Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables
Cañas Herrera, G.

Citation
Cañas Herrera, G. (2022, October 19). *Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables*. *Casimir PhD Series*. Retrieved from https://hdl.handle.net/1887/3483658

Chapter 5

Reconstruction of the propagation and origin of gravitational waves

5.1 Introduction

The first direct detection of gravitational waves (GWs) by (B. P. Abbott et al., 2016) triggered a rapidly increasing interest in exploiting this new field for cosmological information. GWs alone are not particularly useful because the data provides only a sky position and a measure of the luminosity distance to the source. However, GW sources can serve as powerful cosmological probes when combined with electromagnetic (EM) data, from which redshifts can be extracted. This idea dates back to (Schutz, 1986) and has two main variations.

The first, simplest, possibility is the observation of so-called *standard sirens* (Holz & Hughes, 2005), GW sources with EM counterparts from which a redshift can be observed. As an example, a population of binary neutron stars, such as the already observed GW170817 (B. P. Abbott et al., 2017c), can be used to reconstruct the luminosity-redshift function and constrain cosmological observables. In this context, the aforementioned observation has already lead to promising results in constraining the Hubble constant $H_0 = 100h$ km/s/Mpc (B. P. Abbott et al., 2017b; H.-Y. Chen, Fishbach, & Holz, 2018).

The second possibility is to use *dark sirens* with a statistical counterpart. The likely counterpart is identified using the rough sky-localization offered by current GW detectors in combination with deep galaxy catalogs. This technique, sometimes combined with the first, has also been used to extract a measurement of H_0 (Soares-Santos et al., 2019; B. P. Abbott et al., 2021) and a refined version of it is the focus of this work. Assuming that both galaxies and the hosts of GW mergers are biased

tracers of the same cosmological structure, it is possible to measure a non-zero cross-correlation signal between the two (see section 5.3). At the linear level, in particular, the description of this signal is especially simple. It should be noted that this idea has been explored by others before, sometimes using a different formalism (Namikawa, Nishizawa, & Taruya, 2016; S. Mukherjee & Wandelt, 2018; Scelfo et al., 2018) and that similar methods have already been employed to provide a measurement of H_0 (Finke, Foffa, Iacovelli, Maggiore, & Mancarella, 2021).

In addition to these two methods, modelling of higher-order effects in the merger dynamics or knowledge of intrinsic source distribution can also be used to extract the luminosity distance-redshift relation from GW observations alone. Because GW signals provide a measurement of the redshifted chirp-mass of the system, features dependent only on the mass of the system (e.g. tidal effects) can be used to extract the cosmological redshift present in the signal (Messenger & Read, 2012). Furthermore, expected features in the source distribution, such as, e.g. the pair instability mass gap (Farr, Fishbach, Ye, & Holz, 2019) or the peak of the star formation rate (Ye & Fishbach, 2021) can also be used to extract this relation from catalogs of GW events.

We also mention that, in principle, the same idea can be used in the absence of resolved events. Previous works, however, have shown that a detection of this effect from a stochastic background signal is unlikely to happen soon due to the low signal-to-noise ratio (SNR) attainable by current and proposed experiments covering the optimal wavelength range (Cañas-Herrera et al., 2020; Alonso, Cusin, Ferreira, & Pitrou, 2020).

In this chapter, we discuss the possibility of using the non-zero correlation between the distribution of EM galaxies and resolved GW mergers to jointly extract information about the two main quantities in the field of GW cosmology: the luminosity distance as a function of redshift, describing the propagation of gravitational waves across cosmic time, and the linear bias of GW sources with respect to the underlying dark matter distribution of the Universe, describing their clustering properties. The linear bias, in this context, is a multiplicative factor relating the spatial correlation function of a source population to the correlation function of the matter distribution of the Universe.

It is already established that GWs carry the potential of constraining the fundamental laws of gravity. This is the case because propagation of GWs in modified gravity scenarios differs from predictions of the general theory of relativity (GR) in multiple ways (Deffayet & Menou, 2007; Garoffolo, Tasinato, Carbone, Bertacca, & Matarrese, 2020; Ezquiaga & Zumalacárregui, 2020). One of the clear signatures is the speed of tensor modes which can be both sub- and superluminal as opposed to the GR case, where GWs propagate at the speed of light. The tight constraints on speed deviations imposed by the multimessenger observations of GW170817 (B. P. Abbott et al., 2017a), for example, have ruled out a wide parameter space of otherwise viable scalar-tensor theories of gravity (Pettorino & Amendola, 2015; Lombriser & Taylor, 2016; Sakstein & Jain, 2017; Ezquiaga & Zumalacárregui, 2017; Creminelli & Vernizzi, 2017; T. Baker et al., 2017). Similarly, implications for bi-gravity models have also been demonstrated in the literature (Amendola, Könnig, Martinelli, Pet-

torino, & Zumalacarregui, 2015; Max, Platscher, & Smirnov, 2017, 2018; Akrami, Brax, Davis, & Vardanyan, 2018, June; Belgacem et al., 2019).

Another striking difference between modified gravity and GR is the modified friction of GWs (Amendola, Sawicki, Kunz, & Saltas, 2018; Belgacem, Dirian, Foffa, & Maggiore, 2018). This feature arises in models with non-minimal coupling of a scalar field and gravity which manifests itself as a redshift-dependent gravitational coupling. As a result, the inferred luminosity distance to GW sources differs from the corresponding EM luminosity distance. This interesting phenomenon has already been constrained using the aforementioned multimessenger detection of GW170817 (Arai & Nishizawa, 2018; Lagos, Fishbach, Landry, & Holz, 2019) and the mass distribution of existing GW catalogs (María Ezquiaga, 2021). In this work, we will investigate the possibility of testing this hypothesis using the spatial clustering of GW sources and galaxies.

