Maximum redshift of gravitational wave merger events

Savvas M. Koushiappas 

Department of Physics, Brown University, 182 Hope St., Providence, Rhode Island 02912, USA and
Institute for Theory and Computation, Harvard University, 60 Garden Street, Cambridge, Massachusetts, 02138, USA

Astronomy Department, Harvard University, 60 Garden Street, Cambridge, Massachusetts, 02138, USA

(Dated: November 1, 2017)

Future generation of gravitational wave detectors will have the sensitivity to detect gravitational wave events at redshifts far beyond any detectable electromagnetic sources. We show that if the observed event rate is greater than one event per year at redshifts \( z \geq 40 \), then the probability distribution of primordial density fluctuations must be significantly non-Gaussian or the events originate from primordial black holes. The nature of the excess events can be determined from the redshift distribution of the merger rate.

PACS numbers: 95.35.+d, 98.80.-k, 04.30.Db, 04.30.?w

The discovery of gravitational waves from merging pairs of massive black holes \([1–4]\) has opened a new window to the astrophysics of black holes, their formation, and cosmic evolution. Black holes of stellar masses have been observed with LIGO \([1–4]\) and supermassive black holes in galaxies are expected to be detected by LISA over the next couple of decades \([5–7]\). The sensitivity of the next generation of ground-based gravitational wave detectors is expected to improve by at least an order of magnitude \([8]\), thus allowing the detection of merging black holes events out to the highest redshifts, potentially exceeding the reach of electromagnetic observations which respond to amplitude squared and not amplitude.

The expected rates of black hole mergers has been calculated based on the number and properties of the few events discovered to-date (see, e.g. \([9–11]\)). The rate depends on a multitude of factors: black holes must be formed and they must find a way to get close enough so that gravitational waves can take-over as the dominant energy loss mechanism. The redshift distribution encodes information about the origin of black hole pairs. If black holes originate from massive stellar progenitors then the redshift distribution should relate to the formation, accretion, and cooling of gas in galaxies. If on the other hand the black holes are primordial \([12–16]\), then the redshift distribution will extend to earlier cosmic times due to primordial binaries \([17]\).

A key difference between these two scenarios is that in the case of a baryonic origin, black holes must form out of cold gas, which accreted into a dark matter gravitational potential well, and then cooled to form black hole progenitors. This path follows the abundance of appropriate potential wells.

In this letter we calculate the maximum redshift of expected black hole merger events that have baryonic origin in the standard cosmological model. That is, the black holes are formed in galaxies as opposed to primordial black holes, or black holes that are formed in non-standard cosmological scenarios, e.g., cosmologies with a significant non-Gaussianity in the primordial density fluctuations of the dark matter.

The significance of this calculation is two-fold: first, it defines a maximum redshift over which baryonic structures can form, and second any detection above the derived bound will signify the presence of either non-Gaussianities that control the formation of baryonic structures at unexpectedly high redshifts, or that black hole events may be due to primordial black holes.

In the following we make two key assumptions in the derivation of a maximum redshift of baryonic black hole gravitational wave events. First, we conservatively assume that black hole pairs merge instantaneously, i.e., there is no time lag between the formation of black holes, the evolution of the binary and the subsequent sequence of events that leads to a merger. Second, we conservatively assume that all gas accreted in dark matter halos end up in stars that end up in black holes. Realistically, both of these assumptions are vastly optimistic. However, their application guarantees that the derived maximum redshift is indeed a very hard limit and thus any observation that violates this bound will be of enormous scientific significance.

We begin by calculating the number of observed gravitational wave events per year \( > z \) as the integral of the rate of black hole mergers per redshift interval

\[
N(> z) = \int_z^\infty \frac{dR}{dz} \, dz, \tag{1}
\]

where \( dR/dz \) is the rate of merger events per redshift.
Here, \(dN/dM dV\) is the comoving density of dark matter halos of mass \(M\) at redshift \(z\), \(C_{NG}(M, z)\) is a correction to the mass function in the case where non-Gaussianity is present (with \(C_{NG}(M, z) = 1\) in the standard ΛCDM cosmology), \(\dot{M}_g(M, z)\) is the rate of accreted gas in halos of mass \(M\) at \(z\), \((\epsilon(M, z))\) is the efficiency of converting gas to black holes of mass \(m_{\text{BH}}\), \(dV/dz\) is the comoving volume per redshift interval and the \((1 + z)\) factor in the denominator is to convert the rest frame rate to the observed rate.

