Gravitational wave emission from binary supermassive black holes

A Sesana
Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institut, Am Münchner 1, D-14476 Golm, Germany
E-mail: alberto.sesana@aei.mpg.de

Received 13 July 2013, in final form 29 October 2013
Published 29 November 2013
Online at stacks.iop.org/CQG/30/244009

Abstract
Massive black hole binaries (MBHBs) are unavoidable outcomes of the hierarchical structure formation process, and, according to the theory of general relativity, are expected to be the loudest gravitational wave (GW) sources in the Universe. In this paper I provide a broad overview of MBHBs as GW sources. After reviewing the basics of GW emission from binary systems and of MBHB formation, evolution and dynamics, I describe in some details the connection between binary properties and the emitted gravitational waveform. Direct GW observations will provide an unprecedented wealth of information about the physical nature and the astrophysical properties of these extreme objects, allowing to reconstruct their cosmic history, dynamics and coupling with their dense stellar and gaseous environment. In this context I describe ongoing and future efforts to make a direct detection with space based interferometry and pulsar timing arrays, highlighting the invaluable scientific payouts of such enterprises.

PACS numbers: 04.70.−s, 98.65.Fz, 04.30.−w, 04.30.Db, 04.30.Tv, 04.80.Nn
(Some figures may appear in colour only in the online journal)

1. Introduction

Today, massive black holes (MBHs)\(^1\) are ubiquitous in the nuclei of nearby galaxies [1], and we see them shining as quasars along the whole cosmic history up to redshift \(z \approx 7\) [2]. In the last decade, MBHs were recognized as fundamental building blocks in hierarchical models of galaxy formation and evolution, but their origin remains largely unknown. In fact, our current knowledge of the MBH population is limited to a small fraction of objects: either those that are active, or those in our neighborhood, where stellar- and gas-dynamical measurements are

\(^1\) It is customary to use the adjective ‘supermassive’ for the \(10^6−10^9\, M_\odot\) black holes powering quasars; conversely, we deal here with all the mass spectrum of these objects, down to their \(\sim 10^4\, M_\odot\) seeds. We therefore refer to them generally as ‘massive black holes’ throughout the paper.
possible. According to the current paradigm, structure formation proceeds in a hierarchical fashion [3], in which massive galaxies grow by accreting gas through the filaments of the cosmic web and by merging with other galaxies. As a consequence, the MBHs we see in today’s galaxies are expected to be the natural end-product of a complex evolutionary path, in which black holes (BHs) seeded in proto-galaxies at high redshift grow through cosmic history via a sequence of MBH-MBH mergers and accretion episodes [4, 5]. In this framework, a large number of MBH binaries (MBHBs) naturally form following the frequent galaxy mergers.

According to Einstein’s theory of general relativity (GR), accelerating masses cause modifications of the spacetime that propagate at the speed of light, better known as gravitational waves (GWs). However, in Einstein’s equations, the matter–metric coupling constant is of the order of $G/c^4$ (where $G$ is the gravitational constant and $c$ is the speed of light), which is $\sim 10^{-50}$! As a matter of fact the spacetime is extraordinarily stiff, therefore only massive, compact astrophysical object can produce a sizable strain that would be observable with advanced technology [6]. Two MBHs orbiting each other in a bound binary (MBHB) system carry a huge time varying quadrupole moment and are expected to be the loudest GW sources in the Universe [7]. The frequency spectrum of the emitted radiation covers several orders of magnitude, from the sub-nano-Hertz up to the milli-Hertz. The $10^{-4}–10^{-1}$ Hz window is going to be probed by spaceborne interferometers like the recently proposed European eLISA [8–10]. At $10^{-9}–10^{-7}$ Hz, joint precision timing of several ultrastable millisecond pulsars (i.e. a pulsar timing array, PTA) provides a unique opportunity to get the very first low-frequency detection. The European Pulsar Timing Array (EPTA) [11], the Parkes Pulsar Timing Array (PPTA) [12] and the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [13], joining together in the International Pulsar Timing Array (IPTA) [14], are constantly improving their sensitivities, getting closer to their ambitious target.

This focus issue contribution aims at covering all the relevant aspects of MBHBs intended as GW sources, in the spirit of providing a broad overview. Given the extent of the topic, we will just skim through its several facets, providing the appropriate references for in-depth reading. In section 2 we introduce the concept of GWs at a very basic level, defining the relevant astrophysical scales implied by MBHBs. A general overview of MBH formation and evolution (both their masses and spins), together with a brief description of MBHB dynamics is provided in section 3. We return on GWs in section 4, where we describe in deeper detail the expected signal from a MBHB and its dependences on the relevant parameters of the system. There, we also introduce some basics of parameter estimation theory, showing how the rich astrophysical information enclosed in individual signals can be recovered. Section 5 is then devoted to the scientific payouts of GW detection both in the milli-Hertz and in the nano-Hertz regime. We wrap-up in section 6 with some brief final remarks.

2. Gravitational waves: basics

The existence of GWs was one of the first predictions of Einstein’s GR, since they arise as natural solutions of the linearized Einstein equations in vacuum. Expanding the metric tensor as

$$ g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{1} $$

where $\eta_{\mu\nu}$ represent the Minkowski flat metric and $\| h_{\mu\nu} \| \ll \| \eta_{\mu\nu} \|$, and switching to the appropriate Lorentz gauge, the perturbation $h_{\mu\nu}$ satisfies

$$ \Box h_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}, \tag{2} $$
where □ is the d’Alambertian operator and the source term $T_{\mu\nu}$ is the stress–energy tensor. Equation (2) represents a set of wave equations and therefore admits wave solutions. These solutions are ripples in the fabric of the spacetime propagating at the speed of light: GWs. GWs are transverse, i.e. they act in a plane perpendicular to the wave propagation, and (at least in GR) have two distinct polarizations, usually referred to as $h_+$ and $h_\times$ (see cartoon in [6] and section 4.2). By expanding the mass distribution of the source into multipoles, conservation laws enforce GWs coming from the mass monopole and mass dipole to be identically zero, so that the first contribution to GW generation comes from the mass-quadrupole moment $Q$. The GW amplitude is therefore proportional to the second time derivative (acceleration) of $Q$.

Moreover, energy conservation enforces the amplitude to decay as the inverse of the distance to the source, $D$. A straightforward dimensional analysis shows that the amplitude of a GW is of the order of [15]

$$h = \frac{G}{c^2} \frac{1}{D} \frac{d^2 Q}{dt^2},$$

(3)

In order to generate GWs we therefore need accelerating masses with a time varying mass-quadrupole moment. The prefactor $G/c^4$ implies that these waves are tiny, so that the only detectable effect is produced by massive compact astrophysical objects.

We now specialize to the case of a binary astrophysical source. From now on, we shall use Geometric units $c = G = 1$; in such units 1 M⊙ = 4.927 × 10^6 s and 1 pc = 9.7 × 10^5 s. For sources at cosmological distances (i.e., with non-negligible redshift $z$), all the following equations are valid if all the masses are replaced by their redshifted counterparts (e.g., $M \rightarrow M_z = M(1+z)$), the distance $D$ is taken to be the luminosity distance (i.e., $D \rightarrow D_L = D(1+z)$) and $f$ is kept to be the observed GW frequency (the frequency in the source emission restframe is then $f(1+z)$). Keeping this in mind, consider a binary system of masses $M_2 < M_1$, mass ratio $q = M_2/M_1$ and total mass $M = M_1 + M_2$, in circular orbit at a Keplerian frequency $f_K$ at a distance $D$ to the observer. A detailed calculation (in the quadrupole approximation) shows that the system emits a monochromatic wave at a frequency $f = 2f_K$, with inclination–polarization time averaged GW strain given by [16]:

$$h = \sqrt{\frac{32}{5}} \frac{\mathcal{M}^{5/3}(\pi f)^{2/3}}{D},$$

(4)

where we introduced the chirp mass $\mathcal{M} = (M_1M_2)^{3/5}/(M_1 + M_2)^{1/5}$. For a pair of Schwarzschild BHs, the maximum frequency of the wave is emitted at the innermost stable circular orbit (ISCO) and can be written as:

$$f_{\text{ISCO}} = (\pi 6^{3/2} M)^{-1},$$

(5)

where $M_6 = M/10^6$ M⊙. GWs carry away energy from the system, with total luminosity given by [16]:

$$L_{\text{gw}} = \frac{dE_{\text{gw}}}{dt} = \frac{32}{5} (\pi \mathcal{M} f)^{10/3}. $$