The discovery of the first LIGO-Virgo binary black hole has been used to motivate alternative scenarios where the binary did not represent the endpoint of stellar evolution, but originated either as a pair of primordial black holes (PBH) or some exotic compact objects (Bird et al., 2016; Sasaki, Suyama, Tanaka, & Yokoyama, 2016; Bustillo et al., 2021). In particular, the last half-decade has seen a resurgence in interest for PBHs (Clesse & García-Bellido, 2017; Sasaki et al., 2018; Raccanelli, Kovetz, Bird, Cholis, & Munoz, 2016; Raccanelli, Vidotto, & Verde, 2018). The main difference between the PBH and stellar evolution scenarios is the spatial distribution of GW mergers, measurable both in the redshift and sky distribution of the sources. In subsection 5.3.1 we expand on how to model this difference through the linear bias and present our model.

In addition to presenting our formalism, we investigate a possible method to precisely reconstruct the redshift evolution of clustering and modified gravity effects in the upcoming decades. Our study is based on Gaussian processes (GPs), a well-known hyper-parametric regression method (Rasmussen & Williams, 2005). The structure and implementation of this pipeline are presented in section 5.4.

GPs have been widely used in the literature to reconstruct the shapes of physical functions such as the dark energy equation of state $w(z)$ (Shafieloo, Kim, & Linder, 2012; Gerardi, Martinelli, & Silvestri, 2019), the primordial inflaton's speed of sound (Cañas-Herrera, Torrado, & Achúcarro, 2021) or the mass function of the merging binary black hole systems (Li et al., 2021). GPs are very useful for such functional reconstructions due to their flexibility and simplicity, relying only on a handful of hyper-parameters. In general, binned reconstructions (Crittenden, Pogosian, & Zhao, 2009; Crittenden, Zhao, Pogosian, Samushia, & Zhang, 2012; Zhao, Crittenden, Pogosian, & Zhang, 2012), as well as parametric reconstructions with high-degree polynomials are similar in spirit to GPs while also offering more flexibility in the extrapolation (Colgáin & Sheikh-Jabbari, 2021). The chapter is organized as follows. In section 5.2 we summarize the essential concepts concerning the GW propagation in modified gravity models. In section 5.3 we detail the modelling of the clustering correlations between the GW source population and galaxies. In section 5.4 we explain our reconstruction pipeline for the GW luminosity distance and the bias. Our findings are presented in section 5.5 and further discussed in section 5.6.

Unless stated otherwise, our fiducial cosmology is based on the best fit results from Planck 2018 (Planck Collaboration, Aghanim, Akrami, Arroja, et al., 2020). In our analysis, we use `COLOSSUS` (Diemer, 2018) and `Astropy` (Astropy Collaboration et al., 2013, 2018) for cosmological calculations, `sklearn` (*Scikit-learn 0.19.1 documentation: Gaussian Processes*, 2018) for the GP implementation, `emcee` (Foreman-Mackey et al., 2013) as our posterior sampler and `GetDist` (Lewis, 2019) to plot the final contours. Our analysis pipeline is made publicly available, see (Vardanyan, Cañas-Herrera, & Contigiani, 2021).

5.2   Gravitational wave propagation

In GR, and around a background Friedmann-Lemaitre-Robertson-Walker (FLRW) metric, the amplitude of GWs evolves according to

$$h_\alpha'' + 2\mathcal{H}h_\alpha' - \vec{\nabla}^2 h_\alpha = 0, \tag{5.1}$$

where h_α denotes the amplitude of either polarization ($\alpha \in [\times, +]$), primes denote derivatives with respect to the conformal time, and $\mathcal{H}$ is the conformal Hubble function. In this equation, the prefactor of the Laplacian term controls the propagation speed, which we have set to coincide with the speed of light in $c = 1$ units.

The second term is the standard cosmic friction term which causes the strain amplitude to decay as $h_{\alpha(z)} \propto D_{\mathrm{L}}^{-1}(z)$, with D_L being the FLRW luminosity distance:

$$D_{\mathrm{L}(z)} = (1 + z) \int_0^z \frac{\mathrm{d}\tilde{z}}{H(\tilde{z})}, \tag{5.2}$$

where the Hubble function $H(z)$ is given in terms of the Hubble constant H_0, present-day dark matter abundance Ω_{m} and dark energy abundance $\Omega_{\mathrm{DE}(z)}$ as

$$H(z) = H_0 \left[\Omega_m(1 + z)^3 + \Omega_{DE}(z) \right]. \tag{5.3}$$

Throughout this chapter we assume a constant equation of state w_0 for dark energy, such that its energy density is given by

$$\Omega_{DE}(z) = (1 - \Omega_{\mathrm{m}})(1 + z)^{3(1+w_0)}. \tag{5.4}$$

The standard ΛCDM cosmology corresponds to $w_0 = -1$.

It is now established that modifications of GR can affect the propagation of GWs. The important effect for us is the modified friction term with respect to the GR expectation in equation (5.1),

$$h_\alpha'' + [2 + \alpha_M(z)] \, \mathcal{H}h_\alpha' - \vec{\nabla}^2 h_\alpha = 0, \tag{5.5}$$

where we have again imposed the GW speed to be unity as suggested by observations. The modified friction term introduces a new scaling $h_{\alpha(z)} \propto 1/D_{\mathrm{L,GW}}(z)$, with

$D_{\mathrm{L,GW}}(z) \neq D_{\mathrm{L}(z)}$ for non-zero $\alpha_M(z)$. The luminosity distance to GW events can be written as:

$$\frac{D_{\mathrm{L,GW}}}{D_{\mathrm{L,EM}}}(z) = \exp\left\{ -\frac{1}{2} \int_0^z \mathrm{d}\tilde{z}\, \frac{\alpha_M(\tilde{z})}{(1+\tilde{z})} \right\}. \tag{5.6}$$

In this work, we assume that the luminosity distance for EM sources $D_{\mathrm{L,EM}}$ is unaffected and is equal to the expression in equation (5.2). The function α_M corresponds to the running of the effective Planck mass, i.e.,

$$\alpha_M = \frac{d\log(M_{\mathrm{eff}}/M_{\mathrm{P}})^2}{d\log a}, \tag{5.7}$$

where M_{P} is the Planck mass and M_{eff} is its effective value at redshift $z = 1/a - 1$. This function encodes information about extensions of GR such as scalar-tensor theories (Horndeski, 1974; Bellini & Sawicki, 2014) or, more broadly, quantum gravity (Calcagni et al., 2019). The modified friction term is also a natural prediction of non-local modifications of gravity (Dirian, Foffa, Kunz, Maggiore, & Pettorino, 2016)

From an effective field theory point of view $\alpha_M(z)$ is a free function of order unity. In practical studies of modified gravity and dark energy, however, α_M is often assumed to take simple parametric forms. The main guiding principle is the assumption that its effects should be negligible in the early universe, which prompts to choose $\alpha_M(z)$ to be proportional either to the dark energy abundance or simply to some power of the scale factor a.

