Identifying decaying supermassive black hole binaries from their variable electromagnetic emission

Zoltán Haiman\textsuperscript{1}, Bence Kocsis\textsuperscript{2,3}, Kristen Menou\textsuperscript{1}, Zoltán Lippai\textsuperscript{4} and Zsolt Frei\textsuperscript{4}

\textsuperscript{1} Department of Astronomy, Columbia University, New York, NY, USA
\textsuperscript{2} Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
\textsuperscript{3} School of Natural Sciences, Institute for Advanced Study, Princeton, NJ, USA
\textsuperscript{4} Institute of Physics, Eötvös University, Budapest, Hungary

Received 30 October 2008, in final form 9 December 2008
Published 20 April 2009
Online at stacks.iop.org/CQG/26/094032

Abstract
Supermassive black hole binaries (SMBHBs) with masses in the mass range \( \sim (10^4–10^7) M_\odot/(1+z) \), produced in galaxy mergers, are thought to complete their coalescence due to the emission of gravitational waves (GWs). The anticipated detection of the GWs by the future Laser Interferometric Space Antenna (LISA) will constitute a milestone for fundamental physics and astrophysics. While the GW signatures themselves will provide a treasure trove of information, if the source can be securely identified in electromagnetic (EM) bands, this would open up entirely new scientific opportunities, to probe fundamental physics, astrophysics and cosmology. We discuss several ideas, involving wide-field telescopes, that may be useful in locating electromagnetic counterparts to SMBHBs detected by LISA. In particular, the binary may produce a variable electromagnetic flux, such as a roughly periodic signal due to the orbital motion prior to coalescence, or a prompt transient signal caused by shocks in the circumbinary disc when the SMBHB recoils and ‘shakes’ the disc. We discuss whether these time-variable EM signatures may be detectable, and how they can help in identifying a unique counterpart within the localization errors provided by LISA. We also discuss a possibility of identifying a population of coalescing SMBHBs statistically, in a deep optical survey for periodically variable sources, before LISA detects the GWs directly.

The discovery of such sources would confirm that gas is present in the vicinity and is being perturbed by the SMBHB—serving as a proof of concept for eventually finding actual LISA counterparts.

PACS numbers: 95.30.Sf, 04.30.Tv, 98.80.Es, 98.54.Aj, 98.62.Js, 98.62.Mw, 04.50.Kd, 04.80.Cc, 04.80.Nn

(Some figures in this article are in colour only in the electronic version)
1. Introduction

The anticipated detection by LISA of gravitational waves emitted during the coalescence of supermassive black holes (SMBHs) in the mass range \((10^4 - 10^7) M_\odot/(1 + z)\) will constitute a milestone for fundamental physics and astrophysics. While the GW signatures themselves are a rich source of information, if the GW source produces electromagnetic (EM) radiation, and if the object can be securely identified in EM bands, this would open up entirely new scientific opportunities. The simultaneous study of photons and gravitons from a single source could probe fundamental aspects of gravitational physics [1, 2]. The GW sources can also be used as self-calibrated standard sirens [3], and cosmological parameters can be determined if the source redshift is identified [4, 5]. Finally, for many events in the above mass and redshift range, LISA will be able to measure the masses and spin vectors of the SMBHs, their orbital parameters and their luminosity distance, to a precision unprecedented in any other type of astronomical observation [6, 7]. If a counterpart is known, then the Eddington ratios and other attributes of black hole accretion physics can be studied in exquisite detail [2, 5].