(6)

Equating the energy loss to the shrinking of the binary semimajor axis $a$,

$$\frac{1}{E} \frac{dE_{\text{gw}}}{dt} = -\frac{1}{a} \frac{da}{dt},$$

(7)

and converting $a$ into frequency using Kepler’s law yields

$$\frac{df}{dt} = \frac{64}{5} \pi^{8/3} \mathcal{M}^{5/3} f^{11/3}.$$ 

(8)

The integral of equation (8) from $f_0$ to $f_{\text{ISCO}}$ defines the remaining lifetime of a binary emitting at a frequency $f_0$ before its final coalescence.
Putative eccentricity plays an important role in the evolution of the binary and in the emitted GWs [17]. In this case, the luminosity at a fixed MBHB semimajor axis is boosted to $L_{gw} = L_{gw,c} F(e)$, where $L_{gw,c}$ is given by equation (6) and

$$F(e) = (1 - e^2)^{-7/2} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right).$$

(9)

Accordingly, the evolution of the binary orbit is a factor $F(e)$ faster than in the circular case. The energy is radiated in form of a rather complicated GW spectrum, covering the spectral range $nf_k$, where $n$ is an integer index (see, e.g., [18], and section 4.2 for more details). In particular, the emission is stronger close to the binary periastron (there, the acceleration is larger, and so is the derivative of the quadrupole moment of the source), which leads to efficient circularization according to

$$\frac{d e}{d t} = -\frac{304}{15} \frac{M_1 M_2 M}{a^4 (1 - e^2)^{5/2}} \left( 1 + \frac{121}{304} e^2 \right).$$

(10)

Examples of GW driven circularization are illustrated in the MBHB evolutionary paths shown in panels ‘a3’ and ‘b3’ of figure 2.

Normalizing equation (4) to typical astrophysical MBHB values gives a strain of

$$h \approx 2 \times 10^{-18} D_9 \frac{M_6^{5/3}}{f_4^{1/3}}.$$

(11)

where $D_9 = D/10^9$ pc and $f_4 = f/10^{-4}$ Hz. If we require (somewhat arbitrarily) a coalescence timescale $\sim T_9 = T/10^9$ yr, the integral of equation (8), together with equation (5) implies a relevant frequency range

$$f_{\text{min}} = 3.54 \times 10^{-8} T_9^{-3/8} M_6^{-5/8} \text{Hz} < f < f_{\text{ISCO}} = 4.4 \times 10^{-3} M_6^{-1} \text{Hz}.$$

(12)

If we now estimate the frequency change in an observation time $T$ as $\Delta f \approx \dot{f} T$, we find

$$\Delta f = 5 \times 10^{-4} M_6^{5/3} f_4^{-1/3} T_9 \text{Hz}.$$

(13)

where now $T_9 = T/10^9$ yr. Equations (11)–(13) define the properties of typical MBHB signals. A light $10^3 M_\odot$ binary spans a frequency range of $10^{-7} - 10^{-2}$ Hz, in a Gyr before the coalescence. For a source at 1 Gpc, the strain is $h \approx 2 \times 10^{-18}$ at $10^{-3}$ Hz, and $\Delta f \gg f$, implying a chirping signal rapidly sweeping through the frequency band during a putative observation. On the other hand, a massive $10^9 M_\odot$ binary covers a frequency range of $10^{-10} - 10^{-6}$ Hz, in a Gyr before the coalescence. For a source at 1 Gpc, the strain is $h \approx 5 \times 10^{-16}$ at $10^{-8}$ Hz, and $\Delta f \approx 10^{-13}$ Hz, implying a non-evolving, monochromatic source. It is likely that many such sources accumulates at these low frequency, resulting in an incoherent superposition of monochromatic waves. MBHBs are therefore loud primary GW targets in the nano-Hertz–milli-Hertz regime, and they can manifest themselves both as rapidly chirping signals (at milli-Hertz frequencies) as well as incoherent superposition of monochromatic waves (at nano-Hertz frequencies).

3. Massive black hole binaries

The mechanism responsible for the formation of the first seed BHs is not well understood. These primitive objects started to form at the onset of the cosmic dawn, around $z \sim 20$, according to current cosmological models [19]. At an epoch of $z \sim 30 - 20$, the earliest stars formed in small, metal-poor protogalactic halos may have had masses exceeding $100 M_\odot$ [20], ending their lives as comparable stellar mass BHs, providing the seeds that would later grow into MBHs [21]. However, as larger, more massive and metal enriched galactic discs progressively formed, other paths for BH seed formation became viable (see [22] for a review).
Global gravitational instabilities in gaseous discs may have led to the direct formation of $10^5 M_\odot$ BH seeds [23], or to the formation of quasi-stars of $10^3$–$10^6 M_\odot$ that later collapsed into seed BHs [24]. Alternative scenarios are the collapse of massive stars formed in run-away stellar collisions in young, dense star clusters [25] or the collapse of unstable self-gravitating gas clouds in the nuclei of gas-rich galaxy mergers at later epochs [26]. Thus, the initial mass of the seeds remains one of the largest uncertainties in the present theory of MBH formation. However, once formed, these seed BHs inevitably took part in the hierarchical structure formation process, growing along the cosmic history through a sequence of mergers and accretion episodes [4, 5]. Figure 1 shows examples of expected MBHB merger rates as a function of redshift for a sample of selected seed BH models ([5, 24, 27], see [29] for details). The uncertainty is large, with numbers ranging from ten to several hundreds events per year. Multiplying by the Hubble time and dividing by the number of galaxies within our Hubble horizon ($\approx 10^{11}$), figures imply that each galaxy experienced few to few hundred mergers in its past life, placing mergers among the crucial mechanisms in galaxy evolution.

3.1. Mass and spin evolution

Astrophysical BHs are extremely simple objects, described by two quantities only, namely their mass, $M_{\text{BH}}$, and angular momentum, $S$. The magnitude of the latter can be expressed by the dimensionless parameter $a = S/S_{\text{max}} = cS/GM_{\text{BH}}$. By definition $0 \leq a \leq 1$. Along the cosmic history, MBH mass and spin inevitably evolve according to three principal evolution mechanisms: (i) merger with other MBHs, (ii) episodic accretion of compact objects, disrupted stars, or gas clouds, and (iii) prolonged accretion of large supplies of gas via accretion discs. As

---

Figure 1. Differential MBHB merger rate as a function of redshift for different seed formation scenarios. Adapted from [28].
In this experiment, a parameter $F$ (given in each panel) defines the fraction of events in the southern hemisphere (defined with respect to the orientation of the MBH spin). $N \times (1 - F)$ accretion events are then isotropically distributed in the northern hemisphere, and $N \times F$ in the southern hemisphere ($F = 0.5$ for an isotropic distribution on the sphere). In each panel, the black line refers to the mean over 500 realizations, and red and orange shaded areas enclose intervals at 1-$\sigma$ and 2-$\sigma$ deviations, respectively (from [35]). Shown are the evolution of the eccentricity and binary semimajor axis in time (‘1’ and ‘2’ panels) and the evolution of the eccentricity versus the orbital frequency (‘3’ panels). In each panel we show tracks for binaries with initial eccentricity of 0.1 (solid), 0.3 (short-dashed) and 0.6 (long-dashed). The very high eccentricities achieved in the stellar driven phase imply non-negligible eccentricities in the milli-Hertz regime (probed by future space based interferometers like eLISA, shaded area in panel ‘a3’), and potentially extremely high eccentricities in the nano-Hertz regime (targeted by PTAs, shaded area in panel ‘b3’).

