Massive black hole binaries in LISA: multimessenger prospects and electromagnetic counterparts

Alberto Mangiagli,[1, *] Chiara Caprini,[2, 3] Marta Volonteri,[4] Sylvain
Marsat,[5] Susanna Vergani,[6] Nicola Tamanini,[5] and Henri Inchauspé[1, 7]

[1]*Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France*

[2]*Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics,*
24 quai Ernest-Ansermet, CH-1211 Genéve 4, Switzerland

[3]*CERN, Theoretical Physics Department, 1 Esplanade des Particules, CH-1211 Genéve 23, Switzerland*

[4]*Institut d'Astrophysique de Paris, CNRS & Sorbonne Université,*
UMR 7095, 98 bis bd Arago, F-75014 Paris, France

[5]*Laboratoire des 2 Infinis - Toulouse (L2IT-IN2P3),*
Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France

[6]*LUTH, GEPI and LERMA, Observatoire de paris, CNRS,*
PSL University, 5 place Jules Janssen, 92190, Meudon, France

[7]*Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*

(Dated: November 23, 2022)

In the next decade, the Laser Interferometer Space Antenna (LISA) will detect the coalescence of massive black hole binaries (MBHBs) in the range $[10^4, 10^8]$ M$_\odot$, up to $z \sim 10$. Their gravitational wave (GW) signal is expected to be accompanied by an electromagnetic counterpart (EMcp), generated by the gas accreting on the binary or on the remnant BH. In this work, we present the number and characteristics (such as redshift and mass distribution, apparent magnitudes or fluxes) of EMcps detectable jointly by LISA and some representative EM telescopes. We combine state-of-the-art astrophysical models for the galaxies formation and evolution to build the MBHBs catalogues, with Bayesian tools to estimate the binary sky position uncertainty from the GW signal. Exploiting additional information from the astrophysical models, such as the amount of accreted gas and the BH spins, we evaluate the expected EM emission in the soft X-ray, optical and radio bands. Overall, we predict between 7 and 20 EMcps in 4 yrs of joint observations by LISA and the considered EM facilities, depending on the astrophysical model. We also explore the impact of the hydrogen and dust obscuration of the optical and X-ray emissions, as well as of the collimation of the radio emission: these effects reduce the number to EMcps to 2 or 3, depending on the astrophysical model, again in 4 yrs of observations. Most of the EMcps are characterised by faint EM emission, challenging the observational capabilities of future telescopes. Finally, we also find that systems with multi-modal sky position posterior distributions represent only a minority of cases and do not affect significantly the number of EMcps.

I. INTRODUCTION

The Laser Interferometer Space Antenna (LISA)[1] is planned for launch in 2034 and will detect gravitational waves (GWs) in between $[10^{-4}, 10^{-1}]$ Hz. Among other sources, LISA will detect the coalescence of massive black hole binaries (MBHBs) in the entire Universe up to redshift $z \sim 20$ before the epoch of re-ionization [2–8] and in the range of masses $\sim [10^4, 10^8]$ M$_\odot$ in the nearby Universe. Detecting GWs from these sources will allow to reconstruct the merger history of MBHBs, disentangling the astrophysical processes and mechanisms driving their formation and evolution [9–17], perform tests of general relativity [18, 19] and constrain cosmological scenarios [20–25] (see [26, 27] for recent reviews on cosmological and fundamental physics implications of LISA).

Compact binaries emitting GWs can be considered "standard sirens", because they provide access to the source luminosity distance d_L. The latter is indeed encoded in the waveform and can be extracted directly, without resorting to a cosmic distance ladder, as necessary for type Ia supernovae (SNIa) [28]. However, GWs alone do not provide the redshift of the source. In the presence of an electromagnetic (EM) counterpart, the redshift can be obtained identifying and observing the host galaxy with EM facilities; this information can then be used to construct the $d_L - z$ diagram and constrain cosmological parameters [20, 29] (see also [30–32] for ground-based detectors). If no EM counterpart is present, statistical methods can be employed to infer the cosmological parameters, if enough GW sources are available [33–39].

In the case of MBHBs, the presence of an EM counterpart accompanying the GW signal has long been discussed in the literature and the situation is still unclear, mostly due to the lack of observational evidence. In the presence of a sufficient amount of gas in the close envi-

rons of the binary, an EM counterpart can be triggered by the accretion of the gas onto the binary during the inspiral, merger or ringdown [40–44]. The binary motion is expected to excavate a cavity in the circumbinary disk, while gas streams from the inner edge of the disk should form minidisks surrounding each BH, contributing to spectral features and variable EM emission at various wavelengths previous to merger [45–49]. Moreover, the orbital motion of the binary is expected to imprint a modulation on the EM counterpart from minidisks in phase with the GW signal, allowing for the possible identification of the host galaxy in the field of view provided by LISA [50–52]. Additional features can appear at or after merger, for instance an increase in jet power [53], high accretion rate episodes similar to Active Galactic Nuclei (AGN) emission [42], spectral or transient features caused by gravitational recoil [54, 55].

In this work, we present different scenarios for the EM counterpart of MBHB mergers, exploring the potential of multimessenger observations with LISA and future EM facilities. This is the first of a series of papers, sharing the common objective of upgrading the analyses of Klein *et al.* [15] and Tamanini *et al.* [20] (hereafter T16), to provide up to date forecasts on the ability of LISA to constrain MBHBs parameters (especially the sky position and the luminosity distance) with the final aim of probing the expansion of the Universe. For this reason, we present here detection strategies *always including the redshift determination.* Only in Sec. VII B, for comparison, we provide the predicted numbers of MBHBs mergers with associated EM emissions without imposing the redshift determination. Among the other papers, one will focus on the construction of the MBHBs standard sirens catalogues and on the inference of the cosmological parameters, while in the others we will discuss extensively the parameter estimation of the GW signal for this type of sources.

II. GENERAL STRATEGY

In order to provide updated forecasts for multimessenger detection of MBHBs with LISA, we improve and complement the EM counterpart types proposed in T16, as well as the MBHBs parameter estimation with LISA.

Concerning EM counterparts, as put forward in T16, several options can be envisaged. First of all, if either the AGN or the host galaxy are sufficiently bright, and the LISA sky localization error small enough, the system can be identified and its redshift directly measured. Another possibility is the formation of a radio jet or flare during/after the merger, to be detected in the sky localization area provided by LISA. This would allow us to pinpoint the GW source sky position, with subsequent identification of the host galaxy. The source redshift could then be estimated either spectroscopically or photometrically with an optical telescope. Similarly, the X-ray emission associated to the MBHBs could also be used

to identify the GW source sky position, and in turn to determine the host galaxy.

In the context of the counterpart types described above, we consider in this work specific EM observatories. For the direct optical identification of an AGN at the time of the MBHB merger, we consider the Vera C. Rubin Observatory [56, 57]. We assume that the identification via the radio emission is performed by the future radio telescope Square Kilometre Array (SKA) [58]. In addition, we also explore the possibility to detect the X-ray EM counterpart with the Advanced Telescope for High ENergy Astrophysics (Athena) [59–61]. Once the galaxy is identified from the radio or X-ray emission in the sky localization error region provided by LISA, the redshift measurement can be obtained with the Extremely Large Telescope (ELT) [62] as an example of a telescope with a $\gtrsim 30\,\mathrm{m}$ mirror or, if possible, directly with the Rubin Observatory.

In summary, we will analyse 3 observational scenarios:

(a) the Rubin Observatory alone (both identification and redshift)

(b) SKA (identification) + ELT (redshift)

(c) Athena (identification) + ELT (redshift)

with several variations, detailed in Section IV and V to bracket the uncertainties.

As a starting point, we need a population of merging MBHBs. Following T16, we adopt the result of semi-analytical models (SAM) [14, 63–65] to track the evolution of BH masses, spins and surrounding gas across the cosmic time. Specifically we consider three models to explore different seed and time-delay prescriptions:

1. Pop3: a light-seed model with delays included where BHs form from very massive metal-poor stars at high redshift [66, 67];

2. Q3d: a heavy-seed model with delays included where MBHs originate from the collapse of proto-galactic disks [68, 69];

3. Q3nd: a similar heavy-seed model [68, 69] but without delays between the galaxy and the BH merger, leading to more events but skewed toward higher redshift. In this sense, this scenario can be considered as optimistic in the number of predicted MBHB mergers.

These models predict the merger rate, the intrinsic binary properties (masses, spin magnitudes and orientations, luminosity distance) and the properties of the host galaxy (amount of mass in gas and stars, mass in the disk, etc.). For each model, we use a catalogue containing 90 years of data. We assume 4 years of LISA observations, corresponding to an overall mission duration of 5 years with 80% duty cycle of data taking. We further complete the catalogues by assigning randomly to each event the

sky position (uniform over the sky sphere), orbit inclination (random in $[0, \pi]$), polarization (random in $[0, 2\pi]$), coalescence phase (random in $[0, 2\pi]$), and merger time (random over $[0, 1]$ years [70]).

Given the simulated MBHB population, in order to reproduce the actual observational process, the next step should be to perform parameter estimation of the GW signal for each of the MBHBs in the catalogues, to infer the sky localization error. If the latter is small enough, one would then turn to evaluating the detectability of the EM counterpart.

In this work we use the Bayesian Markov Chain Monte-Carlo (MCMC) approach of [71] for the LISA parameter estimation, improving on the Fisher forecast of T16. However, this implies that the parameter estimation of the GW signal is the most computationally expensive step. Moreover, not every system in the catalogues is expected to produce a detectable EM counterpart, as the emission might be too faint, or the merger might happen in a dry environment. Therefore, we choose to assess the detectability of the EM counterpart in the first place. We select the systems whose fluxes (or magnitudes) are greater (smaller) than the corresponding threshold values for each of the listed EM facilities, and we run the parameter estimation only on this subset of events. Applying an initial cut in the EM detectability allows us to reduce the number of sources for which we have to perform the parameter estimation, limiting the computational effort.

For the subset of systems with detectable EM counterpart, we simulate the full inspiral-merger-ringdown GW signal in LISA using the waveform model PhenomHM for circularized binaries with aligned spins [72], further select those with signal-to-noise ratio (SNR) above 10, and we estimate the binary parameters of these systems with the MCMC.

In order to detect an EM counterpart (especially if they are transients close to merger [41]), telescopes must be pointed to the expected sky position of the GW source inferred by LISA (similarly to the alerts provided by LIGO/Virgo). For this reason, we apply a cut in the sky localization uncertainty. Together with the above mentioned SNR level and the detectability of the EM counterpart, we further require that the binary systems satisfy $\Delta\Omega < 10\,\mathrm{deg}^2$ to guarantee detection with the Rubin Observatory and SKA, or $\Delta\Omega < 0.4\,\mathrm{deg}^2$ to guarantee detection with Athena (see also Section IV C for another possible strategy).

Following this procedure, we define as *multimessenger candidate* (hereafter MMcand in the figures) any MBHB system within the catalogue that satisfies the following two conditions:

> *Multimessenger candidate*:

1. The system has a detectable EM counterpart;

2. The system has GW SNR > 10.

According to this definition, a multimessenger candidate is a system that can be detected by LISA and by any EM facility, but for which we impose no restrictive requirement on the sky localization. In other words, multimessenger candidates are mergers detectable by LISA for which the EM emission would be observable if the sky position were known with a certain accuracy.

We then define as *GW event with EM counterpart* (hereafter EMcp) any system that satisfies the following two conditions:

> *GW event with EM counterpart*:

1. The system is a *multimessenger candidate*;

2. The system is localized by LISA with $\Delta\Omega < 10\,\mathrm{deg}^2$ if detectable by the Rubin Observatory and SKA, and/or $\Delta\Omega < 0.4\,\mathrm{deg}^2$ if detectable by Athena.

We stress again that, according to the definition of the three observational strategies at the beginning of this section, we require the redshift determination both for multimessenger candidates and EMcps.

An important caveat of our analysis is that the cuts in the number of GW events performed to implement the observability of the EM counterpart concern the magnitude level and the sky localization, but not the event sky position. This is particularly relevant for Earth-based observatories (i.e. the Rubin Observatory, SKA and ELT), which cover only a fraction of the sky. The only way to implement a sky fraction cut would have been to apply an overall reduction factor on the number of GW events with EM counterpart, corresponding to the sky fraction covered by each facility, and elaborate some technique to account for the detection of the same event by multiple telescopes. Instead of adopting this crude method, we have decided to neglect the observable sky fraction as a whole when we implement the detection with Earth-based telescopes. This rather optimistic choice can lead to an overestimation of the number of predicted GW events with EM counterpart, with respect to those that might effectively be observable with actual data. We consider this a minor issue, compared with the uncertainty inherent in astrophysical MBHB evolution scenarios.

The paper is organised as follows. In Section III we describe the MBHBs catalogues and the physics of the SAM that affects the formation and evolution of the MBHBs. In Section IV we present how we model the EM counterpart to MBHB mergers for different wavelengths and telescopes. In Section V we describe a scenario where the EM flux is reduced by the surrounding gas. In Section VI we describe the tools we adopted to simulate the GW signal and to perform the parameter estimation. Our main results are reported in Section VII. In Section VIII we analyse multi-modal systems, i.e. systems whose sky localization inferred by LISA is dislocated in several portion of the sky. In Section IX we conclude with some final remarks and comments. In Appendix A we compare our results with previous works in the literature, identifying the reasons of the discrepancies, when applicable. In Appendix B we discuss briefly the role of the SNR and other binary parameters in determining the number of EMcps.

Finally in Appendix C we present some figures useful for discussion.

III. CATALOGUE OF MBHBS

The MBHB populations adopted in this work are based on a semi-analytical galaxy formation and evolution model [14, 63–65] (the same model is employed also in Belgacem *et al.* [25] and in T16). We refer the interested readers to the original papers and here we summarise only the general features of the model.

The model evolves dark matter merger trees from a Press-Schechter formalism along with the galactic baryonic structures, accounting for the complex interplay between the multiple components (intergalactic medium, interstellar medium, disk and nuclear cluster properties and the massive BH). Among the physical aspects that affect the mass distribution and the merger rate of MBHBs, we focus on two that have been shown to have a strong effect: the seed prescription, which defines the starting point for the MBH growth, and the time delay between the galaxy merger and the MBHB coalescence.

The light-seed prescription assumes that the first BHs form at high redshift ($z > 15 − 20$) in the range $\sim [10^2, 10^3]\,\mathrm{M_\odot}$ from the collapse of heavy Pop3 stars in the most metal-poor dark matter halos. We will refer to this seed model as 'Pop3'. In the heavy-seed prescription, most of the mass in a protogalactic disk collapses into a supermassive star or a quasistar leaving behind a MBH in the range $\sim [10^4, 10^5]\,\mathrm{M_\odot}$ at $z \sim 8 − 15$. The heavy seeds are considered to be rarer than the light seeds due to their particular birth environmental conditions [73, 74].

The time delays represent the time between the merger of the galaxies and the coalescence of the MBHBs. During a galaxy merger, the two MBHs migrate toward the center owing to dynamical friction [75] operating on the individual MBHs. Dynamical friction generated by the interaction between a massive perturber and the stellar and gaseous distribution decelerates the perturber. If dynamical friction is sufficiently effective, the MBH orbits will decay until they find each other at the center of the galaxy merger remnant and form a bound binary. At this point the MBHs have typical separations of sub-parsec to a few parsecs, depending on the binary mass, still far from the orbital scale at which GWs can efficiently subtract energy and orbital angular momentum to the binary leading to the final coalescence ($\simeq 10^{-3}\,\mathrm{pc}$). Additional processes are therefore needed to further shrink the orbit: energy exchanges in three-body interactions between the MBHB and nearby stars (a process referred to as stellar hardening), gas torques in circumbinary discs, and scattering between the MBHB and a third incoming MBH are included in the semi-analytical models of [65], which are used for this paper in order to account for how the MBHB crosses from parsec to milliparsec scales. There are large uncertainties in all these steps and assessing the efficiency of these mechanisms is beyond the scope of this

work. Here we summarise briefly how time delays have been derived in [65] and we refer to T16 and [65] for more details. In a gas-rich environment with $M_{\rm res} > M_{\rm tot}$, where $M_{\rm res}$ is the mass of the reservoir gas available for accretion [76] and $M_{\rm tot} = m_1 + m_2$ is the binary mass (in our work we set $m_1 > m_2$ with m_1 the mass of the primary BH), the time delays are set by the viscous time of the nuclear gas (see Eq. 29 in [65]), that is supposed to bring the MBHB to coalescence in $\lesssim 10^{7-8}$ yr. In the case of a gas-poor environment with $M_{\rm res} < M_{\rm tot}$, stars would be responsible for bringing the two MBHs together to the scale where GW emission is efficient (see Eq. 30-31 in [65]). Furthermore in [65], the authors took also into account the role of three-body interactions that could lead to coalescence binaries that have stalled (see Eq. 32 in [65]).

The SAM tracks also the evolution of BHs spins. The spin magnitudes can increase (decrease on average) under coherent (chaotic) accretion. Moreover, after the merger, the spin magnitude and orientation of the remnant BH is computed according to the formalism described in [77]. Since we model the GW signal with the PhenomHM waveform [72] (see discussion in Sec. VI), we neglect the binary precession in the GW analysis.

IV. EM COUNTERPART

In this section we discuss the models of the EM counterparts. We upgrade and improve T16 in what concerns the optical emission of the AGN, which can be detected with the Rubin Observatory, the radio jets, which can be detected with SKA, and the near-IR galaxy emission, which can be detected with ELT. To these counterpart types, we further add a model of the X-ray emission, to be detected by Athena.

The AGN bolometric luminosity $L_{\rm bol}$ is computed using the conventions in [78]:

$$L_{\rm bol} = \min\left(\epsilon_{\rm rad}\dot{M}_{\rm acc}c^2, L_{\rm Edd}\right) \qquad (1)$$

$$\dot{M}_{\rm acc} = \min\left(\frac{M_{\rm res}}{t_\nu}, \frac{L_{\rm Edd}}{\epsilon_{\rm rad}c^2}\right) \qquad (2)$$

where $\epsilon_{\rm rad}$ is the radiative efficiency, $\dot{M}_{\rm acc}$ is the mass accretion rate and $L_{\rm Edd} = 4\pi G M_{\rm tot} m_p c/\sigma_T$ is the Eddington luminosity, m_p the proton mass, σ_T the Thompson cross section.

The SAM predicts the amount of gas surrounding the binary at merger. If the binary does not accrete at Eddington, $\dot{M}_{\rm acc} = M_{\rm res}/t_\nu$, where t_ν is the viscous timescale. The latter is computed as

$$t_\nu \simeq \mathrm{Re}\, t_{\rm dyn} \qquad (3)$$

where $\mathrm{Re} \approx 10^3$ is the critical Reynolds number and $t_{\rm dyn}$ is the dynamical timescale

$$t_{\rm dyn} = G M_{\rm tot}/\sigma^3 \qquad (4)$$

with σ the velocity dispersion

$$\sigma = \sqrt{\frac{GM_{\rm b}}{3r_{\rm b}}} + \sqrt{\frac{GM_{\rm dyn}}{3(2r_{\rm h})}}. \quad (5)$$

In the above equation, $M_{\rm b}$ is the total bulge mass from stars and gas, $M_{\rm dyn}$ is the sum of the nuclear star cluster (NSC) mass and of the reservoir of gas feeding the MBH, $r_{\rm b}$ is the scale radius (assumed to be the same for gas and stars) and $r_{\rm h}$ is the NSC half light radii scale. The scale radius $r_{\rm b}$ is related to the half-light radius $R_{\rm eff}$ as $r_{\rm b} = R_{\rm eff}/1.8153$ and $R_{\rm eff}$ is computed following Eq. 32 in [79]. The NSC half light radii scale $r_{\rm h}$ is computed instead as [80]

$$r_{\rm h} = 3{\rm pc} \max\left[\sqrt{\frac{M_{\rm dyn}}{10^6\,{\rm M_\odot}}}, 1\right]. \quad (6)$$

If the binary accretes at Eddington, on the other hand, in Eq. 2 we limit the mass accretion rate to $\dot{M}_{\rm acc} = L_{\rm Edd}/\epsilon_{\rm rad}c^2$. The radiative efficiency $\epsilon_{\rm rad}$ describes the fraction of rest frame energy that can be extracted by the matter accreting onto the MBH and depends on the accretion efficiency η (the *maximum* amount of energy that can be extracted) and on the accretion geometry. In other words, the radiative efficiency takes into account that not all the available energy is radiated but it can be used to produce jets or winds and its value might depend on structural changes in the accretion disc. The accretion efficiency η depends on the spin of the MBH, a_{bh}, and ranges from 0.057 for Schwarzschild MBHs to ~ 0.4 for maximally rotating MBHs. If the merger is in a gas-rich environment with $M_{\rm res} > M_{\rm tot}$, we assume

$$\eta(a) = 1 - E_{\rm ISCO}(a), \quad (7)$$

where $E_{\rm ISCO}$ is the specific energy around a rotating MBH with spin $a \in [-1, 1]$ [81]. If the binary is in a dry environment with $M_{\rm res} < M_{\rm tot}$, the accretion might happen in prograde and retrograde orbits with the same probability so the accretion efficiency becomes

$$\eta(a) = 1 - \frac{1}{2}\left[E_{\rm ISCO}(a) + E_{\rm ISCO}(-a)\right]. \quad (8)$$

In our approach we choose $a = \chi_1$ where χ_1 is the spin component of the primary MBH along the angular momentum. The radiative efficiency is then computed as

$$\epsilon_{\rm rad} = \begin{cases} \eta(a) & \dot{m} \geq \dot{m}_{\rm cr} \\ \eta(a) \times \left(\frac{\dot{m}}{\dot{m}_{\rm cr}}\right) & \dot{m} < \dot{m}_{\rm cr} \end{cases} \quad (9)$$

being $\dot{m}_{\rm cr} = 0.01$ and $\dot{m} = \eta M_{\rm res}c^2/(L_{\rm Edd}t_\nu)$ [78].