Such parametrizations make it possible to find a closed form expression for the ratio in equation (5.6) and have inspired a widely used parametrization of the ratio as a monotonic deviation which goes to 1 at present day (Belgacem et al., 2018)

$$\frac{D_{\mathrm{L,GW}}}{D_{\mathrm{L,EM}}}(z) = \Xi_0 + \frac{1 - \Xi_0}{(1+z)^n}. \tag{5.8}$$

In this expression, Ξ_0 and n are two constant parameters, which are typically ~ 1.

5.3 Angular power-spectra

5.3.1 GW sources

We consider GW mergers with a distribution in redshift written as

$$n_{\mathrm{GW}}(z) = \frac{n_0}{1+z}, \tag{5.9}$$

where n_0 corresponds to the comoving number density of observed events as a function of redshift, and the term $(1+z)$ takes into account the cosmological time dilation. In our analysis, we use a constant value of $n_0 \approx 3 \times 10^{-6}\ h^3\mathrm{Mpc}^{-3}$ (with h denoting the usual normalized Hubble constant), motivated by current LIGO constraints (R. Abbott et al., 2021).

For a given selection of sources along the line of sight, the average number of projected sources can be written using the comoving distance $\chi(z)$:

$$\bar{n}_{\mathrm{gw}} = \int_0^\infty dz \frac{\chi^2(z)}{H(z)} S(z) n_{\mathrm{GW}}(z).$$
(5.10)

The function S encodes the selection and the scatter due to observational errors. In this chapter, simple bins in a range $[D_{\mathrm{L,min}}, D_{\mathrm{L,max}}]$ are used and we assume a lognormal distribution with fixed scatter $\sigma_{\ln D}$ for the individual sources (Oguri, 2016). In this case, S can be written as:

$$S(z) = \frac{1}{2} \left[x_{\min}(z) - x_{\max}(z) \right],$$
(5.11)

with

$$x_{\min}(z) = \mathrm{erfc} \left[\frac{\ln D_{\mathrm{L,min}} - \ln D_{\mathrm{L,GW}}(z)}{\sqrt{2}\sigma_{\ln D}} \right],$$
(5.12)

and similarly for $x_{\max}$. Including this effect makes S resemble a top-hat function with damping tails dictated by $\sigma_{\ln D}$.

The angular power spectrum of these sources can be written using the Limber approximation

$$
\begin{aligned}
C_{\mathrm{GW}}(\ell) = \int_0^\infty dz \ \frac{H(z)}{\chi^2(z)} W_{\mathrm{GW}}^2(z) \\
b_{\mathrm{GW}}^2(z) P \left(\frac{\ell + 1/2}{\chi(z)}, z \right),
\end{aligned}
$$
(5.13)

where $P(k, z)$ is the matter power-spectrum at redshift z and comoving scale k, b_{GW} is the bias of the GW sources, and the window function can be written as

$$W_{\mathrm{GW}}(z) = \frac{\chi^2(z)}{H(z)} \frac{n_{\mathrm{GW}}(z)}{\bar{n}_{\mathrm{GW}}} S(z).$$
(5.14)

For the purpose of illustration, we will make use of a few simple parametrization for the GW bias. We will consider either a constant bias b_{GW} with a value of order unity or a more complex form:

$$b_{\mathrm{GW}}(z) = b_0 \left(1 + \frac{1}{D(z)} \right),$$
(5.15)

where $D(z)$ represents the growth factor. The first model, with its low constant value, mimics a PBH origin for the mergers (Bird et al., 2016; Raccanelli et al., 2016), while the second mimics the stellar evolution case by tracking the galaxy linear bias (Oguri, 2016).

5.3.2 Galaxies

Similarly to the GW population, we again assume a constant comoving number density of galaxies. Throughout our analysis we fix

$$n_{\text{gal}}(z) = 10^{-3} h^3 \text{Mpc}^{-3}, \tag{5.16}$$

and we write the autocorrelation signal of galaxies under the Limber approximation as

$$
\begin{aligned}
C_{\text{gal}}(\ell) = \int_0^\infty dz\, \frac{H(z)}{\chi(z)^2} W_{\text{gal}}^2(z) \\
b_{\text{gal}}^2(z) P\left(\frac{\ell + 1/2}{\chi(z)}, z\right).
\end{aligned}
\tag{5.17}
$$

In this expression the definition of W_{gal} is the same as W_{GW} used in the previous section except for using $n_{\text{gal}}(z)$, a different selection function, and $b_{\text{gal}}(z)$ is the linear galaxy bias. In our analyses, we assume a known galaxy bias in the form of

$$b_{\text{gal}}(z) = 1 + \frac{1}{D(z)}. \tag{5.18}$$

In general, this function is expected to be accurately measured from the galaxy autocorrelation signal alone.

In this chapter, we employ a top-hat selection function for W_{gal}, which assumes no uncertainty in galaxy redshift estimates. This choice mimics a spectroscopic galaxy survey or a general redshift survey with negligible uncertainties. As an example, another choice commonly found in the literature is a Gaussian distribution $\mathcal{N}(z, \sigma_{\text{gal}})$, where σ_{gal} should be much larger than the expected redshift uncertainty for each individual galaxy.

