Multimessenger astronomy with pulsar timing and X-ray observations of massive black hole binaries

A. Sesana,1⋆ C. Roedig,1 M. T. Reynolds2 and M. Dotti3,4

1Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institut, Am Mühlenberg 1, 14476 Golm, Germany
2Department of Astronomy, University of Michigan, 500 Church Street, Ann Arbor, MI 48109, USA
3Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching b. München, Germany
4Dipartimento di Fisica G. Occhialini, Università di Milano Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

Accepted 2011 October 28. Received 2011 October 28; in original form 2011 July 14

ABSTRACT
In the decade of the dawn of gravitational wave astronomy, the concept of multimessenger astronomy, combining gravitational wave signals to conventional electromagnetic observation, has attracted the attention of the astrophysical community. So far, most of the effort has been focused on ground- and space-based laser interferometer sources, with little attention devoted to the ongoing and upcoming pulsar timing arrays (PTAs). We argue in this paper that PTA sources, being very massive (>10^8 M⊙) cosmologically nearby (z < 1) black hole binaries (MBHBs), are particularly appealing multimessenger carriers. According to current models for massive black hole formation and evolution, the planned Square Kilometre Array will observe thousands of such massive systems, being able to individually resolve and locate in the sky several of them (maybe up to a hundred). MBHBs form in galaxy mergers, which are usually accompanied by strong inflows of gas in the centre of the merger remnant. By employing a standard model for the evolution of MBHBs in circumbinary discs, with the aid of dedicated numerical simulations, we characterize the gas–binary interplay, identifying possible electromagnetic signatures of the PTA sources. We concentrate our investigation on two particularly promising scenarios in the high-energy domain, namely the detection of X-ray periodic variability and double broad Kα iron lines. Up to several hundreds of periodic X-ray sources with a flux >10^{-13} erg s^{-1} cm^{-2} will be in the reach of upcoming X-ray observatories; in the most optimistic case, a few of them may be already being observed by the MAXI detector placed on the International Space Station. Double relativistic Kα lines may be observable in a handful of low-redshift (z < 0.3) sources by proposed deep X-ray probes, such as Athena. The exact figures depend on the details of the adopted MBHB population and on the properties of the circumbinary discs, but the existence of a sizeable population of sources suitable to multimessenger astronomy is a robust prediction of our investigation.

Key words: accretion, accretion discs – black hole physics – gravitational waves – pulsars: general – X-rays: general.

1 INTRODUCTION
Within this decade the detection of gravitational waves (GWs) may be a reality, opening a completely new window on the Universe. While signals coming from compact stars and binaries fall in the observational domain of operating and planned ground-based interferometers [such as the Laser Interferometer Gravitational Wave Observatory (LIGO), Virgo and the proposed Einstein Telescope], massive black hole (MBH) binaries (MBHBs) are expected to be among the primary actors of the upcoming low-frequency stage, where the 10^{-4} to 10^{-1} Hz window is going to be probed by GW interferometry in space [see e.g. the proposed Laser Interferometer Space Antenna (LISA)\(^1\)]. Moving to even lower frequencies, long-term monitoring of an array of millisecond pulsars (forming a so-called pulsar timing array, PTA) may unveil the characteristic

\(^{1}\) Following the termination of the ESA–NASA partnership at the base of the LISA project, a new design study is being submitted to ESA, for a smaller mission called ELISA/New Gravitational wave Observatory (NGO). More information can be found at https://lisa-light.aei.mpg.de/bin/view.
Electromagnetic counterparts to GW events have attracted a lot of attention in the last few years in the context of LISA observations of coalescing MBHBs (see Schnittman 2011, and references therein). The final coalescence being a violent event, many possible counterparts related to shocks, flares and transients in the surrounding ambient gas have been proposed. However, the bulk of the LISA sources is likely to cluster at $3 \pm 1 \, \text{deg}$; Lang & Hughes (2008), the number of candidate hosts in the putative GW error box is likely to be very high (order of hundred of thousands); Petitie, Babak & Sesana (2011). Moreover, our incomplete understanding of the observational signatures of such events, together with their intrinsic weakness (most LISA sources are expected to be MBHBs with masses $M \gtrsim 10^{4} \, M_{\odot}$; Sesana et al. 2007), will make host identification extremely challenging in the vast majority of the cases.

On the other hand, very little attention has been devoted to possible counterparts of PTA sources. An obvious advantage with respect to LISA sources is that those are expected to be very massive ($M > 10^{6} \, M_{\odot}$), cosmologically nearby ($z < 1$) systems (Sesana, Vecchio & Volonteri 2009; Lee et al. 2011). Any characteristic electromagnetic signature would be therefore within the sensitivity range of current and future observational capabilities. PTA sources are inspiralling MBHBs still far from coalescence (emitting in the $10^{-9}$ to $10^{-6}$ Hz frequency range), that can be considered, for any practical purpose, stationary during typical observation time-scales of decades (Sesana & Vecchio 2010). Fairly high eccentricities (covering a substantial range from 0.1 up to 0.8) should be the norm, as shown by MBHB hardening studies in both stellar (Sesana 2010; Khan, Just & Merritt 2011; Pretto et al. 2011; Sesana, Gualandris & Dotti 2011) and gaseous (Armitage & Natarajan 2005; Cuadra et al. 2009; Roedig et al. 2011) environments. If the binary is surrounded by a circumbinary gaseous disc, eccentricity triggers periodic inflows (Artymowicz & Lubow 1996; Hayasaka, Mineshige & Sudou 2007; Roedig et al. 2011), opening a wide range of possibilities for distinctive binary fingerprints that can be observationally identified.

In this paper, we study the prospects of conducting multimessenger astronomy with PTA observations of MBHBs. We quantify the number of sources possibly detectable both in the GW and in the electromagnetic realm. We investigate plausible electromagnetic signatures and we forecast identification prospects by means of future X-ray observatories. Starting from the Millennium Run (Springel et al. 2005) data base, we construct detailed MBHB populations satisfying all the currently available constraints in terms of local MBH mass function (Marconi et al. 2004), MBH–host relations (Gültekin et al. 2009) and MBHB merger rates as inferred from close galaxy pair counts (e.g. Bell et al. 2006). We assume that after a galaxy merger, cold gas is funnelled in the centre of the remnant, forming a circumnuclear disc (Mihos & Hernquist 1996; Mayer et al. 2007). The newly formed sub-sparse MBHB excavates a hollow region in the centre of the disc (which we will referred to as ‘cavity’), and the dynamics of the system is governed by the mutual MBHB–disc torques. We compute the detailed population of GW-emitting sources at each frequency, by adopting a simple analytical model for their evolution under the dynamical effect of gaseous drag and GW emission (see Kocsis & Sesana 2011, for details). After identifying the prototypical MBHB observable with PTAs, we simulate its dynamics numerically using a modified version of the smooth particle hydrodynamics (SPH) code GADGET-2 (Springel 2005; Cuadra et al. 2009), to track the fate of the material leaking through the cavity and captured by the MBHs. The outcomes of the simulation are then used to model analytically the accretion on to the two MBHs and the emitted radiation.

Several distinctive signatures of inspiralling MBHBs have been presented in the literature. Proposed scenarios range from the radio up to the X-rays, including variability of the radio (O’Shaughnessy et al. 2011) and optical (e.g. Haiman et al. 2009a) continua, spectral shifts in the broad-line emission (Begelman, Blandford & Rees 1980; Eracleous et al. 2011; Tsalmantza et al. 2011), peculiar flux ratios between optical/UV broad emission lines (Montuori et al. 2011), precessing or X-shaped jets (Liu 2004; Lobanov & Roland 2005), periodic outbursts (Sillanpaa et al. 1988) and astrometric measurement of the source motion (Sudou et al. 2003). Yet, the only unambiguous observed candidate is a double compact radio core at a 7 pc projected separation (Rodriguez et al. 2006). We identify here a number of possible characteristic signatures and we concentrate our investigation on two particularly promising scenarios in the high-energy domain, namely the detection of X-ray periodic variability and of double broad Kα iron lines. For both scenarios, we quantify the population of potentially observable sources, and the joint detection prospects with future X-ray observatories and PTAs.

The paper is organized as follows. In Sections 2 and 3, we model and describe in detail the MBHB population relevant to PTA observations, whereas in Section 4 we investigate the dynamics of the typical source by combining high-resolution SPH simulations to an analytical model for the accreted matter. In Section 5, we model the emitted spectrum, identifying possible characteristic signatures, and Sections 6 and 7 are devoted to periodic variability and double broad Kα lines, respectively. We discuss strategies for joint PTA and X-ray detection and draw our conclusions in Section 8.

---

2 This is particularly true if we consider the new ELISA/NGO. To fit within the available budget, the proposed instrument features four-laser links only (i.e. a single Michelson interferometer), with an inevitable degradation of the distance and sky location measurements of the source.

3 http://www.mpa-garching.mpg.de/galform/virgo/millennium/
2 THE PULSAR TIMING MASSIVE BLACK HOLE BINARY POPULATION

Our main goal is to quantify and characterize the population of MBHBs that is accessible to observation both in GWs – through pulsar timing – and in the electromagnetic realm via distinctive signatures. Formally, what we need to estimate is the population of MBHBs observable today in the Universe as a function of MBH masses, redshift and orbital frequency (or, alternatively, semimajor axis): $\frac{d^4N}{dM_1dM_2d\omega d\beta}$. Here $M_1 > M_2$ are the masses of the two binary components, $\omega$ is the source redshift and $f_{\nu,k}$ is the rest-frame Keplerian frequency of the system. This can be formally written as the product of (i) $\frac{d^3N}{dM_1dM_2d\omega f_{\nu,k}}$ and (ii) $df/\ln f_{\nu,k}$. Item (i) is nothing but the cosmological coalescence rate of MBHBs, and it depends on the general clustering history of structures in the Universe; item (ii) describes the time evolution of the binary frequency and is determined by the detailed physical processes driving the binary dynamics prior to coalescence. For a given coalescence rate, $df/\ln f_{\nu,k}$ quantifies the number of binaries as a function of orbital frequency that has to be present at any time in the sky to sustain that particular rate. This item can be estimated taking into account that we are interested in subparsec MBHB, for which we can employ simple evolutionary models based on disc–binary mutual torques and GW emission, as we will see in Section 2.2. Once the MBHB population has been determined, the requirement of observability through PTA selects a subset of systems which we will refer to as the PTA-MBHB population. In the following, we will treat items (i) and (ii) separately (Sections 2.1 and 2.2). We will then describe the characteristic timing residuals induced in a PTA, defining the PTA-MBHB population (Section 2.3).