From the bolometric luminosity we can compute the absolute magnitude through [82]

$$M_{\rm band} = M_{\rm band,\odot} - 2.5 \log_{10}\left(\frac{1}{BC}\frac{L_{\rm bol}}{L_\odot}\right), \quad (10)$$

where $M_{\rm band,\odot}$ is the absolute magnitude of the Sun in a certain band, $L_\odot$ the Sun luminosity and BC the bolometric correction. From the absolute magnitude we can infer the apparent magnitude $m_{\rm band}$:

$$m_{\rm band} = M_{\rm band} + 5 \log_{10}\left(\frac{d_L}{\rm pc}\right) - 5 + k_{\rm band}(z), \quad (11)$$

where the k-correction $k_{\rm band}(z)$ takes into account how the galaxy's radiation is redshifted during the propagation from the source to the observer. In the optical band, we assume that the AGN spectrum is flat in νf_ν (see Fig.1 in [83]), i.e $f_\nu \sim \nu^{-1}$. This approximation leads to $k_{\rm band} = 0$ so we neglect the last term in Eq. 11.

A. The Vera C. Rubin Observatory

The Rubin Observatory is an optical telescope with a $8.4\,{\rm m}$ mirror for observations in the u, g, r, i, z, y bands and $9.6\,{\rm deg}^2$ field of view (FOV).

We envisage two possible counterpart detection strategies with the Rubin Observatory. The Rubin Observatory Legacy Survey of Space and Time (LSST) is expected to reach a final depth of $m \sim 27.5$ after 10 years of operations in r band [56]. We therefore assume that we can identify the counterpart in the archival data, searching for a possible modulation due to the proper motion of the binary before merger. Moreover, LSST will finish its current scientific objectives in $\sim$2032, so we do not exclude the possibility to carry a target of opportunity (ToO) observation for LISA candidates with the Rubin Observatory with an observation time of few hours. For simplicity we assume the same apparent magnitude threshold for the ToO as for the entire survey. We also use the subscript 'Rubin' when the equations refer both to the LSST or ToO case.

For the AGN spectral energy distribution (SED) in optical bands we assume a bolometric correction $BC = 10$ (Fig. 2 in [83] shows that the bolometric correction is almost constant around 10 in optical bands) and compute the absolute magnitude following Eq. 10 as

$$M_{\rm AGN,Rubin} = 4.64 - 2.5 \log_{10}\left(\frac{0.1L_{\rm bol}}{L_\odot}\right), \quad (12)$$

where we have inserted $M_{\rm band,\odot} = 4.64$, the Sun AB magnitude in r band [84]. The apparent magnitude m becomes then

$$m_{\rm AGN,Rubin} = M_{\rm AGN,Rubin} + 5 \log_{10}\left(\frac{d_L}{\rm pc}\right) - 5. \quad (13)$$

We fix the threshold magnitude for detection with LSST and with ToO with the Rubin Observatory at $m_{\rm AGN,Rubin,\,lim} = 27.5$ and claim detection of the multimessenger candidate if

$$m_{\rm AGN,Rubin} < m_{\rm AGN,Rubin,\,lim}. \quad (14)$$

Once the source is identified, we can get an accurate redshift determination via photometric measurements, with error on the redshift $\Delta z = 0.031(1 + z)$ [85]. Concerning the sky localization threshold, we adopt the value of $\Delta\Omega = 10\,\mathrm{deg}^2$, close to the Rubin Observatory FOV.

B. The Square Kilometre Array

When LISA will be taking data, the SKA will be the largest radio telescope on Earth, with more than a square kilometer of collecting area. We therefore include the possibility of ToO with SKA in the proposed scenarios for EM counterpart detection.

During the MBHB merger, the interaction between the surrounding plasma and the magnetic fields is expected to produce radio emission [53, 86–88]. In particular, the motion of the binary is expected to twist the magnetic field lines, leading to flare emissions, while the Blandford-Znajeck effect [89] could be responsible for strong radio jets, depending on the amount of accreted material.

Following T16, we model the flare emission as

$$L_{\mathrm{flare}} = \frac{\epsilon_{\mathrm{Edd}}\epsilon_{\mathrm{radio}}}{q^2} L_{\mathrm{Edd}} \qquad (15)$$

where $\epsilon_{\mathrm{Edd}} = L_{\mathrm{bol}}/L_{\mathrm{Edd}}$ is the Eddington ratio, $q = m_1/m_2$ is the binary mass ratio and $\epsilon_{\mathrm{radio}} = 0.1$ is the portion of bolometric EM radiation emitted in radio band. In our model, ϵ_{Edd} is either computed from the reservoir amount of gas surrounding the binary at merger, with a floor value of $\epsilon_{\mathrm{Edd}} = 0.02$, or it is set equal to 1 if the binary accretes at Eddington.

Furthermore, the jet luminosity is modelled as [14, 90]

$$L_{\mathrm{jet}} = \begin{cases} 0.8 \times 10^{42.7}\,\mathrm{erg\,s^{-1}}\,m_9^{0.9}\left(\frac{\dot{m}_{\mathrm{jet}}}{0.1}\right)^{6/5}\left(1 + 1.1a_1 + 0.29a_1^2\right), & \mathrm{if}\ 10^{-2} \leq \epsilon_{\mathrm{Edd}} \leq 0.3 \\ 3 \times 10^{45.1}\,\mathrm{erg\,s^{-1}}\,m_9\left(\frac{\dot{m}_{\mathrm{jet}}}{0.1}\right)g^2\left(0.55f^2 + 1.5fa_1 + a_1^2\right) & \mathrm{otherwise} \end{cases} \qquad (16)$$

where $m_9 = m_1/(10^9\,\mathrm{M_\odot})$, $\dot{m}_{\mathrm{jet}} = \dot{M}_{\mathrm{acc}}/(22\,m_9\,\mathrm{M_\odot\,yr^{-1}})$ is the accretion rate in Eddington limit unit, a_1 is the spin magnitude of the primary black hole, $f = 1$ and $g = 2.3$ are dimensionless parameters regulating the angular velocity and azimuthal magnetic field of the system. The jet luminosity is the only emission that could reach $\sim 10L_{\mathrm{Edd}}$. L_{flare}, as well as the other emissions considered in this work (in optical and X-ray) are Eddington limited by construction.

Radio jets are expected to be beamed with an opening angle $\theta = 1/\Gamma$, Γ being the Lorentz factor [91, 92]. Two competing effects come into play:

- If the line of sight is outside the cone angle of the jet, the radio emission cannot be detected;

- Collimated emission increases the flux received from the observer, allowing for the detection of fainter and farther systems.

For typical AGNs, $\Gamma \simeq 5 - 15$ leading to a corresponding opening angle of $\theta \simeq 6°$. Yuan *et al.* 92, however, assumed instead a fiducial value of $\Gamma = 2$ based on the simulations of [86], corresponding to $\theta \simeq 30°$. If the emission is beamed, the luminosity increases as [93]

$$L_{\mathrm{beamed}} = L_{\mathrm{isotropic}}\delta^n\,, \qquad (17)$$

$$\delta = \Gamma^{-1}\left(1 - \beta\cos\iota\right)^{-1}\,, \qquad (18)$$

$$\beta = \sqrt{1 - \frac{1}{\Gamma^2}}\,, \qquad (19)$$

where β is the beam speed in the AGN frame in units of the speed of light, ι is the inclination angle between the binary angular momentum and the line of sight and $n = 2$ is an index describing the geometry and spectral index of the jet. For simplicity, we assume that the jet is aligned with the binary angular momentum. We consider three scenarios for the radio counterparts:

- 'Isotropic flare': the flare emission is isotropic and the jet is collimated with $\Gamma = 2$. In this case we compute the total luminosity in the radio band as

$$L_{\mathrm{radio}} = \begin{cases} L_{\mathrm{flare}} + L_{\mathrm{jet}}^{\mathrm{beamed}} & \mathrm{if}\ \iota < 30°\ \mathrm{or}\ \iota > 150° \\ L_{\mathrm{flare}} & \mathrm{otherwise;} \end{cases}$$
$$(20)$$

- '$\Gamma 2$': both the jet and the flare are beamed with $\Gamma = 2$. The total luminosity is the sum of the flare and jet collimated emissions, or it is zero if we are outside the cone:

$$L_{\mathrm{radio}} = \begin{cases} L_{\mathrm{flare}}^{\mathrm{beamed}} + L_{\mathrm{jet}}^{\mathrm{beamed}} & \mathrm{if}\ \iota < 30°\ \mathrm{or}\ \iota > 150° \\ 0 & \mathrm{otherwise;} \end{cases}$$
$$(21)$$

- '$\Gamma 10$': the same as '$\Gamma = 2$' but setting $\Gamma = 10$. The total luminosity is computed as:

$$L_{\mathrm{radio}} = \begin{cases} L_{\mathrm{flare}}^{\mathrm{beamed}} + L_{\mathrm{jet}}^{\mathrm{beamed}} & \mathrm{if}\ \iota < 6°\ \mathrm{or}\ \iota > 174° \\ 0 & \mathrm{otherwise.} \end{cases}$$
$$(22)$$

For each of the three scenarios above, we use L_{radio} to calculate the flux

$$F_{\mathrm{radio}} = \left(\frac{L_{\mathrm{radio}}}{\mathrm{erg/s}}\right)\left(\frac{d_L}{\mathrm{cm}}\right)^{-2}\,. \qquad (23)$$

To detect the multimessenger candidate with SKA, we impose that

$$F_{\rm radio} \geq 4\pi 10^{-20} F_{\rm lim}^{\rm SKA}, \tag{24}$$

where $F_{\rm lim}^{\rm SKA} = \nu_{\rm SKA} F_{\nu,\,\rm lim}^{\rm SKA}$ is the minimum flux detectable with SKA, $\nu_{\rm SKA} = 1.7\,{\rm GHz}$ is the typical frequency at which we expect the bulk emission and $F_{\nu,\,\rm lim}^{\rm SKA} = 1\mu{\rm Jy}$ is the flux limit in the same band. The final configuration of SKA is expected to reach this sensitivity in a FOV of $\sim 10\,{\rm deg}^2$ and ~ 10 minutes of integration time. We note that this is an optimistic approach that assumes that all the Poynting flux is converted into jet radio emission, which is not guaranteed [94].

Note that the radio jet luminosity can be also computed from the fundamental plane relation [95]. Ongoing work is evaluating how the fundamental plane relation compares with Eq. 16 [96]. Here we have verified that both methods to calculate $L_{\rm jet}$ lead to a similar number of radio EM counterparts, since the flare luminosity on its own is generally detectable with an SKA flux limit of $1~\mu{\rm Jy}$.

For SKA's sky localization threshold we adopt the same limit as in T16, i.e. $\Delta\Omega = 10\,{\rm deg}^2$.

C.   Athena

Being the 2nd Large class mission of the European Space Agency, the X-ray telescope Athena will observe the most energetic events from the high-redshift Universe thanks to its Wide Field Imager (WFI) with a FOV of $0.4\,{\rm deg}^2$ [97].

Together with the radio and optical/IR emission, MBHs produce copious amount of X-ray radiation. In this work we consider a late-stage X-ray emission produced by the gas accretion into the newly formed remnant MBH, possibly leading to a re-brightening of the source [42]. The increase in X-ray flux would allow the identification of the binary, among the possible X-ray transient candidates in the LISA sky localization area. In addition, internal shocks between the inner disk, closer to the remnant BH, and the outer disk portion are expected to produce X-ray emissions, even if they might be too faint to be detectable [54].

We focus on X-ray emission in the soft band, between $0.5-2\,{\rm keV}$, because Athena is most sensitive in this band. The bolometric correction in the X-ray band is given by [83]

$$L_{\rm X} = \frac{L_{\rm bol}}{c_1 \left(\frac{L_{\rm bol}}{10^{10}L_\odot}\right)^{k_1} + c_2 \left(\frac{L_{\rm bol}}{10^{10}L_\odot}\right)^{k_2}} \tag{25}$$

where $L_{\rm bol}$ is defined in Eq. (1), $c_1 = 5.712$, $k_1 = -0.026$, $c_2 = 17.67$ and $k_2 = 0.278$. For the MBH accretion, we consider two possibilities:

- The MBH accretion remains at the same level as before the merger, and we estimate it with the left term in Eq. 2;

- The sudden disappearance of the torques from the binary leads to the infall of the gas in the disc, and accretion reaches the Eddington luminosity.

The X-ray flux is

$$F_{\rm X} = \left(\frac{L_{\rm X}}{{\rm erg\,s}^{-1}}\right)\frac{1}{4\pi(d_L/{\rm cm})^2} \tag{26}$$

where $L_{\rm X}$ is the X-ray luminosity calculated with Eq. 25.

We define two possible detection strategies with Athena. As the exposure time increases, Athena will be able to detect fainter sources. We set a reasonable maximum exposure time of $300\,{\rm ks}$ [98] and two different flux limits, chosen together with appropriate LISA sky localization cuts:

- If Athena observes the same FOV of $\Delta\Omega = 0.4\,{\rm deg}^2$ for the entire exposure time, we assume a limiting flux of $F_{\rm X,\,lim} = 4 \times 10^{-17}{\rm erg\,s}^{-1}\,{\rm cm}^{-2}$;

- Athena could also explore a larger sky localization, limited however to brighter flux. For this second strategy, we choose a flux limit of $F_{\rm X,\,lim} = 2 \times 10^{-16}{\rm erg\,s}^{-1}\,{\rm cm}^{-2}$ and a sky localization threshold of $\Delta\Omega = 2\,{\rm deg}^2$. This corresponds to ~ 5 tiles in the sky, each covered twice.

We claim detection of the X-ray counterpart whenever

$$F_{\rm X} \geq F_{\rm X,\,lim}. \tag{27}$$

In the second detection strategy, we do not take into account the slew time necessary to repoint Athena, because the tiles are next to each other (see Section VIII for more details on LISA sky localization). Finally, for simplicity, in this work we do not consider the possibility of pre-merger modulated X-ray emission [51].

D.   The Extremely Large Telescope

Since one of our aims is to use LISA EMcps to test the expansion of the Universe, we also enforce the redshift measurement of the GW source. Radio and X-ray observations will identify the sky localization of the merger event within the LISA sky area, while the redshift information relies on emission lines at optical/IR wavelengths. The simplest option is to detect the emission of the galaxy hosting the MBHB merger.

By the time LISA will be operative, the Extremely Large Telescope (ELT) will be available and other similarly large telescopes, such as the Giant Magellan Telescope and the Thirty Meter Telescope, are also under discussion. We focus the discussion here on ELT as a case study. Among the ELT instruments, the spectrograph MICADO [99] will allow redshift measurements in the IYJHK band with spectroscopy between the OH lines down to an apparent magnitude of 27.2 and with imaging with advanced filters down to an apparent magnitude of

31.3. These values correspond to the 5σ sensitivity for isolated point sources in 5 hours of observation.

We compute the galaxy luminosity in K band as

$$L_{\mathrm{gal,ELT}} = \frac{1}{\Upsilon} L_\odot M_{\mathrm{stars}}, \qquad (28)$$

where M_{stars} is the total mass of the stars in the disk and bulge, and Υ is a fiducial mass-to-light ratio. The latter requires information, such as the star formation history and the metallicity, which is not provided by the MBHB catalogues. We therefore adopt a simple correction. For systems at $z \geq 3$ we assume $\Upsilon = 0.03$, because galaxies are dominated by young stellar populations with low metallicity, so that radiation in the H, J and K bands in the observer's rest frame actually comes from the blue part of the source's restframe spectral energy distribution. For systems with $z < 3$, we assume $\Upsilon = 0.1$ because the stars are older and more metallic, and the observer-frame wavelength comes from the optical part of the restframe spectrum. Following Eqs. 10-11 the apparent magnitude is

$$M_{\mathrm{gal,ELT}} = 5.08 - 2.5\log_{10}\left(\frac{L_{\mathrm{gal,ELT}}}{L_\odot}\right) \qquad (29)$$

$$m_{\mathrm{gal,ELT}} = M_{\mathrm{gal,ELT}} + 5\log_{10}\left(\frac{d_L}{\mathrm{pc}}\right) - 5. \qquad (30)$$

For systems with $m_{\mathrm{gal,ELT}} \leqslant 27.2$ we expect that a spectroscopic redshift measurement will be doable, with precision $\Delta z = 10^{-3}$. This relies on the reasonable assumption that galaxies are star-forming and therefore emission lines are present, since the typical galaxies in the catalogue have masses of $10^7 - 10^{11}\,\mathrm{M}_\odot$ at $z > 1$.

For galaxies with apparent luminosity $27.2 < m_{\mathrm{gal,ELT}} < 31.3$, the redshift measurement and its uncertainty are less straightforward, and depend on the actual galaxy redshift. For high-redshift sources at $z \gtrsim 6.5$, MICADO will be able to detect the Lyman-α (1215.67 Å) break in the I-band, enabling the redshift measurement with error $\Delta z = 0.2$ [100]. For galaxies between $5 < z \lesssim 6.5$, the redshift measurement is more challenging, also because they might resemble ultra-faint galaxies at $z < 0.5$. First of all, we assume that this degeneracy can be broken by using the information on the luminosity distance inferred from the GW measurement, in the context of a given cosmology. Even though the aim is to use the redshift information to *infer* the cosmology, we believe that assuming the cosmology with the only aim of discriminating between sources at $z < 0.5$ and $z > 5$ will not substantially bias the cosmological analysis. Furthermore, the Roman Space Telescope [101] will observe also in the R band, corresponding to the Lyman-α break for sources at $z \gtrsim 5.2$. There will be the possibility to perform ToO follow-up within 2 weeks from notification, but with a limiting magnitude of 28.5 in 1hr (but longer observing time might be possible) [102]. Consequently, assuming observations with the Roman Telescope combined with the luminosity distance information

	$m_{\mathrm{gal,ELT}} < 27.2$	$27.2 < m_{\mathrm{gal,ELT}} < 31.3$
$z \leqslant 0.5$		No redshift information
$0.5 < z \leqslant 5$	$\Delta z = 10^{-3}$	$\Delta z = 0.5$
$z > 5$		$\Delta z = 0.2$

from GW sources, we claim that the redshift identification will also be possible for sources with $z > 5$ with uncertainty $\Delta z = 0.2$, thanks to the detection of the Lyman-α break.

Finally, for systems with $27.2 < m_{\mathrm{gal,ELT}} < 31.3$ and $0.5 < z < 5$, the photometric redshift can be obtained thanks to the Balmer break, with the penalty, however, of observing only in near-IR bands (from I to K with MICADO): missing the optical and UV parts of the spectrum will increase the photometric redshift errors. Unfortunately, we did not find reliable estimates of the redshift measurement error in the literature for this kind of sources with only near-IR observations. Therefore, we adopt an agnostic estimate and set $\Delta z = 0.5$ for these systems [103].

Tab. I summarizes the errors adopted for redshifts measured based on the host galaxy spectra using ELT and the Roman Space Telescope. Note that we do not take into account the possibility of observations with the James Webb Telescope [104], because it may not overlap with LISA.

V. AGN OBSCURATION

The gas and dust surrounding the MBHs are expected to absorb a fraction of the EM emission, reducing the observed flux and the number of detected systems [105, 106]. Similarly, interstellar dust and intergalactic gas can absorb starlight in the galaxy.

In this section we present how we model obscuration at X-ray and optical wavelengths, affecting the detection of the EM emission with Athena and with the Rubin Observatory respectively. We also model absorption from the galaxy gas, affecting the optical emission detectable with ELT (used to determine the EMcp redshift). Below we detail the two in turns, starting with the AGN obscuration.

