Compact Binary Astrophysics

Joseph Cameron Mills

Submitted for the degree of Doctor of Philosophy
School of Physics and Astronomy
Cardiff University

3/8/2021
Summary of thesis

This thesis brings together several projects in compact binary astrophysics that intersect at two main loci. The first being calculations of the science that is possible with future gravitational-wave detectors. On this front, chapter 2 provides localization estimates for a variety of future gravitational-wave network configurations, to inform the science case and observing strategies for future gravitational and electromagnetic detectors. In chapter 3 we predict the constraints that future observations will make on the inclination of the orbital plane of compact binaries. Finally, chapter 7 explores the observability and parameter estimation of seed black holes with future gravitational-wave networks. The second focus of this thesis is determining what is measurable in a gravitational-wave signal emitted in a binary coalescence. In chapter 3, we describe in detail degeneracies in the amplitude of the dominant emission of the two gravitational-wave polarizations, which limits our ability to measure their luminosity distance. Chapter 4 describes how higher harmonics in the waveform can break these degeneracies, and introduces a simple criterion for their observation. Using this criterion, and a related one for precession, in chapters 5 and 6 we explore the evidence for higher harmonics and precessions in the signals detected in the first half of LIGO-Virgo’s third observing run.
## Contents

1 Introduction

2 Localization of binary neutron star mergers with second and third generation gravitational-wave detectors
   2.1 Introduction
   2.2 Future Detectors and Networks
      2.2.1 Future detectors
      2.2.2 Networks
      2.2.3 Network Sensitivity
   2.3 Source Localization
   2.4 Results
      2.4.1 Face on Binary Neutron Star Mergers
      2.4.2 A Population of Coalescing Binary Neutron Stars
      2.4.3 Implication for EM Followup
   2.5 Discussion

Appendices
   2.A Optimal Detector Orientations

3 Constraining the inclinations of binary mergers from gravitational-wave observations
   3.1 Introduction
   3.2 Measuring Distance and Inclination
      3.2.1 Origin of the degeneracy
      3.2.2 Network sensitivity
      3.2.3 Parameter estimation
   3.3 Accuracy of measuring distance and inclination
   3.4 Conclusion and Future Work

4 Measuring gravitational-wave higher order multipoles
   4.1 Introduction
   4.2 The Gravitational Waveform
   4.3 The Significance of Higher Harmonics
      4.3.1 Dependence upon intrinsic parameters
      4.3.2 Dependence upon extrinsic parameters
   4.4 Observing Higher Harmonics
      4.4.1 Measured SNR in higher harmonics
      4.4.2 Expectation due to noise
      4.4.3 Observation of higher harmonics
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>Higher Harmonics in a Population of Binary Mergers</td>
<td>70</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion</td>
<td>71</td>
</tr>
<tr>
<td>Appendixes</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>4.A</td>
<td>Spin-weighted spherical harmonic polarizations</td>
<td>74</td>
</tr>
<tr>
<td>4.B</td>
<td>Derivation of ( p(R_{lm}) )</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>GW190412 &amp; GW190814: The first binaries with measurable higher order multipoles</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>79</td>
</tr>
<tr>
<td>5.2</td>
<td>GW190412</td>
<td>80</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Detection and parameter estimation</td>
<td>80</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Evidence for higher multipoles</td>
<td>82</td>
</tr>
<tr>
<td>5.3</td>
<td>GW190814</td>
<td>86</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Detection and parameter estimation</td>
<td>86</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Evidence for higher multipoles</td>
<td>88</td>
</tr>
<tr>
<td>Appendixes</td>
<td></td>
<td>92</td>
</tr>
<tr>
<td>5.A</td>
<td>Waveform models used</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>Precession and higher order multipoles in the first half of the third gravitational wave observing run</td>
<td>93</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>93</td>
</tr>
<tr>
<td>6.2</td>
<td>Method</td>
<td>95</td>
</tr>
<tr>
<td>6.3</td>
<td>Results</td>
<td>99</td>
</tr>
<tr>
<td>6.3.1</td>
<td>GW190412 and GW190814</td>
<td>102</td>
</tr>
<tr>
<td>6.3.2</td>
<td>GW190519 and GW190929</td>
<td>102</td>
</tr>
<tr>
<td>6.3.3</td>
<td>GW190521</td>
<td>104</td>
</tr>
<tr>
<td>6.4</td>
<td>Discussion</td>
<td>105</td>
</tr>
<tr>
<td>Appendixes</td>
<td></td>
<td>107</td>
</tr>
<tr>
<td>6.A</td>
<td>Parameter estimation samples used for ( \rho_{33} )</td>
<td>107</td>
</tr>
<tr>
<td>7</td>
<td>Constraining seed black hole populations with third generation gravitational-wave detectors</td>
<td>109</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>109</td>
</tr>
<tr>
<td>7.2</td>
<td>Building a supermassive black hole: stars, seeds and binaries</td>
<td>110</td>
</tr>
<tr>
<td>7.3</td>
<td>Observing seed black holes</td>
<td>111</td>
</tr>
<tr>
<td>7.4</td>
<td>Parameter estimation</td>
<td>116</td>
</tr>
<tr>
<td>7.4.1</td>
<td>Impact of higher order multipoles on inclination distance degeneracy</td>
<td>118</td>
</tr>
<tr>
<td>7.4.2</td>
<td>Results from full parameter estimation</td>
<td>120</td>
</tr>
<tr>
<td>7.5</td>
<td>Discussion</td>
<td>122</td>
</tr>
<tr>
<td>8</td>
<td>Discussion</td>
<td>123</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Target noise curves for existing and future detectors [1]: advanced LIGO at design sensitivity (aLIGO); LIGO Voyager; Einstein Telescope (two proposed configurations, ET-B and ET-D) and Cosmic Explorer (CE). .......................................................... 10

2.2 Relative sensitivity of the different networks over the sky: Voyager, Voyager-ET, 3ET/3CE and CE-ET. Left: network response as a function of sky position and right: alignment factor as a function of sky position. Also shown are the locations of detectors in each network. Magenta markers are for Voyager detectors, white for ET, and red for CE. Since both $N_R$ and the alignment factor are invariant under an overall scaling in network sensitivity, 3ET and 3CE will have identical patterns. .......................................................... 14

2.3 The localization ellipses at different sky locations for face-on 1.4-1.4 binary neutron star (BNS) binaries at a redshift of: left - $z = 0.2$ (luminosity distance of 1 Gpc) and right - $z = 0.5$ (luminosity distance of 3 Gpc). The red crosses indicate that the BNS at this sky position would not be detected — either due to a network signal-to-noise ratio (SNR) less than 12, or not having SNR $> 5$ in at least two detectors. The blue + symbols indicated sources that would be detected, but not well localized due to being identified in only two detectors. The ellipses give the 90% localization regions for a source from a given sky location. Detector locations are shown with golden markers. .......... 18

2.4 The detection and localization efficiency as a function of redshift and luminosity distance of the Voyager, Voyager-ET and 3ET networks (left column) and the 3ET, CE-ET and 3CE networks (right column). For visual comparison, 3ET is plotted as a solid line in both. Left column: Voyager-ET and Voyager are the dashed and dotted lines respectively. Right column: CE-ET and 3CE are the dashed and dotted lines respectively. From bottom to top, the rows show the fraction of events at a given redshift/distance that will be detected and localized within 100, 10 and 1 deg$^2$. ................................. 20
2.5 The number of BNS observations and localizations per year with future networks, as a function of redshift and luminosity distance. The y-axis is scaled so that the area under the curves gives the number of events per year. This assumes an isotropic redshift distribution that follows the star formation rate \([2]\) shifted to account for the delay between star birth and BNS merger \([3]\) and a distribution of possible delay times \(P(t_D) \propto 1/t_D\), with a minimum delay of \(t_D^{\text{min}} = 0.2\) Myr \([4]\). The local merger rate density is taken as \(\Psi_{BNS}(0) = 1500\ Gpc^{-3} y^{-1}\) \([5]\). Note that the y-axis on 1 deg\(^2\) plot is different from the others. For visual comparison 3ET is plotted as a solid line in both column. **Left column**: Voyager-ET and Voyager are the dashed and dotted lines respectively. **Right column**: CE-ET and 3CE are the dashed and dotted lines respectively. From bottom to top, the rows show the number density of events at a given redshift/distance that will be detected and localized within 100, 10 and 1 deg\(^2\).

3.1.1 The relative contributions of the plus and cross polarizations of a gravitational-wave signal, dependent on the inclination. The red solid line indicates the amplitude of the plus polarization, while the dashed red solid line indicates the amplitude for the plus polarization with a negative phase. The blue solid line indicates the amplitude of the cross polarization. The shaded regions show the percent differences between the plus and cross polarizations. The red portion represents when the plus and cross polarization are less than 1% different. The blue region represents where the polarizations are between 1% and 5% different. The grey region represents where the polarizations are between 5% and 10% different.

3.2.1 The relative sensitivity of detector networks to the second polarization, as encoded in the parameter \(\alpha\), defined through \(F_x = \alpha F_+\) (in the dominant polarization frame where the network is maximally sensitive to the plus polarization). The left plot shows the expected distribution of \(\alpha\) for second-generation gravitational-wave networks, while the right plot shows the distribution for potential third generation networks. In both cases, the distribution is the expected normalized distribution for a population of events, distributed uniformly in volume, and observed above threshold in the detector network. Thus, directions of good network sensitivity are more highly weighted. In the legends, \(\bar{\alpha}\) denotes the mean \(\alpha\) for each network. The second generation networks considered are LIGO Hanford and Livingston (HL); two LIGO detectors and Virgo (HLV); LIGO-Virgo and KAGRA (HLVK) and LIGO-Virgo-KAGRA with LIGO-India (HLVKI). As more detectors are added to the network, the average sensitivity to the second polarization increases. The right plot shows results for the Einstein Telescope (ET), which is comprised of three 60-degree interferometers, ET and three LIGO-Voyager detectors (Voyager-ET) and ET with either one or two Cosmic Explorer detectors (1CE-ET and 2CE-ET). As the ET detector has good sensitivity to both polarizations, networks where ET is the most sensitive detector will have large values of \(\alpha\). Third generation target noise curves are taken from \([1]\).
3.2.2 The marginalized posterior distribution for the distance and inclination of the binary neutron star system GW170817, detected with a signal-to-noise ratio $\rho \sim 32$ and network alignment factor $\alpha \sim 0.13$. The left plot was generated using only the data from gravitational-wave detectors, while the right plot also uses the independent distance measurement (40.7 Mpc, $\pm 2.4$ Mpc at 90% confidence) from electromagnetic observations. The coloured portion of the plot shows the probability distribution obtained using our approximate analysis, normalized such that the peak probability is 1. The orange contours represent the 90% and the 50% confidence intervals obtained by performing the full analysis of the LIGO-Virgo data (posterior samples are publicly available here: https://dcc.ligo.org/LIGO-P1800061/public) [6].

3.2.3 The progression of the probability distributions over a $\cos \iota$ and distance parameter space for a signal detected with alignment factor $\alpha = 0.1$ and signal to noise ratio $\rho = 12$. The white star represents the injected signal. The top panel shows the distribution for a face-on signal. The bottom panel shows the distribution for an edge-on signal. The leftmost plots are the distribution for only the likelihood. This is generated by calculating the SNR fall-off over the parameter space. Since we have not yet marginalized over the phase $\phi_0$ and polarization $\psi$, the orientation angles are set equal to the values used in the injection, in this case zero for both $\phi_0$ and $\psi$. The middle plots show how these distribution change when marginalizing over $\psi$ and $\phi_0$. Lastly, the rightmost plots are the complete probability distribution, calculated by applying a distance-squared weighting to the likelihood. This is to account for the expectation that binary systems are distributed uniformly in volume. Recall that $\alpha = 0.1$ is the mode sensitivity for the Hanford-Livingston network. The white star represents the hypothetical signal. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. From these plots, we can see that at this $\alpha$, a side-on signal is indistinguishable from a face-on/faceway signal.

3.2.4 The probability distribution over a $\cos \iota$ and distance parameter space for a signal detected with alignment factor $\alpha = 0.5$ and signal to noise ratio $\rho = 12$. The white star represents the injected signal. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. A face-on signal (where $\cos \iota = 1$) returns a nearly identical probability distribution of the parameter space as a signal from a binary with an inclination of about 66 degrees ($\cos \iota = 0.4$). For inclinations in the range $0.1 < \cos \iota < 0.4$, though the distribution now peaks at the correct inclination, there is support extending across from face-on to an inclination of $\iota \sim 80^\circ - 90^\circ$. In these cases it is not possible to distinguish the binary inclination. The signal is only clearly identified as not face-on after $\cos \iota < 0.1$. 

-- viii --
3.3.1 The un-normalized marginalized posterior for \( \cos \iota \) for a face-on source as measured for three networks with alignment factors \( \alpha = 0.1, \alpha = 0.5, \alpha = 1.0 \) and signal to noise ratio \( \rho = 12, \rho = 12, \rho = 50 \) respectively. The solid line shows the expected \( \cos^3 \iota \) form of the likelihood.

3.3.2 This plot shows a detector network’s ability to constrain the inclination of a face-on signal with 90% confidence. The x-axis shows the network alignment factor \( \alpha \), whereas the y-axis shows the signal-to-noise ratio (SNR) of the hypothetical gravitational-wave signal. The colour represents the upper limit on the inclination angle. For weak signals or for networks which are not very sensitive to the cross polarization, the network can only constrain the inclination to being less than about 45°. Even for the most sensitive detector network detecting the loudest hypothetical signals, the network would be unable to constrain the inclination to being less than 30°. However, we note that at these SNRs, the detector network may be able to identify higher order modes, which would break the degeneracy between distance and inclination, allowing more precise measurement of the inclination.

3.3.3 Marginalized posterior distribution for a 10\( M_\odot - 10\ M_\odot \) binary black hole at redshift \( z = 10 \) detected by the Einstein Telescope in the most sensitive part of the sky, i.e. directly above the detector. Here, the alignment factor is \( \alpha = 1 \) and the signal-to-noise ratio is \( \rho = 20 \). The white star represents the injected signal at three different inclinations: \( \iota = 66° \), \( \iota = 60° \) and \( \iota = 0° \). The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. We use a prior that is a uniform in comoving volume with a rest frame rate density that follows the star formation rate [2]. At this redshift the prior varies by a factor of \( \sim 12 \) across the degeneracy and now favours more inclined binaries. Thus binaries that are face-on will be recovered as being more inclined. The redshift uncertainty \( \Delta z/z \sim 40\% \) dominates the statistical error in the recovery of the binary chirp mass. All conversions between luminosity distance and redshift assume standard cosmological parameters [7].

4.2.1 The absolute amplitude factors \( A^{lm} \) (see Eq. 4.4) of the (2, 2), (2, 1), (3, 3), (3, 2) and (4, 4) harmonics as a function of the inclination \( \iota \) of the binary. The (2, 2) multipole moment is normalized to unity at \( \iota = 0° \) while other multipoles are normalized to unit amplitude in the + polarization at \( \iota = 90° \). Left: + polarization Right: \( \times \) polarization.
4.3.1 Ratio of the intrinsic amplitude, $\alpha_{lm}$, (defined in equation (4.13)) of signal harmonics to the (2,2) harmonic as a function of the total (detector frame) mass and mass ratio of the system, in a detector with sensitivity matching the Advanced LIGO detectors during O3 [8]. Upper left: the (3,3) harmonic; Upper right: the (4,4) harmonic; lower left: the (2,1) harmonic; lower right: the (3,2) harmonic. In all cases, the spins of the black holes are set to zero. The solid white line corresponds to $\alpha_{lm} = 5.3/20$ and the dashed line to $\alpha_{lm} = 2.1/20$, which correspond, approximately, to the threshold for the higher harmonics being confidently/marginally observable for a signal with SNR=20 in the (2,2) multipole. Note that the colorbar is normalized differently between the top and bottom row to improve the visibility of the weaker harmonics.

4.3.2 Absolute value of the noise-weighted inner product between multipoles, evaluated using the Advanced LIGO (O3) sensitivity, as a function of mass ratio and total mass for non-spinning black holes ($\chi_{\text{eff}} = 0$). Left: the overlap between the (2,2) and (2,1) multipoles; Right: the overlap between (2,2) and (3,2) multipoles.

4.3.3 Distribution of $R_{33}(\iota)$ and $R_{44}(\iota)$ for all binaries (Universe) as well as that subset that would be detected above a fixed SNR threshold for the (2,2) harmonic (Detected). We show both the results from a Monte-Carlo simulation as well as the analytical prediction.

4.4.1 Posterior probability distribution for $\hat{\rho}_{33}^\perp$, the orthogonal optimal signal-to-noise ratio of the (3,3) multipole. The simulated waveform corresponds to a system with $m_1 = 40M_\odot$, $m_2 = 10M_\odot$ and $\cos \iota = 0.7$. The two solid histograms show the posterior distribution when the (3,3) multipole is included in the waveform, either with (green) or without (orange) Gaussian noise. The dotted histogram shows the posterior on the SNR in the (3,3) harmonic inferred from the measurement of the (2,2) harmonic alone. The vertical line indicates the simulated value of $\hat{\rho}_{33}^\perp$ and the dashed lines indicate the expected distribution, based on a non-central $\chi$ distribution with two degrees of freedom and assuming a flat prior (as derived in Section 4.4.2), in the presence and absence of a signal in the (3,3) harmonic. Also shown as a solid brown line is the predicted posterior distribution for a signal with $\hat{\rho}_{33}^\perp = 5.5$, but instead using the inferred distribution from the measurement of the (2,2) harmonic (the dotted blue histogram) as a prior. We see general agreement between the predicted and measured posteriors, with the measured values slightly smaller than predicted (see Section 4.4.3 for discussion).

4.4.2 2D Posterior probability distribution for left: inclination and distance, right: polarization and phase at coalescence for a signal model containing top: only the dominant (2,2) multipole and bottom: the (2,2) and (3,3) multipoles. The simulated waveform corresponds to a system with $m_1 = 40M_\odot$, $m_2 = 10M_\odot$ and $\cos \iota = 0.7$. The solid (dashed) white contours denote 90% (50%) credible regions. These are not shown for polarization–phase for the (2,2) waveform, due to the clear degeneracy.
4.4.3 Posterior probability distribution for \( \hat{\rho}_{33} \), the orthogonal optimal signal-to-noise ratio of the \((3, 3)\) multipole, when the simulated signal contains the \((3, 3)\) multipole for a variety of models and noise realizations. The injected parameters are \( m_1 = 40 M_\odot \), \( m_2 = 10 M_\odot \) at \( \cos \iota = 0.7 \).

4.5.1 The mass ratio, total mass and orientation distribution for a simulated population of black hole binaries, and for that subset of systems for which the \((3, 3)\) or \((4, 4)\) harmonics are visible. The population is modelled as a power law in masses, with isotropic distribution of orientations, further details of the population are given in the text. The distribution is shown as a function of \textit{Left}: mass ratio; \textit{Middle}: total mass; \textit{right}: orientation. The subset of sources for which the \((3, 3)\) or \((4, 4)\) multipole moment is observable are shown by the dashed/dotted lines respectively, and those with observable \((4, 4)\) multipole where the \((3, 3)\) multipole is not observable are shown by the dot-dashed line. We show the observed population divided by 10 as a grey solid line, and on the mass ratio plot, without reweighting as a black line. The total number of observed binaries in the simulation is 78,000.

5.2.1 GW190412: The posterior distribution for mass ratio, \( m_2/m_1 > 1 \). The inclusion of higher multipoles constrains the binary away from equal mass. All models shown here are part of the Phenom family.

5.2.2 GW190412: The posterior distribution for the luminosity distance, \( D_L \), and inclination, \( \theta_{JN} \) (angle between the line-of-sight and total angular momentum), of GW190412. We illustrate the 90\% credible regions. By comparing models that include either the dominant multipole (and no precession), higher multipoles and no precession, or higher multipoles and precession, we can see the great impact higher multipoles have on constraining the inclination and distance. All models shown here are part of the EOBNR family. This plot is reproduced from Ref. [9].

5.2.3 GW190412: Posterior distributions for the orthogonal optimal SNR in the \((\ell, m) = (3, 3)\) multipole, \( \rho_{33} \) (see section 5.2.2 and Ref. [10]), the inclination angle between the line-of-sight and orbital angular momentum, \( \iota \), mass ratio \( q = m_2/m_1 < 1 \) and total mass in the detector frame under different waveform model assumptions. Results for the precessing higher multipole waveform IMRPhenomPv3HM model are shown as red contours, while grey contours denote the precessing waveform IMRPhenomPv2 which contains only the \((2,2)\) multipole in the co-precessing frame. Both more unequal masses and more planar viewing angles result in a larger amplitude \((\ell, m) = (3, 3)\) multipole. For this reason the binary inclination angle is degenerate with the mass ratio. Shading shows the 1\(\sigma \), 3\(\sigma \) and 5\(\sigma \) credible regions. This plot was made with PESummary [11].
5.2.4 GW190412: The probability distribution of the precessing SNR, $\rho_p$ (blue) and the orthogonal optimal SNR, $\rho$, contained in the strongest higher multipole, $(\ell, m) = (3, 3)$ (orange). We also show the expected distribution from Gaussian noise (dotted line) and the 3-$\sigma$ level (dashed line). The results indicate that there is marginal support for precession, but the posterior supports a clearly measurable higher multipole. This plot is reproduced here from Ref. [9].

5.3.1 GW190814: The marginalized posterior distribution for the secondary mass obtained using a suite of waveform models. The vertical lines indicate the 90% credible bounds for each waveform model. The labels Phenom/EOBNR PHM (generic spin directions + higher multipoles), Phenom/EOBNR HM (aligned-spin + higher multipoles) and Phenom/EOBNR (aligned-spin, quadrupole only) indicate the different physical content in each of the waveform models. This plot is reproduced here from Ref. [12].

5.3.2 GW190814: Posterior distributions for the orthogonal optimal SNR in the $(\ell, m) = (3, 3)$ multipole, $\rho_{33}$ (see section 5.2.2 and Ref. [10]), the inclination angle between the line-of-sight and orbital angular momentum, $\iota$, mass ratio $q = m_2/m_1 < 1$ and total mass in the detector frame under different waveform model assumptions. Results for the precessing higher multipole waveform IMRPhenomPv3HM model are shown as green contours. Red contours denote results for the non-precessing higher mode model IMRPhenomHM. Grey contours denote the precessing waveform IMRPhenomPv2 which contains only the (2,2) multipole in the co-precessing frame. Here we see that including precession has the largest effect on reducing the uncertainty in mass ratio. Shading shows the 1$\sigma$, 3$\sigma$ and 5$\sigma$ credible regions. This plot was made with PESummary [11].

5.3.3 GW190814: Two-dimensional posterior probability for the tilt-angle and spin-magnitude for the primary object (left) and secondary object (right) based on the Combined samples. The tilt angles are 0° for spins aligned and 180° for spins anti-aligned with the orbital angular momentum. The tiles are constructed linearly in spin magnitude and the cosine of the tilt angles such that each tile contains identical prior probability. The color indicates the posterior probability per pixel. The probabilities are marginalized over the azimuthal angles. This plot is reproduced here from Ref. [12].

5.3.4 GW190814: The posterior distribution for the luminosity distance, $D_L$, and inclination, $\theta_JN$ (angle between the line-of-sight and total angular momentum), of GW190814. We illustrate the 90% credible regions. By comparing models that include either the dominant multipole (and no precession), higher multipoles and no precession, or higher multipoles and precession, we can see the great impact higher multipoles have on constraining the inclination and distance. This plot is reproduced here from Ref. [12].

5.3.5 GW190814: Posterior distributions for the precessing SNR, $\rho_p$ (green) and the optimal SNR in the (3,3) sub-dominant multipole moment, $\rho$ (orange). The grey dotted line shows the expected distribution for Gaussian noise. This plot is reproduced here from Ref. [12].
6.3.1 Plot showing the Left: $\rho_{33}$ and Right: $\rho_p$ distributions for all observations in the second gravitational-wave catalogue (grey). In red we show the $\rho_{33}$ and $\rho_p$ distribution averaged across events. In black we show the average of the median expected noise distribution for Left: higher multipoles and Right: precession. Events which are discussed in the text are colored.

6.3.2 $\rho_p$ distributions for First row: GW190412, Second row: GW190915, Third row: GW190814. The blue line shows the expected distribution of $\rho_p$ in a stretch of noisy data under the assumption that the source is non-precessing, $\rho_{NP}^p$. The blue shaded region shows the 1$\sigma$ uncertainty of $\rho_{NP}^p$. The black line shows the average $\rho_{NP}^p$ across all events.

6.3.3 $\rho_{33}$ for Left: GW190519 Right: GW190929. The blue solid lines are the publicly available posteriors obtained with a waveform model that includes higher order multipoles. Solid orange lines indicate posteriors obtained with models that include only the dominant harmonic. Treating this as a prior, we multiply by the likelihood modelled as a non-central chi distribution with two degrees of freedom and non-centrality parameter $\rho_{MF}^{33} = 2.1$ to obtain an estimate of a posterior that could be obtained when the power in the (3,3) multipole is noise alone. These estimates are shown as the grey dashed lines.

6.3.4 Corner plot showing the inferred mass ratio and $\rho_{33}$ for the reanalysis of GW190521 by Nitz et al. [13] compared to the results in GWTC-2 [14, 15, 16]. Shading shows the 1$\sigma$, 3$\sigma$ and 5$\sigma$ confidence intervals. This plot was made with PESummary [11].

7.3.1 Distribution of BBH coalescence events in the redshift $z$--$m_{BH,T}$ diagram. Data points describe cosmologically-driven BH mergers with mass ratio $q \geq 0.1$, triggered only by triple interactions among galaxy halos. Grey triangles, blue squares and red circles denotes the total mass and redshift of the coalescences extracted from the simulation of a $10^{13}M_\odot$ over-density, forming a $\sim 10^9M_\odot$ SMBH at $z_{QSO} = 6.4$, 2 and 0.2 (represented with stars in the plot). Symbols with white edges indicate mergers involving at least one heavy seed. Color-coded areas represent lines of constant S/N ratios for ET (yellow/red) and LISA (azure/blue) computed for non spinning binaries assuming a mass ratio $q = 0.5$, which corresponds to the mean value of the merging binaries extracted from our samples. The ensemble of the color-coded areas for a given detector is often referred to ”waterfall” plot and provides averaged values of the S/N ratio at which a GW source is detected. Reproduced from Ref. [17].
7.3.2 The GW and EM landscape. Color-coded areas give the average GW horizon computed for a detection threshold equal to $S/N = 10$: contour lines refer to binaries with mass ratios $q = 1, 0.5, 0.3, 0.1$ both in the ET and LISA bandwidth. Upper limits (shown as thick horizontal bars) indicate the sensitivity of the deepest pointing, in the $[0.5 − 2]$ keV observed band, by *Athena* (orange) and *Lynx* (black) given the limiting fluxes of $2.4 \times 10^{-17}$ and $10^{-19}$ erg s$^{-1}$ cm$^{-2}$, respectively. The upper limits are inferred assuming that BHs are emitting at the Eddington limit and adopting a bolometric correction ($L_X/L_{bol}$) of 10%. Ellipses highlight the islands in the $z − m_{BH,T}$ plane where light (blue) and heavy (white) seeds are expected to form as well as where light seeds are expected to grow via accretion and mergers (yellow). The transit to the SMBH domain covers the entire LISA area and EM observations are key to discover the high-mass tail of the SMBH distribution. The light-grey ellipse below $z ∼ 5$ marks the population of long-living ”starved” seeds. Note that in this island, coordinated multi-band observations are possible having LISA the capability to first follow the early inspiral in intermediate-mass black holes and ET the merger phase, enhancing the ability to carry on precise measurements of the source parameters also at $z ∼ 5$ [18]. The islands have overlap with the GW horizon, but an empty inaccessible region is present between ET and LISA, corresponding to the Deci-Hz GW domain. The island corresponding to the stellar realm is included, on the left, for comparison. Reproduced from Ref. [17].