2.1 Cosmological coalescence rate

To construct the MBHB cosmological coalescence rate, we follow the procedure described in Section 2 of Sesana et al. (2009). We refer the reader to that paper for full details. In few words, we extract catalogues of merging galaxies from the semi-analytical model of Bertone, De Lucia & Thomas (2007) applied to the Millennium Run (Springel et al. 2005). We then associate a central MBH to each merging galaxy in our catalogue. We adopt the recent fit to the $M-\sigma$ relation given by Gültekin et al. (2009), and we consider accretion to be efficient on to both MBHs before the final coalescence of the binary.

Assigning an MBH to each galaxy, we obtain a catalogue of all the mergers occurring in the $(500\ h^{-1} \text{Mpc})^3$ comoving volume of the simulation, labelled by MBH masses and redshift. From this, we numerically generate the cosmic merger rate, $\frac{d^4N}{dM_1dM_2d\omega df_{\nu,k}}$, by weighting each event with the observable comoving volume shell at every redshift.

2.2 MBHBs in circumbinary discs

We assume that MBHBs evolve in geometrically thin circumbinary accretion discs (Haiman, Kocsis & Menou 2009b). Even though such an assumption is enforced by the requirement of a distinctive electromagnetic signature, we find that most of the merging galaxy pairs extracted from the Millennium Run involve at least one massive, gas-rich, spiral galaxy. Significant inflows of cold gas are therefore expected in the remnant nucleus (Mihos & Hernquist 1996), feeding the circumbinary disc.

Before entering the dynamics in detail, we define all the quantities used for describing the MBHB–disc system. The MBHB has total mass $M = M_1 + M_2$ ($M_1 > M_2$), mass ratio $q = M_1/M_2$, symmetric binary mass ratio $q_s = 4q/(1 + q^2)$, semimajor axis $a$, rest-frame orbital Keplerian frequency $f_{\nu,k}$ and eccentricity $e$. The thin disc is described in terms of the $\alpha$ viscosity parameter (Shakura & Sunyaev 1973), the accretion rate $\dot{M}$ and the accreted matter/radiation conversion efficiency $\epsilon$. Following Roedig et al. (2011), we define $\delta$ to be the relative size of the cavity to the binary semimajor axis, and we take a fiducial value $\delta = 2$. Throughout the paper, expressions are given in units of $M_0 = M/10^8 M_\odot$, $a_0 = a/0.3$, $\epsilon_{0.1} = \epsilon/0.1$ and $m_{0.3} = m/0.3$, where $m = M/M_{\text{Edd}}$ is the accretion rate normalized to the Eddington rate [$M_{\text{Edd}} = L_{\text{Edd}}/(\epsilon c^2)$, where $L_{\text{Edd}}$ is the Eddington luminosity]. In the equations below, the subscript ‘3’ refers to lengths given in units of $10^3 R_\odot$, where $R_\odot = 2GM/c^2$ is the Schwarzschild radius associated with the total mass of the binary.

The picture we adopt for the MBHB evolution is the following. At separations relevant to our work ($a < 0.03 \text{pc}$), the binary has excavated a cavity (see e.g. Artyomowicz & Lubow 1994) in the circumbinary accretion disc. Both viscous torques exerted by the disc (Goldreich & Tremaine 1980) and GW emission (Peters & Mathews 1963) dissipate the binary binding energy, causing its orbital decay. The secondary MBH is, in general, more massive than the local disc and the problem is analogous to secondary-dominated Type II migration in planetary dynamics. A self-consistent solution to this problem was provided by Syer & Clarke (1995); we refer the interested reader to their original paper and to Haiman et al. (2009b) and Kocsis & Sesana (2011) for a detailed discussion in the context of MBHB evolution. This solution applies only to discs with decreasing surface density as a function of the radius. Standard $\alpha$-discs (Shakura & Sunyaev 1973), with viscosity proportional to the total pressure, meet such a requirement only in the gas pressure dominated zone, but not in the inner radiation pressure dominated zone, $\beta$-discs, with viscosity proportional to the gas pressure only, fulfill this condition at all radii. We therefore assume $\beta$-discs as our fiducial disc model for describing the binary-disc migration. All the relevant features of $\beta$-discs are reviewed by Haiman et al. (2009b); here we merely introduce the quantities relevant to the practical computation of the PTA-MBHB population. The relevant time-scales in the system are as follows:

(i) the viscous time-scale

$$ t_v = 6.09 \times 10^9 \text{yr} a_{0.3}^{-4/3} \left( \frac{\dot{m}_{0.3}}{\epsilon_{0.1}} \right)^{-2/5} M_8^{6/5} \delta^{7/5} a_3^{7/5}, \quad (1) $$

(ii) the migration time-scale

$$ t_m = 2.09 \times 10^6 \text{yr} a_{0.3}^{-1/3} \left( \frac{\dot{m}_{0.3}}{\epsilon_{0.1}} \right)^{-5/8} M_8^{1/4} q_{s}^{5/8} \delta^{5/8} a_3^{7/8}, \quad (2) $$

(iii) the GW shrinking time-scale

$$ t_{gw} = 7.84 \times 10^3 \text{yr} M_8 \dot{M}_{\text{Edd}}^{-1} a_3^{1/2} F(\epsilon)^{-1}, \quad (3) $$

where

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

These time-scales define two characteristic binary separations. Equating $t_m$ to $t_{gw}$, we get the decoupling separation

$$ a_3^{1/3} = 0.31 a_{0.3}^{-4/25} \left( \frac{\dot{m}_{0.3}}{\epsilon_{0.1}} \right)^{-1/5} M_8^{2/5} q_{s}^{11/25} \delta^{7/25} F(\epsilon)^{8/25}, \quad (5) $$

which is the separation below which the GW-driven migration and the binary evolution is dictated by GW emission only. This is the relevant separation that has to be considered when computing the frequency distribution of
inspiralling MBHBs. Equating $t_r$ to $t_{gw}$, we get the ‘disc freezing’ (or detachment) separation

$$a^{de}_{r} = 0.22 a_{0,3}^{4/13} \left( \frac{m_{c,3}}{e_{0,1}} \right)^{-2/13} M_8^{1/13} q_{4}^{4/13} \delta_{7/13} f (e)^{5/13},$$

below which the binary shrinking time-scale is faster than the viscous inward diffusion of the inner edge of the disc. After this point, the disc can no longer follow the binary, and it is basically ‘frozen’ during the quick subsequent inspiral leading to the MBHB coalescence. This latter characteristic radius discriminates between binaries attached to their discs and binaries detached from their discs, and it has to be considered in computing the population of ‘observable systems’. In general, $a^{r} \ll a^{dec}$, defining the following three different phases relevant to our study.

(i) Phase I, $a > a^{dec}$. The binary is coupled to the disc which regulates its dynamical evolution.

(ii) Phase II, $a^{r} < a < a^{dec}$. The binary is dynamically decoupled from the disc and it is driven by GW emission. However, the viscous time is short enough that the disc can follow the binary in its inspiral.

(iii) Phase III, $a < a^{r}$. The binary leaves the disc behind and quickly coalesces.

During phase I, the binary maintains a constant limiting eccentricity given by (Roedig et al. 2011)

$$e_{0} \approx 0.66 \sqrt{\ln \delta} - 0.65 + 0.19.$$  

(7)

After decoupling, in phase II and phase III, the eccentricity can numerically be solved by computing the implicit equation

$$\frac{f_{k}}{f_{dec}} = \left( \frac{1 - e_{dec}^{2}}{1 - e^{2}} \right)^{12/19} \left[ \frac{1 + \frac{121}{80} e^{2}}{1 + \frac{121}{80} e_{dec}^{2}} \right]^{870/2299}$$

obtained in the quadrupole approximation, by coupling the GW orbital decay rate to the eccentricity decay rate (Peters & Mathews 1963).

To compute the frequency distribution of inspiralling MBHBs, following Kocsis & Sesana (2011), we express the frequency evolution of the binary in terms of ‘residence time’:

$$\frac{dt_r}{d \ln f_{k}} = \frac{dt_{r}}{d \ln a} \frac{d \ln a}{d \ln f_{k}} = \frac{2}{3} t_{res},$$

(9)

In phase I, when the binary is driven by the disc, $t_{res} = t_m$, while in phase II and phase III GW emission takes over and $t_{res} = t_{gw}$.

Putting the pieces together, for any given MBHB–disc parameters, we compute $a^{dec}$, assuming $e_0$ given by equation (7); $dt_r/d \ln f_{k}$ is then obtained from equation (9) by plugging in equations (2) and (3), where $e(f_{k}, k)$ after decoupling is numerically obtained at any frequency from equation (8).