Evaluating the fraction of obscured AGNs with respect to the total number of AGNs is generally challenging, and there is no consensus on how it evolves as function of luminosity or redshift [107–109]. Nevertheless, we adopt here a recent modelling of obscuration in X-rays, taken from [110]. We further include absorption in the optical, assuming that dust follows gas.

For each event, we start by computing the hydrogen column density around the MBH from equations 3-6 in [110]. For completeness, we detail here the entire procedure. We introduce the $\psi(L_{X,h}, z)$ parameter that corresponds to the fraction of absorbed Compton-thin AGNs

with respect to the total Compton-thin AGNs:

$$\psi(L_{X,h}, z) = \min[\psi_{\max}, \max[\psi_{43.75}(z) \\ - \beta(\log_{10}[L_{X,h}/(\text{erg/sec})] - 43.75), \psi_{\min}]], \quad (31)$$

where $\psi_{\max} = 0.84$, $\psi_{\min} = 0.2$, $\beta = 0.24$. The source redshift z and the X-ray luminosity in hard band $L_{X,h}$ are the only input parameters, and we infer them from the information contained in the MBHBs catalogues. To compute $L_{X,h}$ we use Eq. 25, with coefficients $c_1 = 4.073$, $k_1 = -0.026$, $c_2 = 12.60$ and $k_2 = 0.278$ [83]. The quantity $\psi_{43.75}(z)$ represents the fraction of Compton-thin AGNs with $\log_{10}[L_{X,h}/(\text{erg/s})] = 43.75$ at z, and it takes the form

$$\psi_{43.75}(z) = \begin{cases} 0.43\,(1+z)^{0.48} & z < 2, \\ 0.43\,(1+2)^{0.48} & z \geq 2. \end{cases} \quad (32)$$

The distribution of the hydrogen column density N_H can be expressed as

$$f(L_{X,h}, z; N_H) = \quad (33)$$
$$= \begin{cases} 1 - \frac{2+\epsilon}{1+\epsilon}\psi(L_{X,h}, z) & 20 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 21 \\ \frac{1}{1+\epsilon}\psi(L_{X,h}, z) & 21 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 23 \\ \frac{\epsilon}{1+\epsilon}\psi(L_{X,h}, z) & 23 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 24 \\ \frac{f_{\text{AGN}}}{2}\psi(L_{X,h}, z) & 24 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 26 \end{cases}$$
$$\text{if } \psi(L_{X,h}, z) < \frac{1+\epsilon}{3+\epsilon},$$

and

$$f(L_{X,h}, z; N_H) = \quad (34)$$
$$= \begin{cases} \frac{2}{3} - \frac{3+2\epsilon}{3+3\epsilon}\psi(L_{X,h}, z) & 20 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 21 \\ \frac{1}{3} - \frac{\epsilon}{3+3\epsilon}\psi(L_{X,h}, z) & 21 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 22 \\ \frac{1}{1+\epsilon}\psi(L_{X,h}, z) & 22 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 23 \\ \frac{\epsilon}{1+\epsilon}\psi(L_{X,h}, z) & 23 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 24 \\ \frac{f_{\text{AGN}}}{2}\psi(L_{X,h}, z) & 24 \leq \log_{10}\left(\frac{N_H}{\text{cm}^2}\right) < 26 \end{cases}$$
$$\text{if } \psi(L_{X,h}, z) \geq \frac{1+\epsilon}{3+\epsilon},$$

where $f_{\text{AGN}} = 1$ is the fraction of Compton-thick AGNs to the absorbed Compton-thin AGNs, and $\epsilon = 1.7$. Given the hard X-ray luminosity and redshift of each binary in the catalogues, we derive, using the equations above, the corresponding hydrogen column density distribution f. We then sample this distribution, in order to extract a value for the hydrogen column density N_H surrounding the binary.

We then need to evaluate the absorption. We consider the soft X-ray band and compute the luminosity after absorption as

$$L_{X,\text{abs}} = L_X\,e^{-\tau_X} \quad (35)$$

with $\tau_X = \sigma_X N_H$, where σ_X is the X-ray cross section, for which we take, following [111],

$$\sigma_X = \left(120.6 + 169.3(E/\text{keV}) - 47.7(E/\text{keV})^2\right) \\ (E/\text{keV})^{-3} \times 10^{-24}\text{cm}^2. \quad (36)$$

For the energy, we choose $E = 1\,\text{keV}$, in the middle of the soft X-ray band. The X-ray flux after obscuration can be computed from Eq. 26, substituting the original L_X with $L_{X,\text{abs}}$. We are applying a statistical correction based on observational samples of AGNs. An alternative approach could be to model obscuration based on the intrinsic properties of the source. For instance Ricci et al. 112 found in a low-redshift AGN sample a relation between the obscuration fraction and the Eddington ratio that can be used to compute the corresponding hydrogen column density, under the assumption that radiation pressure is the dominant factor modulating obscuration, caused by material very close to the black hole. In high redshift sources, however, the interstellar medium can also contribute to obscuration [113, 114]. We prefer therefore to base our correction on empirical results based on observational samples covering a broad range of redshifts.

Similarly as for X-ray, the AGN emission in the optical band after absorption is

$$L_{\text{bol,abs}} = L_{\text{bol}}\,e^{-\tau_{\text{opt}}}, \quad (37)$$

where L_{bol} is the bolometric luminosity from Eq. 1, and $\tau_{\text{opt}} = \sigma_{\text{opt}} N_{\text{dust}}$ is the optical depth, with σ_{opt} the optical cross section and N_{dust} the dust column density. In order to evaluate the optical depth, we start from the galaxy mass-metallicity relation Z_{gas} [115]

$$\log_{10}(Z_{\text{gas}}/Z_\odot) = 0.35[\log_{10}(M_{\text{stars}}/M_\odot) - 10] \\ + 0.93\text{e}^{-0.43z} - 1.05. \quad (38)$$

From the hydrogen column density, calculated as described above, and Z_{gas}, one obtains the dust column density [116]:

$$N_{\text{dust}} = \frac{Z_{\text{gas}}}{0.02}N_H \quad (39)$$

and the optical cross section, given by

$$\sigma_{\text{opt}} = \sigma_0 \sum_{i=1}^{7} F(\lambda/\lambda_i, a_i, b_i, p_i, q_i), \quad (40)$$

where $\sigma_0 = 3 \times 10^{-22}\,\text{cm}^2$ and the fitting function F as

$$F(x, a, b, p, q) = \frac{a}{x^p + x^{-q} + b}, \quad (41)$$

with coefficients $\lambda_i, a_i, b_i, p_i, q_i$ reported in Tab. II. We choose to compute the cross section at a reference wavelength of $\lambda = 0.62\,\mu\text{m}$, at the center of r band. The magnitude after absorption can be inferred from Eq. 10, substituting L_{bol} with $L_{\text{bol,abs}}$.

At last, we turn to the absorption of the optical galactic emission from interstellar dust and intergalactic gas, to be accounted for in the detection of the host galaxy by ELT. Based on Fig. 1 in [117], we adopt a constant hydrogen column density $\log_{10}(N_H/\text{cm}^2) = 22$ and compute the absorbed luminosity in K band following Eqs. 37-39,

TABLE II. Coefficients entering the optical cross section in Eq. 40

Term	$\lambda_i \, [\mu m]$	a_i	b_i	p_i	q_i
1	0.042	185	90	2	2
2	0.08	27	15.5	4	4
3	0.22	0.005	-1.95	2	2
4	9.7	0.01	-1.95	2	2
5	18	0.012	-1.8	2	2
6	25	0.03	0	2	2
7	0.067	10	1.9	4	15

with $\lambda = 2.2 \, \mu m$, at the center of K band. The absorption of the host galaxies emission does not affect significantly the EMcps rates, and it is always included in the following results.

VI. GRAVITATIONAL WAVE SIGNAL AND PARAMETER ESTIMATION

The waveform of a MBHB with aligned spins and circular orbit depends on 11 parameters: the primary and secondary source-frame masses, m_1 and m_2, the two spins magnitudes along the orbital angular momentum, χ_1 and χ_2, the sky latitude β and the longitude λ, the luminosity distance d_L, the inclination ι of the binary angular momentum with respect to the line of sight, the phase at coalescence (or at a reference frequency) ϕ, the time at coalescence t_c, and the polarization angle ψ. We can also define the mass-ratio $q = m_1/m_2 \geqslant 1$ and the chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$. We model the GW signal with the inspiral-merger-ringdown waveform PhenomHM [72] that ignores the binary spins precession but includes higher order harmonics.

If the source is located at cosmological distance, the signal is affected by the expansion of the Universe. As a consequence, in the detector-frame we measure redshifted quantities, i.e. $\mathcal{M}_z = \mathcal{M}(1 + z)$, where z is the source redshift. However, in this work we will refer to rest-frame quantities, unless otherwise stated. We assume a fiducial ΛCDM cosmology with $h = 0.6774$, $\Omega_m = 0.3075$ and $\Omega_\Lambda = 0.6925$.

For each system we compute the signal-to-noise ratio as

$$\text{SNR} = 4 \int_{f_{\min}}^{f_{\max}} \frac{|\tilde{h}(f)|^2}{S_n(f)} df \,, \tag{42}$$

where $\tilde{h}(f)$ corresponds to the Fourier transform of the time-domain signal, and $S_n(f)$ is the noise power spectral density, for which we take the estimate "SciRDv1" described in [118]. We set $f_{\min} = 10^{-5} \, \text{Hz}$ and $f_{\max} = 0.5 \, \text{Hz}$. We assume 5 years of overall mission duration, with 80% duty cycle. We add to the LISA noise PSD the one of the confusion background from unresolved galactic binaries, according to the fits presented in [119]. The amplitude of the background is taken for three years of mission duration, as a representative average value over the total duration of 5 years.

For each event we obtain the posterior distribution $p(\overline{\theta}|d)$ for a set of parameters $\overline{\theta}$ following Bayes theorem,

$$p(\overline{\theta}|d) = \frac{\mathcal{L}(d|\overline{\theta})\pi(\overline{\theta})}{p(d)} \,, \tag{43}$$

where $\mathcal{L}(d|\overline{\theta})$ is the likelihood of the realisation d with the parameters $\overline{\theta}$, $\pi(\overline{\theta})$ corresponds to our prior on the binary parameters, and $p(d) = \int d\overline{\theta}\mathcal{L}(d|\overline{\theta})\pi(\overline{\theta})$ is the evidence. In this work, we assume the so-called zero-noise approximation, i.e. $d = h(\overline{\theta})$ where h denotes the waveform. We refer the interested reader to Marsat et al. [71] for more detail.

The GW signal parameter estimation is performed using the response and likelihood code of [71], together with the parallel-tempered ensemble sampler ptemcee [120] for the Bayesian parameter estimation. We initialize our chains around the simulated signal with a covariance computed from the Fisher matrix, and use enriched proposals that allow to jump to the known potentially degenerate modes in the sky position, inclination and polarization [71]. For the prior distributions, we assume the sky position angles, inclination and polarization uniformly distributed over the sphere. For all the other parameters we adopt uniform flat priors.

We run the MCMC chain for each system for 2000 iterations with 64 walkers [121] and 10 temperatures [122]. This provides a first set of parameter posteriors. Among them, we are particularly interested in the sky area, to ensure the EM detection. We therefore rerun the chains, for 10^5 iterations, for all the systems whose the first run produced a posterior with an initial sky localization error of $5 < \Delta\Omega/\deg^2 < 40$ at 90% confidence level. This way, we ensure the convergence of the MCMC chains for the interesting systems. Moreover, for all the systems with error in the sky localization $\Delta\Omega < 10 \deg^2$ at 90% confidence level, i.e. those which we use in the rest of the paper, we further check the convergence of the parameter estimation for the parameters β, λ and d_L, studying the evolution of the chains as a function of the iterations. At the end of the process, we obtain three catalogues of EMcps for each of the MBHB astrophysical formation models, for which we are confident that the parameter estimation has converged.

VII. RESULTS

For each of the three MBHB formation astrophysical models described in Sections II and III, we use catalogues simulating 90 years of data. They contain a total number of 15546, 692 and 10700 MBHBs for Pop3, Q3d and Q3nd respectively. The following results are presented for 4 years of LISA observations, i.e. 5 years of mission duration with 80% duty cycle. All catalogues start at $z = 20$.

TABLE III. Average number of intrinsic and GW-detected (i.e., SNR > 10 in LISA) MBHBs in 4 yrs, for different astrophysical models.

	Total catalogue	SNR > 10
Pop3	690.9	129.3
Q3d	30.7	30.4
Q3nd	475.5	471.1

As described in Section II, we first select, among all the events in the catalogues, those with a detectable EM emission, on which we then run the GW signal parameter estimation, to extract the EMcps. As we shall demonstrate, the number of EMcps strongly depends on how the EM emission is modeled, on the specific instrument adopted to detect it, and on the detection strategy, other than on the astrophysical population model. Several choices for how to combine these variables are possible, leading to different configurations for the EMcps observations. To simplify the presentation of the results, we focus on two specific models, one maximising the number of EMcps, and one minimising it. They are defined as follow:

- Maximising

    - AGN obscuration neglected
    - Collimated jet emission with $\Gamma = 2$ and isotropic flare
    - Eddington accretion for X-ray emission

- Minimising

    - AGN obscuration included
    - Collimated jet and flare emissions with $\Gamma = 2$
    - Catalogue accretion for X-ray emission

For these two models, we present the rates of both multimessenger candidates and EMcps. We also discuss possible variations to the two selected models separately. Among all the variables, the jet opening angle is the factor that most affects the EMcps rates, while, e.g., taking the accretion from the catalogues or assuming it at Eddington does not change the rates significantly. This is because the SKA+ELT combination dominates the rates in the case of an isotropic flare emission, as we will show below.

A.   General distributions

In this section we discuss the distributions in redshift and chirp mass of the average number of MBHBs events, multimessenger candidates, and EMcps, in the two models labelled respectively "maximising" and "minimising" (c.f. the beginning of Section VII). Furthermore, we also report the average number of multimessenger candidates

and EMcps for the other observational scenarios, in Tables V and IV. The average numbers are intended in 4 yrs of observations with LISA, and are obtained by multiplying the total numbers provided by the 90 yrs of catalogues by 4/90.

In Fig. 1 we present the average number of merging binaries as a function of redshift. Models Pop3 and Q3nd predict a large fraction of mergers at $z \gtrsim 10$, while in the Q3d model all the systems merge at $z \lesssim 12$. Removing the systems with SNR < 10 in LISA leads to the loss of $\sim 80\%$ of high-redshift sources in the Pop3 catalogue, caused by their low mass (see Fig. 2). The systems of the Q3nd catalogue are on average more massive, therefore, the SNR cut does not alter their number. The systems of the Q3d catalogue are also all detected by LISA with SNR > 10: this is expected, since they have a mass distribution similar to Q3nd, and they merge at smaller redshifts. The average number of intrinsic and GW-detected events for each of the three astrophysical models is reported in Tab. III.

Among the systems with SNR > 10, we further select the multimessenger candidates, i.e. those with a detectable EM counterpart. In Fig. 1, we show their distributions in the maximising and minimising models.

The additional requirement of EM detectability selects systems at even smaller redshift: for all the three astrophysical scenarios, multimessenger candidates have $z < 10$. Within the maximising model, we predict in total 24.0 (35) {37.6} multimessenger candidates for Q3d (Pop3) {Q3nd} in 4 years. As expected, if we include obscuration and collimated radio emission, the multimessenger candidates number decreases to 3.6 (6.6) {4.2}, and only systems at $z \lesssim 8$ can be detected. This reduction of about $81 - 89\%$ in the number of multimessenger candidates with respect to the maximising model, is similar in all astrophysical models.

At last, we impose a cut in the sky localization of the systems, to select only the EMcps. We obtain 14.8 (6.4) {20.7} EMcps for Q3d (Pop3) {Q3nd} in 4 yrs in the maximising model, and nothing but 3.3 (1.6) {3.5} if we include AGN obscuration and collimated flare and jet emission (minimising model). The Pop3 scenario predicts the largest number of multimessenger candidates, however, only $\sim 20\%$ among them are promoted to EMcps: LISA will not localize these sources accurately enough, due to their intrinsic low chirp mass and high redshift. On the other hand, the Q3d and Q3nd models predict fewer multimessenger candidates, but $\sim 60\%$ among them are EMcps. Within the minimising model, even though the total number of both multimessenger candidates and EMcps decreases, the fraction of multimessenger candidates promoted to EMcps is higher for all astrophysical models: 24%, 92%, and 83% for Pop3, Q3d and Q3nd respectively, as opposed to 18%, 62%, and 55% in the maximising model. We interpret this fact as follows: including obscuration and collimated radio emission effectively removes the tails of the distributions, selecting the bulk of the "best" events: those with redshift low
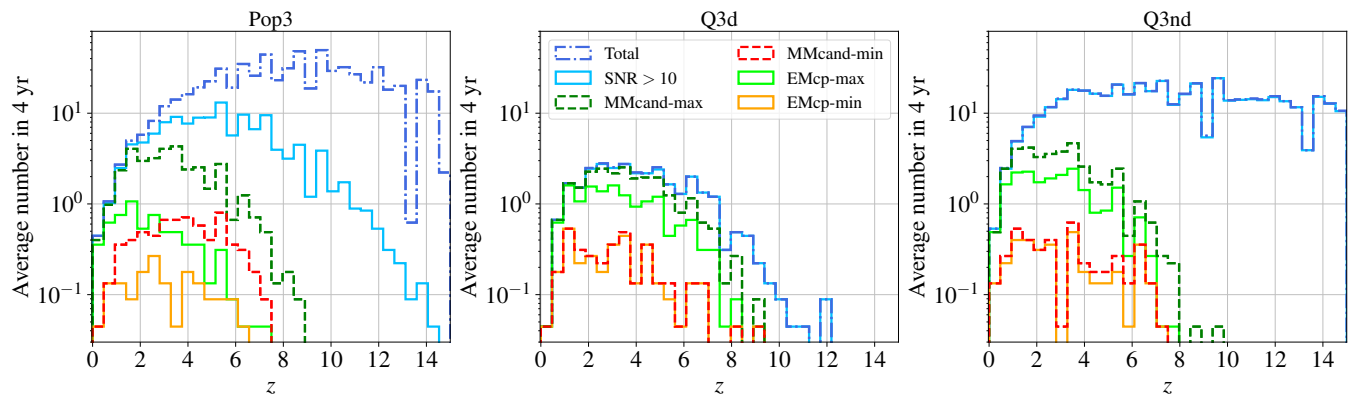
FIG. 1. Average number of *i)* MBHB mergers directly from the catalogues (dark blue dotted-dashed line), *ii)* MBHB mergers detected in LISA with SNR > 10 (light blue solid line), *iii)* multimessenger candidates (dark green and red dashed lines), and *iv)* EMcps (light green and yellow solid lines), as function of redshift, for the three astrophysical MBHBs formation models, and assuming 4 yrs of LISA observations. In particular, the dark green and the red dashed lines correspond respectively to the multimessenger candidates (c.f. definition in Section II) in the maximising model, without absorption and isotropic radio flare emission, and in the minimising model, with absorption and $\Gamma = 2$ (c.f. definitions at the beginning of Section VII). The light green and yellow lines correspond to the EMcps (c.f. definition in Section II) distributions also in the maximising and minimising models respectively. Applying the requirement of EM detectability and imposing the sky localization threshold select only the closest events, while including absorption and collimated radio emission decreases the overall number of both multimessenger candidates and EMcps.
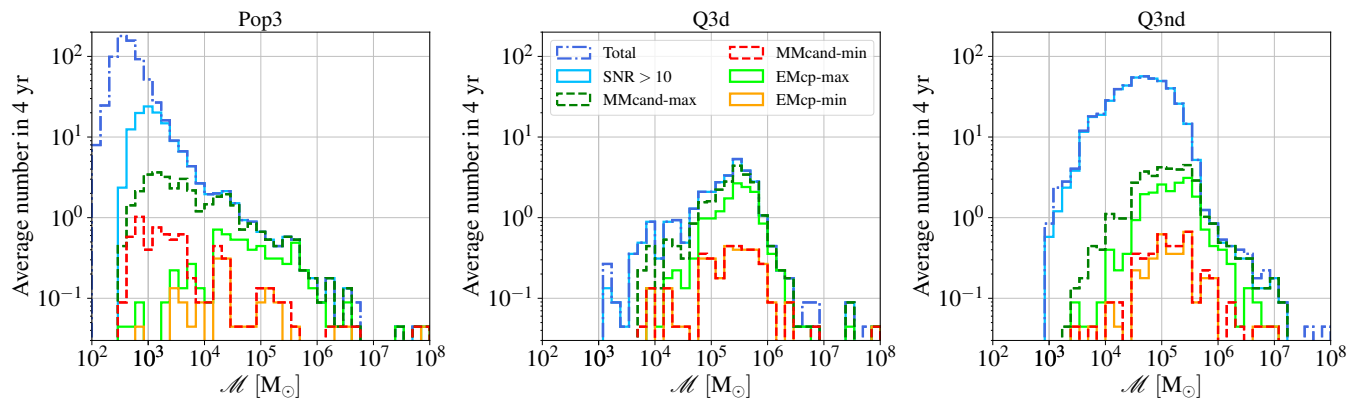

FIG. 2. Same as Fig. 1 as function of chirp mass $\mathcal{M}$. LISA sensitivity selects only systems with $10^4 \lesssim \mathcal{M}/\mathrm{M}_\odot \lesssim 10^6$ as EMcps.

enough to have good LISA parameter estimation, but high enough to be sufficiently numerous. Further imposing the sky localization cut has therefore a minor effect on this subset of events.