Motivated by the above possibilities, several recent studies have addressed the question of whether finding a counterpart will be feasible, given LISA’s localization errors [2, 4, 5, 8, 9]. The crucial uncertainty, of course—besides the uncertain rate of merger events detectable by LISA itself (e.g. [12] and references therein)—is the nature (luminosity, spectrum and time-evolution) of any EM emission produced by coalescing SMBHBs during the GW-inspiral stage. Such emission would have to be related to gas in the vicinity, and possibly accreting onto the coalescing SMBHs. The gas around the BHs would likely settle into a rotationally supported, circumbinary disc. In a geometrically thin disc, the torques from the binary create a central cavity, nearly devoid of gas, within a region about twice the orbital separation [13] (for a nearly equal-mass binary), or a narrower gap around the orbit of the lower-mass BH in the case of unequal masses \(q \equiv M_1/M_2 \ll 1\) and larger orbital separations [14]. In the latter case, the lower-mass hole could ‘usher’ the gas inward as its orbit decays, producing a prompt and luminous signal during coalescence. In the former case, residual gas flows into the cavity and onto the BHs, such as suggested in numerical simulations [15–17], may still produce non-negligible EM emission. Around the time of coalescence, the gravitational waves shear the circumbinary gas and could brighten its emission detectably [18]. Finally, SMBHBs recoil at the time of their coalescence due to the emission of GWs, at speeds up to 4000 km s\(^{-1}\) (e.g. [19] and references therein). The gas disc will respond promptly (on the local orbital timescale of the disc material) to such a kick, which may produce prompt shocks, and transient EM emission after coalescence [20].

The complex processes involved in ultimately producing any EM emission remain poorly understood, and the level of the luminosity produced, as well as its spectrum and time-evolution, are essentially unknown. However, any emission during and promptly following the inspiral stage is likely to be variable. In this contribution, we discuss three issues related to using variability to identify EM counterparts. Can LISA events be localized to within the field of view of astronomical instruments (several deg\(^2\)), hours to weeks prior to coalescence (section 2)? What is the response of the circumbinary gas to the gravitational recoil (‘kick’) of the SMBHB at coalescence (section 3)? Can we identify coalescing SMBHBs before the launch of LISA, as variable sources, due to periodic perturbations in the circumbinary gas (section 4)?

5 Such brightening due to kick-induced disc heating can persist on much longer timescales, and could produce detectable emission—leading to recent proposals that this can help identify a population of such sources before LISA’s launch. This possibility will be briefly discussed further in section 4.
2. Monitoring the 3D LISA error box prior to coalescence

The first and most fundamental question in searching for any EM counterpart by looking for variable emission, during the final stages of coalescence, is the accuracy to which the LISA source can be localized at various look-back times prior to the coalescence. (Coalescence is taken to occur when the binary separation reaches the innermost stable circular orbit; ISCO.) Here we present time-dependent localization errors, obtained by the harmonic mode decomposition method [8]. This technique uses the restricted, post-Newtonian approximation for the GW waveform, and applies the Fisher matrix technique to the Fourier transform of the waveform, to forecast parameter uncertainties. Orbits are assumed to be circular, and spins are neglected. These assumptions are conservative for circular, nearly equal mass sources that are in the LISA band for at least several weeks, since in this case, higher-order harmonics, spin precession and the GW signal from the ring-down stage can all improve localization errors. In practice, these improvements generally may not be significant until the last day or so of the merger, if the signal is in the LISA band for at least a month. This was shown to be the case when including spin precession but neglecting higher harmonics and eccentricity [9, 10], and also when including higher harmonics but neglecting spin precession and eccentricity [11] (see related recent discussions in [2] and [12]). The seventeen-dimensional parameter space describing the general binary inspiral is split into ‘slow’ and ‘fast’ parameters, based on the timescales on which they modulate the waveform, and the two sets are assumed to be decoupled. The angular coordinates of the source, and its luminosity distance—representing the 3D localization of the source—are ‘slow’ Fisher parameters, and information on these parameters is derived from the annual motion of LISA around the Sun. The reader is referred to [8] for a full list of assumptions and details about the method.

The time-evolving 1σ errors on the two-dimensional sky position are shown for an equal-mass binary, \( M_1 = M_2 = 10^6 M_{\odot} \), at redshift \( z = 1 \) in figure 1 (corresponding roughly
to the optimal choice of mass/redshift combination; other masses and redshifts yield poorer localization). The HMD method, by construction, approximates the sky position errors by ellipses. Figure 1 shows the gradual improvement in localization, in the form of the major axis ($2a_N$), the minor axis ($2b_N$) and the equivalent diameter ($2r_N = \sqrt{4a_N b_N}$). Figure 1 displays results for three separate cumulative probability distribution levels, 90%, 50%, 10%, so that 10% refers to the best 10% of all events, as sampled by the random distribution of five angular parameters. The evolution of errors scales steeply with look-back time for $t_f \gtrsim 40$ days. For smaller look-back times, errors essentially stop improving in the ‘worst’ (90% level) case, improve with a relatively shallow slope for the ‘typical’ (50% level) case and improve more steeply in the ‘best’ case (10% level among the realizations of fiducial angular parameters). Similar evolutionary trends are seen for the luminosity distance $d_L(z)$ errors (shown in the right panel in figure 1).