7.3.3 100-100 $M_\odot$ binary at redshift 10 (i.e. with detector frame total mass of 2200 $M_\odot$) Left: time domain at different frequencies. Computed with the model SEOBNRv2 [19]. Right: Frequency domain for (2,2) and (4,4) multipole inclined at 40 degrees, computed with IMRPhenomHM [20]. Normalized such that area between the waveform and noise is indicative of SNR [21]. A version of the time domain plot appeared in the science case document for third generation detectors [22].

7.3.4 Ratio of the intrinsic amplitude, $\alpha_{lm}$, of signal harmonics to the (2,2) harmonic as a function of the total (detector frame) mass and mass ratio of the system, in the Einstein telescope. *Top:* the (3,3) harmonic; *Bottom:* the (4,4) harmonic. In all cases, the spins of the black holes are set to zero. The solid white line corresponds to $\alpha_{lm} = 5.3/20$ and the dashed line to $\alpha_{lm} = 2.1/20$, which correspond, approximately, to the threshold for the higher harmonics being confidently/marginally observable for a signal with SNR=20 in the (2,2) multipole. Dotted lines are at $\alpha_{lm} = 1$. 

---

xvii
7.4.1 Progression of the cos \( \iota \)-distance probability distribution for a 100\( M_\odot \) – 100\( M_\odot \) binary black hole at redshift \( z = 10 \) detected overhead the Einstein Telescope. The white star represents the injected signal. The top row shows the inference with a waveform model that includes only the dominant (2,2) multipole. The bottom panel shows the distribution inferred when the (4,4) multipole is also included. The leftmost column show the distribution for the un-marginalized likelihood, generated by calculating the SNR fall-off over the parameter space. Since we have not yet marginalized over the phase \( \phi_0 \) and polarization \( \psi \), these angles are taken to be equal to the values used in the injection, in this case zero for both \( \phi_0 \) and \( \psi \). The middle column show how these distributions change when marginalizing over \( \psi \) and \( \phi_0 \). Lastly, the rightmost column show the posterior distribution, calculated by multiplying the marginalized likelihood by the prior. We use a uniform in \( \cos \iota \) and comoving volume prior. The white contours represent the 50\% and 90\% confidence intervals obtained from our simplified model. From these plots, we can see that including the (4,4) multipole in the waveform model breaks the degeneracy between distance and inclination, resulting in a posterior peaked at the correct value and excluding a face-off orientation.

7.4.2 Inferred posteriors for the (clockwise from top left) redshift \( z \), mass ratio \( q < 1 \), total mass in source frame \( M_{source}^{total} \) and inclination \( \cos \theta_{JN} \) of a non-spinning 100 – 100\( M_\odot \) binary black hole at redshift 10. The signal is detected with SNR \( \sim 67 \) by a network including ET located at the site of Virgo, and two Cosmic Explorers located at LIGO India and LIGO Livingston. We compare posteriors inferred with waveform models containing the dominant multipole (IMRPhenomPv2) with those inferred with additional higher multipoles (IMRPhenomHM). Dashed vertical lines indicate the 90\% credible bounds for each waveform model.

7.4.3 Posterior distribution for the component masses of the fiducial seed black hole binary described in Fig 7.4.2’s caption. We illustrate the 90\% credible regions. The red cross indicates the injected values. Comparing IMRPhenomPv2, a model which only includes the dominant multipole with one including higher multipoles, we see the great impact higher harmonics have for constraining the source frame masses.
List of Tables

2.1 Performance metrics for Voyager, Voyager-ET, 3ET, CE-ET and 3CE networks for a population of BNS coalescences distributed uniformly in comoving volume with an intrinsic merger rate of $1500 \text{Gpc}^{-3}\text{y}^{-1}$. *From top to bottom:* The number of sources per year that are detected and localized by each network; the number of sources localized per year within 1, 10 and 100 deg$^2$ respectively; the median localization area of all detected sources and the fraction of localized sources whose position is restricted to a single patch in the sky. Note, the median source is not localized by the two heterogeneous networks CE-ET and Voyager-ET.  .......................................................... 24

2.A.1The locations and orientations of the detectors used in this study. All numbers are given in degrees and the orientation angle is defined clockwise relative to a hypothetical L-shaped interferometer with arms due North and East. ................................. 29

3.3.1 The table shows the ability of various networks to distinguish the orientation of a population of binary mergers with given inclination, $\iota$. For each network and range of $\iota$, we give the percentage of binaries for which the posterior on the inclination peaks at $\iota = 0$ or $180^\circ$ (face-on) and this peak contains over 90% of the probability; those binaries for which the recovered inclination peaks at the correct value, and greater than 90% of the probability is consistent with this peak (inclined); and those for which the posterior includes significant contributions for both face-on and inclined orientations (uncertain). For all networks, essentially all binaries with $\iota < 45^\circ$ will be recovered face-on. As the inclination increases further, the ability to clearly identify the binary as inclined increases significantly with the number of detectors in the network as this improves the average sensitivity to the second gravitational-wave polarization. ................................. 46

5.2.1 Inferred parameter values for GW190412 and their 90% credible intervals, obtained using precessing models including higher multipoles. Symbols are described in the text. ................................. 81
5.3.1 Source properties of GW190814: We report the median values along with the symmetric 90% credible intervals for the SEOBNRv4PHM (EOBNR PHM) and IMRPhenomPv3HM (PHENOM PHM) waveform models. The primary spin magnitude and the effective precession is given as the 90% upper limit. The inclination angle is folded to $[0, \pi/2]$. The last column is the result of combining the posteriors of each model with equal weight. The sky location of GW190814 is shown in Figure 2 of Ref. [12].

5.A.1 Waveform models used for GW190814 and GW190412. We indicate which multipoles are included for each model. For precessing models, the multipoles correspond to those in the co-precessing frame.

6.3.1 Table showing the SNR in the $(\ell, m) = (3, 3)$ multipole moment $\rho_{33}$ and the SNR from precession $\rho_p$ for all events in the second gravitational wave catalogue [23] plus the reanalysis of GW190521 by Nitz et al. [13]. For each event we also show two Jensen Shannon Divergences (JSDs); $D_{33}^{33}$ and $D_p^{33}$ compare the calculated $\rho_{33}$ and $\rho_p$ distributions to the expected noise distributions respectively. Events with a larger JSD show greater evidence for higher order multipoles and/or precession. For events where the JSD and $\rho_{33}/\rho_p$ could not be calculated, due to the lack of publically available posterior samples, we add a hyphen. Where applicable we report the median values along with the 90% symmetric credible intervals.

6.A.1 Waveform model assumed for parameter estimation results used to calculate $\rho_{33}$ for each event in O3a. For brevity we omit events which used SEOBNRv4PHM, which was the most commonly used waveform model. All parameter estimation results are publicly available [24].
Aark aark duh laark
Frum kaairdiff aarms park
Iyll aav a claarksee pi
Aan an aarf aan aarf uh daark

Topher Mills
Acknowledgments

Slowly it has dawned on me that the acknowledgments section, like the other chapters, will always, necessarily, be incomplete. Sorry to everyone I have missed, and to those who deserve a better message.

First thanks go to Steve. You are an inspirational scientist and person. Wicked smart and a powerful force for good. Thank you for sharing your time, wisdom, and bag of tricks with me. It has been an enormous privilege.

It has been such a joy to work alongside so many talented people in the Cardiff gravy wavers group. Aldo, Alex, Ali, Alistair, Andrew, Bernard, Cecilio, Charlie, Chimlay, Chris, Dave, Duncan, Ed, Eleanor, Fabio, Francesco, Frank, Gernot, Hartmut, Hong, Iain, Jakob, Jonathan, Kate, Laura, Lionel, Lorenzo, Mark, Michael, Patrick, Paul, Penny, Phil, Rhys, Ronaldas, Sam H, Sam U, Sander, Sathyia, Seb, Soichiro, Tessa, Vaibhav, Valeriu, Vassilis, Virginia, Vivien, Will, and the rest... I have memories with so many of you that I will treasure. I hope we create many more.

Everyone upstairs in 53, across the road in North building, and at SUSSP73 – thanks for the good times. Members of the LVK and wider astrophysics community – thanks for the many lessons learned. Thanks to my MSc Physics cohort, and of course to Richard and Paul. You guys are phenomenal.

To my mum, this is for you. Your love gave me the self confidence and reckless abandon to pursue whatever I thought was interesting. To my dad, thanks for trying to turn me into a rock star – please don’t give up. To Steve, thanks for doing everything in your power to support my progression as a scientist. Without your firm encouragement, I’m not sure I’d have ever done a masters. Big love to Steph and Con for putting up with me, and pulling me through this last part. Congrats to Chris for publishing his book first. Matt, Kharys, Meg, Gazza, Dale, Firas, Tim, and many others, thanks for joining me on this adventure. Looking forward to the next one.

Huge thanks to European society for funding this research (European Research Council Consolidator Grant 647839ST/L000962/1). Thank you to members of the public who showed an interest in this research, your fascination and curiosity inspires me. I am grateful for computational resources provided by the Cardiff University and LIGO Laboratory and supported by STFC grant ST/I006285/1 and National Science Foundation Grants PHY-0757058 and PHY-0823459.

In addition to contributions by many of those already named above, I would also like to thank Duncan Brown, Monica Colpi, Rosa Valiante, Rana Adhikari, Lisa Barsotti, Yanbei Chen, Matt Evans, Salvatore Vitale and Tito Dal Canton for discussions and suggestions that helped to improve this thesis.
Chapter 1

Introduction

The observation of gravitational wave signals from coalescing black holes and neutron stars has established the field of gravitational wave astronomy [21, 5, 25, 26, 27, 28, 29, 30, 31, 16, 32]. As well as representing a completely new domain for enquiry, this new science has provided a fresh perspective on old problems in a variety of fields. For instance, to highlight just a few examples, gravitational wave observations have led to new measurements of the nuclear equation of state [33, 34], the discovery of a major site for nucleosynthesis of elements heavier than iron [35, 36, 37, 38], strong-field tests of the theory of general relativity [39], measurements of the Hubble constant [40] and fresh questions about the origin and nature of compact objects [31, 12, 14, 15, 13]. These scientific outputs, and many others, depend on our ability to infer the nature of the compact system from the gravitational wave data: its location on the sky, the inclination of its orbital plane with respect to earth, the mass, spin, and tidal deformation of its components. There is a complex interplay between the way the properties of the colliding objects are encoded in the gravitational wave signal, how this signal interacts with gravitational-wave antennas here on earth to produce a data stream, and the techniques used to reconstruct the source properties from this data stream. As a result, some properties are easier to measure than others.

This thesis will review some of the possibilities and limitations of these measurements, with a view to contributing to the interpretation of events detected in the third observing run of the LIGO-Virgo network. In addition, we make predictions about the prospects for future observations. In doing so, we inform the science case for future gravitational wave networks, including Voyager [41], the Einstein Telescope [42, 43, 44], Cosmic Explorer [1], and LISA [45].

We begin in chapter 2 with an assessment of the localization ability of future ground-based gravitational wave networks. It is expected that future networks of gravitational wave detectors will possess great potential in probing various aspects of astronomy. An important consideration for successive improvement of current detectors or establishment on new sites is knowledge of the minimum number of
detectors required to perform precision astronomy. We attempt to answer this question by assessing the ability of future detector networks to detect and localize binary neutron stars mergers on the sky. Good localization ability is crucial for many of the scientific goals of gravitational wave astronomy, such as electromagnetic follow-up, measuring the properties of compact binaries throughout cosmic history, and cosmology. We find that although two detectors at improved sensitivity are sufficient to get a substantial increase in the number of observed signals, at least three detectors of comparable sensitivity are required to localize majority of the signals, typically to within around 10 deg$^2$ — adequate for follow-up with most wide field of view optical telescopes.

Next, in chapter 3, we examine the capacity of ground-based gravitational wave networks to accurately constrain the inclination of the orbital plane of compact binaries relative to the line-of-sight. We discuss in detail a degeneracy between the measurement of the binary distance and inclination which limits our ability to accurately measure the inclination using gravitational waves alone. This degeneracy is exacerbated by the expected distribution of events in the universe, which leads us to prefer face-on systems at a greater distance. We use a simplified model that only considers the binary distance and orientation and show that this gives comparable results to the full parameter estimates obtained from the binary neutron star merger GW170817. For the advanced LIGO-Virgo network, it is only binaries which are close to edge-on, i.e. with inclinations $\iota \gtrsim 75^\circ$, that will be distinguishable from face-on systems. Extended networks which have good sensitivity to both gravitational-wave polarizations will only be able to constrain the inclination of a face-on binary at signal-to-noise ratio 20 to $\iota \lesssim 45^\circ$. Even for loud signals, with signal-to-noise ratio of 100, face-on signals will only be constrained to have inclinations $\iota \lesssim 30^\circ$. In the absence of observable higher modes or orbital precession, this degeneracy will dominate the mass measurements of binary black hole mergers at cosmological distances.

A possible way to break this degeneracy is through the observation of higher harmonics of the waveform. In chapter 4, we investigate the observability of higher harmonics in gravitational wave signals emitted during the coalescence of binary black holes. We decompose each harmonic into an overall amplitude, dependent upon the masses and spins of the system, and an orientation-dependent term, dependent upon the inclination and polarization of the source. Using this decomposition, we investigate the significance of higher multipole moments over the parameter space and show that the $\ell = |m| = 3$ harmonic is most significant across much of the sensitive band of ground-based interferometric detectors, with the $\ell = |m| = 4$ harmonic having a significant contribution at high masses. We introduce the higher harmonic signal-to-noise ratio (SNR), and show that a simple threshold on this SNR can be used as a criterion for observation of higher harmonics. Finally, we investigate observability in a population of binaries and show that higher harmonics will
be observable in a few percent of binaries, typically those with unequal masses and viewed close to edge-on. We find that the majority of binaries with mass ratio greater than 4:1 will have an observable $\ell = |m| = 3$ harmonic.

LIGO and Virgo’s third observing run brought the first clear detection of higher harmonics, with the detection of GW190412\[9\] and GW190814\[12\]. In chapter 5, we quantify these detections using the method presented in chapter 4. We find that in both cases the $\ell = |m| = 3$ harmonic is confidently detected at the three and five sigma level respectively. This has an important impact on the parameter estimation of these systems. GW190412 was the first observation of a compact binary with confidently asymmetric masses. Higher harmonics in the signal were decisive here, removing any lingering support in the posterior for comparable masses. For the first time, and again thanks to the (3,3) multipole in the data, the inclination of the orbital plane could be constrained away from face-on/face-off. A similar impact on the inclination measurement by the (3,3) multipole was witnessed in GW190814. This binary was interesting for another reason. Its lighter component was constrained to be both less massive than any black hole and more massive than any neutron star previously detected in a compact binary. We find that a waveform including both precession and higher multipoles bring about the most stringent constraint on the mass.

Precession of the orbital plane is another physical effect that had not been observed prior to 2019. In chapter 6 we discuss the evidence for precession and higher multipoles for all binary black hole mergers detected in the first half of LIGO-Virgo’s third observing run\[16\]. We find that GW190412’s signal exhibits evidence for precession, and one further event shows evidence for a (3,3) harmonic in the signal, both with p-values less than 3%. Further we highlight the potential for strong higher harmonics GW190521, which may have been missed in the original analysis\[16\] but are indicated by the posterior distributions presented in Ref.\[13\].

Finally in chapter 7 we return to the science case for the next generation of detectors, with an investigation into the prospects of observing and measuring the properties of seed black holes, the hypothetical progenitors of supermassive black holes. The hypothesis is that light black hole ”seeds” of several $10^2 M_\odot$ grow through successive mergers and accretion to swiftly transit into the supermassive regime, some within the first billion years of cosmic time. Accretion of these systems produces electromagnetic radiation too faint to be detected by even the deepest future facilities. Mergers, however, will emit gravitational waves detectable with next generation networks at a broad range of redshifts $2 < z < 15$. The large detector frame masses of these binaries means they are only in band for just a few cycles, making precise estimation of parameters challenging. Observation of the dominant harmonic oscillating at twice the orbital frequency typically results in broad constraints on both mass ratio and the inclination angle between the orbital angular momentum and line of sight. This latter parameter is important as it is degenerate.
with redshift and therefore also source frame masses. We find that higher harmonics ringing at 1.5 and 2 times the frequency of the dominant will be observable for the majority of these systems, allowing a greater fraction of these seeds to be detected. These additional harmonics also result in tighter constraints on both mass ratio and inclination angle. This improves the prospects for differentiating light seed black hole candidates from black hole mergers of stellar origin.
Chapter 2

Localization of binary neutron star mergers with second and third generation gravitational-wave detectors

*Code:* The code that produced the main results in this chapter will be made publicly available here: https://github.com/sfairhurst/simple-pe/

*Collaboration:* This chapter was published in Ref. [46]. CM was responsible for most of the results and writing. SF contributed to the writing and conception throughout. VT contributed the majority of section 2.2.1 on future detectors, in addition to contributing ideas and edits throughout, and calculations at an earlier stage of the project.

2.1 Introduction

One hundred years after gravitational waves were predicted [47], the first detection of a binary black hole (BBH) coalescence by the advanced LIGO detectors [48] heralded the beginning of the era of gravitational wave (GW) astronomy. Less than two years later, and with the advanced Virgo detector also in operation, the detection of GWs from a BNS merger marked another landmark event [5]. Over the coming years, the sensitivity of the advanced LIGO and Virgo instruments will improve, and the KAGRA and LIGO India detectors will join the global network [49, 50, 51, 52, 53]. This network of advanced gravitational-wave detectors is expected to observe many more BBH and BNS mergers, as well as GWs emitted during the merger of neutron star–black hole (NSBH) binaries [54, 55]. Additionally, GWs emitted by non-symmetric neutron stars, core-collapse supernovae, and other astrophysical transient events may be observed [56, 57].

The observation of electromagnetic (EM) counterparts to GWs is a major goal
for astronomy, and will be critical for extracting the maximum science from future events. Despite the expectation that stellar mass BBH mergers don’t produce electromagnetic signals, there was broadband follow-up of GW150914 [58]. This demonstrated the willingness of the wider astronomical community to engage in multi-messenger observation of GW sources. In contrast to BBHs, compact binary systems composed of at least one neutron star have plausible EM counterparts across gamma, x-ray, optical, infrared, and radio bandwidths (for possible counterparts see [59, 60, 61, 62, 63, 64, 65, 66, 67]). Indeed, the follow-up campaign for the BNS merger GW170817 was phenomenally successful, measuring counterparts across the EM spectrum [68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 36, 85, 86, 87, 35, 87, 88, 89]. This multimessenger data provided convincing answers to many outstanding questions. For instance, the detection of a short Gamma Ray Burst (GRB) 1.7 seconds after GW170817 [69, 70, 71], and subsequent kilonova [73, 74, 75, 76, 72, 81, 35, 82, 83, 84, 36, 85, 86, 87], confirmed that BNS mergers are a progenitor of these events. Lanthanide signatures in the kilonova light curves also showed BNS mergers to be a major site for nucleosynthesis of elements heavier than iron [35, 36, 37, 38]. Furthermore the measurements of the EM redshift and, from the GW signal, the luminosity distance, allowed an independent estimate of the Hubble constant to be made [90], thus demonstrating a thirty year old prediction [91]. Crucial to these scientific results was the localization provided by the LIGO and Virgo interferometers.

Our ability to measure counterparts, and statistically identify host galaxies when no counterparts are present, depends on the GW localization. Unlike most EM telescopes, GW detectors are not pointing instruments, and localization is achieved primarily by measuring the differences in arrival times of the signal in different detectors [92]. Consequently, searching the relatively large GW localization regions (\(O(100 - 1000 \text{deg}^2)\) for the first detections [56, 93, 94]) represents a challenge for even wide field of view UV, optical and infrared telescopes. These telescopes have fields of view on the order of 10 deg\(^2\) or less [95, 96, 97]. The addition of Virgo has improved the localization ability of the network by about an order magnitude [98, 5]. Extra detectors in India and Japan will further reduce localization regions, allowing many signals to be localized to within tens of square degrees [51, 99, 100, 101].

There are plans for future gravitational-wave detectors that will be significantly more sensitive than the current generation of advanced detectors. These include upgrades to the existing detectors, such as A+ and LIGO Voyager [41], which gives the best possible sensitivity within the current LIGO infrastructure. Additionally, entirely new detectors have been proposed. The Einstein Telescope is a next-generation European gravitational-wave observatory [42, 43, 44], and Cosmic Explorer [1] is a proposed US-based future detector, both of which improve on the advanced detector sensitivity by a factor of ten or more. As well as revealing new sources of gravitational waves, these detectors will allow us to observe BBH merg-
ers throughout most of the history of the universe [102] and BNS to cosmological distances [103, 104, 105, 106]. Furthermore, the nearby signals will be very loud in these detectors, allowing for unprecedented tests of Einstein’s general relativity, and observation of matter at supra-nuclear density inside neutron stars. As with the advanced detector network, joint GW-EM observations will be vital in fully extracting the science from these observations [107, 64, 60, 108].

The science case for these new facilities is still evolving, and will continue to do so as further gravitational wave observations are made. Estimates of the accuracy with which networks of third and second generation detectors can reconstruct parameters will inform decisions over the viability of new facilities. There have been previous studies of ET that estimate the detection efficiency and the accuracy of mass measurements [104, 109, 105, 106]. Estimates of the localization ability of various third generation networks were also considered as part of a comprehensive parameter estimation study [102], and in analytical studies focussing on the low frequency benefits of 3G detectors [110] and the implications for cosmology [111]. Furthermore, detailed studies of the optimal location of future detectors have been performed [112, 113, 114].

One practical consideration is whether it would be advantageous to accelerate the development of third generation detectors, perhaps at the expense of further upgrades to the second generation, or if the operation of a heterogeneous network of detectors is preferable. To date, there is rather little in the literature on the merits of such networks. Here we investigate the differences between homogeneous and heterogeneous networks of detectors. For concreteness, we focus primarily on the sky coverage of the networks and the accuracy with which they are able to localize sources. We consider the network localization accuracy for both face-on BNS systems at a fixed distance as well as a population of BNS distributed isotropically and with a redshift distribution that follows the star formation rate shifted to account for the delay between star formation and binary merger.

Previous estimates of network localization errors largely fall into two distinct categories: the first being analytical estimates that bypass the full task of parameter estimation and reduce the parameter space by focusing primarily on source localization [92, 113, 115, 116, 117, 118, 119, 120, 121, 122, 99, 123, 124]; the second being full parameter estimation studies that extract detailed parameter estimates using Bayesian statistics [125, 126, 127, 128, 100, 129, 130, 131, 132]. Performing the full analysis has the advantage of being more accurate, but due to the computational cost the number of sources that can be considered is typically small. On the other hand, analytical studies using only the timing information [92, 121] have been shown to overestimate the localization error region [133]. Here, we make use of an improved, analytical method that incorporates amplitude and phase consistency between the sites, as well as timing [134].

This chapter is organized as follows. Section 2.2 will describe the networks used
in this study. Section 2.3 introduces the method for calculating the localization error regions. We present and analyse our main results in Section 2.4 before concluding in Section 2.5.

2.2 Future Detectors and Networks

2.2.1 Future detectors

GW detector sensitivity is limited by a number of fundamental noise sources. These noise sources can be broadly separated into two categories: displacement noise and sensing noise. Displacement noises cause motions of the test masses. Noise sources such as seismic noise and mechanical resonances are in this category. Sensing noises, on the other hand, are phenomena that limit the ability to measure those motions; they are present even in the absence of test mass motion. Shot noise and thermal noise are included in this category. In addition, there are technical noise sources which must be understood and mitigated in order that the detector sensitivity is limited by fundamental noise. Typically, low frequency sensitivity is limited by seismic noise, mid frequencies are limited by thermal noise and higher frequencies are limited by quantum noise. LIGO underwent a series of upgrades from its initial to advanced configuration to address each of the noise sources [135]. Seismic noise is being suppressed by the use of multi-stage mechanical seismic isolation and quadruple pendulum suspension systems. Thermal noise arises in test masses and suspensions and is determined by material properties and beam size. Compared to initial LIGO, advanced LIGO uses a larger beam size. This results in better averaging of beam on a larger surface area which combined with better coating and suspension material results in efficient dissipation of heat. Quantum noise arises due to statistical fluctuations in the detected photon arrival rate. Quantum noise is overcome by increasing the beam power and increasing the weight of the test masses to overcome the increased radiation pressure.

Many technologies have been proposed to further increase the sensitivity of ground based detectors. For example, building detectors underground to suppress gravity gradients [42], improving mirror coatings (Section 5.9.3 in [136]) and cryogenically cooling the mirrors for reducing the thermal noise, and using squeezed light for lowering the noise floor due to quantum noise [137]. A detailed discussion on possible technology improvements is given in [136]. In the following, we briefly introduce several proposed future detector configurations and their corresponding sensitivities (see Fig. 2.1). These are used in the following sections when comparing the performance of different networks.

LIGO Voyager: Various upgrades have been proposed for the advanced LIGO detectors [41] leading to the proposal for an upgrade to A+ in 2020 followed by a further upgrade to LIGO Voyager which is envisioned to be operational around 2025.
Voyager improves on the sensitivity of advanced LIGO by around a factor of three across a broad frequency range. The increased sensitivity is intended to be achieved by improvements in all the departments (seismic isolation system, coatings of mirrors, heavier and larger test masses, increased beam power, etc.) of the advanced LIGO infrastructure combined with frequency dependent squeezing and cryogenic cooling of mirrors [138, 139, 140].

**Einstein Telescope:** Various studies have shown that further increase in sensitivity is required for performing precise gravitational-wave astronomy, testing of general relativity and improving our understanding of exotic phenomenon like the equation of state and tidal deformability of neutron stars [59, 103, 141, 142, 143]. The Einstein Telescope is a proposed next-generation European gravitational-wave observatory [42, 43, 44] with sensitivity an order of magnitude higher than advanced LIGO and extending down to 1Hz. It intends to achieve this improvement through a combination of longer arms and improved technologies. The original design called for a triangular configuration of three interferometers with 10 km arms and 60° angle between the arms. In addition, the proposed *xylophone* configuration allows installation of separate high and low-frequency detectors. High frequency sensitivity is most easily achieved with high laser power, but this generates significant complications at lower frequencies. The divided detector avoids this issues by allowing to pursue different strategies in optimising the noise for each frequency range. Additionally, it also reduces the length of tunnel required (as each tunnel is used by two of the interferometers) and also makes the detector sensitive to both gravitational-wave polarizations [136].

**Cosmic Explorer:** There is also a proposal for a Cosmic Explorer detector [1, 41], which would be around a factor of three more sensitive than ET. The design and technology used is similar to ET but with arm length that can stretch out to between 40 to 50 km. Although the possibility of these detectors only lies in the far future, it is noteworthy that these detectors would see GW150914 like BBH mergers throughout the visible universe.

In figure 2.1 we show the sensitivities of the proposed future detectors [1], as well as the advanced LIGO design sensitivity. We show the ET *xylophone* configuration, called ET-D. Also included for comparison is ET-B, which is an alternative ET configuration where every interferometer is optimized for best overall sensitivity, but at the expense of some low frequency sensitivity. For all ET simulations in this study, ET-D sensitivity is assumed.