2.3 Pulsar timing observability

We are interested here in MBHBs potentially observable with ongoing and forthcoming PTAs. At the leading order, an eccentric MBHB radiates GWs in the whole spectrum of harmonics $f_{k,n} = n f_{k}$ ($n = 1, 2, \ldots$). The sky-polarization-averaged amplitude observed at a frequency $f_{k} = f_{k,n}(1 + z)$ at comoving distance $d$ is given by (Finn & Thorne 2000)

$$h_{n}(f_{k}) = 2 \sqrt{ \frac{32}{5} } \frac{M^{3/5}}{n d} (2\pi f_{k})^{2/5} \sqrt{B(n, e)},$$

(10)

where $M = M_1^{3/5} M_2^{2/5} / (M_1 + M_2)^{1/5}$ is the chirp mass of the system and $g(n, e)$ is a combination of Bessel functions that quantifies the relative power radiated in each single harmonic (see Amaro-Seoane et al. 2010, for a detailed description). If we observe for a time $T_{obs}$, the relevant detectable amplitude of each harmonic is approximately $h_{obs,n} = h_{n} \sqrt{T_{obs}}$, 

$$1/(2 \pi) = t = 4$$

is simply the number of cycles completed by the $n$th harmonic in the observation time. The timing residuals are defined as integrals of the GW during the observation time. As detailed in Sesana et al. (2009), each harmonic induces an average timing residual of the order

$$\delta_{t_{gw}}(f_{k}) = \sqrt{ \frac{8}{15} } \frac{h_{obs,n}}{2 \pi f_{k}},$$

(12)

where $\sqrt{8/15}$ is the average ‘antenna beam pattern’, i.e. the average of the signal performed over all possible source-pulsar relative orientations. The total residual can then be assumed to be of the order

$$\delta_{t_{gw}} = \left( \sum_{n=0}^{\infty} \delta_{t_{gw}}^{2}(f_{k}) \right)^{1/2},$$

(13)

In PTA observations, radio pulsars are monitored weekly for several years (Hobbs 2011). Assuming a time interval $\Delta t$ between adjacent observations during a campaign lasting $T_{obs}$, the maximum and minimum resolvable frequencies are $f_{max} = 1/(2 \pi \Delta t) \approx 10^{-6}$ Hz, corresponding to the Nyquist frequency, and $f_{min} = 1/T_{obs} \approx 3 \times 10^{-9}$ Hz, for a 10 yr observation time. We are therefore interested only in MBHBs producing a significant residual in this frequency window.

Coupling the cosmological coalescence rate to the detailed binary evolution (see Sections 2.1 and 2.2), we numerically obtain a distribution $d^3 N / (d M_1 d M_2 d \ln f_{k})$ so that $d^3 N / (d M_1 d M_2 d \ln f_{k})$, over the appropriate mass, redshift and frequency ranges, gives the average number of sources observable in the Universe taking an ideal snapshot of the whole sky at any time. To quantify the population of relevant PTA-MBHBs, we generate Monte Carlo samples of MBHBs according to such distribution assuming $f_{k,n} > 10^{-10}$ Hz, $M_1$ and $M_2 > 10^{8}$ M$_{\odot}$ and $z < 3$. For each MBHB, we then compute the average timing residual in the observed frequency range $[f_{min}, f_{max}]$ according to equation (13), and we keep only those binaries producing a residual larger than 0.1 ns.

3 DESCRIPTION OF THE PTA-MBHB POPULATION

In our default model, we assume that MBHBs evolve according to the scheme described in Section 2.2. We consider $\beta$-discs, with viscosity proportional to gas pressure only and $\alpha = 0.3$ (see King, Pringle & Livio 2007, for a detailed discussion), $\eta = 0.3$ (see e.g. Kollmeier et al. 2006; Labita et al. 2009) and $e = 0.1$ (appropriate for not-to-mildly rotating black holes). All the relevant features of the resulting PTA-MBHB population, averaged over 100 Monte Carlo realizations, are shown in Fig. 1.

Typical PTA sources are very massive ($M > 10^{8}$), cosmologically nearby ($z < 2$) binaries (see also Lee et al. 2011, for an independent study leading to the same conclusion). Mass ratios are in the range 0.1–1, with a long tail extending up to $10^{-3}$. Given the significant eccentricities, orbital periods contributing to the PTA signals extend up to $> 100$ yr. For the range of masses involved, this corresponds to a broad distribution in orbital semimajor axis in the range $30–10^4$ R$_S$ and circular velocities peaked around $10^4$ km s$^{-1}$ (if the systems were in circular orbits). Typical coalescence times $t_c$, defined as
the integral of equation (9) from the rest-frame binary frequency to coalescence are between $10^4$ and $10^6$ yr (bottom-left panel); PTA-MBHBs are therefore caught in the final few hundred thousand years of their life.

The bottom-central panel shows the distribution of induced residuals. Order of $\gtrsim 1000$ sources contribute to a level of 1 ns or more, which can realistically be considered the ultimate goal of future PTA efforts. Their superposition will result in a confusion-noise-type foreground, similar to the white dwarf binary signal expected for LISA (see e.g. Nelemans, Yungelson & Portegies Zwart 2001). As in this latter case, few (maybe most) of the sources contributing to the signal may be individually resolvable. Preliminary, purely frequency-based, crude estimations (Sesana et al. 2009) forecast resolvability of at least about five to 10 sources, but many additional sources will likely be resolved using the spatial information enclosed in the array (up to $2N/7$ per frequency bin, for an array of $N$ pulsars according to Boyle & Pen 2010, the prescription we will use below). The bottom-right panel shows a rough estimation...
of the X-ray flux distribution on Earth, assuming that the accretion rate of the binary (fuelled by streams flowing through the cavity; see next section) is the same as the one assumed for modelling the outer circumbinary disc (i.e. $m = 0.3$ in this case) and a bolometric correction of 30 (i.e. $\approx 3$ per cent of the bolometric luminosity emitted in X-ray; Lusso 2010). Given their high masses and low redshifts, PTA-MBHBs are quite bright, with fluxes generally above $10^{13}$ erg s$^{-1}$ cm$^{-2}$.

In all panels, blue solid histograms refer to sources at $a > a^d$. In such systems, the gas can follow the binary during its inspiral, and strongly periodic inflows of gas feeding the two black holes are a robust prediction of SPH simulations (Hayasaki et al. 2007; Roedig et al. 2011; see next section). Thus, we can safely count them in the population of PTA-MBHBs suitable for multimessenger astronomy observations. The dashed-red histograms, instead, track binaries at $a < a^d$. For such systems, torques exerted by the binary quadrupole moment on the inner edge of the disc become smaller and smaller as the binary shrinks, losing effectiveness in driving the periodic streaming activity. None the less, Tanaka, Menou & Haiman (2011) showed that also in this detached phase (which we labelled phase III) gas can efficiently leak through the cavity, feeding the central binary; moreover, residual activity related to the consumption of the fossil gas left around the two MBHs can still be present (see e.g. Chang et al. 2010). Therefore, detached systems should also be interesting targets for electromagnetic counterpart identification; however, we do not consider them in the following discussion, and we refer the reader to Tanaka et al. (2011) for a comprehensive analysis of the subject.

### 3.1 Parameter dependence and caveats

The population shown in Fig. 1 relies on a selected MBH–host relation (Gültekin et al. 2009) and on a specific disc model (the $\beta$-disc model) with fixed parameters. Even though the qualitative features of the predicted population are robust against our particular choices, the effective number of sources and the relative population of coupled/decoupled systems are not. Lowering $\alpha$ and $m$ both result in longer migration time-scales in the disc. This, on one hand, increases (up to a factor of a few) the number of PTA-MBHBs. However, in this case, GW emission takes over at slightly larger separations (see the weak dependence of $a$ on $\alpha$ and $m$ in equation 6), and more binaries (especially with periods $<3$ yr) will be already detached from their discs and probably less suitable for multimessenger observations. In an $\alpha$-disc, the viscosity is much larger in the radiation-dominated zone, and, even though there is no self-consistent solution to the binary–disc evolution, it is likely that the decoupling would happen at smaller separations. In any case, $t_\nu$ is surely much smaller in the radiation-dominated zone for an $\alpha$-disc, and the disc ‘freezing’ radius would certainly be smaller, allowing binaries to be in touch with their circumbinary discs for longer time and at smaller separation. In this respect, the $\beta$-disc assumption used here is likely to be conservative.

The impact of the chosen MBH–host relation is also mild, affecting the PTA-MBHB population by a factor of 2–3 at most. Also the binary eccentricity is not very important in terms of the population itself, because in our simple model we assume the disc migration to be independent of it. A population of circular binaries will only decouple at mildly smaller values of $\alpha$, slightly increasing the number of coupled systems. However, in this case, variability related to periodic streams flowing through the cavity would be highly reduced (Roedig et al. 2011), probably making difficult to recognize these systems through periodicity studies.

As a note of caution, we remind the reader that our models assume that all merging MBHBs are active in their last evolutionary stage. Under such a condition (and the assumption that systems are accreting at $m = 0.3$), the source flux distribution shown in Fig. 1 accounts for $\approx 30$ per cent of the soft X-ray active galactic nucleus (AGN) logN–logS as observed by ROSAT (Voges 1999), meaning that one bright X-ray AGN out of three is indeed a massive binary. This, although not strictly ruled out by observations, is a strong statement that can be relaxed if $m < 0.3$, if a certain fraction of the sources is obscured, or simply if only a fraction $F$ of MBHBs is active. In this latter case, it is sufficient to rescale our predictions by a factor $F$, accounting for the fraction of active binaries. Similarly, we assumed that all MBHBs overcome the ‘last parsec problem’ (Milosavljević & Merritt 2001), which, in the light of several recent results about both gas- and stellar-driven MBHB hardening (Escala et al. 2005; Berczik et al. 2006; Dotti et al. 2007; Colpi et al. 2009; Cuadra et al. 2009; Hayasaki 2009; Khan et al. 2011; Preto et al. 2011), we consider reasonable. We will see in the following that even for $F \lesssim 0.1$, the number of predicted sources is sizeable, making multimessenger astronomy with PTA sources an appealing prospect, worthy of deeper investigations.