Figure 2 displays contours of fixed ‘advance warning time’ for typical (50%) events, adopting a 10 deg$^2$ FOV as a reference. The contours are logarithmically spaced, with solid contours every decade and the shaded region highlights the $(M, z)$ region where at least 10 day advance notice will be available. This figure shows that 10 day advance warning to cover the full error ellipsoid with a single LSST pointing is possible for a range of masses and source redshifts, up to $M \sim 3 \times 10^7 M_\odot$ and $z \sim 1.7$.

The results shown in figures 1 and 2 suggest that it will be possible to identify, prior to merger, a small enough region in the sky where any prompt electromagnetic (EM) counterpart to a LISA inspiral event will be located. Given sufficient advance notice, it will then be possible to trigger a world-wide monitoring campaign, to search for EM counterparts as the merger proceeds (and also during the most energetic coalescence phase).

The several square-degree field will, of course, contain a very large number of sources (a few $\times 10^5$ galaxies in total, at the limiting optical magnitude of $\sim 27$ mag that may be relevant; e.g. [21] and discussion below). Having predictions for the spectrum and time dependence of a coalescing SMBH binary, and therefore knowing what to look for, would obviously greatly help in identifying the counterpart, possibly allowing an identification.
even post-merger. Such a prediction, however, would require understanding the complex hydrodynamics and radiative properties of circumbinary gas at the relevant small separations of a few to a few thousand Schwarzschild radii. This is a notoriously difficult problem even in the much simpler case of steady accretion onto a single SMBH [22]. This suggests that the best strategy may be an ‘open-minded’ search for any variable signatures prior and during coalescence.

There will be several ways, however, to cut down on the list of possible counterparts, using the photometric redshifts of the candidates (restricting the search to within a narrow, \( \delta z \lesssim \) few percent, redshift slice; see the right panel in figure 1), the expected luminosity of the source, and the other parameters of the binary provided by \textit{LISA} (for example, a variable EM signal may be much more likely if the spin and the orbital angular momenta are known to be aligned, as this may indicate the presence of circumbinary gas). These, and several other possible cuts are discussed further in [2].

3. Prompt shocks in the gas disc around a recoiling binary

The recent breakthrough in numerical relativity has allowed a direct computation of the linear momentum flux (‘kick’) produced during the coalescence of a SMBH binary. Such kicks may help produce prompt EM counterparts to GW sources detected by \textit{LISA}. If the SMBHB is surrounded by a circumbinary gas disc, the disc will indeed respond promptly (on the local orbital timescale) to such a kick. If this results in warps or shocks, the disturbed disc could produce a transient EM signature [23]. As discussed above, the final sky localization uncertainty from \textit{LISA} is typically a few tenths of a square degree, containing a large number of sources; monitoring this area for transient events after the merger may be another method to securely identify counterparts.