The integral in Equation (2) is performed from a minimum halo mass \(M_{\text{min}}(z)\) to infinity. Throughout the paper we use a cosmological model with a power spectrum with a spectral index \(n_s = 0.967\), a normalization \(\sigma_8 = 0.81\), a present value of the Hubble parameter \(H_0 = 70.4\) km/s/Mpc and dark matter, baryonic and cosmological constant mass density parameters of \(\Omega_{DM} = 0.226\), \(\Omega_b = 0.0455\), and \(\Omega_{\Lambda} = 0.728\), respectively [18].

We next explain how we calculate each of these quantities. The expected number of gravitational wave events depends strongly on the halo mass function which declines exponentially at high redshifts for Gaussian fluctuations. Figure 1 shows the mass function at high redshifts from 15 different numerical simulations [19,31]. The only results that are valid at the high redshifts we consider here are the ones of [20,26,29]. The halo masses of interest at these high redshifts correspond to extremely rare peaks. The abundance of halos is roughly bounded by two functional forms – the analytic form of Press-Schecter [31] (red curve in Figure 1) gives the lowest number of halos while the ellipsoidal collapse model of Sheth, Mo & Tormen [30] gives the maximum (blue curve in Figure 1). All other mass functions, including the more realistic results in [25,26,29] lie in between these two analytic forms.

The presence of a non-Gaussianity in the initial conditions can alter the abundance of dark matter halos, especially in the exponential tail of the mass function (which is the regime of interest here). We therefore modify the mass function to include such features by assuming that the non-Gaussian mass function is the product of the Gaussian mass function multiplied by a correction factor [32] that describes \(f_{\text{NL}}\) cosmologies [33,37] (though it is important to emphasize that \(f_{\text{NL}}\) is just one possible parametrization of non-Gaussianities),

\[
C_{NG}(M) = \left[ \frac{\delta_c^2}{6\Delta d\ln\sigma(M)} + \Delta \right] \exp\left( \frac{S_3 \delta_c^3}{6\sigma^2(M)} \right).
\]

Here, \(\Delta \equiv \sqrt{1 - \delta_c S_3/3}\), \(\delta_c = 1.686\sqrt{a}\), with \(a = 0.9\), and \(S_3 = 3.15 \times 10^{-4} f_{\text{NL}}/\sigma^{0.88}(M)\). The Sheth-Mo-Tormen mass function [30], modified to include the effects of \(f_{\text{NL}}\) non-Gaussianities [33] is shown in Figure 1 as the blue dashed line. We consider this modified mass function to represent the maximum abundance of dark matter halos (repeating the calculations for cosmologies with \(g_{\text{NL}}\) or \(\tau_{\text{NL}}\) within the current limits [35] leads to smaller effects than the effects from the current uncertainties in \(f_{\text{NL}}\)).

It is also important to note that the mass function depends on the normalization of the power spectrum; however, the current percent-level uncertainty of \(\sigma_8\) is negligible for our purposes.

The quantity \(\dot{M}_g(M, z)\) represents the rate of gas inflow in halos of mass \(M\) at \(z\). It has been predicted in simple theoretical grounds [39] and has been measured in hydrodynamical simulations at high redshift [39,40]. We adopt the maximum gas accretion rate [40],

\[
\dot{M}_g(M, z) \approx 10^{-3} M_{\odot} \text{yr}^{-1} \left( \frac{M}{10^7 M_{\odot}} \right)^{1.127} \left( 1 + z \right)^{\eta}.
\]

where \(\eta = 2.5\). Assuming a gas accretion rate as given in [39] results in a rate that is only slightly smaller (see Figure 2). If on the other hand we assume that the redshift dependence is steeper (i.e., \(\eta > 2.5\)) as suggested by high-redshift studies of the growth rate of halos [41,42], the gas accretion rate can be higher, however the falloff at high redshift becomes much steeper. Both of these
assumptions have negligible effects to the scope of this paper.

