

5 | CONSTRAINTS ON THE AGN CHANNEL FOR THE FORMATION OF STELLAR-MASS BINARY BLACK HOLES UP TO $z=1.5$ USING THE QUAIA CATALOGUE

Work by **N. Veronesi**, S. van Velzen, E. M. Rossi,
soon to be submitted to *Monthly Notices of the Royal Astronomical Society*

Abstract

Statistical analyses based on the spatial correlation between the sky maps of detected Gravitational Wave (GW) events and the positions of potential host environments are a valid tool to answer the open question regarding the physical origin of the binaries of which have the mergers been detected by the interferometers of the LIGO-Virgo-KAGRA (LVK) collaboration. One plausible formation channel involves the interaction between the merging compact objects and the accretion discs that power Active Galactic Nuclei (AGN). In this paper we test the efficiency of this so-called AGN channel by analysing the spatial correlation between 159 GW events detected not later than June 1st, 2024, and the AGN contained in the all-sky catalogue Quaia. To do so, we estimate the bolometric luminosities of all the potential hosts and the completeness of Quaia up to redshift $z = 1.5$. We find at a 95 per cent credibility level that un-obscured AGN with a bolometric luminosity higher than $10^{44.5} \text{erg s}^{-1}$ ($10^{45} \text{erg s}^{-1}$) do not contribute to more than the 21 (11) per cent of the mergers detected by the LVK collaboration. A significant fraction of the detected mergers might still come from categories of AGN that are not included in our analysis, like faint or obscured ones.

5.1 Introduction

A consistent one-to-one association between the mergers of compact objects (COs) detected through Gravitational Waves (GWs) by the LIGO-Virgo-KAGRA (LVK) collaboration ([Acernese et al. 2015](#); [LIGO Scientific Collaboration et al. 2015](#); [Akutsu et al. 2021](#)) and their host galaxy is currently out of reach.

This is primarily caused by the typical size of the uncertainties that are associated to the sky position and to the luminosity distance of each GW event. The impossibility of associating each merger to its host environment hampers the chances to answer the open question regarding the physical origin of the binary systems of which the coalescences have been directly measured. Among the possible binary formation scenarios that have been proposed (see [Mapelli 2021](#), for a review), there is one according to which binaries of COs like Binary Black Holes (BBHs), binary Neutron Stars (BNSs), and Neutron Star-Black Hole binaries (NSBHs) can be efficiently assembled and driven to merger in dense environments like Globular Clusters ([Rodriguez et al. 2016](#)) or Nuclear Stellar Clusters ([Chattopadhyay et al. 2023](#)). In such hosts binaries can have their semi-major axes shrunk by interactions with single objects. Single-binary encounters are expected to harden binaries, which is crucial for the formation of systems that are able to merge within a Hubble time due to loss of orbital energy, emitted in the form of GWs. The least massive object of the three is expected to be ejected from the location of the single-binary interaction ([Hills & Fullerton 1980](#); [Ziosi et al. 2014](#)), meaning that the remnant hardened binary will be composed of the two heaviest elements that take part in the encounter.

Another reason why dynamically dense environments might be the hosts of the most massive mergers detected by the LVK collaboration is the fact that their escape velocities can be large enough to retain the recoil-kicked remnant of a merger, turning it into the potential component of a subsequent GW event. This "hierarchical merger" scenario might consist of a physical interpretation for the existence of stellar-mass Black Holes (sBHs) in the Pair-Instability Mass Gap ([Gerosa & Berti 2017, 2019](#); [Yang et al. 2019](#); [Barrera & Bartos 2022](#)). This discontinuity is predicted between $\approx 50M_{\odot}$ and $\approx 120M_{\odot}$ in the sBH mass spectrum as a consequence of the complete disruption of the core during the SuperNova event at the end of the life cycle of stars with very high masses, that are therefore expected to leave no remnant ([Heger & Woosley 2002](#); [Belczynski et al. 2016](#)). However, the astrophysical mass distribution of sBHs predicted by the LVK collaboration shows evidence for the presence of objects in the Pair-Instability Mass Gap ([Abbott et al. 2023a](#)). This fact challenges the hypothesis

according to which all the merging systems detected by the LVK collaboration evolved from an isolated stellar binary.

Accretion discs of AGN are a unique type of potential host environment for the assembly and the merger of binaries of COs. This is mainly because in the so-called "AGN formation channel" the binaries and their components are not expected to interact only with other stellar objects, but also with the gas the accretion disc consists of. Thanks to this interaction, the disc might for example capture COs that have orbits moderately inclined with respect to its plane (Ostriker 1983; Fabj et al. 2020; Nasim et al. 2023). These disc captures have the effect of increasing the number density of COs in the disc, where the binary formation can be gas-assisted (Tagawa et al. 2020; DeLaurentiis et al. 2023b; Rowan et al. 2023). Another process which, as far as the formation of merging CO binaries is concerned, is typical of the AGN scenario is migration, the radial motion of COs orbiting around the central SuperMassive Black Hole (SMBH). Inward migration takes place in AGN discs where the net torque exerted by the gas onto the orbiting CO is negative (Paardekooper et al. 2010; Bellovary et al. 2016), and can increase the number density of sBHs and Neutron Stars (NSs) in the inner part of the accretion disc, facilitating the formation of binaries.

One possible way to put constraints on the fractional contribution of the AGN channel to the total merger rate of CO binaries is through the investigation of the spatial correlation between the sky maps of the events detected by the LVK collaboration and the positions of observed AGN. This approach has been first suggested in Bartos et al. (2017a), where it was estimated that 300 GW detections would have been sufficient to be able reject the hypothesis according to which there is no GW-AGN relation in favour of the hypothesis that states that the AGN channel contributes to half of the total merger rate. In Corley et al. (2019) it was later found that the number of needed detections decreases by a factor ≈ 3 if it is assumed that the inclination of the binary angular momentum with respect to the line of sight is known with an uncertainty of 5° . In the simulated GW detections used in both Bartos et al. (2017a) and Corley et al. (2019) the interferometers of LIGO and Virgo have been assumed to be working at Design Sensitivity.

In Veronesi et al. (2022) the analysis has been repeated assuming a realistic distribution of the sizes of the 90 per cent localisation volumes (V90) of the simulated GW detections. This distribution has been created from mock detections simulated using the sensitivity curves that characterised the LIGO and Virgo interferometers during their third observing run (O3). It was found that the amount of data collected during O3 was enough to reject the no-connection hypothesis in favour of the one according to which the AGN channel contributes to

half of the total merger rate only in the case in which rare AGN with a number density lower than 10^{-7}Mpc^{-3} are taken into account.

