

Large scale structure prior knowledge in the dark siren method

Charles Dalang ^{a,b,*} Bartolomeo Fiorini ^a and Tessa Baker ^a

^a*School of Mathematical Sciences, Queen Mary University of London,
Mile End Road, E1 4NS, London, United Kingdom*

^b*Institute of Cosmology and Gravitation, University of Portsmouth,
Burnaby Road, Portsmouth PO1 3FX, United Kingdom*

E-mail: charles.dalang@ens.fr, bartolomeo.fiorini@doriansrl.it,
tessa.baker@port.ac.uk

ABSTRACT: Gravitational wave dark sirens are a powerful tool for cosmology and inference of compact object population hyperparameters. They allow for a measurement of the luminosity distance to the source, but not their redshift. Galaxy catalogues in the source localization volume can be used to infer the redshift of the source in a statistical manner. Catalogues are, however, limited by their incompleteness, which can be significant at redshifts corresponding to current GW events. In this work, we detail how to implement in practice variance completion, a novel galaxy completion method which uses knowledge of the large scale structure to optimize the potential of dark sirens analyses. We compress the prediction for the missing number of galaxies into a ratio between the predictions of variance completion and the standard homogeneous completion method. This ratio format can be easily incorporated into existing line of sight computations used in dark sirens software; we demonstrate this procedure using the GLADE+ galaxy catalogue and the `gwcs` software package. We discuss the robustness of the method, and apply it to well-localized event GW190814 as a proof of concept. Finally, we apply the method to data from the third observing run of LIGO-Virgo-KAGRA, finding that it yields results that are consistent with homogeneous completion. We also discuss the prospects for an improvement if the GW localization volume shrinks.

KEYWORDS: cosmological parameters from LSS, gravitational waves / theory

ARXIV EPRINT: [2410.03275](https://arxiv.org/abs/2410.03275)

*Corresponding author.

Contents

1	Introduction	1
2	Methodology	3
2.1	Line of sight redshift prior	3
2.2	Discretization	4
2.3	Completeness fraction	5
2.4	Homogeneous completion	7
2.5	Variance completion	8
2.6	Average number density and standard deviation	9
2.7	Implementation in <code>gwcosmo</code>	12
2.8	Cosmology dependence of the line of sight redshift priors	14
3	Results	16
3.1	Updated line of sight redshift priors	17
3.2	Dark siren analysis of the Hubble constant for GW190814	17
3.3	Application to GWTC-3	18
3.4	Discussion	21
4	Conclusion	24
A	Approximate matter power spectrum	25

1 Introduction

After centuries of astronomy based on the electromagnetic spectrum, it has now been almost a decade since the beginning of the era of gravitational wave astronomy. As we will see in this work, the topic of gravitational wave (GW) cosmology sits well at the interplay of both fields. One of the major objectives of GW cosmology is to obtain independent measurements of cosmological parameters, and to infer hyperparameters of the compact object population.

In a similar time frame, the so-called *Hubble tension* has emerged as a topic of interest. The lower value of $H_0 = (67.4 \pm 0.5)\text{km s}^{-1}\text{Mpc}^{-1}$ obtained from the cosmic microwave background (CMB) by assuming a cosmological model [1] is in tension at about $(4 - 6)\sigma$ with local measurements, which cluster around $H_0 = 74 \text{ km s}^{-1}\text{Mpc}^{-1}$ [2–5]. It has been shown to be incredibly difficult to solve the Hubble tension by invoking some form of new physics, without also disrupting constraints from the CMB, big bang nucleosynthesis and large scale structure [6, 7]. At the same time, the tension seems to be equally robust to all known sources of systematics [8]. GWs offer a promising avenue to add a completely independent measurement of the Hubble constant [9], which is however not without challenges [10].

GW methods are attractive because the luminosity distance is directly encoded in the waveform of compact binary coalescences. The tricky step in inferring a Hubble constant in this scenario is to extract the redshift of the source. In rare cases, the GW source can also emit

some electromagnetic radiation, which can be used to pinpoint the host galaxy and extract the redshift spectroscopically. These systems are called bright sirens and one event has been detected so far, leading to a Hubble constant measurement of $H_0 = 70.0_{-8.0}^{+12.0} \text{km s}^{-1} \text{Mpc}^{-1}$ sitting nicely in the middle of the two H_0 schools [11]. In the future, we hope to get more of these bright sirens [12], which also requires to model electromagnetic selection effects [13]. Meanwhile, there are about 100 publicly available dark sirens, i.e. those without electromagnetic counterparts, which can be used to study cosmology using different statistical techniques. On one hand, observations of the redshifted masses, and knowledge of features of the source frame distribution of compact object masses can be used to extract the redshift of the source in a statistical way [14–21]. On the other hand, a galaxy catalogue overlapping the GW localization volume of the source can be used to reconstruct the Hubble constant statistically [22–24]. Both of these methods can now be used simultaneously to infer the Hubble constant and compact object population hyperparameters jointly, using the publicly available codes `gwcosmo`, `icarogw` and `CHIMERA` [24–26].

A major difficulty in using galaxy catalogues to infer the redshift of dark sirens is related to the completeness of galaxy surveys. As surveys are often magnitude-limited, the completeness drops with redshift; for many currently available surveys, a low level of completeness is reached whilst still at redshifts of interest for GW science. Moreover, completeness is highly anisotropic due to the different footprints of surveys, taken with telescopes with various different properties.

Several methods have been introduced in the literature to mitigate incompleteness, including homogeneous, multiplicative and variance completion [27–29]. Homogeneous completion is the simplest technique, and the method currently implemented in `gwcosmo` and `icarogw`. It assumes that the out-of-catalogue galaxies are distributed homogeneously in comoving volumes. While simple, this method ignores the tendency of gravity to cluster galaxies in a way which is well understood. To compensate for this, multiplicative completion was introduced in [28], assuming a proportionality factor between the galaxies that are in and those that are out of the catalogue. This has been shown, however, to overestimate the amplitude of structure at low completeness. Additionally, it lacks a robust physical foundation. Conversely, variance completion has been introduced to make use of the fact that galaxies are clustered in a way which can be estimated, so as to fill in missing galaxies, where we expect them to be, from large scale structure knowledge. This method was tested on simulated data in two dimensions and shown to outperform the other methods, so long as the structure of the catalogued galaxies represents sufficiently well the true distribution [29]. Note that a sophisticated completion method taking into account large scale structure and the magnitude distribution of sources is introduced in [30]. Alternatively, one can focus on events which have very good catalogue support by selecting GW events which overlap well with a volume limited sample of galaxies from DES [31] or DESI [32].

In this work we build upon the theoretical foundations of [29], detailing how to implement variance completion on real data. We implement it on the GLADE+ galaxy survey [33], as an example. We discuss a number of practical adaptations to be made, for example, to incorporate redshift uncertainties in the method. We introduce a ratio function, which can easily be incorporated into the formalism of `gwcosmo`. We discuss and marginalize over

the cosmology dependence where relevant. We analyze GW190814, a well-localized GW event with homogeneous and variance completion and discuss the differences on the Hubble constant posterior. Finally, we apply our variance-completed line of sight priors to dark siren events from the third observing run of LIGO-Virgo-KAGRA (LVK) with sufficient signal to noise ratio (SNR), and show how the method affects the posterior distribution of the Hubble constant. The method works particularly well for well-localized events, of which we expect to see more, as GW detectors improve.

2 Methodology

In this section, we describe in detail how we implement in practice the formalism of variance completion on real data. We start by describing the line of sight redshift prior in section 2.1 to introduce how the incompleteness of galaxy catalogues affects this. We then describe how to discretize the galaxy field in voxels and how we use a magnitude threshold map to identify voxels of similar completeness in section 2.2 and 2.3. We then briefly review homogeneous and variance completion in section 2.4 and 2.5 and detail how to compute large scale structure ingredients required for these methods in section 2.6. We explain how to implement the method in a code such as `gwcosmo` in section 2.7 and discuss the cosmology dependence of the method in section 2.8

2.1 Line of sight redshift prior

One required ingredient for the dark siren method, which is the essence of this work, is the line of sight redshift prior. It encodes the probability density that the GW was emitted from a particular redshift z and direction $\hat{\mathbf{n}}$ on the sky. GWs come from binary compact objects, which live inside galaxies. It is therefore clear that this probability density should depend on the galaxy field. In principle, we would like to weight each galaxy according to their probability to host a GW event. Unfortunately, this quantity is difficult (if not impossible) to model. Alternatively, we can expect this quantity to be correlated with some galaxy indicators, such as luminosity, mass or star formation rate. For the completion method that we implement in this work, we work with luminosity weighting for all galaxies. We will then show how to combine this completion method with catalogues which use some form of galaxy weighting for the galaxies in the catalogue in section 2.7.

One can write that the probability density that a GW was emitted from redshift z and direction $\hat{\mathbf{n}}$ is proportional to the weighted galaxy number density

$$p(z, \hat{\mathbf{n}}) = \frac{\frac{dn_g}{dzd\Omega}(z, \hat{\mathbf{n}})}{\int dzd\Omega \frac{dn_g}{dzd\Omega}(z, \hat{\mathbf{n}})}, \quad (2.1)$$

where $dn_g/(dzd\Omega)(z, \hat{\mathbf{n}})$ denotes the galaxy number density per redshift bin dz and solid angle $d\Omega$. The total number of galaxies in the volume under consideration reads

$$N_g = \int dzd\Omega \frac{dn_g}{dzd\Omega}(z, \hat{\mathbf{n}}) \quad (2.2)$$

One should distinguish between the galaxies which are observed in the catalogue n_c from those which are missed n_m . Those should sum up to the total number of galaxies $n_g = n_c + n_m$

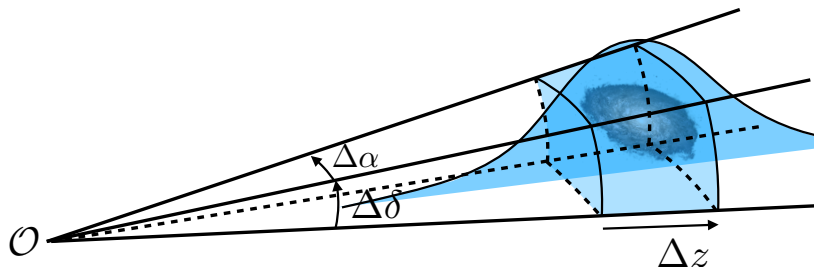


Figure 1. We sketch the contribution of 1 galaxy to a voxel. A galaxy can contribute a fraction of itself to several voxels. This is typically the case for photometric redshift bins, while spectroscopic redshift contribute in general to only 1 voxel.

on any volume $d\Omega dz$. Therefore,

$$p(z, \hat{\mathbf{n}}) = p_c(z, \hat{\mathbf{n}}) + p_m(z, \hat{\mathbf{n}}), \quad (2.3)$$

with

$$p_c(z, \hat{\mathbf{n}}) = \frac{1}{N_g} \frac{dn_c}{dzd\Omega}(z, \hat{\mathbf{n}}), \quad (2.4)$$

$$p_m(z, \hat{\mathbf{n}}) = \frac{1}{N_g} \frac{dn_m}{dzd\Omega}(z, \hat{\mathbf{n}}), \quad (2.5)$$

The galaxies in the catalogue are well localized in direction but less so in redshift space, especially for photometric galaxies. Hence, we include the uncertainty on redshift but neglect the uncertainty on direction. The probability density can be expressed as a sum over the N_c catalogued galaxies

$$p_c(z, \hat{\mathbf{n}}) = \frac{1}{N_g} \sum_{i=1}^{N_c} p(z; \bar{z}^i, \sigma^i) \delta^2(\hat{\mathbf{n}} - \hat{\mathbf{n}}^i). \quad (2.6)$$

For the out-of-catalogue galaxies in eq. (2.5), we will estimate $\frac{dn_m}{dzd\Omega}$ using homogeneous or variance completion, that we describe in section 2.4 and 2.5.

2.2 Discretization

For the purpose of implementing variance completion, it will prove useful to split the galaxies into voxels, i.e. the 3 dimensional counterpart of pixels. The number of galaxies in a volume $\mathcal{V}_{\text{vox}}^i$, defined by its redshift and solid angle support, relate to the galaxy densities via an integral over the voxel volume $\mathcal{V}_{\text{vox}}^i$

$$n_g^i = \int_{\mathcal{V}_{\text{vox}}^i} d\Omega dz \frac{dn_g}{d\Omega dz}(z, \hat{\mathbf{n}}), \quad (2.7)$$

where the index i indicates the voxel number. A similar relation applies to n_c and n_m . We sketch the contribution of a galaxy to a voxel in figure 1.