### 2.2.2 Networks

We will consider five networks of gravitational-wave observatories beyond the advanced detectors that are currently being built, commissioned and operated. Specifically, we consider:
2.2. Future Detectors and Networks

Figure 2.1: Target noise curves for existing and future detectors [1]: advanced LIGO at design sensitivity (aLIGO); LIGO Voyager; Einstein Telescope (two proposed configurations, ET-B and ET-D) and Cosmic Explorer (CE).

(i) A network comprising detectors at the three LIGO sites (Hanford, Livingston and India) where the detectors have been upgraded to LIGO Voyager sensitivity. (Voyager)

(ii) A network comprising the three LIGO Voyager detectors complemented by a triangular ET detector in Europe. (Voyager-ET)

(iii) A network with three L-shaped detectors at ET sensitivity distributed globally. (3ET)

(iv) A network comprising a triangular ET detector and two Cosmic Explorer detectors (CE-ET)

(v) A network of three Cosmic Explorer detectors (3CE).

Networks (i) and (ii) arise naturally from existing proposals, but there is currently no global plan for a third generation network. Although there is no proposal for a network of ET detectors, we include this as configuration (iii), to facilitate comparison with the heterogeneous Voyager-ET network. It can be shown that the triangular ET detector has the same sensitivity to the two gravitational-wave
polarizations as two co-located L-shaped detectors of length 10.6 km whose orientations differ by 45° [42]. For simplicity, we use a network of L-shaped ET detectors with three detectors each of comparable sensitivity to one of the two effective L-shaped detectors in the triangular ET. However, the freedom to orient and place the detectors far apart results in a network typically with improved localization and diminished ability to resolve both polarizations. We also consider a comparable network comprised of Cosmic Explorer detectors as well as a heterogeneous CE-ET network. Both the Voyager-ET and CE-ET networks exhibit substantial heterogeneity of sensitivities with a factor of three difference over a broad frequency range. The majority of previous studies, have assumed that the detectors in the network have identical sensitivity [99, 121, 92, 112].

The locations of future gravitational-wave detectors have not yet been finalized. In this study we make use of the detector locations derived in [112] to optimize the location of future detectors. There, a three part figure of merit is used to determine the optimal location of detectors in a network, comprising equal parts: 1) How equally the network can determine both polarizations; 2) a simple measure of localization ability based on the area of the triangle formed by the detectors and 3) how accurately the chirp mass can be measured. The locations and orientations of all detectors are reported in Appendix 2.A.

For the LIGO Voyager network, the location of the Hanford and Livingston detectors is fixed. Their orientations were chosen so that they were, as much as possible, sensitive to the same gravitational-wave polarization, thus improving the chances of coincident detection [113]. The location of the LIGO India detector has not been announced at this time, so we use the optimal location from [112], which places it in a seismically quiet location. The triangular ET detector is added to this network to form the Voyager-ET network. In [112], it was shown that a location in Slovakia gave maximum flexibility when constructing a global network, so we choose this. Since ET is equally sensitive to both gravitational-wave polarizations, the orientation of the detectors does not affect the results. It should be noted that the precise location in Europe of the triangular ET does not have a significant impact on results.

For the 3ET and CE networks, we are free to optimally site all three of the new detectors. In [112], with the additional requirements that the detectors lie on the land and avoid areas with a high degree of human activity, the authors arrived at two comparable networks for three triangular ET detectors. The best configurations had detectors in either Australia, Central Africa and the USA or in Australia, Europe and South America. Although the optimization was performed for triangular ET

\footnote{This does, however, lead to an increase in 40% in the length of tunnels required. In the case where the cost of constructing the tunnels is dominant, one could instead construct two 7.5km interferometers within the same tunnel, making use of each tunnel twice as is done in the triangular ET design. In this scenario the tunnel length of the two L-shaped detectors is equal to the triangular configuration [144].}
detectors, we use the first set of locations for both the 3ET and CE networks. We then optimize the orientation of the detectors based on part 1 of the figure of merit — sensitivity to both gravitational wave polarizations — as parts 2 and 3 will be largely insensitive to the orientation (this is described further in Appendix 2.A). Finally, for the CE-ET network, we retain the two CE detectors in the USA and Australia and augment the network with a triangular ET detector in Europe.

This by no means covers the full set of proposed future detectors and networks, but is sufficient to allow us to explore the impact of a heterogeneous set of detector sensitivities and compare this to networks where all detectors have the same, or similar, sensitivity.

2.2.3 Network Sensitivity

The response of a detector to the two polarizations of a gravitational wave is given by $F_+$ and $F_\times$, which are functions of the sky location and polarization of the wave [145]. By writing the response time-independently, we assume the strain is constant during the light round-trip time in the interferometer arms. For Voyager this is generally a good approximation. However for CE, and to a lesser extent ET, where the round trips are respectively 20 and 5 times longer, the response pattern of a detector can differ in both amplitude and phase significantly from the static case for GW frequencies $\gtrsim 2$kHz [146]. Nonetheless, for compact binary coalescences (CBCs) the vast majority of the SNR is accumulated at frequencies $\lesssim 2$kHz, and the static response is a good indication of sensitivity to these sources. We also do not account for the motion of the earth. The loudest BNS signals could last hours to days in the detector bandwidth, introducing a time dependence in the detector response. This can be used to improve the localization, particularly when the network is operating with just one, or two, detectors and baseline triangulation is not possible [110, 111]. However for the majority of signals this effect will be small.

For networks of equally sensitive detectors, the network response at a given sky point is given by $\sum_i (F_+^i)^2 + (F_\times^i)^2)^{1/2}$ [119, 112]. However, when dealing with heterogeneous networks, we must generalize the expression to take account of the detector sensitivity. To do so, we introduce a sensitivity measure $\sigma_{h,i}$ defined as [147]

$$\sigma_{h,i}^2 = 4 \int_0^\infty \frac{\tilde{h}_o(f)^2}{S_i(f)} df$$

(2.1)

where $\tilde{h}_o(f)$ is the gravitational-wave strain from a fiducial system placed overhead the detector at a fixed distance and face-on, and $S_i(f)$ is the power spectral density (PSD) of the detector noise. Then $\sigma_{h,i}$ gives the expected SNR for such a signal in detector $i$. For our study, we take $\tilde{h}_o(f)$ to be the signal from a face-on binary

\footnote{This quantity is typically often denoted $\sigma$. To avoid confusion with the signal bandwidth, $\sigma_f$, introduced in the next section, we have introduced the subscript $h$ here.}
neutron star system at 1 Mpc from the detector. Then, we weight the response of each detector by the sensitivity, defining [148]

\[ w_{i,+} = \sigma_{h,i} F_{i,+} \cdot \] (2.2)

The relative sensitivity of the network at a given sky point is then defined as the network response,

\[ N_R = \left( \sum_i \left[ \left( w_{i,+} \right)^2 + \left( w_{i,+} \times \right)^2 \right] / \sum_j \sigma_{w,j}^2 \right)^{1/2}, \] (2.3)

where the indices \( i,j \) run over the detectors. Using this definition, the maximum network response is unity and this will only be achieved when all detectors are aligned to be maximally sensitive to the same sky position. This extends the definition of [119] to a heterogeneous network and is closely related to the network sensitivity to generic transients introduced in [118].

We are also interested in the relative sensitivity to the two gravitational-wave polarizations. To define this unambiguously, we must identify a preferred choice of the + and \( \times \) polarizations or, equivalently, a choice of polarization angle. We define the Dominant Polarization Frame [149, 148], which gives the maximum sensitivity to the + polarization. To do so, we introduce

\[ w_{\text{net},+} = \left( w_{1,+}^1, \ldots, w_{N,+}^N \right). \] (2.4)

The dominant polarization frame, for a given sky location, is the unique frame such that: (1) \( w_{\text{net},+} \cdot w_{\text{net},+} = 0 \); (2) the network is maximally sensitive to the + polarization, thus ensuring \( |w_{\text{net},+}| \geq |w_{\text{net},\times}| \). The ratio of \( |w_{\text{net},\times}| / |w_{\text{net},+}| \) is called the network alignment factor [149] and will vary from one — equal sensitivity to both polarizations — to zero — sensitivity to a single polarization.

In Figure 2.2, we plot the Network Response and Alignment Factor as a function of sky location for the five networks under consideration. The Voyager network has the best sensitivity above and below the location of the two US LIGO detectors, as expected. It has limited sensitivity to the second polarization over large parts of the sky, including the locations with best network sensitivity. In the Voyager-ET network, the ET detector dominates the sensitivity so, as expected, we see the best sensitivity above and below the ET detector. The triangular ET is equally sensitive to both polarizations, and so the Voyager-ET network has good sensitivity to the second polarization over the majority of the sky. Even in regions where ET has poor sensitivity, the second polarization is reasonably well measured by a combination of ET and the three LIGO Voyager detectors.

The 3ET and 3CE networks are comprised of detectors in identical locations, so the relative sensitivity over the sky will be identical for these networks. These networks have good coverage over much of the sky, but the peak sensitivity is no-
Figure 2.2: Relative sensitivity of the different networks over the sky: Voyager, Voyager-ET, 3ET/3CE and CE-ET. Left: network response as a function of sky position and right: alignment factor as a function of sky position. Also shown are the locations of detectors in each network. Magenta markers are for Voyager detectors, white for ET, and red for CE. Since both $N_R$ and the alignment factor are invariant under an overall scaling in network sensitivity, 3ET and 3CE will have identical patterns.
The Sky Coverage [119] of a network is defined as the fraction of the sky for which the response is greater than $1/\sqrt{2}$ of the maximum. The Sky Coverage of the homogeneous ET and CE networks is 79%. Even though the LIGO Voyager network also has three equal sensitivity detectors, the similar orientations of the LIGO Hanford and Livingston detectors lead to a sky coverage of 42%. For the heterogeneous CE-ET and Voyager-ET networks, the sky coverage is 44% and 37% respectively. This confirms what the plots suggest and indicates that the 3ET and 3CE networks have the most uniform response across the sky.

### 2.3 Source Localization

To investigate the ability of different networks to localize sources, we use the formalism introduced in Refs. [92] and [121] and references therein. In those papers, it was shown that localization is primarily determined by the timing accuracy, $\sigma_t$, in each detector which, in turn, is inversely proportional to the signal strength and frequency bandwidth $\sigma_f$ of the signal in the detector. Specifically, given a signal $h(t)$, the effective bandwidth is defined as

$$\sigma_f^2 = \left( \frac{4}{\rho^2} \int_0^\infty df \frac{f^2 |h(f)|^2}{S(f)} \right) - \left( \frac{4}{\rho^2} \int_0^\infty df \frac{f|h(f)|^2}{S(f)} \right)^2,$$

(2.5)

where the SNR, $\rho$, in the absence of noise, is given by

$$\rho^2 = 4 \int_0^\infty \frac{|h(f)|^2}{S(f)} df.$$

The timing accuracy for a signal with SNR $\rho$ is then given by

$$\sigma_t = \frac{1}{2\pi \rho \sigma_f}.$$

(2.6)
Thus, $\sigma_l$ scales inversely with the SNR of the GW, $\rho$, and the effective bandwidth, $\sigma_f$, of the signal in the detector.

Using these expressions, it is possible to calculate the reduction in network SNR due to errors in sky location and derive, at leading order, a relatively simple expression for the localization area. The probability distribution for the location of the source (from a sky location $\mathbf{R}$) is given by

$$p(\mathbf{r}|\mathbf{R}) \propto p(\mathbf{r}) \exp\left[-\frac{1}{2}(\mathbf{r} - \mathbf{R})^T \mathbf{M}(\mathbf{r} - \mathbf{R})\right] \quad (2.7)$$

where $\mathbf{r}$ is the reconstructed position of the source, $p(\mathbf{r})$ is the prior distribution (taken as uniform on the sphere), and the matrix $\mathbf{M}$ describes the localization accuracy and is given by

$$\mathbf{M} = \frac{1}{\sum_k \sigma_k^2} \sum_{i,j} \frac{(\mathbf{D}_i - \mathbf{D}_j)(\mathbf{D}_i - \mathbf{D}_j)^T}{2\sigma_i^2\sigma_j^2} \quad (2.8)$$

and $\mathbf{D}_i$ gives the location of the i-th detector. Thus, the localization is improved by having greater separation between the detectors and good timing accuracy, i.e. high SNR and large bandwidth of the signal in the detectors.

Localization can be improved by accounting for the relative amplitude and phase of the signal observed in each detector. These are necessarily constrained in a network of three or more detectors by the fact that a gravitational wave has only two polarizations. When taken into account, this leads to a more rapid falloff in the network SNR away from the correct sky location which, in turn, leads to an improvement in localization. This has been discussed in detail in [134], and a similar analysis was presented in [120]. The resulting probability distribution for the localization has the same form as eq. (2.7) with a modified expression for the matrix $\mathbf{M}$, which nonetheless remains quadratic in the detector separations $\mathbf{D}_i - \mathbf{D}_j$.

Based on timing information alone, a source observed in three detectors can be localized to two regions in the sky. The two locations lie above and below the plane formed by the three detectors. When we require the signal to be consistent with two gravitational-wave polarizations, this places restrictions on the relative amplitudes and phase differences between the detectors. In many cases, this information can be used to exclude the mirror location and restrict the source to a single sky position. Of course, with four or more sites, timing information alone can be used to localize a source to a single sky location.

In the following studies, we generate a population of events and determine which events would be detected by a given network and how accurately they would be localized. In all instances, we use the above formalism and ignore the effects of noise which would change the recovered SNR and offset the optimal sky location from the expected values. We require that signals would be confidently detected...
Chapter 2. Localization of binary neutron star mergers with second and third generation gravitational-wave detectors

by the detector network. Specifically, we require a network SNR of at least 12 as well as an SNR above 5 in at least two detectors in the network.\(^3\) Furthermore, since the localization methods described above are accurate only to leading order, our localization results are based only upon detectors for which the signal has an SNR greater than 4. As discussed in [92], at lower SNRs the approximations used here break down. Due to these approximations, and our neglect of the changing detector response within a localization region, one should not expect exact agreement between our results and those obtained using existing codes (such as [122]). Detailed comparisons between our method and existing codes are planned in the future, see [134] for details.

The thresholds used mimic those used in the analysis of GW data [151] to obtain events with a false alarm rate of less than 1 per century [122, 121] and are the same as used in previous studies [92, 121]. In addition, they seem appropriate based on the initial gravitational wave observations, where GW150914 and GW151226 both satisfied these requirements while LVT151012 had a network SNR of 10 and was not unambiguously identified as a signal [54]. As GW observations become more common, and searches are further improved [152], it is possible that the detection thresholds will be reduced. While this will change the details of the results presented below, the relative performance of the networks will remain similar.

2.4 Results

2.4.1 Face on Binary Neutron Star Mergers

We first investigate the ability of the networks to localize a given source at a fixed distance, as a function of the sky location of a source. We simulate 1.4 - 1.4 M\(_\odot\) BNS systems that are oriented face on (i.e. with inclination, \(\iota = 0\)) at a fixed distance at each point along a two dimensional 16 by 16 grid of sky coordinates. We repeat the study for sources at redshifts of \(z = 0.2\) (\(D_L = 1\) Gpc) and \(z = 0.5\) (\(D_L = 3\) Gpc). At each sky location, we calculate the expected SNR in each of the detectors in the network. For any signal that meets the detection and localization criteria given above, we calculate the 90% localization region. Since the BNS systems are face on, the GWs are circularly polarized, i.e. both polarizations have the same amplitude. Thus it is the overall sensitivity, and not the relative sensitivity to the two polarizations that will affect the localization ellipses [99]. This small set of sources with fixed distance and inclination allows for simple comparison with previous studies, e.g. [99, 51]. Furthermore, for sources which are inclined at an angle \(\iota \lesssim 45\) deg, the relative amplitude of the two GW polarizations renders them indistinguishable from face-on signals [153, 134] and, indeed, the GW signals seen to date, including GW170817, have all been consistent with face-on binaries [90].

\(^3\)For a discussion of the effects of changing these thresholds and, in particular, removing the
2.4. Results

Figure 2.3: The localization ellipses at different sky locations for face-on 1.4-1.4 BNS binaries at a redshift of: left - $z = 0.2$ (luminosity distance of 1 Gpc) and right - $z = 0.5$ (luminosity distance of 3 Gpc). The red crosses indicate that the BNS at this sky position would not be detected — either due to a network SNR less than 12, or not having SNR $> 5$ in at least two detectors. The blue + symbols indicated sources that would be detected, but not well localized due to being identified in only two detectors. The ellipses give the 90% localization regions for a source from a given sky location. Detector locations are shown with golden markers.

Figure 2.3 shows the localization regions for these BNS sources in the five networks under consideration. In the figures, a red cross indicates that the detection criteria (network SNR $> 12$ and two detectors with SNR $> 5$) were not met for a BNS at this sky position and redshift; a blue plus indicates that the source would be detected but fails our localization criterion (SNR $> 4$ in three or more sites). For signals which would be confidently detected, and observed in at least three sites, the green ellipses show the 90% confidence region for the localization.

For BNS mergers at $z = 0.2$, the LIGO Voyager network would observe the signal over the majority of the sky. There are, however, four patches where the signal would not be found, which correspond to areas of poor sensitivity for the two US LIGO detectors. Furthermore, there are regions where the signals would be detected but not localized, based on our conditions, and these correspond to locations where single detector thresholds see [150].
LIGO India has poor sensitivity. For those signals which are localized, the areas are typically large, as these events will be close to the detection threshold in the network. We can clearly identify a band for which the localizations are extended in one dimension. These points are close to the plane defined by the three detector locations. A large change in sky location, in a direction perpendicular to the plane of the detectors, leads to a relatively small change in the relative arrival times and consequently poor localization. These results are consistent with those obtained for the advanced LIGO network (incorporating LIGO India) given in [99].

The Voyager-ET network is able to detect sources at $z = 0.2$ over the essentially the whole sky. For localization, we require the signal to be observed at three sites; although all three of the detectors in the triangular ET will observe the signal, they provide rather poor localization by themselves. Thus, the network is limited by the requirement that two LIGO Voyager detectors observe the signal. The sky locations where sources are not localized correspond to the locations for which the US LIGO detectors have poor sensitivity, and these sources are only detected in ET and LIGO India. The 3ET network also gives excellent coverage over essentially the whole sky. There are still a handful of points for which localization is not possible. Again, these correspond to points where one of the detectors has close to zero sensitivity. As before, we see the characteristic extended ellipses at locations which lie close to the plane defined by the three ET detectors.

For signals at $z = 0.5$ we consider the three networks comprised of ET and CE detectors. In all cases, the sources are observed over essentially the whole sky. For the 3ET network, there are significant regions where the source is not well localized as it is seen in only two detectors, but the size of these regions shrinks for the CE-ET and 3CE networks due to the increased sensitivity of the CE detector. Finally, as expected, the signals are relatively poorly localized in directions close to the plane defined by the three detectors.

For a two-site observation, the localization is typically restricted to a fraction of a ring in the sky with an area of hundreds of square degrees [48, 54] and we consider these sources to not be localized. The degeneracy along the ring is broken by relative amplitude and phase measurements in the different detectors. For events observed with the triangular ET detector and a single L-shaped detector, the localization may be greatly improved — the triangular detector recovers the amplitude and phase of both GW polarizations so a single, additional observation will provide enough information to break the sky location degeneracy. Furthermore, when there are additional detectors in the network that did not observe the event, this information can be used to further improve the localization. We do not consider these effects here, but note that it would be interesting to examine in detail localization with a network comprised of one triangular and one L-shaped detector.

In these plots we are ignoring the fact that sources detected at three sites may be localized to two distinct patches in the sky, one above and one below the plane.
formed by the three sites. In many cases, the degeneracy can be broken based on consistency of the observed amplitude and phase of the signal in each of the detectors. For the systems at \( z = 0.2 \), both the 3ET and Voyager-ET networks will provide localization to a single region for essentially all sky locations. Voyager localizes to one patch on the sky 70% of the time, and so about a third of the localization ellipses shown below will be augmented by a similar sized region in the mirror location. At \( z=0.5 \), the CE and ET networks all localize to a single patch for at least 95% of sky locations.

### 2.4.2 A Population of Coalescing Binary Neutron Stars

![Detection and localization efficiency as a function of redshift and luminosity distance](image)

Figure 2.4: The detection and localization efficiency as a function of redshift and luminosity distance of the Voyager, Voyager-ET and 3ET networks (left column) and the 3ET, CE-ET and 3CE networks (right column). For visual comparison, 3ET is plotted as a solid line in both. Left column: Voyager-ET and Voyager are the dashed and dotted lines respectively. Right column: CE-ET and 3CE are the dashed and dotted lines respectively. From bottom to top, the rows show the fraction of events at a given redshift/distance that will be detected and localized within 100, 10 and 1 deg\(^2\).

Now, let us consider network localization for a population of BNS coalescences. We follow Singer et. al. [122] in choosing the BNS component masses uniformly in the astrophysically motivated range \( 1.2 - 1.6M_\odot \). This encompasses the masses of all observed neutron stars in binaries and the 1-sigma interval of the initial mass function for a variety of formation mechanisms [154, 155]. The orientation of the sources is uniformly distributed: uniform in polarization, cosine of source inclination and the phase of the GW at merger. We distribute the sources isotropically and assume their redshift distribution to follow the star formation rate density (SFRD) shifted to account for the delay between star birth and BNS merger [3], assuming
Chapter 2. Localization of binary neutron star mergers with second and third generation gravitational-wave detectors

standard cosmology \cite{156},

\[
\frac{dR_{BNS}}{dz} = \frac{dV}{dz} \frac{\Psi_{BNS}(z)}{1 + z} \quad (2.9)
\]

where \(dV/dz\) is the differential comoving volume. The \((1 + z)^{-1}\) factor accounts for the fact that a time interval \(\Delta t_z\) in the rest-frame at \(z\) will be redshifted in the detector frame \((1 + z)\Delta t\). The rest-frame merger rate density is given by

\[
\Psi_{BNS}(z) = \int_{t_\ast}^{t(z)} \psi(z(t_f)) P(t - t_f) dt_f \quad (2.10)
\]

where \(\psi(z(t_f))\) is proportional to the observed cosmic SFRD (Eq. (15) from \cite{2}), evaluated at the redshift the binary formed, and \(P(t - t_f)\) encodes the delay time distribution. We assume the earliest epoch of star formation \(t_\ast\) to be at \(z_f = 20\). The integral is normalised to give a local merger rate density of \(\Psi_{BNS}(0) = 1500 \text{ Gpc}^{-3} \text{yr}^{-1}\) \cite{5}. We assume a distribution of possible delay times between star formation and BNS merger \(P(t_D) \propto 1/t_D\), and a minimum delay of \(t_D^{\text{min}} = 0.2 \text{ Myr}\) corresponding to the smallest time a star can supernova and form a BNS \cite{4}. The choice of \(t_D^{\text{min}}\) has little effect on the overall distribution. The \(t_D^{-1}\) distribution of delay times is motivated both by population synthesis studies \cite{4} and observations of merger times for six neutron star binaries \cite{157} (see \cite{158} and references therein).

Figure 2.4 shows the detection efficiency — the fraction of events that would be observed — for each network as a function of redshift or distance. For those BNS mergers which are detected by a given network, we calculate the 90% confidence sky localization using the prescription given in Section 2.3. We also show the fraction of events that would be localized within \(1, 10,\) and \(100 \text{ deg}^2\) for each network as a function of redshift.

On the left hand side of Figure 2.4, we consider the Voyager, Voyager-ET and 3ET networks. The Voyager-ET and 3ET networks have rather comparable sensitivities, both networks identify over 90% of sources within a redshift of \(z = 0.2\) and the majority of signals within a redshift of \(z = 0.4\). The LIGO Voyager network has good all sky sensitivity within a redshift of \(z = 0.1\), after which it drops rapidly with essentially no sensitivity beyond \(z = 0.4\). Since we require a source to be observed in three sites for good localization, it is unsurprising that the Voyager and 3ET networks are capable of localizing the majority of observed sources — in particular, essentially all sources are localized within \(100 \text{ deg}^2\) and the majority within \(10 \text{ deg}^2\). For the heterogeneous Voyager-ET network, the fraction of sources localized is much lower than the fraction detected. For example, at \(z = 0.4\) over half of all sources are detected but only 10% are localized within \(10 \text{ deg}^2\). These are the events which are too distant to be observed by the LIGO Voyager detectors so, while they are observed by ET they cannot be localized. For all three networks, only a fraction of events will be localized to within \(1 \text{ deg}^2\) and those will be primarily nearby, loud
events. For a 3ET network, half of the events at a redshift of \( z = 0.15 \) will be localized to within 1 deg\(^2\).

The right panel of Figure 2.4 shows the same results for the CE and ET based networks. The results are comparable to those described above: detection efficiency is limited by the second most sensitive detector in the network, while localization requires a third detector to observe the signal. In particular, we note that while the CE-ET and 3CE networks have similar overall detection efficiencies, the 3CE network provides much better localizations. For example, 3CE localizes 50\% of sources at \( z = 2 \) to 100 deg\(^2\) while the CE-ET network is unable to give good localizations for signals at this redshift. We note that those signals which are localized in the heterogeneous CE-ET network are typically localized within 10 deg\(^2\) as they will be recovered with high SNR in the Cosmic Explorer detectors. Finally, it is again only the loudest, nearby signals which are localized within 1 deg\(^2\). The CE-ET and 3CE networks localize half of signals within 1 deg\(^2\) to a redshift of \( z \sim 0.25 \).

Figure 2.5: The number of BNS observations and localizations per year with future networks, as a function of redshift and luminosity distance. The y-axis is scaled so that the area under the curves gives the number of events per year. This assumes an isotropic redshift distribution that follows the star formation rate [2] shifted to account for the delay between star birth and BNS merger [3] and a distribution of possible delay times \( P(t_D) \propto 1/t_D \), with a minimum delay of \( t_D^{\text{min}} = 0.2 \) Myr [4]. The local merger rate density is taken as \( \Psi_{\text{BNS}}(0) = 1500 \) Gpc\(^{-3}\)y\(^{-1}\) [5]. Note that the y-axis on 1 deg\(^2\) plot is different from the others. For visual comparison 3ET is plotted as a solid line in both column. Left column: Voyager-ET and Voyager are the dashed and dotted lines respectively. Right column: CE-ET and 3CE are the dashed and dotted lines respectively. From bottom to top, the rows show the number density of events at a given redshift/distance that will be detected and localized within 100, 10 and 1 deg\(^2\).

Figure 2.5 shows the expected number of observed events as a function of redshift for the five networks, and the overall results are summarized in Table 2.1. In order to obtain these results, we have taken a population of sources assumed to follow the
star formation rate. While the star formation rate is reasonably well constrained up to its peak at a redshift of one to two, observational evidence at higher redshifts, and for the distribution of delay times between binary formation and merger, is relatively scarce. Furthermore, we use an intrinsic merger rate of $1500 \text{ Gpc}^{-3} \text{y}^{-1}$, which lies within the current range of predicted rates [5], and we use this fiducial rate when comparing the network sensitivities. There is at least an order of magnitude uncertainty on the rate of BNS mergers. Overall changes to the merger rate will simply scale the number of observations for all networks equally but changes in the redshift evolution of the rate will affect relative performances.

As expected, the number of events detected by the 3ET and Voyager-ET networks are comparable. The 3ET network is sensitive to 50% more BNS mergers than Voyager-ET, while the Voyager network observes less than a tenth as many mergers. However, due to the differences in localization, the Voyager-ET localizes only a quarter as many events as the 3ET network and, for events localized within 10 deg$^2$ the peak of the redshift distribution is at 0.2 rather than 0.4 (and within 1 deg$^2$ the peak is at 0.06 rather than 0.2). We see similar results for the CE-ET and 3CE networks: they are both able to detect a comparable number of events, but significantly fewer are localized by the heterogeneous network.