The first observational constraint on the efficiency of the AGN channel was put using the spatial-correlation approach in Veronesi et al. (2023). The statistical framework of that work consisted in an improved version of the one first suggested for the investigation of the AGN-GW relation in Bartos et al. (2017a), and then used also in Corley et al. (2019) and in Veronesi et al. (2022). In Veronesi et al. (2023) the completeness of the AGN catalogues used during the cross-match with GW sky maps and the exact position of the potential host environments have been taken into account for the first time in this type of analysis. Moreover, in the same work, the value of the fractional contribution of the AGN channel to the total merger rate, f_{AGN} , has been constrained for the first time via Likelihood maximisation. Using catalogues of AGN with a spectroscopic estimate of redshift of $z \leq 0.3$ obtained from MilliQuasv7.7b (Flesch 2021) and the sky maps of the 30 GW events detected in the same redshift range during the first three observing runs of the LVK collaboration, it was found that the fraction of the detected mergers that have originated in an AGN more luminous than $10^{45.5}\text{erg s}^{-1}$ (10^{46}erg s^{-1}) is not expected to be higher than 0.49 (0.17). A similar statistical investigation has been conducted also in (Veronesi et al. 2024). In this case the sky maps of the GW events detected during O3 have been cross-matched spatially and temporally with the 20 unusual AGN flaring activities detected by the Zwicky Transient Facility (ZTF) (Bellm et al. 2019; Graham et al. 2019) that have been identified in Graham et al. (2023) as potential transient ElectroMagnetic counterparts. The estimate of the spatial correlation analysis is that the hypothesis according to which there is no connection between the GW events and unusual flares is the one favoured by data.

In this work we present new observational constraints on f_{AGN} . We use a spatial-correlation-based method similar to the one presented in Veronesi et al. (2023), and apply it to a larger dataset. In particular we use the sky maps of all the GW events detected by the interferometers of the LVK collaboration up to June 1st, 2024, therefore including data coming from the fourth observing run (O4). This brings the total number of used sky maps to 159, which is more than five times the amount of GW events used in Veronesi et al. (2023). These are cross-matched with the AGN of Quiaia (Storey-Fisher et al. 2024). This all-sky catalogue is the result of a re-analysis of the one containing the objects identified as extra-galactic by the Gaia mission (Gaia Collaboration et al. 2023). The remarkable uniformity and completeness up to $z = 1.5$ of Quiaia make it a valid tool for all-sky spatial correlation analyses like the one presented in this

work.

In Section 5.2 we present the main characteristics of the data used in this work, which include the different GW detections and the Quia catalogues. In Section 5.3 we then describe the likelihood function that we maximise in order to obtain our new constraints on f_{AGN} , which are presented in Section 5.4. Finally, in Section 5.5 we draw the main conclusions regarding the AGN channel that can be inferred from the results of our analysis, and discuss about future developments of statistical spatial-correlation-based approaches. We adopt the cosmological parameters of the Cosmic Microwave Background observations by Planck (Planck Collaboration et al. 2016): $H_0 = (67.8 \pm 0.9) \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\text{m}} = 0.308 \pm 0.012$, and $n_{\text{s}} = 0.968 \pm 0.006$.

5.2 Datasets

In this section we first present the main characteristics of the GW events the sky maps of which are used in our analysis. We then describe the main properties of the Quia catalogue and how we estimate its completeness as a function of redshift for different sub-samples of it, characterised by different thresholds of bolometric luminosity, L_{bol} .

5.2.1 GW events

In this work we use the 159 sky maps of mergers of CO binaries detected by the interferometers of the LVK collaboration up to June 1st, 2024. In particular all the confident GW events from the first three observing runs are included, with the only exclusion of GW200308_173609 and of GW200322_091133. These two events are not taken into consideration because they are poorly localised, and the corresponding value of V_{90} cannot be estimated from the currently available posterior samples. Moreover, we use all the events from the first half of O4 (O4a) as well as all the events from its second half (O4b) that have been detected not later than June 1st, 2024. For this currently ongoing observing run, we select the detections that have a probability of being of terrestrial origin smaller than 1 per cent.

The sky maps of the events detected during the first three observing runs are downloaded from the Gravitational Wave Open Science Center (GWOSC) (Abbott et al. 2023a). We use the posterior samples obtained using the IMR-PHENOMXPHM waveform model (Pratten et al. 2021) for all these events but

GW190425_081805, GW191219_163120, GW200105_162426, and GW200115_042309. For these four events we use the MIXED posterior samples.

The sky maps of the 71 events of O4 we use in the analysis presented in this work are downloaded from the Gravitational-Wave Candidate Event Database (GraceDB) ¹, operated by the LIGO Scientific Collaboration. For each event we use the most recent version of its sky map, which is either obtained from the Bilby localisation algorithm (Ashton et al. 2019) or from the Bayestar one (Singer & Price 2016).

Figure 5.1 shows the luminosity distance of all the 159 GW events used in this work as a function of the correspondent value of V90. The median value of the luminosity distance (2012.3 Mpc) and of V90 ($7.67 \cdot 10^7 \text{Mpc}^3$) are indicated by the horizontal dashed line and by the vertical one, respectively.

5.2.2 Quaia AGN catalogue

In order to obtain precise estimates of f_{AGN} with methods based on spatial correlation like the one used in this work, is important to use an AGN catalogue that has the highest possible value of completeness. This property of a catalogue is in general a function of the redshifts, the sky positions, and the luminosities of the objects that are considered.

In the analyses here presented, the Quaia AGN catalogue (Storey-Fisher et al. 2024) is cross-matched with the sky maps of the GW events described in Section 5.2.1. This is done once the completeness of such a catalogue has been estimated. Quaia is an all-sky catalogue that contains objects identified as extra-galactic in the third data release of the Gaia mission (Gaia Collaboration et al. 2023). The final version of this catalogue is obtained by selecting the objects that have an infrared counterpart in the unWISE catalogue (Lang 2014; Meisner et al. 2019) and by performing cuts in colours and proper motion in order to decrease the amount of contaminants (see Section 3.1 of Storey-Fisher et al. 2024, for details). The result is a catalogue containing 1,295,502 objects with a magnitude in the Gaia G-band $\text{mag}_G < 20.5$. The redshifts of these objects are estimated using a k -Nearest Neighbors model trained on the AGN present in Quaia of which the spectra are in the 16th Data Release (DR16Q) of the Sloan Digital Sky Survey (SDSS) (Lyke et al. 2020).

Figure 5.2 shows the mollweide projection of the sky distribution of the AGN contained in the Quaia catalogue. The colour of each pixel depends on the amount of objects in the correspondent position in the sky. The size of each

