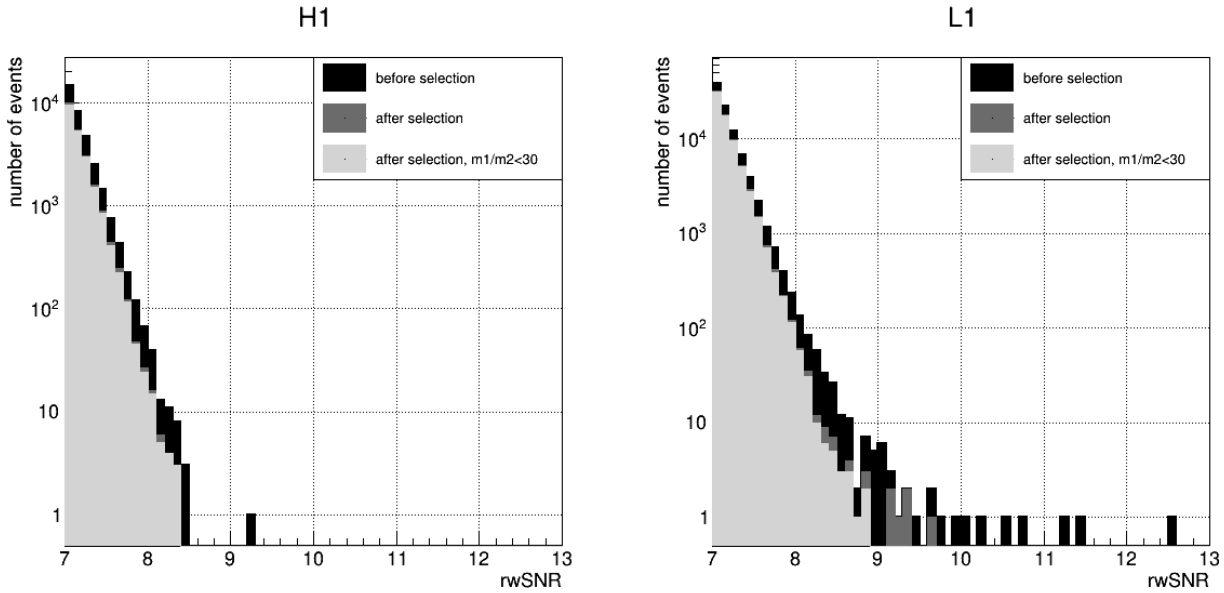


Figure 130: rwSNR distribution of EM bright single detector triggers for the first 46.5 days of O4 (28.4 effective days in H1 and 38.7 effective days in L1). Black is before applying the selection criteria defined in chapter 5. Red is after applying the criteria. Adding an extra selection on the mass ratio $m_1/m_2 < 30$ yields the green distribution. Triggers associated to significant public alerts were removed.

Excess rate cut	H1	L1	Gating	H1	L1
$[-3;+3]$s	89.8%	96.6%	$[-5;+5]$s	99.2%	99.4%
$[-5;+5]$s	88.0%	95.9%	$[-10;+5]$s	98.8%	99.1%
$[-7;+7]$s	**86.3%**	**95.3%**	$[-15;+5]$s	98.4%	98.8%
$[-10;+10]$s	83.8%	94.4%	**$[-20;+5]$s**	**98.0%**	**98.5%**

Table 7: Duty cycle left after applying a selection on the gating or the excess rate. For the analysis on O3 data we used $[-7;+7]$s for the selection on the excess rate and $[-20;+5]$s for the selection on the gating.

134. The remaining duty cycle after applying the selection criteria on the ER or the gating are given in table 7. This gives hope that we may be less stringent on the selection criteria for O4. But we have to keep in mind that we have less than two months of data here. Furthemore, the gating parameters have been changed after several weeks of operation, to better catch the low frequency noise fluctuation. This reduces a bit the high SNR triggers, and could also change the distributions shown in this section when running offline with the new gating.

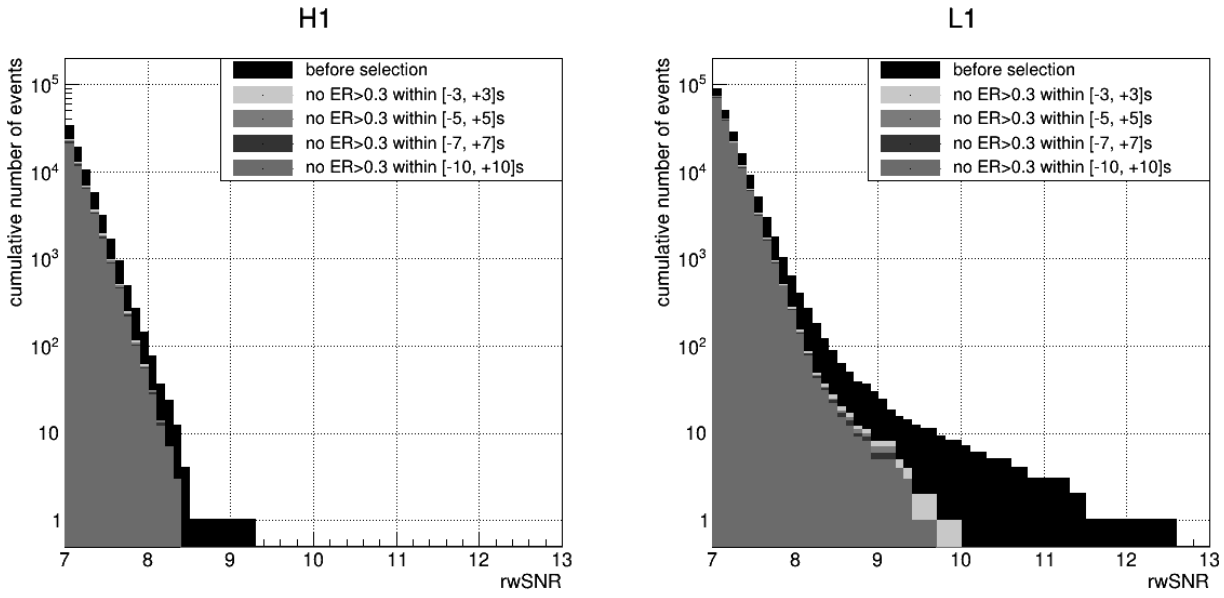


Figure 131: Cumulative rwSNR distribution of EM bright single detector triggers for the first 46.5 days of O4 (28.4 effective days in H1 and 38.7 effective days in L1) for several cuts on the excess rate. No selection on the gating is applied here. Triggers associated to significant public alerts were removed.