By combining the distribution of GW sources and galaxies one can construct a cross-correlation map. In our formalism, we write the cross-correlation between a GW bin i and a galaxy bin j (fully specified by their respective window functions) as:

$$
\begin{aligned}
C_\times^{ij}(\ell) = \int_0^\infty dz \frac{H(z)}{\chi^2(z)} W_{\text{GW}}^i(z) W_{\text{gal}}^j(z) \\
\times\, b_{\text{GW}}(z) b_{\text{gal}}(z) P\left(\frac{\ell + 1/2}{\chi(z)}, z\right).
\end{aligned}
\tag{5.19}
$$

We conclude this section by pointing out that the power spectra in equation (5.17), (5.13) and (5.19) do not include relativistic terms and do not capture the effects of evolution and lensing bias (see, e.g., (Scelfo et al., 2018; Scelfo, Boco, Lapi, & Viel, 2020) for a detailed treatment). Specifically, while the effects of lensing are expected to be negligible at the redshifts considered here (Oguri, 2016; Contigiani, 2020; S. Mukherjee et al., 2020b), the same is not true for relativistic effects. Therefore, we choose not to consider small values of ℓ in our analysis since the signal at these large angular scales is dominated by them.

5.4 Reconstructing GW physics

The primary goal of the chapter is to demonstrate how to reconstruct the properties of GW propagation and source clustering as a function of redshift. We do so by showing how to recover an assumed fiducial model by using mock angular power spectra with cosmic-variance or shot-noise limited uncertainties.
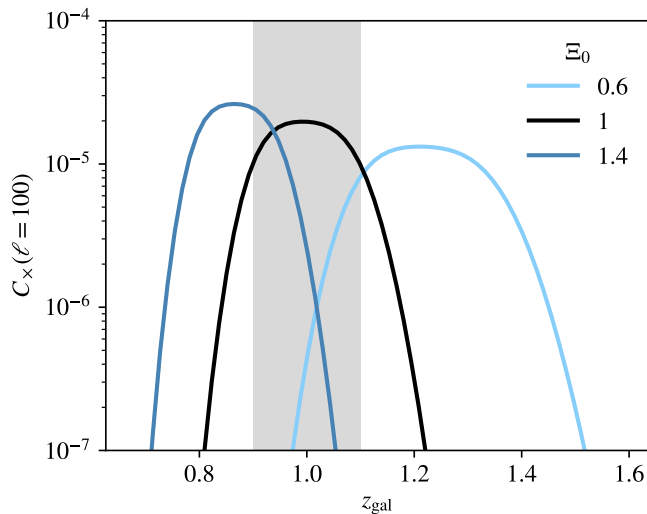

Figure 5.1: The cross-correlation signal between GW sources at $z = [0.9, 1.1]$ (shaded area) and galaxies at different redshifts (z_{gal}). If the luminosity distance ratio $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ in equation (5.8) is different from its GR assumption ($\Xi_0 \neq 1$), the location of the predicted cross-correlation peak is also affected.

Our methodology hinges on the fact that by cross-correlating a GW luminosity distance bin with multiple galaxy redshift bins we can determine the redshift of the GW sources by matching the clustering properties of the two at the true redshift (Oguri, 2016; Bera, Rana, More, & Bose, 2020).

We demonstrate this idea in Figure 5.1, where we have considered GW sources located at redshift $[0.9, 1.1]$ in a GR cosmology where $D_{\mathrm{L,GW}}(z) = D_{\mathrm{L,EM}}(z)$. In this figure, we show the expected cross-correlation signal between the angular distribution of these sources and the angular distribution of galaxies located at various redshifts. As expected, in GR ($\Xi_0 = 1$) the signal peaks inside the correct redshift range (shaded area). However, as we depart from the GW luminosity distance relation, the location of this peak is affected.

5.4.1 Mock data and fiducial model

In this section, we describe the recipe used to generate the mock angular power spectra (C_{gal}, C_{GW}, and $C_{\times}$) that are fed into our reconstruction pipeline together with their error covariance matrix. When describing real data, these angular power-spectra are extracted from the autocorrelation and cross-correlation maps representing the

sky distribution of galaxies and GW sources. The recipe has three main ingredients: the details of the fiducial model, a description of the instrumental configuration and a definition of the dominant source of error.

The first ingredient is the fiducial model. Our decision in this case is based on the results of (T. Baker & Harrison, 2021), where present-day constraints on the function α_M appearing in equation (5.6) are presented. As shown in (Belgacem et al., 2019), the results of the $\alpha_M \propto a$ parametrization found in that work can be mapped to the $\Xi(z)$ function in equation (5.8). Using this transformation, we find that the 3σ upper limit roughly corresponds to

$$\Xi_0 \lesssim 1.4, \tag{5.20}$$

with $n = 1$. Thus, we assume a fiducial model with $\Xi_0^{\mathrm{fid}} = 1.4$ and $n^{\mathrm{fid}} = 1$, representing the limit of our present understanding.

The second ingredient of our forecast is the instrumental configurations. The size of our data vector is given by the number of multipoles ℓ and window functions that we include in our analysis. Since both are largely dictated by observational considerations, in this work we assume an optimistic combination of a network of three Einstein Telescopes (Maggiore et al., 2020; Hall & Evans, 2019) capable of a log-scatter in measured $D_{\mathrm{L,GW}}$ of $\sigma_{\ln D} = 0.05$, and a high-z redshift survey with large sky coverage and negligible redshift uncertainties such as SKA (Weltman et al., 2020).

The range of angular scales that we consider is limited by two factors. On small scales, large multipoles ($\ell > 100$) are excluded due to the angular resolution of about 1 degree expected for our GW detector configuration of choice (Hall & Evans, 2019). On large scales, we do not explore values of $\ell < 10$ because our modelling does not take into account the relativistic effects dominating the signal at these scales. Nevertheless, we stress that these multipoles contribute relatively little information compared to larger multipoles since they are dominated by cosmic variance.

Our window functions are distributed in the redshift range $[0.1, 3]$. We assume $N_{\mathrm{gal}} = 12$ galaxy bins equally spaced in redshift, and $N_{\mathrm{GW}} = 8$ GW luminosity distance bins equally spaced in $D_{\mathrm{L,GW}}$. We mention in particular that this choice is not completely arbitrary. The number of GW bins is motivated by forcing well-defined bins such that their width is at least three times the luminosity distance uncertainty $\sigma_{\ln D}$ that we have assumed. Furthermore, we have also verified that the exact number of galaxy bins does not dominate our results as long as $N_{\mathrm{gal}} > N_{\mathrm{GW}}$.