shown by [29], coalescences of MBHs with random spin directions result in a broad remnant spin distribution; in particular highly spinning MBHs tend to spin-down. Despite the important of MBH-MBH mergers, the dominant role in the mass and spin evolution of MBHs can be attributed to accretion. Continuous Eddington limited accretion implies an exponential mass growth $M_{\text{BH}}(t) = M_{\text{BH}}(0) \exp \left( \frac{t}{t_{\text{Edd}}} \right)$, where $t_{\text{Edd}} = 0.45$ Gyr and $\epsilon$ is the mass-radiation conversion efficiency ($0.06 < \epsilon < 0.4$ for $0 < a < 1$). If this happens in a coherent fashion through, e.g., a thin disc [30], an initially Schwarzschild BH becomes maximally spinning after accreting an amount of mass of the order of $\sqrt{6}M_{\text{BH}}(0)$ [31]. However high spins imply $\epsilon > 0.3$, considerably slowing down the mass growth, making it impossible to produce a MBH of $>10^9 M_\odot$ at $z = 7$ (i.e. in $<10^9$ yrs). The problem is avoided if mass is accreted in a series of small incoherent packets (chaotic accretion [32]). In this case, depending on the angular momentum of the accreted material, the MBH is spun up or down, performing a random walk in spin magnitude that keep it close to zero. However this is true only if the angular momentum direction of the packets is nearly isotropically distributed on the sphere. Real galaxies usually show large coherent gas structures, and a significant amount of rotation (see, e.g., [33, 34]). If the spin vectors of the accreting packets have, on average, a preferential direction (i.e., they angular momenta do not sum up to zero), then the spin evolution is more complicated, and high spin values might still be preferred. This is shown in the left panel of figure 2, taken from [35]. At low masses, the angular momentum of the accreted packet dominates over the MBH spin; in this case the latter always rapidly align with the former, resulting in efficient spin-up.
However, this is not true anymore at larger MBH masses: now the MBH spin is larger than the angular momentum of the accreted packet and alignment does not occur. In this latter regime retrograde accretion causes efficient spin-down, and the final outcome depends on the fraction of prograde versus retrograde events, as shown in the four boxes of the left panel of figure 2 (see [35] for full details). Rapid mass growth is difficult to reconcile with measurements of high spins (although the latter involve galaxies in the local Universe [36]), and the requirement of high spins to power energetic relativistic jets in many theoretical models (e.g., [37]), and a complete joint understanding of the MBH mass and spin evolution is still missing.

3.2. Massive black hole binary dynamics

MBHs become loud sources of GWs when they are in bound, sub–pc binaries, which are the natural end-product of galaxy mergers. However, to get there, the two MBHs need to ‘find each other’ starting from kpc distances, form a bound pair, and shrink to mpc separations, where the emission of GWs become efficient leading to final coalescence. The two MBHs find each other thanks to dynamical friction that helps them sinking to the center of the merger remnant, to the point they start to feel each other gravity and form a Keplerian binary. At this point (around a pc separation), dynamical friction ceases to be effective and the need to find alternative mechanisms to shrink the binary went under the name of ‘last parsec problem’ (see, e.g., [38]). The general consensus is now that interaction of the MBHB with its dense stellar and/or gaseous environment provides effective means to extract the binary energy and angular momentum, leading the system to the efficient GW emission regime (see [39] for an updated review). Such strong MBHB–environment interaction inevitably imprints distinctive signatures in the binary orbital elements and in the individual spins of the holes. We will see later how this information can be recovered by GW observations, which will therefore allow to directly probe the complex physics underlying the evolution of these spectacular objects.

3.2.1. Stellar driven binaries. Ignoring technical details related to the ‘loss cone evolution’ (see D Merritt contribution to this issue), a background of stars scattering off the binary drives its semimajor axis evolution according to the equation [41]

$$\frac{da}{dt} = \frac{a^2 G \rho}{\sigma} H,$$

(14)

where $\rho$ is the density of the background stars, $\sigma$ is the stellar velocity dispersion and $H$ is a numerical coefficient of order 15. The eccentricity evolution in stellar environments has also been tackled by several authors by means of full $N$-body simulations. In general, equal mass, circular binaries tend to stay circular or experience a mild eccentricity increase [42], while binaries that form already eccentric, or with $q \ll 1$ (regardless of their initial eccentricity) tend to grow more eccentric [43, 44], in reasonable agreement with the results of 3-body scattering experiments [41, 45]. Interactions with stars are unlikely to significantly affect the individual spins of the holes. Therefore, star driven binaries are expected to grow to quite high eccentricities, while the spins of the individual holes are likely randomly oriented.

3.2.2. Gas driven binaries. In the case of circumbinary discs, the detailed evolution of the system depends on the complicated and uncertain dissipative physics of the disc itself. The simple case of a coplanar prograde circumbinary disc admits a self-consistent, non-stationary solution that was derived by [46]. In this case, the binary semimajor axis evolution can be approximated as [46, 47]

$$\frac{da}{dr} = \frac{2M}{\mu} (a a_0)^{1/2}.$$

(15)
Here, $\dot{M}$ is the mass accretion rate at the outer edge of the disc, $a_0$ is the semimajor axis at which the mass of the unperturbed disc equals the mass of the secondary MBH, and $\mu$ is the reduced mass of the binary. In the circumbinary disc scenario, eccentricity excitation has been seen in several simulations [48, 49]. In particular, the existence of a limiting eccentricity $e_{\text{crit}} \approx 0.6–0.8$ has been found in [50], in the case of massive self-gravitating discs. If the accretion flow is coherent, and $L_{\text{disc}}$ and the spins $S_i$ of the two MBHs are misaligned, the Bardeen–Petterson effect [51] will act to align $S_i$ to $L_{\text{disc}}$ in a very short timescale ($t_{\text{align}} \ll t_{\text{acc}} \sim 10^8$ yr [52, 53]). Therefore, in gaseous rich environments, mildly eccentric binaries might be the norm, and the MBH individual spins tend to align with the orbital angular momentum.

Compared to the GW driven case, $(da/dt)_{gw} \propto a^{-3}$, equations (14) and (15) have a very different (milder and positive) $a$ dependence. Therefore, equating equations (14) and (15) to $(da/dt)_{gw}$ gives the transition frequency between the external environment driven and the GW driven regimes. For typical astrophysical systems one gets:

$$f_{\text{star/GW}} \approx 1.2 \times 10^{-7} M_6^{-7/10} q^{-3/10} \text{ Hz}$$
$$f_{\text{gas/GW}} \approx 1.6 \times 10^{-7} M_6^{-37/49} q^{-49/98} \text{ Hz}.$$  

(16)

Substituting $M = 10^9$ $M_\odot$, we see that the transition occurs at nano-Hertz frequencies. This is the mass and frequency range of typical PTA sources, which therefore might still be influenced by their environment and be very eccentric ($e > 0.5$, see panel 'b3' in figure 2). Even though GW emission efficiently circularizes binaries (see section 2), low mass systems ($M = 10^4–10^5$ $M_\odot$) decouple only at $\mu$Hz frequencies and can still retain substantial residual eccentricities in the milli-Hertz window where they become eLISA targets ($e > 0.01$, see panel 'a3' in figure 2).

4. MBHB waveforms

Having introduced the basics of GW emission from a binary system in section 2, we turn now in some more detail to the gravitational waveform modeling. In particular we show how eccentricity and spins affect the detectable GW signal and we describe the basic theory of information recovery, that enables us to dig out the parameters of the source from the detected waveform.

4.1. The stages of the binary coalescence

The evolution of MBHBs is customarily divided into three phases: inspiral, merger, and ringdown [55]. The inspiral is a relatively slow, adiabatic process. Different techniques have been employed to describe this stage, ranging from classic Post Newtonian (PN) expansions of the energy-balance equation [56], to non-adiabatic resummed methods in which the equations of motion are derived from an effective one body (EOB) relativistic Hamiltonian [57]. A detailed description of such methods is beyond the scope of this paper, and an excellent overview can be found in [58]. The inspiral is followed by the dynamical coalescence, in which the MBHs plunge and merge together, forming a highly distorted, perturbed remnant. At this stage, all analytical approximations break down, and the system can only be described by directly solving the Einstein equations using numerical simulations [59–61]. The distorted remnant settles into a stationary Kerr BH as it rings down, by emitting gravitational radiation. This latter stage can be, again, modeled analytically using BH perturbation theory [62]. An example of the full waveform with the identification of the various stages is given in figure 3.

In recent years there has been a major effort in constructing accurate waveforms inclusive of all three phases. ‘Complete’ waveforms can be designed by stitching together analytical...
Figure 3. Example of GW signal from two coalescing (circular, non-spinning) BHs as a function of time. The different approximation techniques and their range of validity are indicated. Wavy lines illustrate the regime close to merger where analytical methods have to be bridged by NR (courtesy of Ohme [54]).

PN waveforms for the early inspiral with a (semi)phenomenologically described merger and ring-down phase calibrated against available numerical data (known as PhenomB-PhenomC waveforms, [63]). Alternatively, complete waveforms can be constructed within the EOB formalism by adding free parameters to be calibrated against NR simulations and by attaching a series of damped sinusoids describing the ring-down (known as EOBNR waveforms [64, 65]). A detailed overview is given in [54]. What is relevant to our discussion is that the full evolution of MBHBs can be tackled with a combination of analytical and numerical methods, and accurate waveforms encoding all the parameters of the system can be computed. In the following we concentrate on the inspiral signal only, which is the richest in terms of encoded information.