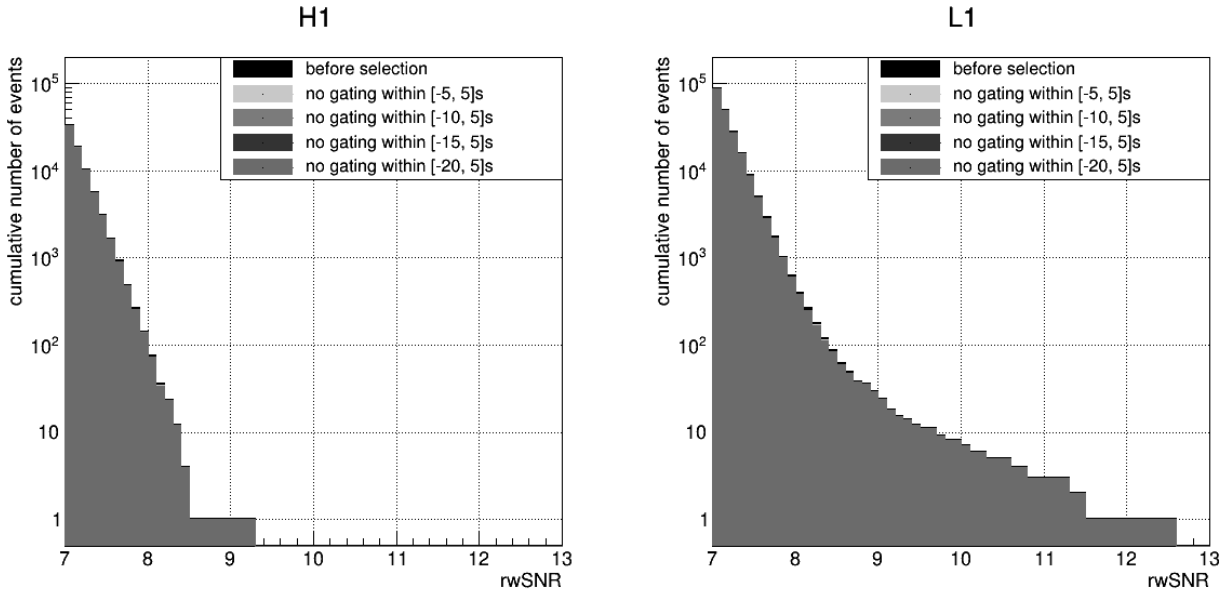


Figure 132: Cumulative rwSNR distribution of EM bright single detector triggers for the first 46.5 days of O4 (28.4 effective days in H1 and 38.7 effective days in L1) for several cuts on the gating. No selection on the ER is applied here. Triggers associated to significant public alerts were removed.

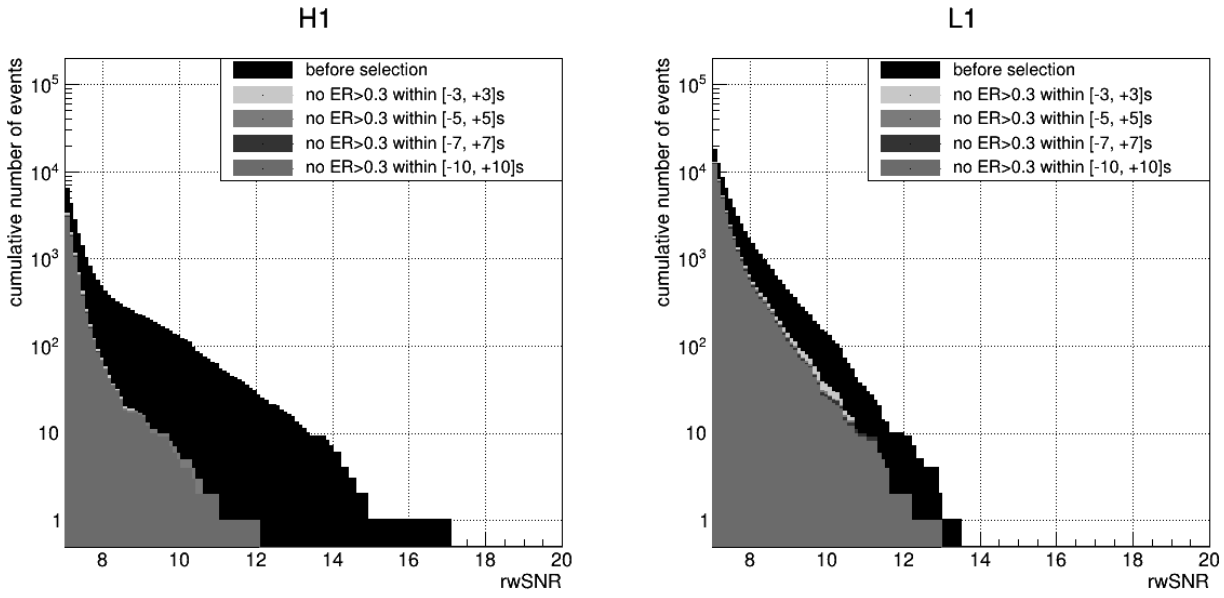


Figure 133: Cumulative rwSNR distribution of long EM dark single detector triggers for the first 46.5 days of O4 (28.4 effective days in H1 and 38.7 effective days in L1) for several cuts on the excess rate. No selection on the gating is applied here. Triggers associated to significant public alerts were removed.