For the results presented above, we have assumed a 100% duty cycle for all detectors in the network. Here we consider how the results would change for a more realistic observing scenario. Experience from previous observing runs tell us that at design sensitivity duty cycle for an individual detector is about 80%, although it is possible that this may be different for third-generation detectors. Thus, we assume an 80% duty cycle for all detectors in the network, with uncorrelated operating times between the detectors, with the exception of the triangular ET where we assume perfect correlation of up-time between the interferometers. For the three detector networks, this means that roughly half the time all three detectors will be operational. For an additional 40% of the time, two detectors will be operating and they will, on average, be sensitive to half as many signals as the three detector network and, based on our criteria for localization, not localize any of them. With the four site Voyager-ET network, we will have all four detectors operating 40% of the time. There will then be 10% of the time when the three Voyager detectors are operating, and an additional 30% when ET plus two of the Voyager detectors are running. Localization with ET and two Voyagers will vary depending on which Voyager detector is offline. For instance, a network of the triangular ET augmented by the two American Voyagers at Hanford and Livingston would localize 70% as many within 100 deg$^2$ as Voyager-ET but only provide 30% as many sub-squared degree localizations. This is due to the fact that the alignment of Hanford and Livingston maximises the chances of the network meeting the localization criteria, but coupled with the smaller baseline between these detectors, the alignment also limits the precision of the localization. On the other hand, when one of the American
### 2.4. Results

<table>
<thead>
<tr>
<th>Network</th>
<th>Voyager</th>
<th>Voy-ET</th>
<th>3ET</th>
<th>CE-ET</th>
<th>3CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detected</td>
<td>6000</td>
<td>83000</td>
<td>120000</td>
<td>130000</td>
<td>160000</td>
</tr>
<tr>
<td>Localized</td>
<td>5500</td>
<td>23000</td>
<td>100000</td>
<td>470000</td>
<td>140000</td>
</tr>
<tr>
<td>Within 1 deg$^2$</td>
<td>100</td>
<td>830</td>
<td>5600</td>
<td>24000</td>
<td>37000</td>
</tr>
<tr>
<td>Within 10 deg$^2$</td>
<td>1900</td>
<td>14000</td>
<td>73000</td>
<td>300000</td>
<td>650000</td>
</tr>
<tr>
<td>Within 100 deg$^2$</td>
<td>4800</td>
<td>23000</td>
<td>95000</td>
<td>450000</td>
<td>1300000</td>
</tr>
<tr>
<td>Median Area (deg$^2$)</td>
<td>22</td>
<td>-</td>
<td>7</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td>Single Patch</td>
<td>45%</td>
<td>99%</td>
<td>84%</td>
<td>97%</td>
<td>90%</td>
</tr>
</tbody>
</table>

Table 2.1: Performance metrics for Voyager, Voyager-ET, 3ET, CE-ET and 3CE networks for a population of BNS coalescences distributed uniformly in comoving volume with an intrinsic merger rate of 1500 Gpc$^{-3}$y$^{-1}$. From top to bottom: The number of sources per year that are detected and localized by each network; the number of sources localized per year within 1, 10 and 100 deg$^2$ respectively; the median localization area of all detected sources and the fraction of localized sources whose position is restricted to a single patch in the sky. Note, the median source is not localized by the two heterogeneous networks CE-ET and Voyager-ET.

Voyager detectors is offline, only about half of the number will be localized within 10 deg$^2$ and 100 deg$^2$ as Voyager-ET, with 60% (40%) as many sub squared degree localizations when the Livingston (Hanford) detector is offline.

#### 2.4.3 Implication for EM Followup

The observation of broad-band electromagnetic emission associated with GW170817 provided conclusive evidence that BNS mergers are a progenitor of short GRBs, and also power kilonova emissions [68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 36, 85, 86, 87, 35, 87, 88, 89]. The primary motivation for accurate localization of future GW signals is to facilitate the observation of electromagnetic counterparts. The requirements on localization will depend upon the strength of the electromagnetic emission accompanying a BNS merger, as well as the ability of wide-field telescopes to cover the error region. While short GRBs can be observed to cosmological distances, they are believed to be rather tightly beamed, so that only a small fraction of BNS mergers would be accompanied by a GRB counterpart [159]. However, observations of GW170817 imply that the GRB emission is structured [70] with a broader, weak emission, possibly powered by a cocoon [37, 160, 161]. Since the GRB emission is likely to be essentially concurrent with the merger, it will be difficult to use GW observations to provide advanced warning to GRB satellites. On the other hand, kilonova emission is expected to last for days or possibly weeks in the optical and near infrared bands [162, 163]. Furthermore while GRBs are beamed, kilonovae are relatively isotropic, and thus are the more likely counterpart to a typical BNS merger [63]. We therefore focus the rest of the discussion on kilonovae.

The neutron-rich ejecta from BNS and NSBH mergers will undergo r-process nucleosynthesis, producing heavy elements which will subsequently decay; this decay
process will power an electromagnetic transient known as a kilonova (see e.g. [164] for details). There are various models for the kilonova emission, which depend upon the mass of the ejecta as well as its opacity [162, 163]. Broadly, the prediction is for an optical or near infrared emission, which will last for days or possibly weeks. The luminosity of the kilonova emission is uncertain, but we take a fiducial model with magnitude 22 emission from a source at 400Mpc, equivalent to the kilonova emission from GW170817 [165]. To date, there has been one other putative near-infrared kilonova observation from GRB 130603B which was observed with a magnitude of 25.8 at redshift of $z \approx 0.35$ (equivalently 22.5 at 400 Mpc) [166, 167], which is broadly consistent with this picture.

Taking our fiducial kilonova model, the current generation of wide-field telescopes, such as DECam [74], Pan-STARRS [84] and VISTA [36], which have limiting magnitudes around 22 would be able to observe kilonova emission to $z \approx 0.1$ or a luminosity distance of 400Mpc. The results in Figures 2.4 and 2.5 show that the Voyager network has good sensitivity within the range of the current generation of telescopes, and would identify and localize the majority of BNS mergers at $z \lesssim 0.1$ to within 10 deg$^2$. All of the other networks are able to detect, and localize within 10 deg$^2$, essentially every event at $z \lesssim 0.1$, thereby enabling followup with one, or a handful, of pointings.

For LSST [168], with a limiting magnitude around 25-26, kilonovae could be observed to $z \approx 0.4$ or 2 Gpc. At these distances, the sensitivity of the Voyager network is sufficient to identify only a minority, let alone provide accurate localizations. The Voyager-ET network would observe the majority of BNS within this range. However, the more distant signals would be observed only by the ET detector and consequently be poorly localized. Again using the fiducial kilonova model, around one third of BNS mergers producing a kilanova observable by LSST would be localized within 10 deg$^2$. A 3ET network would identify a similar number of mergers, and localize twice as many LSST-observable kilonovae as the Voyager-ET network. The networks with CE detectors are complete with LSST's kilonova horizon, and localize the vast majority of sources to within 10 deg$^2$.

Of course, the details of kilonova emission are still uncertain and there are models that predict significantly stronger or weaker emission. For example models, typically with a smaller ejecta mass, predict magnitude 22 emission at 200 Mpc [164]. In this case the Voyager network would be sufficient to detect and localize the majority of signals within the range of the current generation of telescopes. The Voyager-ET network would localize the majority of signals within the LSST range for these signals. Thus, the case for localization capacity of future GW networks remains intimately tied to our knowledge of the range of EM emission from these mergers.

As some of the strongest emissions are predicted from NSBH mergers, it is interesting to briefly consider them. While we have not performed simulations with NSBH systems, it is straightforward to provide approximate sensitivities based on
the BNS results given above. The sensitivity of gravitational-wave detectors scales, at leading order, as $M^{5/6}$, where $M$ is the chirp mass. Consequently, for a signal at a fixed distance, orientation and sky location, the SNR with which NSBH will be observed can be approximated as

$$\rho_{\text{NSBH}} \approx 1.1 \left( \frac{M_{\text{BH}}}{M_{\text{NS}}} \right)^{1/2} \left( 1 + \frac{M_{\text{BH}}}{M_{\text{NS}}} \right)^{-1/6} \rho_{\text{BNS}}$$

(2.11)

where $M_{\text{BH}}$ is the black hole mass and $M_{\text{NS}}$ is the neutron star mass. Thus, the observed SNR for a NSBH with $M_{\text{BH}} = 5M_\odot$ is 1.6 times that of a BNS, and 2.1 times for $M_{\text{BH}} = 10M_\odot$. Consequently, to a reasonable approximation, we can scale the distances in Figure 2.4 by these factors to obtain NSBH sensitivities.

A network’s localization accuracy to NSBHs at a given distance will be improved relative to BNSs, due to higher SNR but will be reduced due to lower frequency bandwidth (as defined in Eq. 2.5). The effective bandwidth of a NSBH binary with a $5M_\odot$ ($10M_\odot$) black hole is 10% (30-40%) less than a BNS. Thus localization areas will be roughly a factor of $1.5^2$ smaller for NSBH than BNS at a given distance. So, we can approximate NSBH localizations by rescaling the distances in Figure 2.4 by a factor of 1.5(515,778),(646,804).

For the most optimistic NSBH kilonova emissions, which predict magnitude 20 emission at 200 Mpc powered by fallback accretion [164], a network with three ET or CE detectors would identify and localize most sources within the LSST range; Voyager augmented by ET would identify but not localize the more distant sources and Voyager alone would have a range comparable to existing wide-field telescopes.

### 2.5 Discussion

We have compared the sensitivity of proposed future gravitational-wave networks to BNS signals, and their ability to accurately localize these events. We find that a minimum of two detectors, which includes the triangular ET, at an improved sensitivity are sufficient to provide a substantial increase in the number of observed sources. For example, the addition of ET to a network of LIGO-Voyager sensitivity detectors could increase the rate of observations by an order of magnitude. However, in order to obtain good source localization, we require a minimum of three sites to observe the event. Consequently, in networks with one or two detectors that are significantly more sensitive than the others, we find that the majority of detected sources are not well localized. In contrast, when the three most sensitive detectors in the network have comparably sensitivity, the majority of signals are well localized with a median localization area around $10 \text{ deg}^2$.

Previously it has been argued that building more detectors further apart improves localization (see for example [118, 121]). However, we find that this is only true when the sensitivities of the detectors in the network are approximately homo-
Chapter 2. Localization of binary neutron star mergers with second and third generation gravitational-wave detectors

geneous, as is often assumed for the advanced detector (2nd generation) networks [92, 121, 99]. In the case of heterogeneous sensitivities, such as the Voyager-ET network, the localization will often be limited since the event cannot be detected by the less sensitive detectors. In such a network, we expect that for a majority of events we will obtain limited directional information.

The interpretation of our results depends critically upon the science question of interest. In particular, the utility of accurate GW localization as a function of redshift will depend upon the strength and spectrum of the associated EM emission, and the sensitivity and field of view of the associated telescopes and satellites. For a standard kilonova model, the LIGO Voyager network provides adequate sensitivity to identify and localize potential kilonova signals for currently operating telescopes but the network must be augmented by at least one ET or CE detector to provide adequate sensitivity to localize all kilonovae that could be observed by LSST. For the models predicting the strongest kilonova emission, a three detector network of ET or CE detectors could increase, by a factor of a few, the number of events observed jointly with GW signals.

In this study, we have neglected a number of factors that affect the size of localization errors in real detector networks. Though the time of arrival and amplitude and phase of GWs carry most of the information relevant to localization, other information can reduce the size of the localization errors. These include, realistic prior distributions on other astrophysical parameters — particularly source inclination and distance, correlations with other parameters such as component masses [133], spin and precession effects. Furthermore, we continue to assume that a signal must be identified in two detectors to be detected, and three to be localized. Ideally, performing a fully coherent analysis of the data [150, 169], would improve the performance of heterogeneous networks when the SNR in the less sensitive detectors is low and would, in effect, remove our requirement of a signal being clearly identified in at least two detectors. While it is possible to localize sources with only two detectors, the first GW observations make it clear that the localization areas will typically be hundreds of square degrees, so our approximation that these sources are not localized is reasonable. However, incorporation of non-detection (or weak detection) in additional detectors, as was done for GW170817 [5], can lead to improved localizations.

We have neglected systematic uncertainties in the results presented here. Errors introduced by mismatches between template waveforms and signals [92] are expected to introduce a similar effect in all detectors and therefore the effect on the time difference, and localization, is likely to be negligible. On the other hand, errors in the calibration of GW detectors [170] will be uncorrelated, and these errors can significantly impact localization. For instance, roughly one third of the localization error budget for GW150914 was due to strain calibration uncertainty [54, 170]. At high SNR, calibration errors are expected to dominate the overall error budget.
2.5. Discussion

for localization [92]. Thus, the ability to achieve the reported sub square-degree localizations predicted here will depend critically on the calibration accuracy of the detectors, with likely requirements of uncertainties under 1% in amplitude and 1° in phase. The impact of calibration on gravitational-wave localization with future networks deserves further study.

Finally, in this chapter we have restricted attention to localization of BNS signals. For many science questions, accurate localization itself is not critical but is required for accurate measurement of the distance to the source. The gravitational wave signal from the inspiral and merger of BNS, NSBH and BBH leads to the accurate measurement of only the luminosity distance \(D_L\) and the redshifted masses \(M(1+z)\). One goal of gravitational-wave astronomy is to map the merger history of black holes and neutron stars through cosmic time. Accurate distance measurements are required not only to infer the redshift of the source, but also to obtain the mass of the source. For tests of cosmology, we require an independent measurement of the redshift. There are numerous methods proposed for this measurement, including identification of a host galaxy from EM counterpart; statistical association with a host galaxy [171]; assumption of a narrow mass range of neutron stars in binaries [158]; observation of post-merger features in the waveform [172]. In all cases, an accurate measurement of the distance (and consequently good source localization) is essential. A detailed investigation of these issues is beyond the scope of this chapter. Nonetheless, it seems likely that a network of three or more detectors of comparable sensitivity will increase scientific returns from a future gravitational-wave network.
Appendix

2.A Optimal Detector Orientations

Table 2.A.1 contains all the final locations and orientations of all detectors used in this study. In order to fix the orientations of the detectors for the ET and CE networks, we make use of the following FoM, taken from Ref. [112],

$$I = \left( \frac{1}{4\pi} \int |w^\text{net}_+ - w^\text{net}_\times|^2 d\Omega \right)^{-1/2} \quad (2.12)$$

Holding the USA detector fixed at 0° we rotated the Central Africa and Australia detectors from 0-90° (due to rotational symmetry of the polarizations all other rotations map to this basis). We optimize the ability of the ET and CE networks to observe both polarizations by choosing the orientation angles to maximize 2.12.

<table>
<thead>
<tr>
<th>Detector</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIGO Livingston Voyager</td>
<td>-90.77</td>
<td>30.56</td>
<td>162.3</td>
</tr>
<tr>
<td>LIGO Hanford Voyager</td>
<td>-119.41</td>
<td>46.46</td>
<td>234.0</td>
</tr>
<tr>
<td>LIGO India Voyager</td>
<td>76.4</td>
<td>14.2</td>
<td>346.8</td>
</tr>
<tr>
<td>European ∆ ET</td>
<td>18.7</td>
<td>48.5</td>
<td>-</td>
</tr>
<tr>
<td>Central Africa CE/ET</td>
<td>17.2</td>
<td>-9.9</td>
<td>82.4</td>
</tr>
<tr>
<td>Australia CE/ET</td>
<td>146.9</td>
<td>-35.8</td>
<td>84.4</td>
</tr>
<tr>
<td>USA CE/ET</td>
<td>-98.4</td>
<td>38.9</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.A.1: The locations and orientations of the detectors used in this study. All numbers are given in degrees and the orientation angle is defined clockwise relative to a hypothetical L-shaped interferometer with arms due North and East.
Chapter 3

Constraining the inclinations of binary mergers from gravitational-wave observations

*Code:* The code that produced the main results in this chapter will be made publicly available here: [https://github.com/sfairhur/simple-pe/](https://github.com/sfairhur/simple-pe/)

*Collaboration:* Published in Ref. [173]. SU and CM produced most of the results and writing. SF contributed to the writing and conception throughout, and was responsible for Fig 3.3.2.

3.1 Introduction

With its ground-breaking detections in the first years of its operation, the upgraded Laser Interferometer Gravitational-Wave Observatory (LIGO) and Virgo detectors have opened up the door to discovering new information about the universe. The collaboration’s many gravitational-wave (GW) detections from binary systems, including GW150914 [48] and GW170817 [174], have allowed us to draw new insights from these astrophysical sources. These developments include constraining the nuclear equation of state [33] and constraining binary black hole populations [175, 176, 177]. With more detections, we hope to learn even more about our universe, such as more accurately measuring the Hubble constant $H_0$ [40, 91] or detailing the opening angle for gamma ray bursts (GRBs) from binary neutron star systems (BNS) [159, 178, 70]. However, both of these measurements rely on the accurate measurement of the distance to the binaries and the inclination of their orbital angular momentum with respect to the line of sight. A degeneracy exists between distance and inclination making the measurement of these two parameters very difficult. Of the compact binary detections made by LIGO and Virgo, only the BNS merger GW170817 has had a tightly constrained inclination and distance. The detection of a kilonova afterglow allowed for an accurate distance measurement [36, 179],...
breaking the degeneracy with inclination. When this type of external information is unavailable, the degeneracy severely limits our ability to measure these parameters.

In this chapter, we will show that this degeneracy is typical for binary mergers. The measured amplitude and phase of the gravitational-wave signal encode the properties of the binary. In particular, it is the differing amplitude of the two polarizations of the gravitational waveform that allow us to determine the binary inclination. However, the plus (+) and cross (×) polarizations have nearly identical amplitudes at small inclination angles (less than 45°) and significantly lower amplitudes at large inclination angles (greater than 45°). This leads to two simple observations: first, that the signal is strongest for binaries which are close to face-on ($\iota \sim 0^\circ$) or face-away ($\iota \sim 180^\circ$) and thus we will be observationally biased to detecting binaries whose orbital angular momentum is well-aligned (or anti-aligned) with the line of sight [180, 119]. Second, for small angles, the amplitudes of the two polarizations are close to equal and we cannot measure distance or inclination separately. Therefore, for the majority of detections, this face-on degeneracy will limit our ability to constrain both electromagnetic (EM) emission models and the Hubble constant. There are various ways to break this degeneracy, such as using the EM measured distance or using jet modelling to constrain the opening angle. These techniques were used to improve the constraints on the inclination and distance for the BNS merger GW170817 [181, 182, 183, 6, 184].

Since an inclined binary system would produce both a high-amplitude plus polarization and a lower-amplitude cross polarization, creating a network of detectors which is sensitive to both the plus and cross polarization has been suggested to constrain the inclination using only gravitational waves [185]. A single L-shaped detector is sensitive to just one polarization of a gravitational wave. Hanford and Livingston are almost aligned, and see essentially the same polarization, while Virgo is anti-aligned and is sensitive to the orthogonal polarization. The addition of KAGRA [53] and LIGO-India [52] would further increase the network’s sensitivity to the orthogonal polarization. Thus it is hoped this five-detector network could better constrain the inclination angle and distance. We examine this possibility of constraining the inclination using only the measurement of the two GW polarizations.

There have been many studies looking at inclination constraints. From the GRB perspective they are largely divided into two groups: the first focuses on exploring the possibility of nailing down the viewing angle by comparing the rate of GRB sources observed in GWs with those in gamma rays [186, 159, 153]. The second focuses on measurements for individual detections, mainly in the case where the event has been three dimensionally localized by an EM counterpart [187, 188]. In [189], it was observed that the inclination measurement is poor for binaries with an inclination less than seventy degrees when there is no redshift information. They attribute this to a combination of the degeneracy between distance and inclination and the prior on the distance. Here we explore the origin of the degeneracy in detail.
and discuss the importance of an additional degeneracy when the binary is circularly polarized \[190\].

Inclination constraints have also been discussed in the context of distance estimates for cosmology \[191, 180, 192\] and as part of wider parameter estimation investigations \[193, 100, 194, 195\]. It was noted in \[180\] that adding detectors to a network did not seem to greatly improve the inclination measurement. Here we investigate the extreme case: a network that measures both polarizations equally as would be expected over the majority of the sky for the Einstein Telescope \[196\].

### 3.2 Measuring Distance and Inclination

#### 3.2.1 Origin of the degeneracy

The degeneracy between distance and inclination arises directly from the dependence of the gravitational waveform on these parameters and has been discussed several times previously \[191, 193, 180\]. The gravitational-wave signal, \(h(t)\), incident on a gravitational-wave detector is \[197\]

\[
\begin{align*}
    h(t) &= F_+(\alpha_s, \delta_s, \chi) h_+(t) + F_\times(\alpha_s, \delta_s, \chi) h_\times(t), \\
    \tilde{h}_\times(f) &= i\tilde{h}_0(f).
\end{align*}
\]  

(3.1)

where \(F_+\) and \(F_\times\) are the detector response to the plus and cross polarizations, respectively. The detector responses depend on the location \((\alpha_s, \delta_s)\) of the source, where we have used subscripts to distinguish the right-ascension of the source \(\alpha_s\) from a later use of \(\alpha\) to represent the alignment factor — a network’s relative sensitivity to the second polarization. In addition, we must specify a polarization angle \(\chi\) to fully specify the radiation frame. It is common \[149, 148\] to define a dominant polarization frame, for which the detector network is maximally sensitive to the plus polarization. With this choice, we can naturally characterize the network by its overall sensitivity and the relative sensitivity to the second polarization \[149, 46\]. This simplifies the comparison of different networks.

For a waveform where it is appropriate to neglect higher order modes and precession, the two polarizations given in Equation 3.1 can be expressed in terms of the two orthogonal phases of the waveform:

\[
\begin{align*}
    h_+(t) &= A^1 h_0(t) + A^3 \tilde{h}_\times(t) \\
    h_\times(t) &= A^2 h_0(t) + A^4 \tilde{h}_\times(t)
\end{align*}
\]  

(3.2)

where \(\tilde{h}_\times(f) = i\tilde{h}_0(f)\). The \(A^i\) are overall amplitude parameters, and depend on
Chapter 3. Constraining the inclinations of binary mergers from gravitational-wave observations

Figure 3.1.1: The relative contributions of the plus and cross polarizations of a gravitational-wave signal, dependent on the inclination. The red solid line indicates the amplitude of the plus polarization, while the dashed red solid line indicates the amplitude for the plus polarization with a negative phase. The blue solid line indicates the amplitude of the cross polarization. The shaded regions show the percent differences between the plus and cross polarizations. The red portion represents when the plus and cross polarization are less than 1% different. The blue region represents where the polarizations are between 1% and 5% different. The grey region represents where the polarizations are between 5% and 10% different.
3.2. Measuring Distance and Inclination

Figure 3.2.1: The relative sensitivity of detector networks to the second polarization, as encoded in the parameter $\alpha$, defined through $F_\times = \alpha F_+$ (in the dominant polarization frame where the network is maximally sensitive to the plus polarization). The left plot shows the expected distribution of $\alpha$ for second-generation gravitational-wave networks, while the right plot shows the distribution for potential third generation networks. In both cases, the distribution is the expected normalized distribution for a population of events, distributed uniformly in volume, and observed above threshold in the detector network. Thus, directions of good network sensitivity are more highly weighted. In the legends, $\bar{\alpha}$ denotes the mean $\alpha$ for each network. The second generation networks considered are LIGO Hanford and Livingston (HL); two LIGO detectors and Virgo (HLV); LIGO-Virgo and KAGRA (HLVK) and LIGO-Virgo-KAGRA with LIGO-India (HLVKI). As more detectors are added to the network, the average sensitivity to the second polarization increases. The right plot shows results for the Einstein Telescope (ET), which is comprised of three 60-degree interferometers, ET and three LIGO-Voyager detectors (Voyager-ET) and ET with either one or two Cosmic Explorer detectors (1CE-ET and 2CE-ET). As the ET detector has good sensitivity to both polarizations, networks where ET is the most sensitive detector will have large values of $\alpha$. Third generation target noise curves are taken from [1].

The distance $d_L$, inclination $\iota$, polarization $\psi$ and coalescence phase $\phi_0$ [198, 199]:

\[
A^1 = A_+ \cos 2\phi_0 \cos 2\psi - A_\times \sin 2\phi_0 \sin 2\psi \tag{3.4}
\]
\[
A^2 = A_+ \cos 2\phi_0 \sin 2\psi + A_\times \sin 2\phi_0 \cos 2\psi \tag{3.5}
\]
\[
A^3 = -A_+ \sin 2\phi_0 \cos 2\psi - A_\times \cos 2\phi_0 \sin 2\psi \tag{3.6}
\]
\[
A^4 = -A_+ \sin 2\phi_0 \sin 2\psi + A_\times \cos 2\phi_0 \cos 2\psi \tag{3.7}
\]

where $A_+$ and $A_\times$ are amplitudes for the plus and cross polarizations in the source frame, which is aligned with the binary’s orbital angular momentum. They are given
Chapter 3. Constraining the inclinations of binary mergers from gravitational-wave observations

by

\[ A_+ = \frac{d_0}{d_L} \frac{1 + \cos^2 \iota}{2}, \]  
\[ A_\times = \frac{d_0}{d_L} \cos \iota, \]  

where \( d_L \) is the luminosity distance and \( d_0 \) is the reference luminosity distance. The variation of the two polarization amplitudes with inclination \( \iota \) is shown in Figure 3.1.1. We note that there is an arbitrary choice of the radiation frame and this will affect the value of the angles \( \psi \) and \( \chi \) and consequently the values of the \( A^i \). However, the signal observed at the detectors is independent of this choice.

In principle, we should be able to measure all four of the amplitude parameters by accurately measuring both the amplitude and phase of both the plus and cross polarizations of a gravitational wave. From here, we could then infer the distance and orientation of the source binary. However, degeneracies limits our ability to accurately measure these parameters.

In order to identify the inclination of the binary system using the polarizations of the gravitational wave, we must distinguish the contributions of the plus and cross polarizations. When the binary system is near face-on or face-away, the two amplitudes \( A_+ \) and \( A_\times \) have nearly identical contributions to the overall gravitational-wave amplitude. In Figure 3.1.1, we see the relative difference between plus and cross is less than 1% for inclinations less than 30° (or greater than 150°) and 5% for inclinations less than 45° (or greater than 135°). This is the main factor that leads to the strong degeneracy in the measurement of the distance and inclination.

3.2.2 Network sensitivity

As we have already described, gravitational-wave detectors with limited sensitivity will preferentially observe signals which are close to face-on or face-off. In addition, when the binary is close to face-on and the emission is circularly polarized, the waveform is described by a single overall amplitude and phase (as the two polarizations are equal, up to a phase difference of ±90°). Thus it is no longer possible to measure both the polarization \( \psi \) and phase at coalescence \( \phi_0 \) of the binary, but only the combination \( \phi_0 \pm \psi \) (with the +/− for face-on/away binaries respectively). This degeneracy, combined with the distance prior, leads to a significantly larger volume of parameter-space which is consistent with face-on, rather than edge-on systems.

To exclude face-on binaries from a marginalized posterior probability distribution on the inclination, the network must accurately measure the amplitude and phase of both of the polarizations. In general, gravitational-wave detectors are not equally sensitive to the two polarizations. For a given sky location, we choose a plus and cross polarization such that the detector network is most sensitive to the plus polarization and the remaining orthogonal component becomes the cross polarization. We can
think of this as a detector network comprised of a long plus-detector and a shorter cross-detector (with a fraction $\alpha$ the length of the arms of the plus-detector). Thus we can estimate the proportional sensitivity to the second polarization, called the network alignment factor \cite{149}, through the relation $F_\times = \alpha F_+$, where $\alpha$ varies between 0 and 1. Therefore the sensitivity of the network to the second polarization can be determined by looking at the values of $\alpha$ over the sky.

