Noisy neighbours: inference biases from overlapping gravitational-wave signals

Andrea Antonelli[1]⋆, Ollie Burke[1,2]†, Jonathan R. Gair[1,2]‡

[1] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam-Golm, 14476, Germany*
[2] *School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh EH9 3FD, UK*

6 April 2021

ABSTRACT

Understanding and dealing with inference biases in gravitational-wave (GW) parameter estimation when a plethora of signals are present in the data is one of the key challenges for the analysis of data from future GW detectors. Working within the linear signal approximation, we describe generic metrics to predict inference biases on GW source parameters in the presence of confusion noise from unfitted foregrounds, from overlapping signals that coalesce close in time to one another, and from residuals of other signals that have been incorrectly fitted out. We illustrate the formalism with simplified, yet realistic, scenarios appropriate to third-generation ground-based (Einstein Telescope) and space-based (LISA) detectors, and demonstrate its validity against Monte-Carlo simulations. We find it to be a reliable tool to cheaply predict the extent and direction of the biases. Finally, we show how this formalism can be used to correct for biases that arise in the sequential characterisation of multiple sources in a single data set, improving the accuracy of the global-fit without the need for expensive joint-fitting of the sources.

Key words: gravitational waves.

1 INTRODUCTION

In the analysis of data from future gravitational-wave (GW) detectors, we will be confronted with the prospect of detecting and performing parameter inference on sources that overlap with other resolved or unresolved signals. The presence of such additional signals in the data or their incomplete removal through inaccurate waveform templates, might lead to biases in the parameter estimates for the source of interest, if they are not properly accounted for. While this possibility is relevant for imminent upgrades of the LIGO-Virgo-KAGRA detectors' network (Abbott et al. 2018), the odds of this happening are higher with future ground-based and space-based detectors such as the Einstein Telescope (ET)(Punturo et al. 2010), Cosmic Explorer (CE) (Reitze et al. 2019) and the Laser Interferometer Space Antenna (LISA)(Amaro-Seoane et al. 2017). The former is expected to detect thousands of GW signals from low-mass black holes and neutron stars (Punturo et al. 2010), the latter is guaranteed to detect tens of thousands of white dwarf binaries in the Milky Way, and is also expected to detect signals from mergers involving

supermassive black holes (Amaro-Seoane et al. 2017). For these future detectors, one will *have to* take into account the possible presence of signals or high-SNR residuals lurking in the data. As this problem is only of peripheral relevance to analyses for the current LIGO-Virgo detector network, it has attracted relatively limited attention in the literature. In the context of ground-based detector networks, the detectability of confusion noise from a population of unresolved signals has been considered (Regimbau & Hughes 2009), but not the impact of the presence of that confusion foreground on parameter estimation for resolved sources. There have also been some recent Bayesian parameter estimation studies for second and third generation detectors, which computed the bias that arises in parameter estimation for a source due to the presence of another source with an overlapping merger (Samajdar et al. 2021; Pizzati et al. 2021; Himemoto, Nishizawa & Taruya 2021; Relton & Raymond 2021), and the impact of simultaneous fitting of two sources on the individual parameter precisions (Smith et al. 2021). These studies were limited to just two sources and did not consider the impact of waveform modelling uncertainties. In the LISA context, there have been studies of the detectability of confusion foregrounds from unresolved extreme mass-ratio inspirals (Bonetti & Sesana 2020), and extensive exploration of the simultaneous global-fit of the thousands of galactic binary sources expected to be present

⋆ andrea.antonelli@aei.mpg.de
† ollie.burke@aei.mpg.de
‡ jonathan.gair@aei.mpg.de

in LISA data (Robson & Cornish 2017; Littenberg et al. 2020; Karnesis et al. 2021). The latter global-fit analyses tackle the problem head-on by considering the simultaneous inference on parameters of an unknown number of sources in the data stream. Clearly, this is a formidable task due to the exceptionally large parameter space and complexity of the likelihood surface. It is thus important to have other approaches to confirm the results.

In this paper, we focus on analytic methods that can be used to cheaply assess when confusion from other sources, or imperfect subtraction of those sources due to waveform errors, is likely to be problematic, in the sense of leading to significant biases in parameter estimation for a source of interest. We leverage existing metrics for the "goodness" of individual waveform models based on the linear signal (Fisher matrix) formalism to derive generic metrics to assess the inference biases on source parameter characterisations. We describe how to apply this approach to several cases of relevance: i) parameter estimation in the presence of "confusion noise" from unfitted signals in the data; ii) parameter estimation for two overlapping signals with approximately coincident coalescence times; iii) parameter estimation for a population of sources using inaccurate waveform models; and iv) the case in which both confusion noise and mismodelling errors contribute to the final biases. Finally, we will show how these results can be used to mitigate biases in a sequential-fitting pipeline for LISA, without the need for expensive simultaneous fitting of identified sources. Our analysis is related to previous work by (Flanagan & Hughes 1998), (Miller 2005) and (Cutler & Vallisneri 2007), in which expressions are provided for the error on parameters due to the presence of noise and due to waveform errors. While their work has been mainly considered in the context of individual signals in the data, two observations make it relevant and easily extendible to the above applications. Firstly, no assumptions are made on the source of the noise appearing in their expressions, meaning that the observed noise can be made into a linear combination of detector noise and confusion noise [with applications to points i) and ii)]. Secondly, no assumptions are made about the dimensions of the parameter space, meaning that expressions relevant to points iii) and iv) can be derived from them.

We illustrate these metrics for several cases of relevance to future ground-based and space-based detectors. We take the ET and LISA instruments as our examples and use simplified, but realistic models for the gravitational waveforms. We consider the following, increasingly more complex, situations:

• The parameter estimation of a single LISA massive black hole source in the presence of other unfitted massive black holes forming a foreground [Sec. (5.1)].
• The parameter estimation of a single ET source in the presence of an overlapping signal with time of coalescence a fraction of a second from the former [Sec. (5.2)].
• The parameter estimation of a single LISA source in the presence of two overlapping sources which have been incorrectly fitted out of the data [Sec. (5.3)].
• The simultaneous inference in LISA of a few overlapping sources, subject to waveform errors, detector noise and unresolved signals present in the data stream [Sec. (5.4)].

We find that unfitted foregrounds or incorrectly removed

sources may lead to significant biases [as discussed in sections Sec. (5.1) and Sec. (5.3)], but that biases from confusion noise and waveform inaccuracies could deconstructively interfere [as discussed in Sec. (5.4)]. We qualitatively confirm one of the main results of (Samajdar et al. 2021; Pizzati et al. 2021; Himemoto, Nishizawa & Taruya 2021; Relton & Raymond 2021) in Sec. (5.2), showing that biases arise when the difference between the coalescence times of two overlapping signals is smaller than a fraction of a second. We find that the formalism herein developed is capable of predicting the biases very well (as confirmed with MCMC analyses), which makes it a useful tool for exploratory studies of future detectors.

Finally, in Section 6, we introduce the *local-fit strategy* as a simplification to the global-fit strategy proposed in the context of LISA data analysis. This method separately fits the parameters of individual sources, and then uses the computationally-inexpensive formalism presented in (3) to correct the biases that result in these estimates from ignoring the other sources in the data. This approach is much more computationally efficient than simultaneously constraining the parameters of all the sources. We believe such an algorithm could be an important alternative method for confirmation of fully-coherent global-fit analyses.

The paper is organised as follows: Sec. (2) contains a review of basic data-analysis concepts needed throughout the paper; Sec. (3) contains the description of the Fisher formalism herein developed; Sec. (4) contains a brief review of our choices of waveform models; Sec. (5) discusses the illustrations of the formalism described above; Sec (6) describes the local-fit strategy; and Sec. (7) summarises our findings and describes some possible future avenues of investigation. In appendix (A), we discuss a geometrical interpretation for the errors from noise and the biases from mismodelling (Cutler & Vallisneri 2007); in appendix (B) we describe the numerical methods used to obtain the results reported in the previous sections; in appendix (C) we describe how we computed the Fisher matrices and how these were verified using MCMC analyses; finally, in appendix (D) we complement the LISA results of Sec. 5.4 with results for ET.

2   DATA ANALYSIS CONCEPTS

The data stream observed by a gravitational wave detector is a superposition of noise $n(t)$ intrinsic to the detector and a gravitational wave signal h_e with "true" parameters $\boldsymbol{\theta}_{\text{tr}}$

$$d(t) = h_e(t; \boldsymbol{\theta}_{\text{tr}}) + n(t). \qquad (1)$$

In general, the gravitational wave component is a combination of the signals from a number of individual sources. The consequences of this will be made explicit in Section 3. In this analysis, we make the usual assumption that the noise $n(t)$ is both stationary and Gaussian with zero mean. As a consequence of stationarity, the covariance of the noise in the frequency domain can be expressed by (Wiener et al. 1930; Khintchine 1934)

$$\langle \hat{n}(f)\hat{n}^{\star}(f') \rangle = \frac{1}{2}\delta(f - f')S_n(f). \qquad (2)$$

Here and throughout this paper, hatted quantities will denote the continuous time Fourier transform. In the above,

δ denotes the Dirac delta function and $\langle \cdot \rangle$ denotes an ensemble averaging process. The quantity $S_n(f)$ denotes the (one-sided) power spectral density, which describes the distribution of power of the noise in the frequency domain.

The "loudness" of a signal can be represented by the optimal matched filtering signal to noise ratio (SNR), which is given by

$$\rho^2 = (h|h) = 4 \int_0^\infty \frac{|\hat{h}(f)|^2}{S_n(f)} df \qquad (3)$$

where we have defined the inner product for real valued time-series,

$$(a|b) = 4\text{Re} \int_0^\infty \frac{\hat{a}(f)\hat{b}^\star(f)}{S_n(f)} df. \qquad (4)$$

To make inference on parameters, one requires a probabilistic model on the data stream for given unknown parameters $\boldsymbol{\theta}$. As the noise $n(t)$ is stationary and Gaussian, the Whittle (log) likelihood can be used (Whittle 1957)

$$\log p(d|\boldsymbol{\theta}) \propto -\frac{1}{2}(d - h_m|d - h_m). \qquad (5)$$

We note that the gravitational wave component of the data stream in Eq. (1), $h_e(t; \boldsymbol{\theta}_{\text{tr}})$, is the true signal which depends on parameters $\boldsymbol{\theta}_{\text{tr}}$ that we wish to infer. In (5), we are denoting the signal by h_m, to allow for the possibility that there is a difference between the approximate waveform templates used to analyse the data, and the true signal, h_e, present in the data stream.

Finally, to quantify the precision of measurements on parameters, we will make use of the the linear signal approximation (LSA) (Finn 1992). By considering a small perturbation $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{tr}} + \Delta\boldsymbol{\theta}$, one can expand the waveform model in the vicinity of the best-fit parameters as

$$h_m(t; \boldsymbol{\theta}) \approx h_m(t; \boldsymbol{\theta}_{\text{tr}}) + \partial_i h_m(t; \boldsymbol{\theta}_{\text{tr}})\Delta\theta^i, \qquad (6)$$

which is valid for $|\Delta\theta^i_{\text{bf}}/\theta^i_{\text{bf}}| \ll 1$. We are using the standard notation $\partial_i = \partial/\partial\theta^i$. Substituting (6) into (5) and restricting to the case that the model and true waveform agree, $h_m = h_e$ for all $\boldsymbol{\theta}$, one obtains

$$-2\log p(d|\boldsymbol{\theta}) = (\Delta\theta^i - \Delta\theta^i_{\text{noise}})\Gamma_{ij}(\Delta\theta^j - \Delta\theta^j_{\text{noise}}). \qquad (7)$$

$$\Delta\theta^i_{\text{noise}} = (\Gamma^{-1})^{ij}(\partial_j h|n), \qquad (8)$$

where Γ_{ij} is the Fisher matrix, with components

$$\Gamma_{ij} = (\partial_i h|\partial_j h). \qquad (9)$$

In the derivation of Eq.(7), we neglected higher order terms which scale like $\mathcal{O}(\rho^{-1})$. Thus this representation of the likelihood is only valid for high SNR. Notice that (7) is Gaussian and centered on $\theta^i_{\text{bf}} = \theta^i_{\text{tr}} + \Delta\theta^i_{\text{noise}}$. Defining the statistic $\widehat{\Delta\theta^i} = \Delta\theta^i_{\text{noise}}$, one observes

$$\mathbb{E}[\widehat{\Delta\theta^i}] = 0, \quad \text{Cov}(\widehat{\Delta\theta^i}, \widehat{\Delta\theta^j}) = (\Gamma^{-1})^{ij} + \mathcal{O}(\rho^{-1}). \qquad (10)$$

This implies that the statistic $\widehat{\Delta\theta^i}$ is unbiased with covariance equal to the inverse of the Fisher matrix. In other words, the shift in the peak of the likelihood due to noise fluctuations is consistent with its width.

In the derivation of (7), we assumed the model template was consistent with the true gravitational waveform in the data set. We can relax this assumption and now consider $h_e \neq h_m$, which leads to a mismodelling error

$\delta h(\boldsymbol{\theta}) = h_e(t; \boldsymbol{\theta}) - h_m(t; \boldsymbol{\theta})$. The maximum of the likelihood function is at a set of parameter values $\boldsymbol{\theta}_{\text{bf}}$ that are a solution to

$$(\partial_i h_m(t; \boldsymbol{\theta}_{\text{bf}})|d - h_m(t; \boldsymbol{\theta}_{\text{bf}})) = 0. \qquad (11)$$

Using the LSA (6) and considering a perturbation $\boldsymbol{\theta}_{\text{tr}} = \boldsymbol{\theta}_{\text{bf}} + \Delta\boldsymbol{\theta}$ and a data stream $d(t) = h_e(t; \boldsymbol{\theta}_{\text{tr}}) + n(t)$ including the true gravitational waveform, one obtains

$$d - h_m = n + \delta h(\boldsymbol{\theta}_{\text{tr}}) + h_m(\boldsymbol{\theta}_{\text{tr}}) - h_m(\boldsymbol{\theta}_{\text{bf}})$$
$$\approx n + \delta h(\boldsymbol{\theta}_{\text{bf}}) - \Delta\theta^i \partial_i h_m(\boldsymbol{\theta}_{\text{bf}}), \qquad (12)$$

where in the last line we take $\delta\vec{h}(\boldsymbol{\theta}_{\text{tr}}) \approx \delta\vec{h}(\boldsymbol{\theta}_{\text{bf}})$. With all waveform models evaluated at the best-fit parameters, we deduce that (Cutler & Vallisneri 2007)

$$(\partial_i h_m|d - h_m) \approx (\partial_i h_m|n) + (\partial_i h_m|\delta h) - \Delta\theta^j\Gamma_{ij} = 0,$$
$$\iff \Delta\theta^i = (\Gamma^{-1})^{ij}\left[(\partial_j h_m|n) + (\partial_j h_m|\delta h)\right]. \qquad (13)$$

where we now separate $\Delta\theta$ into an error from instrumental noise, $\Delta\theta^i_{\text{noise}}$ and a theoretical bias $\Delta\theta^i_{\text{sys}}$,

$$\Delta\theta^i_{\text{noise}} = (\Gamma^{-1})^{ij}(\partial_j h_m|n), \qquad (14)$$
$$\Delta\theta^i_{\text{sys}} = (\Gamma^{-1})^{ij}(\partial_j h_m|\delta h). \qquad (15)$$

This expression for systematic errors first appeared in (Flanagan & Hughes 1998), see their Eq. (6.11), although the implications were not studied in that paper. A much more through analysis was given in (Cutler & Vallisneri 2007). A geometrical intuition for the origin of Eqs. (14,15) is given in appendix A. A `python` tutorial on how to use equations (14) and (15) can be found here.