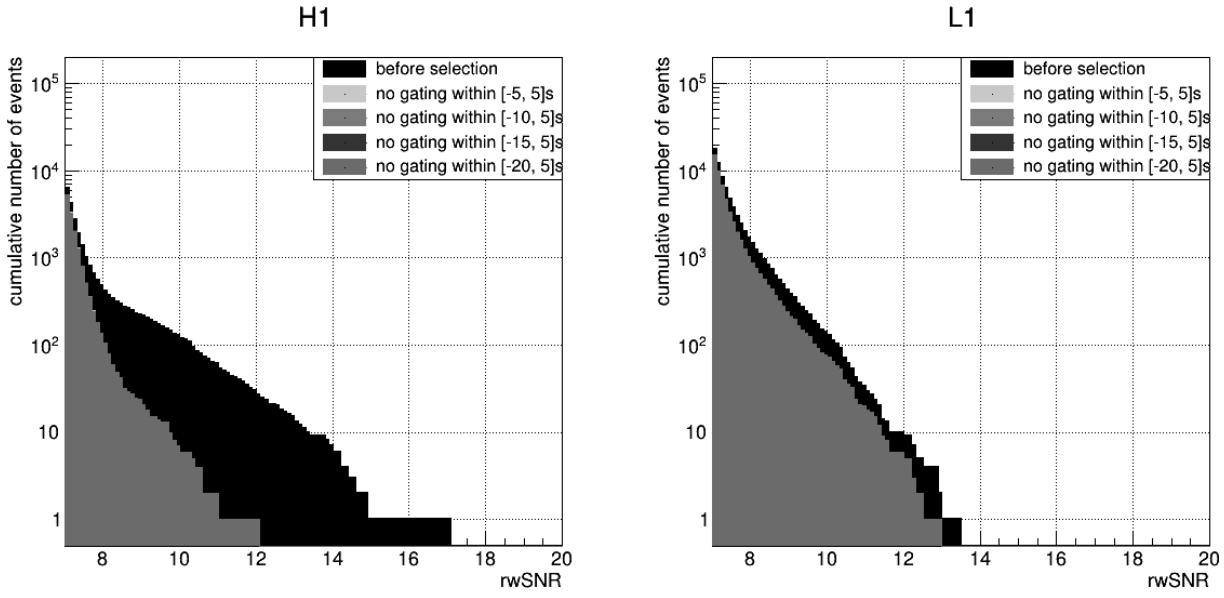


Figure 134: Cumulative rwSNR distribution of long EM dark single detector triggers for the first 46.5 days of O4 (28.4 effective days in H1 and 38.7 effective days in L1) for several cuts on the gating. No selection on the ER is applied here. Triggers associated to significant public alerts were removed.

9.2 Background estimation for EM bright single detector triggers

We now want to check that the method to compute background in order to estimate the FAR of candidates works satisfyingly. We use for the estimation of the background ~ 2.6 days of effective observing time in H1 and ~ 3.0 in L1, from GPS = 1370250000 to GPS = 1370550000. Figure 135 shows the background computed using trigger-random noise coincidences and trigger-trigger coincidences. We see that on this period of time, the "lower bound" computed from trigger-random noise coincidences is actually higher than the observed single detector triggers. It is not clear at this point whether this is simply a statistical effect due to the short duration used for the computation, or if it is a systematic effect. This will require more investigation. On the other side, the "upper bound", computed from trigger-trigger coincidences, behaves as expected from the work on O3 data. The scaling factor in this is 6.3e-05 in H1 and 8.6e-05 in L1. Both methods yield a background distribution of exponential shape with similar slope, we can extrapolate them.

Figure 136 shows the cumulative rwSNR distribution for the observed single detector triggers and the upper bound for the estimation of background. The computed background is fitted with an exponential function $\mathrm{FAR}_{8.6} \times \exp(\mathrm{slope}(\mathrm{rwSNR} - 8.6))$. The fit parameters are given in table 8. We can also parametrize the rwSNR distribution of EM bright single detector triggers for the 46.5 days we showed previously. The fitted function is not shown in figure 136 for readability but the parameters of the fit are given in table 8. Although over-estimated, it allows to reach very low values for high rwSNR triggers.

Computed background (upper bound), fit from 8.5 to 9.5

Detector	$\mathrm{FAR}_{8.6}$ (days^{-1})	slope	IFAR_8 (days)	IFAR_9 (years)	IFAR_{10} (centuries)
H1	0.320 ± 0.004	-7.32 ± 0.09	0.039 ± 0.002	0.160 ± 0.006	2.4 ± 0.3
L1	0.456 ± 0.004	-6.56 ± 0.07	0.043 ± 0.002	0.083 ± 0.002	0.59 ± 0.06

Observed EM bright singles (46.5 days), fit from 7.5 to 8.5

Detector	$\mathrm{FAR}_{8.6}$ (days^{-1})	slope	IFAR_8 (days)	IFAR_9 (years)	IFAR_{10} (centuries)
H1	0.24 ± 0.02	-7.01 ± 0.08	0.062 ± 0.002	0.19 ± 0.02	2.1 ± 0.4
L1	0.33 ± 0.02	-6.29 ± 0.07	0.071 ± 0.002	0.10 ± 0.01	0.6 ± 0.1

Table 8: Results of the fit of the computed upper bound for the background and observed background over 46.5 days using $\mathrm{FAR}_{8.6} \times \exp(\mathrm{slope}(\mathrm{rwSNR} - 8.6))$.