¹<https://gracedb.ligo.org/>

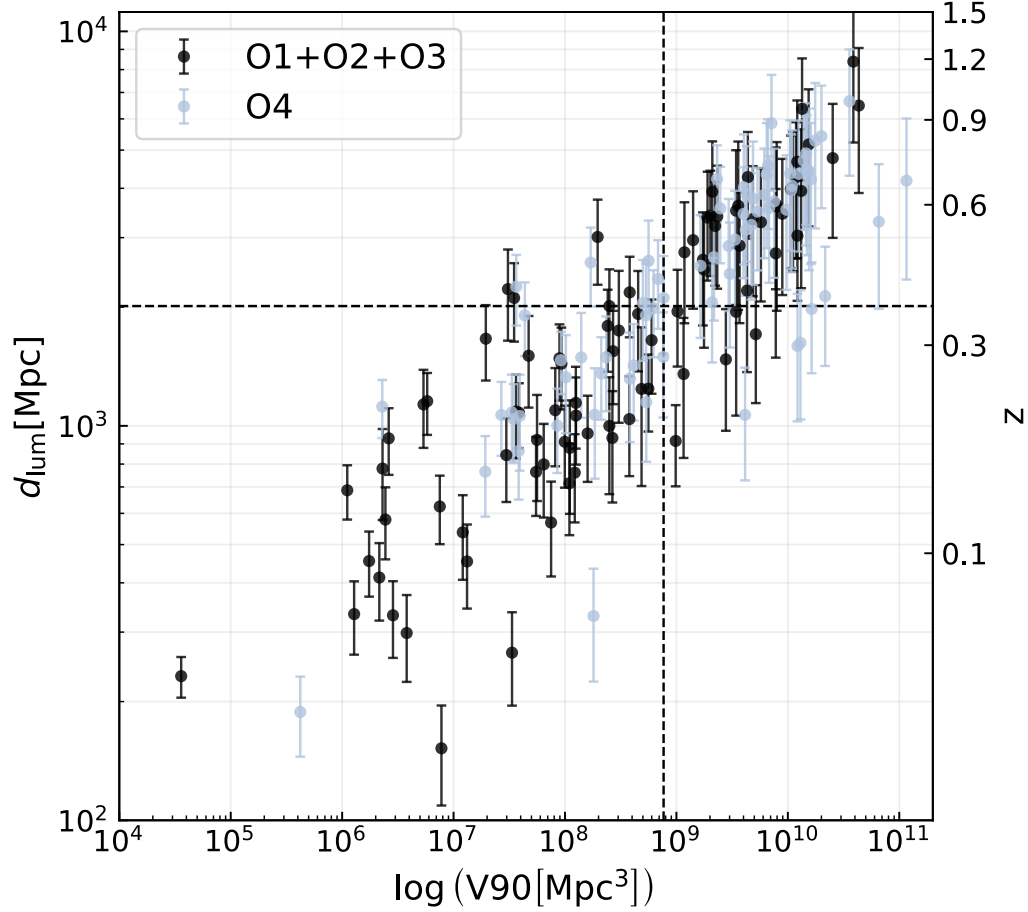


Figure 5.1: Luminosity distance as a function of V90 for the 159 GW events used in the analysis presented in this work. The error bars represent an uncertainty of one standard deviation around the mean of the estimated distance. While black markers indicate the mergers detected during the first three observing runs of the LVK collaboration, the light blue one represent the ones detected during O4, up to June 1st, 2024. The horizontal and the vertical dashed lines mark the median value of the all luminosity distance estimates and of the V90 distribution, respectively. The values reported on the vertical axis on the right-hand side of the panel indicate arbitrarily chosen values of redshift.

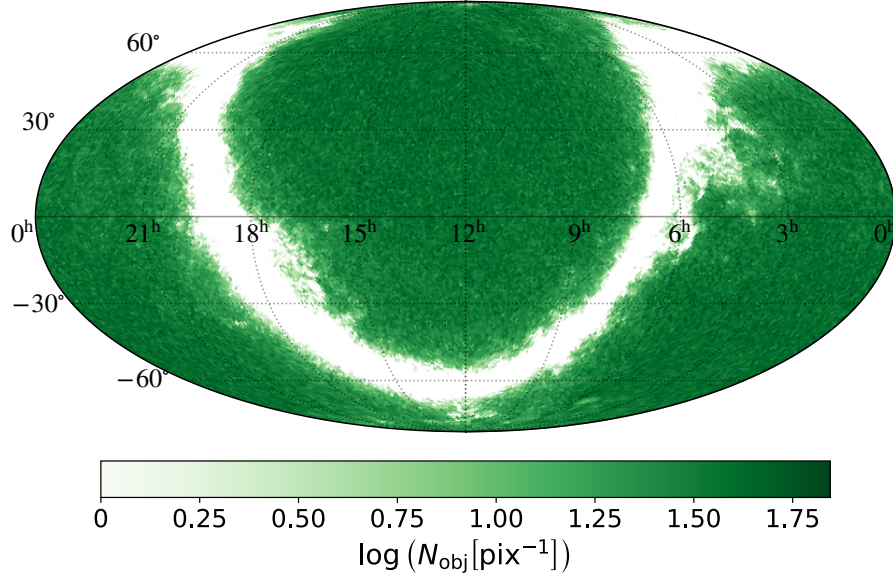


Figure 5.2: Mollweide projection of the sky positions of the 1,295,502 objects contained in the Quia catalog with a magnitude $G < 20.5$. Different shades of green represent different sky-projected number densities. The resolution is the one of an HealPix map with NSIDE=64.

pixel is $\approx 2.557 \cdot 10^{-4}$ steradians. Its remarkable uniformity outside the region containing the Milky Way galactic plane makes Quia a valid tool for spatial correlation analyses between AGN and GW events.

The AGN used during the cross-match with the GW sky maps are the 660,031 objects present in Quia that have a redshift estimate not larger than $z = 1.5$ and that are in the region of the sky where the galactic latitude b is either higher than 10° or smaller than -10° . The redshift selection is performed since all the GW events used in this work are expected to have taken place in the $0 \leq z \leq 1.5$ range (see Figure 5.1). We choose not to use the objects with a small absolute value of galactic latitude for simplicity. In our analysis in fact we consider the density of objects in the catalogue not to depend on their Right Ascension or Declination, as far as they are outside of the region of the galactic plane. This selection on sky position removes only ≈ 1.6 per cent of all the objects within $z = 1.5$ and is not expected to decrease significantly the constraining power of the analysis presented in this work. We further refer to the catalogue obtained after performing these selections in redshift and in sky

position as $\text{Quaia}_{\text{cut}}$.

5.2.3 Completeness estimation

The likelihood maximisation method presented in this work requires an estimate of the completeness of AGN catalogue that is used. We therefore estimate this property of $\text{Quaia}_{\text{cut}}$ as a function of redshift and luminosity. We do so in three steps:

- Estimation of the bolometric luminosity of each object in Quaia with a redshift estimate not higher than $z = 1.5$. This sub-section of Quaia is later referred to as $\text{Quaia}_{z < 1.5}$. The luminosities are evaluated starting from the Vega magnitudes in the G_{RP} band ($\text{mag}_{G_{\text{RP}}, \text{VEG}}$);
- Correction for the host-galaxy contribution to the total luminosity using as reference the AGN with a counterpart in SDSS DR16Q, and the values of their bolometric luminosities presented in [Wu & Shen \(2022\)](#);
- Comparison between the number of objects present in $\text{Quaia}_{\text{cut}}$ in a specific redshift bin and the ones predicted by the AGN luminosity function presented in [Kulkarni et al. \(2019\)](#). This is done five times, considering each time a different threshold in bolometric luminosity.

5.2.3.1 Evaluation of bolometric luminosity from $\text{mag}_{G_{\text{RP}}, \text{VEG}}$:

To obtain an estimate of the bolometric luminosity of all the objects contained in $\text{Quaia}_{z < 1.5}$, we first calculate the flux density in photo-electrons per second using the following standard relation:

$$f_{\nu_{G_{\text{RP}}}} \left[\frac{e^-}{s} \right] = 10^{-\frac{m_{G_{\text{RP}}, \text{VEG}} - ZP_{G_{\text{RP}}, \text{VEG}}}{2.5}}, \quad (5.1)$$

where $ZP_{G_{\text{RP}}, \text{VEG}} = 24.7479$ is the photometric zero-point in the Vega system for the G_{RP} band. The flux density in Jansky is then calculated by multiplying $f_{\nu} \left[\frac{e^-}{s} \right]$ by the conversion factor $c_{\nu} = 3.299 \cdot 10^{-36} \text{Jy} \cdot \text{s}/e^-$. The values of $ZP_{G_{\text{RP}}, \text{VEG}}$ and of c_{ν} are taken from the online documentation regarding the calibration of Gaia data ².