In Fig. 2 we report the same quantities as a function of the rest-frame chirp mass $\mathcal{M}$. While the distributions in the massive models (Q3d and Q3nd) peak at $\mathcal{M} \sim 10^5 - 10^6\,\mathrm{M}_\odot$, for the Pop3 model the peak of the distribution is at $\mathcal{M} \lesssim 10^3\,\mathrm{M}_\odot$, due to the different BH formation processes. The SNR > 10 cut therefore operates similarly to what already observed for the redshift distribution, i.e. it excludes low-mass events in the Pop3 scenario while leaving the Q3d and Q3nd practically unaffected.

In the maximising model, all the systems with $\mathcal{M} \gtrsim 10^5\,\mathrm{M}_\odot$ have detectable EM emission, while at lower masses a significant fraction of Pop3 and Q3nd can be

detected by LISA with SNR > 10 but do not have observable EM emission due to the low BH mass or high redshift of the systems. Adding the further requirement on the sky localization results in an overall rescaling of the multimessenger candidate distributions for the massive astrophysical models, while it selects only the heaviest binaries in the Pop3 model. As already observed for the distributions as a function of redshift, the reduction in the number of events when one includes obscuration and the collimated jet is higher than the one obtained when one imposes the cut in the sky localization.

In Fig. 3 we show the EMcps distributions for the Rubin Observatory, SKA+ELT and Athena+ELT strategies separately in the maximising scenario. SKA+ELT is the only combination to provide EMcps at $z \gtrsim 4$ while the Rubin Observatory and Athena+ELT can observe only
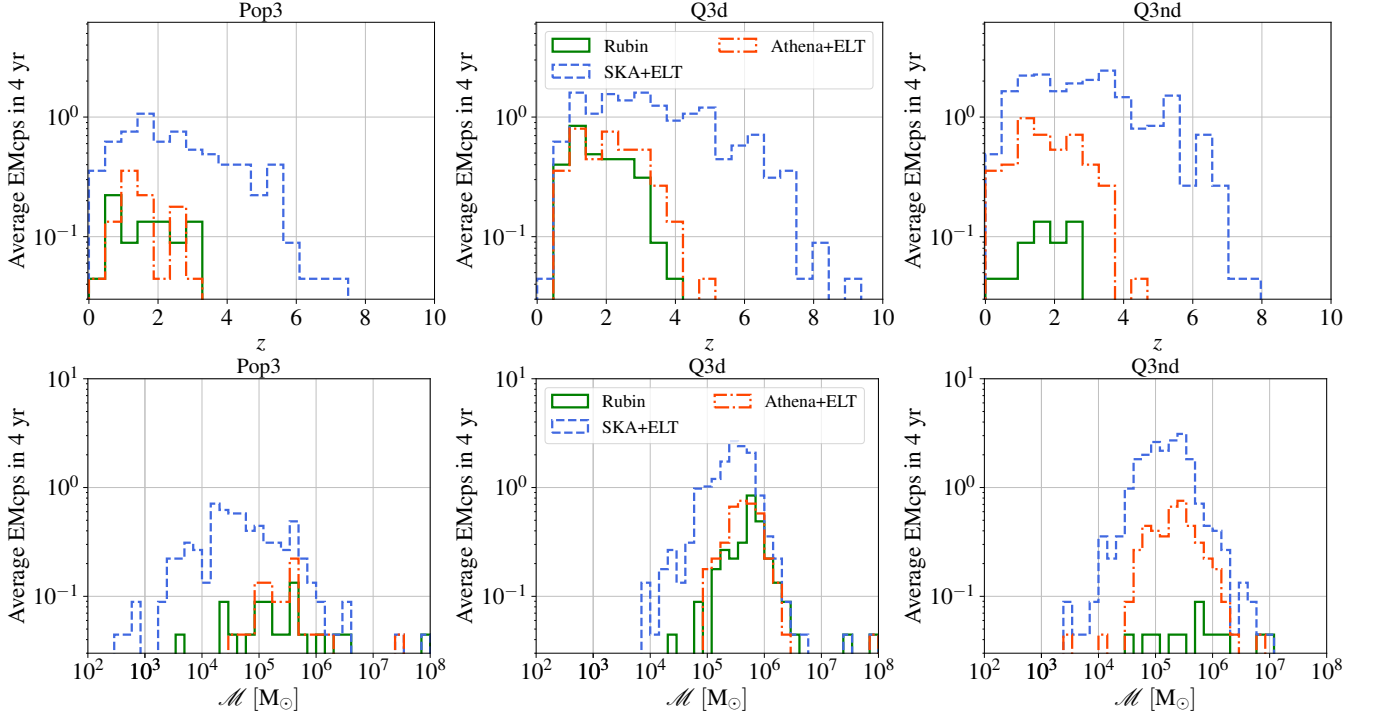
FIG. 3. Average number of EMcps in each observational scenario, as clarified by the legend, in the maximising case as function of redshift (upper panels) and chirp mass (lower panels) for the three astrophysical models assuming 4 yr of LISA time mission. SKA+ELT provides more EMcps in the maximising scenario thanks to the isotropic flare emission. We stress that we can not simply sum the y-axis for each instrument combination because the same system might be observed by different telescopes at the same time.

TABLE IV. Average number of multimessenger candidates in 4 yrs, in each observational scenario. Fluxes are in units of $\mathrm{erg\,s^{-1}\,cm^{-2}}$ (more detail in the text).

| | Rubin | SKA+ELT | | | Athena+ELT | | | | |
| | | Isotropic flare | Γ2 | Γ10 | Catalogue | | Eddington | | |
					$F_{\mathrm{X,lim}} = 4e\text{-}17$	$F_{\mathrm{X,lim}} = 2e\text{-}16$	$F_{\mathrm{X,lim}} = 4e\text{-}17$	$F_{\mathrm{X,lim}} = 2e\text{-}16$	
	1.3	34.4	6.27	0.13	2.35	0.62	3.95	1.82	Pop3
No-obsc.	3.33	24.0	2.89	0.04	5.42	1.64	8.53	3.02	Q3d
	0.84	34.5	3.78	0.04	1.6	0.44	15.9	6.31	Q3nd
	0.35	34.4	6.27	0.13	0.18	0.13	0.4	0.31	Pop3
Obsc.	0.8	24.0	2.89	0.04	0.35	0.17	0.53	0.13	Q3d
	0.49	34.5	3.78	0.04	0.27	0.09	1.42	0.53	Q3nd

TABLE V. Average number of EMcps in 4 yrs in each observational scenario. Fluxes are in units of $\mathrm{erg\,s^{-1}\,cm^{-2}}$ (more detail in the text). The average number of EMcps obtained combining observations with multiple facilities are reported in Tab. VI.

	Rubin	SKA+ELT			Athena+ELT				
		Isotropic flare	Γ2	Γ10	Catalogue		Eddington		
					$F_{\mathrm{X,lim}} = 4e\text{-}17$	$F_{\mathrm{X,lim}} = 2e\text{-}16$	$F_{\mathrm{X,lim}} = 4e\text{-}17$	$F_{\mathrm{X,lim}} = 2e\text{-}16$	
		$\Delta\Omega = 10\,\mathrm{deg}^2$			$\Delta\Omega = 0.4\,\mathrm{deg}^2$	$\Delta\Omega = 2\,\mathrm{deg}^2$	$\Delta\Omega = 0.4\,\mathrm{deg}^2$	$\Delta\Omega = 2\,\mathrm{deg}^2$	
	0.84	6.4	1.51	0.04	0.49	0.27	1.02	0.84	Pop3
No-obsc.	3.07	14.8	2.71	0.04	2.67	1.38	3.87	2.13	Q3d
	0.53	20.3	3.2	0.04	0.58	0.31	4.4	3.24	Q3nd
	0.13	6.4	1.51	0.04	0.04	0.04	0.13	0.17	Pop3
Obsc.	0.75	14.8	2.71	0.04	0.22	0.13	0.18	0.09	Q3d
	0.35	20.3	3.2	0.04	0.18	0.04	0.27	0.31	Q3nd

TABLE VI. Average number of EMcps in 4 yrs combining the different observational models. The case *without obscuration* (*with obscuration*) is reported in the left (right) table. Fluxes are in units of $\mathrm{erg\,s^{-1}\,cm^{-2}}$. The EMcps rates corresponding to the *maximising* (*minimising*) case are in boldface in the left (right) table.

Left table (without obscuration): Rubin $\Delta\Omega = 10\,\mathrm{deg}^2$ No obsc.; SKA $\Delta\Omega = 10\,\mathrm{deg}^2$; Athena No obsc.; columns Pop3 / Q3d / Q3nd.

SKA	Athena	Flux	Pop3	Q3d	Q3nd
Isotropic flare	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	6.4	14.8	20.4
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	6.4	14.8	20.4
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	**6.4**	**14.8**	**20.7**
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	6.4	14.8	20.6
$\Gamma2$	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	2.31	6.18	3.9
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	2.18	5.5	3.6
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	2.8	7.0	6.9
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	2.67	6.0	5.9
$\Gamma10$	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	1.07	4.04	0.9
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	0.9	3.2	0.58
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	1.6	5.2	4.7
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	1.5	3.9	3.5

Right table (with obscuration): Rubin $\Delta\Omega = 10\,\mathrm{deg}^2$ with obsc.; SKA $\Delta\Omega = 10\,\mathrm{deg}^2$; Athena with obsc.; columns Pop3 / Q3d / Q3nd.

SKA	Athena	Flux	Pop3	Q3d	Q3nd
Isotropic flare	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	6.4	14.8	20.4
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	6.4	14.8	20.4
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	6.4	14.8	20.4
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	6.4	14.8	20.4
$\Gamma2$	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	**1.6**	**3.3**	**3.5**
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	1.6	3.3	3.5
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	1.8	3.5	3.6
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	1.8	3.4	3.7
$\Gamma10$	Catalogue	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	0.18	0.80	0.49
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	0.18	0.80	0.40
	Eddington	$F_{\mathrm{X,lim}}=4\text{e-}17$, $\Delta\Omega=0.4\,\mathrm{deg}^2$	0.31	0.98	0.67
		$F_{\mathrm{X,lim}}=2\text{e-}16$, $\Delta\Omega=2\,\mathrm{deg}^2$	0.35	0.84	0.71

the closest events with the latter reaching slightly higher redshifts than the former. Moreover we note that we have no detections with the Rubin Observatory above $z > 4$ so we can safely use the r band without worrying about absorption. Moving to the chirp mass, the SKA+ELT scenario is able to probe the lightest systems in our catalogues, especially for Pop3, while Athena+ELT and the Rubin Observatory detect the EM emission from systems with $10^4 < \mathcal{M}/M_\odot < 10^6$.

The average number of multimessenger candidates and EMcps for each observational scenario is reported in Tab. IV and Tab. V. Overall, the observational strategy providing the most multimessenger candidates and EMcps in 4 yrs is SKA+ELT, if we assume that the radio flare emission is isotropic. Accounting for a beamed emission with $\Gamma = 2$ provides numbers which are closer to those obtained when observing with the Rubin Observatory, or with Athena+ELT. If we further decrease the opening angle ($\Gamma = 10$), observations with SKA+ELT become irrelevant. While the beamed emission allows us to detect systems that are farther away from us, imposing that the observer has to be on-axis excludes the vast majority of the systems.

Observing with the Rubin Observatory provides ~ 3 EMcps in 4 yrs in the Q3d model without accounting for obscuration, while in all the other astrophysical cases the rates are below 1. Concerning the combination Athena+ELT, as expected the Eddington accretion leads to more multimessenger candidates and more EMcps, since the EM emission is brighter. Moreover, in general the observational strategy where one observes a single region of $\Delta\Omega = 0.4\,\mathrm{deg}^2$ allows for the detection of slightly more EMcps than the strategy in which one observes a region of $\Delta\Omega = 2\,\mathrm{deg}^2$ at a higher flux threshold, because there are more systems at fainter fluxes compared to systems with poorer localization.

In general, the two models with massive progenitors predict more EMcps than the Pop3 one, due to the aforementioned difficulties in localising light events with LISA. Indeed, one can appreciate that the Pop3 astrophysical formation model leads to more multimessenger candidates than the Q3nd one, when observing with the Rubin Observatory and SKA+ELT. However, most of them do not satisfy the sky localization requirement and consequently are not accounted for as EMcps.

We highlight that, in order to get the total average number of multimessenger candidates and/or EMcps, one should combine the different EM facilities, while taking care not to count the same event twice (since the same system can be detected with different instruments - see following paragraph). The total average number of EMcps is reported in Tab. VI.

In Tab. VII we report the number of EMcps that can be observed simultaneously by: (i) the Rubin Observatory and SKA ('Rubin+SKA'); (ii) the Rubin Observatory and Athena ('Rubin+Athena'); (iii) SKA and Athena ('SKA+Athena'); (iv) the three instruments ('All'). In the maximising scenario, the Q3d model predicts $\sim 2-4$

TABLE VII. Average number of EMcps observed simultaneously by multiple instruments, in 4 yrs and for different combinations. Upper (lower) table refers to the maximising (minimising) scenario.

	Maximising (multiple instruments)			
	Rubin+SKA	Rubin+Athena	SKA+Athena	All
Pop3	0.84	0.31	1.02	0.31
Q3d	3.07	1.73	3.9	1.7
Q3nd	0.5	0.27	4.0	0.22

	Minimising (multiple instruments)			
	Rubin+SKA	Rubin+Athena	SKA+Athena	All
Pop3	<0.04	0.04	<0.04	<0.04
Q3d	0.13	0.22	0.13	0.13
Q3nd	0.09	0.09	0.04	<0.04

TABLE VIII. Average number of MMcands and EMcps in 4 yrs, for different astrophysical models for the scenario SKA alone. For the Athena only case we obtain the exact same numbers reported in Tab. IV and Tab. V.

	MMcands			EMcps		
	Isotropic flare	$\Gamma 2$	$\Gamma 10$	Isotropic flare	$\Gamma 2$	$\Gamma 10$
Pop3	85.8	26.8	0.58	6.5	1.51	0.04
Q3d	29.3	3.4	0.04	15.4	2.84	0.04
Q3nd	125	18.0	0.93	24.1	4.13	0.13

EMcps in 4 yr (depending on the instruments considered), and about ~ 2 EMcps should be observable by all the instruments simultaneously. As expected, the combination SKA+Athena provides the largest numbers, since both SKA and Athena can observe sources at higher redshift than the Rubin Observatory (c.f. Fig. 3). Moving to the minimising case, we find that < 0.3 EMcps in 4 yrs can be observed by multiple instruments simultaneously, regardless of the astrophysical model.

B. MMcands and EMcps without redshift measurement

In this section we relax the requirement of the redshift determination, i.e. we present the predicted number of multimessenger candidates and EMcps, but without imposing that their redshift should be measured independently. Indeed, interesting information on how the radio or X-ray emissions are produced can also be inferred exclusively by the detection of the EM emission. The redshift can then be determined from the GW-measured luminosity distance by assuming the standard model cosmology (it will not be possible, though, to use these EMcps as standard sirens). Relaxing the redshift determination requirement does not change significantly the number of EMcps; it only affects the number of MMcands, which are, however, less interesting because their

sky localization is unknown. In the following, the requirements for the identification of the MMcands and EMcps in terms of SNR, flux and sky localization remain the same as in the rest of the paper.

First, the number of MMcands and EMcps detectable with the Rubin Observatory with and without imposing redshift determination remains the same because the threshold magnitude that we adopted for spectroscopy, $m_{\mathrm{AGN,Rubin,lim}} = 27.5$, is close to the photometric limit of the survey.

Second, let us focus on the MMcands and EMcps detectable with Athena only. If redshift is not needed, this amounts to dropping the requirement of detectability with ELT, which was imposed exclusively for the redshift determination. However, as can be appreciated from Fig. 3, Athena can only detect sources up to $z \lesssim 4$ while ELT is sensitive up to $z \lesssim 8$ (we justify this value confronting the results for the multimessenger candidates cases in Fig. 1 with the 'SKA alone' configuration in Fig. 19 - see discussion at the end of this subsection). Therefore all the sources detectable with Athena can also be observed with ELT, so that the number of MMcands and EMcps with and without redshift determination is identical.

Moving to SKA, removing the requirement on the redshift determination increases the number of MMcands for Pop3 and Q3nd, respectively, by a factor ~ 2.5 and ~ 3.6 in the isotropic flare scenario. In the models with beamed emissions, the ratio between the number of MMcands without and with redshift rises to ~ 4.3 and 4.8 for Pop3 and Q3nd respectively in the case $\Gamma 2$ and to ~ 4.5 and ~ 23 in the $\Gamma 10$ scenario. For Q3d the increase is less significant in all configurations. The limiting factor is that ELT reaches lower redshifts than SKA, as can be appreciated comparing Fig. 1 and Fig. 19. The number of EMcps does not change significantly instead, because the requirement on the sky localization selects lower redshift systems which can be detected by ELT.

In Tab. VIII we present the numbers of MMcands and EMcps observable with SKA alone, to be compared with the values reported in Tab. IV and Tab. V which include detection with ELT. The distributions of MMcands for SKA in terms of redshift and chirp mass are reported in Fig. 19.

C. Magnitudes and fluxes distributions for EMcps

In this section we present the average number of EMcps as a function of magnitude (relevant for detection with the Rubin Observatory and ELT) and flux (relevant for detection with SKA and Athena).

The magnitude distributions for the EMcps detectable with the Rubin Observatory are reported in Fig. 4. In the absence of AGN obscuration, most of the systems have $m_{\mathrm{AGN,Rubin}} > 25$ and accumulate toward higher magnitudes. From the distribution it is also evident that the Q3d model provides the highest average num-
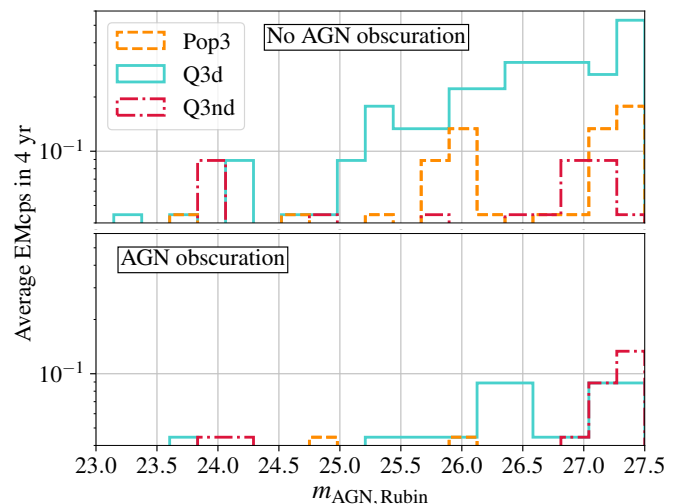
FIG. 4. Magnitude distributions for the EMcps detected with the Rubin Observatory for the three astrophysical scenarios, as clarified by the legend. The upper (lower) panel corresponds to the case without (with) AGN obscuration.
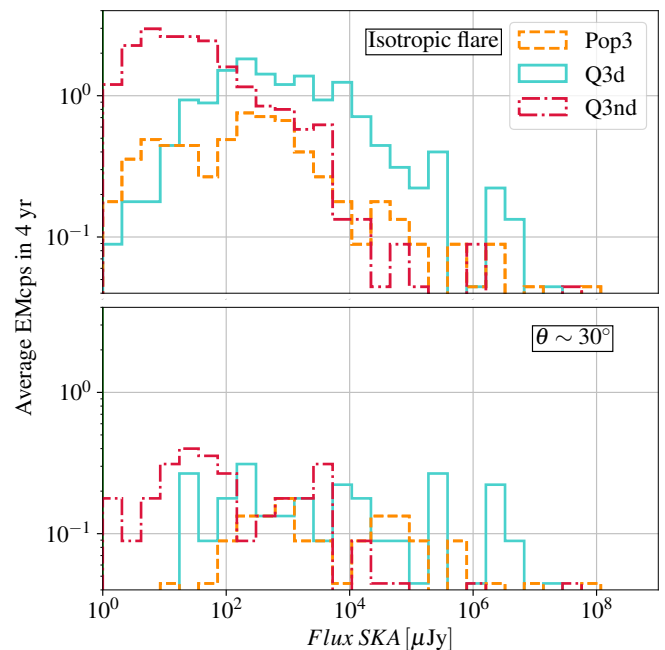
FIG. 5. Flux distributions for the EMcps detected with the SKA+ELT configuration for the three astrophysical scenarios, as clarified by the legend. The upper (lower) panel corresponds to the case with isotropic flare ($\Gamma = 2$ jet) radio emission.

ber of EMcps compared to Pop3 and Q3nd. If we account for AGN obscuration, the number of EMcps diminishes appreciably and the typical magnitude increases to $m_{\mathrm{AGN,Rubin}} \gtrsim 26$, while Q3d still remains the most promising scenario.