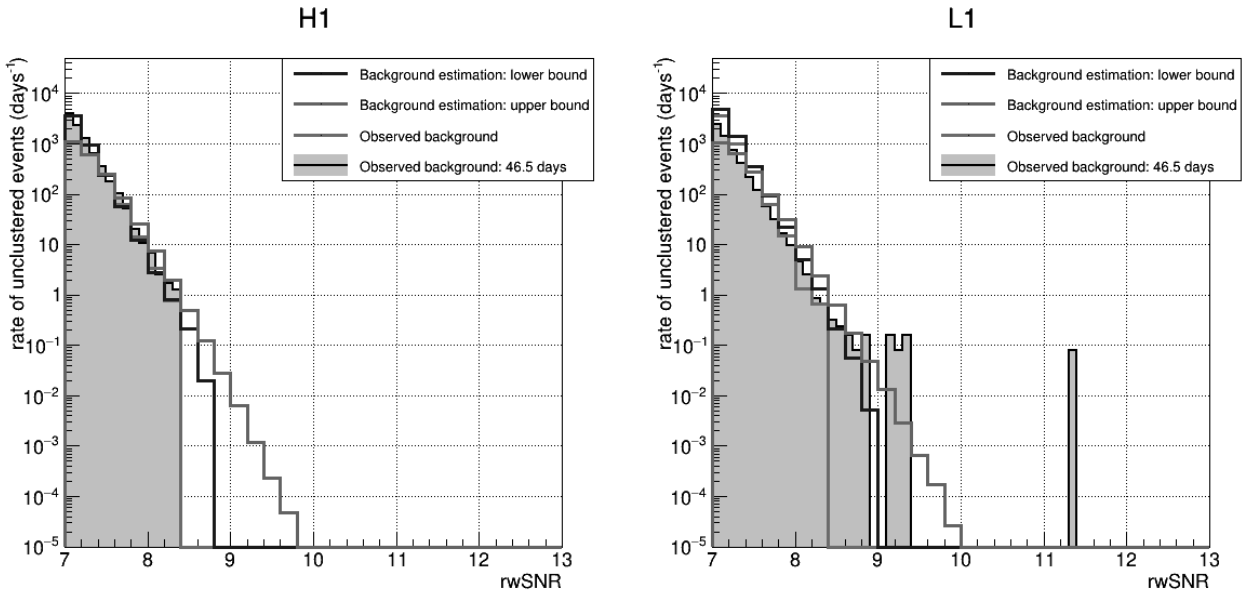


Figure 135: Comparison of the computed lower bound, upper bound and observed background for H1 (left) and L1 (right). Computation was done over ~ 2.6 and ~ 3.0 days for H1 and L1 respectively. All background are given for the EM bright population and the selection criteria are applied.

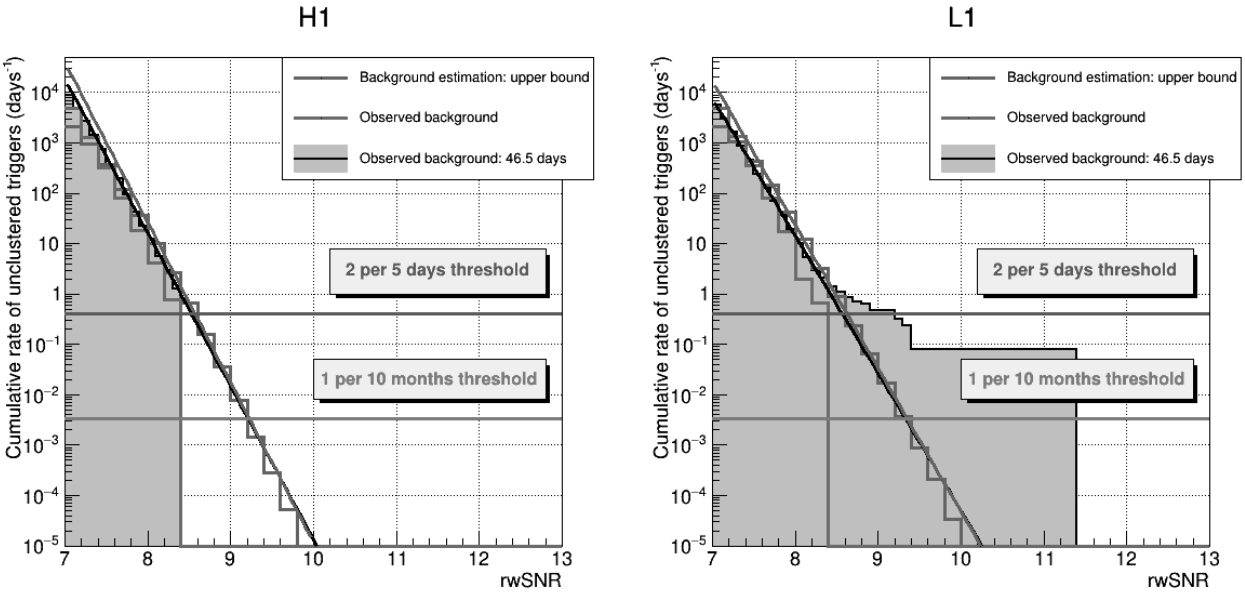


Figure 136: Comparison of the cumulative distributions of the observed background and computed upper bound for the background in H1 (left, for ~ 2.6 days of data) and L1 (right, for ~ 3.0 days of data) fitted with an exponential function. All background are given for the EM bright population and the selection criteria are applied.

10 Conclusion

The primary goal of this PhD work was to develop an analysis of the single detector triggers for the MBTA pipeline. There were two main challenges to face. The first one, discussed in chapter 5, was to reduce the background in order to have the best possible separation between astrophysical signals and background in terms of ranking statistics. The second one, detailed in chapter 6 was to estimate the significance of the candidates through the computation of a false alarm rate.

To reduce the background, the first decision was to focus on a smaller region of the parameter space. This region, called EM bright and defined as hasRemnant $> 0.1\%$, is associated to longer templates and encompasses the triggers that have a chance to have an electromagnetic counterpart. The background associated to this EM bright population was already much lower, both in terms of number of triggers and rwSNR values. In order to prepare future studies, we still considered the EM dark population (triggers complementary to the EM bright population). This population is associated to much shorter templates. It was shown that some short templates produce more triggers with sometimes high ranking statistics. We then used noise rejection tools to define selection criteria for the single detector triggers. These criteria allow to significantly reduce the background and have a clearer separation between astrophysical signals and noise triggers. The losses of duty cycle due to the criteria are 9.1% in H1, 15.4% in L1 and 9.8% in V1. They are largely compensated by the gain in volume for the search.