Generally speaking, a waveform model h_m is "good enough" for parameter estimation if and only if $\Delta\theta^i_{\text{sys}} \lesssim \Delta\theta^i_{\text{noise}}$. The quantity $\Delta\theta^i_{\text{noise}}$ is a zero-mean random variable, so this inequality should hold in an average sense. The 1σ deviation of $\Delta\theta^i_{\text{noise}}$ is $\Delta\theta^i_{\text{stat}} = \sqrt{\Gamma^{-1})^{ii}}$, so we define the function

$$\mathcal{R}(\Delta\theta) := |\Delta\theta^i/\Delta\theta^i_{\text{stat}}|, \qquad (16)$$

and consider biases on the parameter θ arising from systematic effects to be significant whenever $\mathcal{R}(\Delta\theta) > 1$. To conclude this section, we note that the statistical error $\Delta\boldsymbol{\theta}_{\text{stat}} \sim \rho^{-1}$, while the systematic error $\Delta\boldsymbol{\theta}_{\text{sys}} \sim \rho^0$. This implies that biases from modelling errors are independent of the SNR, while statistical errors become smaller as the SNR increases. Therefore, we expect systematics to become more important for loud sources.

3 GENERALISATIONS

We now generalise the formalism represented by Eqs. (14) and (15) to two new cases, the first being the presence of *confusion noise* from signals that have not been fitted for in parameter estimation, and the second being the inclusion of multiple signals in the data stream that are incorrectly modelled with approximate waveforms.

The likelihood (5) only assumes that the noise $n(t)$ is both stationary and Gaussian (with zero mean). The noise $n(t)$ is usually assumed to be instrumental and modelled through the PSD via (2). However, in third-generation or

space-based detectors there may be additional astrophysical contributions to the data stream from unresolved foregrounds of other GW signals (Crowder & Cornish 2007; Błaut, Babak & Królak 2010; Robson & Cornish 2017; Roebber et al. 2020; Korol et al. 2020; Samajdar et al. 2021; Pizzati et al. 2021; Karnesis et al. 2021). This confusion noise ΔH_{conf} can be represented as part of the signal component of the data stream (1),

$$d(t) = h_e(t; \boldsymbol{\theta}_{\mathrm{tr}}) + n(t) + \Delta H_{\mathrm{conf}}(t; \boldsymbol{\theta}^{(i)}) . \tag{17}$$

To understand when such confusion foregrounds can lead to biases, one may consider it to be a (deterministic) superposition of N signals,

$$\Delta H_{\mathrm{conf}}(t; \boldsymbol{\theta}^{(i)}) = \sum_{i=1}^{N} h_e^{(i)}(t; \boldsymbol{\theta}^{(i)}) . \tag{18}$$

Equation (12) now becomes

$$d - h_m = n + \Delta H_{\mathrm{conf}} + \delta h(\boldsymbol{\theta}_{\mathrm{tr}}) + h_m(\boldsymbol{\theta}_{\mathrm{tr}}) - h_m(\boldsymbol{\theta}_{\mathrm{bf}}) , \tag{19}$$

from which we deduce the extra contribution to the biases (14) and (15) that originates from the source confusion term is

$$\Delta \theta_{\mathrm{conf}}^i = (\Gamma^{-1})^{ij}(\partial_j h_m | \Delta H_{\mathrm{conf}}). \tag{20}$$

By analogy with (16), source confusion from unfitted signals can be said to bias parameter estimates when its size exceeds the 1σ deviations arising from instrumental noise fluctuations, which is true if $\mathcal{R}(\Delta\theta_{\mathrm{conf}}) > 1$. To summarise, when inferring the parameters of a single source, the total error is given by the sum of statistical error from noise fluctuations and the biases from source confusion and waveform errors through

$$\Delta \theta^i = \Delta \theta_{\mathrm{noise}}^i + \Delta \theta_{\mathrm{sys}}^i + \Delta \theta_{\mathrm{conf}}^i, \tag{21}$$

with the above terms from left to right given by Eqs.(14,15,20) respectively.

In general, the confusion noise contribution to (21) depends on the particular sources from the unresolved population that are present in the data and so it is a random quantity. The correct way to handle this is to marginalise the likelihood of the corrected data stream, $d(t) - \Delta H_{\mathrm{conf}}(t; \boldsymbol{\theta}^{(i)})$, over the distribution of possible confusion backgrounds, $p(\Delta H_{\mathrm{conf}})$. This is a computationally expensive procedure and it is therefore difficult to obtain insights in that way. An alternative avenue to understanding when confusion is important is to use the formalism described here, working with the bias induced by the confusion noise, $\Delta\theta_{\mathrm{conf}}^i$, which is also a random quantity. We can characterise it at the order of the linear signal approximation through its mean and variance. Since the total confusion noise contribution is a superposition of contributions from N independent sources, the mean and variance of the total contribution is N times the mean and variance of the contribution from a single source, which are

$$\mu_{\mathrm{conf}}^i = \int (\Gamma^{-1})^{ij}(\partial_j h_m | h_e(\boldsymbol{\theta}_{\mathrm{conf}})) \, p_{\mathrm{pop}}(\boldsymbol{\theta}_{\mathrm{conf}}) \, d\boldsymbol{\theta}_{\mathrm{conf}} , \tag{22}$$

$$\Sigma_{\mathrm{conf}}^{ij} = \int (\Gamma^{-1})^{ik}(\partial_k h_m | h_e(\boldsymbol{\theta}_{\mathrm{conf}})) \times$$
$$(\Gamma^{-1})^{jl}(\partial_l h_m | h_e(\boldsymbol{\theta}_{\mathrm{conf}})) \, p_{\mathrm{pop}}(\boldsymbol{\theta}_{\mathrm{conf}}) \, d\boldsymbol{\theta}_{\mathrm{conf}}$$
$$- \mu_{\mathrm{conf}}^i \mu_{\mathrm{conf}}^j. \tag{23}$$

Here, $p_{\mathrm{pop}}(\boldsymbol{\theta}_{\mathrm{conf}})$ is the probability density function of the population of confusion sources. We would normally expect the mean to be close to zero, since for some sources in the population the bias would be positive and others negative and so it averages to zero (though this is not guaranteed to be the case). Regardless, the variance does not vanish, driving the total error to grow like a random walk as the total number of sources contributing to the confusion background increases. For large N we expect the total bias to scale like

$$\Delta\theta^i - N\mu_{\mathrm{conf}}^i \sim \sqrt{N}(\Sigma_{\mathrm{conf}}^{ii})^{1/2}. \tag{24}$$

This point will be investigated further in Sec.(5.1). When the confusion noise is generated by a very large population of sources, it is common to treat it analogously to the instrumental noise by writing

$$\langle \widehat{\Delta H_{\mathrm{conf}}}(f) \rangle = 0, \tag{25}$$

$$\langle \widehat{\Delta H_{\mathrm{conf}}}(f) \widehat{\Delta H_{\mathrm{conf}}^{\star}}(f') \rangle = \frac{1}{2}\delta(f - f') S_{\mathrm{conf}}(f), \tag{26}$$

$$\langle \widehat{\Delta H_{\mathrm{conf}}}(f) \widehat{\Delta H_{\mathrm{conf}}}(f') \rangle = 0 \tag{27}$$

Under these assumptions, the mean bias is zero and the covariance from the confusion background takes the alternative form [1]

$$N\Sigma_{\mathrm{conf}}^{ij} = (\Gamma^{-1})^{ik}(\Gamma^{-1})^{jl}$$
$$\left\langle \int_{-\infty}^{\infty} \frac{(\partial_k^* h_m(f)\widehat{\Delta H}(f) + \partial_k h_m(f)\widehat{\Delta H}^*(f))}{S_n(f)} \, df \right.$$
$$\left. \int_{-\infty}^{\infty} \frac{(\partial_l^* h_m(f')\widehat{\Delta H}(f') + \partial_l h_m(f')\widehat{\Delta H}^*(f'))}{S_n(f')} \, df' \right\rangle$$
$$= (\Gamma^{-1})^{ik}(\Gamma^{-1})^{jl}$$
$$\int_{-\infty}^{\infty} \frac{(\partial_k h_m^{\star}(f)\partial_l h_m(f) + \partial_l h_m^{\star}(f)\partial_k h_m(f)) S_{\mathrm{conf}}(f)}{S_n^2(f)} \, df. \tag{28}$$

If we use this prescription within the formalism we have here described, we can calculate the total covariance in the parameter estimates arising from instrumental noise and source confusion, which is $\langle (\Delta\theta_{\mathrm{noise}}^i + \Delta\theta_{\mathrm{conf}}^i)(\Delta\theta_{\mathrm{noise}}^j + \Delta\theta_{\mathrm{conf}}^j) \rangle = \Gamma^{-1} + N\Sigma_{\mathrm{conf}}$ (since $\langle \Delta\theta_{\mathrm{noise}}^i \Delta\theta_{\mathrm{conf}}^i \rangle = 0$ as the instrumental and astrophysical noises should not depend on one another). By making use of Eqs. (4), (23) and (26), the expression for the total variance can be simplified to

$$\left[\Gamma^{-1} + \Sigma_{\mathrm{conf}} \right]^{ij} = (\Gamma^{-1})^{ik} \Sigma_{\mathrm{mix}}^{kl} (\Gamma^{-1})^{jl}, \tag{29}$$

where

$$\Sigma_{\mathrm{mix}}^{ij} = 4\mathrm{Re} \int_0^{\infty} \frac{(\partial_k \hat{h}_m(f)\partial_l \hat{h}_m^{\star}(f))(S_{\mathrm{conf}}(f) + S_n(f))}{S_n^2(f)}. \tag{30}$$

In contrast to this, the standard approach when modelling the confusion background is to combine the instrumental and confusion noises into a single noise term, $N = n + \Delta H_{\mathrm{conf}}$. Then the standard parameter estimation formalism can be used, with the substitution $S_n(f) \rightarrow S_n(f) + S_{\mathrm{conf}}(f)$

---

[1] Note that S_{conf} describes the contribution from the whole astrophysical population, while Σ_{conf} defined in Eq. (23) was the contribution from a single source in the population. For consistency, we therefore denote the total covariance by $N\Sigma_{\mathrm{conf}}$ in Eq. (28).

in the inner product (4). In this case the inference uncertainties are given by the inverse of the Fisher matrix, $\Gamma^{-1}_{\mathrm{n+conf}}$, where

$$\Gamma^{ij}_{\mathrm{n+conf}} = 4\mathrm{Re}\int_0^\infty \frac{(\partial_k\hat{h}_m(f)\partial_l\hat{h}_m^\star(f))}{S_{\mathrm{n}}(f) + S_{\mathrm{conf}}(f)}. \tag{31}$$

The variance given by Eq. (29) is, in general, larger than that predicted by Eq. (31). This is because it has been derived by maximizing the standard likelihood as an estimator of the parameters, which is no longer the correct likelihood when random confusion noise is included in the model. Expression (31) gives the variance of the true maximum likelihood estimator, which is known to be the minimum variance unbiased estimator and must therefore be smaller than (29). Incorporating the confusion noise uncertainty into the PSD is the correct thing to do when Eq. (26) is known to be a good approximation, but the formalism described here can be used when that equation is not valid, and to assess when confusion noise is likely to be problematic for parameter estimation. As a final remark, we note that in the limit that there are a large number of sources contributing to the confusion background, the central limit theorem allows us to approximate the probability distribution of the parameter bias correction, $p(\Delta\boldsymbol{\theta}_{\mathrm{conf}})$, as a Gaussian with mean $\boldsymbol{\mu}_{\mathrm{conf}}$ and covariance Σ_{conf}. The correct statistical procedure of marginalising the likelihood for $d(t) - \Delta H(t)$ over the confusion noise distribution thus amounts, in the linear signal approximation, to shifting the mean by $\boldsymbol{\mu}_{\mathrm{conf}}$ and adding Σ_{conf} to the covariance. The results described here can therefore be used not only to assess when confusion is important but also to compute leading order corrections to posterior parameter estimates arising from the presence of confusion.

We now generalise Equations (17) and (21) to the case of inference on multiple sources within the data stream. Similar ideas can be found in (Robson & Cornish 2017) for the case of massive black holes and galactic binaries in LISA. Here, we extend their discussion and include a prescription for the effect of waveform errors and confusion noise, generalising their results to multiple source types with an arbitrary number of sources. We suppose there are J different types of source in the data. We suppose that there are N_j sources of type j in the data stream, indexed by i, which each depend on a set of m_j parameters, denoted by $\boldsymbol{\theta}_i^{(j)}$, which determine the corresponding gravitational waveform, $h^j(t;\boldsymbol{\theta}_i^{(j)})$. The complete data stream can be written as

$$d(t) = h(t;\boldsymbol{\Theta}) + n(t) + \Delta H_{\mathrm{conf}}$$
$$= \sum_{j=1}^{J}\sum_{i=1}^{N_j} h_e^{(j)}(t;\boldsymbol{\theta}_i^{(j)}) + n(t) + \Delta H_{\mathrm{conf}}. \tag{32}$$

Here we have introduced a composite vector of parameters, $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_i^{(j)}\}_{i=1,\ldots,N_j}^{j=1,\ldots,J}$, such that $\Theta_{N_{<j}+(i-1)m_j+k} = (\boldsymbol{\theta}_i^{(j)})_k$, where $N_{<j} = \sum_{l=1}^{j-1} N_l m_l$. For any given parameter in $\boldsymbol{\Theta}$, there is exactly one waveform in the above sum that depends on that parameter. Thus the derivatives of the signal reduce to derivatives of the specific waveform template. The combined Fisher matrix has a block structure, with the on-diagonal blocks being the Fisher matrices for the individual sources, and the off-diagonal blocks being formed from overlaps of waveform derivatives of one source with waveform derivatives of another source. Through calculating the Fisher matrix on parameters $\boldsymbol{\Theta}$, one is able to estimate the

expected precision of measurements on individual parameters, taking into account *all* parameter correlations. This is (an estimate for) the precision that would be achieved in a simultaneous coherent fit to all sources in the data.