The distributions of the radio fluxes of the EMcps de-

tectable with SKA+ELT are reported in Fig. 5 for the 'isotropic flare' and 'Γ2' scenarios. In the isotropic flare case, the distributions are characterised by a peak around $\lesssim 10^2\,\mu$Jy, for all astrophysical models. In the Γ2 case, the distributions appear instead to be flatter.

By inspecting how the radio luminosities are distributed in the catalogues (c.f. Fig. 17 in Appendix C) we found that the flare emission occurs typically at lower luminosity than the jet one: it is therefore subdominant with respect to the jet emission. However, the jet is pointing in the direction of the observer only in a small fraction of cases. Therefore, the peak in the distribution observed in the 'isotropic flare scenario' corresponds to the flare emission. In the 'Γ2' scenario, on the other hand, the jet luminosity dominates, and the peak characteristic of the isotropic flare scenario is absent.

In Fig. 6 we report the number of EMcps observable with Athena as a function of their X-ray flux. We account both for accretion evaluated from the amount of gas surrounding the MBHB (estimated by the SAM), or at Eddington. Furthermore, we show the two sky localization thresholds, within the corresponding flux limits. If we consider the scenario where the accretion rate is derived from the catalogue, the Q3d model provides the highest EMcps number, while Pop3 and Q3nd give similar results. From Fig. 18 in Appendix C, it can be appreciated that, in the Pop3 scenario, the Eddington ratio is about the same as in the Q3d scenario for systems with $\mathcal{M} > 10^4 - 10^5\,\mathrm{M}_\odot$, however, the number of systems with high mass is intrinsically lower with respect to Q3d, and therefore the overall EMcps rate is lower. Furthermore, in Q3nd there are overall more events, but the Eddington ratio is reduced (in other words, without delays there is not time to accumulate enough gas to be accreted on the binary) which leads to fewer EMcps. In the entire catalogue, we also find that the fraction of systems accreting at Eddington is $\sim 80\%$, $\sim 50\%$ and $\sim 1\%$ for Pop3, Q3d and Q3nd respectively. However, if we consider only the subset of EMcps, these fractions increase to $\sim 100\%$, $\sim 88\%$ and $\sim 53\%$ because of the requirement on the detectability of the EMcp.

The lower panels of Fig. 6 confirm that the trade-off between sky localization and limiting flux penalises the scenario with the $\Delta\Omega = 2\,\mathrm{deg}^2$ threshold as the AGN are generally faint. Assuming accretion at Eddington, the number of EMcps increase, as expected. In particular, the Q3nd scenario provides slightly more EMcps than Q3d, as the luminosity depends only on the mass of the binary and not on the amount of gas available for the accretion.

The number of EMcps as a function of the magnitude of their host galaxies, observable with ELT, are shown in Fig. 7. In the maximising case, most of the systems have $m_{\mathrm{gal,ELT}} > 17.5 - 20$ but the inclusion of obscuration and jet pushes this value up to $m_{\mathrm{gal,ELT}} > 22.5$. As we move to larger apparent magnitudes, the number of fainter sources increase for all astrophysical models in a similar way. Most of LISA sources are hosted in faint galaxies, pushing the boundaries toward populations that are challenging to observe. Even if there is always a bright fraction of EMcps the bulk of the population is at the limit of the magnitudes currently observed.

Flux-limited samples are always dominated by faint sources, but this statement is specifically motivated by the actual physical properties of the sources: MBHs of mass $10^3 - 10^7\,\mathrm{M}_\odot$ hosted in galaxies with mass $10^7 - 10^{10}\,\mathrm{M}_\odot$ at high redshift. For sources with SNR> 10 the median MBH mass is $10^{3.57}$, $10^{5.22}$ and $10^{5.75}$ solar masses (Pop3, Q3nd, Q3d), the median galaxy masses are $10^{8.24}$, $10^{7.20}$ and $10^{8.63}\,\mathrm{M}_\odot$ (Pop3, Q3nd, Q3d) and the median redshift 5.18, 8.26 and 3.90. In terms of absolute magnitudes the median AGN absolute magnitude is -13.01, -16.67 (Pop3, Q3d) and the median galaxy absolute magnitude is -15.95, -18.57 (Pop3, Q3d), which according to normal definition are faint sources, in line with the galaxies being dwarfs, based on their masses. Since most of the mergers are at high redshift the corresponding apparent magnitudes also are very faint, but we want to stress that the intrinsic faintness is really a property of LISA sources: small MBHs in high-z dwarf galaxies. Concerning Q3nd, since the Eddington ratios are incredibly small (c.f. Fig. 18) the median absolute magnitudes of the AGN is actually a positive value. The median absolute magnitudes of the galaxies is also very very faint, -12.25, since some of the mergers occur in galaxies with a baryonic mass smaller than $10^6\,\mathrm{M}_\odot$.

D. Magnitudes and fluxes distributions for the entire catalogue

In this section we present the magnitude and flux distributions for the entire catalogues, i.e. without the requirement on the SNR or on the sky localization.

In Fig. 8 we show the magnitude distributions of the sources in our catalogues that can be potentially observed with the Rubin Observatory. Similarly to the EMcps case, the Q3d model predicts more events than Pop3 and Q3nd at all magnitudes. Moreover, the Rubin Observatory does not contribute significantly to the number of multimessenger candidates even at $m_{\mathrm{AGN,Rubin}} > 27.5$ due to the intrinsically low fluxes expected from these systems.

Moving to SKA, in Fig. 9 we present the radio fluxes for the isotropic flare case. For all the astrophysical models, the peak at lower fluxes arises from the flare emission. For the Q3d model, most of the sources are above the detection threshold and the distribution from the entire catalogue is similar to the EMcps one (c.f. Fig. 5).

In Fig. 10 we show the flux distributions in soft X-ray assuming the accretion at Eddington or from the values computed in the catalogues and in the case with and without AGN obscuration. Starting from the latter case, the X-ray fluxes for the three astrophysical models are similar above the threshold but they show different behaviour at fluxes $< 10^{-18}\mathrm{erg\,s}^{-1}\,\mathrm{cm}^{-2}$ due to the differ-
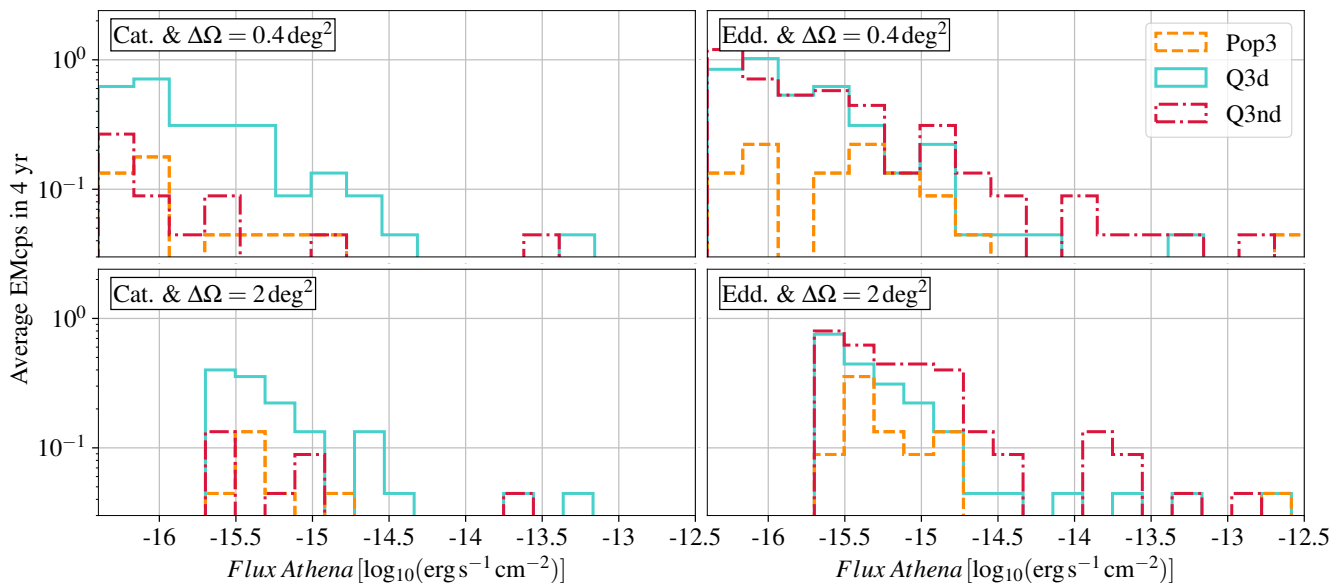
FIG. 6. Flux distributions for the EMcps detected with the Athena+ELT configuration for the three astrophysical scenarios and different configurations.
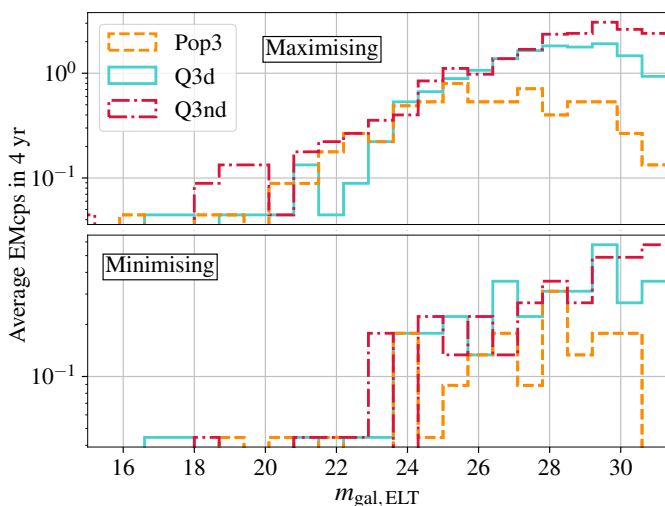

FIG. 7. Magnitude distributions for the EMcps observable with the SKA+ELT and Athena+ELT. The upper (lower) panel corresponds the maximising (minimising) scenario.
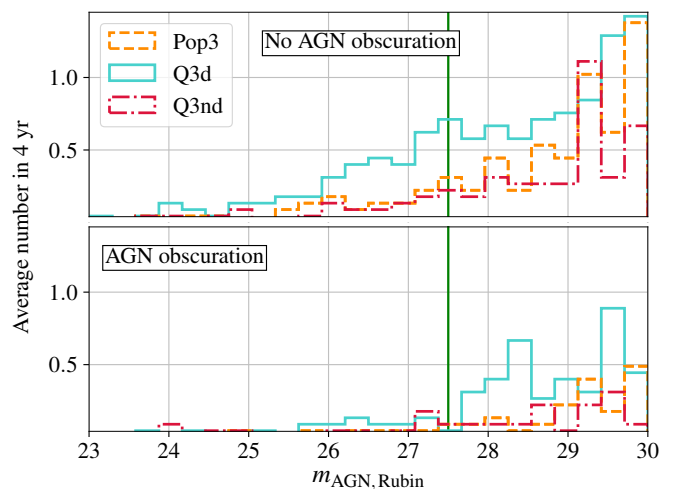
FIG. 8. Same as Fig. 4 but for entire catalogue. The green vertical line represents the limiting magnitude $m_{\mathrm{AGN,Rubin,lim}}$.

ent Eddington ratio values and BH masses.

If we assume Eddington accretion, the flux depends on the binary total mass: since Pop3 has the lightest binaries, it also produces the faintest sources, while Q3nd and Q3d produce brighter emission. The inclusion of AGN obscuration leads to a reduction in the global number of systems with stronger tails extending at fainter fluxes.

In Fig. 11 we report the galaxy magnitude distributions. All three astrophysical models are similar up to $m_{\mathrm{gal,ELT}} \sim 30$. Above this value, the Q3d model starts decreasing due to the lack of sources while Pop3 and Q3nd proceed with the same trend. Both Pop3 and Q3nd

model distributions reach the peak at magnitudes that are too small to be detectable by any planned instruments so we limit the x-axis to $m_{\mathrm{gal,ELT}} = 35$.

VIII. MULTI-MODAL SYSTEMS

LISA ability to accurately localize the gravitational waves source in the sky will strongly depend on the system's parameters, leading to a distribution of sky position uncertainties that spans several orders of magnitude [123]. Moreover, there has been also evidence [71]
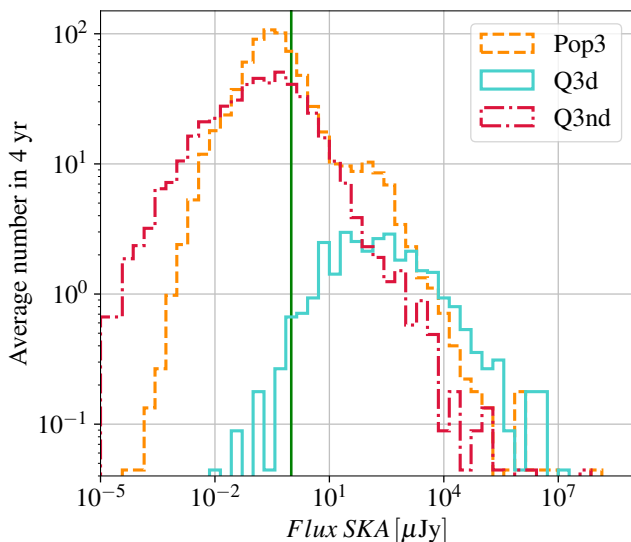
FIG. 9. Same as the upper panel of Fig. 5 but for the entire catalogue. The vertical green line corresponds to 1μJy.

of systems whose sky position posterior distributions are multi-modal in the sky, i.e. they peak not only at the true binary position but also in other regions, symmetrically distributed in the sky. The emergence of these 'multi-modal events' is due to the intrinsic degeneracy in the LISA pattern functions. The degeneracy can be broken only if enough signal is accumulated at low frequencies, where the orbital motion of the detector provides additional information, or at high frequency, thanks to the frequency dependence of the detector response function.

Multi-modal sky position posteriors pose a serious challenge to the search of EM counterparts, since telescopes have to search in a larger region of the sky. In addition, under some conditions that we will discuss, the probability of the "spurious" modes is similar to the probability of the real mode (the actual binary position), further challenging the detection of a counterpart. Fortunately, as we will see, multi-modal EMcps are relatively rare (c.f. Tab. IX).

In [71, 124] it has been shown that, defining (β_T, λ_T) the true binary latitude and longitude, the spurious modes appear at $\beta = -\beta_T$ and $\lambda = \lambda_T + k\pi/2$ with $k = 0, 1, 2, 3$, for a maximum of 8 modes in the sky (one true and seven spurious). For events with only one spurious mode, this secondary point is generally the reflected mode with $k = 0$, i.e. $(-\beta_T, \lambda_T)$. In a minority of cases (two out of the entire catalogue) it is, instead, the antipodal one, i.e. $\beta = -\beta_T$ and $\lambda = \lambda_T + \pi$.

Among the degenerate modes, the reflected one deserves a separate discussion [125]. Without going into the details (we refer the interested readers to [71]), the reflected mode is exactly degenerate with respect to high-frequency effects in the response, and only LISA's motion is expected to break the degeneracy. Thus, the reflected mode appears in the sky position posteriors of signals that are short enough for the LISA motion to be unimportant. Furthermore, the other modes also appear in the sky localization posteriors of systems that are massive enough for their waveform not to reach high frequencies. The degeneracy leading to the other modes, in fact, are usually broken during the merger by the frequency-dependence of LISA's response function. As a consequence, the other modes also become more common in the parameter estimations performed pre-merger.

To define multi-modal events, we introduce the concept of probability for each mode, defined as the ratio between the number of samples in a mode over the total number of samples in the MCMC analysis. We then define as a *1mode* system, a binary whose sky localization posterior has a probability larger than 5% only in a single sky-region. A *2modes* system is such that the probability in the reflected mode is at least 5%, and a *8modes* system is such that the sum of the probability of the other six modes (the total number of modes minus the true binary position and the reflected spot) is at least 5%.

An example of these three cases is reported in Fig. 12. Unimodal events are typically well localized and the Fisher analysis provides a similar result to the Bayesian inference. For *2modes* systems, two spots symmetrical with respect to the equatorial plane of LISA's orbit appear and, for *8modes* systems, the sky position posterior distribution presents eight different peaks located symmetrically. By construction, the Fisher approach, which is a local Gaussian approximation to the posterior, is not able to recover posteriors with multiple peaks.

Some multi-modal events are potential EMcps candidates, and we must include them in our analysis. Two key factors need to be taken into account: the sky localization area of each mode, and the corresponding mode probability. First of all, we want to eliminate events with too wide sky localization region, because telescopes can not explore large areas in the sky. This cut can be performed unambiguously for unimodal systems, but for multi-modal events there are different approaches: the cut can be applied only to the sky localization of the primary mode, or one can choose to combine the sky area of all the modes, assuming the telescope is going to re-point to other locations. This choice influences the number of EMcps. For example, if we assume a threshold of $\Delta\Omega = 10\,\text{deg}^2$ and want to cut all events with larger sky localization region, a bimodal system where the primary and secondary modes have $\Delta\Omega = 8\,\text{deg}^2$ each, is an EMcp in the former approach (with a 50% probability of missing it if the telescope does not point to the right location), but not in the latter, because the total sky area - $\Delta\Omega = 16\,\text{deg}^2$ - would be above threshold. Second, one can also include a requirement on the probability of the modes: for example, one could consider as viable EMcps only the events for which the probability in the primary mode is higher than 50%; or one could argue that modes which probability is less than a given threshold percentage can be discarded, and the EMcp treated like an unimodal one, as far as EM telescopes are
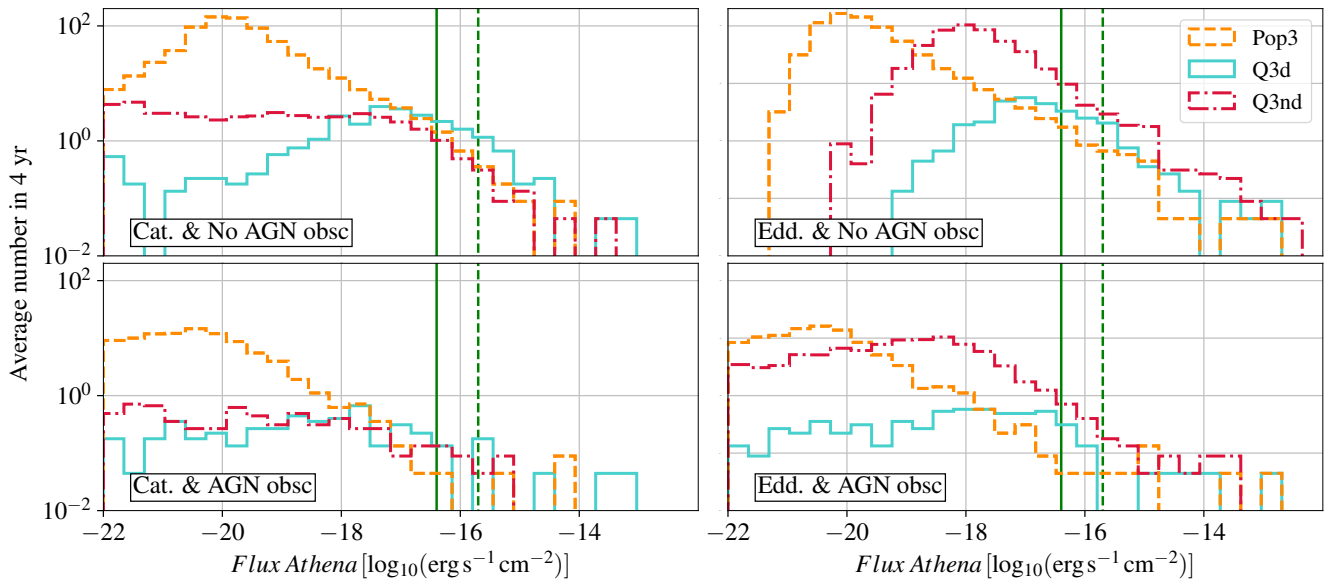
FIG. 10. Flux distributions for the entire catalogue, for the three astrophysical scenarios and different configurations. The left (right) panels correspond to the case where the accretion is computed from the catalogues (assumed at Eddington). The vertical solid (dashed) green line corresponds to $F_{X,\,lim} = 4 \times 10^{-17} \mathrm{erg\,s^{-1}\,cm^{-2}}$ ($F_{X,\,lim} = 2 \times 10^{-16} \mathrm{erg\,s^{-1}\,cm^{-2}}$).
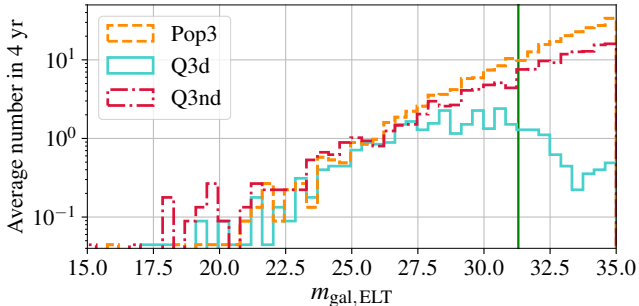
FIG. 11. Same as Fig. 7 but for the entire catalogue. The vertical green line corresponds to 31.3

TABLE IX. Number of *1mode*, *2modes* and *8modes* EMcps in 4 yrs of LISA observation, in the maximising case, and for the three astrophysical models. The sum of all events in each astrophysical model corresponds to what given in Table VI.