Once the background was under control, we looked into ways to compute the false alarm rate of the candidates. Since the background has an exponential shape after we apply the selection criteria, a straightforward way is to parametrize the rwSNR distribution with an exponential function. However this method is not guaranteed to be robust in case of some bad days with lots of noise. We therefore considered a different approach. Motivated by the method used for the coincidence search, we use MBTA's 2-band structure to build an estimation of the background by making fake coincidences with the signals observed in each band. The computed background also has an exponential shape and therefore validate the extrapolation which could be made on the observed background distribution to reach low false alarm rate values.

Two additional studies were conducted to pave the way for possible future improvements of the pipeline. One was an attempt at improving a signal consistency test to recover more astrophysical signals and make them more significant, described in chapter 7. A known issue was that this signal consistency test tends to downgrade too heavily very loud events. This means that loud astrophysical events can become much less significant after application of the test. This behaviour comes from the discrete nature of the template bank. To mitigate this effect a modified test was defined by including a term dependent on the SNR. It allows to "flatten" the behaviour of the test result as a function of the SNR. This was followed-up with some work to tune the parameters of the reweighting formula used to downgrade the candidates in order to maximize the recovery of injections while not increasing the background. The results were not convincing enough to decide to proceed with the changes for the fourth observing run but it could be used by future studies.

The other work carried out in section 8 focused on understanding the differences between single detector triggers obtained from real data and single detector triggers obtained by analyzing simulated Gaussian noise (strain). We showed that single detector triggers tend to be detected at times where there was a drop in the detector sensitivity, meaning that their SNR is over-estimated.

We then explored the possibility of using the detector range variations to apply a correction to the SNR in order to mitigate the effect. Finally we also highlighted differences between the range computed on real data and the range computed on Gaussian noise.

All of this work allowed to start safely the single detector triggers search at the begining of O4. There is still room for improvements in all aspects to make the search even more sensitive during O4 as we gradually increase our knowledge of the background.

References

[1] LIGO Scientific Collaboration and Virgo Collaboration. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116 (6 Feb. 2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102.

[2] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "Search for gravitational-lensing signatures in the full third observing run of the LIGO-Virgo network". In: (2023).

[3] Oleg Bulashenko and Helena Ubach. "Lensing of gravitational waves: universal signatures in the beating pattern". In: *Journal of Cosmology and Astroparticle Physics* 2022.07 (July 2022), p. 022. DOI: 10.1088/1475-7516/2022/07/022.

[4] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "All-sky search for continuous gravitational waves from isolated neutron stars using Advanced LIGO and Advanced Virgo O3 data". In: *Physical Review D* 106.10 (Nov. 2022). DOI: 10.1103/physrevd.106.102008.

[5] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "All-sky search for short gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run". In: *Physical Review D* 104.12 (Dec. 2021). DOI: 10.1103/physrevd.104.122004.

[6] Nelson Christensen. "Stochastic gravitational wave backgrounds". In: *Reports on Progress in Physics* 82.1 (Nov. 2018), p. 016903. DOI: 10.1088/1361-6633/aae6b5.

[7] S. Hanany et al. "MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10 arcminutes to 5 degrees". In: *The Astrophysical Journal* 545.1 (Dec. 2000), pp. L5–L9. DOI: 10.1086/317322.

[8] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "Search for anisotropic gravitational-wave backgrounds using data from Advanced LIGO and Advanced Virgo's first three observing runs". In: *Physical Review D* 104.2 (July 2021). DOI: 10.1103/physrevd.104.022005.

[9] "Search for the isotropic stochastic background using data from Advanced LIGO's second observing run". In: *Physical Review D* 100.6 (Sept. 2019). DOI: 10.1103/physrevd.100.061101.

[10] The LIGO Scientific Collaboration and the Virgo Collaboration. "Search for Tensor, Vector, and Scalar Polarizations in the Stochastic Gravitational-Wave Background". In: *Physical Review Letters* 120.20 (May 2018). DOI: 10.1103/physrevlett.120.201102.

[11] Clifford M. Will. "The Confrontation between General Relativity and Experiment". In: *Living Reviews in Relativity* 17.1 (June 2014). DOI: 10.12942/lrr-2014-4.

[12] G. Fanizza et al. "Comparing the luminosity distance for gravitational waves and electromagnetic signals in a simple model of quadratic gravity". In: *General Relativity and Gravitation* 52.11 (Nov. 2020). DOI: 10.1007/s10714-020-02760-5.

[13] Bruce Allen et al. "FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries". In: *Physical Review D* 85.12 (June 2012). DOI: 10.1103/physrevd.85.122006.

[14] Alexandre Le Tiec and Jerome Novak. "Theory of Gravitational Waves". In: *An Overview of Gravitational Waves: Theory, Sources and Detection* (July 2016). DOI: 10.1142/10082.

[15] Péter Mészáros et al. “Multi-messenger astrophysics”. In: *Nature Reviews Physics* 1.10 (Oct. 2019), pp. 585–599. DOI: 10.1038/s42254-019-0101-z.

[16] The IceCube Collaboration et al. “Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A”. In: *Science* 361.6398 (July 2018). DOI: 10.1126/science.aat1378.

[17] IceCube Collaboration. “Evidence for neutrino emission from the nearby active galaxy NGC 1068”. In: *Science* 378.6619 (Nov. 2022), pp. 538–543. DOI: 10.1126/science.abg3395.

[18] The ANTARES collaboration. “Search for neutrino counterparts to the gravitational wave sources from LIGO/Virgo O3 run with the ANTARES detector”. In: *Journal of Cosmology and Astroparticle Physics* 2023.04 (Apr. 2023), p. 004. DOI: 10.1088/1475-7516/2023/04/004.

[19] IceCube Collaboration. *A Search for IceCube sub-TeV Neutrinos Correlated with Gravitational-Wave Events Detected By LIGO/Virgo*. 2023.

[20] B. P. Abbott et al. “Multi-messenger Observations of a Binary Neutron Star Merger”. In: *The Astrophysical Journal* 848.2 (Oct. 2017), p. L12. DOI: 10.3847/2041-8213/aa91c9.