Figure 3.2.1 shows the distribution of alphas for various detector networks. As might be expected, the sensitivity to the second polarization increases as more detectors are added to the network. For the two LIGO detectors, the typical value is $\alpha \sim 0.1$ because the two detectors have very similar orientations. When the Virgo detector is added to the network, the mode is $\alpha \sim 0.3$ and this increases to $\alpha \sim 0.5$ when KAGRA and LIGO-India join the network. The Einstein telescope (ET) is a proposed future detector with a triangular configuration \cite{196}. For an overhead source, ET is equally sensitive to both polarizations, giving $\alpha = 1$. While ET does not have equal sensitivity to both polarizations over the whole sky, the majority of signals will be observed with $\alpha > 0.9$. For the future networks, we consider an ET detector complemented by either the advanced LIGO detectors with sensitivity improved by around a factor of three (LIGO Voyager), or by one or two Cosmic Explorer (CE) detectors \cite{1, 46}. When the ET detector dominates the network’s sensitivity, we have excellent measurement of both polarizations but, in the CE-ET networks where CE is more sensitive, the relative sensitivity to the second polarization is comparable to the current networks.

### 3.2.3 Parameter estimation

When a gravitational-wave signal is observed in the data from the LIGO and Virgo instruments, the goal is to obtain estimates for the parameters that describe the waveform. Typically, Bayesian inference \cite{200, 201, 202} is used to obtain a posterior distribution for the parameters of the system $\theta$ given the observed data $d$. As described in detail in \cite{203}, the likelihood of obtaining data $d$ given the presence of a signal $h(\theta)$ and under the assumption of Gaussian noise characterized by a power spectrum $S(f)$ is

$$\Lambda(d|\theta) \propto \exp \left[ -\frac{1}{2} (d - h(\theta)) (d - h(\theta))^\dagger \right].$$

(3.10)

Here, we have introduced the weighted inner product

$$\langle a|b \rangle := 4 \text{Re} \int_0^{f_{\text{max}}} \frac{\tilde{a}(f) \tilde{b}(f)^\ast}{S(f)} df.$$  

(3.11)

The likelihood for a network of detectors is simply the product of likelihoods for
Chapter 3. Constraining the inclinations of binary mergers from gravitational-wave observations

Figure 3.2.2: The marginalized posterior distribution for the distance and inclination of the binary neutron star system GW170817, detected with a signal-to-noise ratio $\rho \sim 32$ and network alignment factor $\alpha \sim 0.13$. The left plot was generated using only the data from gravitational-wave detectors, while the right plot also uses the independent distance measurement ($40.7$ Mpc, $\pm 2.4$ Mpc at 90% confidence) from electromagnetic observations. The coloured portion of the plot shows the probability distribution obtained using our approximate analysis, normalized such that the peak probability is 1. The orange contours represent the 90% and the 50% confidence intervals obtained by performing the full analysis of the LIGO-Virgo data (posterior samples are publicly available here: https://dcc.ligo.org/LIGO-P1800061/public) [6].

The individual detectors:

$$
\Lambda(d|\theta) \propto \exp \left[ -\frac{1}{2} \sum_{i \in \text{dets}} (d_i - h_i(\theta))^2 \right].
$$

The posterior distribution for parameters $\theta$ given the data $d$ is given as

$$
p(\theta|d) \propto \Lambda(d|\theta)p(\theta),
$$

where $p(\theta)$ is the prior distribution for the parameters. The posterior distributions are typically calculated by performing a stochastic sampling of the distribution [204, 201, 202, 205, 206]. Distributions for a subset of parameters are obtained by marginalizing, or integrating out, the additional parameters.

In this analysis, we are interested in obtaining the joint distribution of the luminosity distance $d_L$ and binary inclination $i$. This is calculated as

$$
p(d_L, \cos i|d) = \int d\mu \Lambda(d|\mu, d_L, \cos i)p(\mu, d_L, \cos i)
$$

Typically, $\mu$ contains all parameters describing the system, including the masses, spins, sky location, orientation and parameters describing the nuclear equation of state. For our work, we consider a simplified model, for which the only additional
parameters contained in $\mu$ are the binary’s polarization $\psi$ and coalescence phase $\phi_0$. We choose uniform priors on these parameters, as well as a uniform prior on $\cos \iota$, which leads to a uniform distribution of binary orientation. Furthermore, we use a uniform-in-volume prior for the distance $p(d_L) \propto d_L^2$. For binaries at greater distance, we need to take into account cosmological effects and use a prior with sources uniform in comoving volume and merging at a constant local rate. At even greater distances, the local merger rate would follow the star formation rate [2], which peaks at redshift $z \sim 2$. We take this into account later in this chapter for binary black hole systems, (BBH), which can be detected throughout the universe with future detectors.

In our approximation, we fix the sky location and arrival time of the signal, as well as the masses and spins of the system. Fixing the sky location is reasonable, as one of the main motivations for this work is to investigate the accuracy of gravitational-wave measurements of distance and inclination after the signal has already been identified and localized by the detector network. We also investigate how inclination measurements from gravitational-wave observations can be combined with electromagnetic observations. An unknown sky location will only lead to larger uncertainties in the distance and inclination measurements arising from varying detector sensitivities over the sky.

While the masses and spins of the binary will not be known, in most cases these parameters have little impact on the inferred distance and inclination. This is especially true for low mass systems such as binary neutron stars which typically have component masses in the range $1.2 - 1.6M_\odot$ [154, 155]. These systems are in-band for a large number of cycles, $\mathcal{O}(10^4 - 10^6)$, allowing the accurate measurement of the phase evolution of the binary. In these cases the chirp mass $\mathcal{M}$ — the parameter determining the leading order phase evolution — is measured with great precision. Therefore, though $\mathcal{M}$ also appears in the amplitude, the uncertainty in $\mathcal{M}$ will be negligible relative to the total uncertainty in the amplitude and can be safely ignored for when considering uncertainties in measurement of the distance and inclination. For higher mass binaries such as BBH, typically fewer cycles of the waveform are visible $\mathcal{O}(10^1 - 10^3)$. This results in a less precise measurement of the binary chirp mass. This extra uncertainty in the overall amplitude will widen the posteriors on both the distance and inclination. The estimates here should thus be seen as best-case, and any extra uncertainty in mass and sky position will simply widen the posteriors. In the analysis presented here, we focus only on the dominant gravitational-wave emission at twice the orbital frequency. For unequal-mass systems, the other gravitational-wave harmonics can significantly affect the waveform, particularly when the binary has a high mass ratio, i.e. one of the compact objects is significantly more massive than the other [207]. This can lead to improvements in the measurement of the binary orientation [20].

Spins which are misaligned with the orbital angular momentum lead to precession...
Chapter 3. Constraining the inclinations of binary mergers from gravitational-wave observations

Figure 3.2.3: The progression of the probability distributions over a $\cos \iota$ and distance parameter space for a signal detected with alignment factor $\alpha = 0.1$ and signal to noise ratio $\rho = 12$. The white star represents the injected signal. The top panel shows the distribution for a face-on signal. The bottom panel shows the distribution for an edge-on signal. The leftmost plots are the distribution for only the likelihood. This is generated by calculating the SNR fall-off over the parameter space. Since we have not yet marginalized over the phase $\phi_0$ and polarization $\psi$, the orientation angles are set equal to the values used in the injection, in this case zero for both $\phi_0$ and $\psi$. The middle plots show how these distribution change when marginalizing over $\psi$ and $\phi_0$. Lastly, the rightmost plots are the complete probability distribution, calculated by applying a distance-squared weighting to the likelihood. This is to account for the expectation that binary systems are distributed uniformly in volume. Recall that $\alpha = 0.1$ is the mode sensitivity for the Hanford-Livingston network.

The white star represents the hypothetical signal. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. From these plots, we can see that at this $\alpha$, a side-on signal is indistinguishable from a face-on/face-away signal.

of the binary orbit [208] which can, in principle, lead to an improved measurement of the binary orientation. To date, there is no evidence for precession in the observed GW signals [54, 209, 210, 211, 212], so the approximations discussed here would therefore be applicable. Furthermore, neutron stars observed to merge within a Hubble time [213, 214] are not expected to achieve a spin high enough to have observable precession within ground-based detector bandwidths [208, 6].

To verify that fixing the masses and spins has limited impact on the recovered distance and inclination, we compare results from our model with those from the full parameter estimation of GW170817. We recreate the posterior distribution for the multi-messenger signal GW170817, with and without distance information from the coincident electromagnetic signal, and compare it to the full, Bayesian parameter estimation, with a fixed sky location, using the observed LIGO and Virgo data [6].
Figure 3.2.4: The probability distribution over a $\cos \iota$ and distance parameter space for a signal detected with alignment factor $\alpha = 0.5$ and signal to noise ratio $\rho = 12$. The white star represents the injected signal. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. A face-on signal (where $\cos \iota = 1$) returns a nearly identical probability distribution of the parameter space as a signal from a binary with an inclination of about 66 degrees ($\cos \iota = 0.4$). For inclinations in the range $0.1 < \cos \iota < 0.4$, though the distribution now peaks at the correct inclination, there is support extending across from face-on to an inclination of $\iota \sim 80^\circ - 90^\circ$. In these cases it is not possible to distinguish the binary inclination. The signal is only clearly identified as not face-on after $\cos \iota < 0.1$.

The results in Figure 3.2.2 show an example of the degeneracy in the measured values of the distance and binary inclination. The 50% confidence interval includes both a face-away binary at a distance of 45 Mpc and a binary inclined at 135$^\circ$ at a distance of 35 Mpc. It is only when the gravitational-wave data is combined with the electromagnetically determined distance $40.7 \pm 2.4$ Mpc [181] that the binary inclination can be accurately inferred.

\[\text{We note that the results in [6] show this distribution as a function of inclination } \iota \text{ instead of } \cos \iota. \text{ This leads to a different distribution, and different 90% confidence intervals as these are defined to be the minimum range that contains 90% of the probability, and this is dependent upon variable choice. As we discuss later, there is no evidence in the GW data alone that the signal is not face-off, and since the prior is flat in } \cos \iota \text{ we believe that plotting the distribution against } \cos \iota \text{ leads to a clearer understanding of the distribution.}\]
3.3 Accuracy of measuring distance and inclination

Now that we understand how the degeneracy between inclination and distance arises, we can explore the expected accuracy with which these parameters will be measured in various gravitational-wave detector networks. For concreteness, in the examples that follow, we fix the signal-to-noise ratio (SNR), denoted ρ, of the signals to be 12. While this might seem low, we note that for a detection threshold of 8, the mean SNR observed from a uniform-in-volume population would be 12 [119]. We discuss higher SNR signals later in the chapter. Rather than specifying a network and sky location, we instead investigate the ability to measure distance and inclination as we vary the network’s relative sensitivity to the second polarization, encoded in the variable α. For convenience, we fix the masses of the system to be $1.4M_\odot$ and set the sensitivity of detector network to the plus polarization of GW to be equal to that of a single advanced LIGO detector at design sensitivity for an overhead source. This places a face-on system at approximately 300 Mpc at SNR of 12. For inclined systems, the distance will be smaller to ensure that the network still receives an SNR of 12. While we have fixed the masses and detector sensitivities to make the plots, the results are essentially independent of these choices, up to an overall rescaling of the distance. Thus the results will be applicable to any system for which it is reasonable to neglect precession effects and the impact of higher modes in the gravitational waveform.

Let us begin by considering a network with relatively poor sensitivity to the second GW polarization, with $F_x = 0.1F_\perp$. This is typical for the LIGO Hanford-Livingston network and is common for the LIGO-Virgo network, as described in Figure 3.2.1. We consider two signals, both with SNR of 12, but one which is face-on ($\iota = 0$) at a distance of 300 Mpc while the second is edge-on ($\iota = 90^\circ$) at a distance of 150 Mpc and a polarization angle of $\psi = 0$ so that the GW power is contained in the plus polarization. The first column of figures in Figure 3.2.3 shows the likelihood across the distance-inclination plane. Here, we simply kept the values of $\phi_0$ and $\psi$ equal to zero, the value used in generating the signal. Note that the contours here are calculated for our simplified model and do not represent the results of full parameter estimation analyses, as they did in Figure 3.2.2. As expected, the maximum likelihood occurs at values of distance and inclination which exactly match the signal. We observe a degeneracy in distance and inclination, so that there is some support for the edge-on binary to be face-on (or face-away). There is also degeneracy for the face-on binary, which is marginally consistent with an edge-on binary, but face-away orientation can be excluded. With an SNR of 12 and $\alpha = 0.1$, for a face-on signal we expect an SNR of about 1.2 in the cross polarization. These results show that the presence or absence of this signal is sufficient to down-weight, but not exclude, an edge-on orientation when the source is really face-on, and vice-versa. For a face-away system, the expected signal in the cross polarization is the same
3.3. Accuracy of measuring distance and inclination

amplitude, but entirely out of phase from the face-on system, and this is sufficient to distinguish the two.

In the second column, we show the likelihood marginalized over the polarization and phase angles. This marginalization does not have a significant impact on the face-on binary, but completely changes the distribution for the edge-on binary — with the marginal likelihood now peaked at \( \cos \iota = \pm 1 \). Typically, we would expect to be able to measure the two phase angles with accuracy \( \sim 1/\rho \) thus to a crude approximation, marginalizing over the phase angles would give a contribution \( \approx (1/\rho^2)\Lambda_{\text{max}} \), where \( \Lambda_{\text{max}} \) is the maximum likelihood. When the binary is recovered (nearly) face-on the two amplitudes \( A_{+,x} \) are (nearly) equal. Consequently, the signal is circularly polarized, with the phase determined by \( \phi_0 + \psi \). Changing the value of \( \phi_0 - \psi \) has no effect on the waveform. Thus, when marginalizing over the polarization and phase, we obtain a factor \( \sim (\pi/\rho)\Lambda_{\text{max}} \). Thus, for this signal at SNR 12, marginalizing of the polarization and phase will lead to a relative increase of nearly 40 in favour of the face-on signal.

Finally, in the third column, we include the distance prior by re-weighting by \( d_L^2 \) to place sources uniformly in volume. This gives an additional factor of four weighting in favour of the face-on signal over the edge-on one. Once all these weightings are taken into account, the probability distributions between a face-on and edge-on signal are similar for a network with this sensitivity. The edge-on signal has slightly more support at \( \cos \iota \approx 0 \), and this is still included at 90\% confidence. Additionally, the edge-on signal is consistent with either a face-on or face-away orientation. It may seem strange that we will not recover the parameters of the edge-on system accurately. However, this is appropriate. As we have discussed, the volume of parameter space consistent with a face-on system is significantly larger than for the edge-on case. Thus, even if we observe a signal that is entirely consistent with an edge-on system, it is more likely that this is due to a face-on system and noise fluctuations leading to the observed signal than it is that the signal is coming from an edge-on system. We note that this effect is also seen in [194]. Their figures 4–6 show two LIGO detectors recovering edge on systems as face-on or face-away, which results in the wrong sky location being measured.

Our next example investigates differing inclinations for a signal detected by a network with an \( F_\times = 0.5F_+ \), i.e. a network with half the sensitivity to the cross polarization as the plus polarization. This is the predicted mean sensitivity expected for the best near-future detector network consisting of the Hanford, Livingston, Virgo, KAGRA and LIGO-India detectors. Again, the SNR is set to 12 for all hypothetical signals, and now we consider three different inclinations: \( \iota = 0 \) (face-on) and two inclined signals, one with \( \iota = 66^\circ \) and the other with \( \iota = 78^\circ \). In Figure 3.2.4, we show the posterior distribution for distance and inclination for the three cases. Here, we have marginalized over the phase angles and included the distance prior weighting, so the plots are equivalent to the third column of plots in Figure
Figure 3.3.1: The un-normalized marginalized posterior for $\cos \iota$ for a face-on source as measured for three networks with alignment factors $\alpha = 0.1$, $\alpha = 0.5$, $\alpha = 1.0$ and signal to noise ratio $\rho = 12$, $\rho = 12$, $\rho = 50$ respectively. The solid line shows the expected $\cos^3 \iota$ form of the likelihood.
3.3. Accuracy of measuring distance and inclination

Figure 3.3.2: This plot shows a detector network’s ability to constrain the inclination of a face-on signal with 90% confidence. The x-axis shows the network alignment factor $\alpha$, whereas the y-axis shows the signal-to-noise ratio (SNR) of the hypothetical gravitational-wave signal. The colour represents the upper limit on the inclination angle. For weak signals or for networks which are not very sensitive to the cross polarization, the network can only constrain the inclination to being less than about 45$^\circ$. Even for the most sensitive detector network detecting the loudest hypothetical signals, the network would be unable to constrain the inclination to being less than 30$^\circ$. However, we note that at these SNRs, the detector network may be able to identify higher order modes, which would break the degeneracy between distance and inclination, allowing more precise measurement of the inclination.
3.2.3.

The leftmost plot shows the probability distribution for a face-on signal. This distribution is similar to the one for $\alpha = 0.1$, though now the most inclined and face-away points in parameter space are excluded from the 90% credible region. The second plot is for a binary inclined at $66^\circ$ ($\cos \iota = 0.4$). Here, the peak of the inclination distribution corresponds to a face-on system and, indeed, the posterior is nearly identical to that obtained for the face-on system. Thus, for a typical system with close-to-threshold SNR we will remain unable to distinguish between face-on signals and those inclined at $60^\circ$ based on gravitational-wave observations alone. The best near-future detector therefore would be unable to measure a difference in inclination between these two hypothetical signals. For inclinations in the range $0.1 < \cos \iota < 0.4$ the 90% credible region includes peaks at both the true inclination and face-on. In this region, we cannot tell if the difference in polarization amplitudes arises from an inclined signal or from a signal with parameters in the much larger, and prior-preferred, volume of parameter space near face-on, with a noise fluctuation giving rise to the observed data. For values of $\cos \iota < 0.1$, the posterior is peaked at the correct value of $\iota$ and excludes face-on from the 90% credible region.

The results shown in Figures 3.2.3 and 3.2.4 show the general features of the distance and inclination distribution. It is characterized by three components: one consistent with a face-on signal, one with a face-off signal and a third contribution peaked around the true values of distance and inclination. In all of the cases we have shown, only one or two of the contributions are significant. There are, however, cases where we obtain three distinct peaks in the posterior for the inclination, although these are rare. In Appendix B of [190], an approximate expression for probability associated with each peak was obtained, which is valid for networks sensitive to a range where a $d_L^2$ prior is still appropriate. This provides an analytic expression for the probability associated to each of the three contributions, as a function of SNR, inclination, polarization and the network sensitivity to the second polarization, encoded in the variable $\alpha$.

To get a sense of how accurately binary inclination will be measured, we simulated a set of 1,000,000 events isotropically and uniformly in both volume and orientation, keeping only those which would be observed above the detection threshold of the network (typically leaving 30,000-80,000 events). For each event, we then determine whether the event would be recovered as definitely face-on — over 90% of the probability associated to the face-on (and face-away) components of the distribution — definitely inclined or uncertain. These results are summarized in Table 3.3.1, for a series of networks each with an increasing number of detectors. For all networks, essentially all events with a true inclination less than $45^\circ$ will be recovered face-on. Only for those events with inclination greater than $45^\circ$ do we start to be able to distinguish the orientation. Between $45$ and $60^\circ$, networks with three or more detectors will classify a small fraction of events as inclined, and this fraction increases
3.3. Accuracy of measuring distance and inclination

Table 3.3.1: The table shows the ability of various networks to distinguish the orientation of a population of binary mergers with given inclination, $\iota$. For each network and range of $\iota$, we give the percentage of binaries for which the posterior on the inclination peaks at $\iota = 0$ or $180^\circ$ (face-on) and this peak contains over 90% of the probability; those binaries for which the recovered inclination peaks at the correct value, and greater than 90% of the probability is consistent with this peak (inclined); and those for which the posterior includes significant contributions for both face-on and inclined orientations (uncertain). For all networks, essentially all binaries with $\iota < 45^\circ$ will be recovered face-on. As the inclination increases further, the ability to clearly identify the binary as inclined increases significantly with the number of detectors in the network as this improves the average sensitivity to the second gravitational-wave polarization.

with both the inclination of the system and the number of detectors (which directly effects the typical value of $\alpha$). However, even for events which have an inclination greater than $75^\circ$, the LIGO Hanford–Livingston network would recover half as face-on and only 20% as definitely not. This improves for the five detector network where less than 10% are face-on and 80% are clearly identified as being inclined. We note that similar results have been obtained independently in [189].

Next, let us consider the general accuracy with which we can measure the inclination for a binary which is (nearly) face-on. In this case, the distribution for the inclination angle can be approximated in a simple way. If we begin by assuming that the degeneracy between distance and inclination is exact, then orientations with $|\cos \iota| \approx 1$ are preferred due to the prior on the distance. This can be clearly seen by comparing the second and third columns of plots in Figure 3.2.3. The distribution in the second column (when we do not apply the uniform-in-distance weighting) shows a broad degeneracy with equal probability along lines of constant $A = \cos \iota/d_L$. It is only by applying the distance re-weighting that the peak shifts more to $\cos \iota = 1$.

We can obtain the posterior probability for a fixed value of $\iota$ by integrating over a given distribution, $p(\cos \iota/d_L)$:

$$p(\cos \iota) = \int d_L^2 p(\cos \iota/d_L) dd_L$$

$$\propto \int \cos^3 \iota A^{-4} p(A) dA$$

$$\propto \cos^3 \iota$$ (3.15)

Thus, it follows that, where the degeneracy holds, the posterior on $\cos \iota$ will be
Chapter 3. Constrained the inclinations of binary mergers from gravitational-wave observations

Figure 3.3.3: Marginalized posterior distribution for a $10M_\odot - 10M_\odot$ binary black hole at redshift $z = 10$ detected by the Einstein Telescope in the most sensitive part of the sky, i.e. directly above the detector. Here, the alignment factor is $\alpha = 1$ and the signal-to-noise ratio is $\rho = 20$. The white star represents the injected signal at three different inclinations: $\iota = 66^\circ$, $\iota = 60^\circ$ and $\iota = 0^\circ$. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. Note that these contours do not represent the results of full parameter estimation, as they did in Figure 3.2.2. We use a prior that is a uniform in comoving volume with a rest frame rate density that follows the star formation rate [2]. At this redshift the prior varies by a factor of $\sim 12$ across the degeneracy and now favours more inclined binaries. Thus binaries that are face-on will be recovered as being more inclined. The redshift uncertainty $\Delta z/z \sim 40\%$ dominates the statistical error in the recovery of the binary chirp mass. All conversions between luminosity distance and redshift assume standard cosmological parameters [7].

proportional to $\cos^3 \iota$. In Figure 3.3.1, we show the posterior for three examples of face-on signals: SNR $\rho = 12$ with $\alpha = 0.1$ and 0.5, and SNR $\rho = 50$ with $\alpha = 1$. All three distributions follow the $\cos^3 \iota$ distribution for small inclinations. The high-SNR signal deviates at around $30^\circ$ — at this inclination there is enough difference from a circularly polarized signal for larger inclinations to be disfavoured. However, for the lower-SNR signals (and also lower values of $\alpha$) the approximation remains accurate to greater than $45^\circ$.

We can improve the approximation by noting [190] that the SNR lost by projecting an inclined signal onto a circular signal is

$$\Delta \rho^2 = \frac{\alpha^2 \rho^2}{(1 + \alpha^2)^2} \frac{(1 - \cos \iota)^4}{4}. \quad (3.16)$$

This loss in SNR leads to a reduction in the likelihood associated with the inclined signal, which causes the probability distribution to fall off more rapidly away from $\iota = 1$. In particular, we obtain

$$p(\cos \iota) \propto \cos^3 \iota \exp \left( -\frac{\Delta \rho^2}{2} \right). \quad (3.17)$$

We can use this expression to determine how well a network with sensitivity $\alpha$ would be able to constrain a signal’s inclination $\iota$ given the SNR of the signal. In Figure 3.3.2, we specifically look at how tightly we can constrain a face-on signal.
We can see that for low-SNR signals or for networks with little sensitivity to the cross polarization, GW observations will only be able to constrain the signal to being less than about 45°. Even with an extremely loud signal and a very sensitive detector network, we are only able to constrain the signal to about 30°. This effect has also been observed in the results of full parameter estimation runs in [195] where even with a network of two ETs and for sources nearby (at redshifts z < 3), the median 90% credible interval for the cosine of the inclination was found to be $\Delta \cos \iota \sim 0.15$, which at face-on corresponds to an inclination constraint of 30°. It is important to note here that at these SNRs, higher-order modes or precession in the gravitational-wave signal may be observable. If these are detected, the degeneracy between distance and inclination would be broken, and we would be able to more tightly constrain the inclination.

Finally, it is interesting to consider what effect the inclination distance degeneracy would have on the mass estimate of binary black holes. GW detectors actually measure the redshifted mass $M_{\text{det}} = (1 + z)M_{\text{source}}$ where the subscripts denote detector-frame and source-frame respectively [193]. There is no way to determine the redshift directly from the gravitational waveform of a binary black hole. However, the measured value of the luminosity distance can give the redshift if a cosmology is assumed. In this way, the inclination-distance degeneracy will map to an uncertainty in the rest-frame masses. For the next generation of gravitational-wave detectors which will be sensitive to BBH mergers throughout the universe, the uncertainty in the redshift will likely be the dominant uncertainty in the masses. As such, we explore the inclination measurement with ET for a BBH merger at a redshift of $z = 10$ with intrinsic masses of $10M_\odot$ each, corresponding to a detector frame chirp mass of $M_{\text{det}} = 96M_\odot$. We place the source directly above the detector, in the most sensitive part of the sky. In this case, $\alpha = 1$ and $\rho = 20$, where we have assumed standard cosmology [7].

At these cosmological distances, a $d_L^2$ prior for the distance is no longer appropriate. Rather, we use a distance prior that is uniform in comoving volume where the rest-frame binary merger rate density follows the cosmic star formation rate [2] with a delay between star formation and binary merger $\Delta t$, and a distribution of delay times $p(\Delta t) \propto 1/\Delta t$ [3] (see Section 5 of [46] for details). The new prior peaks at $z \sim 1.4$. Therefore at $z \sim 10$, the nearer, more inclined binaries are a priori more likely.

In Figure 3.3.3, we show the marginalized posterior for three different inclinations: $\iota = 66^\circ$, $\iota = 60^\circ$ and $\iota = 0^\circ$. For the second generation networks in Figure 3.2.4, the $\iota = 66^\circ$ ($\cos \iota = 0.4$) source is recovered as face-on. With the higher signal to noise ratio and improved sensitivity to the second polarization, ET can identify the signal as edge-on. At an inclination of $\iota = 60^\circ$, the degeneracy still extends across $25^\circ < \iota < 70^\circ$, though smaller inclinations are now excluded from the 90% credible interval. This is the effect of the new distance prior which is a factor of 12
larger at redshift 6 than at redshift 10. Thus, though the 90% credible region of the marginalized likelihood extends right up to face-on, the prior is able to partially break the degeneracy. For less inclined binaries $\iota < 60^\circ$, the 90% probability interval extends up to face-on.

