The redshift dependence of black hole mass distribution: Is it reliable for standard sirens cosmology?

Suvodip Mukherjee

1 Perimeter Institute for Theoretical Physics, 31 Caroline Street N., Waterloo, Ontario, N2L 2Y5, Canada

29 July 2022

ABSTRACT
An upper limit on the mass of a black hole set by the pair-instability supernovae (PISN) process can be useful in inferring the redshift of the gravitational wave (GW) sources by lifting the degeneracy between mass and redshift. However, for this technique to work, it is essential that the PISN mass-scale is redshift independent or at least has a predictable redshift dependence. We show that the observed PISN mass-scale can get smeared and the position of the PISN mass-scale is likely to exhibit a strong redshift dependence due to a combined effect from the non-zero value of the delay time between the formation of a star and the merging of two black holes and the metallicity dependence of PISN mass scale. Due to the unknown form of the delay-time distribution, the redshift dependence of the PISN mass cut-off of the binary black holes (BBHs) cannot be well characterized and will exhibit a large variation with the change in redshift. As a result, the use of a fixed PISN mass scale to infer the redshift of the BBHs from the observed masses will be systematically biased. Though this uncertainty is not severe for the third observation run conducted by the LIGO-Virgo-KAGRA collaboration, in the future this uncertainty will cause a systematic error in the redshift inferred from the PISN mass scale. The corresponding systematic error will be a bottleneck in achieving a few percent precision measurements of the cosmological parameters using this method in the future.

Key words: gravitational waves, black hole mergers, cosmology: miscellaneous

1 INTRODUCTION
Inference of the cosmological parameters from the gravitational wave (GW) sources is one of the key science goals of the currently ongoing network of GW detectors (Abbott et al. 2018) such as LIGO (Aasi et al. 2015), Virgo (Acernese et al. 2014), and for the upcoming GW detectors such as KAGRA (Akutsu et al. 2020), LIGO-India (Unnikrishnan 2013), LISA (Amaro-Seoane et al. 2017), Cosmic Explorer (Reitze et al. 2019; Hall & Evans 2019), and Einstein Telescope (Punturo et al. 2010), as it can provide accurate measurement to the luminosity distance of the GW sources within the framework of the general theory of relativity, and without invoking any additional distance calibration. This was shown for the first time in the seminal work by Schutz (1986), which justifies the reason for calling GW sources the standard sirens. However one of the essential requirements for making robust measurements of the cosmological parameters from standard sirens is to be able to make an independent and accurate inference of the redshift of these sources. Standard sirens, though excellent distance tracers, cannot provide the redshift to the sources independently, unless one can break the degeneracy between mass and redshift, by using a known mass scale. For binary neutron star (BNS) sources, one can use the tidal deformability to break the mass-redshift degeneracy (Taylor et al. 2012; Messenger & Read 2012). On the other hand for the binary black holes (BBHs), the existence of a maximum mass of the black holes (BHs), can be used to infer the redshift to the sources (Farr et al. 2019; You et al. 2021; Mastrogiovanni et al. 2021a). From the theory of stellar models of BH formations, it is predicted that there is a maximum mass of the BHs due to the pair-instability supernovae process (PISN) (Bethe & Brown 1998; Portegies Zwart & Yungelson 1998; Heger & Woosley 2002; Belczynski et al. 2002), and the maximum mass of the BH is expected to be between $40 \pm 50 M_\odot$ (Farmer et al. 2019; Renzo et al. 2020). A recent study has also shown (Farmer et al. 2019), that the variation of the maximum mass-scale changes by only about 10% due to a large variation in the stellar metallicity and stellar winds. As a result, if the PISN mass-scale is robust within 10% accuracy, it can be used to break the mass-redshift degeneracy and can be used to infer the cosmological parameters which affect the cosmic expansion history. Even in the absence of an accurate theoretical prediction of the maximum mass value, one can do a joint estimation of the cosmological parameters and the PISN mass-scale (Farr et al. 2019; You et al. 2021; Mastrogiovanni et al. 2021a). So, this particular method can be extremely powerful also for the third-generation GW detectors such as Cosmic Explorer (Reitze et al. 2019; Hall & Evans 2019), and Einstein Telescope (Punturo et al. 2010) to
infer the cosmic expansion history to a high redshift \( z \approx 50 \), which is unlikely from any electromagnetic probes in the same time scale.

However, for the usage of the PISN mass-scale for the cosmological purpose, the mass-scale must be redshift independent, or at least the redshift dependence should be predictable. If the PISN mass scale evolves with cosmological redshift, then our ignorance of the redshift dependence on the PISN mass scale, will lead to a systematic error in the measurement of the true cosmological redshift and hence will affect the inference of cosmological redshift that will use this redshift. Several previous studies assumed that the PISN mass-scale is redshift independent or will exhibit mild redshift dependence based on the theoretical studies (Farmer et al. 2019; Renzo et al. 2020), and as the stellar metallicity at low redshift \( (z < 2) \) varies not very significantly. In this work, we scrutinize whether the assumption that the PISN mass-scale is redshift independent is valid and show whether we can use it reliably to infer the true redshift to the binary BHs. Recent studies have explored the variation of the mass distribution with redshift for a fixed model of cosmology from GWTC-2 (Fishbach et al. 2021) and GWTC-3 (Abbott et al. 2021a) and did not find any statistically significant deviation from the redshift independent scenario.