[21] B. P. Abbott et al. “Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A”. In: *The Astrophysical Journal Letters* 848.2 (Oct. 2017), p. L13. DOI: 10.3847/2041-8213/aa920c.

[22] Clifford M. Will. “Bounding the mass of the graviton using gravitational-wave observations of inspiralling compact binaries”. In: *Physical Review D* 57.4 (Feb. 1998), pp. 2061–2068. DOI: 10.1103/physrevd.57.2061.

[23] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. *Tests of General Relativity with GWTC-3*. 2021.

[24] The LIGO Scientific Collaboration and the Virgo Collaboration. “First Search for Nontensorial Gravitational Waves from Known Pulsars”. In: *Physical Review Letters* 120.3 (Jan. 2018). DOI: 10.1103/physrevlett.120.031104.

[25] Hiroki Takeda et al. “Prospects for gravitational-wave polarization tests from compact binary mergers with future ground-based detectors”. In: *Physical Review D* 100.4 (Aug. 2019). DOI: 10.1103/physrevd.100.042001.

[26] Claudia de Rham. “Massive Gravity”. In: *Living Reviews in Relativity* 17.1 (Aug. 2014). DOI: 10.12942/lrr-2014-7.

[27] Eleonora Di Valentino et al. “In the realm of the Hubble tension—a review of solutions”. In: *Classical and Quantum Gravity* 38.15 (July 2021), p. 153001. DOI: 10.1088/1361-6382/ac086d.

[28] B. P. Abbott et al. “A gravitational-wave standard siren measurement of the Hubble constant”. In: *Nature* 551.7678 (Oct. 2017), pp. 85–88. DOI: 10.1038/nature24471.

[29] The LIGO Scientific Collaboration and the Virgo Collaboration. “A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo”. In: *The Astrophysical Journal* 909.2 (Mar. 2021), p. 218. DOI: 10.3847/1538-4357/abdcb7.

[30] Adam G. Riess et al. “Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with LambdaCDM”. In: *The Astrophysical Journal Letters* 908.1 (Feb. 2021), p. L6. DOI: 10.3847/2041-8213/abdbaf.

[31] Planck Collaboration. "Planck 2018 results. VI. Cosmological parameters". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. DOI: 10.1051/0004-6361/201833910.

[32] F. Douchin and P. Haensel. "A unified equation of state of dense matter and neutron star structure". In: *Astronomy & Astrophysics* 380.1 (Dec. 2001), pp. 151–167. DOI: 10.1051/0004-6361:20011402.

[33] J. M. Lattimer and M. Prakash. "Neutron Star Structure and the Equation of State". In: *The Astrophysical Journal* 550.1 (Mar. 2001), p. 426. DOI: 10.1086/319702.

[34] The LIGO Scientific Collaboration and the Virgo Collaboration. "GW170817: Measurements of Neutron Star Radii and Equation of State". In: *Physical Review Letters* 121.16 (Oct. 2018). DOI: 10.1103/physrevlett.121.161101.

[35] A. W. Steiner et al. "Using neutron star observations to determine crust thicknesses, moments of inertia, and tidal deformabilities". In: *Physical Review C* 91.1 (Jan. 2015). DOI: 10.1103/physrevc.91.015804.

[36] J.M. Lattimer. "Neutron Stars and the Nuclear Matter Equation of State". In: *Annual Review of Nuclear and Particle Science* 71.1 (2021), pp. 433–464. DOI: 10.1146/annurev-nucl-102419-124827.

[37] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*. 2021.

[38] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. *The population of merging compact binaries inferred using gravitational waves through GWTC-3*. 2022.

[39] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "Search for Subsolar-Mass Binaries in the First Half of Advanced LIGO's and Advanced Virgo's Third Observing Run". In: *Physical Review Letters* 129.6 (Aug. 2022). DOI: 10.1103/physrevlett.129.061104.

[40] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "Search for subsolar-mass black hole binaries in the second part of Advanced LIGO's and Advanced Virgo's third observing run". In: (2022).

[41] B. J. Carr. "The primordial black hole mass spectrum." In: *apj* 201 (Oct. 1975), pp. 1–19. DOI: 10.1086/153853.

[42] Guido D'Amico et al. "Massive black holes from dissipative dark matter". In: *Monthly Notices of the Royal Astronomical Society* 473.1 (Sept. 2017), pp. 328–335. DOI: 10.1093/mnras/stx2419.

[43] Sarah Shandera, Donghui Jeong, and Henry S. Grasshorn Gebhardt. "Gravitational Waves from Binary Mergers of Subsolar Mass Dark Black Holes". In: *Physical Review Letters* 120.24 (June 2018). DOI: 10.1103/physrevlett.120.241102.

[44] Jeremie Choquette, James M. Cline, and Jonathan M. Cornell. "Early formation of supermassive black holes via dark matter self-interactions". In: *Journal of Cosmology and Astroparticle Physics* 2019.07 (July 2019), pp. 036–036. DOI: 10.1088/1475-7516/2019/07/036.

[45] Vitor Cardoso et al. "Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale". In: *Physical Review D* 94.8 (Oct. 2016). DOI: 10.1103/physrevd.94.084031.

[46] Alexandre Toubiana et al. “Modeling gravitational waves from exotic compact objects”. In: *Physical Review D* 103.6 (Mar. 2021). DOI: 10.1103/physrevd.103.064042.

[47] F. Acernese et al. “Advanced Virgo: a second-generation interferometric gravitational wave detector”. In: *Classical and Quantum Gravity* 32.2 (Dec. 2014), p. 024001. DOI: 10.1088/0264-9381/32/2/024001.

[48] The Virgo Collaboration. *Advanced Virgo Technical Design Report*. Tech. rep. VIR-0128A-12. Apr. 2012. URL: https://tds.virgo-gw.eu/?call_file=VIR-0128A-12.pdf.

[49] The Virgo Collaboration. *Advanced Virgo Plus Phase II Design Report*. Tech. rep. VIR-0501C-22. Sept. 2022. URL: https://tds.virgo-gw.eu/?call_file=VIR-0501C-22_AdvancedVirgoPlusPhaseIIDesign.pdf.