	1mode	*2modes*	*8modes*
Pop3	6.0	0.31	0.13
Q3d	10.7	3.9	0.18
Q3nd	16.8	3.5	0.4

concerned.

In this work, we have decided to focus only on the sky localization of the primary mode (i.e. the sky-region where the binary actually stands) as the criterion to select viable EMcps, and to apply no requirement on the probability of the other modes. In other words, the number of EMcps is given by the systems with detectable EM counterpart, and a sky localization region below threshold in the primary mode. This simplification is possible because, as we will show below, events with multi-modal sky posteriors are a minority of all cases, and furthermore, for most bimodal posteriors there is a clear hierarchy in the probability of the primary and the secondary mode. Therefore, the final number of EMcps does not depend excessively on the selection criterion. Note that, both eliminating from the catalogues all the events without detectable EM counterpart, and considering the sky localization region emerging only from the post-merger

parameter estimation analysis, help in reducing the number of multi-modal events: in fact, multi-modal posterior distributions in the sky localization are more frequent at high redshift and for parameter estimation analyses performed pre-merger, when the signals are shorter and have lower SNR [124].

In Tab. IX we report the fraction of *1mode*, *2modes* and *8modes* EMcps in the maximising case (the minimising scenario provides similar results, just rescaled). For all astrophysical models, the largest fraction of EMcps is unimodal, while bimodal EMcps contribute 5%, 26% and 17% of the total rates for Pop3, Q3d and Q3nd respectively. By contrast, *8modes* events constitute less than 0.5 EMcps in 4 yrs of LISA observation.

Fig. 13 shows the distribution of *1mode*, *2modes* and *8modes* EMcps, in the maximising case, in the $z - \mathcal{M}$ plane. It appears clearly that the vast majority of the EMcps population is constituted of unimodal events. Furthermore, the redshifted chirp mass $\mathcal{M}_z$, being the quantity that enters the waveform, influences the appearance of multi-modal posteriors. *2modes* events tend to have masses such that $\mathcal{M}_z \gtrsim 10^6 M_\odot$, and relatively high
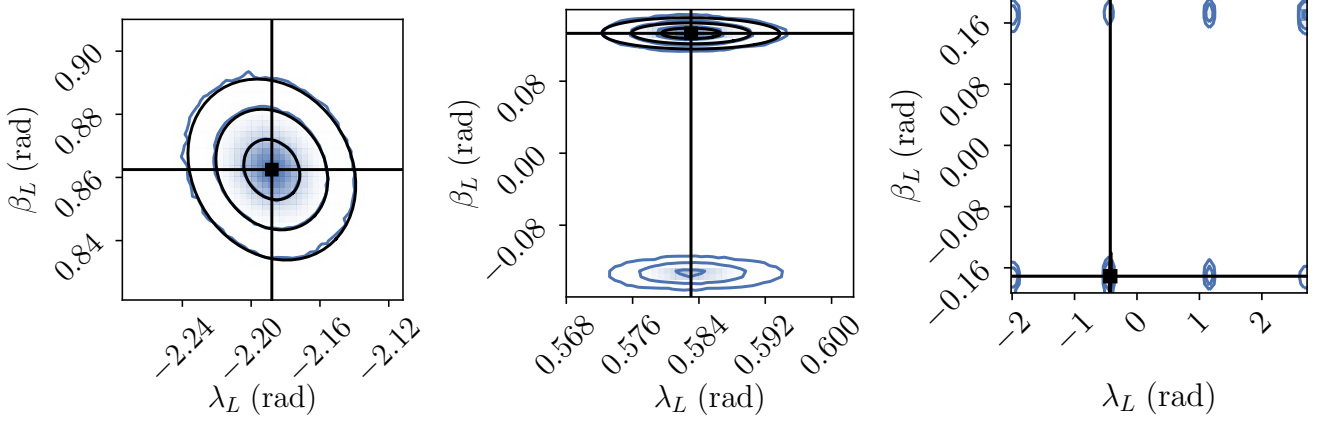
FIG. 12. Examples of multi-modal posterior distributions in the sky localization: from left to right, *1mode*, *2modes* and *8modes* binary systems. The blue contours represent the MCMC results while the black ellipses correspond to the Fisher estimates for 1σ, 2σ and 3σ errors. The black square indicates the true binary position in the sky. The presence of *2modes* and *8modes* events, and the relative probability weight of the modes, can be recovered only with a Bayesian analysis.
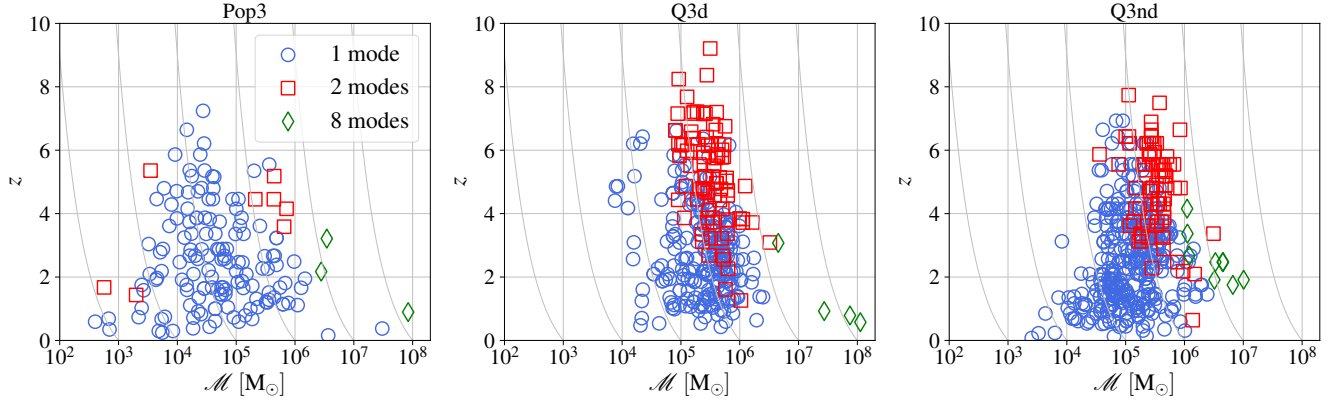

FIG. 13. Distribution of the *1mode*, *2modes* and *8modes* EMcps in the $z - \mathcal{M}$ plane in the maximising case for the three astrophysical scenarios. The grey solid curved lines in background correspond to constant redshifted chirp mass values.

redshift, $z \gtrsim 2$. Events with high chirp mass and redshift, in fact, have GW signals long enough in the LISA band for the high frequency instrument response to play a significant role in breaking the pattern function degeneracy. However, most of the signal is accumulated in the last days, so the effect of LISA motion is insufficient to eliminate the reflected sky position, which remains degenerate. At lower chirp mass and redshift, on the other hand, the combination of the high-frequency response with the motion of the detector fully eliminates the degeneracy and the events are unimodal. However, even if it is possible to identify a general trend, the two sub-populations of *1mode* and *2modes* events do overlap in the $z - \mathcal{M}$ plane, because redshift and $\mathcal{M}_z$ are not the only quantities affecting the parameter estimation, and there is a large dispersion according to the orientation angles. Regardless of the astrophysical model, the *8modes* systems are high chirp mass MBHBs, for which LISA will be able to observe only the merger and ringdown, gathering little

information from the constellation orbital motion; furthermore, their GW signal will not reach high enough frequencies for the frequency-dependence of the detector response to help.

Although *2modes* systems seem to constitute a significant portion of the total EMcps, especially for the massive astrophysical models Q3d and Q3nd, this is partly caused by the fact that, to identify an event as bimodal, we impose a relatively low threshold to the probability of the secondary mode, i.e. 5%. In this regard, it is instructive to look at the probability weight of each mode. In Fig. 14 we present the number of bimodal EMcps as a function of the probability in the primary and secondary modes, for all astrophysical models. It is clear that the primary mode is always more probable than the secondary one, which mitigates the risk, for a substantial fraction of the EMcps, of missing the counterpart if telescopes are pointed only at the primary mode.

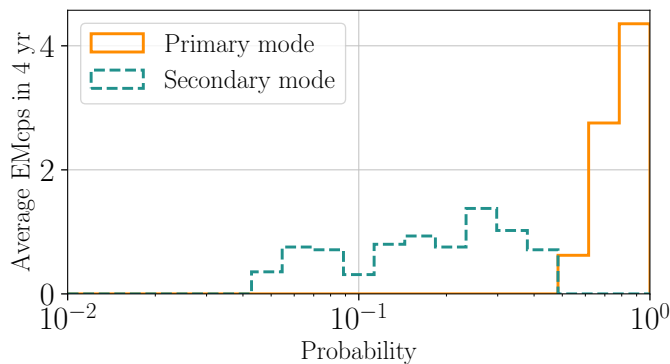In Fig. 15 we show the number of *8modes* EMcps in

FIG. 14. Number of *2modes* EMcps as a function of the probability of the sky position modes, in the maximising scenario, and combining all astrophysical models. The solid (dashed) line corresponds to the probability of the primary (secondary) mode. Note that, by construction, the same EMcp appears twice in this figure. For *2modes* systems, the primary mode is always more probable than the secondary.
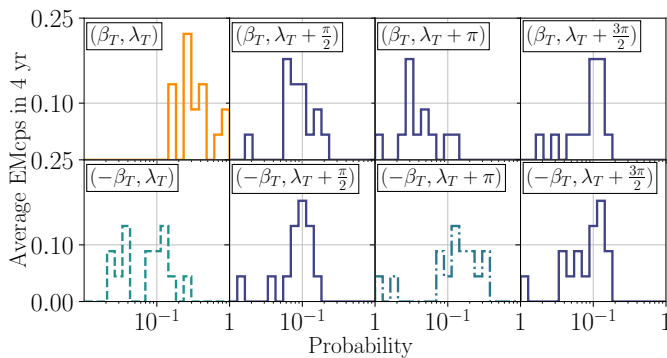

FIG. 15. Number of *8modes* EMcps as a function of the probability of the 8 sky location modes, in the maximising scenario, and combining all the astrophysical models. Each panel shows a different spot in the sky, according to the legend. The solid orange line in the leftmost panel, upper row, corresponds to the real sky position. The dashed and dotted-dashed red lines correspond to the reflected and antipodal sky positions respectively, while the teal solid lines correspond to the other modes. Note that, by construction, the same EMcp appears 8 times in this figure.

each octant of the sky, as a function of the probability of the sky position mode. While the primary mode remains always more probable, the seven spurious modes are rather equiprobable, with probability that can be as large as 10%. This is likely to constitute a serious issue for the search of the EM counterparts of *8modes* events. Fortunately, our results show that *8modes* events are not going to contribute significantly to the total number of EMcps after merger (c.f. Table IX).

Following these results, in this work we have decided to base the selection of multimessenger candidates that can become EMcps only on the requirement that the sky localization region of the primary mode after merger is small enough, as previously described. We decided not to apply, to the so obtained number of EMcps, a correction factor representing the fraction of events that might be missed on average because telescopes are mistakenly pointed at the wrong spot in the sky.

IX. DISCUSSION AND CONCLUSIONS

MBHBs are key sources for multimessenger astrophysics with LISA, as they are expected to merge at the center of galaxies, where the accreting gas might produce an EM counterpart to the GW signal. In this paper we presented the number and characteristics of both multimessenger candidates (i.e. MBHB mergers with a detectable EM emission) and of EMcps (i.e. multimessenger candidates which can be localised well enough via their GW emission). We analysed different astrophysical scenarios for the MBHBs formation, and modelled the EM emission by combining some selected, current and future, EM facilities.

We took as input the results of the SAM code developed in [14, 63–65] to infer the mass and redshift distributions of the merging MBHBs, as well as the properties of their host galaxies. For each MBHB event, we first computed the signal-to-noise ratio in LISA, to assess the number and distribution of the events detectable purely from the GW side. We then exploited the MBH and host galaxy properties to compute the expected EM emission at several wavelengths, from radio to soft X-ray. We considered three observational scenarios: EMcp identification and redshift determination both with the Vera Rubin Observatory; EMcp identification with the SKA and redshift determination with the ELT; EMcp identification with Athena and redshift determination again with the ELT. These observational scenarios cover the entire EM spectrum; naturally, the same system can be detectable by one or more observatories at the same time.

For the subset of events classified as multimessenger candidates, we estimated the errors on the binary parameters, especially the sky localization, inferred by the GW signal with LISA, performing Bayesian parameter estimation as in Marsat *et al.* [71]. We then selected as EMcps only the systems for which the sky localization is smaller than given thresholds, appropriately chosen following the capacities of the EM telescopes. EMcps are therefore those MBHB mergers that can both be exploited for subsequent astrophysical studies, and be used as standard sirens for cosmology, since one can infer their luminosity distance from the GW signal, and their redshift from the EM counterpart.

We focused especially on two scenarios of viable EMcps: one maximising their number, in which we assumed that the radio flare emission is isotropic and that there is no AGN obscuration; and one minimising their number, accounting for beamed radio emission and AGN obscuration. In the maximising scenario we predict 14.9 (6.8) {20.9} EMcps for the Q3d (Pop3) {Q3nd} astrophysical

models respectively, in 4 yrs of LISA observations. In the minimising scenario, these rates decrease to 3.4 (1.7) {3.4}, respectively. The collimation of the radio jet and the AGN obscuration by hydrogen and dust are the two features most affecting our results as far as the detection of the EM emission is concerned.

Removing the requirement of the redshift determination does not change significantly the number of MMcands and EMcps observed with the Rubin Observatory and Athena: the cuts in magnitude imposed by the detection with these facilities already select sources at relatively low redshift. Concerning the SKA, on the other hand, we find that the number of MMcands increases by a factor $\sim 2.5 - 5$ when removing the redshift identification requirement, because SKA can reach higher redshifts than ELT. Note that the number of EMcps always remains the same: these systems have, in fact, low redshift due to the requirement on the sky localisation.

We find that EMcps can be detected up to $z \sim 6 - 8$ and they have typically $\mathcal{M} \sim 10^4 - 10^6 \, \mathrm{M}_\odot$, because these systems have a sufficient amount of gas available for accretion and they are localized with enough accuracy by LISA. Considering each observational scenario separately, we find that the Rubin Observatory is the instrument providing the smallest number of EMcps, with on average less than one event in four years of joint observation with LISA. On the other hand, SKA+ELT provide the largest number of EMcps, if the radio flare emission is isotropic; however, the EMcps rate is drastically reduced if the radio flare and jet emissions are collimated in a jet cone with opening angle $\theta \sim 30°$, and goes to zero for $\theta \sim 6°$. Therefore, under the plausible assumption of a collimated radio emission, only the presence of Athena in conjunction with LISA would re-enable the possibility of having at least a few EMcps in 4 yrs, assuming, though, that the X-ray emission is not affected by dust obscuration. Introducing AGN obscuration leads in fact to a significant drop in the EMcp rates. Interestingly, restricting the sky area to the size of the Athena FOV, thereby reaching deeper fluxes, increases the number of EMcps, rather than scanning a larger area of the sky with higher flux limit. We therefore identify the first observational strategy as the one capable of maximising the opportunity of joint LISA-Athena observations.

The number of multimessenger candidates changes in the same way as the one of EMcps, when changing the observational strategy. We found that the Pop3 astrophysical model leads to more multimessenger candidates than the massive models Q3d and Q3nd, but most of them do not satisfy the sky localization requirements, because they are too light and at high redshift. Consequently, these events cannot be classified as EMcps, so that the final number of EMcps in the Pop3 astrophysical model is overall smaller than in the massive models. It is also important to remark that the vast majority of the MBHB mergers in the SAM catalogues are characterised by an EM emission much fainter than the threshold magnitudes and fluxes of the EM observatories considered.

Since these facilities are representative of the planned future telescopes, this shows, to the best of our present knowledge, how challenging multimessenger MBHB observations with LISA will be.

We also found that a fraction of multimessenger candidates present multi-modal sky posterior distributions, characterised by two or eight spots in the sky, symmetrically distributed over the sphere. We promoted these events to EMcps when their GW-inferred sky localization is smaller than the selected thresholds in the primary mode only, neglecting the other modes. We found that bimodal events can contribute up to $\sim 25\%$ of all EMcps; however, the posterior probability of the primary mode is always larger than the one of the secondary one. *8modes* events, on the other hand, have nearly equiprobable sky-localisation modes, but they contribute less than 0.5 EMcps in 4 yrs of LISA observation, so they play a minor role and do not affect our results.

The present analysis has also some caveats. Concerning the catalogues, our results are based on the same MBH physics described in T16. More recently, new predictions for the MBHB merger rates have been published [126]. In this work, the authors refined the modeling of the time delays, accounting for the baryonic components at galactic merging scale and implementing better prescription for dynamical friction at smaller scales (< 100 pc). They also included the effect of supernova feedback that may reduce the amount of available gas in low-mass galaxies. We compared the two catalogues and found no significant differences in the number of both intrinsic and detected i.e. (SNR > 10) MBHB mergers for the Q3d model (the 'heavy seed' model in the recent work). We note especially that, within the Q3d model, the redshift distributions provided by the new analysis are skewed toward lower redshift, so we expect to recover a similar, or even larger, number of EMcps. On the contrary, the new results predict more (less) MBHB mergers for the Q3nd (Pop3) astrophysical models. Therefore, the results obtained in this work might be overestimated for Q3nd and underestimated for Pop3.

Concerning the modeling of the EM emission and of its detection, several considerations are in order. The effective number of EMcps might be reduced by difficulties linked to the follow-up strategy that we have not addressed, like, for example, how to deal with the situation of two, or more, coincident MBHB mergers in LISA or the presence of other sources with time-varying EM emission in the sky localization region provided by LISA. Concerning the last point, Lops *et al.* [127] explored the possibility to identify the host galaxy of MBHBs mergers starting from a simulated portion of the Universe. Their work can be considered complementary to ours since they focused more extensively on the exploration of galactic fields around MBHBs mergers. Similarly we did not take into account the possibility of a time delay of weeks to years between the GW signal and the peak of the EM emission. Such long delays would seriously impact the possibility to identify unambiguously the host

galaxy since searching for a transient in a large patch of the sky with deep ToOs is challenging. However, a possible solution might be to look for modulated emission in archival data once a transient is detected. Also, we used the post-merger localization of the GW signal by LISA, which is drastically better than the pre-merger one, but at the same time we fully neglected the possibility that some EMcps might be detected already before merger. In order to evaluate the bolometric luminosity in the Eddington emission case, we used the mass of the binary, while for post merger emission it would have been more appropriate to use the mass of the remnant black hole, since the GW emission is expected to subtract energy and angular momentum from the system. However, we found that this leads to no significant difference in our estimates, as the binary and remnant BH masses are very similar [77].

We also chose to be as optimistic as possible and did not apply any correction factor to account for the actual sky coverage of telescopes such as SKA or the Rubin Observatory, which will observe at best half the sky. In the worst case scenario, this would lead to the loss of half of the EMcps, due to Earth orientation.