We show that even though the variation of the PISN mass-scale is below 10% across a large variation in the metallicity range \( (Z \in \{10^{-3} - 10^{-5}\}) \) (Farmer et al. 2019), the delay time between the formation of the stars and merger of the BBHs will play a key role in the observed mass distribution of the BBHs. The observed mass distribution of the BBHs at a redshift will arise from the BBHs formed over a wide range of cosmological redshifts due to a non-zero value of the delay time between the formation of star and merger of BHs. As a result, the mass distribution of the BBHs at any redshift will come from a vast range of redshifts over which the stellar metallicity can vary significantly. In this paper, we show the impact of the observed mass distribution of BBHs on the inferred redshift if a fixed value of the PISN mass scale \( M_{\text{PISN}} \) is assumed.

The paper is organized as follows, in Sec. 2 we introduce the idea behind the redshift dependence of the PISN mass scale and show its impact on the GW source population in Sec. 3. In Sec. 4 we show the inference of the PISN mass scale from the mock GW samples for the network of LIGO-Virgo-KAGRA (LVK) detectors and Cosmic Explorer (CE) and in Sec. 5 we discuss its impact on the GW population and the cosmological parameters. Finally in Sec. 6, we discuss the conclusion and the future prospects.

2 LOWER EDGE OF THE PISN MASS-GAP AND ITS REDSHIFT DEPENDENCE

The gap in the mass distribution of stellar-mass BHs is expected due to the mass loss of the heavy stars due to the PISN process (Fowler & Hoyle 1964; Rakavy & Shaviv 1967; Frayle 1968; Bond et al. 1984; Woosley et al. 2002; Talbot & Thrane 2018; Marchant et al. 2019; Stevenson et al. 2019). The lower limit of the mass gap is expected to be around 45 \( M_\odot \), which is set by the mass loss during the PISN process (Farmer et al. 2019; Renzo et al. 2020). The value of the lower limit of the PISN mass is shown to vary by only about 7% for the variation in the stellar metallicity \( Z \) from \( 10^{-5} \) to \( 3 \times 10^{-3} \), as the mechanism of wind-loss depends strongly on the stellar metallicity (Vink et al. 2001; Mokiem et al. 2007).

Though multiple other parameters such as the nuclear rates can affect the PISN mass-scale \( M_{\text{PISN}} = 45 M_\odot \) by about 35%, this parameter is not expected to evolve with cosmological redshift and hence cannot cause redshift evolution of the PISN mass-scale. So, the stability of the lower edge of the PISN mass gap can be a robust feature for cosmology.

However, whether the PISN mass-scale can be used for reliably inferring the cosmological parameter depends on whether the observed mass distribution is redshift independent or not, or at least whether the redshift dependence of the PISN mass-scale can be well characterized. In this work, we scrutinize the possible redshift dependence of the observed mass distribution of BBHs even in the scenario when the PISN mass scale is well predicted by theoretical studies. Though in reality, there may be uncertainties in the theoretical understanding as well. As a result, the uncertainty that we are exploring in this paper should be considered as only a lower bound. There can be additional errors on top of this as well.

We model the metallicity dependence of the PISN mass cut-off by a relation

\[
M_{\text{PISN}}(Z) = M_{\text{PISN}}(Z_*) - \alpha \log_{10}(Z/Z_*),
\]

where a weak dependence on the metallicity can be captured as a logarithmic correction with a free parameter \( \alpha \). For \( \alpha = 1.5 \) at \( Z_* = 10^{-4} \) and \( M_{\text{PISN}}(Z_*) = 45 M_\odot \), we can capture the variation in the PISN mass scale shown in (Farmer et al. 2019) due to the change in metallicity over the range \( Z \in [10^{-5}, 5 \times 10^{-3}] \). Along with the evolution of the metallicity, the PISN mass scale can also vary due to changes in the fundamental physics such as the reaction rates (Mehta et al. 2022).

The stellar metallicity in the Universe evolves with redshift (Mannucci et al. 2010; Sommariva et al. 2012; Krumholz & Dekel 2012; Dayal et al. 2013; Madau & Dickinson 2014). The metallicity at a high redshift \( (z > 2) \) is much smaller in comparison to the low redshift Universe \( z < 2 \). The first-generation stars contaminate the interstellar medium and cause a chemical evolution of the Universe. We can treat the metallicity evolution with redshift by a relation

\[
\log_{10}(Z(z)) = \gamma z + \zeta,
\]

where \( \gamma \) captures the redshift dependence and \( \zeta \) captures the metallicity value at \( z = 0 \) (Mannucci et al. 2010; Madau & Dickinson 2014). This relation captures the metallicity of the parent star or the gas cloud from which a star has formed. It is written to express only a mean evolution of the metallicity. Along with the mean metallicity evolution of the Universe, there is going to be a scatter in the metallicity depending on the galaxy properties. Such a source of uncertainty brings additional stochasticity to the metallicity relation. Currently, a limited number of observations (Gallazzi et al. 2008; Mannucci et al. 2010; Krumholz & Dekel 2012) are available to explore the environment dependence of the metallicity, and most of our current understandings are based on simulations (Genel 2016; Torrey et al. 2019). These studies show that the overall median metallicity dependence of the galaxies at different can be explained by power form (Pei et al. 1999; Young & Fryer 2007; Torrey et al. 2019). Several
Redshift dependence of PISN mass scale