[50] Brian J. Meers. “Recycling in laser-interferometric gravitational-wave detectors”. In: *Phys. Rev. D* 38 (8 Oct. 1988), pp. 2317–2326. DOI: 10.1103/PhysRevD.38.2317.

[51] K. A. Strain and B. J. Meers. “Experimental demonstration of dual recycling for interferometric gravitational-wave detectors”. In: *Phys. Rev. Lett.* 66 (11 Mar. 1991), pp. 1391–1394. DOI: 10.1103/PhysRevLett.66.1391.

[52] Moritz Mehmet and Henning Vahlbruch. “The Squeezed Light Source for the Advanced Virgo Detector in the Observation Run O3”. In: *Galaxies* 8.4 (2020). ISSN: 2075-4434. DOI: 10.3390/galaxies8040079.

[53] Allocca et al. “Interferometer Sensing and Control for the Advanced Virgo Experiment in the O3 Scientific Run”. In: *Galaxies* 8.4 (2020). ISSN: 2075-4434. DOI: 10.3390/galaxies8040085.

[54] D Estevez et al. “The Advanced Virgo photon calibrators”. In: *Classical and Quantum Gravity* 38.7 (Feb. 2021), p. 075007. DOI: 10.1088/1361-6382/abe2db.

[55] D Estevez et al. “First tests of a Newtonian calibrator on an interferometric gravitational wave detector”. In: *Classical and Quantum Gravity* 35.23 (Nov. 2018), p. 235009. DOI: 10.1088/1361-6382/aae95f.

[56] D Estevez, B Mours, and T Pradier. “Newtonian calibrator tests during the Virgo O3 data taking”. In: *Classical and Quantum Gravity* 38.7 (Mar. 2021), p. 075012. DOI: 10.1088/1361-6382/abe2da.

[57] The Virgo Collaboration. “Calibration of advanced Virgo and reconstruction of the detector strain h(t) during the observing run O3”. In: *Classical and Quantum Gravity* 39.4 (Jan. 2022), p. 045006. DOI: 10.1088/1361-6382/ac3c8e.

[58] Warren G. Anderson et al. “Excess power statistic for detection of burst sources of gravitational radiation”. In: *Physical Review D* 63.4 (Jan. 2001). DOI: 10.1103/physrevd.63.042003.

[59] Hsin-Yu Chen et al. “Distance measures in gravitational-wave astrophysics and cosmology”. In: *Classical and Quantum Gravity* 38.5 (Jan. 2021), p. 055010. DOI: 10.1088/1361-6382/abd594.

[60] Lee Samuel Finn and David F. Chernoff. “Observing binary inspiral in gravitational radiation: One interferometer”. In: *Physical Review D* 47.6 (Mar. 1993), pp. 2198–2219. DOI: 10.1103/physrevd.47.2198.

[61] A. Buikema et al. “Sensitivity and performance of the Advanced LIGO detectors in the third observing run”. In: *Physical Review D* 102.6 (Sept. 2020). DOI: 10.1103/physrevd.102.062003.

[62] F. Acernese et al. *Virgo Detector Characterization and Data Quality during the O3 run*. 2022.

[63] D Davis et al. "LIGO detector characterization in the second and third observing runs". In: *Classical and Quantum Gravity* 38.13 (June 2021), p. 135014. DOI: 10.1088/1361-6382/abfd85.

[64] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA". In: *Living Reviews in Relativity* 23.1 (Sept. 2020). DOI: 10.1007/s41114-020-00026-9.

[65] The LIGO Scientific Collaboration and the Virgo Collaboration. "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". In: *Physical Review Letters* 119.16 (Oct. 2017). DOI: 10.1103/physrevlett.119.161101.

[66] M Punturo et al. "The Einstein Telescope: a third-generation gravitational wave observatory". In: *Classical and Quantum Gravity* 27.19 (Sept. 2010), p. 194002. DOI: 10.1088/0264-9381/27/19/194002.

[67] Pau Amaro-Seoane et al. *Laser Interferometer Space Antenna*. 2017.

[68] G Hobbs et al. "The International Pulsar Timing Array project: using pulsars as a gravitational wave detector". In: *Classical and Quantum Gravity* 27.8 (Apr. 2010), p. 084013. DOI: 10.1088/0264-9381/27/8/084013.

[69] The NANOGrav Collaboration. "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background". In: *The Astrophysical Journal Letters* 951.1 (June 2023), p. L8. DOI: 10.3847/2041-8213/acdac6.

[70] F. Aubin et al. "The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run". In: *Classical and Quantum Gravity* 38.9 (Apr. 2021), p. 095004. ISSN: 1361-6382. DOI: 10.1088/1361-6382/abe913.

[71] The Virgo DetChar group. *"Category 1" (CAT1) data quality vetoes applied to the analysis of the O3 run Virgo data*. URL: https://tds.virgo-gw.eu/?content=3&r=18844.

[72] Derek Davis et al. *LIGO Data Quality Vetoes Applied to the Analysis of O3*. URL: https://dcc.ligo.org/LIGO-T2100045/public.

[73] Florian Aubin. "Recherche à faible latence d'ondes gravitationnelles émises lors de coalescences de binaires compactes durant la troisième période d'observation de Advanced Virgo et Advanced LIGO". 2020CHAMA024. PhD thesis. 2020. URL: http://www.theses.fr/2020CHAMA024/document.

[74] Carl W. Helstrom. *Statistical Theory of Signal Detection*. 1968.

[75] L. A. Wainstein and V. D. Zubakov. *Extraction of Signals from Noise*. 1970.

[76] Matteo Frigo and Steven G. Johnson. "The Design and Implementation of FFTW3". In: *Proceedings of the IEEE* 93.2 (2005). Special issue on "Program Generation, Optimization, and Platform Adaptation", pp. 216–231.

[77] PyCBC team. *Calculating the Effectualness (Fitting Factor) of Template Banks*. URL: https://pycbc.org/pycbc/latest/html/banksim.html.