²<https://www.cosmos.esa.int/web/gaia-users/archive>

The intrinsic luminosity of each AGN emitted at a rest-frame frequency of $\nu_{\text{GRP}} = 10^{14.588} \text{ Hz}$ ³ is then calculated as follows:

$$\nu_{\text{GRP}} L_{\nu_{\text{GRP}}} = \nu_{\text{GRP}} f_{\nu_{\text{GRP}}} (4\pi d_{\text{lum}}^2) (1+z)^{-0.657} , \quad (5.2)$$

where d_{lum} is the luminosity distance correspondent to the redshift z of each object. The last term of Equation 5.2 is used to take into account the shape of the typical Spectral Energy Distribution (SED) of the objects in Quaia. The value of the exponent is calculated from a linear fit of the mean SED of all SDSS quasars presented in Richards et al. (2006). In particular, in the range between ν_{GRP} and $(1+z_{\text{max}})\nu_{\text{GRP}}$ (where for Quaia _{$z < 1.5$} , $z_{\text{max}} = 1.5$), we find that:

$$\log(\nu L_{\nu}) \propto \log(\nu) \cdot 0.657 , \quad (5.3)$$

and therefore

$$\nu_{\text{em}} L_{\nu, \text{em}} = \nu_{\text{GRP}} L_{\nu_{\text{GRP}}} (1+z)^{0.657} , \quad (5.4)$$

where $\nu_{\text{em}} L_{\nu, \text{em}}$ is the luminosity emitted in the rest-frame frequency that is observed at ν_{GRP} .

The bolometric luminosity is then calculated by multiplying $\nu_{\text{GRP}} L_{\nu_{\text{GRP}}}$ by the value that the frequency-dependent bolometric correction presented in Richards et al. (2006) has in correspondence of ν_{GRP} . This value is 14.59.

5.2.3.2 Comparison with the bolometric luminosity estimates of Wu & Shen (2022)

The bolometric luminosities calculated following the procedure described in Section 5.2.3.1 are not corrected for the contribution of the host galaxies, and they are therefore over-estimated. In order to adjust for this, we correct them through the comparison with the estimates listed in the catalogue presented in Wu & Shen (2022). This catalogue contains continuum and emission-line properties of the 750,414 broad-line Quasars of SDSS DR16Q. Among these properties there is the bolometric luminosity of each AGN, estimated from the continuum luminosity at rest-frame wavelengths of 5100, 3000, and 1350 Å. We perform a cross-match between this catalogue and Quaia _{$z < 1.5$} , using two arcsec as the maximum separation on the sky two objects can have in order to be considered as matching. A total of 136,368 matches are found. For each of these AGN is possible to calculate the difference between the logarithm of the bolometric

³this value has been calculated assuming a fiducial wavelength for the GRP band of 7750 Å.

Table 5.1: Partition of the AGN contained in both $\text{Quaia}_{z < 1.5}$ and the catalogue presented in Wu & Shen (2022). This partition has been used to compare the estimates of bolometric luminosity. In the Table we list the minimum and the maximum luminosity of each sub-sample. To perform this subdivision we use the luminosity estimates that are obtained starting from $\text{mag}_{\text{GRP,VEG}}$. In order to correct these estimates according to the values contained in the catalogue of Wu & Shen (2022), we divide each sub-sample of $\text{Quaia}_{z < 1.5}$ in linear redshift bins. The number of these bins for each sub-sample and the maximum value of the redshift used in this division, which has been done using the Quaia redshift estimates, are listed in the last two columns of the Table.

$\log(L_{\text{bol,min}} [\text{erg s}^{-1}])$	$\log(L_{\text{bol,max}} [\text{erg s}^{-1}])$	$N_{\text{bins,z}}$	z_{max}
46.5	-	4	1.5
46	46.5	5	1.5
45.5	46	8	1.5
45	45.5	9	1.1
-	45	3	0.5

luminosity estimated from $\text{mag}_{\text{GRP,VEG}}$ and the logarithm of the bolometric luminosity taken from the catalogue of Wu & Shen (2022). The mean value of the distribution of this differences is 0.073, and its standard deviation is 0.213. As expected, the bolometric luminosities obtained from $\text{mag}_{\text{GRP,VEG}}$ are on average slightly over-estimated. We then proceed to correct this over-estimation.

We divide the AGN that are both in $\text{Quaia}_{z < 1.5}$ and in the catalogue presented in Wu & Shen (2022) into five bolometric luminosity bins. This partition is performed using the estimates calculated starting from $\text{mag}_{\text{GRP,VEG}}$. Each of these sub-samples is then divided in redshift bins, using the Quaia estimate of redshift for each AGN. For each luminosity sub-sample we choose a different number of linear redshift bins and a different value of the maximum redshift to consider, to ensure that in each of these bins there are at least 10 objects. In Table 5.1 we list the minimum and the maximum bolometric luminosity for each of the five sub-samples, the number of redshift bins in which it is divided, and the maximum redshift considered in this division.

For each bin in luminosity and redshift, we perform a linear fit between the logarithm of the bolometric luminosities estimated starting from $\text{mag}_{\text{GRP,VEG}}$, and the bolometric luminosities taken from the catalogue of Wu & Shen (2022). We therefore obtain one set of best-fit parameters for each bin in luminosity and redshift. We finally use these best-fit parameters to correct the bolometric luminosity of each object contained in $\text{Quaia}_{\text{cut}}$ that was estimated from

$\text{mag}_{\text{GRP,VEG}}$. These adjusted estimates are the ones used in the rest of the analysis presented in this paper.

5.2.3.3 Comparison with the AGN luminosity function of Kulkarni et al. (2019)

In order to have an estimate of the completeness of $\text{Quaia}_{\text{cut}}$ as a function of redshift we compare the number of objects it contains with the expectation value calculated from the AGN luminosity function presented in Kulkarni et al. (2019). In particular we use the best-fit double power law function in which the parameters evolve as a function of redshift according to Model 1 (for the analytical expression see Equations 7, 13, 16, 17, and 18 of Kulkarni et al. 2019, while the best-fit parameters are listed in the first column of Table 3 of the same paper).

We compare the observed number of AGN with the predicted one for five different sub-samples of $\text{Quaia}_{\text{cut}}$. Each of these sub-samples is characterised uniquely by a different threshold in bolometric luminosity. We divide each sub-sample into 8 linear redshift bins between $z = 0$ and $z = 1.5$. The expected number of objects in each of these bins is calculated integrating the luminosity function of Kulkarni et al. (2019) in the proper range of redshift for each value of luminosity threshold used also in the subdivision of $\text{Quaia}_{\text{cut}}$. During the integration, in order to convert the absolute monochromatic AB magnitudes at a rest frame of 1450\AA , in terms of which the luminosity function is expressed, into bolometric luminosities, we use the bolometric correction presented in Runnoe et al. (2012).