Figure 1. A schematic diagram showing the source of variation in the PISN mass of the detected GW sources merging at a fixed redshift $z$. BHs formed at different redshifts can merge at the same redshift due to a delay time distribution with a non-zero value of the minimum delay time. The sources at high redshift can have lower metallicity and can have a larger value of the PISN mass scale than the BHs formed at low redshift. So, the PISN mass of the detected GW sources will exhibit a variation with redshift.

The observed mass distribution of the BBHs at a redshift $z_m$ is going to contribute from the BHs which have formed at an earlier redshift $z < z_m$ due to a non-zero value of the delay time $t_d$. The formation time of the individual companion objects will be at a different redshift. As a result, the metallicity dependence of the star corresponding to that will play a role. So, for a probability distribution $P(t_d)$ of the delay time, there is going to be mixing between the BHs forming at different redshifts. The corresponding window function of masses of the BHs mergers can be written as

$$W(m(z_m)) = \mathcal{N} \int_{z_m}^{\infty} P(t_d(z_m, z')) W_s(m(z')) dz',$$

where $\mathcal{N}$ is a factor to normalize the window function. The term $P(t_d(z_m, z'))$ denotes the probability for a BBH source to be at a redshift $z_m$ formed at a redshift $z$ with a delay time $t_d$, and the term $W_s(m(z'))$ denotes the window function of the BH masses at a redshift $z'$ with the mass $m$ denoting the mass of the individual objects in their source frame. The convolution of the probability distribution of the delay time and the probability distribution of the mass distribution gives the source frame mass distribution at redshift $z_m$. The mass distribution of BBHs at redshift $z'$ can be modified depending on the metallicity evolution in the Universe. We show a schematic diagram in Fig. 1 explaining the variation in the PISN mass of the BHs merging at a fixed redshift. At a fixed redshift (denoted by $z$), the merging BBHs could have formed at different redshifts ($z_1', z_2', z_3'$, and $z_4'$) from the parent stars at redshifts (denoted by redshifts $z_1, z_2, z_3,$ and $z_4$). So, the PISN mass scale of the BHs formed at different redshifts. The time lapsed between the formation of a star and the BH can be negligible in comparison to the cosmic time scale of it to merge (which is a few hundred Myr to Gyr) for the heavier black holes that contribute to the PISN mass scale. The redshift of the two parent stars are nearly the same $z_2 \sim z_1$ (and also $z_4 \sim z_3$). However, the difference between the redshifts like $z_3 - z_1$ of individual binary pairs can be large.

1 The time lapsed between the formation of a star and the BH can be negligible in comparison to the cosmic time scale of it to merge (which is a few hundred Myr to Gyr) for the heavier black holes that contribute to the PISN mass scale.
redshifts will be different for the heavier BH. The probability distribution of the delay time with a non-zero minimum value leads to the mixing of BHs formed at a different cosmic time to merge with different PISN masses. As a result, BBHs will not have a fixed PISN mass scale and can show a large variation even if the variation due to a change in the metallicity of the PISN mass is small. The signature of the delay time distribution on the BHs mass distribution can arise in all those scenarios of binary formation for which the delay time distribution can go beyond a few hundreds of Myrs. Only for the scenarios where BBHs are merging very fast (less than a few tens on Myrs), the masses of the BBHs may not show a significant difference, if the metallicity of host galaxies are similar.

The delay-time distribution and the corresponding value of the minimum delay time are not well known. Several theoretical studies are made to understand the delay time distribution and its relation with the BBH formation channels (O’Shaughnessy et al. 2010; Banerjee et al. 2010; Dominik et al. 2012; Dominik et al. 2015; Mandel & de Mink 2016; Lamberts et al. 2016; Cao et al. 2018; Vitale et al. 2019; Eldredge et al. 2018; Eldridge et al. 2019; Callister et al. 2020a; du Buisson et al. 2020; Santoliquido et al. 2021; Safarzadeh et al. 2018; Toffano et al. 2019; Farmer et al. 2019). In a uniform in the log-space distribution and its relation with the BBH formation channels (O’Shaughnessy et al. 2010; Banerjee et al. 2010; Dominik et al. 2012; Dominik et al. 2015; Mandel & de Mink 2016; Lamberts et al. 2016; Cao et al. 2018; Vitale et al. 2019; Eldredge et al. 2018; Eldridge et al. 2019; Callister et al. 2020a; du Buisson et al. 2020; Santoliquido et al. 2021; Safarzadeh et al. 2020; van Son et al. 2022). For a uniform in the log-space distribution of the separation between the binaries, there is going to be a power-law form of the delay time distribution with \( P(t_d) \propto t_d^{-\kappa} \) with the value of \( \kappa = 1 \) for \( t_d > t_d^{\text{min}} \), where \( t_d^{\text{min}} \) denotes the minimum value of the delay time. For \( t_d < t_d^{\text{min}} \), the probability distribution is zero. However, studies have shown that the probability distribution of the delay time can have a different functional form, apart from the unknown value of the minimum delay time denoted by \( t_d^{\text{min}} \).

