Electromagnetic counterparts to black hole mergers

Jeremy D Schnittman
Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
and
NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
Received 15 October 2010, in final form 19 January 2011
Published 18 April 2011
Online at stacks.iop.org/CQG/28/094021

Abstract
During the final moments of a binary black hole (BH) merger, the
gravitational wave (GW) luminosity of the system is greater than the combined
electromagnetic (EM) output of the entire observable universe. However,
the extremely weak coupling between GWs and ordinary matter makes these
waves very difficult to detect directly. Fortunately, the inspiraling BH system
will interact strongly—on a purely Newtonian level—with any surrounding
material in the host galaxy, and this matter can in turn produce unique EM
signals detectable at Earth. By identifying EM counterparts to GW sources,
we will be able to study the host environments of the merging BHs, in turn
greatly expanding the scientific yield of a mission like LISA. Here we present a
comprehensive review of the recent literature on the subject of EM counterparts,
as well as a discussion of the theoretical and observational advances required
to fully realize the scientific potential of the field.

PACS numbers: 95.30.Sf, 98.54.Cm, 98.62.Js, 04.30.Tv, 04.80.Nn

1. Introduction
Prompted by recent advances in numerical relativity (NR), there has been an increased interest
in the astrophysical implications of black hole (BH) mergers (see [1] for a sample of related
White Papers submitted to the recent Astro2010 Decadal Report). Of particular interest is
the possibility of a distinct, luminous electromagnetic (EM) counterpart to a gravitational
wave (GW) signal. If such an EM counterpart could be identified with a LISA\(^1\) detection
of a supermassive BH (SMBH) binary in the merging process, then the host galaxy could
likely be determined. For BHs with masses of \(10^6 M_\odot\) at a redshift of \(z = 1\), LISA should
be able to identify the location of the source within \(\sim 10\) deg\(^2\) a month before merger, and
better than \(\sim 0.1\) deg\(^2\) with the entire waveform, including merger and ringdown [2–8]. Like
the cosmological beacons of gamma-ray bursts and quasars, merging BHs can teach us about

\(^1\)http://lisa.nasa.gov

0264-9381/11/094021+10$33.00 © 2011 IOP Publishing Ltd Printed in the UK & the USA
relativity, high-energy astrophysics, radiation hydrodynamics, dark energy, galaxy formation and evolution, and even dark matter. A large variety of potential EM signatures have recently been proposed, almost all of which require some significant amount of gas in the near vicinity of the merging BHs. In this paper, we review the recent literature on EM signatures, and propose a rough outline of the future work, both observational and theoretical, that will be needed to fully realize the potential of GW astronomy.

2. Diversity of sources

From a theoretical point of view, EM signatures can be categorized by the physical mechanism responsible for the emission, namely stars, hot diffuse gas, or circumbinary/accretion disks. In figure 1, we show the diversity of these sources, arranged according to the spatial and time scales on which they occur.

It is important to note that, while the BHs themselves are of course extremely relativistic objects, most of the observable effects occur on distance and time scales that are solidly in the Newtonian regime. While one of the most interesting NR results in recent years has been the prediction of large recoil velocities originating from the final merger and ringdown of binary BHs [9], the astrophysical implications of these large kicks are for the most part entirely Newtonian.
2.1. Stellar signatures

On the largest scales, we have strong circumstantial evidence of SMBH mergers at the centers of merging galaxies. From large optical surveys of interacting galaxies out to redshifts of $z \sim 1$, we can infer that 5–10% of massive galaxies are merging at any given time, and the majority of galaxies with $M_{\text{gal}} \geq 10^{10} M_{\odot}$ have experienced a major merger in the past 3 Gyr [10–13], with even higher merger rates at redshifts $z \sim 1–3$ [14]. At the same time, high-resolution observations of nearby galactic nuclei find that every large galaxy hosts a SMBH in its center [15]. Yet we see a remarkably small number of dual active galactic nuclei (AGN) [16–18], and only one known source with an actual binary system where the BHs are gravitationally bound to each other [19]. Taken together, these observations strongly suggest that when galaxies merge, the merger of their central SMBHs inevitably follows, and likely occurs on a relatively short time scale, which would explain the apparent scarcity of binary BHs. The famous ‘M-sigma’ relationship between the SMBH mass and the velocity dispersion of the surrounding bulge also points to a merger-driven history over a wide range of BH masses and galaxy types [20].