The dashed lines in Figure 5.3 show the number of expected AGN in all the different redshift bins, obtained through the integration of the luminosity function. Different colours correspond to different bolometric luminosity thresholds. The round markers correspond to the number of objects brighter than the same threshold luminosities that are contained in $\text{Quaia}_{\text{cut}}$ in the different redshift bins, divided by $1 - \sin(10^\circ)$ to take into account that the objects in the catalogue are not distributed over the entirety of the sky area. The crosses show the number of objects in SDSS DR16 for each redshift bin and luminosity threshold, divided by $A_{\text{SDSS}}/A_{\text{sky}}$, where $A_{\text{SDSS}} = 9376 \text{ deg}^2$ is the area of the footprint of SDSS, and $A_{\text{sky}} = 41,253 \text{ deg}^2$ is the total area of the sky. We take the bolometric luminosities of the AGN in the SDSS DR16Q from the catalogue presented in Wu & Shen (2022).

Finally, the completeness of $\text{Quaia}_{\text{cut}}$ for each redshift bin and for each different cut in bolometric luminosity is calculated dividing the sky-area-corrected

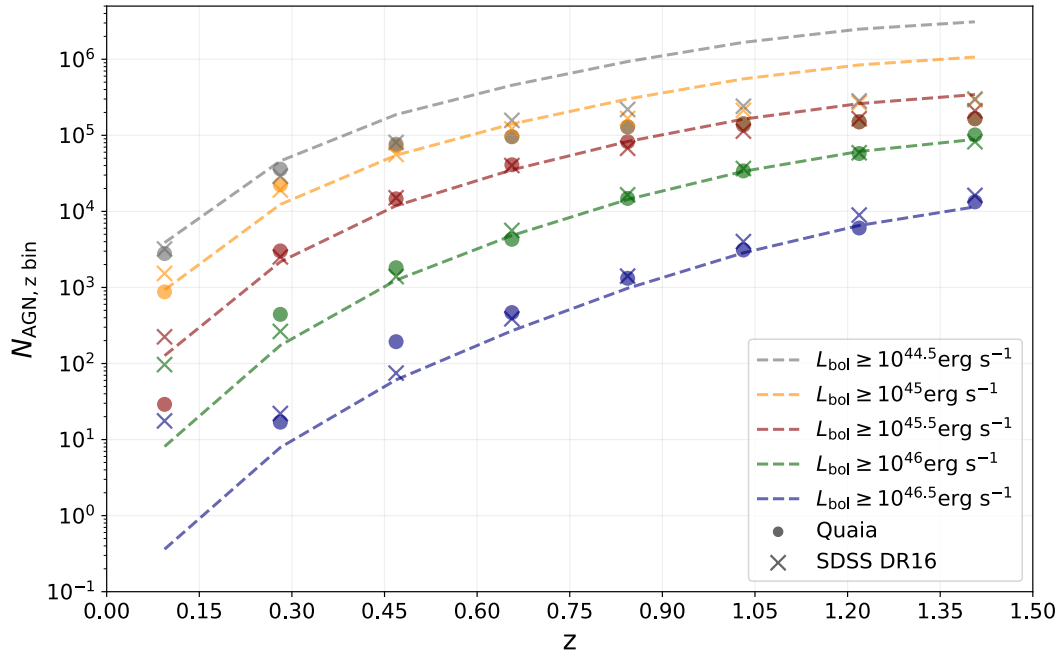


Figure 5.3: Comparison between the expected number of AGN in specific redshift bins as a function of different bolometric luminosity thresholds (dashed lines) and the observed number of objects contained in $\text{Quiaia}_{\text{cut}}$ (round markers) and in SDSS DR16 (cross markers). The expected numbers have been calculated integrating the redshift-evolving luminosity function presented in [Kulkarni et al. \(2019\)](#). Only the objects that have an absolute value of galactic latitude greater than 10° are in $\text{Quiaia}_{\text{cut}}$, and the area of the SDSS footprint is ≈ 22.7 per cent of the entire sky area. The plotted numbers of observed AGN have been renormalized according to this.

number of observed AGN by the expected one. Whenever the former is greater than the latter, the completeness is set to one. Because of the selection in galactic latitude $Quaia_{\text{cut}}$ is obtained from, its completeness in the region where $|b| < 10^\circ$ is zero. The values of the estimated completeness for all the different redshift bins and the different bolometric luminosity thresholds we consider in our analysis are listed in Table 5.2.

5.3 Method

Once the bolometric luminosity has been calculated for each AGN of $Quaia_{\text{cut}}$ and the completeness of such a catalogue has been estimated as a function of redshift and luminosity, we calculate the posterior on f_{AGN} . To do so, we use a likelihood function similar to the ones used in Veronesi et al. (2023) and in Veronesi et al. (2024). The general analytical expression of this function is the following:

$$\begin{aligned} \mathcal{L}(f_{\text{AGN}}) &= \prod_{i=1}^{N_{\text{GW}}} \mathcal{L}_i(f_{\text{AGN}}) \\ &= \prod_{i=1}^{N_{\text{GW}}} [c_i \cdot 0.90 \cdot f_{\text{AGN}} \cdot \mathcal{S}_i + (1 - c_i \cdot 0.90 \cdot f_{\text{AGN}}) \mathcal{B}_i] \quad , \end{aligned} \quad (5.5)$$

where $N_{\text{GW}} = 159$ is the total number of mergers of binaries of COs we consider in the analysis, and c_i is the average completeness of $Quaia_{\text{cut}}$ in the region occupied by $V90_i$, the 90 per cent localisation volume of the i -th GW event. Each value of c_i has been calculated assuming a null completeness in the region of the sky where $|b| < 10^\circ$.

For every GW event we calculate one value of c_i for each different bolometric luminosity threshold we use in our analysis. Figure 5.4 show the Cumulative Distribution Function (CDF) of c_i for all these different sub-samples of $Quaia_{\text{cut}}$. Different colours indicate different luminosity thresholds. The sub-samples of $Quaia_{\text{cut}}$ created with a threshold luminosity of $10^{46} \text{erg s}^{-1}$ and of $10^{46.5} \text{erg s}^{-1}$, and that therefore contain only the most luminous and rarest AGN, have a null completeness at low redshift. This is the reason why for ≈ 10 per cent of the GW events, $c_i = 0$ when those two catalogues are considered.

In Equation 5.5, \mathcal{S}_i is the signal probability density and is calculated as

Table 5.2: Estimated completeness of $\text{Quaia}_{\text{cut}}$ in the region of the sky with a galactic latitude greater than 10° or smaller than -10° . All the values are between 0 and 1, rounded up to the third decimal digit, and are listed for 8 different linear redshift bins as a function of 5 different bolometric luminosity thresholds.