Recently, constraints on the delay time are obtained from the individual events (Fishbach & Kalogera 2021) and the stochastic GW background (Mukherjee & Silk 2021). In the future, it will be possible to measure this quantity to much higher precision by correlating the astrophysical properties of the GW sources with the star formation and metallicity properties of the host galaxies, which can be inferred from different emission lines signatures (Mukherjee & Dizgah 2021). Gravitational lensing of GW can also provide an independent way to measure the delay time distribution from the current generation detectors (Mukherjee et al. 2021e).

For large delay time scenarios, the component mass of the BHs can arise from a very high redshift. At high redshifts, the PISN mass-scale is going to be different and those sources contribute to the BBH mergers. For a probability distribution of the delay time \( P(t_d) = t_d^{-\kappa} \) for \( t_d \geq t_d^{\text{min}} \), we will witness a mixing of the BHs from different redshifts to contribute at the redshift \( z_m \). The probability mass distribution is going to depend on parameters \( \kappa, t_d^{\text{min}}, \gamma, \alpha \). We show the probability distribution of mass at different redshifts for a few different values of the parameters such as \( \kappa, t_d^{\text{min}}, \gamma, \alpha \) in Fig. 2. For the values chosen for the parameters \( \kappa, t_d^{\text{min}}, \gamma, \alpha \), the variation in the PISN mass scale is well within the variation in the mass range explored previously (Belczynski et al. 2002; Belczynski et al. 2016; Mapelli et al. 2017; Giacobbo et al. 2018; Toffano et al. 2019; Farmer et al. 2019).

Fig. 2 shows that for different redshifts, the window function for the GW mass-scale is not the same at all redshifts. At higher redshifts, the PISN mass-scale exhibits a larger value of the PISN mass cut-off than at a lower redshift. This variation in the mass scale is more if the metallicity evolution with redshift is large or the delay time distribution is large or with a shallow evolution of the probability distribution (\( \kappa \) close to zero). The shift in the PISN mass scale \( M_{\text{PISN}} \) and the maximum mass \( M_{\text{max}} \) with redshift is shown in Fig. 3. It indicates that for different scenarios, there is a significant variation in the mass scales. This implies that the observed delay time distribution of the BBHs is going to be significantly different because of the variation of the parameters related to metallicity and delay time distribution. For the scenarios considered in this analysis, the variation in PISN mass scale \( M_{\text{PISN}} \) can be between 10\% to 50\%. The variation in the maximum mass scale \( M_{\text{max}} \) can be also around 25\%. A previous study explored the metallicity dependence of the \( M_{\text{max}} \) (Safarzadeh & Farr 2019) but did not show the impact of delay time on the mass distribution and how it can naturally produce a redshift dependence of the PISN mass scale and the smearing property of the BH mass distribution. The corresponding source frame probability mass distribution merging at a redshift \( z_m \) can be written as

\[
\mathcal{P}(m(z_m)) = W(m(z_m))\mathcal{P}_s(m(z_m)),
\]

where \( \mathcal{P}_s(m(z_m)) \) is the probability distribution of the stellar compact objects that source the BH formation. This probability distribution can be considered to be motivated by a simple power-law form \( m^{-\kappa} \). We assume in this paper that \( \mathcal{P}_s(m(z_m)) \) is redshift independent. But a redshift-dependent scenario is not outside the theoretical scope. The observed masses of the GW sources merging at redshift \( z_m \) will be redshifted which can be written as

\[
m_{\text{det}} = (1 + z_m)m.
\]

We show the probability distribution of the masses, for this simple power-law form in Fig. 4 for two different redshift values \( z = 1.1 \) and \( z = 6.1 \). So, even though the mass distribution of the BHs is extended to higher masses, the mass window function \( W(m(z_m)) \) suppresses the masses above the mass scales set by the window functions at those two redshifts. The mass spectrum shows that the cutoff happens at different mass values at different redshifts. As result, the observed mass distribution of BHs is no more redshift independent.

In summary, even a marginal dependence of the PISN mass scale of a BH can lead to a much broad distribution of the PISN mass scale due to the mixing between the BHs formed at a different cosmic time due to a non-zero value of the delay time between the formation of a star and merger of a black hole. The probability distribution of the delay time leads to redshift dependence in the observed PISN mass scale and the maximum mass of a BH. We summarise below the key aspects of the redshift dependence of the mass distribution.

**The key aspects of this mass modeling**

- This model predicts one of the potential sources of redshift dependence of the PISN mass scale \( M_{\text{PISN}} \) and shows how it is correlated with the quantities such as stellar metallicity, delay time distribution, and the value of the minimum delay time.
- This model naturally predicts a smearing behavior between the PISN mass scale and the maximum value of the mass distribution and how it evolves with redshift.
- This model predicts a redshift dependence of the maximum mass value denoted by \( M_{\text{max}} \) and also its correlation.
Redshift dependence of PISN mass scale

Figure 2. We show the variation in the window function for different redshifts by varying the stellar metallicity and delay time distribution.

with the PISN mass scale $M_{\text{PISN}}$. The maximum mass scale also depends on the metallicity evolution with redshift and on the delay time distribution.