We investigated the response of a circumbinary disc to the kick [20]. We adopted the following simplified picture for the disc around a fiducial, equal-mass, \( M = M_1 + M_2 = 10^6 \, M_\odot \) binary. The disc has an inner edge at \( 100 \, r_S \) (Schwarzschild radii), inside which it is empty (with the gas evacuated due to torques from the binary; e.g. [23]) and an outer edge at \( 10000 \, r_S \). For reference, the orbital speeds at \( 100 \, r_S \) and \( 10^4 \, r_S \) are \( v_{\text{orb}} = 21 \, 000 \, \text{km s}^{-1} \) and \( v_{\text{orb}} = 2100 \, \text{km s}^{-1} \), respectively; the corresponding radii and orbital timescales for a BH with mass \( M = M_6 \times 10^6 \, M_\odot \) are \( (10^{-5} - 10^{-3}) \, M_6 \, \text{pc} \) and \( (1 - 10^3) \) days. Outside this radius, there may still be gas, but it may be unstable to fragmentation, and it will, in any case, evolve slowly (the behavior of this gas will then not be relevant for a \textit{LISA} counterpart search, but if emission is produced in this gas, it could help identify SMBHBs independently; see discussion in the following section). The scale-height and temperature at the inner edge are \( h/r = 0.46 \) and \( T = 1.7 \times 10^6 \, \text{K} \), respectively. The scale-height remains constant with radius out to \( 2000 \, r_S \), beyond which it increases nearly linearly (\( h \propto r^{21/20} \)). The temperature varies with radius as \( T \propto r^{-9/10} \) (see [20] for further details).

The important features of such a disc (as well as other proposed variants of thin \( \alpha \)-discs) are the following: (i) orbital motions in thin discs are supersonic, so that the gas is susceptible to shocks if disturbed; (ii) at the relevant radii outside \( 100 \, r_S \) the viscous timescale is long, and the orbits are near Keplerian; (iii) gas near the inner edge of disc is tightly bound to the kicked BHB \( (v_{\text{orb}} \sim 2.1 \times 10^4 \, \text{km s}^{-1}) \), but the outer edge \( (v_{\text{orb}} \sim 2.1 \times 10^3 \, \text{km s}^{-1}) \) can be marginally bound, or even unbound, for large kicks (a rough condition for being bound is \( v_{\text{orb}} \gtrsim v_{\text{kick}} \)) and (iv) the total disc mass within \( 10^4 \, r_S \) is much less than the BHB mass, which justifies ignoring the inertia of the gas bound to the BHB.

For a quantitative assessment of the disc’s response to the kick, we employed the following approximation: the disc particles are assumed to be massless, collisionless and initially on
co-planar, circular orbits. The kick simply adds the velocity $-\vec{v}_{\text{kick}}$ to the instantaneous orbital velocity of each particle (in the inertial frame centered on the BHB). We used $N = 10^6$ particles, distributed randomly and uniformly along the two-dimensional surface of the disk. The kick velocity was varied between $500 \text{ km s}^{-1} < v_{\text{kick}} < 4000 \text{ km s}^{-1}$, and directed either perpendicular or parallel to the initial disc plane. Note that in both the perpendicular and parallel case, we start with a two-dimensional particle distribution (i.e. an infinitely thin disc), but in the perpendicular case, we then follow the orbits in 3D.

Figure 3 shows, as an example, a face-on view of the surface density of the disc 90 days after a kick with $v_{\text{kick}} = 500 \text{ km s}^{-1}$ in the plane of the disc (top panel). The sharp, tightly wound spiral features clearly seen in the figure trace the locus of points where particles cross each other, corresponding formally to a density caustic. The spiral caustic first forms at $\sim 30$ days, and then propagates outward at a speed of $\approx 500 \text{ km s}^{-1}$, so that the outermost caustic at time $t_c$ is located at $r_c \sim t_c v_{\text{kick}}$ (this behavior can be roughly understood using epicycles; [20, 24]). The bottom panel in figure 3 shows, for comparison, the aerial view...
of the 3D particle density one week after a kick with the same velocity, but perpendicular to the disk. The density profile in this case remains azimuthally symmetric, but still develops concentric rings of density fluctuations (although we find the density enhancements are much weaker, at the ten percent level).