Without loss of generality, we illustrate this considering two classes of sources, with one source in the first class ($j = 1$, $N_1 = 1$) and an arbitrary number N_2 of sources in the second ($j = 2$). This split is only made for ease of exposition, and is quite arbitrary as the sources could always be relabelled so that the first source is the source of interest. We want to estimate the impact of confusion due to the presence of the population of (fitted) sources of type 2, on the precision of parameter estimation for source 1. We define the following quantities

$$\Gamma_{jk}^{(1)} = \left(\partial_j h^{(1)}(\boldsymbol{\theta}^{(1)})\middle|\partial_k h^{(1)}(\boldsymbol{\theta}^{(1)})\right) \tag{33}$$

$$\left(\Gamma_i^{(2)}\right)_{jk} = \left(\partial_j h^{(2)}(\boldsymbol{\theta}_i^{(2)})\middle|\partial_k h^{(2)}(\boldsymbol{\theta}_i^{(2)})\right) \tag{34}$$

$$\left(\Gamma_i^{\mathrm{mix}}\right)_{jk} = \left(\partial_j h^{(1)}(\boldsymbol{\theta}^{(1)})\middle|\partial_k h^{(2)}(\boldsymbol{\theta}_i^{(2)})\right). \tag{35}$$

Here $\Gamma^{(1)}$ is the Fisher matrix for the source of type 1, $\Gamma_i^{(2)}$ is the Fisher matrix for the i'th source of type 2 ($i = 1,\ldots,N_2$) and Γ_i^{mix} is the mixed Fisher matrix for the source of type 1 and the i'th source of type 2. In what follows, we find it useful to combine the Fisher matrix contributions of the entire population of sources in a more compact form. One can write Eqs.(33-35) as

$$\left(\Gamma^{(2)}\right)_{m_2(i-1)+j,m_2(l-1)+k} = \left(\partial_j h^{(2)}(\boldsymbol{\theta}_i^{(2)})\middle|\partial_k h^{(2)}(\boldsymbol{\theta}_l^{(2)})\right) \tag{36}$$

$$\Gamma^{\mathrm{mix}}_{j,m_2(i-1)+k} = (\partial_j h^{(1)}(\boldsymbol{\theta}^{(1)})|\partial_k h^{(2)}(\boldsymbol{\theta}_i^{(2)})). \tag{37}$$

The Fisher matrix for the full analysis and its inverse are therefore

$$\Gamma = \left(\begin{array}{cc} \Gamma^{(1)} & \Gamma^{\mathrm{mix}} \\ (\Gamma^{\mathrm{mix}})^T & \Gamma^{(2)} \end{array}\right); \quad \Gamma^{-1} = \left(\begin{array}{cc} \Gamma_{11}^{-1} & \Gamma_{12}^{-1} \\ (\Gamma_{12}^{-1})^T & \Gamma_{22}^{-1} \end{array}\right) \tag{38}$$

with the components of the inverse[2]

$$\Gamma_{11}^{-1} = \left(\Gamma^{(1)} - \Gamma^{\mathrm{mix}}(\Gamma^{(2)})^{-1}(\Gamma^{\mathrm{mix}})^T\right)^{-1}, \tag{39}$$

$$\Gamma_{22}^{-1} = \left(\Gamma^{(2)} - (\Gamma^{\mathrm{mix}})^T(\Gamma^{(1)})^{-1}\Gamma^{\mathrm{mix}}\right)^{-1}, \tag{40}$$

$$\Gamma_{12}^{-1} = -\Gamma_{11}^{-1}\Gamma^{\mathrm{mix}}(\Gamma^{(2)})^{-1}. \tag{41}$$

The components Γ_{11}^{-1} encode the measurement precisions for source 1. If the degree of correlation between the source types is small, i.e., $\Gamma^{\mathrm{mix}} \ll 1$, we can approximate this as

$$\Gamma_{11}^{-1} \approx (\Gamma^{(1)})^{-1} + (\Gamma^{(1)})^{-1}\Gamma^{\mathrm{mix}}(\Gamma^{(2)})^{-1}(\Gamma^{\mathrm{mix}})^T(\Gamma^{(1)})^{-1}. \tag{42}$$

The first term is the measurement precision when there are no sources in the data, while the second represents the degradation in the precision due to confusion with the other sources. We can understand the form of the second term as

---

[2] We note also that

$$\Gamma_{11}^{-1} = (\Gamma^{(1)})^{-1} + (\Gamma^{(1)})^{-1}\Gamma^{\mathrm{mix}}\Gamma_{22}^{-1}(\Gamma^{\mathrm{mix}})^T(\Gamma^{(1)})^{-1}$$
$$\Gamma_{22}^{-1} = (\Gamma^{(2)})^{-1} + (\Gamma^{(2)})^{-1}\Gamma^{\mathrm{mix}}\Gamma_{11}^{-1}(\Gamma^{\mathrm{mix}})^T(\Gamma^{(2)})^{-1}$$

which can sometimes be cheaper to compute than Eq. (40).

follows. If the other sources were ignored when fitting for source 1, the parameter bias would be given by Eq. (15)

$$\Delta\theta_{\rm sys}^{(1),i} = (\Gamma^{(1)})_{ij}^{-1}(\partial_j h_m^{(1)}|\boldsymbol{h}^{(2)}) \tag{43}$$

where we are combining all of the sources of type 2 into the single term $\boldsymbol{h}^{(2)}$. This bias is dominated by the contribution from the true waveform. When we simultaneously fit for the sources of type 2, we imperfectly remove these signals, leaving a residual in the data of the form $\partial_j \boldsymbol{h}^{(2)}\Delta\theta_2^j$, where again we are combining the parameters of all of the sources of type 2 into a single parameter vector, $\boldsymbol{\theta}_2$. The parameter error, $\Delta\boldsymbol{\theta}_2$, is a random variable with covariance matrix $\langle\Delta\theta_2^j\Delta\theta_2^k\rangle = (\Gamma^{(2)})_{jk}^{-1}$. The covariance of the induced systematic error in the parameters of source 1 is then

$$\langle\Delta\theta_{\rm sys}^{(1),i}\Delta\theta_{\rm sys}^{(1),j}\rangle = (\Gamma^{(1)})_{ik}^{-1}(\partial_k h_m^{(1)}|\partial_l\boldsymbol{h}^{(2)})\langle\Delta\theta_2^l\Delta\theta_2^m\rangle$$
$$(\partial_n h_m^{(1)}|\partial_m\boldsymbol{h}^{(2)})(\Gamma^{(1)})_{jm}^{-1}$$
$$= \left[(\Gamma^{(1)})^{-1}\Gamma^{\rm mix}(\Gamma^{(2)})^{-1}(\Gamma^{\rm mix})^T(\Gamma^{(1)})^{-1}\right]_{ij},$$

which is the second term from Eq. (42). There is nothing that can be done to mitigate uncertainties of this type, which arise from an over-abundance of sources in the data. However, as described above, additional uncertainties can arise from their inaccurate modelling. Previous studies have focused on biases from inaccurate modelling of the target source, but it is also important to ask if the inaccurate modelling of a large number of other sources can leave a sufficient residual in the data to cause problems.

To estimate this, we define $\delta h^{(1)} = h_e^{(1)} - h_m^{(1)}$ as the difference between the exact h_e and template h_m waveforms for the source of type 1, and similarly $\delta h_i^{(2)} = h_e^{(2)}(\boldsymbol{\theta}_i^{(2)}) - h_m^{(2)}(\boldsymbol{\theta}_i^{(2)})$ for the i'th source of type 2. We also define $\delta h = \delta h^{(1)} + \sum_{i=1}^{N_2}\delta h_i^{(2)}$ as the combination of all waveform residuals. Let us define the bias vector $\boldsymbol{b}$

$$\boldsymbol{b} = (b_1^{(1)}, \ldots, b_{m_1}^{(1)}, (b_1^{(2)})_1, \ldots,$$
$$(b_1^{(2)})_{m_2}, \ldots, (b_{N_2}^{(2)})_1, \ldots, (b_{N_2}^{(2)})_{m_2})^T,$$

such that $\boldsymbol{b} = [\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}] \in \mathbb{R}^{(m_1+N_2 m_2)\times 1}$ with individual components given by

$$b_j^{(1)} = (\partial_j h^{(1)}(\boldsymbol{\theta}^{(1)})|\delta h), \tag{44}$$
$$(b_i^{(2)})_j = (\partial_j h^{(2)}(\boldsymbol{\theta}_i^{(2)})|\delta h).$$

Note that the bias defined here is only the contribution from modelling errors. The full shift in the peak of the likelihood may be found from a similar expression, with $n(t)$ and $\Delta H_{\rm conf}$ added to δh in the inner products. The quantity $b_j^{(1)}$ for $j = 1, \ldots, m_1$ are the components $\boldsymbol{b}$ for the first source of type 1. The quantity $(b_i^{(2)})_j$ are the jth components of $\boldsymbol{b}$ with respect to the ith source of type 2. The vector $\boldsymbol{b}$ can be written more concisely as

$$b_j = b_j^{(1)} \quad \text{for } j = 1, \ldots, m_1, \tag{45}$$
$$b_{m_1+m_2(i-1)+j} = (b_i^{(2)})_j \text{ for } i = 1, \ldots, N_2; j = 1, \ldots, m_2, \tag{46}$$

The biases computed from Eq. (15) are given by $\Delta\boldsymbol{\Theta} = \Gamma^{-1}\boldsymbol{b}$ and are thus

$$\Delta\boldsymbol{\Theta} := \begin{pmatrix}\Delta\boldsymbol{\theta}^{(1)} \\ \Delta\boldsymbol{\theta}^{(2)}\end{pmatrix} = \Gamma^{-1}\begin{pmatrix}\boldsymbol{b}^{(1)} \\ \boldsymbol{b}^{(2)}\end{pmatrix}, \tag{47}$$

Using Eqs.(36-38) and Eqs.(45,46), the bias in the source parameters of the signal of type 1 is

$$\Delta\theta_i^{(1)} = (\Gamma_{11}^{-1})^{ij}b_j + (\Gamma_{12}^{-1})^{im}b_{m_1+m}, \tag{48}$$

with components of (Γ_{11}^{-1}) and (Γ_{12}^{-1}) defined in Eqs.(39,41). Using the approximation that led to Eq. (42), that the elements of $\Gamma^{\rm mix}(\Gamma^{(2)})^{-1}(\Gamma^{\rm mix})^T$ are much smaller than those of $\Gamma^{(1)}$, we can approximate Eq. (48) as

$$\Delta\theta_i^{(1)} \approx [(\Gamma^{(1)})^{-1}]^{ij}b_j - $$
$$[(\Gamma^{(1)})^{-1}]^{ij}(\Gamma^{\rm mix})_{jl}[(\Gamma^{(2)})^{-1}]^{lm}b_{m_1+m} \tag{49}$$

We see that there are two contributions to the parameter bias on the single source of type 1: the standard CV bias (15) arising from mismodelling of that source; and an extra correction due to mismodelling of overlapping sources. If the sources from each source type are orthogonal, $\Gamma^{\rm mix} \to 0$ and the presence of other sources does not contribute a parameter bias.

In testing the formalism below, we drop the source type indices for simplicity. The waveform and shift in the peak of the likelihood will be denoted

$$h(t; \boldsymbol{\Theta}) = \sum_{i=1}^{N} h_e(t; \boldsymbol{\theta}_i) \quad \text{for } \boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N\},$$
$$\Delta\Theta^i = (\Gamma^{-1})^{ij}\left(\frac{\partial h}{\partial\Theta^j}\bigg|n(t) + \delta h + \Delta H_{\rm conf}\right)$$
$$= \Delta\Theta_{\rm noise}^i + \Delta\Theta_{\rm sys}^i + \Delta\Theta_{\rm conf}^i, \tag{50}$$

with total theoretical error δh and Fisher matrix denoted

$$\Gamma_{ij} = \left(\frac{\partial h}{\partial\Theta^i}\bigg|\frac{\partial h}{\partial\Theta^j}\right), \quad \delta h = \sum_{i=1}^{N}(h_e^{(1)}(t; \boldsymbol{\theta}_i^{(1)}) - h_m^{(1)}(t; \boldsymbol{\theta}_i^{(1)})).$$

The Γ appearing in (50) is the joint Fisher matrix $\Gamma \in \mathbb{R}^{(N\times m)\times(N\times m)}$, with m the dimension of each parameter space $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$. Equation (50) is separated into a noise induced error, $\Delta\Theta_{\rm noise}^i$, and biases split into a confusion noise contribution, $\Delta\Theta_{\rm conf}^i$, and a contribution from theoretical errors, $\Delta\Theta_{\rm sys}^i$. Biases are then significant whenever $\mathcal{R}(\Delta\Theta_{\rm conf}^i + \Delta\Theta_{\rm sys}^i) > 1$.

4 MODELLING SIGNALS AND NOISE

To illustrate the above formalism, we will consider a number of simplified scenarios. For all of these we will model the signals using the `TaylorF2` waveform model

$$\hat{h}(f) = \mathcal{A}\left(\frac{\pi GMf}{c^3}\right)^{-7/6}e^{-i\psi(f)}, \tag{51}$$

$$\mathcal{A} = -\sqrt{\frac{5}{24}}\frac{c}{D_{\rm eff}\pi^{2/3}}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/6}. \tag{52}$$

Here, $M_c := M\eta^{3/5}$ is the chirp mass and $D_{\rm eff}$ the effective distance. Notice that we ignore the effect of the detector response function, which only affects the SNR of the binary and not its frequency evolution. For this reason $D_{\rm eff}$ should be treated effectively as an overall scaling factor, and not as a physical distance parameter. We retain only the leading-order amplitude $\mathcal{A}$ (Allen et al. 2012) in the waveform. The

phase is PN-expanded in the velocity $v := (\pi M G f/c^3)^{1/3}$ and reads

$$\psi(f) = 2\pi f t_c - \phi_c + \frac{3v^{-5}}{128\eta}\left(1 + \sum_{n=2}^{n=7} v^n \psi_{\frac{n}{2}\mathrm{PN}}\right), \quad (53)$$

with coefficients up to 3.5PN as given in Sec.IIIB of (Cutler & Vallisneri 2007). The constant portion of the phase depends on the time and phase at coalescence, t_c and ϕ_c. We have only included spin-orbit interactions in the 1.5PN phase through the spin parameter β, defined in (Berti, Buonanno & Will 2005). We remark that β satisfies the inequality $|\beta| \lesssim 9.4$. We take the above `TaylorF2` model to be the exact waveform $\hat{h}_e(f; \boldsymbol{\theta})$. In these examples, for simplicity we will treat the phase, ϕ_c, time of coalescence, t_c, and distance, D_{eff}, as perfectly-known parameters. The purpose of these examples is to illustrate the formalism of Sec.(3) and so this is not a serious restriction.

Finally, to represent modelling inaccuracies, we represent the approximate waveform by modifying the smallest contribution in the 3.5PN phase contribution

$$\psi_{3.5\mathrm{PN}}^{(\epsilon)} := \pi\left(\frac{77096675}{254016} + \frac{378515}{1512}\eta - (1-\epsilon)\frac{74045}{756}\eta^2\right), \quad (54)$$

(for $\epsilon \in [0,1]$). The true PN waveform has $\epsilon = 0$ and we will take a (fixed) value of $\epsilon \neq 0$ to represent the approximate model. Finally, we model confusion noise as a superposition of `TaylorF2` models, unless otherwise specified (see Sec. 5.4).

We generate detector noise in both ET and LISA using Eq. (2) and the PSDs found in (Robson, Cornish & Liu 2019) (LISA) and (Regimbau et al. 2012) (ET). More details on how we generate our signals and noise realisations are found in Appendix B. In Appendix C we describe how the waveform derivatives (and Fisher matrices) are calculated, and outline the MCMC techniques used to verify them.

5 RESULTS

In this section, we present four illustrations for the formalism described in Sec. 3. The first one concerns confusion and detector noise only. The second concerns the overlap of two signals with coincident coalescence. The third concerns theoretical errors from incorrectly removed waveforms only. The fourth considers all of the above combined.

5.1 Biases from detector and confusion noise

In this exploration, we consider a single reference signal in the LISA band and a confusion noise ΔH_{conf} of binaries that follow a realistic mass distribution. Our aim is to understand how much the combined effects of the confusion signals affect recovery of the parameters of the reference signal, and whether we can predict the biases using the formalism described above. The data stream we consider is

$$\hat{d}(f) = \hat{h}_e(f; \boldsymbol{\theta}_{\mathrm{tr}}) + \Delta H_{\mathrm{conf}}(f; \boldsymbol{\theta}) + \hat{n}(f). \quad (55)$$

We recover the reference signal perfectly by modelling it with the exact waveform $\hat{h}_e(f; \boldsymbol{\theta}_{\mathrm{tr}})$ of (51) in both the Fisher matrix and the MCMC sampling algorithms. We therefore

expect no biases from modelling errors. We use the following configuration of true (injected) parameters,

$$\boldsymbol{\theta}_{\mathrm{tr}} = \{\log \mathcal{M}_c = 83.34, \eta = 0.210, \beta = 5.00\}, \quad (56)$$

which correspond to a spinning binary of total mass $M = 2 \times 10^6 M_\odot$. We complete the full set of parameters by choosing an effective distance $D_{\mathrm{eff}} = 1\,\mathrm{Gpc}$ and phase at coalescence $\phi_c = 0$, with time at coalescence given by the chirp time (see appendix B). We begin observing the binary at $f_0 = 0.25\,m\mathrm{Hz}$ and stop at $f_{\mathrm{max}} = 2.2\,m\mathrm{Hz}$, corresponding to the ISCO frequency in a Schwarzschild spacetime for the chosen total mass. That is, we observe the binary until it chirps ~ 4.4 days after we have started observing it. These choices lead to an SNR of $\rho \sim 4200$ for this signal, for which we expect the Fisher formalism to be a very good approximation.