It is worth mentioning that the delay time distribution also impacts the redshift evolution of the merger rate of the BBHs (Dominik et al. 2012; Dominik et al. 2015; Mapelli et al. 2017; Giacobbo et al. 2018; Vitale et al. 2019; Toffano et al. 2019; Fishbach & Kalogera 2021; Mukherjee & Dizgah 2021). So, joint estimation of the parameters related to the delay time distribution and the cosmological parameters can be made using their mass distribution and their distribution in luminosity distance.

In light of the recent results from the LVK observations (Abbott et al. 2019, 2020, 2021b), we have detected a few events whose masses are higher than the usual PISN scale considered as $45 \, M_{\odot}$. As we have shown in Fig. 2, due to a non-zero value of the delay time and evolution of the stellar metallicity, the mass distribution of the BHs can go up to high redshift. The window function $W(m)$ gets a slope in its distribution, instead of a sharp cutoff. The slope of the mass window function can go beyond the value $45 \, M_{\odot}$, which can allow for heavier masses of the binary mergers. High mass systems can be also be possible from several astrophysical scenarios as shown previously (Di Carlo et al. 2020; Farrell et al. 2021; Costa et al. 2022; Briel et al. 2022). It is also worth mentioning that apart from the first-generation BHs, second-generation BHs due to hierarchical mergers can lead to heavier BHs. This can lead to additional variation in the BH mass distribution (Liu & Lai 2021; Barrera & Bartos 2022). In summary, all these astrophysical uncertainties can make the mass distribution of the BHs redshift dependent which can make it more difficult to use for the cosmological purpose.

3 IMPACT ON THE MASS DISTRIBUTION OF THE GW SOURCES

The redshift evolution of the PISN mass scale with redshift will lead to a change in the mass distribution of BBHs with redshift. We study the impact on the observed mass distribution of the binary sources for two different network configurations, the current generation detectors such as LIGO (Aasi et al. 2015), Virgo (Acernese et al. 2014), and the upcoming GW detectors such as KAGRA (Akutsu et al. 2020) and LIGO-India (Unnikrishnan 2013), and the third generation GW detectors such as Cosmic Explorer (Reitze et al. 2021).
Figure 3. The variation in the PISN mass scale $M_{\text{PISN}}$ (top) and maximum mass $M_{\text{max}}$ (bottom) as a function of the redshift is shown for different astrophysical parameters. The blue and red markers are overlapping with the magenta markers at all the redshifts in the lower panel.

2019; Hall & Evans 2019), and Einstein Telescope (Punturo et al. 2010). The sources detectable from the current generation detectors and the next generation detectors will have a redshift reach up to $z = 1$ and $z \sim 80$ respectively. In our analysis, we primarily focus on the GW sources for redshift up to $z = 7$ for the third-generation detectors, since the expected merger rate at higher redshifts is not well known and difficult to model for the astrophysical BHs which are motivated by the Madau-Dickinson star formation rate (Madau & Dickinson 2014).
We sample GW sources with a mass distribution of the power-law form $m^{-\alpha}$ with $\alpha = 2.35$ and a cut-off at a mass scale according to the mass window function which is a function of redshift. Along with a power-law form, we consider a bump at the $M_{\text{PISN}}$ scale with the relative height of the peak in comparison with the low mass value as $\Lambda_d = 0.1$, which is in the agreement with the GWTC-3 observation from LVK collaboration (Abbott et al. 2021d,a,b,c) and also with the previous population synthesis models (Stevenson et al. 2019). The sources at higher redshifts have a mass cutoff at a higher value resulting in a broader mass spectrum at a high redshift than at a low redshift. We use a network of four GW detectors (LIGO-Hanford, LIGO-Livingston, Virgo, and KAGRA) to estimate the posteriors on the mass samples for GW sources distributed up to redshift $z = 1.1$ which is possible to access from the current generation detectors (Abbott et al. 2018).

We use an approximated form of the likelihood to infer the posteriors on the GW source parameter, following the previous analysis (Farr et al. 2019; Mastrogiovanni et al. 2021a). This assumption will not change the main conclusion of the paper on the incorrect estimation of the redshift using the PISN mass scale. The matched filtering signal to noise ratio (SNR), $\rho$ is estimated for BBHs with detector-frame chirp mass $M_d$ at the luminosity distance of $d_l$, using the relation (Farr et al. 2016; Mastrogiovanni et al. 2021b)

$$\rho = \rho_\ast \Theta \left(\frac{\mathcal{M}}{M_d^\ast}\right)^{5/6} \left(\frac{d_l^\ast}{d_l}\right),$$

where $\Theta$ is the detector projection factor that we assume as uniform between $[0, 1]$. We have made a pessimistic choice of the projection factor to show the impact of the bias in inferring the redshift using GW sources. For sources, with a higher value of the projection factor, the bias in the redshift estimation can be even more noticeable. The value of $\rho_\ast = 8$ is set for the parameters $M_d^\ast$ and $d_l^\ast$ for the LVK design sensitivity (Abbott et al. 2018) and CE (Hall & Evans 2019). The network matched filtering SNR is obtained by using $\rho_{\text{det}} = \sum_i \rho_i^2$. We choose BBHs having $\rho_{\text{det}} \geq 10$ as detected events. The estimated mass distribution as a function of redshift shows the variation due to the evolving mass distribution. This indicates that the assembling of the heavier objects takes place at higher masses for higher redshifts. The corresponding redshift evolution of the mass distribution is shown in Fig. 5 only for the events which are close to the PISN mass scale.