### 4 DYNAMICAL MODELLING OF THE PROTOTYPICAL PTA-MBHB

Having studied the MBHB population relevant to our investigation, we model in this section the dynamics of the prototypical PTA-MBHB surrounded by a circumbinary thin accretion disc. We use a syncretic approach, combining analytical accretion disc models to a high-resolution SPH simulation. The former provide the global set-up of the system, while the latter is necessary to gather a better insight of the dynamics of the torqued material streaming through the inner edge of the circumbinary disc.

#### 4.1 Standard model for a PTA binary

We model an MBHB of a total mass $M = M_1 + M_2$ surrounded by a thin Keplerian gaseous disc of mass $M_d$. An episodic inflow of low angular momentum gas could result in an accretion disc either corotating or counter-rotating with respect to the binary (Nixon et al. 2011a). However, in the late stages of a relatively ‘wet’ galaxy merger, the binary is surrounded by a corotating coplanar disc, as shown by several SPH simulations (Mayer et al. 2007; Dotti et al. 2009a). We therefore consider only the case of a corotating binary–circumbinary disc system.4 We set primary mass $M_1 = 2.6 \times 10^8 M_\odot$, mass ratio $q = M_2/M_1 = 0.35$ and semimajor axis $a_0 = 0.012$ pc, consistent with the typical range of source parameters found in the previous section (see Fig. 1). Following Roedig et al. (2011), we set the initial binary eccentricity to a value $e_0 = 0.6$. The binary has angular velocity $\Omega_0 = (GM/a_0^3)^{1/2}$ resulting in a period $P = f_0^{-1} = 2\pi/\Omega_0 \approx 6$ yr ($f_0$ is the orbital frequency). Although Newtonian simulations are in principle scale-free, the expected properties of the circumbinary disc depend on the assumed mass and semimajor axis of the binary. We therefore need to work out an a priori scaling of the SPH simulation. A truncated circumbinary $\beta$-disc has mass

---

4 We refer the reader to Nixon, King & Pringle (2011b) for a study of the dynamical evolution of MBHBs in counter-rotating discs.
\[ M_\text{d} = 8.63 \times 10^5 M_\odot \left( \frac{a_{0.3}}{0.1} \right)^{4/5} \left( \frac{R_{\text{in}}}{a_{0.1}} \right)^{7/10} \left( M_\bullet^{11/5} \left( R_{\text{out}}^{5/4} - R_{\text{in}}^{5/4} \right) \right), \]

where the two limiting radii of the disc, \( R_{\text{in}} \) and \( R_{\text{out}} \), are expressed in units of \( 10^3 R_\odot \) and, in our simulation, correspond to \( R_{\text{in}} = 2a_0 \) and \( R_{\text{out}} = 10a_0 \), respectively. Plugging in our default disc model \((a_{0.3} = m_{0.3} = \epsilon_{0.1} = 1)\) and the binary parameters assumed above, we obtain \( R_{\text{in}} \approx 0.6, R_{\text{out}} \approx 3 \) and \( M_\text{d} \approx 5 \times 10^6 M_\odot \approx 1.5 \times 10^{-3}M_\odot \). Comparing the viscous time-scale of the binary to the GW-driven orbital decay (equations 1 and 3, in Section 2.2), we find that the MBHB is still attached to the circumbinary disc, and we can study the dynamics in the Newtonian approximation, neglecting relativistic effects.

4.2 Numerical realization

We use the SPH code GADGET-2 (Springel 2005) as done in Roedig et al. (2011); we refer the reader to that paper for details, and we only outline the essentials here. As initial condition, we use the binary outlined in the previous section surrounded by an 8 million particle circumbinary disc, and let it relax over nine orbits. The self-gravity of the particles is accounted for in the standard tree-code fashion (Springel 2005), whereas the dynamics of the MBHs, which are modelled as Newtonian point masses, is followed with a fixed time-step, equal to 0.01 \( \Omega_o^{-1} \), i.e. they are not part of the tree. The gas is allowed to cool on a time-scale which is proportional to the local dynamical time of the disc; thus, we define a cooling parameter \( \beta = t_{\text{cool}}/t_{\text{dyn}} = 10 \). In order to confine the gas in the central cavity to a thin geometry, we assume that the small amount of gas present in the inner cavity \((r \lesssim 1.75a_0)\) is isothermal, with an internal energy per unit mass \( u \approx 0.14(GM/R) \). The softening for the particles is adaptive with a minimum of 0.001\( a_0 \), while the MBHs are not softened.\(^5\) The sink radius, i.e. the radius below which any passing particle is counted as ‘accreted’ (Bate, Bonnell & Price 1995), is set to 0.005\( a_0 \), which corresponds to \( \approx 6 R_\odot \) for the secondary MBH. The necessity of such a high resolution is given by the high mass ratio between the MBHs and disc. Furthermore, the Eddington accretion rate of the MBHB is \( M_{\text{Edd}} \approx 8 M_\odot \text{yr}^{-1} \), therefore we need to have very light particles to resolve significantly sub-Eddington accretion rates. In our choice of units, the mass of a particle is \( M_{\text{SPH}} = 0.79 M_\odot \), resulting in \( M_\text{d} = 1.58 \times 10^{-2}M_\odot \). To ease comparison with similar previous work by Hayasaki et al. (2007), the scaleheight \( H \) of the disc is defined as the half thickness at which the density decreases by a factor of \( e^{-1/2} \), giving an aspect ratio of the disc of \( H/r \in [0.026, 0.013] \) for \( 3a_0 < r < 6a_0 \). The radial structure of this geometrically thin disc is well resolved at all times, since the azimuthally averaged smoothing length \( h \) is always much smaller than \( r \); for \( 3.6a_0 < r < 6a_0 \), we find \( h/r \in [0.017, 0.0106] \). The vertical structure is only marginally resolved, with \( h/H \in [0.756, 0.824] \). The gravitational stability of the disc can be assessed via the Toomre parameter \( Q \). A disc is stable to linear perturbations if

\[ Q = \frac{\nu c_s}{\pi G \Sigma} > 1, \]

\(^5\) Note that this does not introduce any artificial effect in the computation of the gravitational forces, since the MBHs are modelled as sink particles with sink radius larger than the typical softening of the gas particles.

\(^6\) Fig. 2 is made using splash (Price 2007).
well within the respective Hill radii of the two MBHs, ensuring disc stability against tidal stripping. The minidisc average density is \( \rho = 2 \times 10^{-15} \) g cm\(^{-3}\), giving an average mass of \( m_{\text{md}} \sim 3 M_\odot \) for the primary and \( m_{\text{md}} \sim 0.5 M_\odot \) for the secondary.

The detailed structure of the inflows and their physical properties strongly depend on the disc thermodynamics; moreover, their behaviour close to the MBHs is likely to be affected by the comparable size of the sink radius, the smoothing length and the physical size of the minidiscs. None the less, we can consider the rate at which particles cross the sink radii of the two black holes (what we call numerical accretion rate) as a good proxy to the rate at which material streaming from the outside circumbinary disc feeds to inner minidiscs. Such a rate is shown in Fig. 3 over 60 orbital periods. The orbital periodicity is quite striking, with a maximum-to-minimum ratio larger than a factor of 3. The maximum accretion rate declines from \( m \sim 1 \) to \( \sim 0.4 \), mostly as a result of the relaxation of the disc from its initial condition which is slightly out of equilibrium. Throughout the simulation, the outer disc shows a mild decretion behaviour. Its density peak moves outwards by \( \sim 7 \) per cent, while decreasing by \( \sim 6 \) per cent. As a consequence, the feeding rate to the cavity slowly declines in time reaching a quasi-steady value which is just slightly smaller (by less then 10 per cent) than the accretion rate on to the two MBHs. We note that the inflows occur on the MBHB orbital period, and are triggered by the secondary pulling material off the inner edge of the circumbinary disc at its apopesis passages. Their nature is therefore primarily gravitational, and should be weakly dependent on the exact thermodynamics of the disc. Our simulation therefore suggests that gravitational torques are effective in feeding periodically the central cavity at a rate which is a substantial fraction of the Eddington rate. This is in agreement with 2D Newtonian hydro-simulations by MacFadyen & Milosavljević (2008) and with full general relativistic hydro-simulations studying the evolution of MBHBs at closer separations (Farris, Liu & Shapiro 2011), in which similar streams feeding the two MBHs are observed down to \( \sim 10 M_\odot \) (\( \approx 3 \times 10^{-4} \) pc for our MBHs).

### 4.4 Analytical modelling of the minidiscs

As stated in the previous section, the detailed behaviour of the inflows close to the MBHs is likely to be affected by the comparable size of the sink radius, the smoothing length and the physical size of the minidiscs. As this is the case, our simulation does not have the resolution needed to properly handle the physics on such small scales. One possible solution would be to follow Hayakawa et al. (2008) and perform a higher resolution simulation of the inner region, using the inflows as boundary conditions to feed the minidiscs. This approach is extremely useful to highlight some of the physical processes driving the formation and evolution of the minidiscs. However, as we discuss in the following, the gas dynamics on small scales strongly depends on the details of the fuelling, and in particular on the angular momentum distribution of the inflowing gas. Running a comprehensive suite of high-resolution simulations would therefore be unfeasible. Here we decide to follow a different approach. We build a simple analytic parametric model for the minidiscs, using the inflow dynamics as a guidance to infer their physical size. We assume that the infalling material is periodically fed from the outer disc at some given rate \( \dot{m} \). To simplify the notation, we do not distinguish between the two MBHs; all the quantities appearing in the equations below are intended to be specific of the MBH under consideration, and the results apply to both MBHs (and both attached minidiscs).