There is also indirect evidence for SMBH mergers in the stellar distributions of galactic nuclei, with many elliptical galaxies showing light deficits (cores), which correlate strongly with the central BH mass [21]. The cores are evidence of a history of binary BHs that scour out the nuclear stars via three-body scattering [22–25], or even post-merger relaxation of recoiling BHs [26–29].

While essentially all massive nearby galaxies appear to host central SMBHs, it is quite possible that this is not the case at larger redshifts and smaller masses, where major mergers could lead to the complete ejection of the final BH via large GW recoils. By measuring the occupation fraction of BHs in distant galaxies, one could infer merger rates and the distribution of kick velocities [30–34]. The occupation fraction will of course also affect the LISA event rates, especially at high redshift [35]. An indirect signature for kicked BHs could potentially show up in the statistical properties of active galaxies, in particular in the relative distribution of different classes of AGN in the ‘unified model’ paradigm [36, 37]. On a smaller scale, the presence of intermediate-mass BHs in globular clusters also gives indirect evidence of their merger history [38].

Another EM signature of BH mergers comes from the population of stars that remain bound to a recoiling BH that gets ejected from a galactic nucleus [39–41]. These stellar systems will appear similar to globular clusters, yet with smaller spatial extent and much larger velocity dispersions, as the potential is completely dominated by the central SMBH.

2.2. Gas signatures: accretion disks

Gas in the form of accretion disks around single massive BHs is known to produce some of the most luminous objects in the Universe. However, very little is known about the behavior of accretion disks around two BHs, particularly at late times in their inspiral evolution. In Newtonian systems, it is believed that a circumbinary accretion disk will have a central gap of much lower density, either preventing accretion altogether, or at least decreasing it significantly [42–44]. When including the evolution of the binary due to GW losses, the BHs may also decouple from the disk at the point when the GW inspiral time becomes shorter than the gaseous inflow time at the inner edge of the disk [45]. This decoupling should effectively stop accretion onto the central object until the gap can be filled on an inflow timescale. However, other semi-analytic calculations predict an enhancement of accretion power as the evolving
binary squeezes the gas around the primary BH, leading to a rapid increase in luminosity shortly before merger [46, 47].

Regardless of how the gas can or cannot reach the central BH region, a number of recent papers have shown that if there is sufficient gas present, then an observable EM signal is likely. Krolik [48] used analytic arguments to estimate a peak luminosity comparable to that of the Eddington limit, independent of the detailed mechanisms for shocking and heating the gas. Using relativistic magneto-hydrodynamic simulations in 2D, O'Neill et al [49] showed that the prompt mass loss due to GWs may actually lead to a sudden decrease in luminosity following the merger, as the gas in the inner disk temporarily has too much energy and angular momentum to accrete efficiently. Full NR simulations of the final few orbits of a merging BH binary have now been carried out, including the presence of EM fields in a vacuum [50–52] and also gas, treated as test particles in [53] and as an ideal fluid in [54] and [55]. The simulations including matter all suggest that the gas can get shocked and heated to high temperatures, thus leading to bright counterparts in the event that sufficient gas is in fact present in the immediate vicinity of the merging BHs.

If the primary energy source for heating the gas is gravitational, then typical efficiencies will be on the order of ~1–10%, comparable to that expected for standard accretion in AGN. However, if the merging BH binary is able to generate strong magnetic fields [50–52], then highly relativistic jets may be launched along the resulting BH spin axis, converting matter to energy with a Lorentz boost factor of \( \gamma \gg 1 \). Even with purely hydrodynamic heating, particularly bright and long-lasting afterglows may be produced in the case of very large recoil velocities, which effectively can disrupt the entire disk, leading to strong shocks and dissipation [56–64]. For systems that open up a gap in the circumbinary disk, an EM signature may take the form of a quasar suddenly turning on as the gas refills the gap, months to years after the BH merger [45, 65, 66].

For those systems that also received a large kick at the time of merger, we may observe quasar activity for millions of years after, with the source displaced from the galactic center, either spatially [67–72] or spectroscopically [73–76]. However, large offsets between the redshifts of quasar emission lines and their host galaxies have also been interpreted as evidence of pre-merger binary BHs [77–80] or as being due to the large relative velocities in merging galaxies [81–84], or 'simply' as extreme examples of the class of double-peaked emitters, where the line offsets are generally attributed to the disk [85–89].