4 INFERENCW OF THE SOURCE REDSHIFT USING THE PISN MASS CUTOFF

The PISN mass scale can be used to identify the redshift of the GW sources if the mass of the PISN mass scale is redshift independent. However, as we have discussed in the previous sections, the PISN mass scale is going to vary with redshift. As a result, if we would like to infer the redshift of the GW sources for a fixed value of the PISN mass scale, then there is going to be a systematic error in the redshift estimation. The redshift dependence of the PISN mass scale depends on several astrophysical parameters (discussed above) and cannot be characterized uniquely.

We show the inferred redshift of the value assuming a fixed PISN mass scale indicated by $z_{\text{PISN}}$ and the true redshift of the sources in the mock in Fig. 6 along with the uncertainty due to the inferred mass distribution for 2.5 years of LVK observations and 0.5 years of Cosmic Explorer for the case with $t_d^\text{min} = 0.5$ Gyr, $\kappa = 1$, $\gamma = -0.44$, $\alpha = 1.5$. This scenario is chosen as it has the minimum shift in the PISN mass scale in comparison to the other cases. So, it will describe the minimum error for the scenarios considered in this analysis. The deviation is much larger for the third-generation detectors such as Cosmic Explorer and Einstein Telescope than for the current generation GW detectors such as LVK. The systematic error in the inferred redshift is going to vary for different choices of the astrophysical parameters $t_d^\text{min}$, $\kappa$, $\gamma$, $\alpha$ considered in this analysis. For the cases with large variation in the $M_{\text{PISN}}$ (see Fig. 3), the corresponding systematic error in the inferred redshift $z_{\text{PISN}}$ obtained using assuming a fixed of $M_{\text{PISN}}$ will be larger than the case shown in Fig. 6. The difference between the inferred redshift and true redshift is shown in Fig. 6 bottom panel for four different cases. The
error bars on the redshift are the 68% C.I. obtained from the inferred redshift by combining the mass posteriors. This indicates that with the increase in more events, the statistical error in the mass inference will be subdominant than the systematic error due to the unknown redshift dependence of the PISN mass scale. We find that there is going to be about 10% variation for this model parameters. The variation can go as large as 25 – 30% for different values of the parameters $t_q, \gamma, \alpha$ considered in this analysis (see Fig. 2). Moreover, this systematic uncertainty is only a minimum uncertainty. Incorrect astrophysical modeling, variation of metallicity with galaxies, multiple delay time distributions, and inaccurate prescriptions of stellar winds can increase this even further.

5 IMPACT ON GW SOURCE POPULATION INFERENCE AND COSMOLOGY

From the above discussion, it is evident that the redshift inference from the PISN mass scale is not accurate and cannot be well predicted due to two major sources of uncertainties (i) redshift dependence of the metallicity and (ii) redshift dependence of the delay time distribution. Even in the scenario when the PISN mass scale can be correctly modeled $^2$, then also the observed mass population will not be redshift independent. This will have a significant impact on the inference of the GW source population as well as on the inference of cosmological results.

Impact on the GW source population: The redshift dependence of the GW mass distribution will have a significant impact on understanding the true mass distribution of the sources. Our study shows that the position of the PISN mass cutoff and the mass range over which the distribution can be smeared both are going to depend on redshift through the delay time distribution and metallicity. As a result, it is important to include the redshift dependence of the mass distribution of the GW sources in the analysis. Moreover, the redshift dependence of the mass distribution of the BBHs and also their merger rates are going to be correlated due to their dependence on the delay time distribution. So, a joint analysis of the GW mass distribution and the merger rates will be appropriate to infer the PISN mass scale, metallicity dependence, and the delay time distribution. In a recent work (Karathanasis et al. 2022), we made a joint estimation of these quantities from the GWTC-3 data of the LVK collaboration. An independent way to infer the PISN mass distribution is using the cross-correlation technique after marginalizing the GW bias parameters (Mukherjee & Wandelt 2018; Mukherjee et al. 2020, 2021b; Diaz & Mukherjee 2021). A proper inference of the mass distribution of the BBHs will also be useful for the analysis of the stochastic GW background (Mukherjee & Silk 2021; Callister et al. 2020b; Mukherjee & Silk 2021; Mukherjee et al. 2021a) and lensing event rates (Oguri 2018; Mukherjee et al. 2021e). The amplitude of the stochastic GW background and also the lensing event rates strongly depend on the mass distribution of the BBHs. Inclusion of the redshift dependence of the PISN mass distribution explored in this analysis will have a vital role in correctly estimating these signals from the data. Moreover, constraints on