These results suggest that strong density enhancements can form promptly after a supersonic kick in the plane of the circumbinary disc, within a few weeks of the coalescence of a \( \sim 10^6 \, M_\odot \) BHB. Because the disc is cold, and caustics are formed when particles first cross each other along their orbits, this implies that corresponding shocks could occur in a gas disk. For hydrodynamical shocks to occur within a finite-pressure gas, the relative motions \( v_c \) between the neighbouring particles that produce the caustic must exceed the sound speed. At the outermost radius where the disk is marginally bound to the BHB, one expects \( v_c \sim v_{\text{kick}} \sim v_{\text{orb}} \); relative motions will be slower further inside. The relative speed should roughly correspond to covering the epicyclic amplitude \( \sim (v_{\text{kick}}/v_{\text{orb}}) r_c \) in the caustic-formation time \( t_c \sim r_c/v_{\text{kick}} \), yielding \( v_c \sim v_{\text{kick}}^2/v_{\text{orb}} \). For \( v_{\text{kick}} = 500 \, \text{km s}^{-1} \), this predicts \( v_c \sim 25 \, \text{km s}^{-1} (r/1000r_S)^{1/2} \); we have verified in our simulations that particles cross the caustics at speeds within \( \sim 30\% \) of this predicted value. Compared with the sound speed \( c_s \approx 25 \, \text{km s}^{-1} (r/1000r_S)^{-9/20} \), this suggests that the density waves produced by the kick in the gas beyond \( \sim 700r_S \) will indeed steepen into shocks. We also found that the inclination of the kick may be important in determining the strength and timing of such shocks.

The nature of the emission resulting from the shocks or density enhancements will have to be addressed in future work, by incorporating the effects of hydrodynamics, computing the heating rate at the location of the spiral shocks, and modeling the overall disk structure and vertical radiation transport\(^6\). However, in the fiducial case discussed above, if we assume that the shocked gas is heated to temperatures corresponding to \( v_c \), and \( t_c \) is the timescale on which the corresponding thermal energy is converted to photons, then we find that the luminosity may be a small but non-negligible fraction, 0.2 percent to 5 percent of the Eddington luminosity of the central BH, and would increase with time roughly as \( L_{\text{kick}} \propto t^2 \). This suggests that the afterglows may be detectable, at least for nearby BHs and/or for the most massive BHs in LISA’s range (which extends up to \( M_{\text{BH}} \approx 10^7 \, M_\odot \)). We may also speculate on the spectral evolution of the ‘kick after-glow’, based on the characteristic shock velocity at each radius. We find that the luminosity is dominated by the outermost shocked shells, with the spectrum peaking at the characteristic photon energy corresponding to \( kT_{\text{shock}} \propto v_c^2 \propto v_{\text{orb}}^{-2} \propto r \). The shocks could therefore result in an afterglow, starting from \( 700r_S/v_{\text{kick}} \sim 50 \) days, first peaking in the UV band (\( \sim 3 \, \text{eV} \) or \( \sim 0.3 \, \mu\text{m} \)), corresponding to \( \sim 25 \, \text{km s}^{-1} \), and then hardening to the EUV/soft x-ray (\( \sim 50 \, \text{eV} \) range after \( \sim 2 \) years. The detection of such an afterglow would help identify EM counterparts to LISA events.

4. Searching for gravitational binary inspirals before LISA

Numerical simulations suggest that the central cavity of the circumbinary disc is not completely empty, and that there can be non-negligible gas inflow into this cavity from the disc outside \([15–17]\). The perturbations of the circumbinary gas by the rotating quadrupole potential of the binary make this inflow rate fluctuate, and these fluctuations roughly track the orbital period \([16, 17]\). We must emphasize that the level and nature of EM emission, associated with this inflow (or with other effects from the gas in the vicinity of the binary during the late stages of binary evolution), remain essentially unknown. Furthermore, even in the most

---

\(^6\) Our preliminary work, based on hydrodynamical simulations, does indicate that gas discs with realistic pressure profiles still develop strong shocks \([25]\).
optimistic scenario, in which a clearly periodic emission is indeed produced, at a significant and detectable flux level, there remains the (potentially severe) observational challenge of distinguishing such variability from other possible sources of periodic variations.

Despite these uncertainties and caveats, however, it is interesting to ask the following question: if, indeed, non-negligible emission is produced during the late stages of binary evolution, and the luminosity varies periodically on the orbital timescale, could such periodic sources possibly be identified in EM surveys, even before LISA becomes operational?