One galaxy can spread over multiple voxels depending on its redshift uncertainty. This allows for n_c^i to capture these redshift uncertainties. When combining information from different surveys, some galaxy redshifts have widely different error bars. Some are photometric,

others are spectroscopic. We take the redshift error bars σ_z^i into account by assuming that the redshift posteriors are well modeled by Gaussian PDFs, which is not necessarily a good approximation. Hence, the number of galaxies in each voxel n_c^i reads

$$n_c^i(z, \hat{\mathbf{n}}) = \sum_j \int_{z_i - \delta z/2}^{z_i + \delta z/2} dz p(z; \bar{z}^j, \sigma_z^j), \quad (2.8)$$

where the index j runs over all catalogued galaxies in the pixel corresponding to the direction $\hat{\mathbf{n}}$ and $p(z; \bar{z}^j, \sigma_z^j)$ are normal distributions centered on \bar{z}^j with standard deviation σ_z^j . Therefore, the number of galaxies in each voxel does not need to be an integer.

We shall fix the angular size of voxels throughout each analysis. We shall also fix redshift bins δz implying that voxels change volume with redshift. The comoving volume of voxels at redshift z reads

$$\mathcal{V}_{\text{vox}}(z) = \int_{z - \delta z/2}^{z + \delta z/2} dz \frac{dV_c}{dz}(z), \quad (2.9)$$

where $dV_c/(dzd\Omega) = cr^2(z)/H(z)$, such that $r(z)$ is the comoving distance and $H(z)$ the Hubble rate at redshift z . Those are given, for a flat Λ CDM cosmology by

$$r(z) = \int_0^z \frac{dz' c}{H(z')}, \quad (2.10)$$

and

$$H(z) = H_0 \sqrt{\Omega_{m0}(1+z)^3 + (1 - \Omega_{m0})}, \quad (2.11)$$

where H_0 indicates the present day Hubble constant and Ω_{m0} denotes the present day energy fraction of matter in the Universe. This is the first instance in the method, which requires a cosmological model. We discuss the cosmology dependence of the entire procedure in section 2.8. For a fixed pixel size, two effects affect the volume of a voxel as redshift grows. The depth of the voxel decreases slightly, while the width perpendicular to the line of sight grows. These two effects combine to a net increase in the voxel volume as redshift grows.

We have identified a way to divide the galaxy field into voxels taking into account galaxy redshift uncertainties. In the following section, we discuss the completeness fraction as a way to identify voxels of similar completeness levels, which is a required ingredient of variance completion.

2.3 Completeness fraction

For homogeneous and variance completion, it is necessary to estimate the completeness fraction of galaxies as a function of direction and redshift. For homogeneous completion, it allows to estimate the amplitude of the out-of-catalogue contribution. For variance completion, it allows to determine voxels of similar completeness, on which the algorithm can apply.

For a given galaxy catalogue, a magnitude threshold map $m_{\text{th}}(\hat{\mathbf{n}})$ can be determined to establish groups of pixels with similar levels of completeness at each redshift [24]. We call these groups, *classes*. The estimate of the magnitude threshold is based on the median apparent magnitude of galaxies in the relevant (B- or K- band) band and returns $-\infty$ for

empty pixels, i.e. those which have less galaxies than a certain threshold, fixed to 10 galaxies. This allows the magnitude threshold map to be relatively insensitive to outliers. The threshold of 10 galaxies per pixel prevents an estimate of an imprecise magnitude threshold from the median of the apparent magnitude from a too small sample of galaxies, which may propagate to the completeness. We order pixels as a function of $m_{\text{th}}(\hat{\mathbf{n}})$ and split them in classes of 1024 voxels. For $n_{\text{side}} = 32$, which divides the sphere in 12,288 equal area pixels of about 0.0010 sr, we count 12 classes of 1024 `healpix` pixels [34, 35].¹ We shall apply the variance completion algorithm on these different classes for the B-band. The first class ordered by completeness levels has a few pixels which are observed very deeply. This results in significant variations in the completeness levels among the pixels of this class and we cannot trust the results of variance completion. Likewise the last four classes ordered by completeness levels have significant variations in the magnitude threshold map. They span mostly the regions close to the Milky way and we cannot trust the results of variance completion if the completeness varies a lot from pixel to pixel in the same class. We discard these 5 classes to which we apply homogeneous completion. We note that for the 7 other classes, the difference between the maximum and minimum $m_{\text{th}}(\hat{\mathbf{n}})$ is smaller than 0.25 mag. In figure 2, we plot the magnitude threshold map. We also show in figure 3 a map of the 12 classes of pixels with similar completeness. The completeness is calculated as

$$\hat{f}_S = \frac{\int_{L(m_{\text{th}}(\hat{\mathbf{n}}))}^{L_{\text{max}}} dL L\phi(L)}{\int_{L_{\text{min}}}^{L_{\text{max}}} dL L\phi(L)}, \quad (2.12)$$

where the luminosity function is assumed to be a Schechter function [36]

$$\phi(L) = \frac{\phi_*}{L_*} \left(\frac{L}{L_*}\right)^\alpha \exp(-L/L_*), \quad (2.13)$$

with K-band values $L_{*,\text{K}}$, $L_{\text{min},\text{K}}$ and $L_{\text{max},\text{K}}$ calculated from $M_{*,\text{K}} = -23.39$, $M_{\text{min},\text{K}} = -27$ and $M_{\text{max},\text{K}} = -19$ and $\alpha_K = -1.09$ [37]. For the B-band, we use $M_{*,\text{B}} = -19.66$, $M_{\text{min},\text{B}} = -22$ and $M_{\text{max},\text{B}} = -16.5$ and $\alpha_B = -1.21$ [38]. The conversion from apparent magnitude m to absolute magnitude M is done via

$$M(m) = m - 5 \log_{10}(d_L(z)) - 25 - k(z), \quad (2.14)$$

with k -correction given by $k_K(z) = -6 \log_{10}(1+z)$ for $z \leq 0.25$ for the K-band² [37] and $k_B(z) = (z + 6z^2)/(1 + 15z^3)$ [38] for the B-band. The luminosity distance as a function of redshift is obtained via

$$d_L(z) = (1+z)r(z), \quad (2.15)$$

where $r(z)$ is the comoving distance as a function of redshift, given in eq. (2.10). The luminosities can be obtained from an absolute magnitude using

$$L(M) = L_0 \cdot 10^{-2M/5}, \quad (2.16)$$

¹<http://healpix.sourceforge.net>

²Note that we apply the variance completion algorithm to redshifts lower than this.

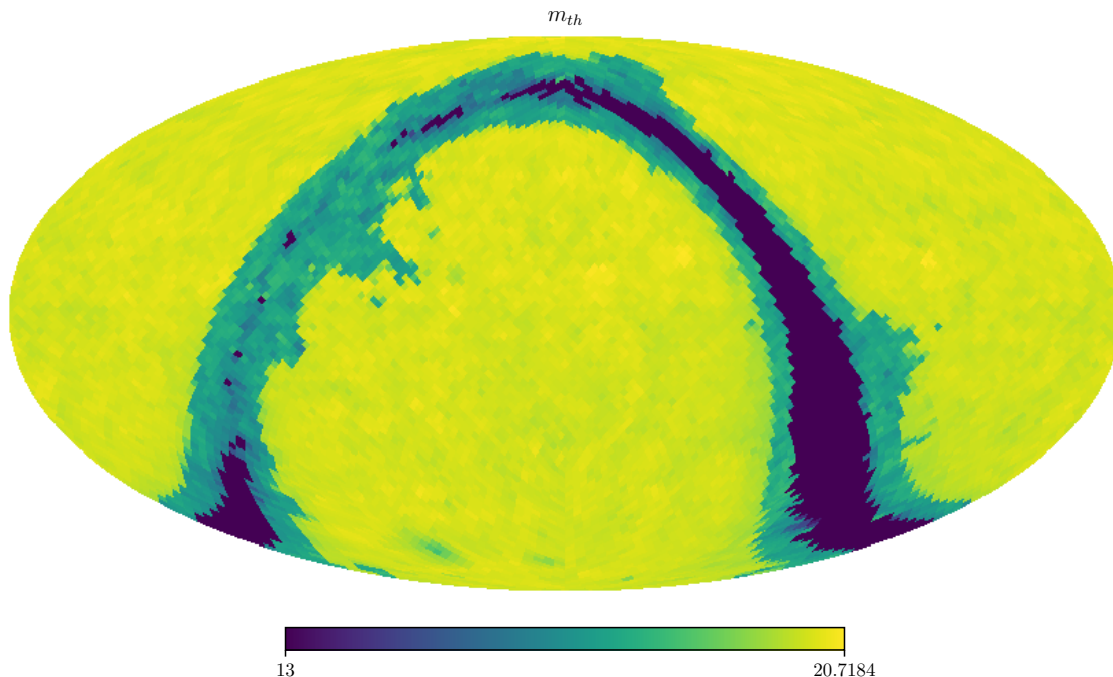


Figure 2. Here, we show the magnitude threshold map $m_{\text{th}}(\hat{\mathbf{n}})$ in the K-band for the GLADE+ galaxy catalogue for $n_{\text{side}} = 32$, from which we can compute the completeness fraction in each voxel. Pixels for which $m_{\text{th}}(\hat{\mathbf{n}}) < 13$ are plotted with the same color as $m_{\text{th}}(\hat{\mathbf{n}}) = 13$. The minimum magnitude threshold determined is 10.378, while pixels left blank have $m_{\text{th}} = -\infty$. The higher the magnitude threshold, the better we estimate a pixel to have been observed.

with $L_0 = 3.0128 \cdot 10^{28}$ W. One can use these relations to convert the apparent magnitude threshold to a luminosity threshold and compute the completeness fraction in each voxel from eq. (2.12). This completeness fraction is required to apply homogeneous or variance completion, as detailed in the following two sections.

2.4 Homogeneous completion

In this section, we review homogeneous completion, introduced in [27], on which variance completion builds up. In the previous section, we determined classes \mathcal{S} of pixels which have similar completeness at fixed redshift. Given a region \mathcal{S} of N_k voxels with the same estimated completeness fraction $\hat{f}_{\mathcal{S}} \neq 0$, one can estimate the number of galaxies in each voxel as

$$\hat{n}_{\text{g}} = \frac{\sum_{i=1}^{N_k} n_{\text{c}}^i}{N_k \hat{f}_{\mathcal{S}}}. \quad (2.17)$$

This comes from the expectation that over sufficiently many voxels, the completeness should satisfy $f_{\mathcal{S}} = \sum_{i=1}^{N_k} n_{\text{c}}^i / \sum_{i=1}^{N_k} n_{\text{g}}^i$. Homogeneous completion estimates the number of missing galaxies in each voxel as

$$\hat{n}_{\text{m};\text{hom}}^i = \hat{n}_{\text{g}}(1 - \hat{f}_{\mathcal{S}}) = \frac{1 - \hat{f}_{\mathcal{S}}}{\hat{f}_{\mathcal{S}}} \cdot \frac{\sum_{i=1}^{N_k} n_{\text{c}}^i}{N_k}, \quad (2.18)$$

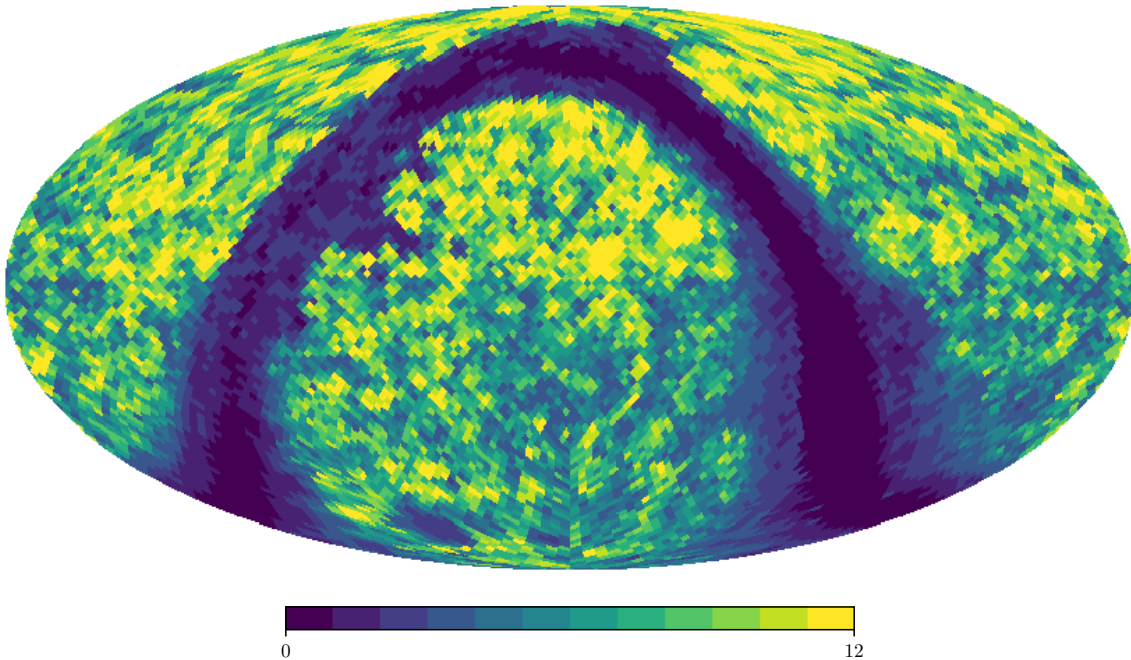


Figure 3. Here, we show a map of the 12 classes of pixels with similar estimated completeness fractions \hat{f}_S . This map is found by ranking pixels according to their magnitude threshold, which is presented in figure 2.

where the completeness fraction can be estimated from the catalogue, as discussed in the previous section. The first expression is useful to estimate the number of missing galaxies even when $\hat{f}_S = 0$, in which case, the second expression is undetermined.