In addition to the many potential prompt and afterglow signals from merging BHs, there has also been a significant amount of theoretical and observational work focusing on the early precursors of mergers. Following the evolutionary trail from the upper-left part of figure 1, we see that shortly after a galaxy merges, dual AGN may form with typical separations of a few kpc [16, 17], sinking to the center of the merged galaxy on a relatively short timescale (≤1 Gyr) due to dynamical friction [90]. This merger process is also expected to funnel a great deal of gas to the galactic center, in turn triggering quasar activity [91–94]. At separations of \(~1\) pc, the BH binary (now 'hardened' into a gravitationally bound system) could stall, having depleted its loss cone of stellar scattering and not yet reached the point of gravitational radiation losses [95]. Gas dynamical drag from massive disks \( (M_{\text{disk}} \gg M_{\text{BH}}) \) leads to a prompt inspiral \((~1–10 \text{ Myr})\), in most cases able to reach sub-parsec separations, depending on the resolution of the simulation [96–102].

At this point, a proper binary quasar is formed, with an orbital period of months to decades, which could be identified by periodic accretion [103–106] or redshifted broad emission lines as mentioned above [107–109]. Direct GW stresses on the circumbinary disk might also lead to periodic variations in the light curve, although with very small amplitude [110].
2.3. Gas signatures: diffuse gas, ‘other’

In addition to the many disk-related signatures, there are also a number of potential EM counterparts that are caused by the accretion of diffuse gas in the galaxy. For BHs that get significant kicks at the time of merger, we expect to see quasi-periodic episodes of Bondi accretion as the BH oscillates through the gravitational potential of the galaxy over millions of years, as well as off-center AGN activity [111–114]. On larger spatial scales, the recoiling BH could also produce trails of overdensity in the hot interstellar gas of elliptical galaxies [115]. In a similar way, rogue SMBHs in gas-rich galaxies could leave trails of star formation in their wake [116]. It is even possible that the same density enhancements could be detected via off-nucleus gamma-ray emission from annihilating dark matter particles [117]. Also on kpc–Mpc scales, X-shaped radio jets have been seen in a number of galaxies, which could possibly be due to the merger and subsequent spin flip of the central BHs [118].

Another potential source of EM counterparts comes not from diffuse gas, or accretion disks, but from the occasional capture and tidal disruption of normal star> by the merging BHs. This tidal disruption, which also occurs in ‘normal’ galaxies [119–121], may be particularly easy to identify in off-center BHs following a large recoil [39]. Tidal disruption rates may be strongly increased by the merger process itself [122–125], while the actual disruption signal may be truncated by the pre-merger binary [126]. These events are likely to be seen by the dozen in coming years with PanSTARRs and LSST [127]. In addition to the tidal disruption scenario, in [125] we showed how gas or stars trapped at the stable Lagrange points in a BH binary could evolve during inspiral and eventually lead to enhanced star formation, ejected hyper-velocity stars, highly-shifted narrow emission lines and short bursts of Eddington-level accretion coincident with the BH merger.

A completely different type of EM counterpart can be seen in the radio. Namely, nanosecond time delays in the arrival of pulses from millisecond radio pulsars is direct evidence of extremely low-frequency (nano-Hertz) GWs from massive ($\gtrsim 10^6 M_\odot$) BH binaries [128–135]. By cross-correlating the signals from multiple pulsars around the sky, we can effectively make use of a GW detector the size of the entire Milky Way galaxy.

3. Game plan

In the coming years, a number of theoretical and observational advances will be required in order to fully realize the potential of GW/EM multi-messenger astronomy. Some of the central questions that need to be answered include:

* What is the galaxy merger rate as a function of galaxy mass, mass ratio, gas fraction, cluster environment, and redshift?
* What is the mass function and spin distribution of the central BHs in these merging (and non-merging) galaxies?
* What is the central environment around the BHs, prior to merger?
  - What is the quantity and quality (temperature, density, composition) of gas?
  - What is the stellar distribution (age, mass function, metallicity)?
  - What are the properties of the circumbinary disk?
* What is the time delay between galaxy merger and BH merger?