[78] Chris Van Den Broeck et al. "Template banks to search for compact binaries with spinning components in gravitational wave data". In: *Physical Review D* 80.2 (July 2009). DOI: 10.1103/physrevd.80.024009.

[79] Francesco Messina et al. "Quasi-5.5PN TaylorF2 approximant for compact binaries: Point-mass phasing and impact on the tidal polarizability inference". In: *Physical Review D* 99.12 (June 2019). DOI: 10.1103/physrevd.99.124051.

[80] R Sturani et al. "Complete phenomenological gravitational waveforms from spinning coalescing binaries". In: *Journal of Physics: Conference Series* 243 (Aug. 2010), p. 012007. DOI: 10.1088/1742-6596/243/1/012007.

[81] Alejandro Bohé et al. "Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors". In: *Physical Review D* 95.4 (Feb. 2017). DOI: 10.1103/physrevd.95.044028.

[82] S Babak et al. "A template bank to search for gravitational waves from inspiralling compact binaries: I. Physical models". In: *Classical and Quantum Gravity* 23.18 (Aug. 2006), pp. 5477–5504. DOI: 10.1088/0264-9381/23/18/002.

[83] T. Cokelaer. "Gravitational waves from inspiralling compact binaries: Hexagonal template placement and its efficiency in detecting physical signals". In: *Physical Review D* 76.10 (Nov. 2007). DOI: 10.1103/physrevd.76.102004.

[84] I. W. Harry, B. Allen, and B. S. Sathyaprakash. "Stochastic template placement algorithm for gravitational wave data analysis". In: *Physical Review D* 80.10 (Nov. 2009). DOI: 10.1103/physrevd.80.104014.

[85] Soumen Roy, Anand S. Sengupta, and Nilay Thakor. "Hybrid geometric-random template-placement algorithm for gravitational wave searches from compact binary coalescences". In: *Physical Review D* 95.10 (May 2017). DOI: 10.1103/physrevd.95.104045.

[86] Soumen Roy, Anand S. Sengupta, and Parameswaran Ajith. "Effectual template banks for upcoming compact binary searches in Advanced-LIGO and Virgo data". In: *Physical Review D* 99.2 (Jan. 2019). DOI: 10.1103/physrevd.99.024048.

[87] N Andres et al. "Assessing the compact-binary merger candidates reported by the MBTA pipeline in the LIGO–Virgo O3 run: probability of astrophysical origin, classification, and associated uncertainties". In: *Classical and Quantum Gravity* 39.5 (Feb. 2022), p. 055002. DOI: 10.1088/1361-6382/ac482a.

[88] Francois Foucart, Tanja Hinderer, and Samaya Nissanke. "Remnant baryon mass in neutron star-black hole mergers: Predictions for binary neutron star mimickers and rapidly spinning black holes". In: *Physical Review D* 98.8 (Oct. 2018). DOI: 10.1103/physrevd.98.081501.

[89] Cosima Breu and Luciano Rezzolla. "Maximum mass, moment of inertia and compactness of relativistic stars". In: *Monthly Notices of the Royal Astronomical Society* 459.1 (Mar. 2016), pp. 646–656. ISSN: 0035-8711. DOI: 10.1093/mnras/stw575.

[90] Isaac Legred et al. "Impact of the PSR J0740+6620 radius constraint on the properties of high-density matter". In: *Physical Review D* 104.6 (Sept. 2021). DOI: 10.1103/physrevd.104.063003.

[91] Stefano Ascenzi et al. "Electromagnetic counterparts of compact binary mergers". In: *Journal of Plasma Physics* 87.1 (Feb. 2021). DOI: 10.1017/s0022377820001646.

[92] The LIGO Scientific Collaboration and the Virgo Collaboration. *GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*. 2021.

[93] Vincent Juste. *Improving the MBTA rwSNR for high SNR triggers*. Tech. rep. VIR-0833A-22. Aug. 2022. URL: https://tds.virgo-gw.eu/ql/?c=18479.

UNIVERSITÉ DE STRASBOURG

Vincent Juste

IPHC Institut Pluridisciplinaire Hubert CURIEN STRASBOURG

Search for compact binary coalescence using a single gravitational waves detector with the MBTA data analysis pipeline

Résumé

La quatrième période d'observation des détecteurs LIGO/Virgo a démarré le 24 Mai 2023. Afin de la préparer au mieux, les détecteurs tout comme les chaines d'analyse des données ont évolués en utilisant les connaissances acquises lors de la prise de données précédente. L'une de ces chaînes est MBTA.

Parmis les changements effectués sur MBTA, il a été décidé de développer une recherche des évènements vus dans un seul détecteur alors que seules les coïncidences étaient jusque là publiées. Cette recherche permet d'augmenter le temps effectif d'observation et donc les chances de détections. Ce travail de thèse porte sur le développement de cette recherche.

Les deux principaux objectifs de ce travail de thèse sont la réduction du bruit de fond et l'estimation de la significativité des évènements. Le premier objectif a été remplis en appliquant des critères de sélection basés sur la qualité des données. Le deuxième a été accompli en mettant au point une méthode de calcul d'un taux de fausse alarme pour les évènements vu dans un unique détecteur qui permet de valider l'extrapolation du bruit de fond observé.

Mots-clés : ondes gravitationnelles, CBC, système binaire compact, MBTA, détecteur unique, LIGO, Virgo

Abstract

The fourth observing run of the LIGO/Virgo detector start on May 24 2023. In order to prepare it as best as possible, detectors and analysis pipelines evolved using the knowledge acquired during the previous observing run. One of those pipeline is MBTA.

Among the changes planned for MBTA, it was decided to develop a search for single detector triggers. Until then, only coincidences were published. This search allows to increase the effective observing time and therefore the chances of detections. This PhD work is about the development of this search.

The two main goals of this work are the reduction of the background and the estimation of the significance of events. The first goal was achieved by applying selection criteria based on data quality. The second goal was completed by developping a method to compute a false alarm rate for single detector triggers which allows to validate the extrapolation of the observed background.

Keywords : gravitational waves, CBC, compact binary coalescence, MBTA, single detector triggers, LIGO, Virgo