Regarding Athena, our results suggests that observing a single tile (for example, the one with the highest posterior probability resulting from the LISA parameter estimation), might be a successful strategy to detect the X-ray counterpart. However, this result needs confirmation via a realistic end to end simulation of the observing strategy, which is behind the scope of the current work. For example, each tile and the corresponding observing time might be weighted with the posterior probability from the GW measurement.

Regarding the LISA data analysis, we did not account for data gaps in the detector output. This might be justified, as at least the gaps expected in connection with the standard LISA maintenance can be avoided by triggering a protected period around the predicted MBHB merger time, if the binary is detected sufficiently early before the merger. However, data gaps might still be problematic for massive systems entering in LISA band already in the merger or ringdown phase, since there will be no time to postpone a scheduled maintenance. Moreover in this work, we perform the parameter estimation only for the MBHBs with detectable EM emission. In the future, we plan to perform the parameter estimation of all the binaries present in the SAM catalogues, independently of whether a detectable EM emission is associated to them or not [124], to have a more comprehensive view.

In this work, we restrict ourselves to MBHBs with circularized orbits and spins aligned to the orbital angular momentum, thus neglecting the possibility of misaligned spins inducing precession, as well as the possible presence of eccentricity. We leave the investigation of the localization of sources in the presence of these effects for future work.

Finally, the actual distribution of observed MBHBs might arise from a combination of the three astrophysical models we adopted in this study (see for example [128, 129]). For simplicity, here we did not considered mixed astrophysical formation models.

ACKNOWLEDGMENTS

[1] P. Amaro-Seoane *et al.*, ArXiv e-prints (2017), arXiv:1702.00786 [astro-ph.IM].

[2] M. Volonteri, The Astronomy and Astrophysics Review **18**, 279 (2010), arXiv:1003.4404 [astro-ph.CO].

[3] J. L. Johnson and F. Haardt, Publications of the Astronomical Society of Australia **33**, e007 (2016), arXiv:1601.05473.

[4] H. Umeda, T. Hosokawa, K. Omukai, and N. Yoshida, The Astrophysical Journal Letters **830**, L34 (2016), arXiv:1609.04457 [astro-ph.SR].

[5] R. Valiante, R. Schneider, M. Volonteri, and K. Omukai, Monthly Notices of the Royal Astronomical Society **457**, 3356 (2016), arXiv:1601.07915.

[6] M. Habouzit, M. Volonteri, and Y. Dubois, Monthly Notices of the Royal Astronomical Society **468**, 3935 (2017), arXiv:1605.09394 [astro-ph.GA].

[7] K. Inayoshi, E. Visbal, and Z. Haiman, Annual Review of Astronomy and Astrophysics **58**, 27 (2020), arXiv:1911.05791 [astro-ph.GA].

[8] R. Valiante *et al.*, Monthly Notices of the Royal Astronomical Society **500**, 4095 (2021), arXiv:2010.15096 [astro-ph.GA].

[9] A. Sesana, J. Gair, E. Berti, and M. Volonteri, Phys. Rev. D **83**, 044036 (2011), arXiv:1011.5893 [astro-ph.CO].

[10] A. H. Jaffe and D. C. Backer, Astrophys. J. **583**, 616 (2003), arXiv:astro-ph/0210148 [astro-ph].

[11] J. S. B. Wyithe and A. Loeb, Astrophys. J. **590**, 691 (2003), arXiv:astro-ph/0211556 [astro-ph].

[12] M. Enoki, K. T. Inoue, M. Nagashima, and N. Sugiyama, Astrophys. J. **615**, 19 (2004), arXiv:astro-ph/0404389 [astro-ph].

[13] A. Sesana, F. Haardt, P. Madau, and M. Volonteri, Astrophys. J. **623**, 23 (2005), arXiv:astro-ph/0409255 [astro-ph].

[14] E. Barausse, Monthly Notices of the Royal Astronom-

ical Society **423**, 2533 (2012), arXiv:1201.5888 [astro-ph.CO].

[15] A. Klein, E. Barausse, A. Sesana, A. Petiteau, E. Berti, S. Babak, J. Gair, S. Aoudia, I. Hinder, F. Ohme, and B. Wardell, Phys. Rev. D **93**, 024003 (2016).

[16] H. Padmanabhan and A. Loeb, Journal of Cosmology and Astroparticle Physics **2020**, 055 (2020), arXiv:2007.12710 [astro-ph.CO].

[17] P. Amaro-Seoane *et al.*, arXiv e-prints , arXiv:2203.06016 (2022), arXiv:2203.06016 [gr-qc].

[18] E. Berti, V. Cardoso, and C. M. Will, Phys. Rev. D **73**, 064030 (2006).

[19] S. A. Hughes and K. Menou, Astrophys. J. **623**, 689 (2005), arXiv:astro-ph/0410148 [astro-ph].

[20] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, Journal of Cosmology and Astroparticle Physics **2016**, 002 (2016), arXiv:1601.07112 [astro-ph.CO].

[21] R.-G. Cai, N. Tamanini, and T. Yang, Journal of Cosmology and Astroparticle Physics **2017**, 031 (2017), arXiv:1703.07323 [astro-ph.CO].

[22] C. Caprini and N. Tamanini, Journal of Cosmology and Astroparticle Physics **2016**, 006 (2016), arXiv:1607.08755 [astro-ph.CO].

[23] M. Corman, A. Ghosh, C. Escamilla-Rivera, M. A. Hendry, S. Marsat, and N. Tamanini, Phys. Rev. D **105**, 064061 (2022), arXiv:2109.08748 [gr-qc].

[24] L. Speri, N. Tamanini, R. R. Caldwell, J. R. Gair, and B. Wang, Phys. Rev. D **103**, 083526 (2021), arXiv:2010.09049 [astro-ph.CO].

[25] E. Belgacem *et al.*, Journal of Cosmology and Astroparticle Physics **2019**, 024 (2019), arXiv:1906.01593 [astro-ph.CO].

[26] P. Auclair *et al.*, arXiv e-prints , arXiv:2204.05434 (2022), arXiv:2204.05434 [astro-ph.CO].

[27] K. G. Arun *et al.*, Living Reviews in Relativity **25**, 4 (2022), arXiv:2205.01597 [gr-qc].

[28] W. Hillebrandt and J. C. Niemeyer, Annual Review of Astronomy and Astrophysics **38**, 191 (2000), https://doi.org/10.1146/annurev.astro.38.1.191.

[29] D. E. Holz and S. A. Hughes, The Astrophysical Journal **629**, 15 (2005).

[30] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, The Astrophysical Journal **725**, 496 (2010).

[31] W. Del Pozzo, Phys. Rev. D **86**, 043011 (2012).

[32] B. P. Abbott *et al.*, Nature (London) **551**, 85 (2017), arXiv:1710.05835 [astro-ph.CO].

[33] B. F. Schutz, Nature **323**, 310 (1986).

[34] A. Petiteau, S. Babak, and A. Sesana, Astrophys. J. **732**, 82 (2011), arXiv:1102.0769 [astro-ph.CO].

[35] W. Del Pozzo, A. Sesana, and A. Klein, Monthly Notices of the Royal Astronomical Society **475**, 3485 (2018), arXiv:1703.01300 [astro-ph.CO].

[36] K. Kyutoku and N. Seto, Phys. Rev. D **95**, 083525 (2017).

[37] N. Muttoni, A. Mangiagli, A. Sesana, D. Laghi, W. Del Pozzo, and D. Izquierdo-Villalba, arXiv e-prints , arXiv:2109.13934 (2021), arXiv:2109.13934 [astro-ph.CO].

[38] D. Laghi, N. Tamanini, W. Del Pozzo, A. Sesana, J. Gair, S. Babak, and D. Izquierdo-Villalba, Monthly Notices of the Royal Astronomical Society **508**, 4512 (2021), arXiv:2102.01708 [astro-ph.CO].

[39] L.-G. Zhu, Y.-M. Hu, H.-T. Wang, J.-d. Zhang, X.-D. Li, M. Hendry, and J. Mei, Physical Review Research **4**, 013247 (2022), arXiv:2104.11956 [astro-ph.CO].

[40] A. De Rosa *et al.*, New Astronomy Reviews , 101525 (2020).

[41] P. J. Armitage and P. Natarajan, The Astrophysical Journal **567**, L9 (2002).

[42] M. Milosavljevic and E. S. Phinney, Astrophys. J. **622**, L93 (2005), arXiv:astro-ph/0410343 [astro-ph].

[43] M. Dotti, R. Salvaterra, A. Sesana, M. Colpi, and F. Haardt, Mon. Not. Roy. Astron. Soc. **372**, 869 (2006), arXiv:astro-ph/0605624 [astro-ph].

[44] B. Kocsis, Z. Frei, Z. Haiman, and K. Menou, The Astrophysical Journal **637**, 27 (2006).

[45] Y. Tang, Z. Haiman, and A. MacFadyen, Monthly Notices of the Royal Astronomical Society **476**, 2249 (2018).

[46] C. Roedig, J. H. Krolik, and M. C. Miller, Astrophys. J. **785**, 115 (2014), arXiv:1402.7098 [astro-ph.HE].

[47] S. d'Ascoli, S. C. Noble, D. B. Bowen, M. Campanelli, J. H. Krolik, and V. Mewes, The Astrophysical Journal **865**, 140 (2018).

[48] B. D. Farris, P. Duffell, A. I. MacFadyen, and Z. Haiman, Monthly Notices of the Royal Astronomical Society **447**, L80 (2015), arXiv:1409.5124 [astro-ph.HE].

[49] A. Franchini, A. Lupi, and A. Sesana, The Astrophysical Journal Letters **929**, L13 (2022).

[50] D. B. Bowen, V. Mewes, M. Campanelli, S. C. Noble, J. H. Krolik, and M. Zilhão, The Astrophysical Journal **853**, L17 (2018).

[51] T. Dal Canton, A. Mangiagli, S. C. Noble, J. Schnittman, A. Ptak, A. Klein, A. Sesana, and J. Camp, Astrophys. J. **886**, 146 (2019), arXiv:1902.01538 [astro-ph.HE].

[52] S. McGee, A. Sesana, and A. Vecchio, Nature Astronomy **4**, 26 (2020), arXiv:1811.00050 [astro-ph.HE].

[53] C. Palenzuela, L. Lehner, and S. L. Liebling, Science **329**, 927 (2010), arXiv:1005.1067 [astro-ph.HE].

[54] E. M. Rossi, G. Lodato, P. J. Armitage, J. E. Pringle, and A. R. King, Monthly Notices of the Royal Astronomical Society **401**, 2021 (2010), arXiv:0910.0002 [astro-ph.HE].

[55] J. D. Schnittman and J. H. Krolik, Astrophys. J. **684**, 835 (2008), arXiv:0802.3556 [astro-ph].

[56] Ž. Ivezić *et al.*, Astrophys. J. **873**, 111 (2019), arXiv:0805.2366 [astro-ph].

[57] https://www.lsst.org/about.

[58] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, IEEE Proceedings **97**, 1482 (2009).

[59] K. Nandra, D. Barret, X. Barcons, A. Fabian, J.-W. den Herder, L. Piro, M. Watson, C. Adami, J. Aird, J. M. Afonso, and et al., arXiv e-prints (2013), arXiv:1306.2307 [astro-ph.HE].

[60] L. Piro *et al.*, Experimental Astronomy (2022), 10.1007/s10686-022-09865-6, arXiv:2110.15677 [astro-ph.HE].

[61] L. Piro *et al.*, in preparation.

[62] E-ELT, https://www.eso.org/sci/facilities/eelt/.

[63] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, Astrophys. J. **794**, 104 (2014), arXiv:1402.7088 [astro-ph.CO].

[64] F. Antonini, E. Barausse, and J. Silk, The Astrophys-

ical Journal Letters **806**, L8 (2015), arXiv:1504.04033 [astro-ph.GA].

[65] F. Antonini, E. Barausse, and J. Silk, Astrophys. J. **812**, 72 (2015), arXiv:1506.02050 [astro-ph.GA].

[66] P. Madau and M. J. Rees, The Astrophysical Journal Letters **551**, L27 (2001), arXiv:astro-ph/0101223 [astro-ph].

[67] M. Volonteri, F. Haardt, and P. Madau, Astrophys. J. **582**, 559 (2003), arXiv:astro-ph/0207276 [astro-ph].

[68] M. Volonteri, G. Lodato, and P. Natarajan, Monthly Notices of the Royal Astronomical Society **383**, 1079 (2008), arXiv:0709.0529 [astro-ph].

[69] M. C. Begelman, M. Volonteri, and M. J. Rees, Monthly Notices of the Royal Astronomical Society **370**, 289 (2006), arXiv:astro-ph/0602363 [astro-ph].

[70] The orbits of LISA will not be perfectly periodic due to some fluctuations in the motion of the satellites and the degradation of the onboard instrumentation. However, for simplicity, we assume periodic orbits which allow us to rescale any arbitrary merger time in the interval [0, 1] yr. Moreover most of these systems will last less than 1 month in LISA band so the choice of the coalescence time does not impact significantly the parameter estimation. We also did not take into account signals truncated at the beginning and at the end of LISA mission.

[71] S. Marsat, J. G. Baker, and T. D. Canton, Phys. Rev. D **103**, 083011 (2021), arXiv:2003.00357 [gr-qc].

[72] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, Phys. Rev. Lett. **120**, 161102 (2018).

[73] J. A. Regan, E. Visbal, J. H. Wise, Z. Haiman, P. H. Johansson, and G. L. Bryan, Nature Astronomy **1**, 0075 (2017), arXiv:1703.03805 [astro-ph.GA].

[74] S. Chon, S. Hirano, T. Hosokawa, and N. Yoshida, Astrophys. J. **832**, 134 (2016), arXiv:1603.08923 [astro-ph.GA].

[75] S. Chandrasekhar, Astrophys. J. **97**, 255 (1943).

[76] The radiation from young stars in the bulge is expected to bring the gas in a low-angular momentum disk surrounding the MBHB [130] so, for our purposes, the accretion of the reservoir mass follow the star formation rate in the bulge.

[77] E. Barausse, V. Morozova, and L. Rezzolla, Astrophys. J. **758**, 63 (2012), arXiv:1206.3803 [gr-qc].

[78] A. Merloni and S. Heinz, Monthly Notices of the Royal Astronomical Society **388**, 1011 (2008), arXiv:0805.2499 [astro-ph].

[79] S. Shen, H. J. Mo, S. D. M. White, M. R. Blanton, G. Kauffmann, W. Voges, J. Brinkmann, and I. Csabai, Monthly Notices of the Royal Astronomical Society **343**, 978 (2003), arXiv:astro-ph/0301527 [astro-ph].

[80] W. E. Harris, Astronomical Journal **112**, 1487 (1996).

[81] J. M. Bardeen, Nature (London) **226**, 64 (1970).

[82] L. S. Sparke and J. S. Gallagher III, *Galaxies in the Universe: An Introduction*, 2nd ed. (Cambridge University Press, 2007).

[83] X. Shen, P. F. Hopkins, C.-A. Faucher-Giguère, D. M. Alexander, G. T. Richards, N. P. Ross, and R. C. Hickox, Monthly Notices of the Royal Astronomical Society **495**, 3252 (2020), arXiv:2001.02696 [astro-ph.GA].

[84] C. N. A. Willmer, The Astrophysical Journal Supplement Series **236**, 47 (2018).

[85] C. Laigle *et al.*, Monthly Notices of the Royal Astronom-

[86] R. Gold, V. Paschalidis, M. Ruiz, S. L. Shapiro, Z. B. Etienne, and H. P. Pfeiffer, Phys. Rev. D **90**, 104030 (2014), arXiv:1410.1543 [astro-ph.GA].

[87] B. J. Kelly, J. G. Baker, Z. B. Etienne, B. Giacomazzo, and J. Schnittman, Phys. Rev. D **96**, 123003 (2017), arXiv:1710.02132 [astro-ph.HE].

[88] F. Cattorini, B. Giacomazzo, F. Haardt, and M. Colpi, Phys. Rev. D **103**, 103022 (2021), arXiv:2102.13166 [astro-ph.HE].

[89] R. D. Blandford and R. L. Znajek, Monthly Notices of the Royal Astronomical Society **179**, 433 (1977).

[90] D. L. Meier, The Astrophysical Journal Letters **548**, L9 (2001), arXiv:astro-ph/0010231 [astro-ph].

[91] R. D. Blandford and A. Königl, Astrophys. J. **232**, 34 (1979).

[92] C. Yuan, K. Murase, B. T. Zhang, S. S. Kimura, and P. Mészáros, The Astrophysical Journal Letters **911**, L15 (2021), arXiv:2101.05788 [astro-ph.HE].

[93] M. H. Cohen, M. L. Lister, D. C. Homan, M. Kadler, K. I. Kellermann, Y. Y. Kovalev, and R. C. Vermeulen, Astrophys. J. **658**, 232 (2007), arXiv:astro-ph/0611642 [astro-ph].

[94] P. Moesta, D. Alic, L. Rezzolla, O. Zanotti, and C. Palenzuela, The Astrophysical Journal Letters **749**, L32 (2012), arXiv:1109.1177 [gr-qc].

[95] K. Gültekin, A. L. King, E. M. Cackett, K. Nyland, J. M. Miller, T. Di Matteo, S. Markoff, and M. P. Rupen, Astrophys. J. **871**, 80 (2019), arXiv:1901.02530 [astro-ph.HE].

[96] C. A. Páez *et al.*, in preparation.

[97] At the moment of writing this paper, there have been rumors concerning the possibility of Athena mission being descoped. Currently nothing is yet decided so we chose to adopt the current Athena configuration but we plan to review our result if any major issue will rise. We also discuss the value of having LISA and Athena operational at the same in Sec. IX.

[98] M. Guainazzi, personal communication.

[99] R. Davies *et al.*, in *Ground-based and Airborne Instrumentation for Astronomy III*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, edited by I. S. McLean, S. K. Ramsay, and H. Takami (2010) p. 77352A, arXiv:1005.5009 [astro-ph.IM].

[100] J. S. Dunlop, in *The First Galaxies*, Astrophysics and Space Science Library, Vol. 396, edited by T. Wiklind, B. Mobasher, and V. Bromm (2013) p. 223, arXiv:1205.1543 [astro-ph.CO].

[101] D. Spergel *et al.*, arXiv e-prints , arXiv:1503.03757 (2015), arXiv:1503.03757 [astro-ph.IM].

[102] We report this value for completeness and we do not use this value in the analysis. Therefore we do not require the detectability also with the Roman Space Telescope. More details can be found at https://roman.gsfc.nasa.gov/instruments_and_capabilities.html.

[103] In paper II we will provide more detail on this point and show that our cosmological results are robust against this, somewhat arbitrary, choice of the redshift error for systems with $27.2 < m_{ELT} < 31.3$ and $0.5 < z < 5$.

[104] J. P. Gardner *et al.*, Space Science Reviews **123**, 485 (2006), arXiv:astro-ph/0606175 [astro-ph].

[105] C. Ramos Almeida and C. Ricci, Nature Astronomy **1**,

679 (2017), arXiv:1709.00019 [astro-ph.GA].

[106] J. Buchner and F. E. Bauer, Monthly Notices of the Royal Astronomical Society **465**, 4348 (2016), https://academic.oup.com/mnras/article-pdf/465/4/4348/10254411/stw2955.pdf.

[107] R. Gilli, A. Comastri, and G. Hasinger, Astronomy&Astrophysics **463**, 79 (2007), arXiv:astro-ph/0610939 [astro-ph].

[108] A. Merloni *et al.*, Monthly Notices of the Royal Astronomical Society **437**, 3550 (2014), arXiv:1311.1305 [astro-ph.CO].

[109] J. Buchner, A. Georgakakis, K. Nandra, M. Brightman, M.-L. Menzel, Z. Liu, L.-T. Hsu, M. Salvato, C. Rangel, J. Aird, A. Merloni, and N. Ross, Astrophys. J. **802**, 89 (2015), arXiv:1501.02805 [astro-ph.HE].

[110] Y. Ueda, M. Akiyama, G. Hasinger, T. Miyaji, and M. G. Watson, Astrophys. J. **786**, 104 (2014), arXiv:1402.1836 [astro-ph.CO].

[111] R. Morrison and D. McCammon, Astrophys. J. **270**, 119 (1983).

[112] C. Ricci *et al.*, Nature (London) **549**, 488 (2017), arXiv:1709.09651 [astro-ph.HE].

[113] M. Trebitsch, M. Volonteri, and Y. Dubois, Monthly Notices of the Royal Astronomical Society **487**, 819 (2019), arXiv:1901.01261 [astro-ph.GA].