We have rough predictions for some of these questions from cosmological $N$-body simulations, but the uncertainties and model dependences are quite large [136–138]; for a comprehensive comparison of the leading methods, see [139]. Similarly, observational constraints on the merger rates are relatively weak and often open to widely varying interpretations [11–13, 17, 140–142].
3.1. Theory

With respect to the questions outlined above, improved cosmological simulations will certainly help improve our estimates for galactic and BH merger rates, as well as the gas environments expected in the central regions. Particularly promising are multi-scale simulations that can zoom in on regions of interest, going to higher resolution and more realistic physics closer to the BHs [137]. To model more accurately the interaction between the circumbinary disk and the BHs, grid-based methods (as opposed to smoothed particle hydrodynamics, SPH) will be necessary, especially at the inner edge where steep density and pressure gradients are likely to be found. The accurate treatment of this region is critical to understanding the gas environment immediately around the BHs at the time of merger, and thus whether any bright EM signal is likely to be produced.

The natural product of these (Newtonian) circumbinary magneto-hydrodynamic (MHD) simulations would be a set of reasonable initial conditions to be fed into the much more computationally intensive NR codes that compute the final orbits and merger of the BHs, now including matter and magnetic fields. The results of [50--55] are extremely impressive from a computational point of view, but their astrophysical relevance is limited by our complete ignorance of the likely initial conditions. Even with perfect knowledge of the initial conditions, the value of the MHD simulations is also limited by the lack of radiation transport and accurate thermodynamics, which are only now being incorporated into local Newtonian simulations of steady-state accretion disks [143]. Significant future work will be required to incorporate the radiation transport into a fully relativistic global framework, required not just for accurate modeling of the dynamics, but also for the prediction of EM signatures that might be compared directly with observations.

3.2. Observations

Even with the launch of LISA a decade or more away, many of the EM counterparts discussed above should be observable today, in some cases even giving unambiguous evidence for merging BHs. On the largest distance and time scales, dual AGN candidates can be identified with large spectroscopic surveys like SDSS\(^3\), then followed up with high-resolution imaging and spectroscopy. This has already been done to successfully estimate the fraction of merging galaxies, but to get a merger rate, more information is needed about the expected lifetimes of dual AGN [13, 17, 142]. Combined with surveys of galaxy morphology and pairs, the distribution of dual AGN will help us test theories of galactic merger rates as a function of mass and redshift, as well as the connection between gas-rich mergers and AGN activity. Spectroscopic surveys should also be able to identify many candidate binary AGN, which may be confirmed or ruled out with subsequent observations over relatively short timescales (~1–10 yrs), as the line-of-site velocities to the BHs change by an observable degree. Here, a number of candidates have already been identified, but as yet, none have been confirmed to be binaries by multi-year monitoring [75, 88, 144].

Many of these spectroscopic surveys have also been used to look for recoiling BHs [73, 74, 76], where the line redshifts are not expected to change in time, but they may be confirmed or ruled out as high-recoil candidates via high-resolution imaging of the host galactic nucleus. Long-lived afterglows could be discovered in existing multi-wavelength surveys, but successfully identifying them as merger remnants as opposed to obscured AGN or other bright unresolved sources would require improved pipeline analysis of literally millions of point sources, as well as extensive follow-up observations.

\(^3\) http://www.sdss.org
Particularly promising as unambiguous examples of recoiling BHs would be the measurement of large velocity dispersions in nearby ($d \lesssim 20$ Mpc) globular clusters [40]. With multi-object spectrometers on large ground-based telescopes, this is also technically realistic in the immediate future, though a directed campaign has not been carried out yet. Perhaps the most exciting direction for the coming decade of astronomy is in the time domain. Optical telescopes like PTF and PanSTARRS are already taking data from huge areas of the sky with daily and even hourly frequency. These time-domain surveys are ideally suited for looking for variability from binary BH systems as precursors to merger. Especially promising would be the detection of long-period variable AGN, ideally suited to extensive multi-wavelength follow-up observations.

Acknowledgments

This work was supported in part by the Chandra Postdoctoral Fellowship Program. We acknowledge helpful conversations with Joan Centrella, John Kormendy, Julian Krolik, David Merritt, Cole Miller and Greg Shields.

© 2011 US Government Copyright.

References

Nandra K 2009 arXiv:0903.0547
Schutz B F et al 2009 arXiv:0903.100

[6] Lang R N and Hughes S A 2009 Class. Quantum Grav. 26 094035
Tichy W and Marrone P 2007 Phys. Rev. D 76 061502
[89] Gaskell M C 2010 Nature 463 E1
[105] Haiman Z, Kocsis B, Menci K, Lipari Z and Frei Z 2009 Class. Quantum Grav. 26 094032