4.2. The adiabatic inspiral: impact of eccentricity and spin

Expanding the Einstein’s equations and the covariant conservation of the stress–energy tensor $T_{\mu\nu}$ perturbatively in powers of $v/c$ (where $v$ is the relative velocity of the two objects), one gets a PN series for the binary acceleration of the form [66]:

$$a = a_N + a_{\text{PN}} + a_{\text{SS}} + a_{\text{RR}},$$

(17)

where $a_N$, $a_{\text{PN}}$, and $a_{\text{2PN}}$ are the Newtonian, (post)1-Newtonian, and (post)2-Newtonian contributions to the equations of motion, $a_{\text{RR}}$ is the contribution due to the radiation-reaction force, and $a_{\text{SO}}$ and $a_{\text{SS}}$ are the spin–orbit and spin–spin coupling contributions. In the same way, the radiated wave (in the far field zone, see [56]) can be similarly expanded in the form:

$$h^{ij} = \frac{2}{D} [\ddot{Q}^{ij} + P^{0.5} \dot{\dot{Q}}^{ij} + P\ddot{Q}^{ij} + P\ddot{Q}_{SO}^{ij} + P^{1.5} \dot{\dot{Q}}^{ij} + P^{2.0} \dddot{Q}_{SO}^{ij} + P^{2.5} \dddot{Q}_{SS}^{ij}]$$

(18)

where $\dddot{Q}^{ij}$ is the second time derivative of the standard quadrupole moment of the source, $i, j = 1, 2, 3$ define the spatial components of the perturbation tensor, and the subscripts have the same meaning as in equation (17).

A practical way to derive expressions for the acceleration and the waveform, is to employ the adiabatic approximation, in which the inspiral of the system is treated as a quasistationary sequence of orbits. For circular binaries, the evolution of the adiabatic inspiral is completely determined by the energy-balance equation that relates the derivative of the energy function $\dot{E}$ to the gravitational flux $\mathcal{F}$ radiated away:

$$\frac{d\dot{E}(v)}{dt} = -\mathcal{F}(v)$$

(19)

For eccentric binaries, an angular momentum balance equation must also be imposed to compute the evolution in eccentricity.
from which we may derive the binary acceleration and phase evolution as:

$$\frac{d\Phi}{dt} = \frac{2v^3}{M}, \quad \frac{dv}{dt} = -\frac{\mathcal{F}(v)}{M\mathcal{E}(v)/dv}. \tag{20}$$

For orbital velocities $v \ll c$, $\mathcal{E}$ and $\mathcal{F}$ can be expanded in powers of $v^2n$ to a given order in $n$. Similarly, the resulting waveform is written as an expansion of the form (18), where now $h^{11}$ do not explicitly depend on the secular adiabatic inspiral. Choosing the appropriate orthonormal radiation frame, $h^{11}$ can be written as two independent polarizations only; $h_+$ and $h_\times$. Assuming a circular binary, in the adiabatic approximation, to the leading quadrupole order one obtains the familiar form

$$h_+(t) = 2\frac{\lambda f^{5/3}}{D} - [{\pi f(t)}]^{2/3}(1 + \cos^2 \iota) \cos \Phi(t)$$

$$h_\times(t) = 2\frac{\lambda f^{5/3}}{D} - [{\pi f(t)}]^{2/3}2\sin^2 \iota \cos \Phi(t) \tag{21}$$

where $\iota$ (usually referred as inclination) is the angle defined by the line of sight with respect to the orbital angular momentum vector, $\Phi(t) = 2\pi \int f'(t') dt'$, and $f = 2f_K$ as defined in section 2.

The eccentricity $e$ enters directly in the computation of $Q^{ij}$ since it affects the velocity $v$ of the MBHs along the orbit. In fact, $e$ affects the computation of $h^{ij}$ at all orders, starting from the simple quadrupole term, by ‘splitting’ each polarization amplitude $h_+(t)$ and $h_\times(t)$ into harmonics according to (see, e.g., equations (5) and (6) in [67] and references therein):

$$h_+(t) = A \left( - (1 + \cos^2 \iota)u_n(e) \cos \left[ \frac{n}{2} \Phi(t) + 2\gamma(t) \right] 
- (1 + \cos^2 \iota)v_n(e) \cos \left[ \frac{n}{2} \Phi(t) - 2\gamma(t) \right] + \sin^2 \iota w_n(e) \cos \left[ \frac{n}{2} \Phi(t) \right] \right),$$

$$h_\times(t) = 2A \cos \iota \left( u_n(e) \sin \left[ \frac{n}{2} \Phi(t) + 2\gamma(t) \right] + v_n(e) \sin \left[ \frac{n}{2} \Phi(t) - 2\gamma(t) \right] \right). \tag{22}$$

The amplitude coefficients $u_n(e)$, $v_n(e)$, and $w_n(e)$ are linear combinations of the Bessel functions of the first kind $J_n(ne)$, $J_{n\pm 1}(ne)$ and $J_{n\mp 2}(ne)$, and $\gamma(t)$ is an additional precession term to the phase given by $e$. For $e \ll 1$, $|u_n(e)| \gg |v_n(e)|, |w_n(e)|$ and we recover the circular limit given by equation (21). As GW emission tends to decrease eccentricity, this is likely to be mostly important at large separations (i.e., $f_K \ll f_{\text{ISCO}}$, see panels ‘a3’ and ‘b3’ in figure 2). An example of an eccentric waveform is shown in the upper panel of figure 4.

Turning now to spins, equation (17) shows that spins enter as higher order corrections (namely at the 1.5PN order, i.e. $O(v/c)^3$) in the computation of the acceleration and, consequently, of the waveform. Therefore they become important only when $v/c$ is large. Such corrections generate additional small terms in the phase evolution, but most importantly, cause the orbital angular momentum $\mathbf{L}$ to precess around the total angular momentum. At the leading order, the latter is conserved in its Newtonian form $\mathbf{J} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2$ (in which $\mathbf{S}_1$ and $\mathbf{S}_2$ the spins of the two MBHs) and $\mathbf{L}$ precesses according to the equation [68]

$$\dot{\mathbf{L}} = \frac{1}{\kappa^3} \left( 2 + \frac{3}{2} q \right) \mathbf{J} \times \mathbf{L} \tag{23}$$

(in this case $\alpha$ is the binary semimajor axis). The typical precession timescale is given by

$$T_p = \frac{\mathcal{L}}{\mathcal{L}} \approx M^{1/3} \mu^{-1} \left[ 2 + \frac{3}{2} q \right]^{-1} (\pi f)^{5/3} \approx 2.5 \times 10^3 \pi^{5/3} M_6^{1/3} \left( \frac{f}{f_{\text{ISCO}}} \right)^{-5/3} \tag{24}$$
where we used $J \approx L$, and we assumed equal mass binaries in the last equality. It is clear that at, say, $f = 10^{-3} f_{\text{ISCO}}$, $T_p$ is of the order of several tens of years, making precession effects negligible on plausible observation timescales. An example of waveform modulated by plane precession in the late inspiral is shown in the lower panel of figure 4.

The most general detectable signal from a spinning eccentric binary is a function of 17 parameters (some describing the intrinsic properties of the binary and some others related to the relative binary-detector position and orientation): two combination of the redshifted masses, $\mathcal{M}$ and $\mu$;\(^5\) six parameters defining the individual spin vectors (two magnitudes $a_1$ and $a_2$ and four angles); two parameters related to the eccentricity of the orbit, initial eccentricity $e_0$ and an additional angle defining the line of nodes; source inclination with respect to the line of sight, $\iota$; polarization angle, $\psi$; sky location, two angle usually labeled $\theta$ and $\phi$;\(^6\) luminosity distance $D_L$; initial orbital phase $\Phi_0$; and, depending on the type of signal, initial frequency $f_0$ or time to coalescence $t_c$ (the two are related, in the quadrupole approximation, by equation (8)). We saw examples of how some of these parameters are imprinted in the waveform, in the following subsection we turn to the problem of how accurately they can be extracted given some signal observation.

\(^5\) As discussed in section 2, for sources at cosmological distances, the GW depends on the redshifted masses, the intrinsic ones can be extracted by measuring $D_L$ and then converting it into $z$ according to a cosmological model, or by obtaining an independent measurement of $z$ through, e.g., the identification of an electromagnetic counterpart to the GW signal.