The infalling material is captured in the Roche lobe of the 4th MBH with an average specific angular momentum \( J \) with respect to that MBH. This is converted to a characteristic size of a Keplerian orbit around the MBH simply as \( r_{\text{md}} = J^2 / GM \). If \( r_{\text{md}} < 3 R_\odot \), the inflow is radial and accretion proceeds in a Bondi-like fashion. Otherwise, the infalling material settles at a characteristic radius \( r_{\text{md}} \) around the MBH and dissipates its angular momentum through viscous processes, diffusing inwards and eventually being accreted by the two holes (a process similar to the debris fallback in tidal disruption induced accretion; Rees 1988). The time needed for an annulus of material at radius \( r \) to be accreted is

\[
\begin{align*}
\tau_{\text{acc}, \alpha} &= 0.31 \text{ yr} \, \alpha_0^{-1} \left( \frac{\rho_{0,3}}{\kappa_{0,1}} \right)^{-2} M_8 r_1^{7/5}, \\
\tau_{\text{acc}, \beta} &= 683 \text{ yr} \, \alpha_0^{-4/5} \left( \frac{\rho_{0,4}}{\kappa_{0,1}} \right)^{-2/5} M_8^{6/5} r_1^{7/5},
\end{align*}
\]

for \( \alpha \) and \( \beta \) discs, respectively. Here \( r_1 \) is the size of the minidisc in units of \( 10 R_\odot \).

If \( \tau_{\text{acc}} \) is longer than the binary orbital period \( P \), then a pair of persistent, periodically fed minidiscs is formed around the two MBHs. If, conversely, \( \tau_{\text{acc}} \) is shorter than \( P \), then the streaming periodicity is reflected in episodes of periodic accretion. For \( \alpha \)-discs, this happens for

\[
r_{\text{md}, \max} < 20 R_8 \, \alpha_0^{2/7} \left( \frac{\rho_{0,3}}{\kappa_{0,1}} \right)^{4/7} M_8^{-2/7} \, P_5^{2/7}
\]

8 For example, \( M_8 = M_{\text{f}} / 10^8 M_\odot \) is the mass of the 4th MBH in units of \( 10^8 M_\odot \); \( \alpha_0 = \alpha_0 / 0.3 \) is the viscosity parameter of the 4th disc in units of 0.3, and so on.

9 To sketch the situation, we describe the minidiscs as small steady accretion discs, even though the condition of stationarity of the system is not met because of the periodicity of the inflowing streaming.
(\(P_{\text{e}}\) is the MBHB period in years), while it never happens for \(\beta\)-discs. Fig. 4 shows the maximum radius \(r_{\text{md,max}}\) and the corresponding maximum mass \(m_{\text{md,max}}\) of an \(\alpha\)-minidisc. For average feeding rates of \(m \approx 0.3\) (consistent with the output of our SPH simulation), maximum disc sizes are in the range 10–40\(R_{\odot}\).

We can compare our model to the SPH numerical implementation of Hayasaki et al. (2008). They initially build discs of size \(\lesssim 0.2\alpha\) around the two (equal mass) MBHs. Matter is then added on eccentric orbits from moving injection points following the phase of the binary. Injected particles have much lower angular momentum than the outer edge of the minidisks, and are consequently readily accreted by the two MBHs, in a periodic fashion corresponding to the periodicity of injection. Their very eccentric injection is consistent with our radial accretion hypothesis, or with an \(\alpha\)-minidisc with \(t_{\text{acc}}\) shorter than \(P\). However, if the streams are readily accreted, it is difficult to justify the presence of the persistent, quite large minidisks they initially set around the two MBHs. Such long-lived discs require fueling streams with larger angular momentum (for which \(t_{\text{acc}} > P\)). Hayasaki et al. (2008) simulations show that if the infalling material has enough angular momentum to initially form persistent minidisks around the two MBHs, such discs are stable even in presence of tidal interactions, and a standard steady state analytical description is appropriate.

In our specific simulation, the average angular momentum of the particles captured by the MBHs within their respective Roche lobes implies \(r_{\text{md}} \approx 10–20R_{\odot}\), meaning that the accretion process must happen through viscous inspiral rather than through radial infall, disfavouring the Bondi-like accretion scenario. Moreover, \(r_{\text{md}} < r_{\text{md,max}}\) for both MBHs; for \(\alpha\)-discs, this implies period-related variability in the accretion and thus in the emitted luminosity.

5 Emitted Spectrum and Electromagnetic Signatures

In optically thick disc theory, each disc annulus emits blackbody radiation corresponding to its particular temperature. For a standard optically thick geometrically thin disc, this is (Frank, King & Raine 2002)

\[
T(r) = \left[\frac{3GM M f(r)}{8\pi r^3\sigma_{\text{SB}}}\right]^{1/4},
\]

where \(f(r) = (1 - 3R_{\odot}/r)^{1/2}\) (we assume a Schwarzschild black hole), and \(\sigma_{\text{SB}}\) is the Stefan–Boltzmann constant. The emitted luminosity at frequency \(\nu\) is given by the integral over the disc extension of the blackbody radiation emitted by each annulus, i.e.

\[
L_\nu = \frac{32\pi h_\nu v^3}{c^3} \int_{r_{\text{in}}}^{r_{\text{out}}} \frac{r^2dr}{e^{h_\nu/kT(r)} - 1},
\]

where \(r_{\text{in}}\) and \(r_{\text{out}}\) are the inner and outer edges of the disc respectively, \(h_\nu\) is the Planck constant and \(k\) is the Boltzmann constant.

In our model, each of the discs (the circumbinary plus the two minidisks) emits a thermal component according to equation (20). We therefore infer the presence of three distinctive continuum components: (i) a thermal component peaked in the optical/IR emitted by the circumbinary disc; (ii) two thermal components peaked in the UV associated with the inner minidisks; (iii) two X-ray power law associated with the tenuous hot electron plasma (corona; Galeev, Rosner & Vaiana 1979) surrounding the inner minidisks. These latter components are generated by the inverse-Compton scattering of UV photons emitted by the inner radii of the minidisks against the diffuse hot electron plasma embedding the inner part of the discs. The upscattered photon spectrum can be modelled as a power law spanning the soft/hard X-ray domain.

For the qualitative nature of our discussion, it is sufficient to consider a flat spectrum in \(L_\nu\), (Haardt & Maraschi 1993), normalized to give \(L_{0.5–10\text{keV}} = 0.03L_{\text{bol}}\) (Lusso 2010). The general features of the predicted continuum are shown in Fig. 5 for two selected systems. The presence of a cavity is reflected by the characteristic double bump; the one in the IR is due to the circumbinary disc, while the one in the UV is produced by the two inner minidisks.

On top of the continuum, several sets of emission lines are also expected, as in common AGNs, in particular (i) optical broad emission lines caused by the photoionization of the tenuous infalling material and the outer circumbinary disc (\(r > 0.01\) pc) by the inner ionizing UV source (the minidisks, \(r < 10^{-3}\) pc) and (ii) two \(6.4\) keV fluorescence broad emission iron lines (Fe \(K\alpha\) lines) produced by the reflection of the X-ray photons upscattered in the corona, on the surface of the inner accretion minidisks at only few Schwarzschild radii (Matt, Perola & Piro 1991). Moreover, several other spectral features may be present: (i) some weak optical/UV emission associated with the streaming material (Bogdanović, Eracleous & Sigurdsson 2009; Dotti et al. 2009b) and (ii) X-ray hotspots related to the infalling material shocking on to the outer edge of the inner minidisks (if \(t_{\text{acc}} > P\); see Section 5.1).

5.1 Characteristic signatures

Depending on the strength and angular momentum of the inflows, the features outlined above may or may not be present. We identify the following three different scenarios.
resulting periodic ionizing photon flux will cause optical broad- 
line strength to fluctuate by a factor larger than 2 over the orbital 
period. However, identification of such optical variability may be 
challenging, since it would require an all sky periodic spectroscopic 
monitoring. Large photometric surveys like the LSST (LSST Sci- 
ence Collaborations et al. 2009) will be sensitive to the luminosity 
of the optical continuum, which, depending on the characteristics 
of the system, may be dominated by the steady circumbinary disc 
(see Fig. 5). None the less, depending on the amount of reddening, 
colour selection may identify candidates showing a distinctive 
optical bump across the different wavelength bands. Note that in 
such close binaries the broad-line emission region is likely related 
to the circumbinary disc; the typical size of the broad-line emission 
region in AGNs is ~0.01–0.1 pc, much larger than the size of the 
minidisces, and correlates with the AGN luminosity (Kaspi et al. 
2005). Detection of separate sets of broad optical emission lines 
related to accretion on to the individual MBHs or the detection of 
broad emission lines shifted with respect to the narrow emission 
lines associated with the host (see e.g. Decarli et al. 2010) is there- 
fore unlikely in this case. Periodic soft-to-hard X-ray emission in 
response to the periodic continuum may be associated with the hot 
corona, and variable relativistic Fe Kα lines may also be present. 
This is the scenario suggested by our SPH simulation.

(iii) \( r_{\text{md}} > R_{\text{max}} (\alpha \text{-discs}) \) or \( r_{\text{md}} > 3R_5 (\beta \text{-discs}) \). The streams 
cannot be swallowed by the MBHs within an orbital period. Steady 
thin accretion minidisces form around the two MBHs. The accretion 
rate may be still quite variable, but redistribution of angular mo- 
mentum within the minidisces makes it problematic to directly infer 
accretion variability following the periodicity of the streams. Two 
superoosed relativistic Fe Kα lines associated with the minidisces 
are likely to be observable. Moreover, shocks created by the collision 
of the instreaming material and the minidisces at their outer edge may 
result in X-ray hotspots varying on the orbital period. The 
luminosity and electromagnetic frequency at which these regions irradiate 
depend on the relative velocity (\( v_{\text{rel}} \)) of the streams with respect to 
the minidisces. We can estimate \( v_{\text{rel}} \) for \( M_2 \), when the secondary is 
at the apocentre, assuming that its minidisc fills the whole Roche 
lobe. In this case, the emission is expected to peak in the hard X-ray, 
at about \( \sim v_{\text{rel}}^2 \mu \beta / k \gtrsim 10 \text{keV} \) (\( \mu \beta \) is the mean molecular weight 
of the plasma). The luminosity depends on the time-scale \( \tau_{\text{shock}} \) over 
which the shock irradiates its internal energy. Assuming \( \tau_{\text{shock}} \) to be 
of the order of the orbital period of the gas at the edge of the mini- 
disc (\( \tau_{\text{dis}} \)), we obtain \( L_{\text{shock}} \approx \mu v_{\text{rel}}^2 P_{\text{disc}} \gtrsim 10^{-3} L_{\text{edd}} \). Smaller 
\( \tau_{\text{shock}} \) and larger \( v_{\text{rel}} \) result in larger luminosities, but increasing 
\( v_{\text{rel}} \) shifts the peak of the emitted luminosity at considerably higher 
frequencies. Note that low values of \( v_{\text{rel}} \) are the most favourable 
conditions to efficiently perturb and bind large amounts of gas.