As the orbit of each individual binary shrinks, its orbital period, and therefore, by assumption, its variability timescale decreases. The evolution of a binary embedded in a thin circumbinary disc generically proceeds through three distinct stages (see [26]). (i) First, the binary is strongly coupled to the circumbinary disc and is driven by viscosity (analogous to ‘disc-dominated’ planetary migration), and the radius of the gap follows the binary. As the binary separation shrinks below $\sim 10^3 r_S$ ($10^3 r_S$) for mass ratios $q \equiv M_1/M_2 \sim 1$ ($q \sim 0.01$), the binary mass starts to dominate over the local disc density, and the binary evolves more slowly, according to so-called ‘secondary-dominated’ type-II migration. During this stage, the GW emission is negligible. (ii) Later, within the radius $\sim 500 r_S$, the binary starts to be driven primarily by GWs but the inner edge of the gap can still follow the binary decay. In this case, the GW timescale is already less than the secondary dominated type-II migration timescale, but is still larger than the viscous time. (iii) Finally, within $\sim 100 r_S$ the binary is entirely driven by GWs and the binary falls in much more quickly than the inner edge of the gap is able to move inward. The ordering of these events is valid for a broad range of binary and disc parameters. Note that the outcome of the gas inside the binary’s orbit is left unspecified above (see [14] for a possible outcome for $q \ll 1$).

In general, the rate at which the binary’s orbit decays is a function of the orbital separation $r_{\text{orb}}$, with the decay accelerating as the orbital separation shrinks. The residence time $T$ that the binary spends at each $r_{\text{orb}}$ thus gets progressively shorter at smaller $r_{\text{orb}}$ (or equivalently, at shorter orbital time $t_{\text{orb}}$). At large separations, where the evolution is due to viscous processes (stage i), the decay accelerates relatively slowly, and the residence time $T$ has a relatively flat dependence on $t_{\text{var}}$. For near-equal mass binaries, the scaling is found to be between $T \propto t_{\text{var}}^{5/51}$ and $T \propto t_{\text{var}}^{7/12}$, depending on whether the opacity is dominated by free–free absorption or electron scattering, respectively. For $q \ll 1$, the scaling is slightly steeper, between $T \propto t_{\text{orb}}^{5/6}$ and $T \propto t_{\text{orb}}^{14/15}$ for free–free absorption or electron scattering opacity, respectively (see the appendix in [26] for an extensive discussion of these scalings). When GWs begin to contribute significantly to the decay (stage ii) the decay rate accelerates more rapidly, and the residence time becomes a much steeper function of the orbital time. In the last, purely GW-driven regime (stage iii), the scaling should asymptotically reach the power-law $T \propto t_{\text{orb}}^{5/3}$, predicted by the leading-order quadrupole formula for energy losses in GWs during the inspiral [27].

Among sources that have similar BH masses, the abundance $N$ of objects that show near-periodic variability on the timescale $t_{\text{var}} \approx t_{\text{var}}$ is proportional to the residence time $T$, $N \propto T$, where $T$ is a strong function of $t_{\text{var}}$. In particular, the incidence rate of sources in the final GW-driven stage would follow $N \propto t_{\text{var}}^{5/3}$. In the left panel of figure 4, we show that the time spent at each orbital period can be interestingly long, i.e. a non-negligible fraction of the expected lifetime of quasar activity of a few $\times 10^7$ years.