\(^6\) Of particular interest is the source sky localization errorbox $\Delta \Omega$, defined, following [69], in terms of $\theta$ and $\phi$ as: $\Delta \Omega = 2\pi \sqrt{(\sin \theta \Delta \theta \Delta \phi)^2 - (\sin \theta)^2 (\Delta \phi)^2}$ (according to the notation used in section 4).
4.3. Parameter extraction

We briefly review the basic theory regarding the estimate of the statistical errors that affect the measurements of the source parameters. For a comprehensive discussion of this topic we refer the reader to [70]. The data \( d \) collected in a detector is given by the superposition of the noise \( n \) and a signal \( x \) determined by a vector of parameters \( \lambda \):

\[
d(t) = n(t) + x(t; \lambda).
\]

In the following, we make the usual (simplifying) assumption that the marginal likelihood or evidence. Under the assumption of Gaussian noise

\[
\hat{p}(\lambda|d) = \frac{p(\lambda)p(d|\lambda)}{p(d)},
\]

where \( p(d|\lambda) \) is the likelihood function, \( p(\lambda) \) is the prior probability density of \( \lambda \), and \( p(d) \) is the marginal likelihood or evidence. Under the assumption of Gaussian noise

\[
p(d|\lambda) = p(n = d - x(\lambda)) \propto \exp \left\{ -\frac{1}{2} (d - x(\lambda)|d - x(\bar{\lambda})) \right\}
\]

where the inner product between two functions \( (g|h) \) is defined as the integral

\[
(g|h) = 2 \int_0^\infty \tilde{g}^*(f) \tilde{h}(f) + \tilde{g}(f) \tilde{h}^*(f) \frac{S_n(f)}{S_n(f)} df,
\]

applied to the Fourier transform of the functions, e.g.,

\[
\tilde{g}(f) = \int_{-\infty}^{+\infty} g(t) e^{-2\pi i ft}.
\]

In the neighborhood of the maximum-likelihood estimate value \( \hat{\lambda} \), the likelihood function can be approximated as a multi-variate Gaussian distribution,

\[
p(\lambda|d) \propto p(\lambda) \exp \left\{ -\frac{1}{2} \Gamma_{ab} \Delta \lambda_a \Delta \lambda_b \right\},
\]

where \( \Delta \lambda_a = \hat{\lambda}_a - \lambda_a \) and the matrix \( \Gamma_{ab} \) is the Fisher information matrix (FIM). Here \( a, b = 1, \ldots, N \) label the components of \( \lambda \) (i.e., the parameters defining the shape of the signal), and we have used Einstein’s summation convention (and we do not distinguish between covariant and contravariant indices). The FIM is simply related to the derivatives of the GW signal with respect to the unknown parameters integrated over the observation:

\[
\Gamma_{ab} = \left( \frac{\partial x(t; \lambda)}{\partial \lambda_a} \right) \left( \frac{\partial x(t; \hat{\lambda})}{\partial \lambda_b} \right).
\]

In terms of the inner product \(<.,.>\) the maximal signal-to-noise ratio (SNR) at which a signal can be observed is obtained by matched filtering of the data against a template equal to the waveform signal, \( x(\lambda) \). The optimal matched filtering SNR achievable in this way is \( \langle x| \rangle \). In the limit of large SNR, \( \hat{\lambda} \) tends to \( \bar{\lambda} \), and the inverse of the FIM provides a lower limit to the error covariance of unbiased estimators of \( \lambda \), the so-called Cramer–Rao bound. The variance–covariance matrix is simply the inverse of the FIM, and its elements are

\[
\sigma^2_a = (\Gamma^{-1})_{aa}, \quad \epsilon_{ab} = \frac{(\Gamma^{-1})_{ab}}{\sqrt{\sigma^2_a \sigma^2_b}}.
\]
where \(-1 \leq c_{ab} \leq +1\) are the correlation coefficients. We can therefore interpret \(\sigma^2\) as a way to quantify the expected uncertainties on the measurements of the source parameters. We refer the reader to [71] and references therein for an in-depth discussion of the interpretation of the inverse of the FIM in the context of assessing the prospect of the estimation of the source parameters in GW observations. When combining \(\alpha = 1, \ldots, N\) different pieces of independent (i.e., having uncorrelated noise) information (for example, by observing several pulsars, or by combining the inspiral and the ring-down portion of a signal), the FIM that characterizes the joint observations in equation (30) is simply given by the sum of the matrices of all the individual pieces

\[ \Gamma_{ab} = \sum_{\alpha} \Gamma_{ab}^{(\alpha)}, \]

(33)

and all the theory follows unchanged.

### 5. The gravitational wave landscape: observations and scientific payouts

As discussed in section 2, GW emission from MBHBs covers several decades in frequency, ranging from sub-nano-Hertz to milli-Hertz. As shown in figure 5, this range is (or it will be) covered by multiple probes. The ground-based network of advanced interferometric detectors (three LIGO detectors, VIRGO [72], and the Kamioka Gravitational wave Detector, KAGRA [73]) and possibly the third-generation Einstein Telescope (ET, [74]) will observe inspiralling binaries up to around \(\lesssim 100 M_\odot\). The milli-Hertz regime will be the hunting territory of spaced based detectors such as eLISA. Acceleration noise however, will severely affect the sensitivity below 0.1 mHz, which, according to equation (5), roughly corresponds to the ISCO.
Figure 6. Constant-contour levels of the sky and polarization angle averaged SNR for eLISA, for equal mass non-spinning binaries as a function of their total rest frame mass, $M$, and cosmological redshift, $z$. The tracks represent the mass-redshift evolution of selected MBHs: two possible evolutionary paths for a MBH powering a $z \approx 6$ quasar (starting from a massive seed, blue curve, or from a Pop III seed from a collapsed metal-free star, yellow curve); a typical $10^9 \, M_\odot$ MBH in a giant elliptical galaxy (red curve); and a Milky Way-like MBH (green curve). Circles mark MBH-MBH mergers occurring along the way [75]. The gray transparent area in the bottom right corner roughly identifies the parameter space for which MBHs might power phenomena that will likely be observable by future electromagnetic probes.

of a $10^7 \, M_\odot$ binary. On the other hand, PTAs start to be sensitive at $10^{-6}$ Hz (this upper limit is set by the typical weekly cadence of millisecond pulsar observations) and extend their observational window all the way down to the nano-Hertz range, where inspiralling $10^8$–$10^9 \, M_\odot$ systems should be abundant. Therefore, space based interferometers and PTAs will provide a complementary, complete census of the MBHB population throughout the Universe. In this section we focus on the prospects of GW observation in these two bands and on the related scientific payouts.

5.1. The milli-Hertz regime: science with space based interferometry

Space based interferometry will open a revolutionary new window on the Universe. In the following we refer to the eLISA design presented in [10] to describe the extraordinary scientific payouts of milli-Hertz GW observations. Figure 6 highlights the exquisite capabilities of eLISA in covering almost all the mass-redshift parameter space relevant to MBH astrophysics. GW observations will catch sources with $M \sim 10^4 \, M_\odot$ at early cosmological times, prior to reionization. A binary with $10^4 \lesssim M \lesssim 10^7 \, M_\odot$ can be detected out to $z \sim 20$ with a SNR $\gtrsim 10$, making an extensive census of the MBH population in the Universe possible.

As detailed in section 4, detected waveforms carry information on all the relevant source parameters, including redshifted masses and spins of the individual BHs prior to coalescence, the distance to the source and its sky location. The left panel of figure 7 shows error distributions in the source parameter estimation, for events collected in a meta-catalog of $\sim 1500$ sources, based on state of the art MBH evolution models (see [8] for details). Here, circular precessing spinning binary were considered (i.e. waveforms determined by 15 parameters), and ‘hybrid’
Left panel: eLISA parameter estimation accuracy; meta-catalog of MBHBs described in [8]. Top panels show errors on the redshifted masses (left) and spins (right). Red solid lines are for the primary and blue dashed lines are for the secondary MBH. The bottom panels show the error distribution on the luminosity distance $D_L$ (left), and the sky location accuracy $\Delta \Omega$ (in deg$^2$, right).

Right panel: eLISA capabilities of selecting among different MBH formation routes as a function of observation time. Plotted is the fraction of realizations in which one of the four investigated models (SE) is chosen over each of the other three models (LE (solid green), LC (long-dashed blue) and SC (short-dashed red)) at 95% confidence level, as a function of observation time. No spin information is included.