As for the last ingredient, we assume cosmic-variance or shot-noise limited uncertainties. In this case, we can write the covariance matrix of the auto-correlation and cross-correlation signals defined in equation (5.13), (5.17) and (5.19) as the following:

$$\mathrm{Cov}\left[C^{ij}(\ell)C^{mn}(\ell')\right] = \frac{\delta_{\ell\ell'}}{(2\ell+1)f_{\mathrm{sky}}} \times \left(\tilde{C}^{im}\tilde{C}^{jn} + \tilde{C}^{in}\tilde{C}^{jm}\right), \tag{5.21}$$

where the indices i, j, m, n can represent both galaxy or gravitational wave bins. The terms $\tilde{C}^{im}$ contain the shot-noise contribution when they represent the autocorrela-

tions in the same bin:

$$\tilde{C}^{im}(\ell) = C^{im}(\ell) + \frac{\delta_{im}}{\bar{n}}, \tag{5.22}$$

where $\bar{n}$ is the average density of projected objects from equation (5.10). In this work, we assume a survey covering a sky fraction equal to $f_{\rm sky} = 0.5$.

Let us point out that we do not use the cross-bin correlations for the bins of the same type as a signal. However, we properly take into account the $\tilde{C}'^{im}$ terms for overlaps between two different GW bins. Similar terms for galaxy bins are completely negligible as there are no overlaps between the spectroscopic redshift bins.

For the setup described in this section, we find a total SNR of the GW-gal and GW-GW angular power-spectra of ~ 37. This value is dominated by the GW-gal cross-correlations since the GW-GW auto-correlations are not well measured (SNR$\lesssim 6$).

To generalize our choices, in subsection 5.5.2 we expand on how different combinations of instrumental specifications can affect the precision of the reconstruction.

5.4.2   Reconstruction Pipeline

In this section, we describe how the mock data presented in the previous section can be used to reconstruct $D_{\rm L,GW}$ and $b_{\rm GW}$ as a function of z. For ease of interpretation and visualization, in our analysis we do not fit these functions directly, but instead focus on the ratios $D_{\rm L,GW}/D_{\rm L,EM}(z)$ and $b_{\rm GW}/b_{\rm gal}(z)$. We point out that we do not marginalize over different possibilities for $b_{\rm gal}(z)$. This is because we are not interested in exploring the properties of the galaxy population, which are expected to be very well constrained by the galaxy-galaxy correlation signal alone.

As emphasized earlier, our approach makes use of a GP regression. This method is often used when the function of interest is directly measured at certain redshifts. These measurements are used as a training sample of the GP model, which then can predict the values of the reconstructed function at redshifts lacking any direct measurements (Belgacem, Foffa, Maggiore, & Yang, 2020; Renzi et al., 2021; P. Mukherjee & Mukherjee, 2021; Levi Said, Mifsud, Sultana, & Zarb Adami, 2021; Perenon et al., 2021).

However, this is not directly applicable for our current problem as neither the $D_{\rm L,GW}/D_{\rm L,EM}(z)$ nor the $b_{\rm GW}/b_{\rm gal}(z)$ are directly observable. Instead, these functions determine the auto- and cross-angular power spectra, which constitute our direct observables. In order to use GPs for our problem, we consider a certain number of redshift nodes for the two functions, referred to as *training nodes* with a slight abuse of terminology. The amplitudes of the nodes are free and, given a node configuration, we consider GPs which pass through all of these nodes exactly. To render our scenario computationally feasible and not consider many functions for each node configuration, we use the GPs regressor of the `python` package `sklearn` to output the best fit and use this as our function.

Our use of GPs can be thought of as a binning of the functions of interest in redshift space, and imposing a certain prior correlations between the bins. These correlations are specified by the GP kernel function, which in our case is chosen to
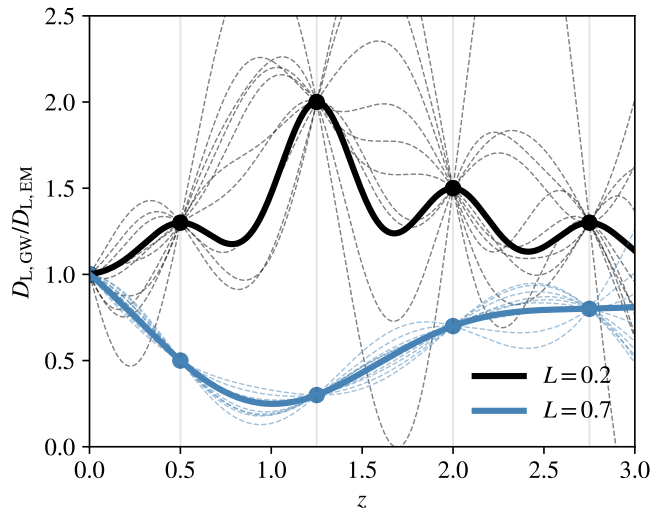
Figure 5.2: Example of GP reconstruction. The function $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ is constructed using 4 nodes at fixed redshifts (filled black dots). By varying the amplitudes of the nodes and the correlation length (L), it is possible to obtain different functional forms. The dashed lines represent a few possible realizations, while the thick lines represents the GP best fit from `sklearn` used in our sampling. Larger correlation lengths produce smoother lines.

be

$$\kappa(z_i, z_j;\, L) \propto \exp\left\{-\frac{1}{2}\left(\frac{|z_i - z_j|}{L}\right)^2\right\}, \qquad (5.23)$$

where L is the so-called correlation length. This kernel is flexible enough for our purposes, and we do not expect the detailed choice to have any significant impact on our results. For computational purposes, we generate the GPs using a baseline around $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z) = 1$. This baseline makes the GPs reconstruction to efficiently return to $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z) = 1$ when not pushed towards other values by the training nodes.

This process is described in Figure 5.2 for two values of correlation length L. While pictured in this example, for physical reasons in our analysis we exclude negative-valued functions when exploring $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$, and also non-monotonic realizations of $D_{\mathrm{L,GW}}(z)$.

The goal of our statistical analysis is to explore the possible constraints on the shape of both $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$. For that, we aim to sample the posterior distributions of the amplitudes of the nodes as well as the cosmological parameters. The correlation length L can in principle be fixed based on theoretical priors. Lacking such priors in our case, we only impose a wide uniform prior on L and consider it as a free parameter (see Table 5.1).