	$\log(L_{\text{bol}} [\text{erg s}^{-1}]) \geq 46.5$	$\log(L_{\text{bol}} [\text{erg s}^{-1}]) \geq 46.0$	$\log(L_{\text{bol}} [\text{erg s}^{-1}]) \geq 45.5$	$\log(L_{\text{bol}} [\text{erg s}^{-1}]) \geq 45$	$\log(L_{\text{bol}} [\text{erg s}^{-1}]) \geq 44.5$
$0.0000 < z \leq 0.1875$	0	0	0.229	0.945	0.718
$0.1875 < z \leq 0.3750$	1	1	1	1	0.781
$0.3750 < z \leq 0.5625$	1	1	1	1	0.408
$0.5625 < z \leq 0.7500$	1	0.891	1	0.681	0.211
$0.7500 < z \leq 0.9375$	1	1	0.994	0.429	0.138
$0.9375 < z \leq 1.1250$	1	1	0.837	0.258	0.085
$1.1250 < z \leq 1.3125$	0.927	0.940	0.576	0.179	0.060
$1.3125 < z \leq 1.5000$	1	1	0.482	0.155	0.053

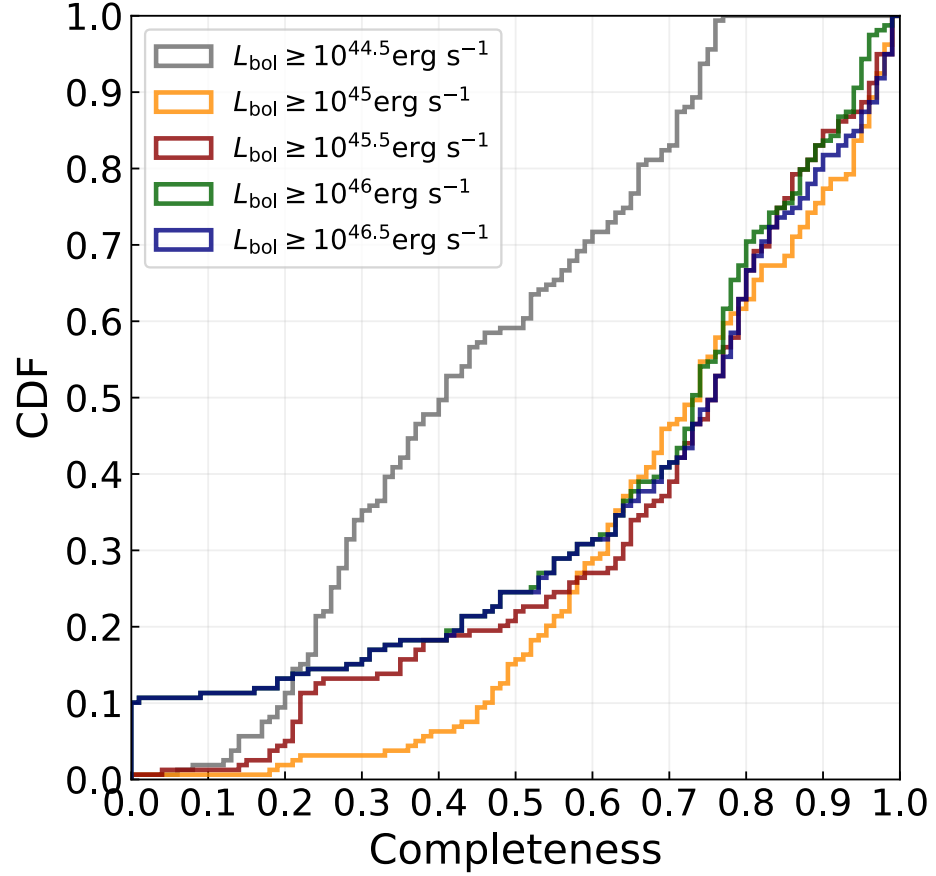


Figure 5.4: CDF of the values of average completeness of $\text{Quaia}_{\text{cut}}$ in the region contained within the 90 per cent Credibility Level localisation volume of each of the 159 GW events used in this work. Different colours correspond to different sub-samples of the catalogue. Each of these sub-samples is uniquely characterised by a different value of bolometric luminosity threshold.

follows:

$$\mathcal{S}_i = \frac{1}{V90_i} \sum_{j=1}^{N_{\text{AGN}, V90_i}} \frac{p_j}{n_{\text{AGN}, j}} , \quad (5.6)$$

where $N_{\text{AGN}, V90_i}$ is the number of AGN located within $V90_i$, p_j is measured in Mpc^{-3} and is the probability density representing how likely it is that the i -th GW event has happened in the exact position of the j -th AGN. Finally, $n_{\text{AGN}, j}$ is the number density of the AGN catalogue in the redshift bin where the j -th AGN is. The value of this number density for each AGN has been obtained by dividing the total number of objects in the redshift bin it is by the comoving volume enclosed in such a bin, excluding the region in which $|b| < 10^\circ$. This has been done for each of the sub-samples of $\text{Quaia}_{\text{cut}}$, separately.

The background probability density is calculated as follows:

$$\mathcal{B}_i = \frac{0.9}{V90_i} , \quad (5.7)$$

where, in analogy to what has been done in [Veronesi et al. \(2023\)](#) and in [Veronesi et al. \(2024\)](#), the 0.9 factor ensures that \mathcal{B}_i and \mathcal{S}_i have the same normalisation.

We cross-match the sky maps of the 159 GW events with the 5 different sub-samples of $\text{Quaia}_{\text{cut}}$ separately, using the `postprocess.crossmatch` function of the package `ligo.skymap`. The results of these cross-matches are used to evaluate $\mathcal{L}(f_{\text{AGN}})$ using Equation 5.5. We then calculate the posterior probability distribution $\mathcal{P}_{f_{\text{AGN}}}$, normalising the likelihood function and assuming a uniform prior on f_{AGN} in the $[0, 1]$ range.

5.4 Results

The posterior probability on f_{AGN} peaks at $\hat{f}_{\text{AGN}} = 0$ independently on which sub-set of $\text{Quaia}_{\text{cut}}$ we consider. In Figure 5.5 is shown in blue the region of the parameter space investigated in this work that our analysis rejects with a credibility of 95 per cent. We show the comparison with previous results, obtained in [Veronesi et al. \(2023\)](#) using a more limited dataset.

In Table 5.3 we list, as a function of the threshold on bolometric luminosity, how many objects are considered during the cross-match, what fraction of $\text{Quaia}_{\text{cut}}$ they consist of, and the upper limits we put on f_{AGN} at 68, 90, and 95 per cent credibility.

Figure 5.6 shows the logarithm of the ratio between the single-event likelihoods (\mathcal{L}_i) calculated at $f_{\text{AGN}} = 1$ and the ones calculated at $f_{\text{AGN}} = 0$, as a