The quantity $\langle \epsilon(M,z) \rangle$ represents the fraction of infalling gas that turns into black holes. In order to compute an extremely conservative upper bound on the number of black holes, we assume that all stars formed out of infalling gas will be converted to black holes that will merge within the Hubble time at $z$. We assume that this function is bracketed from above by the ratio of stellar mass to baryon mass in dark matter halos, i.e., $\langle \epsilon(M,z) \rangle = M_{\text{stellar}}/M_\eta$, where $\eta = \Omega_b/\Omega_{DM}$, is the baryon fraction, and we assume the stellar mass as a function of host halo mass and redshift is given by extrapolating (beyond $z \approx 8$) the results of [20].

In Figure 2, we show the logarithm of the product $\langle \epsilon(M,z) \rangle M_g$ in units of $M_\odot \text{yr}^{-1}$, from $10^{-1} - 10^{-15}$ (in declining factors of ten from left to right). Thin lines correspond to the analytic prediction of [43] while thick lines are the numerical results of [40]. The contours show that at redshifts $z \leq 30$ the gas infall rate increases with mass, and that for a fixed mass it decreases with increasing redshift at $z \approx 30$.

At each redshift $z$, we integrate Equation (2) from $M_{\text{min}}(z)$ to infinity. The lower mass limit of the integral, $M_{\text{min}}(z)$, is the minimum halo mass in which stars can form. This is set by the requirement of the formation of molecular hydrogen [45, 46]. Recently, it has been argued that the tight coupling of baryons to photons prior to recombination gives rise to a velocity component that becomes important once baryons decouple [47]. Figure 2 shows this effect on the minimum mass: the thin solid black curve is the standard case where baryons are assumed to follow dark matter, while the thick solid black curve corresponds to the numerical results of [44] where there is a velocity difference between dark matter and baryons.

Since the minimum mass of molecular hydrogen cooling is roughly constant with redshift, star formation is severely suppressed at increasing redshifts for two reasons: the minimum mass corresponds to extremely rare peaks in the density field (see dashed grey lines in Figure 2 that show the rarity of mass scales as a function of redshift) while at the same time the rate of gas accretion decreases rapidly. The combination of these two effects introduces a sharp cutoff to the abundance of stars beyond $z \approx 40$.

Integrating the rate of merger events (Equation 2) from redshift $z$ to infinity gives the total number of events per year greater than redshift $z$ (Equation 1). Figure 3 shows the result of this calculation. The blue curve corresponds to the maximal mass function [33], a lower $M_{\text{min}}$ (i.e., ignoring the suppressing effects of a relative speed between dark matter and baryons [44]), and the maximal value of gas accretion [10]. The dashed blue curve makes the same assumptions as above, but with a modified mass function that includes a correction owing to the presence of non-Gaussianity at the current upper bound of $f_{\text{NL}} = 43$ [33]. The red curve is the opposite of the aforementioned case, where the mass function assumed is at its minimum [31], the minimum mass is the largest (including relative velocities between baryons and dark matter [44]) and a low gas accretion rate [39]. The shaded area represents everything in between these two extreme cases.

We define the maximum redshift $z_{\text{max}}$ such that the observed event rate is $N(z = z_{\text{max}}) = 1 \text{yr}^{-1}$. We find that the maximum redshift of expected gravitational wave events cannot exceed $z_{\text{max}} \approx 40$. All assumptions leading to this result are such so that the maximum redshift is maximized: largest abundance of halos (even including current limits on non-Gaussianity), lowest minimum mass for the formation of stars in halos at high redshifts, the assumed gas infall in halos is the maximum measured in numerical simulations, all stars formed in all halos end up in black hole pairs and all black hole pairs merge instantaneously. This confluence of maximizing all assumptions makes the result that the maximum redshift of expected gravitational wave sources of $z_{\text{max}} \approx 40$
The number of gravitational wave events of $m_{\text{BH}} = 30M_\odot$ black hole pairs originating from redshifts greater than $z$ (Equation 1) as a function of redshift. The blue curve corresponds to the upper limit on the halo mass function [30], a low value of $M_{\text{min}}$ (i.e., ignoring the effects of a relative speed between dark matter and baryons [44]) and a high value of gas accretion [10]. The dashed blue curve makes the same assumptions as above, but with a modified mass function that includes a correction corresponding to non-Gaussianity with $f_{\text{NL}} = 43$ [33]. The red curve assumes the lower limit on the mass function [34], a large minimum mass (assuming relative velocities between baryons and dark matter [44]) and a low gas accretion rate [29]. The shaded area represents everything in between these two extreme cases. The two vertical lines correspond to the $5\sigma$ and $10\sigma$ sensitivity to $m_{\text{BH}} = 30M_\odot$ black hole pairs with the future gravitational wave detector, Cosmic Explorer [8].