The advantage of this method is that it can be computed once for all the voxels in the relevant class \mathcal{S} which have completeness \hat{f}_S . In this sense, it is very simple. We approximate the number density by dividing the number of galaxies by the corresponding redshift bin Δz^i and solid angle $\Delta\Omega^i$

$$\frac{dn_{m;\text{hom}}}{dzd\Omega}(z^i, \hat{\mathbf{n}}^i) \simeq \frac{\hat{n}_{m;\text{hom}}^i}{\Delta z^i \Delta\Omega^i}, \quad (2.19)$$

The downside of this method is that it underestimates the amplitude of structure [29]. This does not make optimal use of the information brought into dark siren analyses by galaxy catalogues. In the next section, we review variance completion, which is designed to take into account large scale structure, to complete a galaxy catalogue.

2.5 Variance completion

In ref. [29], we presented *variance completion*, a way to complete a galaxy catalogue which makes use of knowledge from large scale structure. This method is quite sophisticated on its own and the subject of an entire article [29]. Here, we give only a very crude reminder of the method and defer the details to the original paper [29].

On top of the knowledge of the mean galaxy number density \bar{n}_g , variance completion also requires knowledge of the standard deviation in the galaxy field σ_g . Given these two ingredients, determining the number density of missing galaxies n_m^i boils down to minimizing a well-chosen function. It was shown that minimizing the following function with respect to all the n_m^i worked very well [29]:

$$\mathcal{L}[n_m^1, n_m^2, \dots, n_m^{N_k}] = A_0 \sum_{i=1}^{N_k} (n_m^i)^2 + B_k \left[\left(\frac{1}{N_k} \sum_{i=1}^{N_k} (n_m^i + n_c^i)^2 \right) - (\bar{n}_g^2 + \sigma_g^2) \right]^2, \quad (2.20)$$

where $A_0, B_k \in \mathbb{R}$ are real constants. All number densities n_m^i, n_c^i are understood to belong to the same class \mathcal{S} of N_k pixels, at the same redshift z . Fixing the redshift ensures that the volume of each of the voxels involved in this sum is the same and that comparing number densities in each voxels of this class makes sense. Intuitively, the minimization of the first term ensures that only the minimum number of missing galaxies are introduced. The second term ensures that the standard deviation of the estimated galaxies $\hat{n}_g = \hat{n}_m + n_c$ respects the expectation from the standard deviation in the galaxy field σ_g .

The minimum can be found by setting the gradient of eq. (2.20) to zero, i.e. $\nabla \mathcal{L} = \mathbf{0}$. This results in a complicated set of nonlinear coupled equations, which can be linearized around an initial guess $n_m^i = \bar{n}_m^i + \delta n_m^i$. This initial guess is provided by eq. (2.18) from homogeneous completion, i.e. $\bar{n}_m^i \rightarrow n_{m;\text{hom}}^i$. The linear system can be written in matrix form as $A \cdot \delta \mathbf{n}_m = \mathbf{b}$, where the form of A and \mathbf{b} are given in [29] and depend on the initial guess \bar{n}_m^i . The linear system can then be inverted to find $\delta \mathbf{n}_m = A^{-1} \mathbf{b}$. These solutions can then be fed into a better initial guess as $\bar{n}_m^i \rightarrow \bar{n}_m^i + \delta n_m^i$. This iteration can be made until the solution converges to the desired accuracy. The values of A_0 and B_k can be calibrated such that the total number of missing galaxies in \mathcal{S} averages to the total expected number of galaxies from homogeneous completion in \mathcal{S} , i.e. $N_k \bar{n}_{m;\text{hom}}$ within 1%. Finally, the variance completion estimators are given by $\hat{n}_{m;\text{var}}^i = \bar{n}_m^i + \delta n_m^i$. We can approximate the number density by dividing by the redshift and solid angle volume of the voxels

$$\frac{dn_{m;\text{var}}}{dz d\Omega}(z^i, \hat{\mathbf{n}}^i) \simeq \frac{\hat{n}_{m;\text{var}}^i}{\Delta z^i \Delta \Omega^i} \quad (2.21)$$

In the rest of this article, we aim to implement this method in practice to reconstruct the line of sight redshift prior (given in eq. (2.3)) in a way which is compatible with `gwcs`. Furthermore, we check whether homogeneous or variance completion gives a more informative posterior on the Hubble constant. In the following, we detail how the average and standard deviation in the galaxy field can be estimated.

2.6 Average number density and standard deviation

As hinted in the previous section, variance completion requires knowledge of the standard deviation in the number of galaxies per voxel σ_g on top of the average number of galaxies per voxel \bar{n}_g , which is also an ingredient of homogeneous completion. In this subsection, we detail how they are determined.

For the average number density of galaxies, one can use the results of [39], which provide a formula for the comoving number density as a function of cosmic time, which is however,

cosmology dependent. Instead, we work with the average number density \hat{n}_g calculated from eq. (2.17), determined from the catalogue itself. As we shall see, this quantity ends up canceling when implemented in `gwcsmo`.

The variance in the galaxy number counts requires more work. It can be computed from the variance $\sigma^2(R)$ in the smoothed dark matter density contrast $\delta_s(z, \mathbf{x}, R)$, where R represents the smoothing scale.

$$\sigma^2(R) \equiv \langle \delta_s^2(z, \mathbf{x}, R) \rangle - \underbrace{\langle \delta_s(z, \mathbf{x}, R) \rangle^2}_{=0}, \quad (2.22)$$

where the smoothed density contrast reads

$$\delta_s(z, \mathbf{x}, R) = \int_{\mathbb{R}^3} d^3 \mathbf{x}' \delta_{\text{dm}}(z, \mathbf{x}') W_s(|\mathbf{x}' - \mathbf{x}|, R). \quad (2.23)$$

There, the window function $W_s(|\mathbf{x}' - \mathbf{x}|, R)$ smooths the dark matter density contrast $\delta_{\text{dm}}(z, \mathbf{x}) \equiv (\rho_{\text{dm}}(z, \mathbf{x}) - \bar{\rho}_{\text{dm}}(z))/\bar{\rho}_{\text{dm}}(z)$ on a scale R around the position \mathbf{x} . The local dark matter energy density $\rho_{\text{dm}}(z, \mathbf{x})$ averages on spacelike hypersurfaces of constant cosmic time defined by the cosmological redshift z to $\bar{\rho}_{\text{dm}}(z)$. In principle, the window function support domain should take the shape of the voxels. For simplicity, we assume that the voxels are equivalent in shape to spheres, and neglect any errors arising from not capturing their full shape. As we shall see, we only need an approximate estimate of the variance. For this purpose, we extract the redshift-dependent radius from the voxel volume $\mathcal{V}_{\text{vox}}(z) = 4\pi R^3(z)/3$, which can be calculated using eq. (2.9) and compute the corresponding window function as follows

$$W_s(x, R(z)) = \begin{cases} \frac{3}{4\pi R^3(z)}, & \text{if } x \leq R(z), \\ 0, & \text{otherwise.} \end{cases} \quad (2.24)$$

Its Fourier transform reads

$$\tilde{W}_s(k, R(z)) = \frac{3j_1(kR(z))}{(kR(z))}, \quad (2.25)$$

where $j_1(x)$ is a spherical Bessel function. A straightforward computation gives

$$\sigma^2(z, R(z), h) = \int_0^{+\infty} \frac{dk k^2}{2\pi^2} b^2(z, k) P(z, k) \tilde{W}_s^2(k, R(z)), \quad (2.26)$$

where in practice we apply an infrared and UV cut-off to the integral. The matter power spectrum $P(z, k)$ is defined as

$$\langle \delta_{\text{dm}}(z, \mathbf{k}) \delta_{\text{dm}}(z, \mathbf{k}') \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}') P(z, k), \quad (2.27)$$

and the bias parameter $b(z, k)$ relates the galaxy and matter density contrast as $\delta_g(z, k) = b(z, k) \cdot \delta_{\text{dm}}(z, k)$. Note that we are interested in the variance in the galaxy field and not in the dark matter field, since we assume that GWs are emitted from compact objects, which live inside galaxies. This issue can be circumvented by using the fact that we are looking for a lower bound on σ_g . Indeed, while knowing the exact value of σ_g would be best, introducing any amount of variance in the galaxy field would give a better estimate than

homogeneous completion which amounts to setting $\sigma_g = 0$. We assume that $b(z, k) \geq 1$, which is a reasonable assumption if we adopt the following argument: in principle, the galaxy bias also depends on the halo mass. However, most stars, which eventually become GW dark siren sources, live in galaxies of $\sim 10^{11} M_\odot$ (see the bottom panel of figure 12 of [40]), for which the bias is expected to be larger than one [41], at least for large scales. This assumption may break down if the volume of voxels is not large enough compared to the typical size of a galaxy. This introduces a lower redshift cut z_{\min} below which the calculation of σ_g cannot be trusted. We assume that the bias is larger than a certain threshold b_{th} for all relevant galaxies which can host GW sources. We assume that the threshold is $b_{\text{th}} = 1$, although one can easily adapt it, if it turns out to be lower. In the extreme case $b_{\text{th}} \rightarrow 0$, which corresponds to no clustering, the method converges to homogeneous completion.

We use the following symbolic approximation for the non-linear matter power spectrum

$$\begin{aligned} \ln(P(k, h, z)) = & 4.78h + (0.972 - 1.24z) (-0.152k + \tanh(1.4k) + 1.17) \\ & - \frac{0.217k + \tanh(0.727h - 107.0k)}{\tanh(0.231h + 112.0k)} - \cos(1.37h - 0.201k + 0.892z) \quad (2.28) \\ & + \frac{\cos(0.982h + 0.0826z)}{0.732k + 0.132}, \end{aligned}$$

which is $\sim 10\%$ accurate in the parameter space identified by the constraints: $0.001h/\text{Mpc} < k < 9.4h/\text{Mpc}$, $0 < z < 1$, $0.61 < h < 0.73$.³ Clearly, the dimensionless standard deviation $\sigma(z, R, h)$ given in eq. (2.26) depends on h . We discuss the cosmology dependence introduced by this in section 2.8. We take the value of b which minimizes the amplitude of the standard deviation and define the minimal standard deviation in the galaxy field as

$$\sigma_g(z, h) \equiv \hat{n}_g(z, h) \cdot \min_{b \geq b_{\text{th}}} \sigma(z, R, h). \quad (2.29)$$

Note that multiplying with the average number density of galaxy, introduces an extra cosmology dependence via the voxel size and the completeness fraction, which enters in eq. (2.17). However, when implementing this method in `gwcosmo`, mostly the ratio σ_g/\hat{n}_g is relevant, for which the cosmology dependence arising from the volume of voxels cancels out. We will check the impact of this cosmology dependence in section 2.8.

Finally, the integral in eq. (2.26) depends on the UV cut-off k_{max} . For constant $b(z, k)$, the integrand is a damped oscillating function. The dependence is strongest at the lowest redshift where variance completion applies and where voxels are the smallest. The voxel size already constrained the lowest redshift to be such that $z > z_{\min} = 0.015$. At this redshift, we find that σ oscillates as a function of k_{max} with an amplitude of about 10%, which corresponds to the level of precision of our method. At redshift $z = 0.1$, these oscillations have an amplitude of about 3%. In the following, we compute σ using

$$\sigma^2(z, R, h) = \int_{k_{\min}}^{k_{\max}} \frac{dk k^2}{2\pi^2} P(h, k, z) \tilde{W}_s^2(k, R), \quad (2.30)$$

with $k_{\min} = 0.001h/\text{Mpc}$, $k_{\max} = 9.4h/\text{Mpc}$, $P(k, h, z)$ from eq. (2.28), $\tilde{W}_s^2(k, R)$ from eq. (2.25) and ignore the weak dependence on the UV cut-off as well as the uncertainty on the power spectrum which are of similar order.

³See appendix A for more information on how this expression was obtained.

Eq. (2.29) gives an estimate of the minimal standard deviation in the galaxy number density field that can be used in variance completion. In the following section, we explain how to introduce variance completion in a redshift line of sight prior from `gwcs`.