Each step of the sampling process consists of producing two GP curves – one for each $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$ – given the current set of training node

Parameter	Prior
Node amplitudes	[0, 11] (Uniform)
Correlation length (L)	[1, 10] (Uniform)
Ω_m	1% (Gaussian)
h	1% (Gaussian)
w_0	5% (Gaussian)

Table 5.1: Summary of the priors imposed before reconstructing $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$ and $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ using 4 nodes each. The GP hyper-parameters (i.e., the 2 correlation lengths and the 4×2 amplitudes) are explored independently. The fiducial model is given by $\Xi_0 = 1.4, n = 1, \Omega_{\mathrm{m}} = 0.31, h = 0.67, w_0 = -1$.

amplitudes. The curves are used in the calculation of theoretical auto- and cross-correlation power spectra described in section 5.3. The theoretical power spectra enter the Gaussian likelihood together with the generated mock data. The theoretical predictions are computed using our `python` code which is interfaced with the `emcee` sampler. The typical runs with varying cosmology take approximately 10 hours on a modern machine.

After obtaining the posterior distribution of the nodes, we reverse-engineer the problem to obtain confidence contours for each $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$. For all the sampled node amplitudes we generate the corresponding GP profiles on finite but sufficiently many redshift points and calculate the 68% and 95% confidence intervals at each redshift using the statistical `python` package `GetDist`.

To assess the impact of different cosmological backgrounds and clustering properties, we include in our reconstruction three nuisance parameters: the dark matter abundance Ω_{m}, the Hubble constant H_0 and the dark energy equation of state parameter w_0. A summary of our model parameters and priors used in this work is presented in Table 5.1.

We would like to point out that upcoming galaxy surveys will measure these parameters with very high precision. While GWs alone might be able to provide competitive constraints, the focus of our analysis is not in constraining them. Rather, we would like to quantify how accurately the GW luminosity distance and source properties can be measured. These properties are not accessible to generic redshift surveys and can only be measured with the use of GW-specific observables. This reasoning justifies our tight Gaussian priors on the aforementioned cosmological parameters.

5.5 Results

5.5.1 Reconstructions

In our reconstructions, we always impose the $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ to become unity at redshift zero by placing a fixed node at $z = 0$ with an amplitude of 1. Besides this fixed node, we have 4 nodes for each of the reconstructed functions. We have arrived at this number by gradually increasing the number of nodes and monitoring
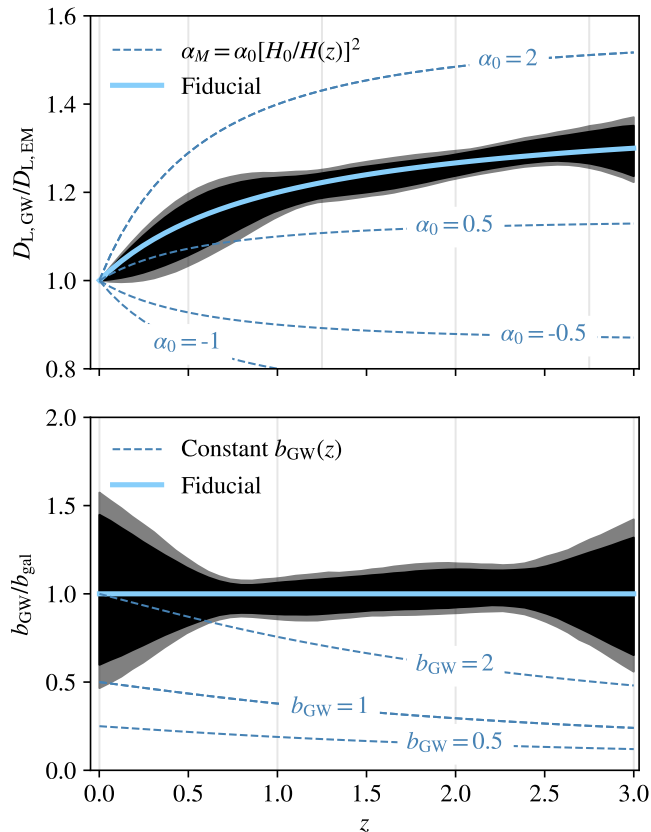
Figure 5.3: Confidence intervals (68%, in black, and 95%, in lighter grey) of the jointly reconstructed functions $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$. Together with the assumed fiducial model, we also plot the expectation for different models (see text for more details). The vertical lines mark the fixed location of the nodes used in the GP reconstruction.

the goodness of fit. In practice, we have monitored the AIC information criterion for 1-, 2- and 3- node setups for the GW luminosity distance. Our experiments suggest that, as expected, the 1-node configuration is significantly worse than the presented 4-node setup. The 2- and 3- node setups have similar performance compared to the 4-node case, but the latter still outperforms the former two. We then use the same number of nodes for the bias reconstruction.

In our main analysis, the redshift positions z_i of the nodes are fixed. We have, however, performed an experimental run to assess the impact of letting them free in reduced uniform prior ranges. The result of this experimentation is that the node locations remain unconstrained, and the final posterior of e.g. $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ does not change when the node redshift locations are being sampled as free parameters. This implies that the exact locations of the GP notes are unimportant given they are uniformly distributed in the redshift range of interest.

It is worth emphasizing that when applying our methodology to real data, the number of nodes, as well as their exact redshift placements, should be constrained

by performing similar, but more systematic experiments. Particularly, more accurate measures, such as the Bayesian evidence ratios, should be employed. Also, if enough data is used, some possible constraints could be found by letting the training nodes be completely free in the entire redshift range of interest. As the resulting posteriors are expected to be multimodal, this should be investigated using *nested sampling* algorithms.

As mentioned earlier, we also explore the GP correlation lengths both for the bias and the luminosity distance. On physical grounds, we are interested in smooth GP functions and have imposed the minimum of the uniform prior range for L to be of the order of the inter-node distance so that the smoothness is maintained. We find that, as expected, both of the correlation lengths remain unconstrained within the imposed wide prior ranges.