In addition, the assumption of a $\delta$-function mass spectrum of black holes is not realistic. A range of black hole masses is most likely present. The effects of such an assumption have been studied in the context of explaining the current rate of observed black hole merger events with LIGO [48–55]. In our case, such a black hole mass function will alter the shape of $N(z)$, but the effect on $z_{\text{max}}$ is negligible since the factors that give rise to the cutoff remain as discussed earlier (namely the shape of the halo mass function and the decline in gas infall at high redshifts).

The prediction of a maximum redshift for black hole merger events can be tested with future gravitational wave detectors. In particular, Cosmic Explorer [8] will have the ability to detect events at these very high redshifts. Given the current design capabilities, Cosmic Explorer will be able to detect the merger of $30M_\odot$ black hole pairs at $10\sigma$ significance out to redshift of $z \approx 36$ and at $5\sigma$ significance to redshift $z \approx 44$ [8]. These two limits are shown as vertical dashed lines in Figure 3.

Any detection of an event rate greater than once a year from a redshift greater than $z_{\text{max}} \approx 40$ will have major implications for cosmology. It would mean that either structure formation is not proceeding in the way that is currently envisioned, or that black hole mergers are due to some exotic phenomenon. Two such possibilities exist: a strange non-Gaussianity that is not parametrized in terms of $f_{\text{NL}}$ (e.g., decay of cosmic strings [56]), or from the merger of primordial black holes [17]. The latter idea has received considerable attention recently in light of the spectacular detection of gravitational waves by LIGO; however at present it seems that other astrophysical constraints make such a possibility less likely [57, 65]. Nevertheless, if events with redshifts greater than $z_{\text{max}} \approx 40$ appear with rates greater than once per year, it may still be possible to disentangle their origin by looking at their redshift distribution as the exact dependence on redshift will be sensitive to the abundance of primordial binaries.

We acknowledge useful discussions with Robert Fisher, Zoltn Haiman, Alex Geringer-Sameth, Darren Reed and Kyriakos Vattis. This work was supported by the Black Hole Initiative, which is funded by a grant from the John Templeton Foundation. SMK acknowledges support from DE–SC0017993 and the Institute for Theory and Computation at the Harvard-Smithsonian Center for Astrophysics. We acknowledge usage of the HMFCalc online tool [66].

[1] LIGO Scientific Collaboration, J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, et al., Classical and Quantum Gravity 32, 074001 (2015), 1411.4547.
[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Physical Review Let-
[3] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Physical Review Let-
[4] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams,
A. C. Becker, D. P. Bennett, K. H. Cook, N. Dalal, A. J. Drake, K. C. Freeman, et al., ApJ 550, L169 (2001), astro-ph/0011506.

[59] P. Tisserand, L. Le Guillou, C. Afonso, J. N. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, J. P. Beaulieu, X. Charlot, et al., A&A 469, 387 (2007), astro-ph/0607207.

[60] D. P. Quinn, M. I. Wilkinson, M. J. Irwin, J. Marshall, A. Koch, and V. Belokurov, MNRAS 396, L11 (2009), 0903.1644.

[61] M. A. Monroy-Rodríguez and C. Allen, ApJ 790, 159 (2014), 1406.5169.

[62] Y. Ali-Haïmoud and M. Kamionkowski, ArXiv e-prints (2016), 1612.05644.

[63] D. Aloni, K. Blum, and R. Flauger, ArXiv e-prints (2016), 1612.06811.

[64] T. D. Brandt, ApJ 824, L31 (2016), 1605.03665.

[65] S. M. Koushiappas and A. Loeb, Physical Review Letters 119, 041102 (2017), 1704.01668.

[66] S. G. Murray, C. Power, and A. S. G. Robotham, Astronomy and Computing 3, 23 (2013), 1306.6721.