For the face-on binary in the rightmost plot, the prior shifts the peak of the posterior away from the true value. Although the value of the likelihood at face-on and redshift 10 is a factor of 12 larger than it is at an inclination of $60^\circ$ and redshift 6, these two points in the parameter space are equally likely after the prior re-weighting. If the detector frame chirp mass of the binary is measured to be $M_{\text{det}} = 96M_\odot$, the degeneracy between the inclination and distance results in $M_{\text{source}} = 96M_\odot$ and $M_{\text{source}} = 61M_\odot$ being equally likely. The detector-frame chirp mass $M_{\text{det}}$ would be determined to an accuracy similar to the accuracy of the GW phase measurement $\Delta M_{\text{det}}/M_{\text{det}} \sim 1/(\rho N_{\text{cycles}})$ [215, 180]. Parameter estimation for GW150914 yielded a precision in the detector-frame mass estimate of $\Delta M_{\text{det}}/M_{\text{det}} \sim 10\%$ for a comparable SNR [216]. Everything else being equal, GW150914 with a detector frame chirp mass of $M_{\text{det}} \sim 30M_\odot$ will have more cycles than the ET binary with detector-frame chirp mass $M_{\text{det}} = 96M_\odot$ above say 20Hz. However ET’s improved sensitivity at low frequencies compared to LIGO means that we can expect the precision of the detector-frame mass estimate of GW150914 and the ET binary to be roughly the same. Thus the broad uncertainty in the intrinsic masses due to the distance inclination degeneracy $\Delta M_{\text{source}}/M_{\text{source}} \sim 40\%$ will dominate the total error budget.

3.4 Conclusion and Future Work

Our work demonstrates that even with a network equally sensitive to both polarizations of the gravitational wave, we would be unable to precisely measure the inclination or distance of a nearly face-on binary due to a strong degeneracy between distance and inclination. However, we have focused on non-spinning binaries and assume that the sky location, masses and arrival times of the detectors are all known. Introducing these parameters would increase the uncertainties. Exploring how these parameters affect the overall measurement of the distance and inclination could give a more accurate summary of a gravitational-wave network’s ability to measure distance and inclination.

The degeneracy between inclination and distance described here could be broken in a few different ways: by using distance or inclination from electromagnetic measurements, by detecting higher order modes [20] and by measuring precession [217]. Binary neutron star systems produce a variety of EM signatures, as were observed for GW170817 [174]. Neutron star-black hole binaries (NSBH) could produce EM signatures should the neutron star be tidally disrupted. However, tidal disruption only happens at relatively small mass ratios [218]. For larger mass ra-
tios, the neutron star falls into the black hole before tidal disruption can produce EM signatures. Interestingly, both precession and higher modes have a larger effect on the gravitational waveform at higher mass ratios [219, 220]. The polarizations of the higher modes have a different dependence on the inclination, and the precession of the orbital plane would result in changing amplitudes for the plus and cross polarizations. These effects can make it easier to identify the inclination angle [20, 220, 221, 222, 217]. For NSBH, the degeneracy can thus be broken by either information from the EM emission or from higher modes or precession. [217] demonstrated that precession would break the distance-inclination degeneracy in NSBH for a few binaries with a few values of the precession angle and large, highly spinning black holes. It would be an interesting follow up to this study to explore this with a realistic distribution of spins, to see when precession plays a significant role in measuring binary parameters.
Chapter 4

Measuring gravitational-wave higher order multipoles

Code: The code that produced the main results in this chapter will be made publicly available here: https://github.com/millsjc/eventhm

Collaboration: Published in Ref. [10]. CM produced most of the results and writing. SF contributed to the writing, conception and presentation throughout.

4.1 Introduction

Gravitational waves emitted during the coalescence of black hole and/or neutron star binaries are emitted predominantly at twice the orbital frequency, during the inspiral phase of the coalescence [223]. However, it is also well-known that the gravitational wave signal cannot be completely characterized by a single harmonic but, rather, is better decomposed as a sum of spin-weighed spherical [224, 225] (or spheroidal [226, 227, 228, 229, 230]) harmonics. The dominant harmonic is the $(\ell, |m|) = (2, 2)$ harmonic, but there is also power in higher harmonics, most notably the $(\ell, |m|) = (2, 1), (3, 2), (3, 3)$ and $(4, 4)$ harmonics \(^1\) [231, 232]. The importance of these additional harmonics increases as the mass ratio between the two black holes increases and also increases for observed signals where the late inspiral and merger of the objects contribute significantly to the SNR [233, 234, 235, 236, 237, 238, 232, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250]. Recent semi-analytical and numerical relativity models have provided expressions for an increasing number of the higher harmonics accurate across the inspiral, merger and ringdown regimes [251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 230, 261, 262, 263].

Clear evidence of higher gravitational-wave harmonics has been observed in two recent observations, GW190412 [9] and GW190814 [12], as well as weaker evidence in the high-mass system GW170729 [264]. These observations provide further evidence that Einstein’s general relativity is an accurate description of gravity, including in

\(^1\)When we refer to a multipole by the label $(\ell, m)$ we always mean $(\ell, |m|)$. 

- 51 -
the strong-field, highly dynamic regime of the merger of two black holes [231, 265]. By incorporating knowledge of the higher harmonics into a search for gravitational waves, the sensitivity of gravitational wave searches can be improved, leading to an increase in the rate of observed systems [245]; furthermore these observations would typically be from less densely populated regions of the parameter space [266], for example high mass binaries and those with unequal mass components. Finally, the observation of higher harmonics enables more accurate measurement of the properties of system [267, 9, 12]. For example, the measurement of multiple harmonics can be used to break well-known degeneracies between the measured distance and orientation of the system [173], or the mass ratio and spins of the black holes [268, 269].

The observation of other features in the gravitational waveform, most notably spin-induced orbital precession in black-hole binaries [270, 271] and matter effects in binaries containing neutron stars [272, 273], will also enable more accurate measurements of the source properties. Measurements of neutron-star structure have been inferred from GW170817 [33, 34], while there is some evidence for precession in the black hole binary mergers GW190521 [14] and GW190412 [9]. The inclusion of matter effects will not affect the applicability of the spin-weighted spherical harmonic decomposition discussed above, and we do not consider these effects in this chapter.

Orbital precession does change the structure of the emitted gravitational waveform. The waveform can still be decomposed into a basis of spin-weighted spherical harmonics, but now each of the harmonics shows the characteristic amplitude and phase modulations associated with time-dependent evolution of the orbital plane [254]. Alternatively, this can be understood as the splitting of each gravitational wave multipole into several precession-induced harmonics, which are offset by multiples of the precession frequency [274, 254, 275]. Thus, a full analysis of the observability of higher multipole moments requires a detailed treatment of precession. However, for the majority of events observed to date [16], there is little, if any, evidence for precession. In this chapter, we restrict attention to aligned-spin systems, which do not exhibit precession. This will provide good insight into the significance of higher gravitational wave multipoles in the observed population.

While the gravitational waveform is comprised of an infinite number of harmonics, it is the unambiguous measurement of a second harmonic, in addition to the \((\ell, |m|) = (2, 2)\) harmonic, which will lead to a step-change in our ability to measure the properties of the system; additional harmonics will then further refine the measurement accuracy. In this chapter, we perform an in-depth investigation of the importance of the higher harmonics across the parameter space and identify regions of the parameter space where particular harmonics are most likely to make a significant contribution. The amplitude of each harmonic depends both upon the intrinsic parameters of the system (its masses and spins, both magnitudes and orientations) as well as the extrinsic parameters (the orientation of the binary and the detector
network’s sensitivity to the two polarizations of gravitational waves). For simplicity, we decompose the harmonics into an overall amplitude factor, dependent only upon the intrinsic parameters, and an orientation dependent term. We then investigate the significance of each harmonic across the parameter space.

Next, we turn to the question of when additional harmonics have been unambiguously observed. From a model selection perspective, this can be addressed by considering the evidence in favour of a waveform containing higher harmonics against one without. Here, we introduce the higher-harmonic signal to noise ratio, and argue that it can be used as an alternative method of establishing the observability of higher harmonics. \(^2\) It is straightforward to calculate the SNR contained in each of the higher waveform harmonics, and compare to the expectation due to noise-only in the higher harmonics. This approach has been used to verify the observation of higher harmonics from the binary mergers observed as GW190412 and GW190814 [9, 12].

The structure of the chapter is as follows. In section 4.2, we provide a brief review of the gravitational waveform, incorporating the higher harmonics, and use this to fix the notation for the remainder of the chapter; in section 4.3 we explore the significance of the higher harmonics over the parameter space, both intrinsic (masses and spins) and extrinsic (binary orientation); in section 4.4 we investigate the observability of higher harmonics and introduce a simple criterion for detection; finally in section 4.5 we investigate observability for a population of events.

### 4.2 The Gravitational Waveform

The measured gravitational wave strain \(h\) can be written as

\[
h = F_+ h_+ + F_\times h_\times, \tag{4.1}\]

where the antenna factors \(F_+\) and \(F_\times\) depend upon the sky location (right-ascension and declination) of the source, as well as the polarization of the source. It is often convenient to explicitly extract the unknown polarization angle \(\psi\) and then consider the detector response to be a known quantity dependent upon only the details of the detector and the direction to the source. Thus, we write the detector response as,

\[
\begin{align*}
F_+ &= w_+ \cos 2\psi + w_\times \sin 2\psi, \\
F_\times &= -w_+ \sin 2\psi + w_\times \cos 2\psi, \tag{4.2}
\end{align*}
\]

where \(w_+\) and \(w_\times\) are the detector response functions in a fixed frame — for a single detector it is natural to choose \(w_\times = 0\) and \(w_+ = 1\), and for a network to work in

\(^2\)A similar prescription has recently been introduced for precessing systems [276, 275].
4.2. The Gravitational Waveform

the dominant polarization, in which for each sky point the polarization angle $\chi$ is chosen to maximize the network sensitivity to $w_\mp$ [277, 278]. The relative amplitude of $w_\times$ to $w_+$ describes the sensitivity of the network to the second gravitational wave polarization. The unknown polarization of the source relative to this preferred frame is denoted $\psi$.

The radiation-frame gravitational wave polarizations $h_+$ and $h_\times$ can be decomposed into multipole moments using spin-weighted spherical harmonics of spin weight $-2$, $-2Y_{\ell m}$, which are functions of the inclination angle $\iota$ and a reference phase $\phi_0$ (see Appendix 4.A for a more detailed discussion of the decomposition).

For binaries with aligned spins, the orbital plane remains unchanged during the merger, and this provides a natural fixed basis for the spherical harmonic decomposition. However, if spin-induced precession is present, the orbital plane changes during the course of the merger (equivalently, the inclination angle becomes time dependent). In this case, it is natural to consider the waveform in a co-precessing frame, i.e. a frame which is locked to the binary’s orbit [274, 279]. While it is still possible to decompose as a series of spin-weighted spherical harmonics, each multipole moment of the waveform in the observer’s frame involves a sum over multiple harmonics in the co-precessing frame [254]. In this chapter, we restrict attention to binaries which do not exhibit precession. While we do expect generic black hole binaries to have spins which are mis-aligned with the orbital angular momentum (and hence will precess), in the majority of observations to date [16] there has been little evidence for precession and low black hole spins are inferred from current observations. Furthermore, as discussed in [275], the $(2, 2)$ multipole moment of a precessing waveform can be decomposed into five precession harmonics with the two leading harmonics providing the majority of the signal power. A similar decomposition is possible for the higher $(\ell, |m|)$ multipoles and, over much of the parameter space, it is only the leading precession harmonic of the higher $(\ell, |m|)$ multipoles that will provide significant power. Consequently, for many signals with non-zero in-plane spins, our analysis will remain valid. We leave the detailed examination of the interplay between precession and higher harmonics to future investigations.

For a binary merger which does not exhibit precession, the waveform can be expressed in the frequency domain, using the stationary-phase approximation, as

$$
\tilde{h}_+(f) = \frac{d_o}{d_L} \sum_{\ell \geq 2} \sum_{m = 0}^{\ell} A^{\ell m}_+(i) e^{i m \phi_0} \tilde{h}_{\ell m}(f)
$$

$$
\tilde{h}_\times(f) = \frac{d_o}{d_L} \sum_{\ell \geq 2} \sum_{m = 0}^{\ell} A^{\ell m}_\times(i) e^{i m \phi_0} \tilde{h}_{\ell m}(f)
$$

where $d_L$ is the luminosity distance, $d_o$ is a fiducial distance used to normalize the waveforms $\tilde{h}_{\ell m}$. In this chapter, we use the IMRPhenomHM model for $\tilde{h}_{\ell m}$, defined in Eq. (2) of Ref. [252]. The amplitude factors $A^{\ell m}$ are functions only of the
Chapter 4. Measuring gravitational-wave higher order multipoles

Figure 4.2.1: The absolute amplitude factors $A_{lm}$ (see Eq. 4.4) of the (2, 2), (2, 1), (3, 3), (3, 2) and (4, 4) harmonics as a function of the inclination $\iota$ of the binary. The (2, 2) multipole moment is normalized to unity at $\iota = 0^\circ$ while other multipoles are normalized to unit amplitude in the $+$ polarization at $\iota = 90^\circ$. Left: $+$ polarization Right: $\times$ polarization.

The inclination angle and are given below for the most significant harmonics:

\begin{align*}
A_{22}^+ &= \frac{1}{2}(1 + \cos^2 \iota) \\
A_{22}^\times &= \cos \iota \\
A_{21}^+ &= \sin \iota \\
A_{21}^\times &= \sin \iota \cos \iota \\
A_{33}^+ &= \sin \iota(1 + \cos^2 \iota) \\
A_{33}^\times &= 2 \sin \iota \cos \iota \\
A_{32}^+ &= 1 - 2 \cos^2 \iota \\
A_{32}^\times &= \frac{1}{2}(\cos \iota - 3 \cos^3 \iota) \\
A_{44}^+ &= \sin^2 \iota(1 + \cos^2 \iota) \\
A_{44}^\times &= 2 \sin^2 \iota \cos \iota
\end{align*}

There is a freedom in choice of overall normalization for these amplitude factors, which corresponds to an overall rescaling of the waveform defining each harmonic, $\tilde{h}_{\ell m}$. For the $(\ell, |m|) = (2, 2)$ multipole moment, it is customary to choose a normalization such that $A_{22}^+ = A_{22}^\times = 1$ for a face-on system, and we use that normalization here. Since many of the higher harmonics vanish for face-on systems, we instead choose a normalization for the higher-harmonic amplitude factors, $A_{\ell m}^{+\times}$, in Eq. (4.4), by requiring that for the plus polarization $A_{\ell m}^{+} = 1$ at $\iota = \frac{\pi}{2}$, i.e. when the system is edge on 3.

Fig. 4.2.1 shows the dependence of the multipole moments on inclination. The

\[3\]The normalization obtained by directly expanding the spherical harmonics from Eq. (4.27) in Eq. (4.31) differs by multiplicative factors of $\frac{1}{4}\sqrt{\frac{5}{\pi}}, \frac{1}{4}\sqrt{\frac{5}{\pi}}, \frac{1}{8}\sqrt{\frac{21}{\pi}}, \frac{1}{8}\sqrt{\frac{21}{\pi}}$, and $\frac{1}{16}\sqrt{\frac{7}{\pi}}$ for the (2,2), (2,1), (3,3), (3,2) and (4,4) multipoles respectively. The calculation is presented in Appendix 4.A.
plus polarization of the (2, 2) harmonic peaks at face-on, while the (2, 1) and (4, 4) harmonics peak at edge-on. The (3, 2) amplitude factor is maximum at both face-on and edge-on orientations while the (3, 3) harmonic peaks at \( \sin \iota = \sqrt{3} \). The different dependence of the harmonics on the binary orientation can lead to the improved measurement of the inclination, when more than one harmonic is observed [9, 267], breaking the well-known degeneracy between distance and inclination angle that arises when observing only the dominant harmonic [173].

During inspiral, the frequency evolution of a multipole, \( \omega_{\ell m} \), is related to the orbital frequency, \( \omega_{\text{orb}} \), as \( \omega_{\ell m} \sim m \omega_{\text{orb}} \) [231]. Phenomenologically, it has been qualitatively observed that during the ringdown the frequency approximately evolves as \( \omega_{\ell m} \sim l \omega_{\text{orb}} \) [227, 280]. Thus it is possible to scale the frequencies of the (2, 2) multipole moment in quite a simple manner to obtain an approximate phase evolution of the \( l = m \) harmonics, for example the phase evolution of the (3, 3) multipole moment is approximately a factor of 1.5 times \( \omega_{22} \) [281].

### 4.3 The Significance of Higher Harmonics

In this section, we investigate the observability of the different \((\ell, m)\) multipole moments, and how this varies across the mass and spin parameter space. For concreteness, we restrict attention to a single detector with a sensitivity comparable to that achieved by the LIGO observatories during their third observing run [8].

The key metric for waveform observability is the optimal SNR defined as

\[
\hat{\rho} = \sqrt{\langle h|\tilde{h} \rangle},
\]

where we have introduced the inner product weighted by noise (characterized by a power spectrum \( S(f) \)) as

\[
(a|b) := 4 \text{Re} \int_0^{f_{\text{max}}} \frac{\tilde{a}(f)\tilde{b}(f)^*}{S(f)} \, df.
\]

Consider the situation where the \((\ell, |m|) = (2, 2)\) harmonic has been observed, and we are interested in obtaining an estimate of the expected SNR in the other harmonics. As is clear from Eq. (4.3), the SNR in the higher harmonics will depend upon the detector sensitivity to the higher harmonic waveform, \( \tilde{h}_{\ell m} \), as well as the amplitude factor \( A_{+x} \).

Let us examine the single-detector case in detail. For simplicity, we choose a detector sensitive only to the + polarization (in the preferred frame), so that \( w_\times = 0 \), and we take \( w_+ = 1 \). Furthermore, we simplify the calculation to consider only two multipole moments, the (2, 2) harmonic and one other generic \((\ell, m)\) harmonic. The amplitude of each multipole depends on both the intrinsic properties of the system and the orientation relative to the network of detectors.
The waveform observed at the detector defined in the preferred frame with $F_+ = \cos 2\psi$ and $F_\times = -\sin 2\psi$ is

$$h = \cos 2\psi(h_+^{22} + h_{\times}^{\ell m}) - \sin 2\psi(h_+^{22} + h_{\times}^{\ell m}),$$

(4.7)

where $h_+^{\ell m}$ and $h_{\times}^{\ell m}$ are the two orthogonal components for the $(\ell, m)$ multipole moment of the waveform. In the frequency domain, these are defined as a single $(l, m)$ index pair from the summation introduced in Eq. (4.3):

$$\hat{h}_+^{\ell m}(f) := \frac{d_\alpha}{d_\ell} A_+^{\ell m}(\alpha)e^{im\phi_0}\tilde{h}_{\ell m}(f)$$

and

$$\hat{h}_\times^{\ell m}(f) := \frac{d_\alpha}{d_\ell} A_{\times}^{\ell m}(\alpha)e^{im\phi_0}\tilde{h}_{\ell m}(f).$$

(4.8)

A simple substitution of Eq. (4.7) into Eq. (4.5) gives the optimal SNR for a signal comprising two harmonics as

$$\hat{\rho}^2 = \cos^2 2\psi \left[ |\hat{h}_+^{22}|^2 + |\hat{h}_{\times}^{\ell m}|^2 + 2(\hat{h}_+^{\ell m} |\hat{h}_+^{22}|) \right]$$

$$+ \sin^2 2\psi \left[ |\hat{h}_\times^{22}|^2 + |\hat{h}_+^{\ell m}|^2 + 2(\hat{h}_+^{\ell m} |\hat{h}_\times^{22}|) \right]$$

$$- 2 \sin 2\psi \cos 2\psi \left[ (\hat{h}_+^{22} |\hat{h}_+^{\ell m}) + (h_+^{22} |\hat{h}_{\times}^{\ell m}) \right]$$

(4.9)

where the cross terms ($\sin 2\psi \cos 2\psi$) between polarizations for a single mode cancel since $(\hat{h}_+^{\ell m} |\hat{h}_+^{22}) = 0$.

The cross terms between different multipole moments, $(\hat{h}_+^{\ell m} |\hat{h}_\times^{22})$, can be both positive or negative, causing constructive or destructive interference between the harmonics. As discussed previously, the frequency during inspiral scales with $m$ while the ringdown frequency has been observed to scale approximately with $\ell$. Consequently, there is typically little overlap between the $(2, 2)$ multipole and multipoles for which both $\ell \neq 2$ and $m \neq 2$. Thus, in many cases, the cross terms between different harmonics will not make a significant contribution. We examine in detail the importance of the overlap between different multipole moments in section 4.3.1. For now, we restrict to the case where these terms can be neglected.

Neglecting the cross terms between harmonics, $(\hat{h}_+^{\ell m} |\hat{h}_+^{22})$, we are able to reexpress the optimal SNR as

$$\hat{\rho}^2 = \cos^2 2\psi |\hat{h}_+^{22}|^2 + \sin^2 2\psi |\hat{h}_\times^{22}|^2$$

$$+ \cos^2 2\psi |\hat{h}_+^{\ell m}|^2 + \sin^2 2\psi |\hat{h}_{\times}^{\ell m}|^2$$

$$= \rho_{22}^2 + \rho_{\ell m}^2.$$ 

(4.10)

We have defined, in the obvious way, the power in the $(2, 2)$ and $(\ell, m)$ multipole moments as $\rho_{22}^2$ and $\rho_{\ell m}^2$ respectively.
Next, using the form of \( h_{\ell m} \) from Eq. (4.8), we can write

\[
\rho_{\ell m} = \left[ (A_{+}^{2})^2 \cos^2 2\psi + (A_{\times}^{2})^2 \sin^2 2\psi \right] |\tilde{h}_{\ell m}|^2 ,
\]

(4.11)

which is also valid for \((\ell, m) = (2, 2)\). The SNR in the \((\ell, m)\) harmonic is given by an orientation contribution, dependent upon inclination and polarization, and a term that depends upon the overall amplitude of the given multipole moment. Consequently, we can express the SNR in the higher harmonic as:

\[
\rho_{\ell m} = \rho_{22} \alpha_{\ell m} R_{\ell m} ,
\]

(4.12)

where \(\alpha_{\ell m}\) encodes the relative, intrinsic amplitude of the \((\ell, m)\) multipole moment relative to the \((2, 2)\) multipole and \(R_{\ell m}\) encodes the relative size of the orientation factors. Specifically,

\[
\alpha_{\ell m}^2 = \frac{(|\tilde{h}_{\ell m}|^2 |\tilde{h}_{\ell m}|)}{(|\tilde{h}_{22}|^2 |\tilde{h}_{22}|)}
\]

(4.13)

and

\[
R_{\ell m}^2 = \frac{[(A_{+}^{2})^2 \cos^2 2\psi + (A_{\times}^{2})^2 \sin^2 2\psi] }{[(A_{+}^{2})^2 \cos^2 2\psi + (A_{\times}^{2})^2 \sin^2 2\psi]}
\]

(4.14)

In general, the relative amplitudes \(R_{\ell m}\) will depend upon both the inclination and polarization angles. However, for the \(\ell = m\) multipole moments, the expression simplifies as the relative amplitudes are the same for both +, \times polarization. In this case, there is no dependence upon the polarization angle and .

\[
R_{33}(\iota) = 2 \sin \iota \\
R_{44}(\iota) = 2 \sin^2 \iota
\]

(4.15)

In Section 4.3.1, we explore the dependence of the relative amplitudes \(\alpha_{\ell m}\) over the mass and spin parameter space and, in Section 4.3.2, the expected distribution of \(R_{\ell m}\) for a population of sources.

### 4.3.1 Dependence upon intrinsic parameters

The two important intrinsic parameters determining the relative power in the higher harmonics are mass ratio and total mass, with spin effects entering at higher post-Newtonian (PN) order for most harmonics . The contribution of a higher harmonic relative to the \((2, 2)\) harmonic generically increases with an increasing mass ratio. The relative amplitudes of the multipole moments are independent of the total mass of the system. However the frequency content of each multipole does depend upon the total mass and thus, depending on the shape of the detector power spectral density, certain higher harmonics might be preferentially observed. In particular,

\[\text{In this case, the cross term between polarizations in Eq. (4.9) also vanishes.}\]
Chapter 4. Measuring gravitational-wave higher order multipoles

Figure 4.3.1: Ratio of the intrinsic amplitude, $\alpha_{\ell m}$, (defined in equation (4.13)) of signal harmonics to the (2,2) harmonic as a function of the total (detector frame) mass and mass ratio of the system, in a detector with sensitivity matching the Advanced LIGO detectors during O3 [8]. Upper left: the (3,3) harmonic; Upper right: the (4,4) harmonic; lower left: the (2,1) harmonic; lower right: the (3,2) harmonic. In all cases, the spins of the black holes are set to zero. The solid white line corresponds to $\alpha_{\ell m} = 5.3/20$ and the dashed line to $\alpha_{\ell m} = 2.1/20$, which correspond, approximately, to the threshold for the higher harmonics being confidently/marginally observable for a signal with SNR=20 in the (2,2) multipole. Note that the colorbar is normalized differently between the top and bottom row to improve the visibility of the weaker harmonics.

The contribution of higher harmonics can become more significant at high masses, for which the merger frequency of the dominant harmonic lies below the optimal sensitivity of the detector.

In Fig. 4.3.1 we show the relative amplitude, $\alpha_{\ell m}$, of the four multipoles that we are considering: the (3,3), (4,4), (2,1) and (3,2) harmonics. The amplitudes have been calculated using the PhenomHM waveform [252], for a signal observed in a detector with LIGO O3 sensitivity [16, 8], as a function of the (detector frame) total mass and mass ratio of the system.

Over much of the parameter space, the (3,3) harmonic is the most significant, with the relative amplitude of the (3,3) harmonic increasing with mass ratio. For example, at a total mass of $50M_\odot$, the (3,3) harmonic has 10% of the amplitude of the leading harmonic at a mass ratio of 2:1 and 20% at 5:1. At high masses, and significant mass ratios, the relative sensitivity to the (3,3) harmonic is greater
4.3. The Significance of Higher Harmonics

Figure 4.3.2: Absolute value of the noise-weighted inner product between multipoles, evaluated using the Advanced LIGO (O3) sensitivity, as a function of mass ratio and total mass for non-spinning black holes ($\chi_{\text{eff}} = 0$). **Left:** the overlap between the (2,2) and (2,1) multipoles; **Right:** the overlap between (2,2) and (3,2) multipoles.

than one third of the (2,2) multipole. The (4,4) multipole is generally the third most significant, after the (2,2) and (3,3) harmonics. However, sensitivity to the (4,4) multipole increases rapidly as the mass of the system increases so that for total mass above $\sim 75 M_\odot$ and mass ratio less than 2:1, the (4,4) multipole is more significant than the (3,3) multipole.

The intrinsic amplitudes of the (2,1) and (3,2) harmonics are always lower than at least one of the (3,3) and (4,4) harmonics, over the mass and mass ratio ranges explored in Figure 4.3.1. As with the other harmonics, their relative importance increases as the mass ratio decreases and also, for the (3,2) harmonic in particular, as the total mass increases. The (2,1) multipole is the only subdominant multipole considered in this chapter which has spin terms in the amplitude at 1 PN order [232]. For this reason, the (2,1) multipole is more significant for binaries with large anti-aligned spins: the intrinsic amplitude roughly doubles for a binary with effective spin $\chi_{\text{eff}} = -0.8$, relative to a non-spinning system. Even then, the relative contribution of the (2,1) harmonic is less than the (3,3) harmonic.