Waveforms of the PhenomC family were used to evaluate uncertainties based on the FIM approximation, as outlined in the previous section. Individual redshifted masses can be measured with unprecedented precision, i.e., with an error of 0.1%–1%, on both components. The spin of the primary hole can be measured with an exquisite accuracy, to a 0.01–0.1 absolute uncertainty. This precision mirrors the big imprint left by the primary MBH spin in the waveform. The measurement is more problematic for $a_2$ that can be either determined to an accuracy of 0.1, or remain completely undetermined, depending on the source mass ratio and spin amplitude. The source luminosity distance error has a wide spread, usually ranging from being undetermined (but see [76] for possible shortcomings of the FIM approximation in these cases) to a stunning few per cent accuracy (note that this is a direct measurement of $D_L$). GW detectors are full sky monitors, and the localization of the source in the sky is also encoded in the waveform pattern. Sky location accuracy is typically estimated in the range 10–1000 deg$^2$.

While measurements of individual systems are extremely interesting and very useful for making, e.g., strong-field tests of GR, it is the properties of the whole set of MBHB mergers that are observed which will carry the most information for astrophysics. GW observations of multiple MBHBs may be used together to learn about their formation and evolution through cosmic history, as demonstrated by [77, 78]. We briefly provide here an illustrative example from [8]. As argued above, in the general picture of MBH cosmic evolution, the population is shaped by the seeding process and the accretion history. [8] therefore consider a set of four models with distinctive properties: (i) small seeds and extended (coherent) accretion (SE), (ii) small seeds and chaotic accretion (SC), (iii) large seeds and extended accretion (LE), (iv) large seeds and chaotic accretion (LC).

These numbers assume a ‘single Michelson’ (four laser links) configuration for eLISA, the full triangular ‘two Michelsons’ (six laser links) configuration results in a significant improvement of the estimation of all parameters, in particular luminosity distance and sky location (by one to two orders of magnitude).
Table 1. Summary of all possible comparisons of the considered MBH evolution models. Results are for one year of observation with eLISA. We take a fixed confidence level of \( p = 0.95 \). The numbers in the upper-right half of each table show the fraction of realizations in which the row model is chosen at more than this confidence level when the row model is true. The numbers in the lower-left half of each table show the fraction of realizations in which the row model cannot be ruled out at that confidence level when the column model is true. In the top table we consider the trivariate \( M, q, \) and \( z \) distribution of observed events; in the bottom table we also include the observed distribution of remnant spins, \( a_r \).

| Without spins | SE | SC | LE | LC |
|---------------|----|----|----|----|
| SE            | ×  | 0.48 | 0.99 | 0.99 |
| SC            | 0.53 | ×  | 1.00 | 1.00 |
| LE            | 0.02 | 0.01 | ×  | 0.79 |
| LC            | 0.02 | 0.02 | 0.22 | ×  |

| With spins | SE | SC | LE | LC |
|------------|----|----|----|----|
| SE         | ×  | 0.96 | 0.99 | 0.99 |
| SC         | 0.13 | ×  | 1.00 | 1.00 |
| LE         | 0.01 | 0.01 | ×  | 0.97 |
| LC         | 0.02 | 0.02 | 0.06 | ×  |

(iv) large seeds and chaotic accretion (LC). Each model predicts a theoretical distribution of coalescing MBHBs. A given dataset \( D \) of observed events can be compared to a given model \( A \) by computing the likelihood \( p(D|A) \) that the observed dataset \( D \) is a realization of model \( A \). When testing a dataset \( D \) against a pair of models \( A \) and \( B \), one assigns probability \( p_A = p(D|A)/(p(D|A) + p(D|B)) \) to model \( A \), and probability \( p_B = 1 - p_A \) to model \( B \). The probabilities \( p_A \) and \( p_B \) are a measure of the relative confidence one has in model \( A \) and \( B \), given an observation \( D \). Setting a confidence threshold of 0.95 one can count what fraction of the 1000 realizations of model \( A \) yields a confidence \( p_A > 0.95 \) when compared to an alternative model \( B \). Results are shown in the top panel of table 1 for all pairs of models, assuming one year observation and circular non-spinning waveforms (i.e., for an extremely conservative waveform model). The majority of the pair comparisons yields a 95% confidence in the true model for almost all the realizations, with the exception of comparisons LE to LC and SE to SC, i.e., comparisons among models differing by accretion mode only. This is because the accretion mode (efficient versus chaotic) particularly affects the spin distribution of the coalescing systems, which is not considered in the circular non-spinning waveform model. It is sufficient to add a measurement of the remnant spin parameter \( a_r \) to make those pairs easily distinguishable (bottom panel of table 1). The right panel of figure 7 shows the evolution of the fraction of correctly identified models as a function of observation time (no spin information included). Small versus large seed scenarios (SE versus LE and SE versus LC) can be easily discriminated after only 1 year of observation.

5.2. The nano-Hertz regime: science with pulsar timing arrays

PTAs are sensitive at much lower frequencies \( (10^{-9} - 10^{-7} \) Hz), where the expected signal is given by a superposition of a large number of massive \( (M > 10^8 M_\odot) \), relatively nearby \( (z < 1) \) sources overlapping in frequency. As argued in section 3, at such low frequencies the properties of the MBHBs are likely to be severely affected by their coupling with their stellar and gaseous environment. In particular binaries can be highly eccentric, which might suppress the low-frequency portion of the spectrum, crucial to PTA detection [79]. Here we consider
Circular GW driven binaries for simplicity. The overall expected characteristic strain $h_c$ of the GW signal can be written as [80]

$$h_c^2(f) = \int_0^\infty dz \int_0^\infty dM \frac{d^3N}{dz dM d \ln f} h^2,$$

(34)

where $d^3N/dz dM d \ln f$, is the comoving number of binaries emitting in a given logarithmic frequency interval with chirp mass and redshift in the range $[M, M + dM]$ and $[z, z + dz]$, respectively; and $h$ is the inclination–polarization averaged strain given by equation (4).

The GW spectrum has a characteristic shape $h_c = A(f/1 \text{ yr}^{-1})^{-2/3}$, where $A$ is the signal normalization at $f = 1 \text{ yr}^{-1}$, which depends on the details of the MBH binary population only. A Monte Carlo realization of the signal is shown in figure 8 for a selected MBHB population model. In the timing residual of each individual pulsar, the signal appears as a structured red noise (left panel), but a representation of the characteristic strain in the Fourier domain reveals the complexity of its nature (right panel). Although several millions of sources contribute to it, the bulk of the strain comes from few hundred sources only. Therefore, the signal is far from being a Gaussian isotropic background [87]; a handful of sources dominates the strain budget, and some of them might be individually identified. Detection techniques have been developed for stochastic signals [88–91] and individual sources [92–94], and more sophisticated schemes accounting for signal anisotropy have recently been proposed [95–97]. In terms of level of the stochastic signal, recent works [81, 85, 87] set a plausible range $3 \times 10^{-16} < A < 4 \times 10^{-15}$, and the newly published PPTA limit of $A = 2.4 \times 10^{-15}$ [84], is already digging into it. This is shown in the left panel of figure 9, where observations are compared to theoretically predictions. Here, the difference between the top-left and the top-right panels is given by the recent upgrades in the MBH mass–host relation [98, 99] to include the overmassive BHs measured in brightest cluster galaxies (BCGs) [100], that boost-up the range of expected signals by a factor of 2. In the lower panels instead, we consider two subset of the models
Figure 9. Left panel: characteristic amplitude of the GW signal. Shaded areas represent the 68%, 95% and 99.7% confidence levels given by [81]. In each panel, the asterisks mark the limits given by NANOGrav[82] (magenta), the EPTA [83] (black), and the PPTA [84] (red). Shaded areas in the upper left panel refer to the 95% confidence level given by [85] (red) and the uncertainty range estimated by [80] (black). Right panel: median expected statistical error on the source parameters. Each point (asterisk or square) is obtained by averaging over a large Monte Carlo sample of MBHBs. In each panel, solid lines (squares) represent the median statistical error as a function of the total coherent SNR, assuming 100 randomly distributed pulsars in the sky; the thick dashed lines (asterisks) represent the median statistical error as a function of the number of pulsars for a fixed total SNR = 10. In this latter case, thin dashed lines label the 25th and the 75th percentile of the error distributions (from [86]).