2.7 Implementation in `gwcs`

Given the level of sophistication of `gwcs`, it turns out to be more convenient to condense the information of variance information into a ratio function $R(z, \hat{\mathbf{n}})$, which multiplies the out-of-catalogue contribution to the redshift line of sight prior of `gwcs` rather than to apply eq. (2.3). This allows to easily use other features of the line of sight redshift prior such as luminosity weighting. Schematically, we write that the probability density of presence of a galaxy at redshift z in the pixel corresponding to direction $\hat{\mathbf{n}}$ is given by

$$p(z, \hat{\mathbf{n}}) = p_c(z, \hat{\mathbf{n}}) + R(z, \hat{\mathbf{n}})p_m(z, \hat{\mathbf{n}}). \quad (2.31)$$

This corresponds to eq. (2.22) of [24] where the first term indicates the in-catalogue contribution, while the second term is proportional to the homogeneous out-of-catalogue contribution $p_m(z, \hat{\mathbf{n}})$. The ratio function $R(z, \hat{\mathbf{n}})$, introduced here, returns 1 whenever variance completion is uninformative and otherwise redistributes power from voxels which are more likely to be underdense to voxels which are more likely to be overdense. The ratio function is computed as follows

$$R(z, \hat{\mathbf{n}}) \equiv \frac{\hat{n}_{m;\text{var}}(z, \hat{\mathbf{n}})}{\hat{n}_{m;\text{hom}}(z, \hat{\mathbf{n}})}. \quad (2.32)$$

This has the benefit to cancel what ever modeling we used for $\hat{n}_{m;\text{hom}}$ since $\hat{n}_{m;\text{var}}$ averages to $\hat{n}_{m;\text{hom}}$. In this context, variance completion therefore boils down to estimate this ratio function in each pixel and redshift.

We apply variance completion to B-band galaxies of GLADE+ and luminosity weight them as per eq. (2.16). The high density to higher redshifts of the B-band with respect to the K-band allows us to successfully apply variance completion and to estimate the ratio function via eq. (2.32). Indeed, there are about 20M galaxies in the B-band against, 1M in the K-band. We then combine this ratio function with `gwcs` lines of sight extracted for the K-band. Note that while this is not strictly necessary, this allows to combine information from different bands. We work with bin sizes of $\delta z = 0.002$ between redshift $z_{\text{min}} = 0.015$ and a maximum redshift $z_{\text{max}}(\hat{\mathbf{n}}) \in [0.07, 0.20]$ which depends on direction. The maximal redshift is set such that the completeness in the B-band is above 10%. We get a nontrivial value for the ratio function for the voxels in the interval $z \in [z_{\text{min}}, z_{\text{max}}(\hat{\mathbf{n}})]$ and we set it to 1 everywhere else.

When the algorithm of variance completion does not converge on 1 or 2 consecutive redshift slices, we interpolate between those (This happens for less than 10% of voxels). If it does not converge on more consecutive slices, we interpolate smoothly with homogeneous completion. The outcomes of $\hat{n}_{m;\text{var}}^i$ which are below a threshold (fixed to $n_t \equiv \hat{n}_{m;\text{hom}}/10$) are filled with a minimum threshold value of n_t (This happens for less than 0.1% of the voxels to which we applied the algorithm). We then use a *Savitzky-Golay* filter to smooth the curves. Because of the postprocessing and numerical inaccuracies in variance completion,

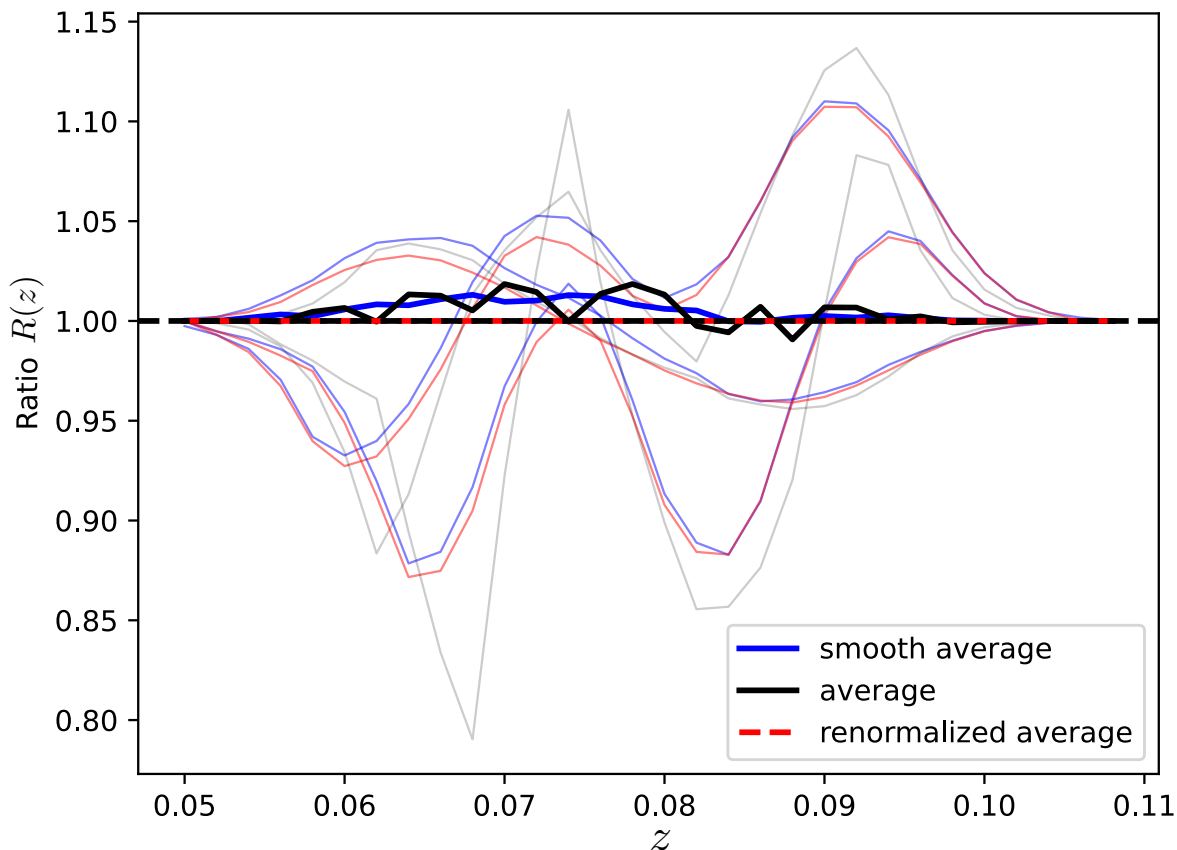


Figure 4. We plot 3 lines of sight at different levels of postprocessing. In black, we show the outcome of variance completion, after filling the gaps and removing values which are below the threshold of $\bar{n}_m/10$. This explains why the average over the region \mathcal{S} (thick black) differs from 1. In blue, we show the same curves after Savitzky-Golay filtering. In red, we show the same curves after renormalizing them such that the average of $R(z)$ over the 1024 lines of sight of this region average to 1 at each redshift z . The thick curves show the average over the region \mathcal{S} at each stage of postprocessing.

averaging the lines of sight ratio functions in a class \mathcal{S} at fixed redshift does not necessarily average to unity. We correct for this by renormalizing each line of sight to ensure that the ratio functions of a class \mathcal{S} at fixed redshift average to one. We show the process from the results of variance completion for the ratio function to a smooth ratio function that we use in `gwcosmo` in figure 4. We show 200 lines of sight ratios in figure 5.

We do not expect the ratio function to affect the normalization of the entire line of sight redshift prior $p(z, \hat{n})$. This is because of the following identity

$$\int_{\mathbb{S}_2} d\Omega R(z, \hat{n}) \cdot p_m(z, \hat{n}) = \sum_i \int_{\mathcal{S}_i} d\Omega R(z, \hat{n}) \cdot p_m(z, \hat{n}) \quad (2.33)$$

$$\simeq \sum_i p_m(z, \mathcal{S}_i) \underbrace{\int_{\mathcal{S}_i} d\Omega R(z, \hat{n})}_{=1} = \int_{\mathbb{S}_2} d\Omega p_m(z, \hat{n}), \quad (2.34)$$

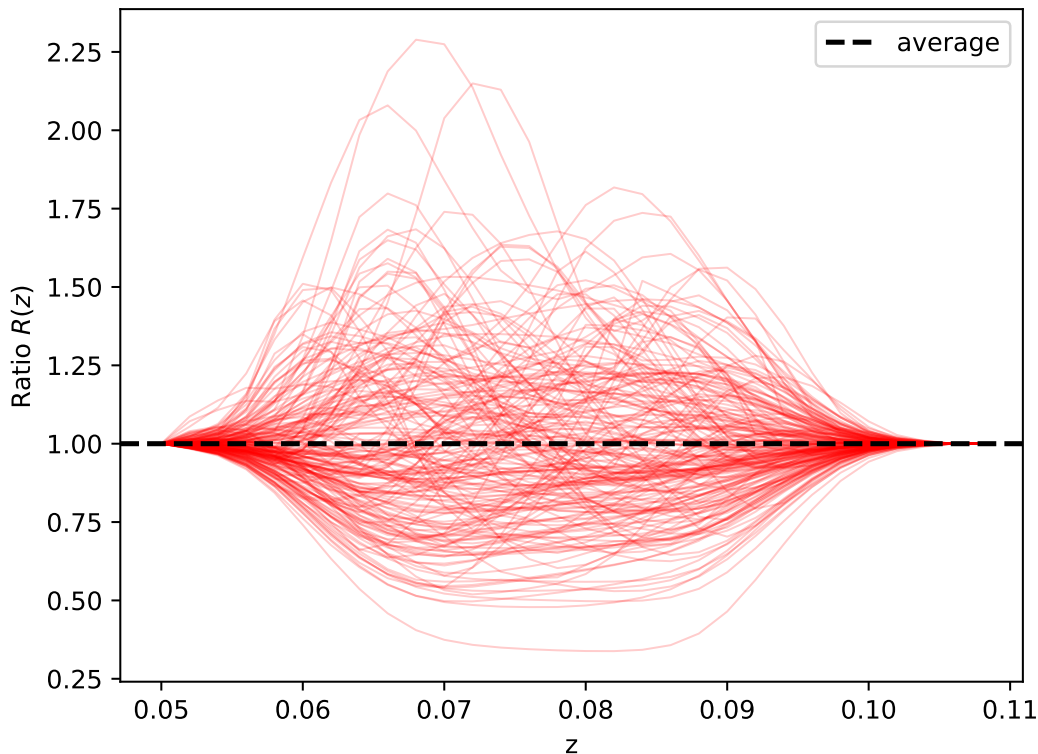


Figure 5. We plot 200 ratio functions $R(z, \hat{n})$ for 200 lines of sight in a same region \mathcal{S} . Some lines of sights have less galaxies over the relevant redshift range, implying that variance completion sets them to a value $R < 1$. However, they are expected to be as complete as other lines of sight, based on their magnitude threshold, which is probably dominated by galaxies seen at lower or (higher) redshift. The peaks at $R(z) \sim 2$ denote that these lines of sight are likely to host an over-dense region such as a cluster or a filament at this redshift. This class of lines of sight has a maximum redshift $z_{\max}(\hat{n}) = 0.105$, where the completeness in the B-band drops below 10%.

where the last approximate equality holds because the out-of-catalogue contributions have similar completeness levels in a fixed region \mathcal{S}_i . This implies that

$$\int_0^{z_m} \int_{\mathbb{S}_2} dz d\Omega p(z, \hat{n}) \simeq \int_0^{z_m} \int_{\mathbb{S}_2} dz d\Omega (p_c(z, \hat{n}) + p_m(z, \hat{n})) \quad (2.35)$$

such that the normalization is approximately preserved.

The ratio function introduced in eq. (2.31) and (2.32) allows to separate completely the determination of variance completion from the traditional line of sight redshift prior computation. This permits to easily implement variance completion in a dark siren method based on the line of sight redshift prior such as `gwcosmo` or `icarogw`.

2.8 Cosmology dependence of the line of sight redshift priors

Ultimately, we aim to use these lines of sight redshift priors to estimate the Hubble constant H_0 . Since we have used H_0 to determine σ_g , the completeness fraction and the voxel volumes,

it is legitimate to ask what is the impact of the choice of H_0 on the lines of sight. For our purpose, we stick to the dependence on the reduced Hubble constant h , defined as $H_0 \equiv 100h$ $\text{km s}^{-1} \text{Mpc}^{-1}$, leaving any dependence on Ω_{m0} and the dark energy equation of state to future work. In the end, we implement the method via the ratio function $R(z, \hat{\mathbf{n}})$, given in eq. (2.32). Because of the many steps required to reach an estimation of $\hat{n}_{m;\text{var}}$ and $\hat{n}_{m;\text{hom}}$ which enter in its definition, it is not entirely obvious what is the cosmology dependence of the ratio function. For instance, the standard deviation $\sigma(z, R, h)$ depends on h in two ways. First, through the power spectrum $P(k, h, z)$ which appears in (2.26). Second, it depends on h implicitly through the sphere equivalent radius R of the voxels, which appears in the same expression. This one depends on cosmology because it is extracted from the comoving volume of voxels of fixed angular size and fixed redshift bin. Another non-obvious example is the dependence of $\hat{n}_{m;\text{hom}}$ on the completeness fraction $\hat{f}_{\mathcal{S}}$ which is made through the apparent magnitude threshold map, converted to a luminosity threshold. This last step depends on cosmology via eq. (2.14)–(2.16). Finally, the estimates $\hat{n}_{m;\text{var}}^i$ are obtained by solving numerically and iteratively a linear system, which depends nonlinearly on σ_g . This makes the cosmology dependence of the ratio function intractable analytically; instead we will assess it numerically.