[114] C. Circosta, C. Vignali, R. Gilli, A. Feltre, F. Vito, F. Calura, V. Mainieri, M. Massardi, and C. Norman, Astronomy&Astrophysics **623**, A172 (2019), arXiv:1901.07108 [astro-ph.GA].

[115] X. Ma, P. F. Hopkins, C.-A. Faucher-Giguère, N. Zolman, A. L. Muratov, D. Kereš, and E. Quataert, Monthly Notices of the Royal Astronomical Society **456**, 2140 (2016), arXiv:1504.02097 [astro-ph.GA].

[116] N. Y. Gnedin, A. V. Kravtsov, and H.-W. Chen, Astrophys. J. **672**, 765 (2008), arXiv:0707.0879 [astro-ph].

[117] N. R. Tanvir *et al.*, Monthly Notices of the Royal Astronomical Society **483**, 5380 (2019), arXiv:1805.07318 [astro-ph.GA].

[118] S. Babak, M. Hewitson, and A. Petiteau, arXiv e-prints , arXiv:2108.01167 (2021), arXiv:2108.01167 [astro-ph.IM].

[119] N. Karnesis, S. Babak, M. Pieroni, N. Cornish, and T. Littenberg, Phys. Rev. D **104**, 043019 (2021), arXiv:2103.14598 [astro-ph.IM].

[120] W. D. Vousden, W. M. Farr, and I. Mandel, Monthly Notices of the Royal Astronomical Society **455**, 1919 (2016), arXiv:1501.05823 [astro-ph.IM].

[121] In the MCMC formalism, a 'walker' represents the point of the chain that is exploring the parameter space.

[122] In the context of parallel tempering, different temperatures are adopted to smooth the parameter space and ease the convergence of the algorithm.

[123] A. Mangiagli, A. Klein, M. Bonetti, M. L. Katz, A. Sesana, M. Volonteri, M. Colpi, S. Marsat, and S. Babak, Phys. Rev. D **102**, 084056 (2020), arXiv:2006.12513 [astro-ph.HE].

[124] S. Marsat *et al.*, in preparation.

[125] In the spurious modes we expect degeneracies also in the binary inclination and polarization but these are not relevant for our scope.

[126] E. Barausse, I. Dvorkin, M. Tremmel, M. Volonteri, and M. Bonetti, Astrophys. J. **904**, 16 (2020), arXiv:2006.03065 [astro-ph.GA].

[127] G. Lops, D. Izquierdo-Villalba, M. Colpi, S. Bonoli, A. Sesana, and A. Mangiagli, arXiv e-prints , arXiv:2207.10683 (2022), arXiv:2207.10683 [astro-ph.GA].

[128] M. Volonteri and M. C. Begelman, Monthly Notices of the Royal Astronomical Society **409**, 1022 (2010), arXiv:1003.5220 [astro-ph.HE].

[129] D. Spinoso, S. Bonoli, R. Valiante, R. Schneider, and D. Izquierdo-Villalba, arXiv e-prints , arXiv:2203.13846 (2022), arXiv:2203.13846 [astro-ph.GA].

[130] Z. Haiman, L. Ciotti, and J. P. Ostriker, The Astrophysical Journal **606**, 763 (2004).

[131] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan, Phys. Rev. D **82**, 064016 (2010).

[132] A. Petiteau, M. Hewitson, G. Heinzel, E. Fitzsimons, and H. Halloin, *LISA Noise Budget*, Tech. Rep. LISA-CST-TN0001 (LISA Consortium, 2016).

[133] M. Punturo *et al.*, Classical and Quantum Gravity **27**, 194002 (2010).

[134] S. Kawamura *et al.*, Journal of Physics: Conference Series **122**, 012006 (2008).

Appendix A: Comparison with previous results

In this appendix we compare our results to previous works in the literature. We start comparing our estimates with Belgacem *et al.* [25] (hereafter 'B19') where the authors adopted the same catalogues but with different LISA sensitivity, GW waveform and EM counterpart modeling. We have seen that the SKA+ELT configuration provides the largest number of EMcps. Therefore, to perform a closer comparison between results, we assume that the total luminosity is always $L_{\rm radio} = L_{\rm flare} + L_{\rm jet}$, i.e. we assume an isotropic jet, as it was computed in the original works.

Assuming 4 yr of LISA observations, they predict 208.6, 30.5 and 471 MBHBs for Pop3, Q3d and Q3nd, respectively. These values are not reported in the original paper but the authors shared the set of data with us. These values can be compared with the values reported in Tab. III. For the massive models the values are similar because these sources are typically detected with large SNR therefore different waveforms or sensitivity curves do not affect significantly the number of detected events.

However, they predict more detections in the Pop3 scenario. In order to understand this discrepancy we focus on: 1) the waveform and 2) LISA sensitivity. In B19 the authors adopted the PhenomC [131] to compute the SNR while we adopted the PhenomHM. The latter includes the contribution of higher harmonics in the GW signal, which are relevant at merger and ringdown. However Pop3 MBHBs are "light" systems so they merge at high frequency in LISA and we expect that the contribution from higher harmonics is negligible.

For the sensitivity curve they adopt the so-called "ESACall v1.1" [132] which performs better at high frequency with respect to the current "SciRDv1". Since the high frequency sensitivity impacts the detectability of light systems, a better sensitivity at those frequencies is expected to increase the number of detected events and might explain the discrepancy.

We can compare the number of EMcps in B19 with our results . This comparison is not trivial since the number of EMcps depends on the parameter estimation so we limit to report the general differences and propose qualitative explanations to address the different expected numbers of EMcps.

Even if we remove the Athena+ELT case, we predict 6.8, 14.9 and 20.9 EMcps, while in B19 they predict 13.6, 14.7 and 28.3 EMcps for Pop3, Q3d and Q3nd, respectively.

For Pop3 we predict $\sim 50\%$ less EMcps with respect to B19 due to the lower number of multimessenger candidates and to the more realistic parameter estimation. Note that the reduction probably is not due to the reduced number of detected events as these systems would be detected with low SNR so the sky localization would not be likely below the thresholds. For Q3nd we predict 30% fewer EMcps, compatible with the loss in the number of multimessenger candidates, while for Q3d the pre-

TABLE X. Numbers of EMcps in 4 yrs considering only systems with $q < 18$. In parenthesis, we report the rates from Tab. VI to ease the comparison.

	Maximising	Minimising
Pop3	5.5 (6.4)	1.3 (1.7)
Q3d	14.58 (14.8)	3.3 (3.4)
Q3nd	19.2 (20.7)	3.0 (3.4)

dictions are similar since these mergers happen at lower redshift.

In principle we can also perform the same comparison with the results in T16. However, in that work, the authors explored LISA ability to constrain cosmological parameters under different LISA configurations. Nowadays those configurations are out-of-date so a direct comparison is non trivial. The configuration closer to the SciRDv1 is the one labeled as 'N2A2M5L6'. Implementing the old LISA curve and redoing the parameter estimation is extremely expensive, so we limit to point out that the old LISA sensitivity performs better than the current one at high frequency, leading to a significant increase in the number of EMcps, especially for Pop3.

Appendix B: GW analysis

In this appendix we focus on the GW analysis of our systems. In Fig. 16 we present the average number of multimessenger candidates and EMcps as function of the SNR and mass ratio q. As a general trend, highest SNR values lead to a better parameter estimation, so the sky localization requirement selects naturally the systems with largest SNR as EMcps. In Pop3 model we have multimessenger candidates even at small SNR, however their parameter estimation is not good enough to promote them at the rank of EMcps while in the massive models the multimessenger candidates have always SNR > 10. Therefore there might be a population of low mass systems that will not be detected by LISA but that might be suitable for multimessenger studies and be accessible with Einstein Telescope [133] or future deci-hertz [134] detectors.

About the mass ratio, the multimessenger candidates and EMcps distributions for the Q3d and Q3nd models are similar to the total distribution from the catalogues. However, for Pop3, the detectability of the EM counterpart selects also systems with $q > 100$: these binaries typically host a massive BH, which explains the large mass ratio, and they are at small redshift because such massive BHs require time to form in Pop3 model.

From our distribution it is evident that there is a subset of MBHBs with very large mass ratio, up to $q \sim 10^2 - 10^3$. As described in Sec. VI, we adopted the PhenomHM for the GW analysis. This waveform is calibrated with numerical relativity simulations performed up to mass ratio 1 : 18 so the signals for $q > 18$ are based on an ex-
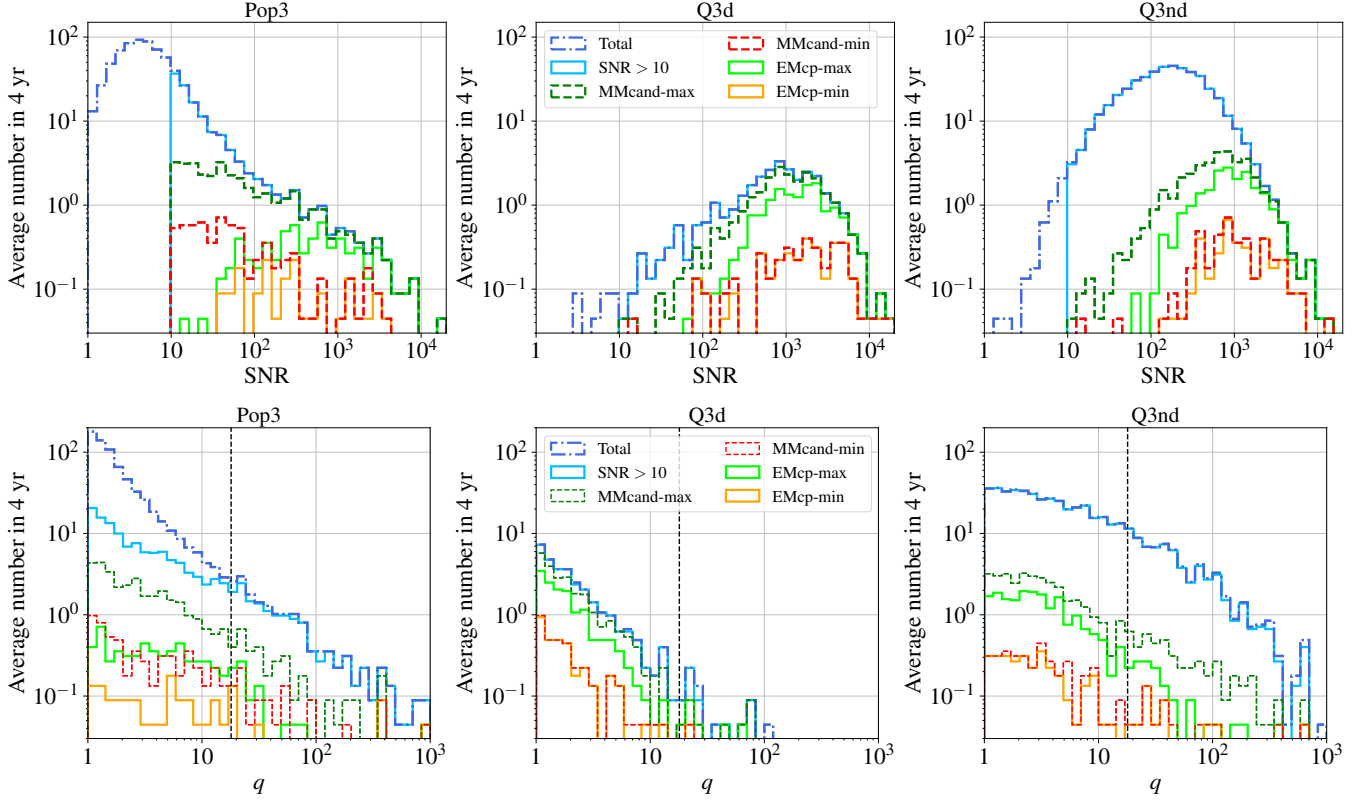
FIG. 16. Same as Fig. 1 for the SNR and q. In the lower panels, the black dashed line is at $q = 18$. The requirements on the sky localization and detectability of the EMcps select the systems with largest SNR and mass ratio close to unity.

trapolation of the current results and might not be representative of the estimates that an accurate waveform would produce in this range. For this reason, in Tab. X we report the number of EMcps in the maximising and minimising models if we consider only the events with $q < 18$. For Q3d the numbers remain unchanged due to the fact that the vast majority of the systems has $q < 18$, however the approximation in the waveform might affect the parameter estimation of $\sim 20\%$ ($\sim 10\%$) of the cases in Pop3 (Q3nd) astrophysical model. Finally the PhenomHM is validated for spin magnitudes up to 0.98 for

equal mass binaries. Even if in our simulations we have MBHBs with $\chi \sim 0.99$, they are a minority and we do not expect significant differences from the parameter estimation.

Appendix C: Useful figures for discussion

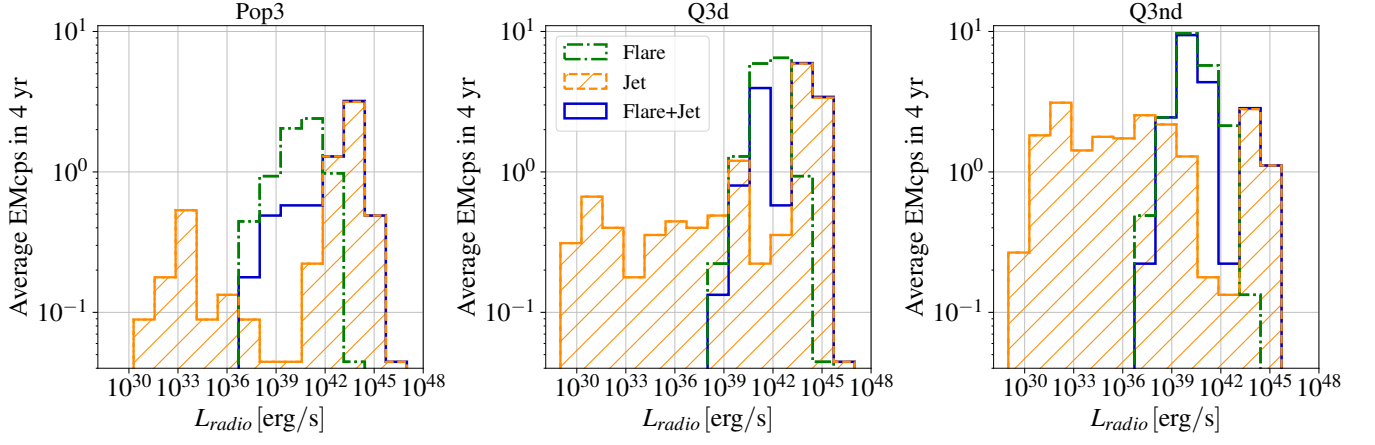In this appendix, we limit to report some figures that are interesting for the discussion in Sec. VII.

FIG. 17. Average number of EMcps as function of the luminosity of the radio emissions without imposing that the observer must be on axis. We show the EMcps emission distributions of the jet, flare and of their sum, detected with SKA+ELT for the three astrophysical scenarios, as clarified by the legend. At high luminosity, the total emission overlaps completely with the jet luminosity, which dominates the bright end of the distribution. The presence of a floor value of $\epsilon_{\text{Edd}} = 0.02$ in the flare expression (Eq. 15) further modulates the flare luminosity, avoiding $L_{\text{flare}} < 10^{36}$ erg/s values.
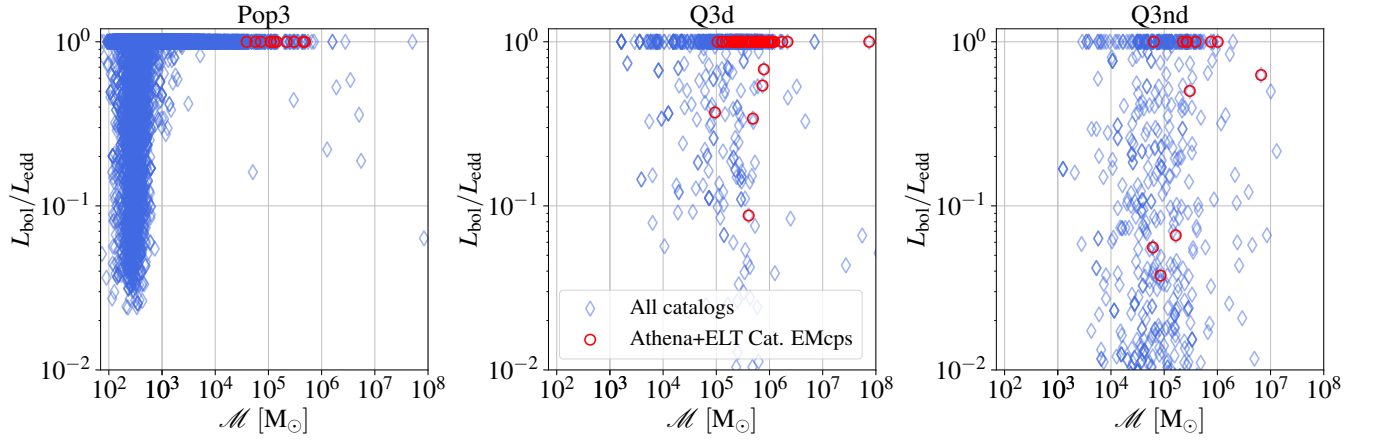

FIG. 18. Eddington ratio values as function of the chirp mass for the entire catalog (blue diamonds) and for EMcps detected by Athena+ELT in the case where the accretion is derived from the catalogues (red circles), as clarified by the legend, for the three astrophysical models.
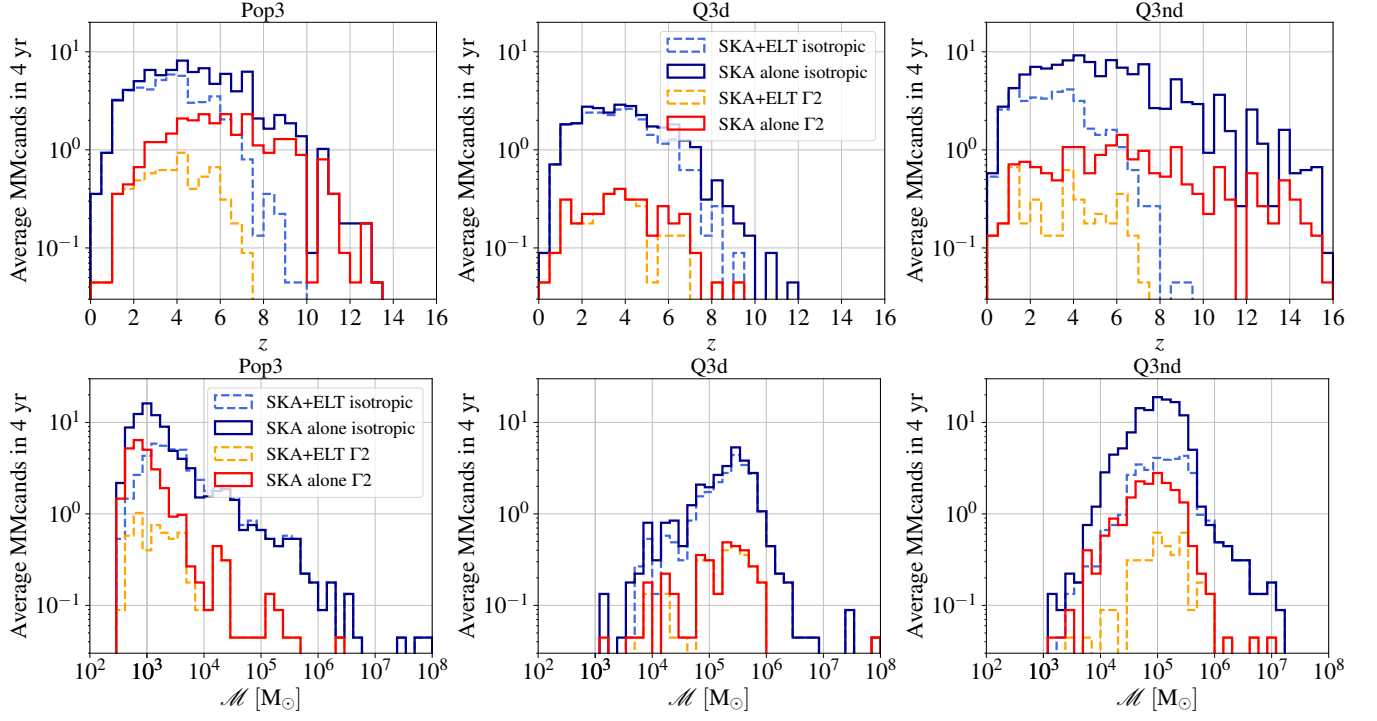
FIG. 19. Average number of MMcands in each observational scenario, as clarified by the legend, as function of redshift (upper panels) and chirp mass (lower panels) for the three astrophysical models assuming 4 yr of LISA time mission.