In the following we will focus on two distinctive signatures in 
the X-ray domain.

(i) Periodic X-ray emission related to the periodicity of the mini-
disc fed through the cavity. The outer circumbinary disc is relatively 
stable, and optical variability related to the streams may be 
overwhelmed by the outer disc continuum. Periodic variability, related 
to (i) periodic accretion of the minidisces, (ii) the associated varying 
flux of UV photons upscattered in the corona and (iii) X-ray hotspots 
created by the periodic streams colliding with the minidisces, may 
instead be easily detectable in UV and X-ray.

(ii) Double relativistic fluorescence Fe Kα lines. Such lines are 
a common feature in AGNs (Nandra et al. 2007; de La Calle Pérez 
et al. 2010), and are produced at only few Schwarzschild radii. 
The line profile strongly depends on the location of the last stable
orbit around the MBH (which depends on the spin magnitude) and on the disc inclination with respect to the observer. If accretion is efficient on two MBHs with different spin parameters (magnitude and/or orientation), a distinctive ‘double Fe K\alpha line’ feature may be observable in the hard X-ray spectrum of the source (see also Yu & Lu 2001).

In the two following sections, we will consider these two possibilities separately, proposing possible strategies for combining X-ray and PTA observations in the final discussion.

6 X-RAY PERIODIC VARIABILITY

6.1 Subpopulation of observable periodic sources

Being interested in variability related to the binary orbit, only a subset of the population shown in Fig. 1, having reasonably short periods, will be suitable targets. We therefore limit our study to the systems with \( P < 5 \) yr. The resulting population, represented in the left-hand plot of Fig. 6 for our default model, retains the mass, mass ratio and redshift distribution of the parent one. Detached systems obviously have a period distributions extending to \( P \ll 1 \) yr, while binaries still attached to their discs (relevant to our investigation) are abundant at \( P > 1 \) yr. Assuming that a fraction of 3 per cent of the bolometric luminosity is emitted in the X-ray,\(^{10}\) we can construct the log\(N\)--log\(S\) (i.e. the cumulative number of sources having a measured 0.5–10 keV flux larger than a certain value \( S \), as a function of \( S \)) of the emitting population. This is shown in the right-hand plot of Fig. 6 for both our default population (upper panel) and an alternative model in which the smaller \( \alpha \) viscosity and accretion rate imply an earlier detachment of the circumbinary disc. At large luminosities, such periodic binaries may add up to 5 per cent of the observed luminous X-ray sources in the sky. The currently operating X-ray all sky monitor MAXI (Matsuoka 2009) may already have detected a handful of them in its surveys. On the other hand, according to our default model, there might be up to 500 MBHBs significantly contributing to the PTA signal, showing periodic variability on a time-scale of 1–5 yr, in the sensitivity reach of the upcoming eROSITA observatory (Predehl 2010). We stress, once again, that our models assume that all merging MBHBs are accreting in their last evolutionary phase prior to coalescence, which may be in fact likely in gas-rich mergers, but none the less it remains an ad hoc assumption. If, for instance, only 10 per cent of the merging systems is active, the number of sources showing emission periodicity drops to 10–50, which is, however, still an interesting, sizeable sample.

6.2 Simulating observations: sampling and statistics

An X-ray luminosity above the instrument sensitivity is certainly not enough to identify a periodic source. The light curve has to be reconstructed with a minimum number of data points per orbit, for at least few orbits. In this subsection, we carry a detailed statistical study of the periodicity observability. The main goal is to provide the minimum required sampling cadence an observatory must have to efficiently identify periodicities. Since astronomical time series are always limited to a certain amount of data points, we address the problem of having a very coarse sampling rate of our light curve. We

\(^{10}\)This picture is appropriate for hot corona reprocessing. Inviscid minidiscs have matter/luminosity efficiency conversions similar to standard accretion discs (\( \epsilon \sim 0.03–0.1 \)), and even though detailed calculation of the spectrum is not available, we may expect comparable X-ray luminosities. If the minidiscs are instead steady, then X-ray luminosity emitted by shock-induced hotspots may be more than an order of magnitude smaller (\( 10^{-3}L_{\text{Edd}} \), compared to 3 per cent of \( 0.3L_{\text{Edd}} \) for corona reprocessing in our default model).
assume that the X-ray light curve mimics the numerical accretion rate of Fig. 3.

The complete light curve of our simulation consists of \( L = 898 \) points, sampling \( \sim 80 \) orbits; each orbit is therefore sampled by \( \sim 11 \) equally spaced points. We create subsamples of this light curve by selecting a random starting zero-point \( \tilde{p} \) and then considering all the subsequent \( n \tilde{p} \) points, where \( n = 1, \ldots, N - 1 \), and \( p \in \{1, 2, 3\} \). A stride of \( p = 1 \) translates into taking 11 points per orbit (points \( \tilde{p}, \tilde{p} + 1, \tilde{p} + 2, \ldots \)), \( p = 2 \) means taking approximately six points per orbit (points \( \tilde{p}, \tilde{p} + 2, \tilde{p} + 4, \ldots \)) and \( p = 3 \) means taking approximately four points per orbit (points \( \tilde{p}, \tilde{p} + 3, \tilde{p} + 6, \ldots \)). We refer as \( 6p0p1 \) a sample covering six orbits with all points included in the sampling (total of \( N = 65 \) points); \( 3p0p2 \) refers to a string of three orbits sampled at every other point (total of \( N = 18 \) points), etc.

The reason we consider different sampling is to investigate the effect of the observation cadence. Assuming a binary with, e.g., 3-year period, \( p1 \) corresponds to a three month cadence, \( p2 \) to six months, etc. If the number of points covered by the typical subsample is \( N \), we can draw a maximum of \( \zeta = L - N \) subsamples.\(^\text{11}\) To account for other possible sources of variability (AGNs are generally variable in the X-ray; Grupe, Thomas & Beuermann 2001), the samples are polluted with Gaussian noise with a dispersion \( \sigma = 10 \) per cent of the maximum accretion rate/luminosity of each specific sample. Each subsample is Fourier transformed using the normalized Lomb–Scargle periodogram (Scargle 1982), and the false alarm probability (FAP) of the highest detected peak is computed. After a cross-check that this highest peak corresponds to the fundamental frequency of the binary \( f_0 = 1P_0 \), the FAPs of the \( f_0 \) frequency are binned logarithmically and shown in the histograms in Fig. 7. The left-hand panel shows that a minimum of \( \sim 15-20 \) data points covering at least three orbits are needed in order to identify a decent fraction (\( \approx 50 \) per cent) of the sources at a 10 per cent FAP level. Increasing the number of data points dramatically increases identification performances; already with \( \approx 30 \) points in the light curve we can detect most of the sources (\( \approx 70 \) per cent) at a FAP level of 1 per cent.

In order to cross-check with a possible PTA identification of the source, it is important to know to which accuracy the fundamental frequency can be recovered from the Fourier spectra. For an observation time \( T \), the frequency resolution bin of the observation is \( 1/T \). Therefore, if we observe a frequency \( f_0 \), the relative accuracy of the observation is \( \Delta f/f_0 = (1/T)f_0 \). However, we can oversample the spectrum to get a better constrain on the observed frequency. For all periodograms shown, we have used an oversampling factor (ofac) of 8. In evenly sampled data, we expect the analysis to be only mildly dependent on this parameter: for some choices of low ofac (\(<4\)) in very short data sets, we observed the appearance of aliasing especially if high-frequency components were not suppressed (hi-fac > 1 as defined in Scargle 1982). We thus concluded, especially in the prospect of real unevenly sampled astrophysical data, that the high-frequency noise should be suppressed using very low hi-fac (\(<1\)) and at the same time using high oversampling (ofac >4).

For the identification of any frequency, we set the requirement \( |f_{\text{detected}} - f_{\text{true}}|/f_{\text{true}} < 5 \) per cent. In Fig. 8, we show the dependence of the number of correctly identified frequencies on the number of points available in the light curve, for different samplings (i.e. for different number of points per orbital period). We consider only frequency peaks identified at a 3\( \sigma \) significance or more. Note the continuity of the \( p2, p3 \) and \( p4 \) curves (six, four and three points per orbit, respectively), meaning that, as long as we sample three or more orbits, the fraction of correctly identified systems

\(^\text{11}\) Note that the subsamples in general overlap with each other, therefore cannot be regarded as uncorrelated. However, a much longer simulation would be needed to draw a large number of uncorrelated samples, which would have been prohibitively time consuming.

Figure 7. Detection statistics of periodic sources. In each plot, we show histograms of the number of samples correctly identified within a certain FAP. Left-hand panel: statistics is constructed with six points per orbit over three orbits (\( 3p0p2 \), red histogram), or with three points per orbit over six orbits (\( 6p0p4 \), blue histogram). Central panel: statistics is constructed with 11 points per orbit over three orbits (\( 3p0p1 \), red histogram), or with six points per orbit over six orbits (\( 6p0p2 \), blue histogram). Right-hand panel: statistics is constructed with 11 points per orbit over five orbits (\( 5p0p1 \), red histogram), or with 11 points per orbit over six orbits (\( 6p0p1 \), blue histogram). In all cases, a Gaussian noise with \( \sigma = 10 \) per cent of the maximum luminosity was added to each individual sample.