Since the out-of-catalogue part is modulated by the ratio function, it is sufficient for our purpose to study the h dependence of the ratio function, which contains in a single function of redshift and pixel, all the practical h dependence of the variance completion method. To understand qualitatively, the behavior of the ratio function under a change in h , we plot 1 line of sight for different values of $h \in \mathcal{H} \equiv \{0.67, 0.68, 0.69, 0.70, 0.71, 0.72, 0.73\}$ in figure 6.

The ratio function of this particular line of sight represents well the fact that we get different results for different values of h . However, the overall structure of the features is preserved. Given that the postprocessing seems to induce just as much change in the ratio function as a change in h and the fact that the power spectrum is precise to about 10%, we conclude that the h dependence is subdominant and acceptably negligible to within our current level of precision. To remove any remaining dependence on h , we marginalize over $h \in \mathcal{H}$. We reconstruct the line of sight using the average ratio function

$$R(z, \hat{\mathbf{n}}) = \langle R(z, \hat{\mathbf{n}}, h) \rangle_h. \quad (2.36)$$

Therefore, strong features survive this average if they appear independently of h . Their amplitude is moderated by that average. We plot an example line of sight for these values of h with the corresponding average and 1σ contours in figure 6. It should be noted that in applying this average, we have used a much narrower prior on h than `gwcsmo` does. Indeed, the later assumes the conservative prior $h \in [0.2, 1.4]$. Given, that this information only feeds in a fraction of the galaxy data, for which there is evidence that h is consistent with the prior \mathcal{H} [42], we do not expect this to affect significantly our conclusions. Another possibility is for each pixel and at each redshift, to take the minimum on h of $R(h, z)$. In this way, one ensures to take the most conservative prior on structure. Alternatively, one may divide the ratio excess from unity by a number larger than 1 to suppress the amplitude of the ratio function. This would also result in a more conservative estimate on structure. Note that this is the analogue of assuming a smaller minimum bias threshold for the calculation of σ_g but applied to the results of variance completion instead of the input.

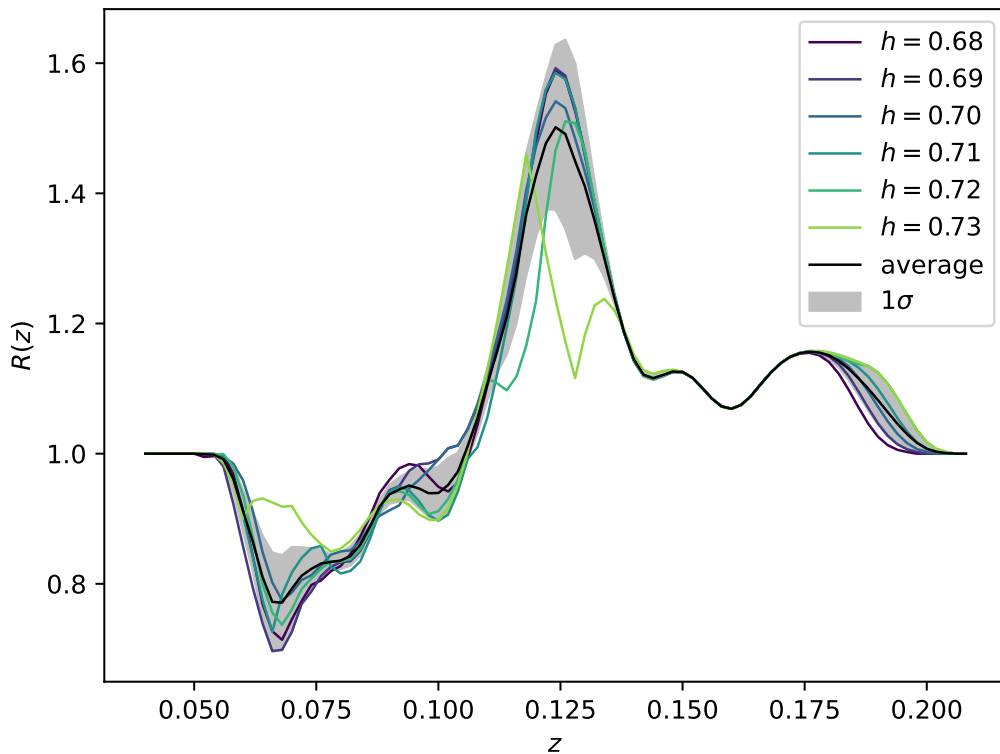


Figure 6. We plot an example of ratio function for different values of h . The grey shaded area indicates the 1 sigma contour around the average, which is in black. First, one can recognize the dominant features from one value of h to another. Second, the relative amplitude of the peaks is not preserved from one value of h to the other. There are several reasons for this. One reason is that variance completion is anyways approximate. Indeed, the variance completion algorithm stops when the total number of missing galaxies introduced matches the expectation from homogeneous completion within 1%. A further reason is the postprocessing (gap filling and smoothing, described in section 2.7) of the lines of sight, which easily modifies the shape of the lines of sight to an extent which is larger than what can be seen here. The variance completion also applies to a higher maximum redshift for higher h . The reason for this is that the completeness is cosmology dependent and we stop applying variance completion beyond a lower threshold on completeness, which is cosmology independent.

Here we have discussed the cosmology dependence of the ratio function and marginalized over the remaining dependence. This allows us to construct a line of sight redshift prior based on variance completion and combine with GW data to extract the Hubble constant, as discussed in the next sections.

3 Results

In this section, we present the redshift line of sight priors determined from variance completion applied to the GLADE+ catalogue. We then perform a dark siren analysis of GW190814,

a well localized event, which has good catalogue support. Finally, we apply the method to the events up to the third observing run, restricting to events with $\text{SNR} > 11$.

3.1 Updated line of sight redshift priors

We apply variance completion with $n_{\text{side}} = 32$ and $\delta z = 0.002$ between redshift $z_{\text{min}} = 0.015$ and a maximum redshift $z_{\text{max}}(\hat{\mathbf{n}}) \in [0.07, 0.20]$, which is fixed such that the B-band completeness remains above 10%. In this interval, the typical variance enforced by variance completion on the out-of-catalogue contribution is such that the standard deviation in the ratio function $R(z)$ is about $\sigma_R \lesssim 0.3$, as may be seen in figure 10. Outside these domains, we smoothly interpolate with homogeneous completion. The complete methodology is explained in details in section 2.

In figure 7, we present 50 lines of sight for variance completion versus homogeneous completion using luminosity weighting in both cases. One clearly sees that they overlap well at lower redshift and start to differ when the completeness drops, at higher redshift. Combined, the lines of sight average to the empty catalogue prior, which is also plotted there. Variance completion redistributes probabilities for the out-of-catalogue contribution to the redshift LOS prior, from voxels which seem to be emptier to those which are expected to contain more galaxies, at fixed redshift. Finally, between redshift 0.15 and 0.20, one can see some lines of light which are interpolated between variance completion and homogeneous completion. In the next section, we apply these line of sight redshift priors to the event GW190814, which is a well localized GW event.

3.2 Dark siren analysis of the Hubble constant for GW190814

With our brand new LOS redshift prior, we analyze the GW event GW190814, which has fairly good catalog support and excellent sky localization (90% confidence regions covers about 19 deg²). This event has also been analyzed for astrophysics and tests of gravity [43]. We present the 13 lines of sight which correspond to the 99.9% sky localization region of GW190814 in figure 8. We also plot the 68% confidence interval for the expected redshift z of GW190814, given the luminosity distance measurement and assuming a Planck prior on H_0 . This region overlaps well with a place where the impact of variance completion begins to be noticeable by eye on individual lines of sight. For single lines of sight, the change can be relatively important. However, combining as few as 13 lines of sight with equal weights, averages out some of the structure brought in by variance completion, as may be seen in figure 9. Indeed, assuming for simplicity that the over- and under-densities of these N lines of sight are gaussian distributed, we expect the standard deviation at a given redshift to be of order $\sigma = \sqrt{\sigma_R^2(z)/N}$. For $\sigma_R(z) \sim 0.3$ as may be seen in figure 10 and $N = 13$, this gives $\sigma \simeq 0.08$. This corresponds roughly to the amplitude of the difference between the two curves in figure 9. It should be clear that if the sky localization covers more pixels, then the information brought in by variance completion averages out even more, although weighting each pixel by their respective probability to host the GW may mitigate this effect. In figure 11, we plot a Hubble constant posterior using the uniform and luminosity weighting with homogeneous completion versus luminosity weighting with variance completion. It turns out that variance completion gives a posterior distribution which is marginally less compact for this particular event. Uniform or

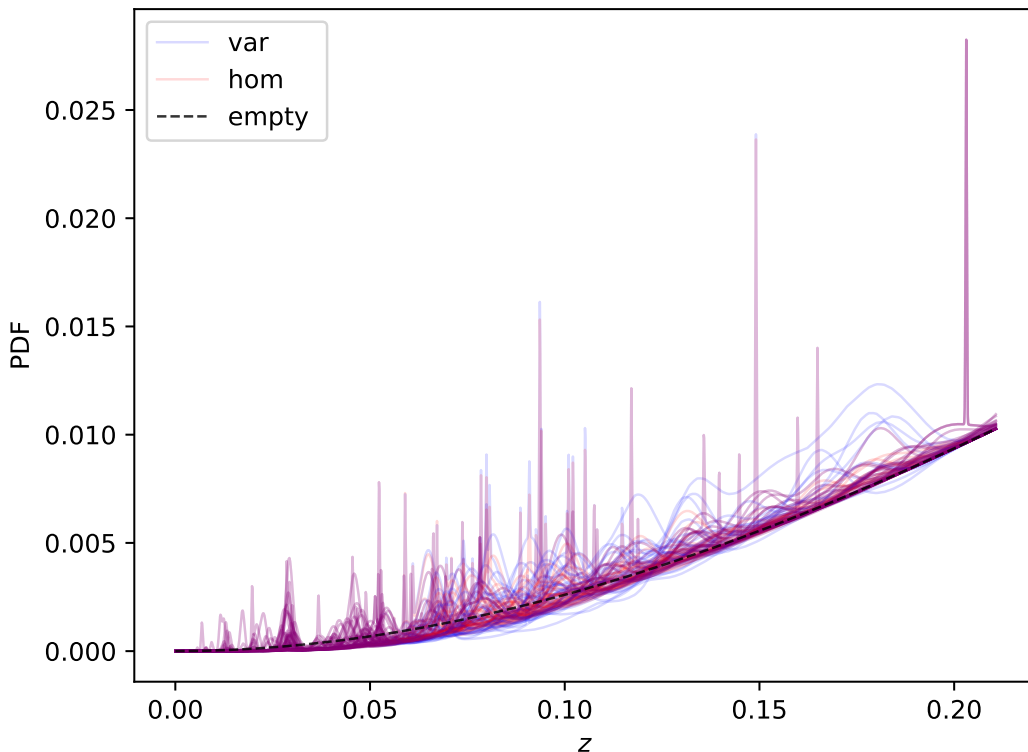


Figure 7. We plot 50 lines of sight for variance completion and homogeneous completion. At low redshift, when the completeness of GLADE+ is relatively high, they overlap well. It is only at higher redshift that variance completion takes a little bit more risk by assessing which of the lines of sight are more likely to be void or clusters depending on their observed number counts and on prior knowledge of large scale structure. Beyond a certain redshift, variance completion smoothly interpolates with homogeneous completion.

luminosity weighting with homogeneous completion and luminosity weighting with variance completion give $H_0 = 79.373^{+40.560}_{-39.054}$, $82.122^{39.153}_{-36.991}$, $80.353^{+40.770}_{-36.213}$ $\text{km s}^{-1} \text{Mpc}^{-1}$ respectively. They are all consistent with each other and with the results of [43, 44], which use `gwcs` which relies on homogeneous completion. They are also consistent with the analysis of [45], which use the Dark Energy Survey, which is complete at the relevant redshift range and as such, does not require any completion technique. In the next section, we apply variance completion to GW events of up to the third observing run of LVK.