The results of our joint reconstruction is presented in Figure 5.3. In the same Figure, we also compare these constraints to different theoretical models. In the case of $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$, we use the parametrization $\alpha_M(z) = \alpha_0 \left[\frac{H_0}{H(z)}\right]^2$ where we use the equation (5.3) with $w_0 = -1$ to obtain the plotted lines (Belgacem et al., 2019). On the other hand, for $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$ we plot the lines corresponding to constant values of $b_{\mathrm{GW}}(z)$, while keeping the galaxy bias fixed to the expression in equation (5.18).

In principle, the output of our sampling can also be used to reconstruct the function $\alpha_M(z)$ by calculating the numerical derivative of $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$. For the purposes of this chapter, however, we chose not to do this. The kernel in equation (5.23) can be interpreted as a smoothness prior and the value of $\alpha_M(z)$ is directly affected by it. Because of this, if one is interested in inferring $\alpha_M(z)$, GPs should be used to sample this function directly.

As expected, we observe how the fiducial models for both $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ and $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$ are well encoded within the reconstructed confidence contours in both panels of Figure 5.3. The constraints at higher redshift ($z \approx 3$) for both reconstructions are broader. This is an effect that could not be seen if a parametric function was used for $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$, for instance, as the parametrization would have fixed the behaviour similarly at low and higher redshifts. Finally, in Figure 5.4 we show the correlation between these functions. We observe weak but non-zero correlations between the GW bias and the luminosity distance. In general, parametric models might induce non-physical correlations. GPs are expected to behave better in this regard, but they can still induce spurious correlations due to finite correlation lengths. For consistency, we have also performed a reconstruction using a redshift binning approach and concluded that we can recover a very similar correlation structure with both of the methods.

5.5.2 Signal-to-Noise scaling

The constraining power of our method crucially depends on a number of observational specifications. The most relevant parameters are 1) the angular sensitivity, specified by the maximum multipole $\ell_{\max}$ of the angular power spectra; 2) the number of GW sources, which is specified by the comoving number density n_{GW}; and 3) the
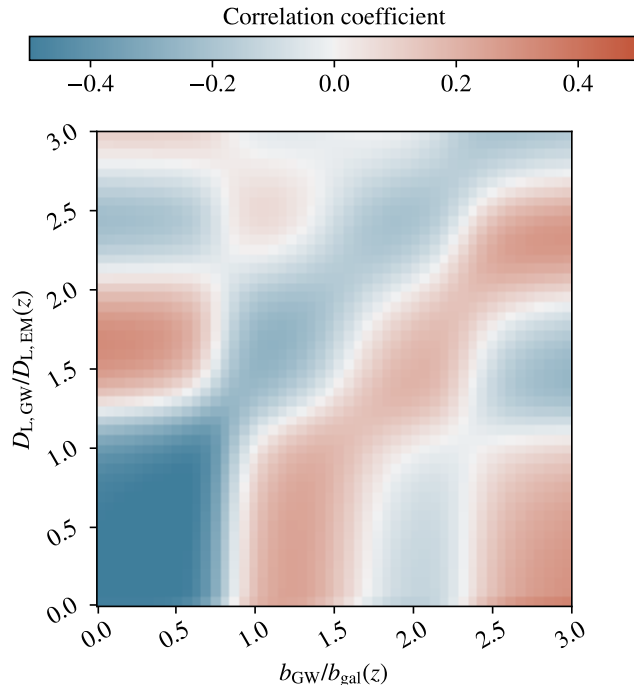
Figure 5.4: Correlation matrix for the reconstructed functions $b_{\mathrm{GW}}/b_{\mathrm{gal}}\left(z\right)$ and $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}\left(z\right)$. The figure displays how mild, but non trivial correlations can arise from a joint reconstruction.

precision of the GW luminosity distance measurements $\sigma_{\ln D}$. In the case of n_{GW}, we adjust the value of n_0 in equation (5.9) as a way to explore different values of the total number of observed GW events, $N = 4\pi f_{\mathrm{sky}}\bar{n}_{\mathrm{GW}}$. This, in principle, should include selection effects not captured by our formalism. Obviously, for a given experimental configuration the mentioned three variables are not independent, but it is still interesting to find the dependence of our results on each one of them separately. This allows us to reach conclusions without relying on specific experiments, and to suggest potential design guidelines for future GW detectors.

To attain such insights, in this subsection we consider constraints on the parametric expression in equation (5.8), as well as the parametric GW bias given by equation (5.15). For simplicity, we fix $n = 1$ and only constrain the parameter Ξ_0.

When varying ℓ_{max} and N we keep the rest of the configuration (including the luminosity distance binning) fixed. Each case of $\sigma_{\ln D}$, on the other hand, is accompanied by an adjustment in the number of luminosity distance bins. This is done to be consistent with our binning strategy, namely that the luminosity distance width of each bin is at least $\mathcal{O}(3)$ times wider than $\sigma_{\ln D}$.

Our results are summarized in Figure 5.5, where we plot the anticipated uncertainties in Ξ_0 (upper panel) and b_0 (lower panel) as a function of the SNR of the cross-correlation in equation (5.19).
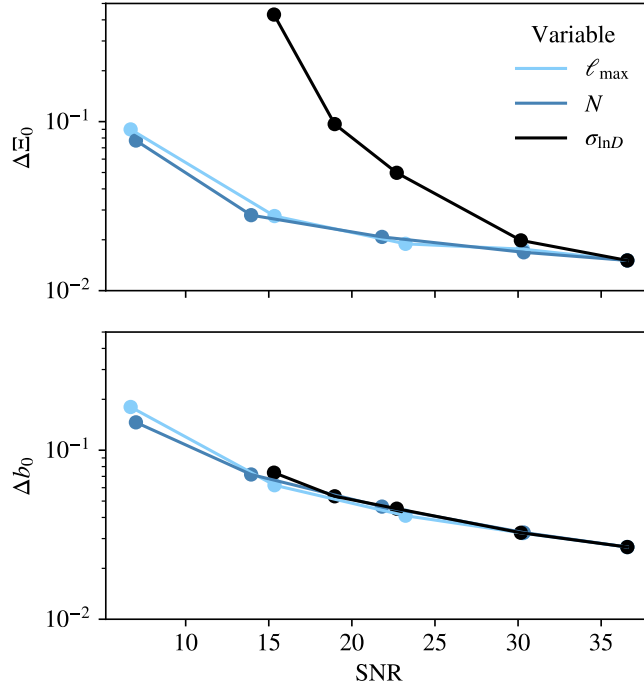
Figure 5.5: Scaling of the observed constraints with the cross-correlation SNR. Using a parametrized model for $b_{\mathrm{GW}}/b_{\mathrm{gal}}(z)$ and $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ we explore the constraining power of our method as a function of the number of observed GW sources $N = [0.7, 4, 7, 13, 20] \times 10^4$, angular resolution $\ell_{\max} = [20, 40, 60, 80, 100]$ and luminosity distance uncertainty $\sigma_{\ln D} = [0.5, 0.3, 0.2, 0.075, 0.05]$. As visible from the figure, the data SNR completely captures the effect on the observed uncertainties Δb_0 and $\Delta \Xi_0$ in the first two cases. In the case of $\sigma_{\ln D}$, we observe that the increase in constraining power for $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$ is steeper due to the larger number of window functions that we can build to sample $D_{\mathrm{L,GW}}(z)$.