To construct $\Delta H_{\mathrm{conf}}(f; \boldsymbol{\theta})$, we first build a mock catalogue of $N = 1000$ sources, which are sampled from uniform distributions

$$
\begin{aligned}
D_{\mathrm{eff}} &\sim U[1,5]\ \mathrm{Gpc} \\
\beta &\sim U[0.001, 9.4] \\
\eta &\sim U[0.001, 0.05] \\
\phi_c &\sim U[0, 2\pi]
\end{aligned}
\quad (57)
$$

and t_c given by the individual chirp times. We let the total masses of the binaries in this catalog follow a standard probability density function for massive black holes (Gair et al. 2011; Gair, Tang & Volonteri 2010; Sesana et al. 2011),

$$\frac{dN}{dM} = \frac{\alpha\, M^{\alpha-1}}{M_{\mathrm{max}}^\alpha - M_{\mathrm{min}}^\alpha}, \quad (58)$$

where the masses' range is $M_{\mathrm{min}} = 10^4 M_\odot < M < 10^7 M_\odot = M_{\mathrm{max}}$ and $\alpha = 0.03$ is the fit in Ref. (Gair, Tang & Volonteri 2010) to the inactive massive black holes of (Greene & Ho 2007). We can directly sample the total masses using

$$\log M = \alpha^{-1} \log\left[(M_{\mathrm{max}}^\alpha - M_{\mathrm{min}}^\alpha)\, u + M_{\mathrm{min}}^\alpha\right], \quad (59)$$

with $u \sim U[0,1]$. For each element of the catalogue, we compute the waveform of the binary using the exact model h_e. For those mass draws for which the frequency array of the binary is longer than that of the reference signal, we cut the former to be of the same length as the latter. Otherwise, we stop the evolution of the binary at its ISCO to avoid introducing an artificial portion of the waveform into the analysis. If the waveform has an observed SNR $\rho_{\mathrm{obs}} = \rho + N(0,1)$ (Sathyaprakash & Schutz 2009), where $N(0,1)$ is a standard normal distribution, such that $\rho_{\mathrm{obs}} < \rho_{\mathrm{threshold}} = 15$ then we consider the binary as "missed", retain the waveform and add it to ΔH_{conf} in a cumulative fashion. In our example, for $N = 1000$ events in the mock catalogue, $N_{\mathrm{U}} = \mathcal{O}(175)$ have SNRs below the threshold and are thus unresolved. The final SNR of ΔH_{conf} is ~ 85 in this case.

Once ΔH_{conf} is obtained and the data stream (55) is thus fully specified, we predict the biases from confusion noise $\Delta\boldsymbol{\theta}_{\mathrm{conf}}$ [namely, using (50) retaining only ΔH_{conf} in the bias vector], which we can compare to the statistical error $\Delta\boldsymbol{\theta}_{\mathrm{noise}}$ [found from (50) with n only]. We show the accumulation of the biases from confusion noise in Fig. 1 by plotting the ratio $\mathcal{R}(\Delta\boldsymbol{\theta}_{\mathrm{conf}})$. In this plot, calculations with different numbers of sources use different noise realisations,
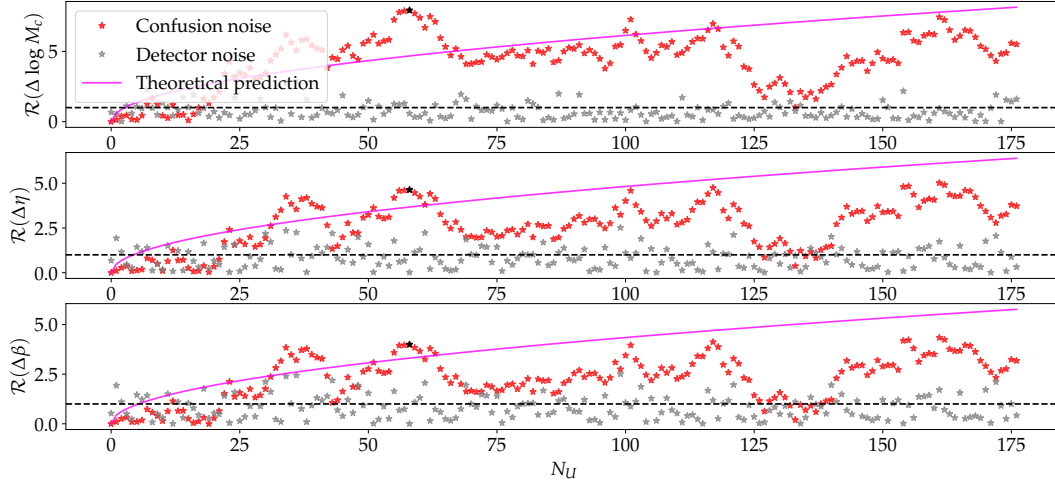
Figure 1. *Accumulation of bias from population of overlapping signals.* In red, the accumulation of bias on the parameters of the reference signal from massive black hole binaries that have not been resolved. In gray, the statistical errors arising from instrumental noise fluctuations. The noise is independently generated for each data set and so we expect the $\mathcal{R}$ values to follow a $N(0,1)$ distribution, which is consistent with what is seen in the figure. In black, the data point with the largest bias in $\mathcal{M}_c$, for which the results were verified using an MCMC simulation, giving the posterior shown in Figure 2. We note that these panels are not independent, as they represent one-dimensional marginals of a three dimensional distribution that has large correlations.

but consistent source catalogues, i.e., the data set with $N+1$ confusion sources includes the same sources as the N confusion sources data set, plus one additional source. The ratio, $\mathcal{R}(\Delta\boldsymbol{\theta}_{\mathrm{noise}})$, of the noise-induced shift in the peak of the likelihood to the expected standard deviation of this quantity, is also shown and can be seen to hover around the value of 1, as expected. Conversely, we find that the formalism predicts significant biases ($\mathcal{R} > 1$) from the accumulation of missed signals drawn from a simple, but realistic distribution of the masses. We plot the theoretical prediction from Eq. (24) of Sec. 3 on top of the found ratios, showing that they follow the expected $\sqrt{N}$ behaviour.

To assess whether these predictions are sound, we confirm them with an MCMC analysis for the data set that gives the largest bias ($\mathcal{R} \sim 8$) in chirp mass, indicated by the black data point in Fig. 1). The result of the MCMC run and the predictions for the shift in the peak of the likelihood due to the confusion sources and noise, computed with Eq. (50), are shown in Fig 2. Even in this most extreme case, we can clearly see that the predictions for the bias match the MCMC posterior very well, demonstrating that the formalism works well in estimating source confusion from missed signals. We remark that in this example the SNR of the residuals is lower than the SNR of the signal we are inferring from the data stream. This is a regime in which we would expect that the linear signal approximation is valid. In scenarios in which the SNR of the "missed" signals is larger than that of the target source, the linear signal approximation might cease to be valid, but this formalism should at least provide an indication that systematic biases are "large".
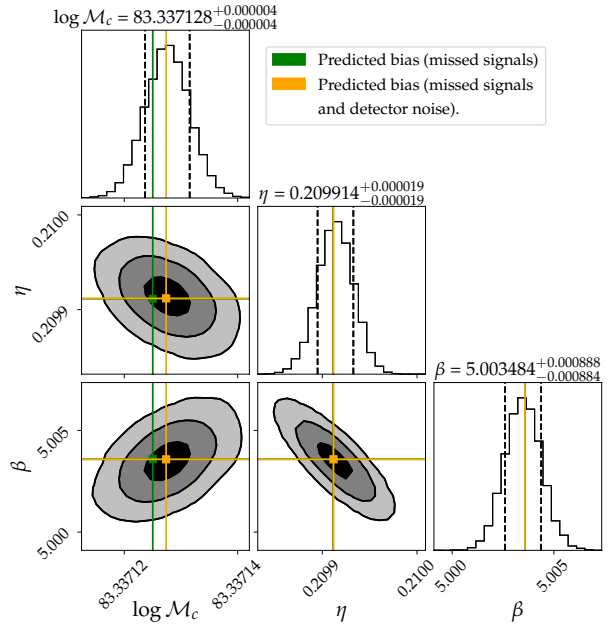
Figure 2. *Biases from source confusion and detector noise.* MCMC posteriors and predictions from the Fisher formalism for the largest-bias case in Fig. 1. The values in green are predictions considering source confusion only. Those in orange combine biases from source confusion and detector noise (which we cannot access in a realistic situation). The true values are well beyond the range of the plot at $\sim 8\sigma$ for each parameter (see Fig. 1).
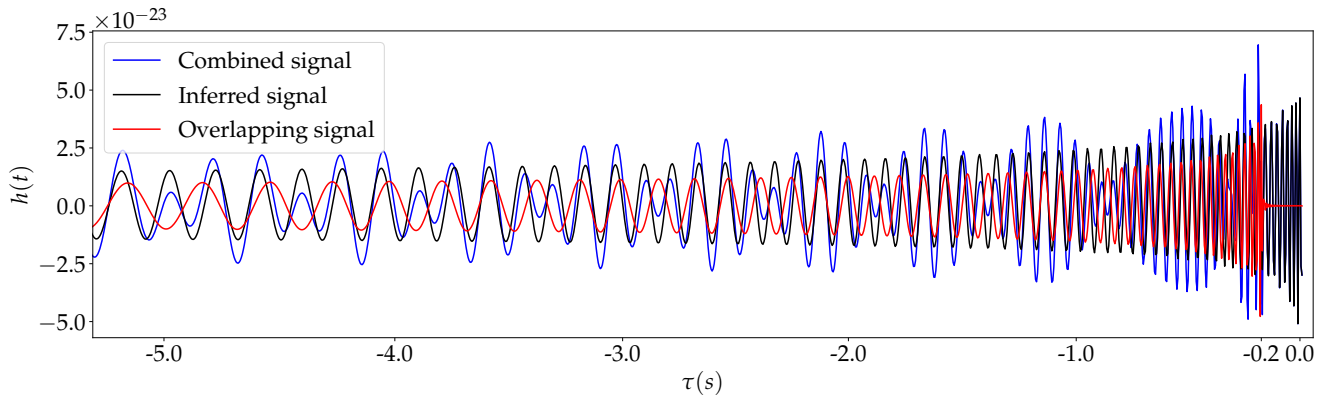
Figure 3. *Waveforms for overlapping signals.* We plot the waveforms for signal $h_e^{(1)}(t)$ in black; this represents the "inferred source" for which we are attempting to recover the parameters. We plot the waveform of the overlapping signal $h_e^{(2)}(t)$ in red; the signal has a coalescence time at $\tau = -0.2s$ relative to the one of the inferred source. The sum of the two signals is shown in blue.

5.2 Biases from overlapping signals with coincident coalescence

A particularly interesting class of overlapping signals that has attracted attention in the recent literature are those where the coalescence times t_c are nearly simultaneous. Such a scenario could be relevant to mergers of massive black holes observed by LISA or to stellar-origin binary black-holes (BBH) observed by ET and Cosmic Explorer (CE), but this will depend on the rate of such mergers and, therefore, the probability that mergers happen within the same time period. Quantitative studies of the rate of overlapping mergers have been carried out for advanced LIGO and CE. In (Samajdar et al. 2021; Pizzati et al. 2021; Himemoto, Nishizawa & Taruya 2021; Relton & Raymond 2021), the authors conclude that coincident (meaning merger times within 2 seconds) mergers of BBH binaries will occur tens of times per year for CE, and binary neutron star (BNS) mergers could occur coincidentally with other BNS or BBH mergers hundreds or even thousands of times per year.

The same papers, as well as (Smith et al. 2021), also present the first Bayesian inference analyses with overlapping signals, with some critical differences. (Smith et al. 2021) studies the simultaneous inference of overlapping neutron star binaries, in such a way that no biases on the parameters are expected from confusion noise. (Relton & Raymond 2021) performs a similar analysis for the second-generation LIGO-Voyager detector, (Samajdar et al. 2021) for pairs of BBH-BBH, BBH-BNS and BNS-BNS systems using `LAL-inference` (Veitch et al. 2015), and (Pizzati et al. 2021) for BBH pairs with `bilby` (Ashton et al. 2019). However, in these last two papers, inference is performed for one binary only, treating the second as confusion noise. They find that biases occur when the difference between the coalescence times $\tau = t_c^{(2)} - t_c^{(1)}$ of signals "(1)" and "(2)" is sufficiently small, roughly $\tau \lesssim 0.5s$. Here we analyse a similar scenario to that of (Pizzati et al. 2021), interpreting the bias as arising from a single confusion source, to see whether the analytic formalism presented here can reproduce that result without the need for expensive Bayesian posterior computation. Notice that a (joint) Fisher-matrix analysis is presented
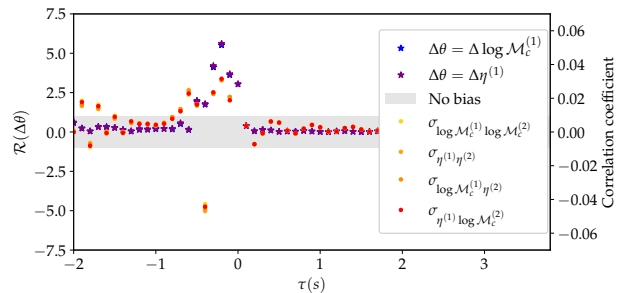
Figure 4. *Biases from an overlapping signal as a function of the difference in coalescence time.* In cold (blue, purple) colors, we plot the bias ratios for the parameters of signal "(1)" due to the unaccounted-for presence of signal "(2)", as a function of the coalescence time difference τ between the two signals. The relevant scale is the y-axis on the left, where we see that biases $\mathcal{R} > 1$ can arise. In gray, we indicate the region where we regard biases as not significant ($\mathcal{R} < 1$). In warm colors, we plot the correlation coefficients (with relevant y-axis on the right), defined in Eq. (61). We see that the largest correlations $\sigma \gtrsim 0.05$ correspond to the largest biases ($\sim 6\sigma$).

in (Himemoto, Nishizawa & Taruya 2021) for a similar scenario, though the similarities end there.

We consider an ET data stream composed of a signal $h^{(1)}$ to be inferred and a missed signal $h^{(2)}$ that creates confusion noise

$$\hat{d}(f) = \hat{h}_e^{(1)}(f; \boldsymbol{\theta}^{(1)}) + \hat{h}_e^{(2)}(f; \boldsymbol{\theta}^{(2)}). \qquad (60)$$

For this example we ignore waveform errors and detector noise. The biases arise solely due to the confusion noise $\hat{h}_e^{(2)}$, and can be predicted from (50) setting $n = \delta h = 0$. The parameter space of the Fisher matrix is $\boldsymbol{\theta}^{(1)} = \{\log \mathcal{M}_c^{(1)}, \eta^{(1)}\}$, with true parameters $\boldsymbol{\theta}_{\rm tr}^{(1)} = \{15.4 M_\odot, 0.243\}$ (corresponding to a binary with component masses $m_1 = 21 M_\odot$ and $m_2 = 15 M_\odot$). We take the signal to be nonspinning ($\beta^{(1)} = 0$) with an effective distance $D_{\rm eff}^{(1)} = 5$Gpc, and phase and times at coalescence $\phi_c^{(1)} = \pi/3$ and $t_c^{(1)} = 0s$. The SNR for this source is $\rho(h^{(1)}) \sim 75$. For the overlapping signal, we pick component masses $m_1 = 25 M_\odot$ and

$m_2 = 20 M_\odot$, a nonspinning configuration $\beta^{(2)} = 0$, an effective distance $D_{\mathrm{eff}}^{(2)} = 10\mathrm{Gpc}$, and phase at coalescence $\phi_c^{(0)} = \pi/3$. We let $t_c^{(2)}$ vary as a free parameter. For a nominal value of $t_c^{(2)} = -0.2s$, the SNR for the overlapping source is $\rho(h^{(2)}) \sim 46$. In Fig. 3, we plot time-domain waveforms for this particular configuration.