An handful of sources might be bright enough to be individually resolved, and in this case some of their parameters can be determined according to the scheme described in section 4. A pioneering investigation was performed by [86] assuming circular, non-spinning monochromatic systems. In this case the waveform is function of seven parameters only: the amplitude $R$, sky location $\theta, \phi$, polarization $\psi$, inclination $\iota$, frequency $f$ and phase $\Phi_0$, defining the parameter vector $\lambda = [R, \theta, \phi, \psi, \iota, f, \Phi_0]$. Results about typical parameter estimation accuracy are shown in the right panel of figure 9 for a detection SNR = 10. The source amplitude is determined to a 20% accuracy, whereas $\phi, \psi, \iota$ are only determined within a fraction of a radian. Sky location within few tens to few deg$^2$ is possible (see also [92, 94]), and even sub deg$^2$ determination, under some specific conditions [103]. Even though this is a large chunk of the sky, these systems are extremely massive and at relatively low redshift ($z < 0.5$), making any putative electromagnetic signature of their presence (e.g., emission periodicity related to the binary orbital period, peculiar emission spectra, peculiar Kα line profiles, etc) detectable [104, 105].

8 For monochromatic signals the two masses and the luminosity distance degenerate into a single amplitude parameter.
6. Conclusions

We provided a general overview of massive black hole binaries (MBHBs) as gravitational wave (GW) sources. MBHs are today ubiquitous in massive galaxies, and power luminous quasars up to \( z > 7 \).Although they are believed to play a central role in the process of structure formation, their origin and early growth is largely unknown. According to our current understanding, MBHBs must form in large numbers along the cosmic history, providing the loudest sources of GWs in the Universe in a wide range of frequencies spanning from the sub-nano-Hertz up to the milli-Hertz. GWs carry precise information about the parameters of the emitting systems. We showed how those parameters are imprinted in the phase (and amplitude) modulation of the wave, and can therefore be efficiently extracted and determined to high accuracy with ongoing and future GW probes. From those we will learn about MBH formation and evolution through cosmic history, about the nature of the first black hole (BH) seeds, their subsequent accretion history, and, more generally, about the early hierarchical structure formation at high redshift. We will also learn about the complex interplay of physical processes, including stellar and gas dynamics and GW emission, that leads to the dynamical formation and evolution of MBHBs. Direct GW detection will open a new era in MBH and MBHB astrophysics.

Acknowledgments

I would like to thank S Babak, M Dotti, F Ohme and A Petiteau for providing useful material, and E Barausse for useful discussions about the PN formalism. This work is supported by the DFG grant SFB/TR 7 Gravitational Wave Astronomy and by DLR (Deutsches Zentrum fur Luft- und Raumfahrt).

References

[1] Magorrian J et al 1998 The demography of massive dark objects in galaxy centers Astron. J. 115 2285–305
[2] Mortlock D J et al 2011 A luminous quasar at a redshift of \( z = 7.085 \) Nature 474 616–9
[3] White S D M and Rees M J 1978 Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering Mon. Not. R. Astron. Soc. 183 341–58
[4] Kauffmann G and Haehnelt M 2000 A unified model for the evolution of galaxies and quasars Mon. Not. R. Astron. Soc. 311 576–88
[5] Volonteri M, Haardt F and Madau P 2003 The assembly and merging history of supermassive black holes in hierarchical models of galaxy formation Astrophys. J. 582 559–73
[6] Thorne K S 1995 Gravitational waves Particle and Nuclear Astrophysics and Cosmology in the Next Millennium ed E W Kolb and R D Peccei (Singapore: World Scientific) p 160
[7] Hughes S A 2002 Untangling the merger history of massive black holes with LISA Mon. Not. R. Astron. Soc. 331 805–16
[8] Amaro-Seoane P et al 2013 eLISA: astrophysics and cosmology in the millihertz regime GW Notes 6 4–110
[9] Amaro-Seoane P et al 2012 Low-frequency gravitational-wave science with eLISA/NGO Class. Quantum Grav. 29 124016
[10] Amaro-Seoane P et al 2013 The gravitational universe arXiv:1305.5720
[11] Ferdman R D et al 2010 The European pulsar timing array: current efforts and a LEAP toward the future Class. Quantum Grav. 27 084014
[12] Manchester R N et al 2012 The Parkes pulsar timing array project arXiv:1210.6130
[13] Jenet F A et al 2009 The North American nanohertz observatory for gravitational waves arXiv:0909.1058
[14] Hobbs G et al 2010 The international pulsar timing array project: using pulsars as a gravitational wave detector Class. Quantum Grav. 27 084013
[15] Hughes S A 2003 Listening to the Universe with gravitational-wave astronomy Ann. Phys. 303 142–78
[16] Thorne K S 1987 Gravitational radiation Three Hundred Years of Gravitation (Cambridge: Cambridge University Press) pp 330–458
[17] Peters P C and Mathews J 1963 Gravitational radiation from point masses in a Keplerian orbit Phys. Rev. 131 435–40
[18] Pierro V, Pinto I M, Spallicci A D, Laserra E and Recano F 2001 Fast and accurate computational tools for gravitational waveforms from binary stars with any orbital eccentricity Mon. Not. R. Astron. Soc. 325 358–72
[19] Tegmark M, Silk J, Rees M J, Blanchard A, Abel T and Palla F 1997 How small were the first cosmological objects? Astrophys. J. 474 1
[20] Bromm V, Coppi P S and Larson R B 1999 Forming the first stars in the Universe: the fragmentation of primordial gas Astrophys. J. Lett. 527 L5–L8
[21] Madau P and Rees M J 2001 Massive black holes as population III remnants Astrophys. J. Lett. 551 L27–30
[22] Volonteri M 2010 Formation of supermassive black holes Astron. Astrophys. Rev. 18 279–315
[23] Haiman Z, Kocsis B and Menou K 2009 The population of viscosity- and gravitational wave-driven supermassive black hole binaries among luminous active galactic nuclei Astrophys. J. 700 1952–69
[24] Ivanov P B, Papaloizou J C B and Polnarev A G 1999 The evolution of a supermassive binary caused by an accretion disc Mon. Not. R. Astron. Soc. 307 79–90
[25] Koushiappas S M, Bullock J S and Dekel A 2004 Massive black hole seeds from low angular momentum material Mon. Not. R. Astron. Soc. 354 292–304
[26] Begelman M C, Volonteri M and Rees M J 2006 Formation of supermassive black holes by direct collapse in pre-galactic haloes Mon. Not. R. Astron. Soc. 370 289–98
[27] Mayer L, Kazantzidis S, Escala A and Callegari S 2010 Direct formation of supermassive black holes via multi-scale gas inflows in galaxy mergers Nature 466 1082–4
[28] Sesana A, Haardt F and Madau P 2006 Interaction of massive black hole binaries with their stellar environment: implications for low-frequency gravitational-wave astrophysics Mon. Not. R. Astron. Soc. 377 1711–6
[29] Hughes S A and Blandford R D 2003 Black hole mass and spin coevolution by mergers Astrophys. J. Lett. 585 L101–4
[30] Shakura N I and Sunyaev R A 1973 Black holes in binary systems. Observational appearance Astron. Astrophys. 24 337–54
[31] Thorne K S 1974 Disk-accretion onto a black hole: II. Evolution of the hole Astrophys. J. 191 507–20
[32] King A R, Lubow S H, Ogilvie G I and Pringle J E 2005 Aligning spinning black holes and accretion discs Mon. Not. R. Astron. Soc. 363 49–56
[33] Kassin S A et al 2012 The epoch of disk settling: z~1 to now Astrophys. J. 758 106
[34] Fabricius M H, Saglia R P, Fisher D B, Drory N, Bender R and Hopp U 2012 Kinematic signatures of bulges correlate with bulge morphologies and Sérsic index Astrophys. J. 754 67
[35] Dotti M, Colpi M, Pallini S, Perego A and Volonteri M 2013 On the orientation and magnitude of the black hole spin in galactic nuclei Astrophys. J. 762 68
[36] Reynolds C S 2013 Measuring black hole spin using x-ray reflection spectroscopy arXiv:1302.3260
[37] Blandford R D and Znajek R L 1977 Electromagnetic extraction of energy from Kerr black holes Mon. Not. R. Astron. Soc. 179 433–56
[38] King A R, Lubow S H, Ogilvie G I and Pringle J E 2005 Aligning spinning black holes and accretion discs Mon. Not. R. Astron. Soc. 363 49–56
[39] Cassisi S A et al 2012 The epoch of disk settling: z~1 to now Astrophys. J. 758 106
[40] Fabricius M H, Saglia R P, Fisher D B, Drory N, Bender R and Hopp U 2012 Kinematic signatures of bulges correlate with bulge morphologies and Sérsic index Astrophys. J. 754 67
[41] Dotti M, Colpi M, Pallini S, Perego A and Volonteri M 2013 On the orientation and magnitude of the black hole spin in galactic nuclei Astrophys. J. 762 68
[42] Reynolds C S 2013 Measuring black hole spin using x-ray reflection spectroscopy arXiv:1302.3260
[43] Blandford R D and Znajek R L 1977 Electromagnetic extraction of energy from Kerr black holes Mon. Not. R. Astron. Soc. 179 433–56
[44] Milosavljević M and Merritt D 2003 Long-term evolution of massive black hole binaries Astrophys. J. 596 860–78
[45] Dotti M, Sesana A and Decarli R 2012 Massive black hole binaries: dynamical evolution and observational signatures Adv. Astron. 2012 940568
[46] Sesana A 2010 Self consistent model for the evolution of eccentric massive black hole binaries in stellar environments: implications for gravitational wave observations Astrophys. J. 719 651–64
[47] Quinlan G D 1996 The dynamical evolution of massive black hole binaries: I. Hardening in a fixed stellar background New Astron. 1 35–56
[48] Merritt D, Mikkola S and Szell A 2007 Long-term evolution of massive black hole binaries: III. Binary evolution in collisional nuclear clusters New Astron. Rev. 51 57–72
[49] Matsubayashi T, Makino J and Ebisuzaki T 2007 Orbital evolution of an IMBH in the galactic nucleus with a massive central black hole Astrophys. J. 656 879–96
[50] Dotti M, Sesana A, Berczik P and Zarnecki R 2011 Fast coalescence of massive black hole binaries from mergers of galactic nuclei: implications for low-frequency gravitational-wave astrophysics Astrophys. J. Lett. 732 L26
[51] Haiman Z, Kocsis B and Menou K 2009 The population of viscosity- and gravitational wave-driven supermassive black hole binaries among luminous active galactic nuclei Astrophys. J. 700 1952–69
Roedig C and Sesana A 2012 Origin and implications of high eccentricities in massive black hole binaries at
Callegari S, Kazantzidis S, Mayer L, Colpi M, Bellovary J M, Quinn T and Wadsley J 2011 Growing massive
Hlavacek-Larrondo J, Fabian A C, Edge A C and Hogan M T 2012 On the hunt for ultramassive black holes
Lee K J, Wex N, Kramer M, Stappers B W, Bassa C G, Janssen G H, Karuppusamy R and Smits R
Sesana A, Roedig C, Reynolds M T and Dotti M 2012 Multimessenger astronomy with pulsar timing and x-ray
observations of massive black hole binaries Mon. Not. R. Astron. Soc. 420 860–77
Tanaka T, Menou K and Haaim Z 2011 Electromagnetic counterparts of supermassive black hole binaries
resolved by pulsar timing arrays Mon. Not. R. Astron. Soc. 420 705–19