Figure 8. Percentage of correctly identified frequencies at 3\( \sigma \) level, within a fractional frequency error <5 per cent, as a function of the number of points in the light curve. For each sampling (shown in different colours; see legend in figure), points from left to right correspond to three, four, five and six full orbits of the observed source. A 10 per cent random Gaussian noise was added to each subsample.
depends strongly on the total number of points in the light curve and only mildly on the number of points per orbit (as long as this is larger than three). The discontinuity with respect to the \( p1 \) series tells us that sampling six orbits with six points per orbit is much better than sampling three orbits doubling the observation frequency. Sampling at least three orbits is a minimum identification requirement, because other sources of uncorrelated noise may wash out the significance of the periodicity if the statistics is too poor. In general, having at least \( \sim 15-20 \) points in the light curve, sampling three or more orbits, is a minimum requirement for a confident identification of the periodicity.

7 DOUBLE IRON LINES

7.1 Relevant source population

Spectral lines formed in relativistic accretion discs are distorted due to Doppler and relativistic effects producing a characteristic shape, which may be modelled to determine the properties of the space–time around the compact object (Fabian et al. 1989; Laor 1991). In particular, observations of supermassive black holes at the centre of galaxies have revealed broad skewed lines, which have allowed constraints to be placed on the spin of the black holes (Tanaka et al. 1995; Nandra et al. 1997, 2007; Miller 2007; de La Calle Pérez et al. 2010). The current status of our knowledge of relativistically broadened iron lines from AGN is summarized in Guainazzi et al. (2011).

Double relativistic Fe Kα lines are more likely to be present in a ‘steady’ environment, where material rising from the accretion discs (driven by heating and vertical stratification of the accreting material, magnetic turbulence, winds, or star–disc interactions) has time to shape a tenuous hot electron plasma corona. They are therefore likely to appear in situations where \( t_{\text{acc}} > P_{\text{rot}} \), but we do not exclude such possibility otherwise. Broad Fe Kα lines appear to be common in AGN (Nandra et al. 2007; de La Calle Pérez et al. 2010), yet their identification requires a large number of collected X-ray photons, i.e. deep, targeted, time-consuming observations. It is therefore reasonable to consider only sources that may be individually resolvable in PTA campaigns, and consequently localized in the sky to some accuracy (Corbin & Cornish 2010; Sesana & Vecchio 2010).

We therefore estimate the population of sources suitable for double Fe Kα line detection by considering individually resolvable sources only. Unfortunately, the concept of resolvability has not been deeply investigated in the PTA observation context (and certainly not for eccentric binaries). In the circular binary case, a rough ‘one bin’ rule estimate provides \( \sim 10 \) bright resolvable binaries (Sesana et al. 2009). However, such estimate does not take into account the spatial information enclosed in the detection with an array of pulsars; two sources at different sky locations contribute differently in each pulsar, and their signals may be disentangled even if their frequencies fall in the same bin. Boyle & Pen (2010) estimated that, exploiting the spatial information enclosed in the signal, \( 2N/7 \) sources per frequency bin would be resolvable by an array of \( N \) pulsars. For our estimate, we therefore (somewhat arbitrarily) pick the 10 strongest GW sources per frequency bin, and impose a further cut at 1 ns (dimmer sources would not be individually detectable anyway). This leaves us with \( \sim 100 \) sources, the precise number being vastly independent of the details of the global population model, but only on the PTA observation time (assumed to be 10 years). Such population is shown in Fig. 9 for our default model. Blue histograms represent MBHBs still attached to their circumbinary discs (i.e. with \( a > a^c \)), and therefore plausibly hosting small accretion minidiscs. We are left with \( \sim 20 \) bright, low-redshift sources.

Note, however, that our estimation of \( a^c \) is quite conservative. In deriving it, we equated the GW shrinking time to the viscous time of an \textit{unperturbed} disc at a radius corresponding to the inner rim of the cavity. Tanaka & Menou (2010), however, showed that the steep density gradient at the inner edge of the disc will shorten the inward diffusion time-scale of the gas, causing a significant delay in the disc–binary detachment, therefore increasing the number of sources with observable minidiscs. Moreover, in the \( \beta \)-disc picture, the consumption time of a minidisc of \( \sim 30R_S \) can easily exceed \( 10^4 \) yr (see equation 17), implying that a relevant fraction of the detached systems may still be significantly accreting (Chang et al. 2010; Tanaka et al. 2011). It is in any case worth noting that the relevant population is sizeable (maybe few to hundred sources) and probably mostly composed by very low redshift systems (\( \varepsilon < 0.2 \); see lower left panel of Fig. 9).

7.2 Simulations of double Fe Kα line observability

We assume the iron lines to be similar to that observed from the archetype MCG\texttextendash6-30-15 (Tanaka et al. 1995; Brenneman & Reynolds 2006; Miniutti 2007). We aim to assess the feasibility of detecting pairs of relativistic iron lines which may be emitted from the inner minidiscs around the MBHs discussed herein, and to use them to constrain the properties of the binary system, i.e. MBH spins, relative radial velocity, inclination, etc. In order to investigate the feasibility of utilizing broad Fe Kα emission from an
MBHB, spectra were simulated with \texttt{xspec v12.6}\textsuperscript{12} (Arnaud 1996). We assume the availability of a next-generation high throughput X-ray telescope, in comparison to current instruments, i.e. \textit{XMM–Newton}. Specifically, we use the response matrices created for the proposed \textit{Athena} mission concept.\textsuperscript{13} The current implementation envisions two focal plane instruments: (i) a wide field CCD imager (WFI) with a 25 arcmin\(^2\) field of view and (ii) a narrow field of view (2.4 arcmin\(^2\)) X-ray calorimeter spectrometer (XMS) providing spectra in the Fe \textit{K}\alpha region with resolution of 5 eV. Hereafter, we assume that the MBHB has been identified, opening the possibility to study it in high spectral resolution with the XMS.

We assume a total mass of \(\sim10^7\) M\(_\odot\) for the binary pair and a mass ratio of 1/3, consistent with the binary population presented in Section 7.1. We further assume a fiducial redshift of 0.1, a bolometric luminosity of 10 per cent \(L_{\text{Edd}}\) and an associated bolometric correction of \(\sim20\) (Vasudevan & Fabian 2009; Lusso 2010), i.e. \(f_{\text{bol}} \sim 10^{-11}\) erg \(s^{-1}\) cm\(^{-2}\), \(f_{\text{bol}} \sim 3 \times 10^{-11}\) erg \(s^{-1}\) cm\(^{-2}\). The parameters of our fiducial system are marked by vertical ticks in Fig. 9. The black hole binary is assumed to have a periodicity consistent with a velocity separation of 10\(^4\) km \(s^{-1}\). For simplicity, we assume the energy of the first iron line to be consistent with 6.4 keV emission at a redshift of 0.1, while the second line is blueshifted by 10\(^3\) km \(s^{-1}\) relative to the first line.

In order to simulate the expected X-ray spectrum from such a system, we first make a number of simplifying assumptions. The broad-band continuum emission from each accreting black hole is assumed to have a power-law shape, where the spectral index has a value of \(\Gamma = 1.8\). The relativistic iron line is modelled using the \(\text{\textit{Laor}}\) line profile (Laor 1991), where the equivalent width of the line is set to be approximately 150 eV, as expected for reflection from a disc surrounding a black hole (George & Fabian 1991). The radial emissivity profile of the disc is fixed at \(R^{-3}\), and the outer radius of the emitting region is assumed to be 40\(R_\odot\) (see Section 4.4). A narrow line consistent with Fe \textit{K}\alpha emission from material at much larger radii is also included. These values are consistent with current available observations (e.g. de La Calle Pérez et al. 2010; Guainazzi et al. 2011). A more thorough and self-consistent modelling of the expected spectrum from an MBHB system, for example, including complex absorbers and self consistently calculating the continuum plus reflected emission is beyond the scope of this work (e.g. see Brenneman et al. 2011), and as such we defer a detailed analysis of this problem for future investigations.

The model described above was defined in \texttt{xspec} as \texttt{phabs((zpo + laor) + (zpo + laor) + gauss)}, and 200 spectra were simulated in each case, using the latest available response matrices. An example of a simulated spectrum is shown in the top panel of Fig. 10. The interstellar absorption component is modelled via the \texttt{phabs}; however, as we are interested in the Fe line region, modest values for the column density (i.e. \(\lesssim10^{22}\) cm\(^{-2}\)) will have a negligible effect. Hence, the column density was held constant in all fits and as such is not discussed further. The spectral index \(\Gamma\) is assumed to be identical in both sources as it is inherently difficult to accurately extract both indices correctly, if the difference between them is small, due to the narrow assumed bandpass (0.5–10.0 keV). After each spectrum was simulated, the \texttt{fit} and \texttt{error} commands were then run in order to measure the best-fitting parameters given the signal-to-noise ratio (S/N) of the spectrum. This results in scatter around the defined model which is indicated in colour in Fig. 11. This model was then removed and a new ‘blind’ model was defined. In this case, only two parameters are known a priori, the redshift and the line-of-sight column density. Additionally, the inclination of the line is fixed at the known value. All other parameters are initialized at reasonable values given the observed spectrum. The model is first fitted with only a single relativistic line and the narrow Fe \textit{K}\alpha line [\texttt{phabs((zpo + laor) + gauss)}]. After the best fit is obtained, a second relativistic line is added to the model and the significance of this line is calculated using a simple \texttt{ftest} (see Fig. 10). In all cases, the inclinations of both lines are tied to each other, as we expect the angular momentum of each black hole to be aligned to the orbital angular momentum of the binary in the late stage of a wet merger (Bogdanović, Reynolds & Miller 2007; Dotti et al. 2010).