We now turn to the problem of predicting the biases on $\boldsymbol{\theta}^{(1)}$. From Eqs. (50) and (16), we compute the bias ratio $\mathcal{R}(\Delta\boldsymbol{\theta}_{\mathrm{conf}}^{(1)})$ due to the presence of confusion noise, varying $\tau := t_c^{(2)} - t_c^{(1)}$ between $\tau = -2.0$ and $\tau = 2.0$. The results are shown in Fig. 4. In this Figure, we plot both the ratios $\mathcal{R}$ and the Pearson correlation coefficients[3], defined as

$$\sigma_{\theta_1 \theta_2} = \frac{(\Gamma^{-1})_{\theta_1 \theta_2}}{\sqrt{(\Gamma^{-1})_{\theta_1 \theta_1} (\Gamma^{-1})_{\theta_2 \theta_2}}} \,. \qquad (61)$$

We notice that non-trivial biases start appearing when $|\tau| \lesssim 0.5$, which correspond to the largest correlation coefficients ($\sigma_{\theta_1 \theta_2} \sim 0.05$). We therefore (qualitatively) confirm the main result of (Pizzati et al. 2021) [and of (Samajdar et al. 2021; Himemoto, Nishizawa & Taruya 2021; Relton & Raymond 2021) indirectly]. Notice that because of our choice of data input and parameters, our comparisons with the results of (Pizzati et al. 2021) can only be qualitative. They consider noise in Advanced LIGO, while we consider ET (picking a noiseless realization in the data stream). Furthermore, they model their signals with a different approximant (IMRPhenomv2), include detector response functions, sample through masses with different true values and include additional parameters in the analysis, specifically the phase, ϕ_c, and time, t_c, at coalescence, and luminosity distance, d_L.

To check the reliability of our bias predictions, we have also compared them against posteriors from an MCMC run for a configuration with the τ leading to the largest biases ($\sim 6\sigma$, for the $\tau = -0.2s$ configuration shown in Fig. 3): we obtain excellent agreement, at the level of the accuracy shown by the (orange) prediction in Fig. 2. This example illustrates the advantage of our formalism, namely that the biases can be cheaply and reliably predicted. Our formalism will be a valuable tool for extending previous Bayesian analyses into regions of parameter space that are difficult to sample with fully Bayesian techniques.
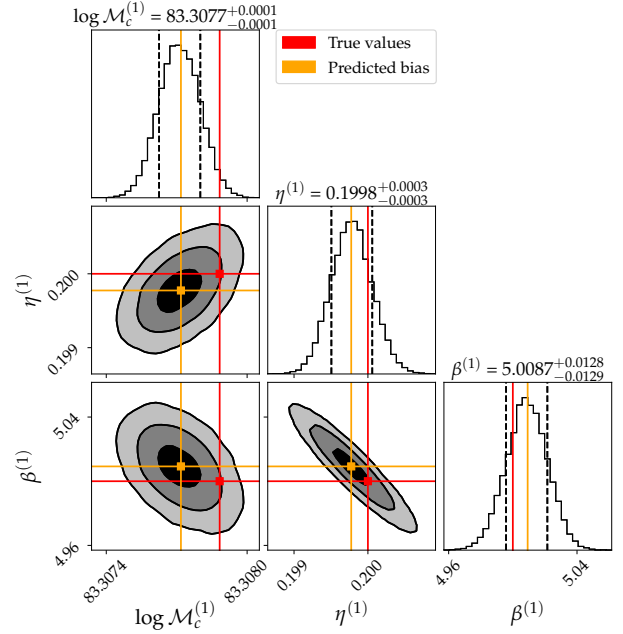
5.3   Biases from the inaccurate removal of signals

We now consider the situation in which the confusion sources are not "missed", but incorrectly fitted out. To simulate this, we consider a LISA data stream,

$$\hat{d}(f) = \hat{h}_e^{(1)}(f; \boldsymbol{\theta}^{(1)}) + \hat{h}_e^{(2)}(f; \boldsymbol{\theta}^{(2)}) + \hat{h}_e^{(3)}(f; \boldsymbol{\theta}^{(3)}) \qquad (62)$$

where the signal "(1)" is our reference signal, which we assume is modelled perfectly, and the other sources

---

[3] These correlations are calculated using the joint Fisher matrix, which fundamentally assumes that we have resolved *both* signals. In this case, we would expect no biases from the overlapping signal. In the bias ratios calculation, we treat the second signal as unfitted, which leads us to the shown biases from confusion noise. Regardless of this difference in treating the Fisher matrix, we conclude that Pearson correlations can be a guide to understand where biases would occur if the overlapping signal were not inferred, as suggested in (Pizzati et al. 2021).

Figure 5. *Biases from the inaccurate removal of loud sources.* Posterior distributions for the parameters of a reference signal, computed using MCMC, when 2 mismodelled overlapping signals are removed from the data (with parameters given in Table 1). We also show the biases predicted using our formalism.

Table 1. Parameter configurations for the signals in Sec. 5.3. We also report the SNR of the source, ρ_h and of the residual $\rho_{\delta h}$. Notice that we do not consider waveform errors for the first (reference) source here, implying its residual is zero. We sample all the sources from $f = 0.5\mathrm{mHz}$ and stop at 2mHz ($T_{\mathrm{obs}} = 0.3$ days), the earliest chirp time for these masses.

i	$M/M_\odot$	η	β	D_{eff}	t_c	ϕ_c	ρ_h	$\rho_{\delta h}$
1	$2 \cdot 10^6$	0.20	5.0	10 Gpc	6 h	0	83	-
2	$1 \cdot 10^6$	0.23	1.0	3 Gpc	48 h	π	790	31
3	$4 \cdot 10^6$	0.08	2.4	2 Gpc	6 h	0.9	2216	76

are incorrectly subtracted using approximate templates $\hat{h}_m(f; \boldsymbol{\theta}^{(2,3)}, \epsilon = 0.3)$. In such a procedure, we expect biases to arise only from the residual that the incorrectly modelled signals leave in the data stream (62),

$$\delta h = \sum_{i=2}^{3} \hat{h}_e^{(i)}(f; \boldsymbol{\theta}^{(i)}) - \hat{h}_m^{(i)}(f; \boldsymbol{\theta}^{(i)}, \epsilon = 0.3) \,. \qquad (63)$$

In this case, the relevant parameter space is $\boldsymbol{\Theta} = \{\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}\}$, where we pick each subset to be $\boldsymbol{\theta}^{(i)} = \{\log \mathcal{M}_c^{(i)}, \eta^{(i)}, \beta^{(i)}\}$. The joint Fisher matrix Γ is therefore a 9×9 matrix (calculated using $\hat{h}_m$). We report the true source parameters in Table 1. We calculate the biases $\Delta\boldsymbol{\theta}^{(1)}$ on the reference signal's parameters using (48)(or equivalently (50)), which leads us to

$$\mathcal{R}(\Delta \log \mathcal{M}_c^{(1)}) = 1.98 > 1$$
$$\mathcal{R}(\Delta\eta^{(1)}) = 0.84$$
$$\mathcal{R}(\Delta\beta^{(1)}) = 0.74 \,. \qquad (64)$$

Biases are then significant for the chirp mass in this case. These predictions can be checked with an MCMC analysis,

see Fig. 5. We find that the formalism can accurately predict the biases from the inaccurate removal of signals.

The fact that each contribution to δh in Eqs. (48,50) affects the parameters of each source equally suggests that residuals effectively behave as missed sources and confusion noise. In fact, we can rewrite the data stream analysed in Fig. 5 in the form

$$\hat{d}(f) = \hat{h}_e^{(1)}(f; \boldsymbol{\theta}^{(1)}) +$$
$$\hat{h}_m^{(2)}(f; \boldsymbol{\theta}^{(2)}, \epsilon = 0.3) + \hat{h}_m^{(3)}(f; \boldsymbol{\theta}^{(3)}, \epsilon = 0.3) + \delta h, \tag{65}$$

which explicitly separates out the modelled part using the models employed by the MCMC analysis and the calculation of the joint Fisher matrix. Doing so leaves an extra term, δh, which plays the role of the confusion noise caused by the residuals. One can check that the biases predicted from the data stream (65) (and obtained using the joint Fisher matrix with $\hat{h}_m$) match the predictions reported in Fig. 5. An important implication of this equivalence of results is that significant biases may arise from the incorrect removal of a very large number of signals drawn from the same population, in direct analogy with the findings of the previous section. We have not checked this directly, since adding a considerable number of fitted sources dramatically increases the dimensionality of $\boldsymbol{\Theta}$, making the implementation of the joint Fisher matrix difficult.

5.4 Waveform errors & confusion noise

We now bring together the ideas described in sections (5.1) and (5.3), and show that the formalism developed in Sec. (3) can accurately predict biases on parameter estimates when we simultaneously fit N_{fit} sources with inaccurate waveforms, while confusion and detector noise are also present in the data stream. We show this here for LISA, but an ET example may be found in Appendix D. The data stream in this case is

$$\hat{d}(f) = \sum_{i=1}^{N_{\text{fit}}} \hat{h}_e^{(i)}(f; \boldsymbol{\theta}_{\text{tr}}^{(i)}) + \Delta H_{\text{conf}} + \hat{n}(f). \tag{66}$$

We assume ΔH_{conf} arises from the galactic foreground of white-dwarf binaries (WDB). LISA is guaranteed to detect WDBs in their thousands or even tens of thousands (Crowder & Cornish 2007; Błaut, Babak & Królak 2010) (depending on the imposed SNR threshold), but there will also be millions of unresolved WDBs radiatig in the LISA band. Here we assume that WDBs with $\rho < 8$ have been folded into the PSD (Błaut, Babak & Królak 2010). We additionally assume that only WDBs with $\rho > 15$ have been detected by dedicated pipelines, which leaves us with missed WDBs with SNRs in the range $8 < \rho < 15$. To simulate these sources, we construct a superposition of signals, see Eq. (18), with frequencies chosen from $f_i \in (10^{-4}, 10^{-3})$Hz. For simplicity, we only retain the leading PN term in the waveform, computed for random masses drawn from $(m_1, m_2) \sim 10^2 \cdot U^2[0.3, 1] M_\odot$. We finally draw effective distances $D_{\text{eff}} \sim 10^4 \cdot U[1, 3]$pc. We discard binaries not in the specified range of SNRs, until $N_U = 1000$ are found. To complete the input data stream, we add $N_{\text{fit}} = 4$ fitted signals with waveform errors $\epsilon = 0.04$ and source parameters $\boldsymbol{\theta}^{(i)}$ given in Tab.2. We choose initial frequencies $f_0 = 10^{-4}$Hz

**Table 2.** Parameters for the simultaneously-fitted LISA signals in Sec. 5.4.

i	$M/M_\odot$	η	β	$D_{\text{eff}}/\text{Gpc}$
1	3×10^6	0.222	6	2
2	6×10^6	0.139	7	3
3	7×10^6	0.204	8	1
4	5×10^6	0.240	9	1

and sample the sources simultaneously with a maximum frequency given by the highest ISCO frequency among the fitted sources. For simplicity, we set $(\phi_c, t_c) = (0, T_{\min})$ for all sources, where $T_{\min}$ is the minimum chirping time allowed over all parameter configurations. The SNRs are of order $\mathcal{O}(10^3)$ for all fitted sources.

Corner plots displaying *all* parameter biases can be found in Fig. 6. We see that the predicted biases from (50) are in remarkable agreement with the posteriors from the MCMC algorithm. Additionally, in Fig.7 we show how the total shift in the peak of the posterior of the parameters $\boldsymbol{\theta}^{(1)}$ of the first source, computed from Eq. (48), breaks down into its constituent contributions. Firstly, we see that biases from confusion noise, unresolved sources or waveform residuals can deconstructively interfere, i.e., the combined contribution can be smaller than the worst of the individual contributions. Secondly, we notice that there are large biases from confusion noise, which implies that if global-fit analyses miss $\mathcal{O}(1000)$ WDBs, this will lead to a significant bias in parameter estimates for other GW sources. We have further explored how biases change when the threshold is taken to be any value $\rho_{\text{th}} \in [8, 15]$. We have tested that, when this threshold is increased towards $\rho_{\text{th}} = 15$, biases tend to increase as the SNR of ΔH_{conf} increases. While the model used here is approximate, it suggests that the completeness of LISA data analysis algorithms needs to be sufficiently high down to sufficiently low threshold SNRs for biases on other parameters to be minimized.

6   GLOBAL-FIT SCHEMES

So far, we have defined the global-fit as the simultaneous search for and parameter estimation of *all* gravitational wave signals in the LISA data stream. In Sec. 5.4, this was achieved by assuming the number of signals (and the associated parameter space) present in the data stream was known precisely. However, in a realistic scenario, we will not know how many signals are present in the data. Furthermore, the number of signals present at any given time may be large, leading to a prohibitively large parameter space. Consider, for example, the simultaneous inference of an extreme mass-ratio inspiral (a small compact object inspiralling into a super massive black hole) and a massive black-hole binary. Both systems will have parameter spaces $\gtrsim 14$ dimensions, requiring parameter estimation algorithms to sample from a $\gtrsim 28$ dimensional posterior. This would stretch the capabilities of current inference techniques and the problem will only worsen as more signals are included in the model. One solution is to use state-of-the-art parameter estimation techniques that are able to efficiently sample such complicated, high-dimensional posterior distributions. The overall aim of this section is to provide alternative ideas to tackle the global fit problem. We begin by proposing an (expensive)
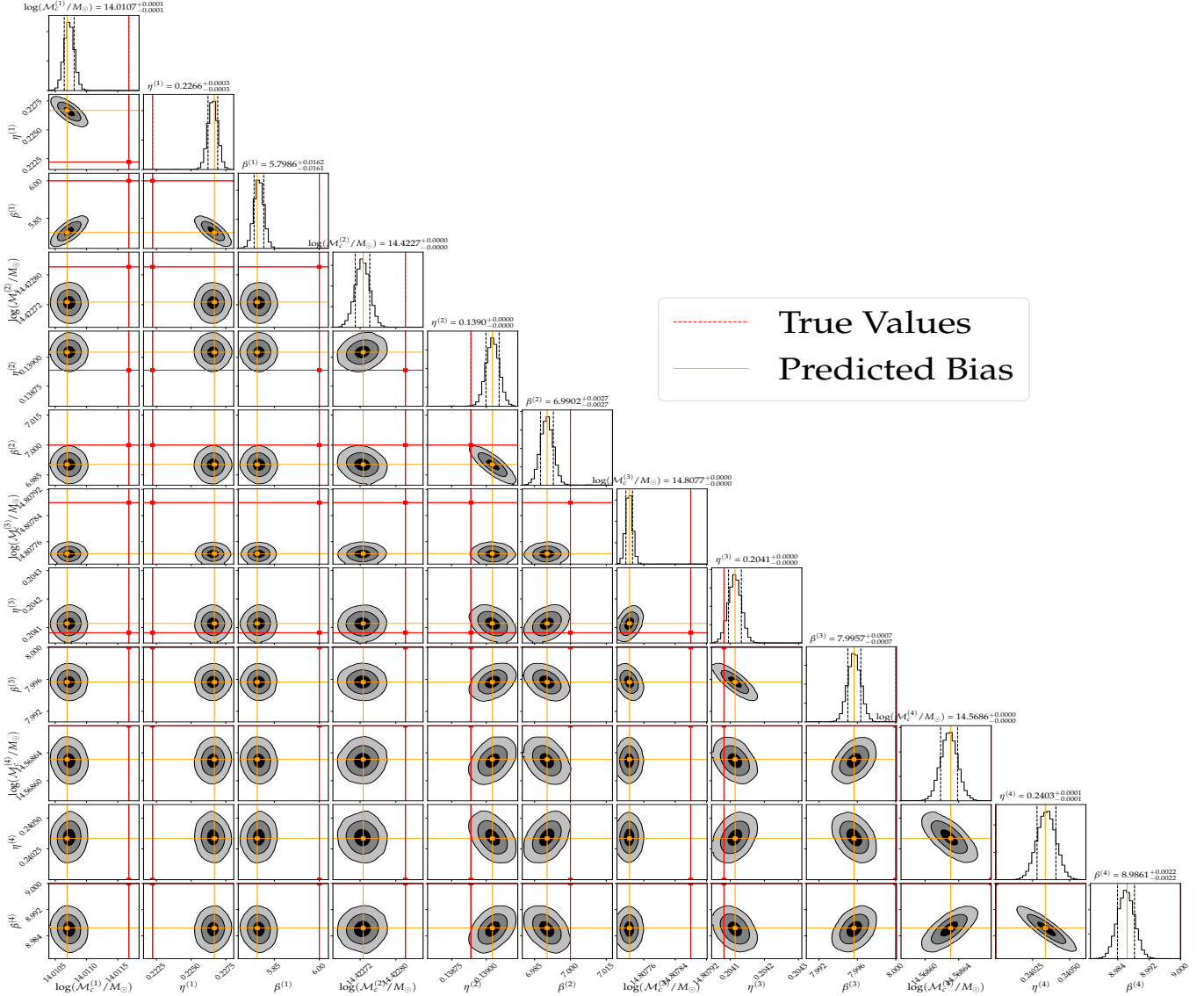
Figure 6. Triangle plot of the one-dimensional (on the diagonal) and two-dimensional marginalised posterior distributions for the inferred parameters in the LISA scenario considered in Sec. 5.4. The red lines indicate the true parameters and orange lines indicate the biases predicted from (50).

iterative approach to sample reduce portions of the parameter space. Then, using the formalism developed above, we illustrate how to reduce the cost by cheaply correcting the biases arising within the first few parameter estimation simulations.