$^2$ This is unlikely the case, as there exist several other unknown factors that can change the PISN mass scale.
redshift dependence of the PISN mass cutoff. Our analysis shows that there can be a 10-30% systematic error in PISN mass scale over the redshift up to $z = 7$. As a result, a few percent level measurement is not possible unless this effect can be properly mitigated. Several recent analyses (Ezquiaga 2021; Mancarella et al. 2021; Hernandez 2021) assume a fixed PISN mass scale and hence these measurements are not reliable and subject to systematic errors. The inference of the cosmological parameters jointly with the frictional term from GW propagation can be made in a robust way using the cross-
Figure 7. The impact of the incorrect estimation of the redshift on the luminosity distance $d_L$ and redshift $z$ plane is shown for LIGO-Virgo-KAGRA (LVK) for 2.5 years and 10 years and Cosmic Explorer (CE) for 0.5 years and 5 years along with the case for the true value of redshift in yellow. We show also a family of curves by changing only the dark energy equation of state parameter $w_0$ from -0.5 to -1.48. This indicates that an incorrect inference of the redshift from the PISN mass scale can lead to a map of an incorrect cosmic expansion history.

correlation technique from BBHs and inferring the clustering redshift (Mukherjee et al. 2021d; Cañas-Herrera et al. 2021).

6 CONCLUSION AND THE WAY FORWARD

In this work, we explore for the first time the redshift dependence of the PISN mass cutoff and its implication on the estimation of the source redshift of BBHs. We show that even if the PISN mass cutoff and its dependence on the stellar metallicity can be well modeled theoretically, the redshift evolution of the metallicity and the delay time distribution of the BBH mergers can lead to a mixing between BHs formed across a large cosmic epoch. This will lead to redshift dependence of the PISN mass cutoff. The unknown form of the probability distribution of the delay time will lead to an unknown amount of systematic shift in the PISN mass cutoff to higher values with a change in the redshift. The mass scales over which the BH distribution will be smeared to zero also depend on the metallicity and the delay time distribution. Ignoring the redshift dependence for any cosmological and GW population analysis can lead to a systematic bias.

We show that the use of a fixed PISN mass cutoff to infer the source redshift of the BBHs can lead to a biased inference. Even for one of the best cases (minimum shift in the PISN mass cutoff) and assuming that one can model the metallicity dependence of the PISN mass cutoff theoretically, we find that there is going to be around 10 – 30% systematic error in the inference of the redshift of the sources (towards high values). Though this error is sub-dominant to the statistical error in the mass measurement of the current LVK data (Abbott et al. 2021d,a,b,c) and the cosmological results inferred from GWTC-3 (Abbott et al. 2021c) are unlikely to be influenced by this, but it is going to be a major source of uncertainty in the future when statistical error will be sub-dominant. This is also in agreement with the recent studies from GWTC-2 and GWTC-3 (Fishbach et al. 2021; Abbott et al. 2021a). We show that for a network of four GW detectors such as LIGO-Hanford, LIGO-Livingston, Virgo, and KAGRA and also in the future from Cosmic Explorer, the systematic error will lead to a systematic bias in the inferred value of redshift at the level of 10-30%. As a result, to reach a percent level measurement of the cosmological parameter using the PISN mass cutoff, it is essential to mitigate this systematic error. Similarly, a wrong inference of the redshift will have an impact on the measurement of any non-GR signatures due to GW propagation.

The redshift dependence of the PISN mass cutoff modeled in this paper will also be useful for understanding the source frame mass distribution of the BBHs. The inclusion of a redshift dependence of the PISN mass cutoff and exploring its dependence on the stellar metallicity and delay time distribution will help us to understand the formation channel of the BBHs. An independent way to infer the redshift dependence of the PISN mass scale is by inferring the clustering
redshift using the cross-correlation technique and marginalizing over the GW bias parameters (Oguri 2018; Mukherjee & Wandelt 2018; Mukherjee et al. 2021b; Scelfo et al. 2020; Diaz & Mukherjee 2021).

In summary, this work shows for the first time the redshift independence of the PISN mass cutoff is not possible in reality. Even when there is no theoretical modeling uncertainty, astrophysical uncertainties will play a significant role in inducing the redshift dependence of the PISN mass cutoff. In addition, there can be also secondary sources, and additional theoretical uncertainties that even increase the budget of the systematic error. As a result, the use of the PISN mass cutoff for inferring redshift to the source, and use that for cosmological analysis can cause a significant systematic bias. Though currently, the statistical error is larger than the systematic error, in the future, an accurate measurement of the cosmological parameters will not be possible unless this can be mitigated by other techniques. In future work, we will develop techniques to also include the redshift dependence of the PISN mass cutoff and its joint estimation with the cosmological parameters.