We focus on two primary cases for the spin of the merging black holes (i) \(a_1 = 0.0, a_2 = 0.9\), and (ii) \(a_1 = 0.3, a_2 = 0.7\). In Fig. 10, we plot an example best-fitting model to one of our simulated spectra in case (i) above. Here the inclination is 30\(^\circ\). In Fig. 11, we display the results of the fit to the relativistic iron lines. The input model is in colour, whereas the subsequent blind fits are indicated by the plus and cross symbols. The exposure time in this case is 200 ks. It is clear that it will be possible to resolve and constrain both Fe lines to high accuracy in both cases. At higher inclinations, our ability to accurately recover the line parameters deteriorates, as expected due to the relative narrowing of the line profile. We note that in all of the lower inclination models (\(\lesssim40\)^\circ\)), all of the blind fits require

\textsuperscript{12} http://heasarc.nasa.gov/xanadu/xspec/
\textsuperscript{13} http://www.mpe.mpg.de/athena/home.php?lang=en
the presence of two relativistic lines at \(>5\sigma\) confidence level. As we move to higher inclinations, this decreases somewhat whereby at 60° only \(\approx 95\) per cent of our realizations require two relativistic lines.

Because of the degeneracy in the line shape with changing inclination and/or spin, the primary source of uncertainty in the spin determination is the inclination of the MBHB. If we relax the initial inclination constraint imposed above, our ability to accurately determine the line parameters is weakened. For example in case (i) above, a second line is required at \(\geq 5\sigma\) level in only \(\approx 85\) per cent of our models, while only \(95\) per cent require a second line at the \(3\sigma\) level. We expect, however, to get some indicative prior on the system inclination from PTA observations. In fact, Sesana & Vecchio (2010) showed that the GW signal analysis will enable a determination of the system inclination within a \(\sim 20°\) accuracy (assuming a source S/N of 10). Models were also created with additional narrow emission/absorption lines consistent with Fe \(xxvi/xxvii\). As in the case of the narrow 6.4 keV line, the exquisite resolution provided by the calorimeter in the Fe K region allows these lines to be easily resolved. We also experimented with decreasing the velocity separation of the broad-line components; however, in this case the results are strongly inclination-dependent. The use of a more accurate relativistic line profile model, e.g. \texttt{kerrdisk} (Brenneman & Reynolds 2006), would help in this case.

There are a number of caveats which should be considered when interpreting the results above. For example, even in the case of MCG–6–30–15 where long high S/N observations exist, there is uncertainty in the parameters of the observed line when modelled by different groups (Fabian et al. 2002; Reynolds et al. 2004; Brenneman & Reynolds 2006). In addition, complex relativistic effects (e.g. light bending; Miniutti & Fabian 2004) may render the line indistinguishable from the continuum in the absence of highly sophisticated spectral modelling (e.g. Bhayani & Nandra 2011). Finally, we note that an alternative physical model has been proposed to explain the relativistic lines observed in numerous AGN. In this scenario, the lines are produced by a combination of non-relativistic fluorescent line emission from distant material and complex absorption by material in the inner accretion flow (for further details, see e.g. the review by Turner & Miller 2009).

8 DISCUSSION AND CONCLUSIONS

The distinctive signatures of PTA-MBHBs highlighted above open interesting scenarios for combining pulsar timing and X-ray observations in the coming years. Even though current PTA efforts (PTA, EPTA, NANOGrav) may eventually succeed in the detection challenge, multimessenger astronomy prospects are particularly promising in the context of the planned SKA. If a nominal 1ns sensitivity level will be achieved, SKA will make GW astronomy with pulsar timing possible. Several hundreds of signals emitted by the low-redshift population of MBHBs will add up to form a confusion noise, on top of which some (possibly a hundred) sources will be individually resolvable and their position in the sky determined.

In this context, X-ray (but also other bands, not considered in this paper) observations may play either a preparatory or a follow-up role. On one hand, the detection of a population of periodic X-ray sources in the present decade may provide useful information for PTA observations. By the time the SKA will be online, we may have identified a catalogue of systems in the sky from which we expect to detect GW signals, being also able to give indications about the expected frequency. This may substantially facilitate the PTA detection and analysis pipelines by reducing the search parameter space for specific signals. On the other hand, several resolvable sources with high S/N may be located in the sky with enough accuracy to make follow-up X-ray monitoring possible. To be conservative, according to Sesana & Vecchio (2010), typical PTA source sky location accuracy is expected to be of the order of tens of square degrees for a source S/N = 10 (but, under specific detection conditions, it can actually be a factor of a few better; see Corbin & Cornish 2010). According to Fig. 9, typical masses are greater than few times \(10^8 M_{\odot}\) at \(z < 0.3\). Such systems should be hosted in massive galaxies with stellar bulges of masses \(>10^{11} M_{\odot}\) (Gültekin et al. 2009), whose number density is \(\lesssim 10^{-3} \) Mpc\(^{-3}\) (Bell et al. 2003). Considering a sky location error of \(10\) deg\(^2\), the comoving volume enclosed in the error box is \(2 \times 10^{-3} \) Gpc\(^3\); maybe up to few thousands candidates falls in the error box. Since the presence of a distinctive counterpart relies on the assumption of accretion, active galaxies only should be selected, which may leave us with few dozens of candidates.
Further down selection may be performed by keeping galaxies that show signatures of recent merger activity. If just one or at most a few systems are identified, ultradeep X-ray exposure may be used to reveal characteristic double relativistic Fe Kα emission lines. Such an investigation will likely be possible in the next decade with future generation of PTAs (in particular with the SKA) and X-ray observatories. In this paper, we quantified for the first time the population of expected sources and their characteristic signatures, and we proposed strategies for coordinating multimessenger observations. We summarize in the following our main results.

(i) About a thousand MBHBs are expected to contribute to the PTA detectable signal in the frequency range $10^{-9}$ to $10^{-6}$ Hz, at a level of 1 ns or more. Uncertainties in the MBHB population model and in the detailed dynamics of the contributing systems may impact such figures by a factor of a few.

(ii) We assumed the standard picture of MBHB migration in circumbinary discs. For popular geometrically thin, optically thick discs described by typical parameters, we find that many of these PTA sources (order of few hundreds, 20–50 per cent of the total population, for $0.1 < \alpha < 0.3$ and $0.1 < \dot{m} < 0.3$) are coupled to their circumbinary discs, making a strong case for looking to possible electromagnetic signatures rising by the disc–binary dynamical interplay.

(iii) Detailed SPH simulations of the eccentric binary–disc interaction highlight periodic streams of gas leaking from the inner edge of the circumbinary disc, through the low-density central cavity. The rate at which such streams feed the central binary can be a significant fraction of the Eddington accretion rate of the MBHB (after initial relaxation of the system, we find an average $\dot{m} \approx 0.4$). The streaming periodicity is extremely sharp, with a variation in the MBHB feeding rate of a factor of 2–4.

(iv) By modelling analytically the dynamics and the emitted spectrum of the infall-fed minidisks forming around the two MBHs, we identified several observational features that may be distinctive of an accreting MBHB. We concentrated on the X-ray domain by separately studying X-ray periodicity and double relativistic Fe Kα lines.

(v) Assuming all merging MBHBs are accreting, we estimate about 100–500 (depending on the detailed property of the circumbinary disc) periodically variable X-ray sources with periods between 1 and 5 years. The observed flux on Earth for most of the sources is larger than $10^{-11}$ erg s$^{-1}$ cm$^{-2}$, within the sensitivity reach of the upcoming X-ray all sky monitor eROSITA. However, a light curve with at least 15–20 points (eROSITA will point each region of the sky at eight different times only) is required for a statistically significant detection of such a periodicity.

(vi) Iron line identification requires deep, targeted observations. We therefore consider as suitable candidate only those PTA sources that are individually resolvable, for which the sky location can be identified to some accuracy. The number of targets is of the order of few tens, depending both on the ability of PTA of resolving sources in the sky and on the details of the disc–MBHB decoupling and the physical nature of the minidiscs.

(vii) Assuming that the accretion flow on to each black hole is capable of generating relativistic Fe Kα emission lines, we have demonstrated that it will in principle be feasible, via high spectral resolution observations with a next generation X-ray observatory (e.g. Athena), to identify double Kα line features. Such features can be used to estimate the properties of the black holes, forecasting, in combination with modelling of the PTA signal, the possibility to constrain the space–time around the black hole to unprecedented accuracy.

(viii) Note that we assumed all MBHBs to be surrounded by a circumbinary disc in their late evolutionary stage, prior to merger. Even though this is likely for gas-rich mergers, it might be a too extreme assumption for the massive, low-redshift sources relevant to our study. Even assuming that only 10 per cent of the systems are surrounded by a circumbinary disc, the number of detectable sources is still interesting: about 10–50 bright, periodic X-ray sources, at least a few individually resolvable PTA sources showing double Fe Kα line profiles.

Even though GW detection of MBHBs remains a challenging task for the present and future astrophysical generations, systematic pulsar timing campaigns are ongoing, and their accuracy and sensitivity will inevitably improve in the coming years. Pulsar timing will therefore provide a safe, open GW window on the low-frequency Universe. The combination with electromagnetic observations, such as the ones proposed in this paper, will help exploiting GW detection capabilities at their best, providing a lot of information about the population and dynamics of MBHBs. The present paper is just a first step into the realm of multimessenger astronomy with pulsar timing, hoping that investigators from both the GW and the X-ray/optical/radio communities will take the challenge following our footsteps.

ACKNOWLEDGMENTS

As we were completing this work, we became aware of a concurrent independent study by Tanaka et al. (2011) addressing similar questions. We thank J. Krolik and Margherita Giustini for insightful discussions, and Francesco Haardt and Elena Rossi for the detailed comments on the manuscript. CR wishes to thank Jorge Cuadra for his modified version of GADGET-2.

REFERENCES