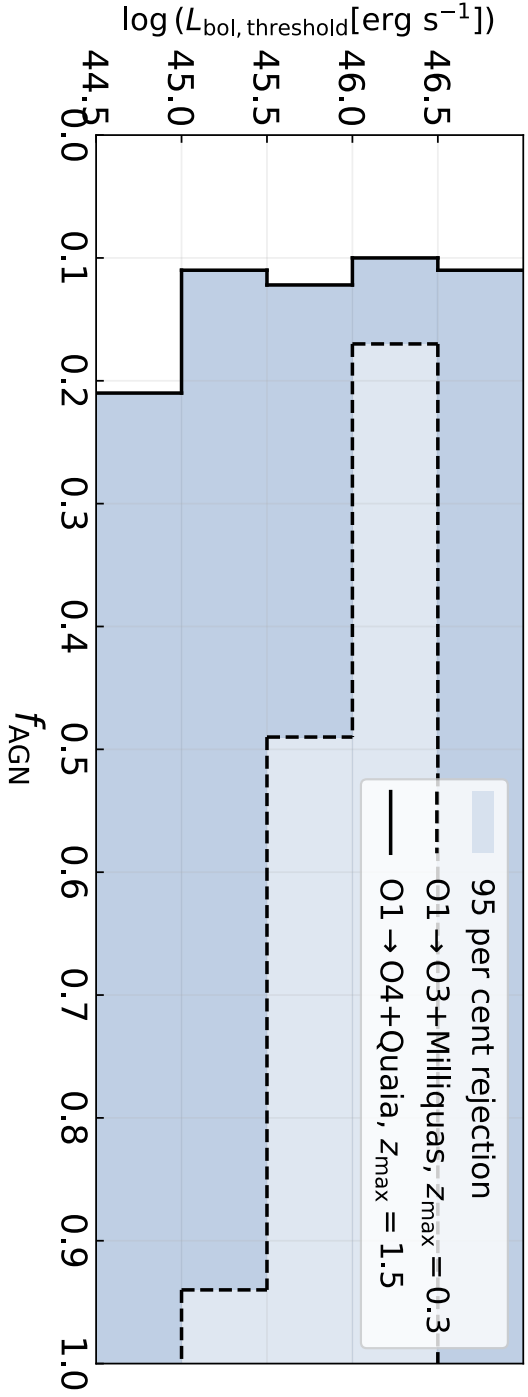


Figure 5.5: Observational constraints on f_{AGN} based on spatial correlation. The blue region of the plot shows the part of the parameter space we investigated that is rejected by our analysis at a 95 per cent credibility level. The region enclosed in the dashed line shows the results obtained with a more limited dataset in Veronesi et al. (2023). Such previous work used the 30 GW events detected in the first three observing runs of the LVK collaboration that are located within $z = 0.3$ at a 90 per cent credibility level, and three different catalogues of AGN in the same redshift range. The region enclosed in the solid line shows the results of this work, which explores a wider range of AGN luminosities and uses all the mergers of binaries of COs directly detected up until June 1st, 2024.

Table 5.3: Upper limits on f_{AGN} we obtain at different levels of credibility, for the three cuts in bolometric luminosity we consider. For each of such cuts we also list the number of AGN used in the analysis ($N_{\text{AGN, cut}}$), and what fraction of the total number of AGN ($N_{\text{AGN, Quia}_{\text{cut}}}$) they consist of. All the values in this table have been rounded up to the second decimal digit.

$\log(L_{\text{bol, threshold}} [\text{erg s}^{-1}])$	$N_{\text{AGN, cut}}$	$N_{\text{AGN, cut}}/N_{\text{AGN, Quia}_{\text{cut}}}$	68 per cent upper limit	90 per cent upper limit	95 per cent upper limit
46.5	20,236	0.03	0.05	0.09	0.11
46	177,117	0.27	0.04	0.08	0.10
45.5	490,628	0.74	0.05	0.10	0.13
45	644,393	0.98	0.05	0.09	0.11
44.5	659,949	1.00	0.08	0.16	0.21

function of $V90_i$. Different panels correspond to different bolometric luminosity thresholds. In each plot the markers are coloured according to the average completeness inside the localisation volume of the corresponding merger. The dashed horizontal lines indicate where the logarithm has a null value, therefore where $\mathcal{L}_i(f_{\text{AGN}} = 1) = \mathcal{L}_i(f_{\text{AGN}} = 0)$. The markers above these lines correspond to GW events according to which the hypothesis of random-chance association between mergers and AGN is disfavoured.

One event in particular, detected during the second half of O4, has a correspondent value of $\mathcal{L}_i(f_{\text{AGN}} = 1) / \mathcal{L}_i(f_{\text{AGN}} = 0)$ which is greater than one, independently on the luminosity cut of the Quaia sub-sample it is cross-matched with. The ID of such event is S240511i, and the corresponding marker is particularly evident in the first two panels of Figure 5.5. In the case of the cross-match with the sub-sample of $\text{Quaia}_{\text{cut}}$ characterised by AGN brighter than $10^{46.5} \text{erg s}^{-1}$ ($10^{46} \text{erg s}^{-1}$), for this particular GW event $\log(\mathcal{L}_i(f_{\text{AGN}} = 1) / \mathcal{L}_i(f_{\text{AGN}} = 0)) \approx 1.32$ (0.72), which means that the value of the single-event likelihood calculated at $f_{\text{AGN}} = 1$ is ≈ 21 (5) times larger than the one at $f_{\text{AGN}} = 0$. This result based solely on spatial correlation hints towards a possible AGN origin of such event. Further follow-up analyses regarding this merger and its intrinsic binary properties will be needed to be able to draw any definitive conclusion.

5.5 Discussion and conclusion

In this work we present new observational constraints on the fractional contribution of the AGN channel to the total observed merger rate of binaries of COs. These constraints are obtained using the same spatial-correlation-based approach used in Veronesi et al. (2023). With respect to our previous work, we make use of a new, larger dataset, which consists of 159 GW events detected by the interferometers of the LVK collaboration not later than June 1st, 2024 and of the all-sky AGN catalogue Quaia. In particular we use all the AGN contained in such a catalogue that have a redshift estimate not larger than $z = 1.5$ and an absolute value of the galactic latitude $|b| \geq 10^\circ$. We call this sub-sample $\text{Quaia}_{\text{cut}}$.

We estimate the bolometric luminosity of every object in $\text{Quaia}_{\text{cut}}$ using the magnitudes in the Gaia G_{RP} band. We also estimate the completeness of such a catalogue as a function of redshift for different values of bolometric luminosity threshold. The average value of the completeness within the 90 per cent Credibility Level localisation volume of each GW event is used during the