For the power in these higher harmonics to be observable, it must be possible to distinguish the signal in the higher harmonic from the (2,2) harmonic. Generally, it is only the contribution which is orthogonal to the (2,2) harmonic which will be observable. Any contribution from the higher harmonics which is proportional to the (2,2) harmonic will simply serve to change the power observed in the (2,2). Consequently, we are interested in knowing whether the waveforms are orthogonal or, equivalently, in the overlap between the harmonics. Here, we define the normalized overlap maximized over the reference phase, $\phi_0$,

$$O(\ell m, 22) = \frac{\operatorname{Max}_{\phi_0} \langle \tilde{h}_{\ell m} | \tilde{h}_{22} \rangle}{| \langle \tilde{h}_{\ell m} | \tilde{h}_{22} \rangle |}. \quad (4.16)$$

The overlap between the (3,3) and (4,4) multipoles with the (2,2) harmonic is
< 10% across the parameter space explored, as expected due to the fact that the frequency evolution of these harmonics differs significantly from the (2,2). However, the overlap of the (2,2) harmonic with the (2,1) and (3,2) multipoles can be significant. These overlaps are shown in Fig. 4.3.2 as a function of total mass and mass ratio. As the (2,1) multipole has approximately the same frequency as the (2,2) multipole during ringdown, we expect a significant overlap at higher masses when the (merger and) ringdown occur within the sensitive band of the detector. Similarly, for the (3,2) multipole, the frequency evolution during the inspiral matches closely with the (2,2) multipole and we therefore expect a significant overlap between the (2,2) and (3,2) multipoles, particularly for low masses. Consequently, it can be difficult to identify these harmonics in the signal. Interestingly, one of the most significant impacts of the (3,2) multipole can be to produce an incorrect estimate of the amplitude of the (2,2) harmonic, and consequently introduce an error in the measured distance, as power from the (3,2) multipole will be mistakenly attributed to the (2,2) harmonic [265].

4.3.2 Dependence upon extrinsic parameters

The observed SNR in the higher harmonics depends upon the orientation of the binary, through the $R_{\ell m}$ factor defined in Eq. (4.14), in addition to the intrinsic amplitude of the harmonics discussed above. We can make several immediate observations from Fig. 4.2.1 or, equivalently, directly from the functional form of the spin-weighted spherical harmonics. The (3,3), (4, 4) and (2, 1) multipoles vanish for a signal observed face-on ($\iota = 0$), so the minimum value of $R_{\ell m}$ for these harmonics is zero; in contrast, there is no orientation for which both polarizations of the (3, 2) harmonic vanishes. Next, there is no orientation where the (2,2) harmonic vanishes, but the other harmonics do not — the (2,2) harmonic only vanishes for the $\times$
4.3. The Significance of Higher Harmonics

polarization for an edge-on system, but all other harmonics we are considering also vanish there. Thus, there is a finite, maximum value of \( R_{\ell m} \) for all harmonics, and it’s easy to see from Fig. 4.2.1 that \( R_{\ell m}^{\text{max}} = 2 \), which occurs for edge-on systems.

In Fig. 4.3.3 we plot the distribution of the geometrical factors \( R_{33} (\iota) \) and \( R_{44} (\iota) \). We have restricted attention to the (3, 3) and (4, 4) multipoles as these are the most significant, as seen in Fig. 4.3.1, and also the expression for \( R_{\ell m} \) is independent of the observed gravitational wave polarization. We consider the distribution of \( R_{\ell m} \) for a population of sources distributed uniformly in volume \(^5\) and with uniformly distributed orientation. We show both the distribution of \( R_{\ell m} \) for a uniform population of sources, as well as the expected observed distribution. In order to obtain the observed distribution for \( R_{33} (\iota) \) and \( R_{44} (\iota) \), it is sufficient to consider selection effects only on the inclination angle. If binaries are uniformly distributed in volume and orientation the distribution of inclinations for signals above an arbitrary detection threshold is known analytically \(^{283}\) (see Eq. (4.34)). This allows us to derive analytical expressions for the \( p(R_{33} (\iota)) \) and \( p(R_{44} (\iota)) \) distributions shown in Fig. 4.3.3, which we do in Appendix 4.B.

For both the (3,3) and (4,4) harmonics, the distribution peaks at \( R_{\ell m} = 2 \), the value for an edge-on system, with mean values of \( \bar{R}_{33} = 1.57 \) and \( \bar{R}_{44} = 1.33 \). However, since the emission in the (2, 2) harmonic is weakest when the system is observed edge on, selection effects serve to significantly reduce the peak in the observed population. For the (3, 3) harmonic, the peak remains at \( R_{33}^{\text{obs}} = 2 \), but the distribution is broad, with significant support over the full range from 0 to 2 and a mean value of \( \bar{R}_{33}^{\text{obs}} = 1.16 \). The mode of the observed \( R_{33}^{\text{obs}} \) distribution is zero, although again there is broad support over the range from 0 to 2 with a mean value of \( \bar{R}_{33}^{\text{obs}} = 0.79 \).

For other harmonics, the expected distribution of \( R_{\ell m} \) will depend upon the sensitivity of the detector network to the two polarizations of the gravitational wave — the distribution for \( R_{\ell m} \) will differ between a single detector, sensitive to only one polarization, and a network with good sensitivity to both polarizations. Nonetheless, the distribution for \( R_{21} \) will share features with \( R_{33} \) and \( R_{44} \), namely it will take values between 0 (face on) and 2 (edge on), with a peak at \( R_{21} = 2 \) which is reduced by selection effects in the observed population. The distribution for \( R_{32} \) will also be bounded between 0 and 2, although, due to the fact that the (3, 2) multipole doesn’t vanish for face-on sources, there is also a significant contribution at \( R_{32} \approx 1 \) from face-on sources.

\(^5\)Realistically, we do not expect sources to be uniformly distributed, due to both cosmological effects and a redshift dependent star formation and, hence, merger rate \(^{282}\). Nonetheless, this simple model provides a reasonably approximation to gain an understanding of the likely values of \( R_{\ell m} \).
4.4 Observing Higher Harmonics

When a gravitational wave signal from a binary merger is observed, it is natural to ask whether the higher multipoles have been observed. Typically, the searches that identify events do not use higher harmonics to extract events from the data [284, 285, 286] (although see [245] for ways to incorporate them). However, parameter estimation routines do incorporate higher harmonics into the recovery of parameters, and a natural way to ask whether higher harmonics have been observed is to calculate the Bayes factor (or odds ratio) between parameter recovery with and without higher multipoles in the waveform [9, 12]. In this Section, we show that the SNR in higher harmonics is also an effective tool in determining observability of higher harmonics.

4.4.1 Measured SNR in higher harmonics

We assume that the (2, 2) harmonic has been observed and consider the SNR contained in the higher multipoles. As in Eq. (4.7), we consider only two harmonics, the (2, 2) harmonic and a single additional harmonic. Since the (2, 2) harmonic has been identified, it is straightforward to calculate the SNR in the (ℓ, m) harmonic, by generating the \( \tilde{h}_{ℓm} \) waveform, with the same masses, spins and arrival time, and filtering it against the data. If the overlap between the (ℓ, m) and (2, 2) harmonics is non-zero, then this will pick up power contained in the (2, 2) harmonic, and it is necessary to remove it by first computing the orthogonal component,

\[
\tilde{h}_{ℓm}^⊥ = \tilde{h}_{ℓm} - \frac{1}{|\tilde{h}_{22}|^2} \left[ (\tilde{h}_{ℓm} | \tilde{h}_{22}) \tilde{h}_{22} + (\tilde{h}_{ℓm} | i\tilde{h}_{22}) i\tilde{h}_{22} \right].
\]  

(4.17)

Here, \( \tilde{h}_{22} \) and \( i\tilde{h}_{22} \) denote the two orthogonal phases of the (2, 2) harmonic. Filtering \( \tilde{h}_{ℓm}^⊥ \) against the data, \( s \), gives

\[
(\rho_{ℓm}^⊥)^2 = \frac{1}{|\tilde{h}_{ℓm}^⊥|^2} \left[ (s|\tilde{h}_{ℓm}^⊥)^2 + (s|i\tilde{h}_{ℓm}^⊥)^2 \right].
\]  

(4.18)

When the parameters of the waveform are known, or have been inferred through parameter estimation, we can calculate the expected SNR in the (ℓ, m) multipole as

\[
\hat{ρ}_{ℓm}^⊥ = \hat{ρ}_{ℓm} \sqrt{1 - O(ℓm, 22)^2},
\]  

(4.19)

where \( \hat{ρ}_{ℓm} \) is the expected SNR in the (ℓ, m) harmonic, as defined in Eq. (4.12) and \( O(ℓm, 22) \) is the overlap between the (ℓ, m) and (2, 2) waveforms as defined in Eq. (4.16).

In Fig. 4.4.1 we show the inferred posterior probability distribution for \( \hat{ρ}_{33}^⊥ \) for a binary with masses \( m_1 = 40M_⊙, m_2 = 10M_⊙ \) inclined at \( \cos ι = 0.7 \) (\( ι ≈ 45° \)) and with \( ρ_{22} = 22 \) under a variety of assumptions for signal and model. 6 For the

---

6 All parameter estimates reported in this chapter were obtained with LALInference [200] assum-
4.4. Observing Higher Harmonics

Figure 4.4.1: Posterior probability distribution for \( \hat{\rho}_\perp^{33} \), the orthogonal optimal signal-to-noise ratio of the (3, 3) multipole. The simulated waveform corresponds to system with \( m_1 = 40M_\odot, m_2 = 10M_\odot \) and \( \cos \iota = 0.7 \). The two solid histograms show the posterior distribution when the (3, 3) multipole is included in the waveform, either with (green) or without (orange) Gaussian noise. The dotted histogram shows the posterior on the SNR in the (3, 3) harmonic inferred from the measurement of the (2, 2) harmonic alone. The vertical line indicates the simulated value of \( \hat{\rho}_\perp^{33} \) and the dashed lines indicate the expected distribution, based on a non-central \( \chi \) distribution with two degrees of freedom and assuming a flat prior (as derived in Section 4.4.2), in the presence and absence of a signal in the (3, 3) harmonic. Also shown as a solid brown line is the predicted posterior distribution for a signal with \( \hat{\rho}_\perp^{33} = 5.5 \), but instead using the inferred distribution from the measurement of the (2, 2) harmonic (the dotted blue histogram) as a prior. We see general agreement between the predicted and measured posteriors, with the measured values slightly smaller than predicted (see Section 4.4.3 for discussion).
simulated signal the relative amplitude of the (3, 3) harmonic is $\alpha_{33} \approx 0.18$, and the orientation factor is $R_{33} = 1.4$, which implies $\hat{\rho}_{33} \approx 5.5$ and, since the overlap between (2, 2) and (3, 3) harmonics is small, $\hat{\rho}_{33} \approx \hat{\rho}_{33}$. The recovered distribution of $\hat{\rho}_{33}$, both with and without noise [(3,3) zero noise and (3,3) Gaussian noise in the figure], matches well with the simulated value but is shifted to lower values. We discuss this shift in Section 4.4.3.

We can also infer the power in the (3, 3) harmonic even when we use only the (2, 2) harmonic to recover the parameters, particularly masses and binary orientation, of the system. Unsurprisingly, the distribution of $\hat{\rho}_{33}^\parallel$ no longer matches well with the simulated value and now spreads over a broad range from 0 to 8. In this case, it seems clear that the (3, 3) harmonic has been observed, as its inclusion leads to a significant change in the inferred SNR in the (3, 3) harmonic.

In Fig. 4.4.2 we show the inferred posterior probability distributions of inclination, distance, polarization and phase at coalescence using waveform models that do/do not include the higher harmonics. Although the binary is generated with the orbital plane inclined at an angle of $\iota \approx 45^\circ$, using only the (2, 2) harmonic,
4.4. Observing Higher Harmonics

the system is recovered consistent with face-on, due to well-known degeneracies between distance in inclination [173]. Consequently, the only well measured quantities are the amplitude and phase of the circularly polarized waveform that is recovered: $A_{22} \approx \frac{\cos \iota}{D_L}$ and $\phi_{22} \approx \psi + \phi_o$, with the inclination bounded between $0 \leq \iota \leq 45^\circ$. When the (3, 3) harmonic is added, the degeneracy is broken and the distance, inclination, polarization and phase are all measured with good accuracy.

4.4.2 Expectation due to noise

The question, then, is whether an observed SNR in a given higher harmonic is evidence that the higher harmonic has been observed, or if this is to be expected due to noise alone. Similar questions have been addressed multiple times in the gravitational wave literature, for example in [287, 288, 289, 290, 291]. Here, we follow the methods developed in those earlier papers, focusing on a specific application to higher harmonics. We calculate the expected distribution of $\hat{\rho}_{\ell m}$ under some simplifying assumptions. Specifically, we consider the scenario where measurement of the (2, 2) harmonic has already fixed the parameters which determine the phase evolution of the binary (primarily the chirp mass, but also a combination of aligned spin and mass ratio [291]), the time of arrival and sky location of the system. Furthermore, we assume that the ($\ell, m$) multipole is the second most significant (in many cases, this is the (3, 3) harmonic), and other multipoles do not contribute significantly.

We will treat the overall amplitude and phase of the ($\ell, m$) multipole as unconstrained by the observation of the (2, 2) harmonic. As shown in Fig. 4.4.2, there are degeneracies in the measurement of distance/inclination and polarization/phase when observing only the (2, 2) harmonic. The amplitude of the higher harmonics, and in particular the (3, 3) and (4, 4) harmonics, varies significantly over the range $0^\circ \leq \iota \leq 45^\circ$ and can therefore be treated as unconstrained. Similarly, the overall phase of these multipoles differs from the (2, 2) by a factor of $(m-2)\phi_o$ and is therefore unconstrained by the measurement of the (2, 2) harmonic. Another way to see this is to look at the posterior probability distribution for the (3, 3) amplitude inferred when using a waveform containing only the (2, 2) multipole in Fig. 4.4.1. The distribution is broad and has support across a large range of $\hat{\rho}_{33}$. This argument will only hold for the subdominant harmonic: once the amplitude of a second harmonic is fixed, the four orientation parameters of the binary are determined and, consequently, the amplitude of the remaining harmonics is significantly constrained.

We are interested in obtaining the expected distribution for $\hat{\rho}_{\ell m}$ under the noise-only hypothesis (i.e. only power in the (2, 2) harmonic). In this case, we are fitting

\[ \text{For simplicity of presentation, we drop the } \perp \text{ from the equations in the remainder of the section. Where the harmonic has overlap with the (2, 2) waveform, the calculation should be understood to be performed with the orthogonal component.} \]
Chapter 4. Measuring gravitational-wave higher order multipoles

the data with a template waveform

$$\tilde{h} = a\tilde{h}_{\ell m} + ib\tilde{h}_{\ell m} + c\tilde{h}_{22} + id\tilde{h}_{22},$$

(4.20)

where $\tilde{h}_{\ell m}$ and $i\tilde{h}_{\ell m}$ are the two phases of the waveform of the $\ell m$ harmonic, $a$ and $b$ control the overall amplitude of this harmonic, and $\tilde{h}_{22}$, $c$ and $d$ give the contribution of the dominant harmonic to the waveform. We are interested in the expected distribution of $a$ and $b$ when there is no power in higher harmonics. Based upon the discussion above, we choose a uniform prior $\pi(a, b)$ on $a$ and $b$. In what follows we neglect the terms related to the dominant harmonic as they are unaffected, to the level of our approximation, by the presence of the higher harmonics. The posterior will be

$$p(a, b|s) \propto \Lambda(a, b)\pi(a, b).$$

(4.21)

where the likelihood of a signal $s$ given the amplitudes $a$, $b$ and Gaussian noise is

$$\Lambda(s|a, b) \propto \exp \left[ -\frac{1}{2} \left( s - \tilde{h}(a, b) \right) \left( s - \tilde{h}(a, b) \right)^{\dagger} \right].$$

(4.22)

Using polar variables $\rho_{\ell m} = \sqrt{a^2 + b^2}$ and $\phi_{\ell m} = \arctan(b/a)$, and assuming a uniform prior we can write the posterior probability distribution for the amplitudes $a$ and $b$ given a signal $s$ as

$$p(a, b|s) \propto \rho_{\ell m} e^{\left[ -\rho_{\ell m}^2 + \rho_{\ell m} \left( \cos \phi_{\ell m} - \phi_{\ell m} \right) \right] d\phi_{\ell m}}$$

(4.23)

Defining the matched filter signal-to-noise ratio, $\rho_{\ell m}$ as in Eq. (4.18) and the phase

$$\phi_{\ell m} = \arctan \left( \frac{s|\tilde{h}_{\ell m}}{s|\tilde{h}_{\ell m}} \right)$$

and marginalizing over $\phi_{\ell m}$, we obtain

$$p(\rho_{\ell m}|s) \propto \rho_{\ell m} e^{\left[ -\frac{\rho_{\ell m}^2}{2} + \rho_{\ell m} \cos \phi_{\ell m} - \phi_{\ell m} \right] d\phi_{\ell m}}$$

$$\propto \rho_{\ell m} \exp \left[ -\frac{\rho_{\ell m}^2 + \rho_{\ell m}^2}{2} \right] I_0(\rho_{\ell m})$$

(4.24)

where $I_0$ is the modified Bessel function of the first kind. We recognize Eq. (4.24) as the non-central chi distribution with 2 degrees of freedom and non-centrality parameter equal to $\rho_{\ell m}$. In the absence of signal power in the higher harmonics,
4.4. Observing Higher Harmonics

In Fig. 4.4.3 we show the distribution of $\hat{\rho}_{33}$, in the absence of a signal in the (3, 3) harmonic. First, we have the recovered distribution when performing parameter estimation on a signal containing only the (2, 2) harmonic and the zero instance of the noise distribution. Based upon the calculation above, we expect this to follow the $\chi$ distribution with two degrees of freedom, and we see that it does. We also show the distribution for three different instances of Gaussian noise. In each of these cases, the distribution is expected to follow a non-central $\chi$ distribution, where the non-centrality parameter is given by the matched filter SNR in the (3, 3) harmonic – in this case, there is no signal and any power is simply due to noise. For two of the noise realizations (Gaussian noise 1 and Gaussian noise 3), there was minimal power in the (3, 3) harmonic and the $\hat{\rho}_{33}$ distribution matches closely with the zero-noise case. However, in the Gaussian noise 2 realization, the SNR in the (3, 3) multipole is higher, and the mode of the distribution is moved significantly away from zero.

Figure 4.4.3: Posterior probability distribution for $\hat{\rho}_{33}$, the orthogonal optimal signal-to-noise ratio of the (3, 3) multipole, when the simulated signal contains the (3, 3) multipole for a variety of models and noise realizations. The injected parameters are $m_1 = 40M_\odot$, $m_2 = 10M_\odot$ at $\cos\iota = 0.7$.

The probability distributions for the filters $(s|\tilde{h}_{\ell m})$ and $(s|i\tilde{h}_{\ell m})$ are zero-mean, unit-variance gaussians and $\rho_{\ell m}$ is chi-distributed with 2 degrees of freedom.
We propose a simple test for the observability of higher multipole moments in a gravitational waveform: if the SNR in the second most significant harmonic satisfies $\rho_{\ell m} > 2.1$, this signifies the observation of power in the higher multipole moments. We have argued that the matched filter SNR, in the absence of signal, will be well approximated by a $\chi$ distribution with two degrees of freedom. We expect Gaussian noise to produce an SNR greater than 2.1 less than 10% of the time and therefore require $\rho_{\ell m} > 2.1$. \(^8\)

The estimate of $\rho_{\ell m}$ can be obtained either by matched filtering, or by fitting the measured distribution of $\hat{\rho}_{\ell m}$ from parameter estimation results and obtaining the non-centrality parameter. Based on this criteria, our third noise trial would show marginal evidence for presence of the $(3, 3)$ harmonic. This prescription can be easily extended to a criterion for confident detection of the higher harmonics: a “5-sigma” observation could correspond to $\rho_{\ell m} > 5.3$. In Fig. 4.3.1, we have added contours at values of $\alpha_{\ell m} = 2.1/20$ and 5.3/20. These indicate the approximate boundaries in the mass space where higher harmonics would be marginally/confidently observed for a signal at SNR = 20. Of course, the actual higher harmonic SNR will depend also on the orientation factor $R_{\ell m}$, which varies between 0 and 2, with a median value around 1 for the $(3, 3)$ and $(4, 4)$ harmonics.

Next, we return to Fig. 4.4.1 and note that the distribution of $\hat{\rho}_{33}$ for the signal containing higher harmonics matches well with expectation – a non-central $\chi$ distribution with non centrality parameter 5.5 – but is shifted to lower values. A more accurate prediction for the distribution can be made by revisiting our assumptions. The derivation of the expected distribution Eq. (4.24), assumed a uniform prior for $\hat{\rho}_{33}$. We note that the dotted histogram in Fig. 4.4.1 is the probability of $\hat{\rho}_{33}$ before we probe the likelihood associated with the $(3,3)$ harmonic and so is approximately equivalent to the prior on $\hat{\rho}_{33}$. Thus, the prior distribution is broad but not flat. The solid curve in Fig. 4.4.1 shows the product of the likelihood in Eq. (4.24) with the prior as approximated by the dotted histogram. We see that this results in a more accurate predicted posterior. Remaining differences between the predicted and measured posteriors can be sourced to our assumption that the phase evolution (or equivalently the masses and spins) of the waveform templates are fixed. During inference, as the masses and spins are varied, the matched filter SNR is never larger and is usually lower than the simulated value. By fixing the SNR to the simulated value we therefore overestimate the actual value and the distribution is shifted to larger values of $\hat{\rho}_{33}$.

An alternative method of establishing the observability of higher harmonics is to compare the Bayes factor (or evidence) between a waveform model additionally

\(^8\)In [292], the authors propose an observability criterion specifically for higher harmonics observed during the ringdown of the final black hole. They require an SNR in each harmonic above 5, corresponding to our 5-sigma detection. In addition, require distinguishability between the ringdown mode frequencies which (as can be seen in their Table 1) is virtually always satisfied if the modes can be observed.
containing the (3, 3) multipole and a model with only the (2, 2) multipole [9, 12]. The difference in Bayes factor, obtained by marginalizing the likelihood [200], between the two parameter estimation runs (with and without (3, 3) harmonic) is $\log_{10} BF = 4.7$. We can compare this to our results on the SNR in the higher harmonics by noting that an increase in SNR corresponds to an increase of the likelihood by a factor of $\approx e^{\rho^2/2}$. Our injected value of $\rho_{33} = 5.5$, leads to an estimate of $\log_{10}$ Bayes factor of 6.6 ($\log_{10}$ of the increase in the likelihood). For a more accurate comparison, we should also account for the prior distribution, as well as the width of the posteriors. Since both the (2, 2) only and higher harmonic waveforms are described by the same parameters, the priors are unchanged. However, as is clear from Fig. 4.4.2, the posterior is significantly more peaked when the higher harmonics are included. The improved constraints from the (3, 3) multipole reduce the prior volume by a factor of $\sim 2$ in the distance inclination plane (assuming a uniform in volume prior), and a factor of $\sim 5$ in the polarization phase plane. This implies the Bayes Factor based purely on the increased likelihood be reduced by a factor of $\sim 10$, equivalent to reducing the $\log_{10}$ Bayes Factor by one to 5.6. Finally, as discussed above, the higher harmonic SNR is generally inferred to be smaller than the simulated value. Indeed, from Fig. 4.4.1, the peak occurs at $\hat{\rho}_{33} \sim 5.0$, which corresponds to a $\log_{10}$ Bayes Factor of 4.4. This is in close, but not perfect agreement with the full parameter estimation result.

### 4.5 Higher Harmonics in a Population of Binary Mergers

Here, we consider the likelihood of observing the higher harmonics in signals drawn from a population. To do so, we generate a large number of potential signals from a given population and assess which would be observable above a given threshold and, of those, which would have sufficient power in the (3, 3) and/or (4, 4) harmonics for them to be observable (above the threshold of $\rho_{\ell m} = 2.1$). We choose a mass distribution of black holes in binary systems where the mass of the more massive black hole is taken from a power-law distribution $p(m_1) \propto m_1^{-\alpha}$ and choose the power law parameter of $\alpha = -2.35$, while restricting the mass to lie in the range $[5, 50] \, M_\odot$; the distribution for $m_2$ is taken to be uniform in the range $[5M_\odot, m_1]$. The spins of the individual black holes are assumed to be isotropically distributed, with low spin magnitudes (the magnitude is a triangular distribution peaked at spin magnitude of zero and falling to zero at maximal spin) [177]. Binaries are assumed to be isotropic on the sky, with uniform orientations and distances distributed uniformly in comoving volume.

In Fig. 4.5.1 we show the subset of this population which would be detectable with the HLV network operating at the sensitivities achieved during O3 [8]. More
pertinently, we also plot the subset from this detected population which result in gravitational waves with a measurable signal in the two loudest subdominant multipoles. Overall around 5.5% of binaries are expected to have sufficient power in the higher harmonics for them to be observed. Of these, the vast majority will have an observable (3, 3) harmonic (5.3%), with the (4, 4) harmonic observable in 0.85% of binaries, but for the majority of these, the (3, 3) harmonic will also be observable. Only two to three observable events in 1,000 from this population are expected to have an observable (4, 4) harmonic but not observable (3, 3).

The higher harmonics are preferentially observable in signals with unequal masses and for sources for binaries which are significantly inclined. In particular, for binaries with mass ratio between 4:1 and 10:1, the majority of signals will have observable higher harmonics, and even at a mass ratio of 2:1, around 10% of binaries will have observable higher harmonics. Convolving the observed distribution with the fraction of binaries with significant higher harmonics gives a peak of signals with observable higher harmonics around a mass ratio of 3:1. Interestingly, for binaries close to equal mass, it is the (4, 4) harmonic which is more likely to be observed, and essentially all binaries where the (4, 4) but not (3, 3) is observed have close to equal masses (between 1:1 and 5:4).

4.6 Discussion

We have explored the relative significance of the higher gravitational wave harmonics in binary merger signals. For simplicity, we have decomposed the harmonics
into an overall amplitude — dependent upon the masses and spins of the system — and an orientation-dependent term — dependent upon the inclination and polarization. This allows us to easily identify the most significant harmonics, and the regions of parameter space where they are most likely to be observable. As is well known [231, 232, 267], the higher harmonics are most significant when the binary is observed edge on. As expected, our orientation amplitudes are largest for edge-on systems although, due to selection effects, we observe that the most likely observed configuration is a binary with axes orientated at around 45° to the line of sight. In addition, we show that for much of the binary parameter space, the (3, 3) multipole will be the most significant sub-dominant harmonic, with an amplitude about one tenth of the (2, 2) harmonic for a mass-ratio 2 binary (over a broad range of masses). The (4, 4) multipole becomes more significant at higher masses and, although the relative amplitude is less than 0.1 for much of the parameter space, it is still the most significant sub-dominant harmonic for high-mass systems where the two components have comparable masses.

For signals which are observed at low SNR, it is likely that at most one additional harmonic will be clearly observable. Thus, for simplicity, we have introduced an observability criterion for the second harmonic. In many cases, the amplitude and phase of the second harmonic is largely unconstrained by the observation of the (2, 2) multipole: there are often large degeneracies between the measurement of the distance and inclination of the binary and also the polarization and phase [173]. Consequently, in the absence of a signal, the power in the second most significant harmonic will be \( \chi^2 \) distributed with two degrees of freedom, corresponding to the unknown amplitude and phase of the harmonic. If there is power in the higher harmonic, the distribution will be non-central \( \chi^2 \), where the non-centrality is given by the SNR in the higher harmonic. We have performed a series of simulations that demonstrate this expectation is valid. Using this simple observation, we have introduced a criterion for observation of power in a higher harmonic: if the observed SNR in the second most significant harmonic is above 2.1, then this is unlikely to occur due to noise alone so there is marginal evidence of a higher multipole signal, while an SNR > 5.3 would provide strong (“5-sigma”) evidence.