6.1 Parameter Estimation through local-fits

Let $h^{(\mathcal{A})} \in \mathcal{A}$ and $h^{(\mathcal{B})} \in \mathcal{B}$ denote a set of distinct signals with parameters $\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{A}}$ and $\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{B}}$ we wish to infer. The joint data stream is given by

$$d(\mathcal{A}, \mathcal{B}) = \sum_{\mathcal{A}} h_e^{(\mathcal{A})}(\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{A}}) + \sum_{\mathcal{B}} h_e^{(\mathcal{B})}(\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{B}}) + n(t). \qquad (67)$$

For simplicity, we ignore effects coming from unresolved signals. Global-fit pipelines are concerned with the data stream (67) with the goal to simultaneously infer both signal sets $h^{(\mathcal{A})} \in \mathcal{A}$ and $h^{(\mathcal{B})} \in \mathcal{B}$.

In a local-fit procedure, we consider performing parameter estimation *only* on signal set $\mathcal{A}$ and treat signals from the set $\mathcal{B}$ as missed signals. We write this data stream as

$$d(\mathcal{A}|\mathcal{B}) = \sum_{\mathcal{A}} h_e^{(\mathcal{A})}(\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{A}}) + \Delta H_{\mathrm{conf}} + n(t), \qquad (68)$$

$$\Delta H_{\mathrm{conf}} = \sum_{\mathcal{B}} h_e^{(\mathcal{B})}(\boldsymbol{\theta}_{\mathrm{tr}}^{\mathcal{B}}). \qquad (69)$$

The best fit parameters for $\mathcal{A}$ obtained in this stage can be denoted $\boldsymbol{\theta}_{\mathrm{bf}}^{\mathcal{A}|\mathcal{B}}$, the conditioning on $\mathcal{B}$ indicating that the estimate was obtained with $\mathcal{B}$ present in the data. In the second step, we use the recovered parameters $\boldsymbol{\theta}_{\mathrm{bf}}^{\mathcal{A}|\mathcal{B}}$ to subtract out an estimate of $h^{\mathcal{A}}$ from the joint data stream using our approximate model

$$d(\mathcal{B}|\mathcal{A}_{\mathrm{res}}) = d(\mathcal{A}, \mathcal{B}) - \sum_{\mathcal{A}} h_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\mathrm{bf}}^{\mathcal{A}|B}). \qquad (70)$$

Then one estimates the parameters of signal $\mathcal{B}$ using the
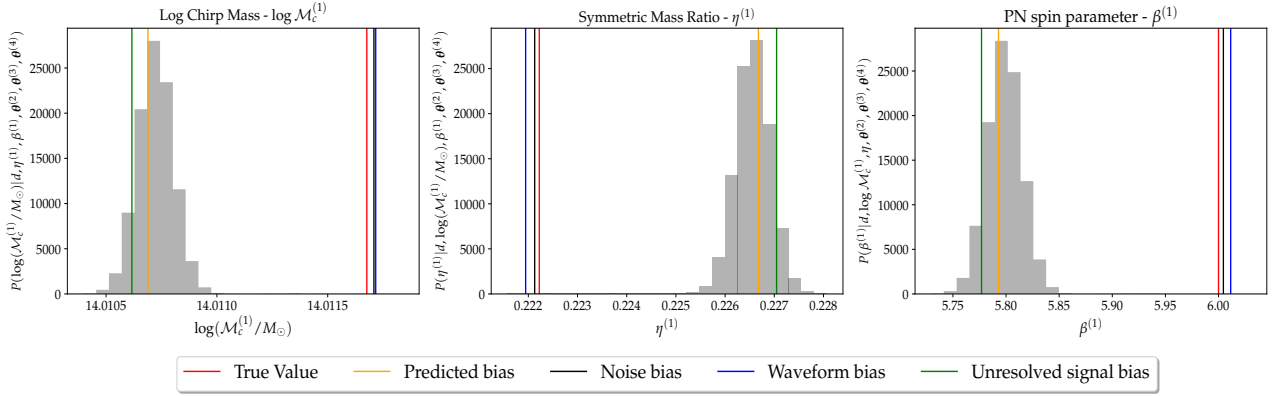
Figure 7. (top/bottom left to right) The grey histograms are the posterior samples for $\log M_c^{(1)}, \eta^{(1)}$ and $\beta^{(1)}$ for the LISA scenario considered in Sec. 5.4. The red lines indicate the true parameters, blue lines the biases arising from the use of inaccurate waveforms as templates, the black ones the noise induced shift in the peak of the likelihood, the green lines the biases due to unresolved signals and the orange lines show the approximation to the total bias computed from (50).

data stream (70) with signal templates representing signals in $\mathcal{B}$. This will yield parameters $\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$, where $\mathcal{A}_{\rm res}$ indicates that this analysis was done on a "residual data set" from which an estimate of $h^{\mathcal{A}}$ had been subtracted. This estimate can be used to *update* the initial data stream $d(\mathcal{A}|\mathcal{B})$, now denoted $d(\mathcal{A}|\mathcal{B}_{\rm res}) = d(\mathcal{A}, \mathcal{B}) - \sum_{\mathcal{B}} h^{(\mathcal{B})}(\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}})$. Again, we can perform parameter estimation on signals $\mathcal{A}$, now with residuals from $\mathcal{B}$ in the data stream, using this updated data array and recovering $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}_{\rm res}}$. These recovered parameters should be *closer* to the true parameters than $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$. We can continue this scheme by then searching over

$$d(\mathcal{B}|\mathcal{A}_{\rm res}) = d(\mathcal{A}, \mathcal{B}) - \sum_{\mathcal{A}} h_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}_{\rm res}}), \qquad (71)$$

recovering parameters, then searching over $d(\mathcal{A}|\mathcal{B}_{\rm res})$, and so on and so forth. What we would find is that the recovered parameters for both $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}}$ and $\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}}$ tend towards the "true" parameters, i.e., the parameters that would have been recovered if a global fit procedure was carried out. An advantage of this procedure is that it sidesteps issues arising from sampling the joint posterior for $\mathcal{A}$ and $\mathcal{B}$, but a clear disadvantage is that it requires a number of repeated parameter inference calculations. Computationally, this is expensive and time consuming. As an alternative, we propose that one can use the algorithm presented in Sec.(3) to correct the biases found above. In doing so, one may be able to get a reliable estimate of the true parameters $\boldsymbol{\theta}_{\rm tr}^{(\mathcal{A})}$ and $\boldsymbol{\theta}_{\rm tr}^{(\mathcal{B})}$ without having to iterate, i.e., using just the first two parameter inference calculations.

6.2 Correcting biases in the local-fit analysis

Before we talk about the details of our algorithm, it is instructive to discuss the source of the biases in parameters $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$ and $\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$. For the data stream (68), the bias in the recovered parameter $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$ is sourced by

$$\delta h^{\mathcal{A}|\mathcal{B}} = \sum_{\mathcal{B}} \hat{h}_e^{(\mathcal{B})}(\boldsymbol{\theta}_{\rm tr}) + \sum_{\mathcal{A}} \left[\hat{h}_e^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{A})}) - \hat{h}_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{A})}) \right] + \hat{n}(f), \tag{72}$$

and similarly the bias in $\boldsymbol{\theta}_{\rm tr}^{\mathcal{B}}$ when performing PE on the data stream (70)

$$\delta h^{\mathcal{B}|\mathcal{A}_{\rm res}} = \sum_{\mathcal{A}} \left[\hat{h}_e^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{A})}) - \hat{h}_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}) \right] + \\ \sum_{\mathcal{B}} \left[\hat{h}_e^{(\mathcal{B})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{B})}) - \hat{h}_m^{(\mathcal{B})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{B})}) \right] + \hat{n}(f). \tag{73}$$

In Eq.(72), the first term is the bias due to missed signals $\mathcal{B}$, the second term the residuals due to incorrect subtraction of the true signals and finally the noise. The noise related bias should be consistent with the width of the posterior. Also, the errors due to inaccurate waveforms should decrease as more accurate waveforms are developed. Thus, we believe it is reasonable to assume that the dominant contribution to the bias comes from the first term in Eq.(72). A similar story can be told for Eq.(73) where we expect the first term will dominate and the latter two will be subdominant corrections. Finally, we do not have access to the true parameters $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$ and $\boldsymbol{\theta}_{\rm tr}^{\mathcal{B}}$, nor the exact models for $h_e^{(\mathcal{A})}$ or $h_e^{(\mathcal{B})}$. We make a further approximation for the $\mathcal{B}$ true parameters $\boldsymbol{\theta}_{\rm tr}^{\mathcal{B}} \approx \boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$ and assume that $h_e \approx h_m$. We have access to these parameters from our first parameter estimation run on signal set $\mathcal{B}$ using the data stream $d(\mathcal{B}|\mathcal{A}_{\rm res})$. From this information, we can approximate both Eqs.(72) and Eq.(73) by

$$\delta h^{\mathcal{A}|\mathcal{B}} \approx \sum_{\mathcal{B}} \hat{h}_m^{(\mathcal{B})}(\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}) \tag{74}$$

$$\delta h^{\mathcal{B}|\mathcal{A}_{\rm res}} \approx \sum_{\mathcal{A}} \left[\hat{h}_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm tr}^{(\mathcal{A})}) - \hat{h}_m^{(\mathcal{A})}(\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}) \right]. \tag{75}$$

A similar complication arises from our lack of access to $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$ in Eq.(75). However, the true parameter $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$ can be estimated by calculating the CV bias using $\delta h^{\mathcal{A}|\mathcal{B}}$ from Eq.(74) with the Fisher matrix and numerical derivatives calculated at parameter values $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$. This will produce an estimate of the bias, $\Delta \boldsymbol{\theta}^{\mathcal{A}|\mathcal{B}}$, which can be subtracted from $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$, to give an updated estimate of $\boldsymbol{\theta}^{\mathcal{A}}$ that should lie closer to the true parameters, $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$. This new parameter $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}} = \boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}} - \Delta \boldsymbol{\theta}^{\mathcal{A}|\mathcal{B}}$ can be used to approximate $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$ in Eq.(75). Finally, using parameter values $\boldsymbol{\theta}_{\rm res}^{\mathcal{B}|\mathcal{A}}$ to evaluate waveform derivatives and

Table 3. This table presents the true parameter values for source 1 ($\mathcal{A}$) and source 2 ($\mathcal{B}$) for the example of the local-fit procedure presented in section 6.3. The SNR of each signal within the data stream $\rho_h^2 = (h_e|h_e)$ is given in the final column.

i	$M/M_\odot$	η	β	$D_{\rm eff}/{\rm Gpc}$	ρ_h
1 ($\mathcal{A}$)	1.2×10^7	0.222	8	2	~ 1850
2 ($\mathcal{B}$)	5×10^6	0.160	7	4	~ 379

Fisher matrices, one can compute a new estimate of the bias in the $\mathcal{B}$ set signal parameters, $\Delta\boldsymbol{\theta}^{\mathcal{B}|\mathcal{A}_{\rm res}}$ by using Eq.(75) with $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}} \approx \boldsymbol{\theta}_{\rm tr}^A$. This new bias can be used to update our best guess for the true parameters if the set of $\mathcal{B}$ signals, namely $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}} = \boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}} - \Delta\boldsymbol{\theta}^{\mathcal{B}|\mathcal{A}_{\rm res}}$. By construction, the parameter values $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}}$ and $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}}$ should lie closer to the true values $\boldsymbol{\theta}_{\rm tr}^{\mathcal{A}}$ and $\boldsymbol{\theta}_{\rm tr}^{\mathcal{B}}$ respectively.

To summarise, the algorithm is as follows

(i) Calculate $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$ and $\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$ by performing PE on signals $\mathcal{A}$ and $\mathcal{B}$ using data streams $d(\mathcal{A}|\mathcal{B})$ then $d(\mathcal{B}|\mathcal{A}_{\rm res})$.

(ii) Calculate

$$\delta h_{\rm conf}^{\mathcal{A}|\mathcal{B}} \approx \sum_{\mathcal{B}} \hat{h}_m(\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}) \tag{76}$$

and then compute an estimate of the bias on the parameters specific to $\mathcal{A}$, denoted $\Delta\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$, evaluating the waveform derivatives at the parameter values $\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$. Set new best fit parameters for $\mathcal{A}$ as $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}} = \boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}} - \Delta\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}}$.

(iii) Then calculate

$$\delta h_{\rm conf}^{\mathcal{B}|\mathcal{A}_{\rm res}} = \hat{h}_m(\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}_{\rm res}}}) - \hat{h}_m(\boldsymbol{\theta}_{\rm bf}^{\mathcal{A}|\mathcal{B}_{\rm res}}) \tag{77}$$

and calculate the CV bias $\Delta\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$ on parameters specific to $\mathcal{B}$ using parameter values $\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$. Now set new parameters $\widehat{\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}} = \boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}} - \Delta\boldsymbol{\theta}_{\rm bf}^{\mathcal{B}|\mathcal{A}_{\rm res}}$.

We illustrate the algorithm above by considering a noisy data stream containing two signals, each of which have waveform errors $\epsilon \neq 0$. We lose no generality here since the algorithm presented above is easily generalised to handle a greater number of signals. Thus we consider

$$\hat{d}(f) = \underbrace{\hat{h}_e^{(1)}(f; \boldsymbol{\theta}^{(1)}, \epsilon = 10^{-3})}_{\mathcal{A}} + \underbrace{\hat{h}_e^{(2)}(f; \boldsymbol{\theta}^{(2)}, \epsilon = 10^{-3})}_{\mathcal{B}} + \hat{n}(f). \tag{78}$$

With parameters for the $\mathcal{A}$ and $\mathcal{B}$ sources given in table 3. The results of applying the local-fit procedure are presented in the next section.