3.3 Application to GWTC-3

We apply the line of sight priors generated with variance completion to the third observing run of LVK using 46 events which have $\text{SNR} > 11$, including binary black hole and neutron star black hole events. Compared to GW190814, analyzed in the previous section, those events span very different levels of uncertainty in the localization solid angles $\Delta\Omega$. The best localized events have $\Delta\Omega = \mathcal{O}(10) \text{ deg}^2$, while the worst localized event has $\Delta\Omega = 14,000$

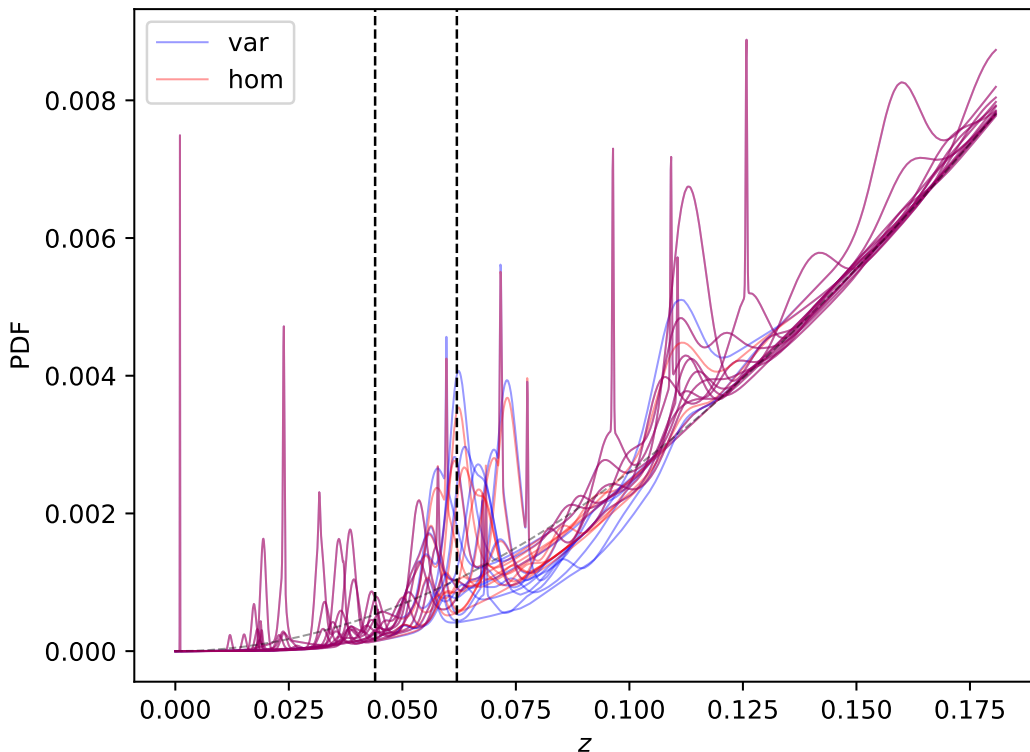


Figure 8. Here, we plot the 13 lines of sight corresponding to the 99.9% confidence sky localization of GW190814 for homogeneous completion (in red) and variance completion (in blue) both with in-catalogue galaxies weighed by their K-band luminosity. The vertical dashed lines shows the $\pm 1\sigma$ contour for the expectation on the redshift of the GW source. The dashed grey line shows the empty catalogue line of sight redshift prior.

deg^2 . We obtain the posterior on H_0 presented in figure 12 for homogeneous and variance completion, using luminosity weighting in both cases. We also plot the result of homogeneous completion with uniform weighting of galaxies. The measurement of H_0 changes from $H_0 = (65.022^{+11.533}_{-12.106}) \text{ km s}^{-1} \text{ Mpc}^{-1}$ for homogeneous completion to $H_0 = (65.076^{+11.298}_{-11.970}) \text{ km s}^{-1} \text{ Mpc}^{-1}$ for variance completion. We also show the result for variance completion when $h = 0.68$ is assumed to generate the ratio function for variance completion. The three measurements are consistent with each other and with the results of [46], which use `gwcosmo` with homogeneous completion. They are also consistent with the results of [47] which use the DESI legacy survey and deal with incompleteness with a data driven approach which avoids having to model the luminosity function. Our results are also consistent with those of [48], which use dark siren events for which the galaxy catalogue DELVE is complete over the corresponding GW sky areas. Finally, the authors of [49] focus on the DESI legacy survey and DELVE and focus on GW events for which they expect their galaxy catalogues to be complete.

The large scale structure prior changes slightly the median value of H_0 and makes the error bars slightly smaller. This is what we expect from adding in information from the

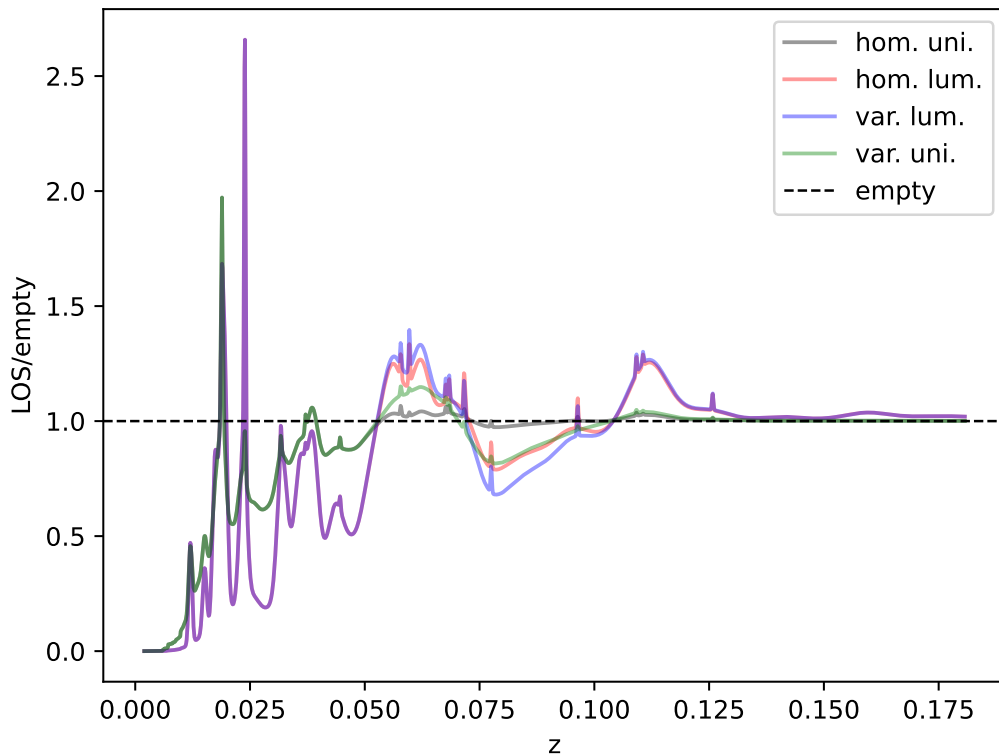


Figure 9. We plot the ratio of the combined line of sight redshift prior for the 13 pixels corresponding to the 99.9% confidence sky localization for different galaxy weightings and completion. We plot homogeneous completion with uniform (black) and luminosity (red) weighting of in-catalogue galaxies. We also plot variance completion with luminosity weighting (in blue) and with uniform weighting in green. The change is less spectacular than for individual lines of sight. This should not be surprising since we are averaging 13 lines of sight. However, for both uniform and luminosity weighting, variance completion is more informative than homogeneous completion, in the sense that the lines of sight differ more from the empty catalogue. On the right of the plot, the catalogue converges to an empty one which explains why the ratio asymptotes to unity.

large scale structure for the out-of-catalogue galaxies. The reason why variance completion does not perform significantly better than homogeneous completion is a reflection of what happens for the best localized event GW190814. As we argued in the previous section on this event, averaging the line of sight redshift prior over as few as 13 lines of sight, already washes significantly the effect of variance completion. For the worse localized event, implementing variance completion over many more lines of sight (about 4,170 for $\Delta\Omega = 14,000 \text{ deg}^2$), the out-of-catalogue contribution simply averages out to the empty catalogue. This means that only a handful of events with good sky localization ($\Delta\Omega = \mathcal{O}(10) \text{ deg}^2$) contribute to the improvement of the line of sight redshift prior.

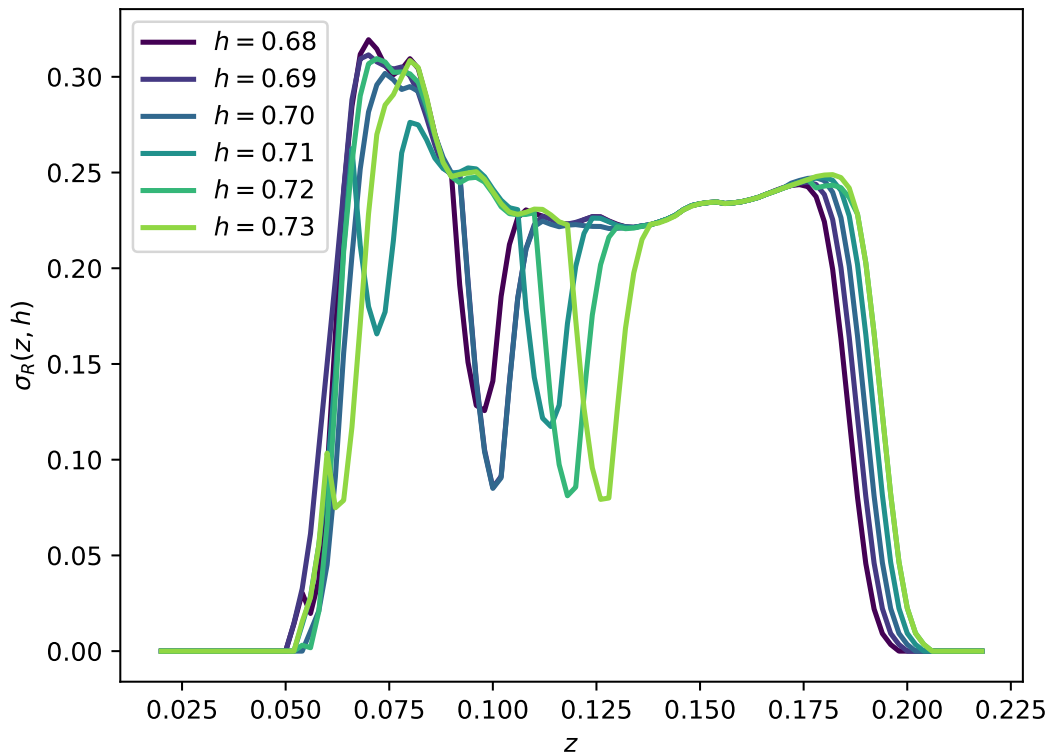


Figure 10. Standard deviation $\sigma_R(z, h)$ of the ratio function $R(z, h)$ as a function of redshift z for different values of h . At fixed redshift z , there exists a value of h which minimizes $\sigma_R(z, h)$, giving a prospective h which minimizes the amplitude of structure at that redshift. However, that value changes with redshift, making the value of h which minimizes structure difficult to identify unequivocally. For certain values of h and at certain redshifts, the amplitude of the dispersion is lower than for the other values of h . That can happen if variance completion converges to a solution with lower dispersion.

3.4 Discussion

The goal of this work is to maximize the potential constraining power of dark siren analyses, by including large scale structure knowledge in the method. Implementing variance completion is a step towards this goal, but to achieve this we had to overcome several challenges.

One such difficulty is the resolution in pixel and in redshift space. Indeed, we had to deal with `healpix` pixels and redshift bins, with voxels which change geometry with redshift. The ideal resolution for variance completion to work is a compromise between two extremes. If the resolution is too low, then the variance in the number counts on large volumes drops to zero, making the method converge to homogeneous completion. If the resolution is too high, the variance in the number count can be significant, potentially maximizing the outcome of variance completion, but four other limitations appear. First, the number of voxels required to approximate a homogeneous patch of the Universe ($\sim 10^6$ Mpc³) on which the method was designed to work becomes too large (> 1000), making the method computationally heavy. Second, at some point the smoothed voxelated galaxy number densities do not accurately

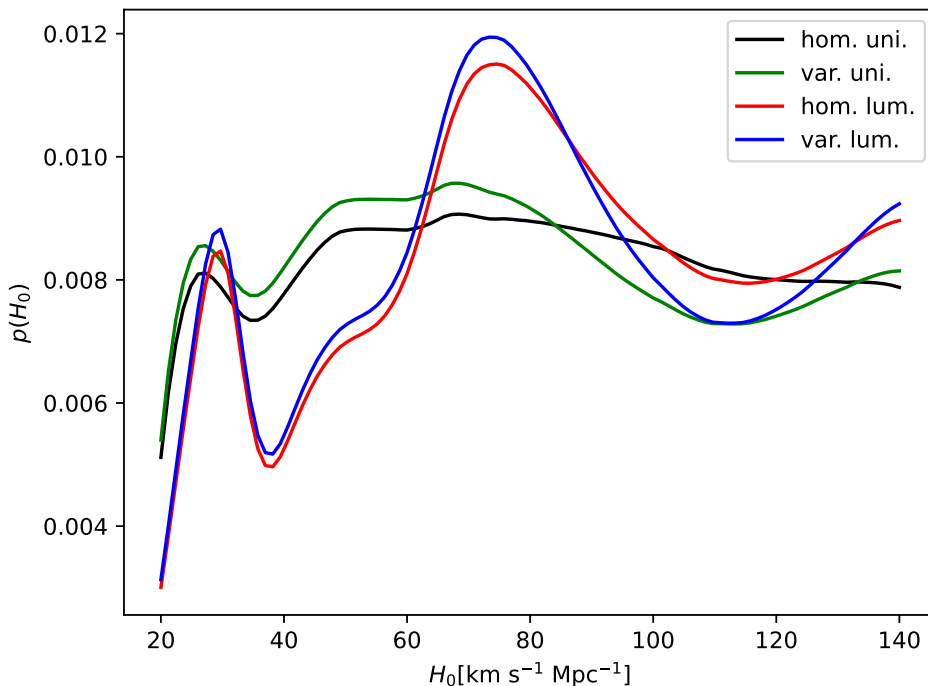


Figure 11. Here, we present the posterior for the Hubble constant H_0 for three different line of sight prescriptions for GW data coming from GW190814. The median and 68.3% confidence intervals for uniform weighting of in-catalogue galaxies, homogeneous and variance completion give $H_0 = 79.373^{+40.560}_{-39.054}, 76.510^{+43.070}_{-37.246}$ km s $^{-1}$ Mpc $^{-1}$, respectively. For luminosity weighting of in-catalogue galaxies, homogeneous and variance completion give $H_0 = 82.122^{+39.153}_{-36.991}$ and $80.351^{+40.770}_{-36.213}$ km s $^{-1}$ Mpc $^{-1}$, respectively. In both the uniform and luminosity weighting cases, variance completion suppresses power at relatively high values (between 90 – 130) of H_0 with respect to homogeneous completion. The curves have been smoothed to remove unphysical features coming from the finite number of $2 \cdot 10^6$ injections.