Can we detect the flux variations from these sources? Luminosity variations corresponding to a fraction $f_{\text{flux}}$ of the Eddington luminosity would appear as a periodic variable component, with the amplitude of the variations $\Delta l \approx 26 + 2.5 \log\left((f_{\text{Edd}}/0.01)(M_{\text{BH}}/3 \times 10^7 M_\odot)^{-1}\right)$ magnitudes in the optical (for a BHB at $z = 2$). Detecting this level of fluctuation requires a long integration, but is feasible.
We can also estimate the expected number of such periodically variable sources, by assuming that there is a one-to-one correspondence between bright quasars and these sources. More specifically, we make the following assumptions: (i) every quasar is directly triggered by a SMBH–SMBH merger; (ii) as the binary orbit decays, goes through the evolutionary stages discussed above (see [26] for more details); (iii) for the last \( t_Q \approx 10^7 \) years of this evolution, the binary maintains a constant steady luminosity \( \bar{L}_Q \), with roughly periodic fluctuations of amplitude \( \Delta L_Q \) and period \( t_{\text{var}} = t_{\text{orb}} \) about this steady mean luminosity. Furthermore, we assume (iv) that the source is bright (\( \bar{L}_Q = f_{\text{Edd}} L_{\text{Edd}} \), with a fiducial \( f_{\text{Edd}} = 0.3 \)) and produces non-negligible fraction variations (\( \Delta L_Q = f_{\text{var}} f_{\text{Edd}} L_Q \) with a fiducial \( f_{\text{var}} = 0.1 \)). The total
number $N_Q(L_Q)$ of quasars (i.e., the luminosity function; LF) is well known from observations. Here we use a recently compiled bolometric quasar LF [28] and assume a width $\Delta z = 1$ when computing the total number $N_Q$ of quasars above the survey detection threshold. Under the above assumptions, the number of periodic variable sources with $t_{\text{var}} = t_{\text{orb}}$, $N_{\text{var}}(L_Q, t_{\text{var}})$ is given simply by the fraction of all quasars residing at a given $t_{\text{orb}}$: $N_{\text{var}}/N_Q = T/t_Q$.

Could a survey find a sufficiently large number of such periodic variable sources to identify them as a population? For illustration, imagine a survey with a sensitivity that corresponds to detecting the periodic variability of BHBs with a mass $M_{\text{min}}$ at $z = 2$, covering a solid angle $\Delta \Omega$. (A real survey, of course, will have a completeness for variability detection that is not a step function.) If the survey volume contains a total of $N_Q$ BHBs with the luminosity $L_Q$, then the periodic variable fraction, $T/t_Q$, can be determined down to the smallest value $\approx N_{\text{var}}^{-1}$ (i.e. to find at least one periodic source). At fixed values of $f_{\text{var}}$, $f_{\text{edd}}$ and $t_Q$ (as well as $z$ and $q$), this corresponds to a minimum variability timescale $t_{\text{orb.min}}$ that can be probed. Let us define the requirement that this minimum is $t_{\text{orb.min}} \leq 20$ weeks. Assuming that the longest variability timescale of interest is around $t_{\text{orb.max}} \sim 1$ year (so that the periodic nature of the variations can be convincingly demonstrated over a multi-year survey), this choice will offer a factor of three range in $t_{\text{var}}$ for mapping out the $N_{\text{var}}$ versus $t_{\text{var}}$ dependence (i.e., with the pure GW-driven scaling $N_{\text{var}} \propto t_{\text{var}}^{-8/3}$, the same survey volume would contain $\approx 20$ sources with a similar luminosity but with a $t_{\text{var}} = 60$ week period).

In the right panel of figure 4, we show the required depth and area coverage for a multi-year survey to detect such periodic sources, achieving this factor of three range in period. While no existing optical variability survey has the required duration, sampling rate, depth and area coverage, many come close. Several planned optical surveys (e.g., those looking for distant supernovae, such as LSST) do fall well above the curves shown in this figure.

Our main, admittedly speculative, conclusion is that the dependence in the abundance of these sources on their period—in particular, a switch from a flat, viscosity-driven power-law between $N \propto t_{\text{var}}^{1/2}$ and $N \propto t_{\text{var}}^{8/3}$—below a characteristic $t_{\text{var}} \sim$ tens of weeks (corresponding to $\sim 10^3$ Schwarzschild radii for equal-mass $\sim 10^6 M_\odot$ binaries), could be uncovered in a future survey; therefore, such a survey should be designed and carried out. If a population of periodic variable sources were indeed identified, it would (1) confirm that the orbital decay for sources below a characteristic $t_{\text{var}}$ is indeed driven by GWs; (ii) that circumbinary gas is present at small orbital radii and is being perturbed by the BHs—serving as a proof of concept that electromagnetic counterparts could be found for actual LISA sources; and (iii) the actual number of periodic sources could be used to yield a robust, empirically calibrated, prediction for the LISA event rates.