6.3 Results

Following the algorithm above, we present results for the marginalised posteriors in Fig. 8. In the top row, the blue histogram is the posterior $p(\mathcal{A}|d(\mathcal{A}|\mathcal{B}), \mathcal{B})$ obtained fitting for source $\mathcal{A}$ with source $\mathcal{B}$ in the data, the orange histogram is the posterior for the global-fit solution $p(\mathcal{A}, \mathcal{B}|d(\mathcal{A}, \mathcal{B}))$, the red lines mark the true parameters and the black line the predicted bias using the formalism. The bottom row of figure 8 show corresponding results for the inference of source $\mathcal{B}$, with, for example, the orange histograms representing posterior samples from $p(\mathcal{B}|d(\mathcal{B}|\mathcal{A}_{\rm res}), \mathcal{A}_{\rm res})$. In each case, the algorithm is able to correct the bias from the poorly

subtracted other signal in the data. In all cases, after subtracting the predicted bias, the true parameters lie within the 1σ width of the posteriors.

In fig.(8), the local fit posterior for source $\mathcal{B}$ appears to provide a more conservative estimate on how well we can constrain each parameter in comparison to the global fit analysis. Shifting the posterior by the amount predicted by the preceding algorithm will therefore yield a posterior that is broader, and hence more conservative than that which would be obtained from a full analysis. We are yet to develop a strategy to correct parameter uncertainties from the prior local fit analysis. This implies that one must retain precision measurement statements on parameters from the first two parameter estimation runs on $d(\mathcal{A}|\mathcal{B})$ and $d(\mathcal{B}|\mathcal{A}_{\rm res})$. Correcting the widths of the local fit posteriors are beyond the scope of this paper and we leave this for future work.

To conclude this section, we make a few important remarks about the algorithm given above. First of all, the algorithm is likely to be less effective if the recovered best fit parameters are far from the true value. This would cause a breakdown of the linear-signal approximation, which is a key assumption in the generalised CV algorithm presented in 3. We also assume that, through many local-fits, we have found all the signals present in the data stream we are studying. Further, the two signals present here are near orthogonal with relatively little correlation between the two signals. If there were significant overlap, then the posteriors for the global-fit procedure would be wider since extra uncertainty would be introduced into the parameters in question. This would mean that the procedure presented here, in which we shift a posterior computed with a single source model into the correct location, but do not modify the posterior width, would yield overly optimistic estimates of the source parameters. There are two approaches to address this shortcoming. Firstly, the correlation between sources identified in the data can be evaluated, and any pairs of source with sufficiently high correlation can be reanalysed jointly. Alternatively, it is possible to generate an updated posterior for the parameters of each source by marginalising over the biases due to the other source. The procedure is similar to the algorithm described here, but rather than shift each sample in the $\mathcal{A}$ source posterior by the same amount, given by the best-fit parameters of source $\mathcal{B}$, we instead shift them by an amount given by Eq. (43) evaluated for the $\mathbf{h}^{(2)}$ waveform computed as a random sample drawn from the $\mathcal{B}$ source distribution. This approach is beyond the scope of the analysis presented here, but we leave it for future work.

We are not advocating that the algorithm above should replace global-fit studies, but that is provides a cheap way to cross-check results and identify where joint fitting of sources is required. A full global-fit parameter estimation study will outperform the algorithm given above since it does not rely on as many approximations.

7   CONCLUSIONS

In this paper, we have generalized the approach in (Cutler & Vallisneri 2007) to provide metrics for the parameter estimation biases on individually resolved sources from the presence of confusion noise from missed signals or incorrectly fitted waveforms. We have illustrated these generalisations
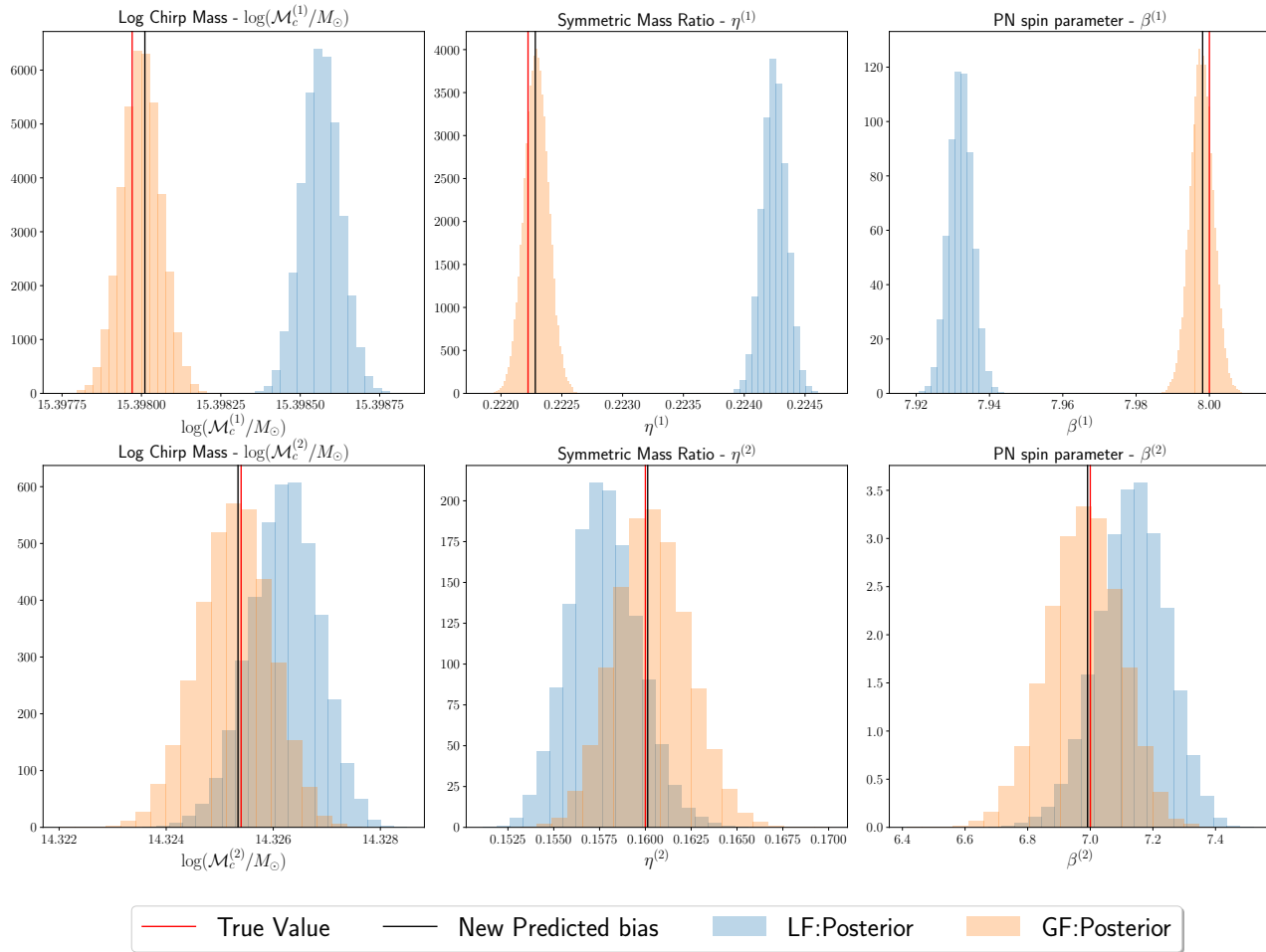
Figure 8. The orange histograms are the global-fit (GF) posteriors from searching the joint data stream $d(\mathcal{A}, \mathcal{B})$ for both $\mathcal{A}$ and $\mathcal{B}$ simultaneously. The red lines are true values and black lines the (corrected) new predicted bias using the generalised CV formalism in section 3. The blue histograms in the top row are posterior samples from the local-fit (LF) $p(\mathcal{A}|d(\mathcal{A}|\mathcal{B}), \mathcal{B})$ for missed signals $\mathcal{B}$. Similarly, the blue histograms in the bottom row are samples from $p(\mathcal{B}|d(\mathcal{B}|\mathcal{A}_{\text{res}}), \mathcal{A}_{\text{res}})$ for signal $\mathcal{B}$.

with simple (yet realistic) scenarios relevant to the LISA and ET detectors, and we can collect several generic findings:

• We find that the presence of altogether missed signals drawn from the same population could lead to significant biases on the parameter estimation of other signals which are instead fitted out of the data.

• We qualitatively confirm one of the main results of (Samajdar et al. 2021; Pizzati et al. 2021; Himemoto, Nishizawa & Taruya 2021; Relton & Raymond 2021). The coincident arrival of two signals in a ground-based detector, with nearly overlapping mergers, may lead to biases when the difference between coalescence times of the signals is less than a fraction of a second.

• We find that residuals in the data arising from the incorrect removal of sources effectively behave like missed signals, and may lead to significant biases.

• We find that biases from confusion noise and waveform inaccuracies may deconstructively interfere with one another.

• Our results suggest that galactic binaries which are missed by dedicated searches (Littenberg et al. 2020), and

not accounted for in confusion noise estimates, may lead to significant biases on the parameter estimation of other typical LISA sources.

• We proposed a proof-of-concept global-fit scheme in which, starting from local-fits of LISA sources, guesses for the true parameters are obtained through bias predictions from previous parameter estimation simulations. We find these guesses lie within the 1σ interval of global-fit posteriors across all sources. This has potential applications to confirm global-fit search algorithms, and as a standalone novel local-fit parameter estimation algorithm.

In all the cases outlined above, the formalism we have developed plays an important role in providing a theoretical ground for the described biases and a solid tool to address them. We believe this formalism could be useful in exploratory studies of future GW detectors, to assess under what circumstances we expect the biases described above to appear. We also believe this formalism is an early but significant step towards an understanding of how to simultaneously infer parameters from multiple signals of differ-

ent nature with future detectors, as we highlight with our global-fit algorithm scheme.

There are several ways in which the application of this formalism could be extended. One could perform systematics studies for realistic populations of missed signals using realistically modelled waveforms. One could check whether inaccurately modelled signals could lead to significant biases when several of them are incorrectly subtracted from the data, which our understanding of residuals as missed signals and the biases they lead to strongly suggests. This is a possibility that we have not explored due to the technical challenge in dealing with very large Fisher matrices and MCMC sampling algorithms to sample over such a large parameter space. Given more complex waveform models and response functions for either detectors, one could use this algorithm to understand the impact on PE studies of non-stationarities in the data stream. Finally, one could explore further the applications of this formalism for global-fit algorithms, which could be extended to take into account significant overlaps between the signals in the data stream, and to explore correcting the width as well as the peak location of the parameter posteriors.

As a final note, the formalism itself can be extended to take into account brighter confusion sources and more pronounced waveform errors (as would happen with different families of waveform models or models within the same family containing different physics). To do so, one could derive higher order terms in the equations present in Sec.3 to describe biases that are farther from the true parameters than those considered in this work.

REFERENCES

Abbott B. P., et al., 2018, Living Rev. Rel., 21, 3
Allen B., Anderson W. G., Brady P. R., Brown D. A., Creighton J. D., 2012, Phys. Rev. D, 85, 122006
Amaro-Seoane P., 2018, Living reviews in relativity, 21, 4
Amaro-Seoane P., et al., 2017
Ashton G., et al., 2019, Astrophys. J. Suppl., 241, 27
Berti E., Buonanno A., Will C. M., 2005, Phys. Rev. D, 71, 084025
Bonetti M., Sesana A., 2020, Physical Review D, 102
Burke O., Gair J. R., Simón J., Edwards M. C., 2020, Phys. Rev. D, 102, 124054
Błaut A., Babak S., Królak A., 2010, Physical Review D, 81
Crowder J., Cornish N. J., 2007, Physical Review D, 75
Cutler C., Vallisneri M., 2007, Phys. Rev. D, 76, 104018

Finn L. S., 1992, Phys. Rev. D, 46, 5236
Flanagan E. E., Hughes S. A., 1998, Phys. Rev. D, 57, 4566
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publ. Astron. Soc. Pac., 125, 306
Gair J. R., Sesana A., Berti E., Volonteri M., 2011, Class. Quant. Grav., 28, 094018
Gair J. R., Tang C., Volonteri M., 2010, Phys. Rev. D, 81, 104014
Gair J. R., Vallisneri M., Larson S. L., Baker J. G., 2013, Living Reviews in Relativity, 16, 7
Greene J. E., Ho L. C., 2007, Astrophys. J., 667, 131, [Erratum: Astrophys.J. 704, 1743–1747 (2009)]
Gupta A., Datta S., Kastha S., Borhanian S., Arun K., Sathyaprakash B., 2020, Phys. Rev. Lett., 125, 201101
Himemoto Y., Nishizawa A., Taruya A., 2021
Johansson F., et al., 2013, mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18). http://mpmath.org/
Karnesis N., Babak S., Pieroni M., Cornish N., Littenberg T., 2021
Khintchine A., 1934, Mathematische Annalen, 109, 604
Korol V., Belokurov V., Moore C. J., Toonen S., 2020
Littenberg T., Cornish N., Lackeos K., Robson T., 2020, Phys. Rev. D, 101, 123021
Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, The journal of chemical physics, 21, 1087
Miller M., 2005, Phys. Rev. D, 71, 104016
Pizzati E., Sachdev S., Gupta A., Sathyaprakash B., 2021
Porter E. K., 2009, arXiv preprint arXiv:0910.0373
Porter E. K., Cornish N. J., 2015, Physical Review D, 91, 104001
Punturo M., et al., 2010, Class. Quant. Grav., 27, 194002
Regimbau T. et al., 2012, Physical Review D, 86, 122001
Regimbau T., Hughes S. A., 2009, Phys. Rev. D, 79, 062002
Reitze D., et al., 2019, Bull. Am. Astron. Soc., 51, 035
Relton P., Raymond V., 2021
Roberts G. O., Gelman A., Gilks W. R., et al., 1997, The annals of applied probability, 7, 110
Robson T., Cornish N., 2017, Class. Quant. Grav., 34, 244002
Robson T., Cornish N. J., Liu C., 2019, Classical and Quantum Gravity, 36, 105011
Rodriguez C. L., Mandel I., Gair J. R., 2012, Physical Review D, 85, 062002
Roebber E., et al., 2020, Astrophys. J. Lett., 894, L15
Samajdar A., Janquart J., Van Den Broeck C., Dietrich T., 2021
Sathyaprakash B., Schutz B., 2009, Living Rev. Rel., 12, 2
Sesana A., Gair J., Berti E., Volonteri M., 2011, Phys. Rev. D, 83, 044036
Shannon C. E., 1949, Proceedings of the IRE, 37, 10
Smith R., et al., 2021
Vallisneri M., 2008, Physical Review D, 77, 042001
Veitch J., et al., 2015, Phys. Rev. D, 91, 042003
Wen L., Gair J. R., 2005, Classical and Quantum Gravity, 22, S445
Whittle P., 1957, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 19, 38
Wiener N., et al., 1930, Acta mathematica, 55, 117
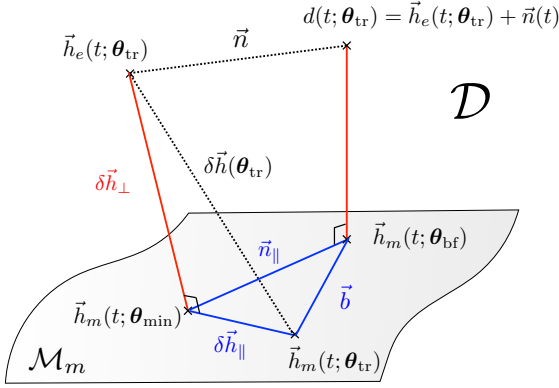
Figure A1. Geometrical setup for the CV biases. Represented is the space of signals $\mathcal{D}$ and the various realisations of model and exact templates, with definitions for the parameters as given in the main text. In red, the perpendicular contributions from waveform mismodelling and noise realisations that affect the detectability of the signal. In blue, the contributions to the shifts to the parameters, corresponding to noise-induced errors ($\vec{n}_{\parallel}$) and theoretical biases ($\delta\vec{h}_{\parallel}$).