For a fixed $\sigma_{\ln D}$ the constraining power on Ξ_0 and b_0 is almost completely determined by the cross-correlation SNR. This fact suggests that no matter how the given SNR is realized (either by increasing the number of sources or by improving the angular sensitivity), the expected constraints will be the same. This implies that the results presented in this chapter can be easily scaled to different configurations. Unsurprisingly, we find that the constraints scale as $1/\mathrm{SNR}$.

The situation is somewhat different for the case of varying $\sigma_{\ln D}$ (and the number of luminosity distance bins). The constraints on the bias still follow the same form (see the lower panel), but the scaling of the Ξ_0 constraints, on the other hand, is much steeper than in the cases of varying $\ell_{\max}$ and n_{GW}, roughly $1/\mathrm{SNR}^3$. This fact can be qualitatively understood by remembering the importance of the relative positions of GW and galaxy window functions demonstrated in Figure 5.1. Sampling this relation with a larger number of window functions increases the precision of our reconstruction.

The results presented in this section quantify the importance of accurate luminosity distance measurements and demonstrate the benefit that smaller values of $\sigma_{\ln D}$ can bring to a binned approach.

5.6   Discussion and Conclusions

In this chapter, we have shown how the combination of GW observations and redshift surveys can be exploited in the era of GW cosmology. We have identified two essential quantities that characterize this new research field: 1) the luminosity distance relation, a clear imprint of modifications to the propagation of tensor modes (see section 5.2), and 2) the bias of the sources, regulating their spatial distribution and betraying their origin.

Proposed GW detectors such as the Einstein Telescope (Maggiore et al., 2020) or Cosmic Explorer (Reitze et al., 2019) are expected to probe a sizeable fraction of the visible Universe and produce large statistical samples. In this context, we point out that the number of sources assumed for our main analysis, 2×10^5, is particularly conservative and differs from expectations by at least an order of magnitude (Maggiore et al., 2020). This difference is primarily due to our assumption of a constant comoving density of events. While we do not explore other assumptions for the distribution of GW sources over cosmic time, we have investigated similar effects in subsection 5.5.2, where we have shown how our results can be rescaled to other instrument configurations or number of observed sources.

Our formalism, based on binned angular power-spectra and sky maps, is optimal for a large number of sources with no known counterpart. Its main advantages are related to the simple modelling of the theoretical signals and their data covariance matrix. Because no reconstruction of the underlying density field is necessary, the predictions display a clear separation of scales. For example, the angular scales that we have considered hare are all within the linear regime ($k \lesssim 0.1$ Mpc^{-1}). Furthermore, because this formalism is well established, our shot-noise limited covariance matrix can be easily generalized to include additional sources of (co-)variance.

Although a comprehensive comparison between multiple approaches is outside the scope of this work, it is worth discussing how our results compare to others found in the literature. We preface this by saying that one-to-one comparisons, however, are often complicated either by significantly different assumptions or the impossibility of directly translating these assumptions from one prescription to another. Despite this, here we draw a parallel between our method and two other methods.

The first method is the one used in (S. Mukherjee, Wandelt, & Silk, 2021), which has also been shown to be extremely successful in measuring both $b_{\mathrm{GW}}(z)$ and $D_{\mathrm{L,GW}}(z)$ using parametric models. Similarly to this work, the information is also extracted from the cross-correlation with redshift sources, but no binning of the GW data is performed. In this case, we have verified that such methods perform significantly better than our map-based approach in the case of a low number density of GW sources and large uncertainty in the measured $D_{\mathrm{L,EM}}(z)$. These features, in particular, make it especially useful for near-future samples of a few tens of objects.

5.6. DISCUSSION AND CONCLUSIONS

The second promising method to measure $D_{\mathrm{L,GW}}(z)$ that has been proposed in the literature is offered by GW sources with known counterparts. Such observations give direct access to $D_{\mathrm{L,GW}}$ as a function of redshift and can be combined with similar measurements in the EM spectrum to obtain $D_{\mathrm{L,GW}}/D_{\mathrm{L,EM}}(z)$. The analysis of (Belgacem et al., 2020) is based on this methodology and, similarly to ours, also employs GPs to reconstruct this ratio from an Einstein Telescope sample with $\sim 10^2$ sources.

Ultimately, we expect this counterpart-based formalism and the one described in this work to be complementary: a direct measure of $D_{\mathrm{L,GW}}(z)$ can be used to break the degeneracy between bias and luminosity distance shown in Figure 5.4. However, because the fraction of events with known counterparts that will be observed is heavily dependent on both the GW source distribution and multiple instrumental setups, we do not attempt to combine the two methods here.

In conclusion, the combination of GW resolved events and the clustering of galaxies is expected to improve our current knowledge of the physical properties of the Universe. Our work shows how to reconstruct these properties as a function of redshift in a generic way, and highlights the need for accurate and precise measurements of $D_{\mathrm{L,GW}}$. This will require control over the instrument calibration uncertainties (Cahillane et al., 2017), but also the degeneracy between the inclination of the source and its luminosity distance (Ghosh, Del Pozzo, & Ajith, 2016). In the future, we aim to apply our current analysis pipeline to the next generation of large scale structure surveys and incoming GW observations.