represent the galaxy field distribution anymore. Third, we need an input on the variance in the galaxy field, which we could bound from below, by assuming that the bias satisfies $b(z, \mathbf{k}) \geq 1$. We cannot apply this bound anymore if we lower the resolution down to arbitrary small scales, in which case, the concept of bias does not make sense anymore. Finally, the lowest resolution that we can apply the formalism to, depends on what scales we manage to model the nonlinear power spectrum. In our case, $k_{\max} = 9.4h/\text{Mpc}$, which corresponds to a wavelength of approximately 1 Mpc. The fact that the redshift range over which the GLADE+ catalogue is sufficiently empty, allowing variance completion to enhance the dark siren method, and not too empty, to allow the method to work, has resulted in constraints on the geometry. Indeed, we fixed the `healpix` resolution to $n_{\text{side}} = 32$ and the redshift bin to $\delta z = 0.002$. The bin size could well vary with redshift and the `healpix` resolution changed at various redshift checkpoints, but we leave these possibilities to future work.

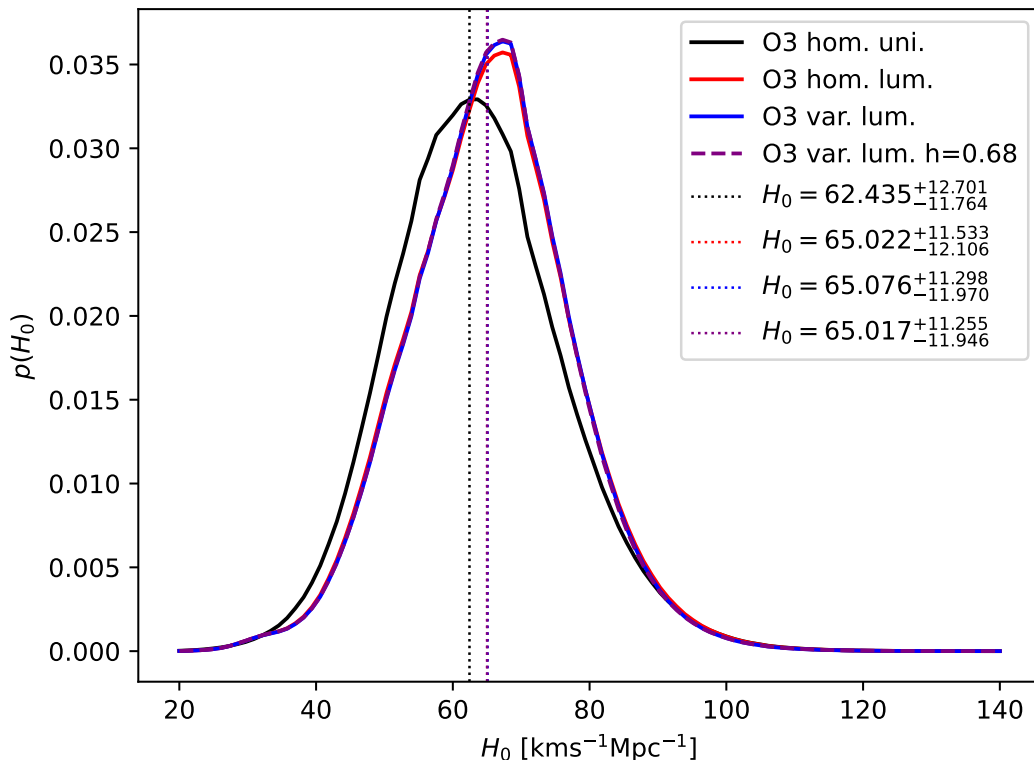


Figure 12. We present the posterior on H_0 for 43 events of O3 for homogeneous completion with uniform weighting (O3 uni), homogeneous completion with luminosity weighting (O3 lum), variance completion with luminosity weighting (O3 var) and variance completion using $h = 0.68$ with luminosity weighting. These last two curves mentioned, sit almost on top of each other. The reported values of H_0 are in units of $\text{km s}^{-1} \text{Mpc}^{-1}$ and correspond to the median and 68.3% confidence intervals.

One obvious difficulty of this method is the modeling required for \bar{n}_g and σ_g . The former, which is estimated from data, eventually drops out of the method when implemented in `gwcsmo`, since it cancels between the numerator and the denominator of the ratio function $R(z, \hat{\mathbf{n}})$ given in eq. (2.32). The variance σ_g^2 on the other hand is estimated from the power spectrum, which has to be modelled. This required sophisticated techniques presented in the appendix A. We also simplified the analysis by neglecting geometric effects and by assuming that it is sufficient to consider the voxels as sphere-equivalent. Even if this approximation turns out to be insufficient, the difference between the sphere-equivalent variance σ_g^2 and the one calculated from the full geometry of a voxel can be absorbed into an assumption of a lower bias prescription than $b(z, \mathbf{k}) > 1$. Finally, we assumed that the bias between galaxies and dark matter is larger than unity, which is expected for large scales and which may break down for smaller scales, although we do not expect to enter this regime, given our minimal voxel size.

Redshift uncertainties can be incorporated in the method by integrating the redshift posterior of galaxies over the voxel redshift range. For simplicity, we have limited ourselves to the assumption of Gaussian uncertainties, although this may be inaccurate, especially for

photometric redshifts. On top of this, below a certain magnitude threshold, photometric redshifts can be biased measurements of redshift, which is a complication beyond the scope of this work.

Between the raw outcome of variance completion and the smooth curves of the ratio function, we had to apply some postprocessing, to fill gaps, smooth the curves and renormalize the ratio function such that it averages to unity among a class of voxels. We also averaged the ratio function over a few values of h . These procedures affect the shape of the lines of sight. As such they may affect the outcome on H_0 , which is however, difficult to quantify. As the change in the posterior on H_0 for 46 events remains small, the effect is negligible for now, but may not remain so in the future.

With the advent of new generations of GW detectors and galaxy surveys, galaxy catalogues will be more complete to higher redshift with better redshift uncertainties and the GW localization volumes smaller. As both of these improve, the method will be relevant to more GW events and should contribute to lower the error bars on the Hubble constant. Meanwhile, the method can always be implemented in parallel to check that homogeneous completion is consistent with the expectation from large scale structure.

While the change may not be far-reaching for now, there may be GW events in the future for which localization volume corresponds only to a few lines of sight, for which variance completion predicts a 2σ over/under-density in the region of interest. This would have a significant impact on the H_0 posterior. Likewise, the identification of multiple images of a strongly lensed GW may allow to constrain their sky localization to a sub-arcsecond level [50]. This subpixel localization would also result in a higher impact of the out-of-catalogue contribution from variance completion.

4 Conclusion

In this work, we have developed variance completion to make it applicable to real galaxy data. After a brief reminder of variance completion, we detailed how to obtain two important ingredients for the method, which are the average number density of galaxies together with its standard deviation. We argued that the former drops out eventually, when implemented in `gwcs` and discussed the cosmology dependence of the latter. We presented the classes of pixels, with similar estimated completeness fractions over which the variance completion algorithm applies. We introduced a ratio function which contains the information on whether we expect a certain voxel to rather be an over or underdensity and with what amplitude. This ratio function can easily be incorporated in the `gwcs` formalism by multiplying the out-of-catalogue homogeneous in comoving volume contribution. We applied some postprocessing to fill gaps, smooth the curves and renormalize them. We showed how this ratio function is not expected to change the overall normalization of the redshift line of sight prior.

We then presented updated lines of sight and applied the method to GW190814, which is a well localized event. The updated lines of sight compare easily with the ones for homogeneous completion and are slightly more informative on some of them, increasing the amplitude of large scale under/over-densities. We then applied the method to extract the Hubble constant and found consistent results whether we applied homogeneous completion or variance completion. The median value shifted a little bit and the error bars as well.

For simplicity, it seems preferable to use homogeneous completion for now, even though there is a marginal gain in constraining power from variance completion. This picture may change for GW events that are localized to about 3 [deg²], which corresponds to one `healpix` pixel of $n_{\text{side}} = 32$. In this case, we can expect variance completion to improve significantly the posterior on the Hubble constant.

In the near future, a wealth of galaxy data is going to be collected from the Large Synoptic Survey Telescope (LSST) [51], Euclid [52] and the Dark Energy Spectroscopic Instrument (DESI) [42]. For relatively low redshift regions, the incompleteness may no longer be a problem, although this always depends on the mask. Nonetheless, there will still be redshift regions for which the galaxy catalogues are incomplete. In this work, we have shown in practice how to apply variance completion to a real survey and take into account knowledge of the large scale structure, as a prior for a dark siren analysis.

A Approximate matter power spectrum

The non-linear matter power spectrum is a non-closed form function that requires expensive cosmological simulations to be accurately predicted for each cosmology of interest. A workaround to this problem is to run a single time a relatively large number of simulations to sample the parameter space of interest and use efficient high-dimensional interpolation techniques to approximate the output of the simulations (in this case the matter power spectrum) in a dense region of the parameter space. This is for example what was done for the Euclid Emulator 2 (EE2) [53], a state-of-the-art cosmological emulator for the matter power spectrum, which was claimed to be percent level accurate in a wide region of the cosmological parameter space. The EE2 relies on polynomial chaos expansion [54] to approximate the output of its *training* simulations. To simplify the interpolation problem, instead of using the full matter power spectrum as target function, the EE2 uses the non-linear boost factor of the power spectrum, i.e. the ratio of the non-linear power spectrum computed from simulations with the linear power spectrum computed from the Boltzmann solver CLASS [55]. Therefore predicting the power spectrum with EE2 requires the evaluation of the CLASS code, which, while much faster than cosmological simulations, still requires about a second to run. In this work we are more interested in having a faster evaluation time than a high accuracy, so we use the power spectrum computed with EE2 in combination with CLASS as target to create a symbolic approximation which can be computed in microseconds, while providing $\sim 10\%$ accurate power spectra.

Our approach consists in spanning a subset of the EE2 parameter space in the variables h , k , z , while fixing the other parameters to the Planck 2018 fiducial cosmology [1], as summarized in table 1. We use a latin hypercube sampling (LHS) technique [56] to uniformly sample the (h, k, z) parameter space with 300 points and apply the symbolic regression algorithm `pyoperon` [57] to obtain candidate symbolic expressions of various complexities, following an approach similar to the one outlined in [58]. We select the model of least complexity that achieves better than 10% accuracy⁴ on an independent test set of 1000 points, also constructed with LHS.

⁴Unlike the common practice in literature of quoting the root mean squared error (RMSE) for the accuracy, we use here the maximum error on the training set, which we find being ~ 3 times larger than the RMSE.

h	[0.61, 0.73]
k [$h \text{ Mpc}^{-1}$]	[0.001, 9.4]
z	[0, 1]
Ω_{m}	0.3111
Ω_{b}	0.04897
m_{ν} [eV]	0.06
n_{s}	0.9667
A_{s}	2.215×10^{-9}

Table 1. Ranges and fiducial values of the cosmological parameters used to create the training set for the approximate matter power spectrum expression (2.28).

Acknowledgments

C.D. and T.B. are supported by ERC Starting Grant SHADE (grant no. StG 949572). B.F. and T. B. are further supported by a Royal Society University Research Fellowship (grant no. URF\R\231006). We thank Rachel Gray, Matthieu Schaller, Konstantin Leyde, Anson Chen for useful discussions. C.D. thanks SPCS of QMUL for hospitality. Some of the results in this paper have been derived using the healpy and HEALPix package [34, 35].