[78] Sesana A, Gair J, Berti E and Volonteri M 2011 Reconstructing the massive black hole cosmic history through
gravitational waves Phys. Rev. D 83 044036
[79] Sesana A 2013 Insights on the astrophysics of supermassive black hole binaries from pulsar timing observations
Mon. Not. R. Astron. Soc. 433 L1–L5
[80] Sesana A, Vecchio A and Colacino C N 2008 The stochastic gravitational-wave background from massive
black hole binary systems: implications for observations with pulsar timing arrays Mon. Not. R. Astron.
Soc. 390 192–209
[81] Sesana A 2013 Systematic investigation of the expected gravitational wave signal from supermassive black
hole binaries in the pulsar timing band Mon. Not. R. Astron. Soc. 433 L1–5
[82] Demorest P B et al 2013 Limits on the stochastic gravitational wave background from the North American
Nanohertz Observatory for gravitational waves Astrophys. J. 762 94
[83] van Haasteren R et al 2011 Placing limits on the stochastic gravitational-wave background using European
pulsar timing array data Mon. Not. R. Astron. Soc. 414 3117–28
[84] Shannon R M et al 2013 Gravitational-wave limits from pulsar timing constrain supermassive black hole
evolution Science 342 334–7
[85] McWilliams S T, Ostriker J P and Pretorius F 2012 Gravitational waves and stalled satellites from massive
galaxy mergers at z ⩽ 1 arXiv:1211.5377
[86] Sesana A and Vecchio A 2010 Measuring the parameters of massive black hole binary systems with pulsar
timing array observations of gravitational waves Phys. Rev. D 81 104008
[87] Ravi V, Wyithe J S B, Hobbs G, Shannon R M, Manchester R N, Yardley D R B and Keith M J 2012 Does a
‘stochastic’ background of gravitational waves exist in the pulsar timing band? Astrophys. J. 761 84
[88] Hellings R W and Downs G S 1983 Upper limits on the isotropic gravitational radiation background from
pulsar timing analysis Astrophys. J. Lett. 265 L39–42
[89] Jenet F A, Hobbs G B, Lee K J and Manchester R N 2005 Detecting the stochastic gravitational wave
background using pulsar timing Astrophys. J. Lett. 625 L123–6
[90] van Haasteren R, Levin Y, McDonald P and Lu T 2009 On measuring the gravitational-wave background using
pulsar timing arrays Mon. Not. R. Astron. Soc. 395 1005–14
[91] Anholm M, Ballmer S, Creighton J D E, Price L R and Siemens X 2009 Optimal strategies for gravitational
wave stochastic background searches in pulsar timing data Phys. Rev. D 79 084030
[92] Ellis J A, Siemens X and Creighton J D E 2012 Optimal strategies for continuous gravitational wave detection
in pulsar timing arrays Astrophys. J. 756 175
[93] Babak S and Sesana A 2012 Resolving multiple supermassive black hole binaries with pulsar timing arrays
Phys. Rev. D 85 044034
[94] Petiteau A, Babak S, Sesana A and de Araújo M 2013 Resolving multiple supermassive black hole binaries
with pulsar timing arrays: II. Genetic algorithm implementation Phys. Rev. D 87 064036
[95] Cornish N J and Sesana A 2013 Pulsar timing array analysis for black hole backgrounds Class. Quantum
Grav. 30 224005
[96] Mingarelli C M F, Sidery T, Mandel I and Vecchio A 2013 Characterising gravitational wave stochastic
background anisotropy with pulsar timing arrays Phys. Rev. D 88 064005
[97] Taylor S R and Gair J R 2013 Searching for anisotropic gravitational-wave backgrounds using pulsar timing
arrays Phys. Rev. D 88 084001
[98] McConnell N J and Ma C-P 2013 Revisiting the scaling relations of black hole masses and host galaxy
properties Astrophys. J. 764 184
[99] Graham A W and Scott N 2013 The MBH-Lspheroid relation at high and low masses, the quadratic growth of
black holes, and intermediate-mass black hole candidates Astrophys. J. 764 151
[100] Hlavacek-Larrondo J, Fabian A C, Edge A C and Hogan M T 2012 On the hunt for ultramassive black holes
in brightest cluster galaxies Mon. Not. R. Astron. Soc. 424 224–31
[101] Callegari S, Kazantzidis S, Mayer L, Colpi M, Bellovary J M, Quinn T and Wadsley J 2011 Growing massive
black hole pairs in minor mergers of disk galaxies Astrophys. J. 729 85
[102] Roedig C and Sesana A 2012 Origin and implications of high eccentricities in massive black hole binaries at
sub-pc scales J. Phys.: Conf. Ser. 363 012035
[103] Lee K J, Wex N, Kramer M, Stappers B W, Bassa C G, Janssen G H, Karuppusamy R and Smits R
2011 Gravitational wave astronomy of single sources with a pulsar timing array Mon. Not. R. Astron.
Soc. 414 3251–64
[104] Sesana A, Roedig C, Reynolds M T and Dotti M 2012 Multimessenger astronomy with pulsar timing and x-ray
observations of massive black hole binaries Mon. Not. R. Astron. Soc. 420 860–77
[105] Tanaka T, Menou K and Haaim Z 2011 Electromagnetic counterparts of supermassive black hole binaries
resolved by pulsar timing arrays Mon. Not. R. Astron. Soc. 420 705–19