Finally, we mention two other possibilities to identify coalescing SMBHBs before LISA’s launch. First, [29] and [24] followed the response of the gas disc around a recoiling SMBHB, similar to our calculation described in section 3, but on much longer timescales ($\sim 10^4$ years). Shields and Bonning [29] found that the shocked gas, thrown out of the disc plane by an oblique kick, may produce bright flares in x-ray lines (assuming it remains optically thin). Schnittman and Krolik [24], on the other hand, assumed that the disc gas is optically thick, and that the shocks are promptly dissipated in the disc plane, and derived a characteristic infrared light curve. The above effects arise from the change in the gravitational potential, following the recoil, i.e. from the ‘shaking’ of the disc. A different possibility is that once the BHs have merged, the torques from the rotating quadrupolar potential cease to act on the gas outside the inner cavity, allowing the gas to accrete onto the merged binary on the viscous timescale [23]. An x-ray ‘afterglow’ may then occur after a delay of $\sim (1 + z)(M/10^6 M_\odot)^{1.32}$ years. This could be detectable in LISA follow-up observations, or perhaps without a LISA trigger for very massive BHs at high $z$, where the timescale is long.
Acknowledgments

ZH thanks the organizers of the conference. The work described here was supported by NASA (grant NNX08AH35G), by the Polányi Program of the Hungarian National Office for Research and Technology (NKTH), and by the Országos Tudományos Kutatási Alap (OTKA) in Hungary through Grant 68228.

References

[1] Deffayet C and Menou K 2007 Astrophys. J. 668 L143
[2] Kocsis B, Haiman Z and Menou K 2008 Astrophys. J. 684 870
[3] Schutz B F 1986 Nature 323 310
[4] Holz D E and Hughes S A 2005 Astrophys. J. 629 15
[5] Kocsis B, Frei Z, Haiman Z and Menou K 2006 Astrophys. J. 637 27
[6] Lang R N and Hughes S A 2006 Phys. Rev. D 74 122001
[7] Vecchio A 2004 Phys. Rev. D 70 042001
[8] Kocsis B, Haiman Z, Menou K and Frei Z 2007 Phys. Rev. D 76 022003
[9] Lang R N and Hughes S A 2008 Astrophys. J. 677 1184
[10] Lang R N and Hughes S A 2009 Class. Quantum Grav. 26 094035
[11] Trias M and Sintes A M 2008 Phys. Rev. D 77 024030
[12] Arun K G et al 2009 Class. Quantum Grav. 26 094021
[13] Artyomowicz P and Lubow S H 1994 Astrophys. J. 421 651
[14] Armitage P J and Natarajan P 2002 Astrophys. J. 567 L9
[15] Artyomowicz P and Lubow S H 1996 Astrophys. J. Lett. 467 L77
[16] MacFadyen A and Milosavljević M 2008 Astrophys. J. 672 83
[17] Cuadra J, Armitage P J, Alexander R D and Begelman M C 2008 Mon. Not. R. Astron. Soc. at press (arXiv:0809.0311)
[18] Kocsis B and Loeb A 2008 Phys. Rev. Lett. 101 041101
[19] Campanelli M et al 2007 Phys. Rev. Lett. 98 231102
[20] Lippai Z, Frei Z and Haiman Z 2008 Astrophys. J. Lett. 676 L5
[21] Madau P 2000 Cosmic Star Formation History and the Brightness of the Night Sky From Extrasolar Planets to Cosmology: The VLT Opening Symposium ed J Bergeron and A Renzini (Berlin: Springer) p 52 (arXiv:astro-ph/9907268)
[22] Melia F and Falcke H 2001 Ann. Rev. Astron. Astrophys. 39 309
[23] Milosavljević M and Phinney E S 2005 Astrophys. J. 622 L93
[24] Schnittman J D and Krolik J H 2008 Astrophys. J. 684 835
[25] Corrales L, MacFadyen A and Haiman Z 2009 in preparation
[26] Haiman Z, Kocsis B and Menou K Astrophys. J. submitted (arXiv:0807.4697)
[27] Peters P C 1964 Phys. Rev. D 136 1224
[28] Hopkins P F, Richards G T and Hernquist L 2007 Astrophys. J. 654 731
[29] Shields G A and Bonning E W 2008 Astrophys. J. 682 758