APPENDIX A: GEOMETRICAL INTERPRETATION OF PARAMETER ERRORS

In this section, we provide a geometrical interpretation for the noise and systematic biases derived in (Cutler & Vallisneri 2007). Consider the vector space $\mathcal{D}$ of outputs $\vec{d}(t; \boldsymbol{\theta})$ depending on parameters $\theta^i \in \boldsymbol{\theta}$. Further define two submanifolds $\mathcal{M}_m$ and $\mathcal{M}_e$ of model $\vec{h}_m(t; \boldsymbol{\theta})$ and fiducial $\vec{h}_e(t; \boldsymbol{\theta})$ templates, representing both the limiting case of no instrumental noise. Next, consider the waveform difference $\delta\vec{h}(\boldsymbol{\theta}_{\mathrm{tr}}) := \vec{h}_e(t; \boldsymbol{\theta}_{\mathrm{tr}}) - \vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{tr}})$ evaluated at the true parameters. This can be split into a perpendicular and parallel component. The former is obtained drawing a perpendicular vector $\delta\vec{h}_{\perp}$ from $\vec{h}_e(t; \boldsymbol{\theta}_{\mathrm{tr}}) \in \mathcal{D}$ onto $\mathcal{M}_m$. The projection point is $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{min}})$, evaluated at the parameters $\boldsymbol{\theta}_{\mathrm{min}}$ that minimise the distance $(\vec{h}_e - \vec{h}_m | \vec{h}_e - \vec{h}_m)$. Starting from $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{min}})$, one can perform a coordinate transformation that maps the model waveform evaluated at $\boldsymbol{\theta}_{\mathrm{min}}$ to the same model evaluated at the true parameters $\boldsymbol{\theta}_{\mathrm{tr}}$. This defines the component $\delta\vec{h}_{\parallel}$, see Fig. (A1). Physically, the $\delta\vec{h}_{\perp}$ component corresponds to a "loss" of SNR that changes the distance (and therefore affects the likelihood and detectability of the signal only), whereas $\delta h_{\parallel} \approx (\theta_{\mathrm{tr}}^i - \theta_{\mathrm{min}}^i)\partial_i h_m$ corresponds to shifts in the parameters. In what follows, we restrict our attention to vectors in $\mathcal{M}_m$ signalling errors and biases in the parameters, leaving out perpendicular components related to the detectability of the source.

In a realistic situation, we are confronted with a detector output $\vec{d}(t; \boldsymbol{\theta}_{\mathrm{tr}}) = \vec{h}_e(t; \boldsymbol{\theta}_{\mathrm{tr}}) + \vec{n}(t)$ that includes noise. We can project $\vec{d}$ onto $\mathcal{M}_m$, which defines the model template $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{bf}})$ evaluated at the best-fit parameters. These are the ones one obtains minimising the argument of the Whittle likelihood, $(\vec{d} - \vec{h}_m | \vec{d} - \vec{h}_m)$. The new element of $\mathcal{M}_m$, $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{bf}})$ is connected to $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{min}})$ through the parallel component of the noise $\vec{n}_{\parallel}$, which can be rewritten as

$n_{\parallel} \approx (\theta_{\mathrm{bf}}^i - \theta_{\mathrm{min}}^i)\partial_i h_m$, and to $\vec{h}_m(t; \boldsymbol{\theta}_{\mathrm{tr}})$ through a (bias) vector $b \approx (\theta_{\mathrm{bf}}^i - \theta_{\mathrm{tr}}^i)\partial_i h_m$, see Fig. (A1). Then, in this realistic situation the total bias on the PE performed with the model template $\vec{h}_m$ is given by $\Delta\theta^i := \theta_{\mathrm{bf}}^i - \theta_{\mathrm{tr}}^i$, which itself is formed by two contributions $\Delta\theta_{\mathrm{noise}}^i := \theta_{\mathrm{bf}}^i - \theta_{\mathrm{min}}^i$ and $\Delta\theta_{\mathrm{sys}}^i := \theta_{\mathrm{tr}}^i - \theta_{\mathrm{min}}^i$. The former is a statistical error from the noise vector (which averages to zero after many draws of $\vec{n}$), and we identify it with Eq. (14). The latter is a contribution from waveform mismodelling ($\delta\vec{h}$) that does not average to zero after many repetitions of the experiment, and we identify it with the CV bias from theoretical errors (15).

APPENDIX B: NUMERICAL ROUTINES

In this appendix, we provide more details on how we sample our signals in the frequency domain. We begin by choosing a starting frequency f_0 and final frequency determined by the last stable orbit in a Schwarzschild spacetime $f_{\mathrm{max}} = c^3/6\sqrt{6}\pi GM$. The calculated time to merger is then predicted through the 3.5PN chirp time (see Eq.(3.5a) of (Allen et al. 2012)). Invoking Shannon's sampling theorem (Shannon 1949), the spacing between time points Δt is chosen to be $\Delta t = 1/(2f_{\mathrm{max}})$. For multiple signals, we choose the minimum sampling interval common to all waveforms for given mass parameters. In doing so, we find the length of the signal $N_t = \lfloor t_{obs}/\Delta t \rfloor$ in the time domain. Combining all these elements, one is able to construct a list of sampling frequencies $f = [0, \Delta f, 2\Delta f, \ldots, \lfloor (N_t - 1)/2 \rfloor \Delta f]$ for $\Delta f = 1/N_t\Delta t$. Given the discrete Fourier frequencies, it is then possible to construct waveforms using (51).

Noise is generated in the frequency domain with real and imaginary parts drawn separately from Gaussian distributions with equal variance and zero mean. Discretising equation (2), it's easy to show that the variance of both real and imaginary parts are equivalent to

$$\sigma^2(f_i) = N_t S_n(f_i)/4\Delta t. \tag{B1}$$

Finally, in order to calculate various quantities involving inner products (Fisher matrices, SNRs and likelihoods), we use the discrete analogue of (4),

$$(a|b) \approx 4\Delta f \,\mathrm{Re} \sum_{i=0}^{\lfloor \frac{N_t - 1}{2} \rfloor} \frac{\hat{a}(f_i)\hat{b}^{\star}(f_i)}{S_n(f_i)}. \tag{B2}$$

APPENDIX C: FISHER MATRICES AND THEIR VALIDATION

The Fisher Matrix (9) can be calculated through inner products of waveform derivatives. We choose to use a second order finite difference method,

$$\frac{\partial h_m(f; \Theta^i)}{\partial \Theta^i} \approx \frac{h_m(f; \Theta^i + \delta\Theta^i) - h_m(f; \Theta^i - \delta\Theta^i)}{2\delta\Theta^i} \tag{C1}$$

Fisher matrices in gravitational wave astronomy have high condition numbers, which influence our ability to obtain reliable parameter precision estimates. We invert our Fisher matrices using the high precision arithmetic Python package mpmath (Johansson et al. 2013). This was done in order to mitigate instabilities arising from computing the inverse

of the potentially badly conditioned matrix (Wen & Gair 2005; Vallisneri 2008; Porter 2009; Rodriguez, Mandel & Gair 2012; Gair et al. 2013; Porter & Cornish 2015; Amaro-Seoane 2018; Burke et al. 2020). A criterion to establish the stability of the inverse Fisher matrix based on the (1-norm) absolute value reads $\left|\Gamma^{-1}\Gamma - \mathtt{I}\right|_1 \leqslant 10^{-3}$, where $\mathtt{I}$ is the identity matrix (Gupta et al. 2020). We used ~ 500 decimal digits and found $\Gamma^{-1}\Gamma = \mathtt{I} - \epsilon_{ij}\mathtt{I}$ with $\max_{i,j}\{|\epsilon_{ij}|\} \approx 10^{-14}$, even with condition numbers $\sim 10^{21}$. This gives us confidence that the numerical inversion of our Fisher matrix is both numerically robust and accurate.

To validate our results, we carry out a Markov-Chain Monte-Carlo (MCMC) with the goal to match our Fisher matrix results in a high-SNR regime. Our Bayesian analyses are carried out using `emcee` (Foreman-Mackey et al. 2013) and an appropriate modification of the code developed in (Burke et al. 2020). The posteriors are sampled with `emcee` using a Whittle log-likelihood (5) and flat priors. A publicly available implementation of the MCMC illustrations carried out with `emcee` can be found at `https://github.com/aantonelli94/GWOP`. The latter code is based on a standard Metropolis-Hastings algorithm (Metropolis et al. 1953). A publicly available implementation can be found at `https://github.com/OllieBurke/Noisy_Neighbours`. For this algorithm, we chose a proposal distribution equivalent to a multivariate Gaussian with covariance matrix equal to a scaled variate inverse of the Fisher Matrix. By pre-multiplying the inverse Fisher matrix Γ^{-1} by N_{sources}, we found better acceptance ratios $\sim 30\%$ [near the optimal acceptance rate for non-single parameter studies (Roberts et al. 1997)].

APPENDIX D: PREDICTING WAVEFORM AND CONFUSION NOISE BIASES WITH ET

In this appendix, we repeat the analysis of Sec. 5.4 for a source in ET. We use the same data stream as (66), modelling $N_{\text{fit}} = 2$ simultaneously-fitted signals in a similar manner. We pick waveform errors $\epsilon = 0.02$ and a starting frequency $f_0 = 5$Hz. As for confusion noise, we construct it with a series of missed signals which we model without errors. We report the parameters for both fitted and missed sources in Tab. (D1). The SNRs of the fitted signals are $\mathcal{O}(10^3)$, those of the missed signals $\lesssim 1000$ (with the lowest ~ 200). The SNRs of the missed signals for ET are noticeably high, and would likely be detected in a future analysis. However, for sake of example, treat these signals as missed signals in the parameter estimation scheme. The predictions for the biases of all parameters, Fig. D, show that the formalism can predict the mean of the posterior as remarkably well as in the case of LISA. The individual bias contributions, Fig. D1, confirm that biases can deconstructively interfere.

Table D1. Parameter configurations for the ET case.

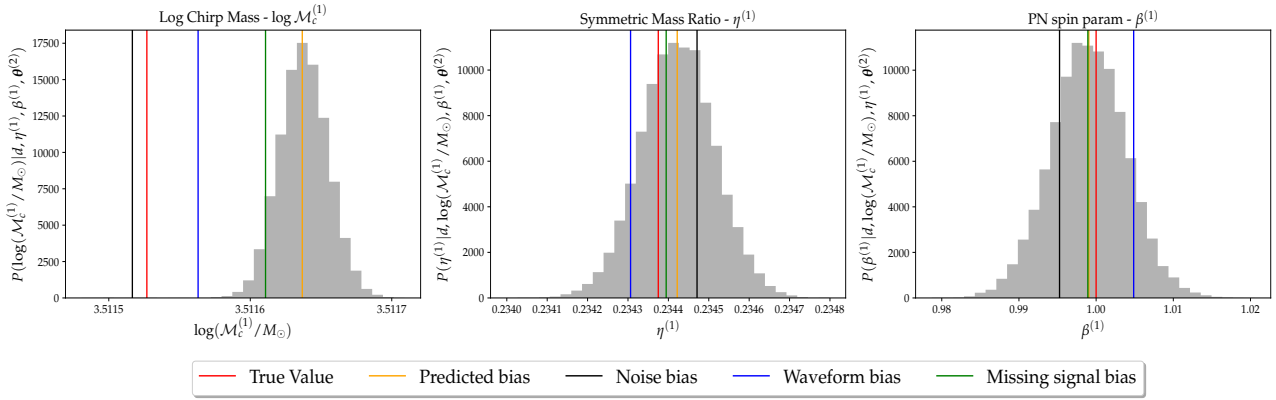|	Fitted			
i	$M/M_\odot$	η	β	$D_{\text{eff}}/\text{Mpc}$
1	80	0.234	1	400
2	70	0.204	5	40
	Missed			
i	$M/M_\odot$	η	β	$D_{\text{eff}}/\text{Mpc}$
1	2.22	2.708	5.04	259.93
2	2.886	0.247	3.882	253.36
3	4.395	0.2264	5.539	324.227
4	6.452	0.1991	4.404	305.828

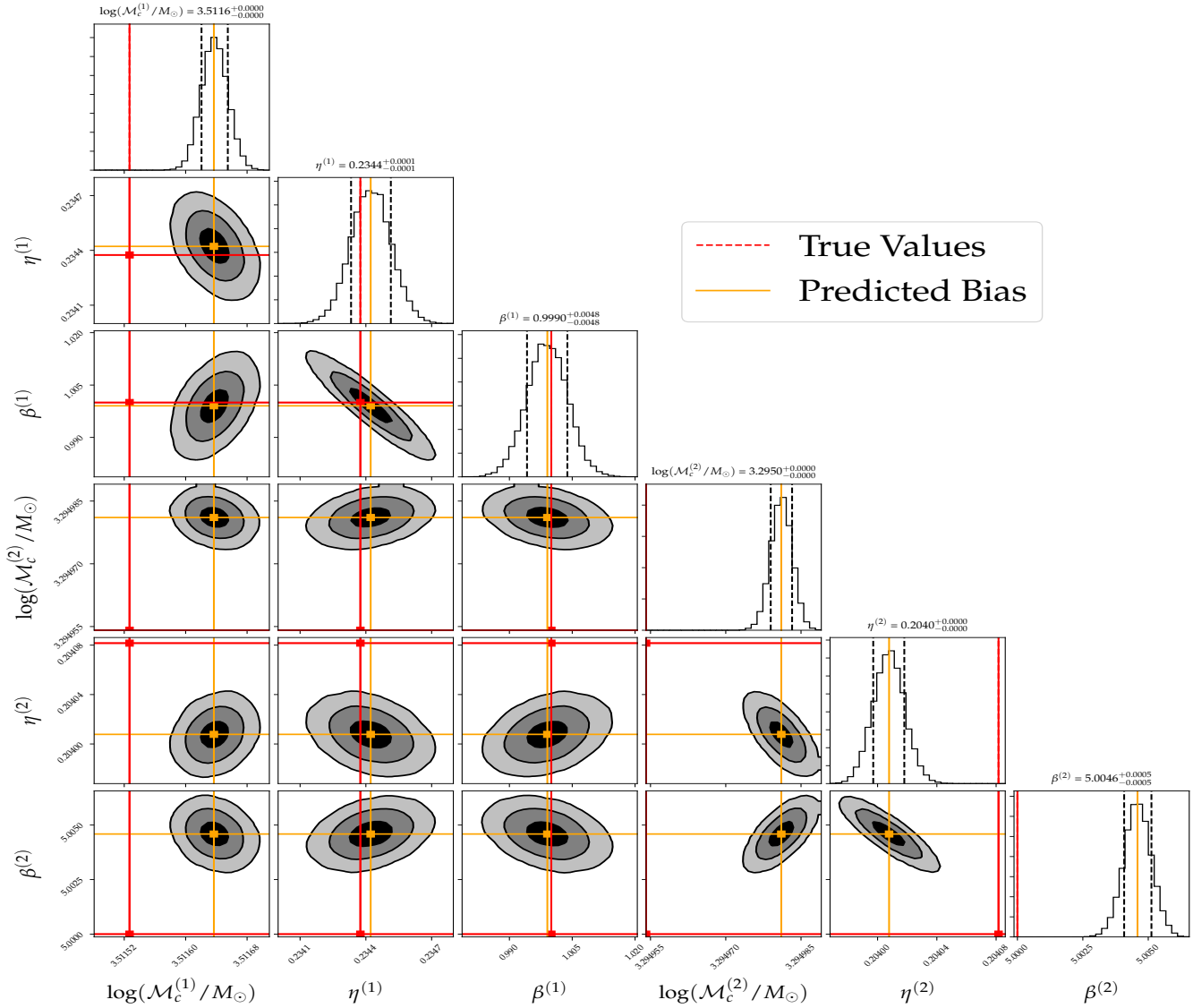**Figure D1.** Same as Figure 7 but for the ET configuration of Appendix D.

Figure D2. Same as Fig. 6 but for the ET configuration of Appendix D.