References

- [1] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [Erratum *ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [2] A.G. Riess et al., *Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond Λ CDM*, *Astrophys. J.* **876** (2019) 85 [[arXiv:1903.07603](#)] [[INSPIRE](#)].
- [3] A.G. Riess et al., *A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s⁻¹ Mpc⁻¹ Uncertainty from the Hubble Space Telescope and the SH0ES Team*, *Astrophys. J. Lett.* **934** (2022) L7 [[arXiv:2112.04510](#)] [[INSPIRE](#)].
- [4] D.W. Pesce et al., *The Megamaser Cosmology Project. XIII. Combined Hubble constant constraints*, *Astrophys. J. Lett.* **891** (2020) L1 [[arXiv:2001.09213](#)] [[INSPIRE](#)].
- [5] H0LiCOW collaboration, *H0LiCOW – XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3 σ tension between early- and late-Universe probes*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 1420 [[arXiv:1907.04869](#)] [[INSPIRE](#)].
- [6] N. Schöneberg et al., *The H0 Olympics: A fair ranking of proposed models*, *Phys. Rept.* **984** (2022) 1 [[arXiv:2107.10291](#)] [[INSPIRE](#)].
- [7] E. Di Valentino et al., *In the realm of the Hubble tension—a review of solutions*, *Class. Quant. Grav.* **38** (2021) 153001 [[arXiv:2103.01183](#)] [[INSPIRE](#)].
- [8] D. Brout et al., *The Pantheon+ Analysis: SuperCal-fragilistic Cross Calibration, Retrained SALT2 Light-curve Model, and Calibration Systematic Uncertainty*, *Astrophys. J.* **938** (2022) 111 [[arXiv:2112.03864](#)] [[INSPIRE](#)].
- [9] B.F. Schutz, *Determining the Hubble Constant from Gravitational Wave Observations*, *Nature* **323** (1986) 310 [[INSPIRE](#)].

- [10] G. Pierra, S. Mastrogiovanni, S. Perriès and M. Mapelli, *Study of systematics on the cosmological inference of the Hubble constant from gravitational wave standard sirens*, *Phys. Rev. D* **109** (2024) 083504 [[arXiv:2312.11627](#)] [[INSPIRE](#)].
- [11] LIGO SCIENTIFIC et al. collaborations, *A gravitational-wave standard siren measurement of the Hubble constant*, *Nature* **551** (2017) 85 [[arXiv:1710.05835](#)] [[INSPIRE](#)].
- [12] B.S. Sathyaprakash, B.F. Schutz and C. Van Den Broeck, *Cosmography with the Einstein Telescope*, *Class. Quant. Grav.* **27** (2010) 215006 [[arXiv:0906.4151](#)] [[INSPIRE](#)].
- [13] M. Mancarella et al., *Accurate Standard Siren Cosmology with Joint Gravitational-Wave and γ -Ray Burst Observations*, *Phys. Rev. Lett.* **133** (2024) 261001 [[arXiv:2405.02286](#)] [[INSPIRE](#)].
- [14] J.M. Ezquiaga and D.E. Holz, *Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries*, *Phys. Rev. Lett.* **129** (2022) 061102 [[arXiv:2202.08240](#)] [[INSPIRE](#)].
- [15] S. Mastrogiovanni et al., *On the importance of source population models for gravitational-wave cosmology*, *Phys. Rev. D* **104** (2021) 062009 [[arXiv:2103.14663](#)] [[INSPIRE](#)].
- [16] A. Finke et al., *Modified gravitational wave propagation and the binary neutron star mass function*, *Phys. Dark Univ.* **36** (2022) 100994 [[arXiv:2108.04065](#)] [[INSPIRE](#)].
- [17] M. Mancarella, E. Genoud-Prachex and M. Maggiore, *Cosmology and modified gravitational wave propagation from binary black hole population models*, *Phys. Rev. D* **105** (2022) 064030 [[arXiv:2112.05728](#)] [[INSPIRE](#)].
- [18] M. Mancarella et al., *Gravitational-wave cosmology with dark sirens: state of the art and perspectives for 3G detectors*, *PoS ICHEP2022* (2022) 127 [[arXiv:2211.15512](#)] [[INSPIRE](#)].
- [19] F. Iacovelli et al., *Modified gravitational wave propagation: information from strongly lensed binaries and the BNS mass function*, in the proceedings of the 56th Rencontres de Moriond on Gravitation, La Thuile, Italy, January 30 – February 06 (2022) [[arXiv:2203.09237](#)] [[INSPIRE](#)].
- [20] C. Karathanasis, S. Mukherjee and S. Mastrogiovanni, *Binary black holes population and cosmology in new lights: signature of PISN mass and formation channel in GWTC-3*, *Mon. Not. Roy. Astron. Soc.* **523** (2023) 4539 [[arXiv:2204.13495](#)] [[INSPIRE](#)].
- [21] L.W.H. Fung, T. Broadhurst and G.F. Smoot, *Pure gravitational wave estimation of Hubble’s constant using neutron star–black hole mergers*, *Mon. Not. Roy. Astron. Soc.* **538** (2025) 711 [[arXiv:2308.02440](#)] [[INSPIRE](#)].
- [22] R. Gray, C. Messenger and J. Veitch, *A pixelated approach to galaxy catalogue incompleteness: improving the dark siren measurement of the Hubble constant*, *Mon. Not. Roy. Astron. Soc.* **512** (2022) 1127 [[arXiv:2111.04629](#)] [[INSPIRE](#)].
- [23] R. Gray et al., *Cosmological inference using gravitational wave standard sirens: A mock data analysis*, *Phys. Rev. D* **101** (2020) 122001 [[arXiv:1908.06050](#)] [[INSPIRE](#)].
- [24] R. Gray et al., *Joint cosmological and gravitational-wave population inference using dark sirens and galaxy catalogues*, *JCAP* **12** (2023) 023 [[arXiv:2308.02281](#)] [[INSPIRE](#)].
- [25] S. Mastrogiovanni et al., *Joint population and cosmological properties inference with gravitational waves standard sirens and galaxy surveys*, *Phys. Rev. D* **108** (2023) 042002 [[arXiv:2305.10488](#)] [[INSPIRE](#)].
- [26] N. Borghi et al., *Cosmology and Astrophysics with Standard Sirens and Galaxy Catalogs in View of Future Gravitational Wave Observations*, *Astrophys. J.* **964** (2024) 191 [[arXiv:2312.05302](#)] [[INSPIRE](#)].

- [27] LIGO SCIENTIFIC and VIRGO collaborations, *A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart*, *Astrophys. J. Lett.* **871** (2019) L13 [[arXiv:1807.05667](#)] [[INSPIRE](#)].
- [28] A. Finke et al., *Cosmology with LIGO/Virgo dark sirens: Hubble parameter and modified gravitational wave propagation*, *JCAP* **08** (2021) 026 [[arXiv:2101.12660](#)] [[INSPIRE](#)].
- [29] C. Dalang and T. Baker, *The clustering of dark sirens' invisible host galaxies*, *JCAP* **02** (2024) 024 [[arXiv:2310.08991](#)] [[INSPIRE](#)].
- [30] K. Leyde, T. Baker and W. Enzi, *Cosmic cartography: Bayesian reconstruction of the galaxy density informed by large-scale structure*, *JCAP* **12** (2024) 013 [[arXiv:2409.20531](#)] [[INSPIRE](#)].
- [31] DES et al. collaborations, *First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary–Black-hole Merger GW170814*, *Astrophys. J. Lett.* **876** (2019) L7 [[arXiv:1901.01540](#)] [[INSPIRE](#)].
- [32] DESI collaboration, *A Dark Siren Measurement of the Hubble Constant with the LIGO/Virgo Gravitational Wave Event GW190412 and DESI Galaxies*, *Res. Notes AAS* **7** (2023) 250 [[arXiv:2311.13062](#)] [[INSPIRE](#)].
- [33] G. Dálya et al., *GLADE+ : an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors*, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 1403 [[arXiv:2110.06184](#)] [[INSPIRE](#)].
- [34] A. Zonca et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, *J. Open Source Softw.* **4** (2019) 1298 [[INSPIRE](#)].
- [35] K.M. Górski et al., *HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys. J.* **622** (2005) 759 [[astro-ph/0409513](#)] [[INSPIRE](#)].
- [36] P. Schechter, *An analytic expression for the luminosity function for galaxies*, *Astrophys. J.* **203** (1976) 297 [[INSPIRE](#)].
- [37] C.S. Kochanek et al., *The k-band galaxy luminosity function*, *Astrophys. J.* **560** (2001) 566 [[astro-ph/0011456](#)] [[INSPIRE](#)].
- [38] 2dGRS collaboration, *The 2dF Galaxy Redshift Survey: The B(J)-band galaxy luminosity function and survey selection function*, *Mon. Not. Roy. Astron. Soc.* **336** (2002) 907 [[astro-ph/0111011](#)] [[INSPIRE](#)].
- [39] C.J. Conselice, A. Wilkinson, K. Duncan and A. Mortlock, *The evolution of galaxy number density at $z < 8$ and its implications*, *Astrophys. J.* **830** (2016) 83 [[arXiv:1607.03909](#)] [[INSPIRE](#)].
- [40] S.P. Driver et al., *Galaxy And Mass Assembly (GAMA): Data Release 4 and the $z < 0.1$ total and $z < 0.08$ morphological galaxy stellar mass functions*, *Mon. Not. Roy. Astron. Soc.* **513** (2022) 439 [[arXiv:2203.08539](#)] [[INSPIRE](#)].
- [41] J.L. Tinker et al., *The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, *Astrophys. J.* **724** (2010) 878 [[arXiv:1001.3162](#)] [[INSPIRE](#)].
- [42] DESI collaboration, *DESI 2024 VI: cosmological constraints from the measurements of baryon acoustic oscillations*, *JCAP* **02** (2025) 021 [[arXiv:2404.03002](#)] [[INSPIRE](#)].
- [43] LIGO SCIENTIFIC and VIRGO collaborations, *GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object*, *Astrophys. J. Lett.* **896** (2020) L44 [[arXiv:2006.12611](#)] [[INSPIRE](#)].

- [44] S. Vasylyev and A. Filippenko, *A Measurement of the Hubble Constant using Gravitational Waves from the Binary Merger GW190814*, *Astrophys. J.* **902** (2020) 149 [[arXiv:2007.11148](#)] [[INSPIRE](#)].
- [45] DES collaboration, *A statistical standard siren measurement of the Hubble constant from the LIGO/Virgo gravitational wave compact object merger GW190814 and Dark Energy Survey galaxies*, *Astrophys. J. Lett.* **900** (2020) L33 [[arXiv:2006.14961](#)] [[INSPIRE](#)].
- [46] LIGO SCIENTIFIC et al. collaborations, *Constraints on the Cosmic Expansion History from GWTC-3*, *Astrophys. J.* **949** (2023) 76 [[arXiv:2111.03604](#)] [[INSPIRE](#)].
- [47] A. Palmese, C.R. Bom, S. Mucesh and W.G. Hartley, *A Standard Siren Measurement of the Hubble Constant Using Gravitational-wave Events from the First Three LIGO/Virgo Observing Runs and the DESI Legacy Survey*, *Astrophys. J.* **943** (2023) 56 [[arXiv:2111.06445](#)] [[INSPIRE](#)].
- [48] V. Alfradique et al., *A dark siren measurement of the Hubble constant using gravitational wave events from the first three LIGO/Virgo observing runs and DELVE*, *Mon. Not. Roy. Astron. Soc.* **528** (2024) 3249 [[arXiv:2310.13695](#)] [[INSPIRE](#)].
- [49] C.R. Bom et al., *A dark standard siren measurement of the Hubble constant following LIGO/Virgo/KAGRA O4a and previous runs*, *Mon. Not. Roy. Astron. Soc.* **535** (2024) 961 [[arXiv:2404.16092](#)] [[INSPIRE](#)].
- [50] O.A. Hannuksela, T.E. Collett, M. Çahşkan and T.G.F. Li, *Localizing merging black holes with sub-arcsecond precision using gravitational-wave lensing*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 3395 [[arXiv:2004.13811](#)] [[INSPIRE](#)].
- [51] LSST collaboration, *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, *Astrophys. J.* **873** (2019) 111 [[arXiv:0805.2366](#)] [[INSPIRE](#)].
- [52] EUCLID collaboration, *Euclid. I. Overview of the Euclid mission*, *Astron. Astrophys.* **697** (2025) A1 [[arXiv:2405.13491](#)] [[INSPIRE](#)].
- [53] EUCLID collaboration, *Euclid preparation: IX. EuclidEmulator2 – power spectrum emulation with massive neutrinos and self-consistent dark energy perturbations*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 2840 [[arXiv:2010.11288](#)] [[INSPIRE](#)].
- [54] D. Xiu and G.E. Karniadakis, *The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations*, *SIAM J. Sci. Comput.* **24** (2002) 619.
- [55] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034 [[arXiv:1104.2933](#)] [[INSPIRE](#)].
- [56] M.D. McKay, R.J. Beckman and W.J. Conover, *Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*, *Technometrics* **21** (1979) 239.
- [57] B. Burlacu, G. Kronberger and M. Kommenda, *Operon C++: an efficient genetic programming framework for symbolic regression*, in the proceedings of the *2020 Genetic and Evolutionary Computation Conference Companion*, Cancún, Mexico, July 8 - 12 (2020) [[DOI:10.1145/3377929.3398099](#)].
- [58] D.J. Bartlett et al., *SYREN-HALOFIT: A fast, interpretable, high-precision formula for the Λ CDM nonlinear matter power spectrum*, *Astron. Astrophys.* **686** (2024) A150 [[arXiv:2402.17492](#)] [[INSPIRE](#)].