ACKNOWLEDGEMENTS

S.M. is grateful to Simone Mastrogiannl for reviewing the manuscript as a part of the LSC review process and providing very useful suggestions. The author acknowledges useful discussions and feedback from Jonathan Gair, Martin Hendry, Gilbert Holder, Matthew Johnson, and Luis Lehner. S.M. thank Robert Farmer for pointing to a typographical error in the paper. S.M. is supported by the Simons Foundation. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. This analysis is carried out at the computing facility of the Perimeter Institute and the Infinity cluster hosted by Institut d’Astrophysique de Paris. We thank Stephane Rouberol for smoothly running the Infinity cluster. We acknowledge the use of following packages in this work: Astropy (Astropy Collaboration et al. 2013, 2018), IPython (P´erez & Granger 2007), Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011), and SciPy (Jones et al. 01 ). The author would also like to thank the LIGO-Virgo-KAGRA scientific collaboration and Cosmic Explorer collaboration for providing the noise curves. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN), and the Dutch Nikhef, with contributions by Polish and Hungarian institutes. This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

DATA AVAILABILITY

The data underlying this article will be shared at request to the corresponding author.

REFERENCES

Aasi J., et al., 2015, Class. Quant. Grav., 32, 074001
Abbott B. P., et al., 2018, Living Rev. Rel., 21, 3
Abbott B. P., et al., 2019, Astrophys. J. Lett., 882, L24
Abbott R., et al., 2020, Phys. Rev. Lett., 125, 101102
Abbott R., et al., 2021c, arXiv: 2111.03604
Abbott R., et al., 2021b, arXiv: 2111.03606
Abbott R., et al., 2021a, arXiv: 2111.03634
Abbott R., et al., 2021d, Physical Review X, 11
Acerese F., et al., 2014, Classical and Quantum Gravity, 32, 024001
Akutsu T., et al., 2020, arXiv:2005.05574
Amaro-Seoane P., et al., 2017, arXiv e-prints, p. arXiv:1702.00786
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Banerjee S., Baumgardt H., Kroupa P., 2010, MNRAS, 402, 371
Barrera O., Bartos I., 2022, ApJL, 929, L1
Belczynski K., Kalogera V., Bulik T., 2002, ApJ, 572, 407
Belczynski K., Holz D. E., Bulik T., O’Shaughnessy R., 2016, Nature, 534, 512
Belgacem E., Dirian Y., Foffa S., Maggiore M., 2018a, Phys. Rev. D, 97, 104066
Belgacem E., Dirian Y., Foffa S., Maggiore M., 2018b, Phys. Rev. D, 98, 023510
Bethe H. A., Brown G. E., 1998, Astrophys. J., 506, 785
Bond J. R., Arnett W. D., Carr B. J., 1984, ApJ, 280, 825
Briel M. M., Stevance H. F., Eldridge J. J., 2022, arXiv:2206.13842
Buscicchio R., Moore C. J., Pratten G., Schmidt P., Bianconi M., Vecchio A., 2020, Phys.Rev.Lett., 125, 141102
Cañas-Herrera G., Contigiani O., Vardanyan V., 2021, ApJ, 918, 20
Callister T., Fishbach M., Holz D. E., Farr W. M., 2020a, ApJL, 896, L32
Callister T., Fishbach M., Holz D., Farr W., 2020b, Astrophys. J. Lett., 906, L32
Cao L., Lu Y., Zhao Y., 2018, MNRAS, 474, 4997
Costa G., Ballone A., Mapelli M., Bressan A., 2022, arXiv:2204.03492
Dayal P., Ferrara A., Dunlop J. S., 2013, MNRAS, 430, 2891
Di Carlo U. N., Mapelli M., Bouffanais Y., Giacobbo N., Santoliquido F., Bressan A., Spera M., Haardt F., 2020, Mon. Not. Roy. Astron. Soc., 497, 1043
Diaz C. C., Mukherjee S., 2021, arXiv:2107.12787
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O’Shaughnessy R., 2012, ApJ, 759, 52
Dominik M., et al., 2015, Astrophys. J., 806, 263
Elbert O. D., Bullock J. S., Kaplinghat M., 2018, Mon. Not. Roy. Astron. Soc., 473, 1186
Eldridge J. J., Stanway E. R., Tang P. N., 2019, Mon. Not. Roy. Astron. Soc., 482, 870
Ezquiaga J. M., 2021, Phys. Lett. B, 822, 136665
Farmer R., Renzo M., de Mink S. E., Marchant P., Justham S., 2019, ApJ, 887, 53
Farr B., et al., 2016, Astrophys. J., 825, 116
Farr W. M., Fishbach M., Ye J., Holz D., 2019, Astrophys. J. Lett., 883, L42
Farrell E. J., Groh J. H., Hirschi R., Murphy L., Kaiser E., Ekström S., Georgy C., Meynet G., 2021, Mon. Not. Roy. Astron. Soc., 502, L40
Fishbach M., Kalogera V., 2021, Astrophys. J. Lett., 914, L30
Fishbach M., et al., 2021, ApJ, 912, 98
Fowler W. A., Hoyle F., 1964, ApJS, 9, 201
Fraley G. S., 1968, ApSS, 2, 96
Gallazzi A., Brinchmann J., Charlot S., White S. D. M., 2008, MNRAS, 383, 1439
Genel S., 2016, ApJ, 822, 107
Giacobbo N., Mapelli M., Spera M., 2018, MNRAS, 474, 2959

MNRAS 000, 1–12 (2021)