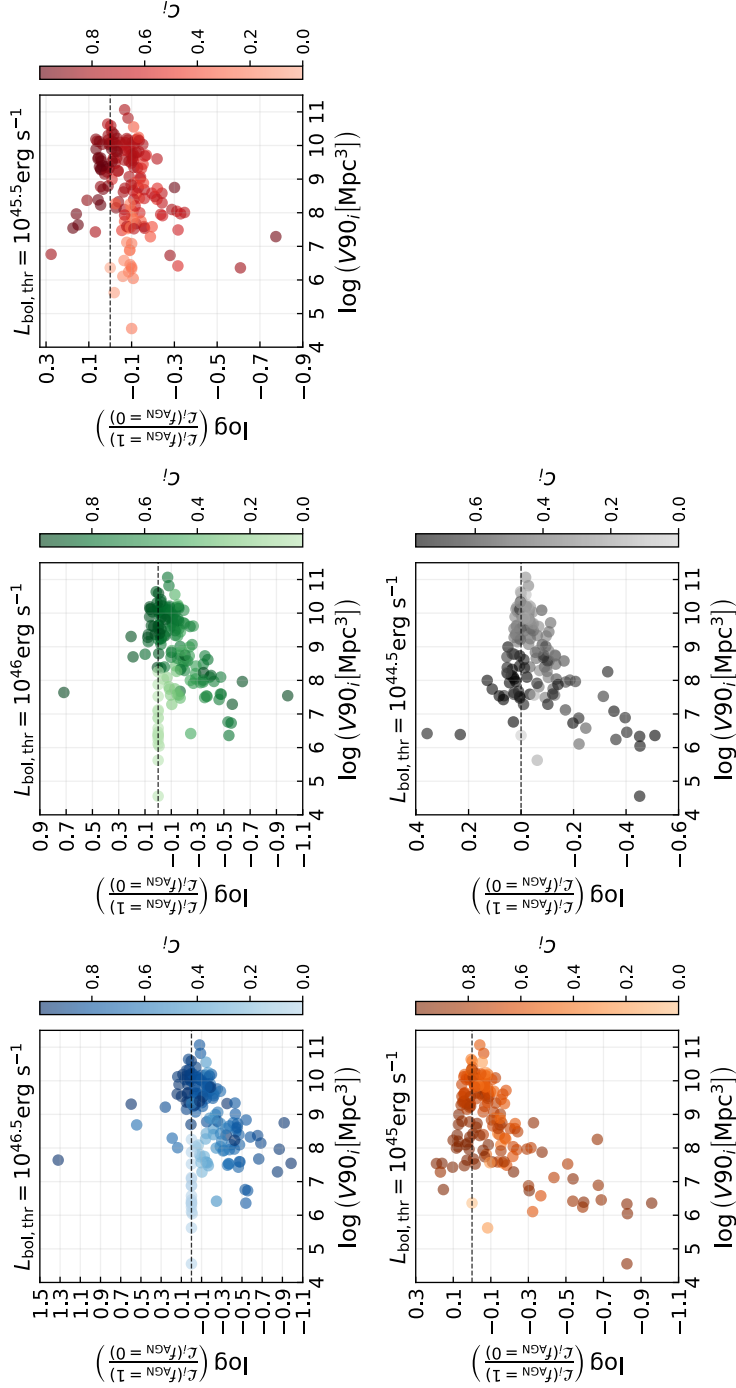


Figure 5.6: Logarithm of the ratio between the value of the likelihood at $f_{\text{AGN}} = 1$ and the one at $f_{\text{AGN}} = 0$ for each GW event, as a function of the corresponding V90. Each panel shows the results for a different bolometric luminosity threshold. In each plot, the markers are coloured as a function of the average completeness of Quiaia_{cut} in the region occupied by V90. The dashed horizontal lines mark where $\mathcal{L}_i(f_{\text{AGN}} = 1) = \mathcal{L}_i(f_{\text{AGN}} = 0)$.

likelihood maximisation process.

We calculate the posterior probability function on the fractional contribution of the AGN channel to the total merger rate for all the different luminosity thresholds, obtaining that this function always peaks at $\hat{f}_{\text{AGN}} = 0$. We calculate the upper limits of the 68, 90, and 95 per cent Credibility Intervals on f_{AGN} . The main results of this work are summarised in Figure 5.5 and in Table 5.3. In particular we estimate that no more than the 21 per cent of the detected GW events used in our analysis originated from an AGN with a bolometric luminosity higher than $10^{44.5} \text{ erg s}^{-1}$ with 95 per cent credibility. Objects brighter than such threshold consist of almost the entirety of the catalogue we consider. Due to the flux limitations of Quiaia, fainter objects are not included in our analysis.

We find a particular GW event for which our analysis suggests a possible AGN origin, especially when objects brighter than $10^{46} \text{ erg s}^{-1}$ are considered. The ID of such event is S240511i, and it has been detected during O4b. Since the statistical framework used in this work is focused on analysing the entire population of GW events, the results concerning this specific event are to be considered as hints, not as statistically significant conclusions. Follow-up analyses, conducted especially when the full catalogue of mergers detected during O4 will be published, will be necessary to confidently assess whether or not this merger has an AGN origin.

Thanks to the increase of a factor ≈ 5 in the number of used GW events with respect to Veronesi et al. (2023) we are able to put much tighter constraints on the efficiency of the AGN channel (see Figure 5.5 for a comparison with our previous work). However, our results have been obtained under some assumptions, which are inevitable when inferring the properties of AGN in large catalogues. First and foremost, in order to calculate the bolometric luminosities of the Objects of Quiaia we have to assume a shape of the typical SED and a value for the bolometric correction. We take the values of these AGN properties from Richards et al. (2006) in order to be consistent with what has been done in Wu & Shen (2022), which we use to adjust the bolometric luminosities obtained from Gaia magnitudes. Different assumptions on the SED and on the bolometric correction might lead to different estimates of the bolometric luminosities, even if these difference are not expected to be large enough to be significant for our statistical analysis.

One more assumption is to be made during the calculation of the completeness of $\text{Quiaia}_{\text{cut}}$. This is in fact done comparing the observed number of AGN with the expected one, which is obtained from the integration of a luminosity function. The best fit parameters of these functions depend on the data used during the fit, and different works might assume different analytical expressions

for them. We choose to use the luminosity function presented in [Kulkarni et al. \(2019\)](#) because in the redshift range considered in our analysis it has been fitted on AGN from SDSS, which is a survey that contains un-obscured AGN detectable in the optical band, like the ones contained in Quaia. Moreover, the objects contained in [Wu & Shen \(2022\)](#), used to obtain the final estimate on the bolometric luminosity of the objects in $\text{Quaia}_{\text{cut}}$, come from SDSS. Comparing the measured number of objects in our AGN catalogue with a different luminosity function might lead to different estimates of its completeness, and the constraining power presented in this work is expected to scale linearly with such a parameter.

Another important caveat to mention is that our analysis and the constraints we are able to put concern un-obscured AGN, which are visible both in the optical and in the infrared band. In order to extend our conclusions to the entire population of AGN, one should take into account what fraction of them is not visible in the wavelength observed by Gaia and by WISE, which are the two surveys from which Quaia has been created from. This obscuration fraction is in general expected to increase as a function of redshift and to decrease as a function of luminosity ([Merloni et al. 2014](#); [Ueda et al. 2014](#)).

The observational constraints presented in this work consist of a generalisation of the ones obtained in [Veronesi et al. \(2023\)](#). Here we investigate a wider range of AGN luminosities and the entirety of the redshift range reached by the interferometers of the LVK collaboration. For this reason, in order to obtain in the future even more general results one will need an all-sky AGN catalogue with a lower threshold in flux with respect to Quaia, to be able to extend the analysis to the faint-end of the un-obscured AGN population. We estimate that no more than one GW event out of five has originated in an un-obscured AGN with a bolometric luminosity higher than $10^{44.5} \text{erg s}^{-1}$, but a more significant fraction might still come from fainter objects, or from obscured ones.

While the results of this work demonstrate that the efficiency of the AGN channel can already be investigated with spatial-correlation analyses using the currently available data, in the next years it will be possible to put even tighter constraints using all the events that will be detected in the rest of O4 as well as the ones that will be detected during the fifth observing run of the LVK collaboration, O5. Using more data could either reduce the upper limits on f_{AGN} , or it could cause a shifting of the value of such parameter that maximised the posterior distribution, moving it away from zero.

Future developments of the statistical method used in this work involve also the introduction of physically-motivate priors on the intrinsic binary properties. Different binary formation channels are indeed expected to produce different

features in the distributions of the masses and the spins of the merging systems, as well as on their eccentricity. The analysis here presented has been kept purposely agnostic as far as the physics of the formation mechanism is concerned. Introducing physically-motivated changes in the likelihood function might result in different constraints on f_{AGN} and will inform us on which are the intrinsic binary parameters that are able to add more information to the analysis.

Acknowledgements

EMR acknowledges support from ERC Grant “VEGA P.”, number 101002511. This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gwosc.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. KAGRA is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea; Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan. *Software*: Numpy ([Harris et al. 2020](#)); Matplotlib ([Hunter 2007](#)); SciPy ([Virtanen et al. 2020](#)); Astropy ([Astropy Collaboration et al. 2013, 2018](#)); Bilby ([Ashton et al. 2019](#)); BAYESTAR ([Singer & Price 2016](#)); Healpy ([Zonca et al. 2019](#)).

Data Availabilty

The data underlying this article will be shared on reasonable request to the corresponding author.