We have identified regions in the parameter space where higher harmonics are most likely to be observed. These regions are those where higher harmonics are likely to be observed, but also which are relatively common in the underlying population of observed gravitational wave signals [266]. We find that these correspond to signals with mass ratios between 2:1 and 5:1 — for more equal masses, the higher harmonics are too weak, more unequal mass binaries are thought to be rarer. Furthermore, the most likely orientation is for the axis to be inclined at between 30° and 60° to the observer — less inclined systems have insufficient power in the higher harmonics while more inclined systems have a weaker overall emission.

There are several applications of the work presented here. As already mentioned,
the criterion for observability of higher harmonics has been in used, along with other methods [281], in establishing the presence of power in the (3, 3) harmonic in the observed signals GW190412 and GW190814 [9, 12]. Furthermore, the method can be used in a straightforward way to determine whether it is likely that the higher harmonics will be observable in a given system, and we have provided an example in the population study presented in section 4.5. This is directly applicable to signals observed using a search for the dominant harmonic. Based upon the observed parameters, we can calculate the expected power in the higher harmonics and identify the expected SNR. If significant SNR is expected in higher harmonics, then it becomes worthwhile to undertake the (computationally costly) parameter estimation with waveforms containing several gravitational wave harmonics. This will lead either to the observation of higher harmonics, and the subsequent improvement of parameter measurement, or the non-observation of higher harmonics and subsequent restriction of the binary parameters to regions of the parameter space where the higher harmonic amplitudes are low.

While the method introduced here is straightforward, there are several clear limitations. Most obviously, the discussion has limited attention to a single observable harmonic. In many cases, this will be a reasonable approximation as there will be one harmonic which is significantly larger than the others (as can be seen from Fig. 4.3.1). Furthermore, from simple parameter counting, it seems likely that the observation of a single higher harmonic will be sufficient to significantly improve parameter recovery, most notably the binary orientation. Nonetheless, the observation of additional multipoles will provide additional improvements. For a detailed understanding of the impact of all of the higher harmonics, a full, Bayesian parameter estimation exploration of the issue will be necessary [267]. Additionally, throughout this chapter, we have used a single waveform model, IMRPhenomHM [252] and checked for consistency with a more recently updated model IMRPhenomXHM [230]; but results are likely to vary somewhat with other models of the higher harmonics (for example, [257, 261, 260]). Finally, we have restricted attention throughout the chapter to non-precessing systems. Recently, [275, 276], an analysis similar to the one presented here was performed on precessing systems, again with a focus on the observability of the two dominant harmonics. For systems where both higher harmonics and precession have an significant impact on the waveform, it will be necessary to combine these approaches to develop a straightforward categorization of precessing systems with observable higher harmonics.
Appendix

4.A Spin-weighted spherical harmonic polarizations

The general form for the spin-weighted spherical harmonics is

\[ sY_{lm}(\iota, \phi_o) = (-1)^m \sqrt{\frac{(l + m)!(l - m)!(2l + 1)}{4\pi(l + s)!(l - s)!}} \sin^2 \left(\frac{l}{2}\right) \times \sum_{r=0}^{l-s} \binom{l-s}{r} \binom{l+s}{r+s-m} (-1)^{l-r-s} e^{im\phi_o} \cot^{2r+s-m} \left(\frac{l}{2}\right), \]  \hspace{1cm} (4.25)

which can be written in terms of the Wigner d-functions \( d_{lm}^{m*}(\iota) \) (implicitly defined here)

\[ sY_{lm}(\iota, \phi_o) = \sqrt{\frac{(2l + 1)}{4\pi}} d_{lm}^{m*}(\iota) e^{im\phi_o}. \]  \hspace{1cm} (4.26)

They have the following symmetries

\[ s\bar{Y}_{lm} = (-1)^{s+m} sY_{l(-m)} \]
\[ sY_{lm}(\pi - \iota, \phi_o + \pi) = (-1)^l sY_{lm}(\iota, \phi_o). \]
The spin-weighted spherical harmonics for the harmonics we are interested in are

\[ -2Y_{22} = \frac{1}{2} \sqrt{\frac{5}{\pi}} e^{i2\phi_o} \cos^4 \left( \frac{\iota}{2} \right) \]

\[ -2Y_{2-2} = \frac{1}{2} \sqrt{\frac{5}{\pi}} e^{-i2\phi_o} \sin^4 \left( \frac{\iota}{2} \right) \]

\[ -2Y_{21} = \frac{1}{2} \sqrt{\frac{5}{\pi}} e^{i\phi_o} \cos^2 \left( \frac{\iota}{2} \right) \sin(\iota) \]

\[ -2Y_{2-1} = \frac{1}{2} \sqrt{\frac{5}{\pi}} e^{-i\phi_o} \sin^2 \left( \frac{\iota}{2} \right) \sin(\iota) \]

\[ -2Y_{33} = \frac{1}{2} \sqrt{\frac{21}{2\pi}} \left( -e^{i3\phi_o} \right) \cos^4 \left( \frac{\iota}{2} \right) \sin(\iota) \]

\[ -2Y_{3-3} = \frac{1}{2} \sqrt{\frac{21}{2\pi}} e^{-i3\phi_o} \sin^4 \left( \frac{\iota}{2} \right) \sin(\iota) \]

\[ -2Y_{32} = \frac{1}{2} \sqrt{\frac{7}{\pi}} e^{i2\phi_o} \cos^4 \left( \frac{\iota}{2} \right) \left( 3\cos(\iota) - 2 \right) \]

\[ -2Y_{3-2} = \frac{1}{2} \sqrt{\frac{7}{\pi}} e^{-i2\phi_o} \sin^4 \left( \frac{\iota}{2} \right) \left( 3\cos(\iota) + 2 \right) \]

\[ -2Y_{44} = \frac{3}{4} \sqrt{\frac{7}{\pi}} e^{i4\phi_o} \cos^2 \left( \frac{\iota}{2} \right) \sin^2(\iota) \]

\[ -2Y_{4-4} = \frac{3}{4} \sqrt{\frac{7}{\pi}} e^{-i4\phi_o} \sin^4 \left( \frac{\iota}{2} \right) \sin^2(\iota) \].

We can write the gravitational wave polarizations as a sum of these spherical harmonics with coefficients \( h_{lm} \)

\[ h_+ - ih_\times = \sum_{l \geq 2} \sum_{m=-l}^{l} -2Y_{lm}(\iota, \phi_o) h_{lm}. \] (4.28)

Three properties of \( h_{lm} \) help to simplify Eq. (4.28). Firstly, specializing to planar (i.e. non-precessing) binaries allows us to write \( h_{l-m} = (-1)^l h_{lm}^{*} \) [223]. Secondly, in the frequency domain, \( \tilde{h}_{l-m}(f) = \tilde{h}_{lm}(-f)^{*} \). Finally we make the further approximation [293] that if we only care about the waveform in direction \( \hat{n} \) we can neglect one side of the frequency spectrum, depending on the sign of \( m \). This approximation is valid in particular where the stationary phase approximation has been used. We therefore assume, with the sign convention on the Fourier transform as \( \tilde{h}(f) = \int dh(t)e^{+i2\pi ft} \), that

\[ \tilde{h}_{lm}(f) \approx 0 \begin{cases} f > 0, m < 0 \\ f < 0, m > 0. \end{cases} \] (4.29)

With these three properties we can obtain explicit expressions for the orientation
dependence of each of the harmonics for positive frequencies

\[
\tilde{h}_+(f) = \frac{1}{2} \sum_{l \geq 2} \sum_{m=-l}^{l} \left[ -2Y_{lm}(\iota, \phi_0)\tilde{h}_{lm}(f) + \frac{(-1)^l}{2}Y_{l-m}^*(\iota, \phi_0)\tilde{h}_{lm}(f) \right]
\]

(4.30)

\[
\tilde{h}_\times(f) = \frac{i}{2} \sum_{l \geq 2} \sum_{m=-l}^{l} \left[ -2Y_{lm}(\iota, \phi_0)\tilde{h}_{lm}(f) - \frac{(-1)^l}{2}Y_{l-m}^*(\iota, \phi_0)\tilde{h}_{lm}(f) \right]
\]

(4.31)

where in both cases, we have neglected the \( m = 0 \) terms in the sums as they are not considered in the models we have used. Finally, we note that we have used a different normalization convention in the main text, Eq. (4.3), than the one typically used in the spin-weighted spherical harmonic decomposition described in this Appendix. This has no impact on the results, but merely changes the values of \( \alpha_{\ell m} \) and \( R_{\ell m} \) while maintaining the same values of the SNR in the higher harmonics.

### 4.B Derivation of \( p(R_{\ell m}) \)

We now derive the probability distributions in Fig. 4.3.3. Assuming no preferred orientation for binaries in the universe, the probability density function for \( \cos \iota \), \( p(\cos \iota) \), is

\[
p_{\text{univ}}(\cos \iota) = \frac{1}{2} \quad (4.32)
\]

However, binaries which emit primarily in the 22 multipole radiate most powerfully in the direction perpendicular to the orbital plane, \(| \cos \iota | \sim 1 \). Consequently, the horizon for the subset of these binaries which are viewed edge-on is much closer and
we preferentially observe face-on binaries. It can be shown [283] that the radiated power of the dominant multipole as a function of inclination is

\[ F(\iota)^{22} = (A_{+}^{22})^2 + (A_{\times}^{22})^2, \quad (4.33) \]

where \( A_{+}\!\!,_{\times}^{22} \) are defined in Eq. (4.4). For a detector sensitive to only one polarization of gravitational wave, the observed power will depend upon the polarization. This will also be the case for a network with different sensitivities to the two polarization, but not for one equally sensitive to both polarizations of the gravitational wave. It is possible to approximately marginalize over the polarization distribution and obtain a probability distribution for the inclinations of detected binaries, assuming sources are distributed uniformly in volume, as [283]

\[ p_{\text{det}}(\cos \iota) \propto F(\iota)^{3/2} = (1 + 6 \cos^2 \iota + \cos^4 \iota)^{3/2}. \quad (4.34) \]

Using these results, it is straightforward to obtain expressions for the distributions for the expected power in the (3,3) and (4,4) multipoles, both for a uniform population of binaries and for those which are observable above a fixed threshold. The distribution for other multipoles can also be obtained but, since in general \( R_{\ell m} \) will depend upon polarization angle, the results will be dependent upon the details of the network and its sensitivity to the two gravitational wave polarizations. For the (3,3) and (4,4) multipole moments, the relative amplitude depends only on the inclination angle \( \iota \).

To obtain an expression for the probability distribution for \( R_{\ell m} \), we change variables

\[ p(R_{\ell m}) = \left( \frac{d \cos \iota}{d R_{\ell m}} \right) p(\cos \iota) \quad (4.35) \]

so that, recalling the functional form of \( R_{33} \) and \( R_{44} \) from Eq. (4.15), we obtain

\[ p_{\text{univ}}(R_{33}) = \frac{R_{33}}{4 \sqrt{1 - \left( \frac{R_{33}}{2} \right)^2}} \]

\[ p_{\text{univ}}(R_{44}) = \frac{1}{4 \sqrt{1 - \left( \frac{R_{44}}{2} \right)^2}}. \quad (4.36) \]

Assuming binaries are detected with (2,2) harmonic-only waveforms, we can apply the same weighting factor as above in obtaining the distributions for the observed binaries, to obtain

\[ p_{\text{det}}(R_{\ell m}) = \left( \frac{d \cos \iota}{d R_{\ell m}} \right) p_{\text{det}}(\cos \iota) \quad (4.37) \]
which gives

\[ p_{\text{det}}(R_{33}) \propto \left( 8 - 2R_{33}^2 + \frac{R_{33}^4}{16} \right) p_{\text{univ}}(R_{33}) \]

\[ p_{\text{det}}(R_{44}) \propto \left( 8 - 4R_{44}^2 + \frac{R_{44}^4}{4} \right) p_{\text{univ}}(R_{44}) \]  

These distributions are plotted in Fig. 4.3.3, and discussed in the surrounding text.
Chapter 5

GW190412 & GW190814: The first binaries with measurable higher order multipoles

This chapter is based on Refs. [9, 12], which are collaboration papers and involve contributions from a (very) large number of people. In particular section 5.2.2 closely follows Ref. [9], within which CM was the primary author of the section describing the (3,3) orthogonal SNR measurement (section IV B of Ref. [9]). CM contributed the (3,3) SNR calculation in these papers which appears in figures 5.2.4 and 5.3.5.

5.1 Introduction

Before 2019, all compact binary systems observed by the LIGO-Virgo gravitational-wave detector network had mass ratios consistent with unity [25]. Here we report on two compact binaries that are the first to buck this trend. In April 2019 GW190412 [9] was the first observation of an unambiguously asymmetric merger. Later that year GW190814 [12] — a binary with even more asymmetric masses — was observed with the unusual property of having a secondary component with a mass heavier than any neutron star, but lighter than any black hole previously detected in a compact binary. The unequal masses of both these binaries results in another novelty for gravitational-wave astronomy. For the first time we observe harmonics of the gravitational waveform oscillating at a frequency other than the dominant emission at twice the orbital frequency. It was the power in these higher harmonics that strengthens, or in the case of GW190412 allows, the characterization as unequal mass systems.

In section 5.2.1 we summarize the detection of GW190412 and introduce its properties. We focus on the most salient properties for the detection of higher
harmonics, which is our main focus here. Section 5.2.2 discusses the evidence for higher mode content. This is quantified with both the Bayesian evidence and the orthogonal optimal signal-to-noise ratio in each harmonic. Section 5.3 repeats a similar structure for GW190814.

5.2 GW190412

5.2.1 Detection and parameter estimation

GW190412 was detected by multiple low latency searches with network SNR $\sim 18-19$ as a triple coincidence with high significance. Inferred parameter estimates are shown in Table 5.2.1. The table shows results for two waveforms with similar physics but obtained using different modelling techniques (for more information on the models used see appendix 5.A and references therein). Combined estimates are obtained by averaging the posterior distributions inferred with the two different waveform models. Unless otherwise stated the properties quoted in this chapter are these combined estimates. The component masses in the source frame are $m_1 = 30.1^{+4.6}_{-5.3} M_\odot$ and $m_2 = 8.3^{+1.6}_{-0.9} M_\odot$, which in the detector frame combine for a total mass of $44.2^{+4.4}_{-4.7} M_\odot$.

The detection claims of both events are based on a bank of template waveforms that include only the dominant $(\ell, m) = (2, 2)$ multipole, and do not include precession. Both of these physical effects leave only small imprints on the waveform for the majority of binaries [276, 275, 10], though their inclusion can reduce parameter uncertainties. Larger higher harmonics are expected for binaries that have more asymmetric masses, and are viewed closer to edge-on.

In figure 5.2.1 we see that the inclusion of higher order multipoles enables us to constrain the mass ratio of GW190412 away from equal mass. We constrain mass ratio to be at least 2:1 with 99% probability using a model including higher harmonics, compared with a confidence level of 91% with a model containing only the dominant harmonic. Figure 5.2.3 demonstrates how the power in the $(3,3)$ multipole is related to the properties of GW190412. Increasing either the binary inclination (equivalent to decreasing $\cos \iota$) or the mass asymmetry will increase the amplitude of the $(3,3)$ harmonic in the model waveform to match the power in the $(3,3)$ multipole in the data. Thus for events like GW190412 there is a degeneracy between inclination and mass ratio that is not observed when the $(3,3)$ multipole is

1For further details about the events see Refs [9, 12].

2Symbols: $m_i$: individual mass; $M = m_1 + m_2$; $M = (m_1 m_2)^{3/5} M^{-1/5}$; superscript “det” refers to the detector-frame (redshifted) mass, while without subscript, masses are source-frame masses, assuming a standard cosmology [294] detailed in Appendix B of [25]; $q = m_2/m_1$; $M_t$, $\chi_t$: mass and dimensionless spin magnitude of the remnant BH, obtained through numerical-relativity fits [295, 296, 297, 298]; $\chi_{eff}$, $\chi_p$: effective and precessing spin parameter; $\chi_1$: dimensionless spin magnitude of more massive BH; $D_L$: luminosity distance; $z$: redshift; $\theta_{JN}$: inclination angle (folded to $[0, \pi/2]$); $\rho_X$: matched-filter SNR for the Hanford, Livingston and Virgo detectors, indicated by subscript. $\rho_{HLN}$: network SNR.
Table 5.2.1: Inferred parameter values for GW190412 and their 90% credible intervals, obtained using precessing models including higher multipoles. Symbols are described in the text.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>EOBNR PHM</th>
<th>Phenom PHM</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1/M_{\odot}$</td>
<td>31.7$^{+3.6}_{-3.5}$</td>
<td>28.1$^{+4.8}_{-4.3}$</td>
<td>30.1$^{+4.6}_{-5.3}$</td>
</tr>
<tr>
<td>$m_2/M_{\odot}$</td>
<td>8.0$^{+0.7}_{-0.7}$</td>
<td>8.8$^{+1.3}_{-1.1}$</td>
<td>8.3$^{+1.6}_{-0.9}$</td>
</tr>
<tr>
<td>$M/M_{\odot}$</td>
<td>39.7$^{+3.0}_{-2.8}$</td>
<td>36.9$^{+3.7}_{-2.9}$</td>
<td>38.4$^{+3.8}_{-3.9}$</td>
</tr>
<tr>
<td>$M/M_{\odot}$</td>
<td>13.3$^{+0.3}_{-0.3}$</td>
<td>13.2$^{+0.5}_{-0.3}$</td>
<td>13.3$^{+0.4}_{-0.4}$</td>
</tr>
<tr>
<td>$q$</td>
<td>0.25$^{+0.06}_{-0.04}$</td>
<td>0.31$^{+0.07}_{-0.07}$</td>
<td>0.28$^{+0.12}_{-0.07}$</td>
</tr>
<tr>
<td>$M_t/M_{\odot}$</td>
<td>38.6$^{+3.1}_{-2.8}$</td>
<td>35.7$^{+3.8}_{-3.0}$</td>
<td>37.3$^{+3.8}_{-4.0}$</td>
</tr>
<tr>
<td>$\chi_f$</td>
<td>0.68$^{+0.04}_{-0.04}$</td>
<td>0.67$^{+0.07}_{-0.07}$</td>
<td>0.67$^{+0.06}_{-0.05}$</td>
</tr>
<tr>
<td>$m_1^{\text{det}}/M_{\odot}$</td>
<td>36.5$^{+4.2}_{-4.2}$</td>
<td>32.3$^{+5.7}_{-5.2}$</td>
<td>34.6$^{+5.4}_{-6.4}$</td>
</tr>
<tr>
<td>$m_2^{\text{det}}/M_{\odot}$</td>
<td>9.2$^{+0.7}_{-0.7}$</td>
<td>10.1$^{+1.6}_{-1.2}$</td>
<td>9.6$^{+1.7}_{-1.0}$</td>
</tr>
<tr>
<td>$M^{\text{det}}/M_{\odot}$</td>
<td>45.7$^{+3.5}_{-3.3}$</td>
<td>42.5$^{+4.4}_{-3.7}$</td>
<td>44.2$^{+4.4}_{-4.7}$</td>
</tr>
<tr>
<td>$\dot{\chi}_{\text{eff}}$</td>
<td>0.28$^{+0.06}_{-0.08}$</td>
<td>0.29$^{+0.08}_{-0.11}$</td>
<td>0.25$^{+0.08}_{-0.11}$</td>
</tr>
<tr>
<td>$\chi_p$</td>
<td>0.31$^{+0.14}_{-0.10}$</td>
<td>0.31$^{+0.24}_{-0.17}$</td>
<td>0.31$^{+0.19}_{-0.16}$</td>
</tr>
<tr>
<td>$\chi_i$</td>
<td>0.46$^{+0.15}_{-0.15}$</td>
<td>0.41$^{+0.22}_{-0.24}$</td>
<td>0.44$^{+0.16}_{-0.22}$</td>
</tr>
<tr>
<td>$D_L$/Mpc</td>
<td>740$^{+120}_{-130}$</td>
<td>740$^{+150}_{-140}$</td>
<td>740$^{+130}_{-140}$</td>
</tr>
<tr>
<td>$z$</td>
<td>0.19$^{+0.02}_{-0.02}$</td>
<td>0.19$^{+0.03}_{-0.04}$</td>
<td>0.15$^{+0.03}_{-0.03}$</td>
</tr>
<tr>
<td>$\theta_{JN}$</td>
<td>0.71$^{+0.23}_{-0.21}$</td>
<td>0.71$^{+0.39}_{-0.29}$</td>
<td>0.71$^{+0.31}_{-0.24}$</td>
</tr>
<tr>
<td>$\rho_H$</td>
<td>9.5$^{+0.1}_{-0.2}$</td>
<td>9.5$^{+0.2}_{-0.3}$</td>
<td>9.5$^{+0.1}_{-0.3}$</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>16.2$^{+0.1}_{-0.2}$</td>
<td>16.1$^{+0.2}_{-0.3}$</td>
<td>16.2$^{+0.1}_{-0.3}$</td>
</tr>
<tr>
<td>$\rho_V$</td>
<td>3.7$^{+0.2}_{-0.5}$</td>
<td>3.6$^{+0.3}_{-1.0}$</td>
<td>3.6$^{+0.3}_{-0.7}$</td>
</tr>
<tr>
<td>$\rho_{\text{HLN}}$</td>
<td>19.1$^{+0.2}_{-0.2}$</td>
<td>19.0$^{+0.2}_{-0.3}$</td>
<td>19.1$^{+0.2}_{-0.3}$</td>
</tr>
</tbody>
</table>
Figure 5.2.1: GW190412: The posterior distribution for mass ratio, $m_2/m_1 > 1$. The inclusion of higher multipoles constrains the binary away from equal mass. All models shown here are part of the Phenom family.

not measured. This degeneracy follows lines of constant $\rho_{33}$. Also shown here is the total mass in the detector frame, which is degenerate with mass ratio along a line of constant chirp mass [291]. We note there appears to be some railing against the lower bound on the total mass prior in the IMRPhenomPv2 run. Setting a smaller lower bound on the total mass would likely result in more accurate posteriors which show more support at lower total mass, and more equal masses. This means that the impact of the higher multipoles is likely even greater for constraining the binary away from equal mass.

5.2.2 Evidence for higher multipoles

As already noted, signal models that include higher multipoles are needed to infer the strongest constraints on GW190412’s source properties. This is because if the data contain significant imprints of higher multipoles, the appropriate models can fit the data better than dominant-mode models, leading to a higher statistical likelihood. Conversely, if the data would not contain imprints of higher multipoles, using more complex models allows us to disfavor configurations in which clear imprints of higher multipoles are predicted [252, 267, 299].

In this section, we analyse how strong the imprints of higher multipoles are in GW190412 and ask if their contributions in the data are significantly stronger than random noise fluctuations. We address this question using two different approaches, each with its unique set of strengths and caveats.
Figure 5.2.2: GW190412: The posterior distribution for the luminosity distance, $D_L$, and inclination, $\theta_{JN}$ (angle between the line-of-sight and total angular momentum), of GW190412. We illustrate the 90% credible regions. By comparing models that include either the dominant multipole (and no precession), higher multipoles and no precession, or higher multipoles and precession, we can see the great impact higher multipoles have on constraining the inclination and distance. All models shown here are part of the EOBNR family. This plot is reproduced from Ref. [9].
Figure 5.2.3: GW190412: Posterior distributions for the orthogonal optimal SNR in the $(\ell, m) = (3, 3)$ multipole, $\rho_{33}$ (see section 5.2.2 and Ref. [10], the inclination angle between the line-of-sight and orbital angular momentum, $\iota$, mass ratio $q = m_2/m_1 < 1$ and total mass in the detector frame under different waveform model assumptions. Results for the precessing higher multipole waveform IMRPhenomPv3HM model are shown as red contours, while grey contours denote the precessing waveform IMRPhenomPv2 which contains only the $(2,2)$ multipole in the co-precessing frame. Both more unequal masses and more planar viewing angles result in a larger amplitude $(\ell, m) = (3, 3)$ multipole. For this reason the binary inclination angle is degenerate with the mass ratio. Shading shows the $1\sigma$, $3\sigma$ and $5\sigma$ credible regions. This plot was made with PESummary [11].
Bayes Factors and Matched-Filter SNR

One answer to the question of which model describes the data best can be given in the Bayesian framework. The ratio of marginalized likelihoods under two competing hypotheses is called the Bayes factor, $B$ [300]. Bayes factors may be used to quantify support for one hypothesis over another. The Bayes factor does not take into account our prior belief in the hypotheses being tested. Within GR, every compact binary coalescence signal includes higher multipoles and the prior odds in favor of their presence in the signal are infinite. We therefore focus on the Bayes factors and do not discuss the odds ratio (which is the Bayes factor multiplied by the prior odds).

We find $\log_{10} B = 3.6$ in favor of IMRPhenomPv3HM over its dominant multipole counterpart IMRPhenomPv2. This indicates strong evidence that the observed signal contains measurable imprints of higher multipoles. In Ref. [9] they show that despite uncertainty of order unity in $\log_{10} B$, they consistently find $\log_{10} B \geq 3$, in favor of higher multipoles, robust to waveform and code choice.

Optimal SNR

A complementary way to quantify the strength of higher multipoles is to use parameter estimation results from a signal model including higher-order multipoles [301, 10]. Each multipole is decomposed into parts parallel and perpendicular to the dominant multipole by calculating the noise-weighted inner product [302, 287] (often referred to as overlap) between it and the dominant multipole. Among the strongest multipoles that are included in our models, the $(\ell, m) = (3, 3), (4, 4)$ and $(4, 3)$ multipoles of GW190412 are close to orthogonal to the dominant $(2, 2)$ multipole within the band of the detector. In contrast, the $(3, 2)$ and $(2, 1)$ multipoles have non-negligible
parallel components. To quantify the strength of the higher multipoles we remove any parallel components from the multipoles and calculate the orthogonal optimal SNR using IMRPhenomHM [252]. We find \((\ell,m) = (3,3)\) to be the strongest sub-dominant multipole.

The templates that include higher multipoles do not allow the amplitude and phase of the \((3,3)\) multipole to be free parameters; they are determined by the properties of the system. An analysis of this event using only the dominant \((2,2)\) multipole recovers posteriors that are consistent with a broad range of inclinations, coalescence phases, and mass ratios, while the same analysis using higher multipoles results in significantly more restricted posteriors (see Fig. 5.2.2). This suggests that by changing those parameters, our models can effectively treat the amplitude and phase of the higher multipoles as tunable parameters that make their contributions more or less pronounced. If the data only contained the dominant quadrupole mode and Gaussian noise, the squared orthogonal SNR in the subdominant multipole will be \(\chi^2\)-distributed with two degrees of freedom [275, 276, 10]. This was verified by analysing an injection with parameters close to GW190412.

This noise-only distribution is shown in Fig. 5.2.4, along with the orthogonal optimal SNR in the \((\ell,m) = (3,3)\) mode. The peak of the SNR distribution is at the Gaussian equivalent three sigma level for the noise-only distribution (i.e., with cumulative tail probability of \(p = 3 \times 10^{-3}\)), making this the most significant evidence for something other than the dominant multipole prior to GW190814 [303].

5.3 GW190814

5.3.1 Detection and parameter estimation

GW190814 was initially detected only by gstLAL, but later offline versions of all matched filter searches were able to recover the event as a triple coincidence with network SNR \(\sim 24 − 25\) and high significance. Table 5.3.1 lists the inferred properties of GW190814. Component masses are \(m_1 = 23.2^{+1.1}_{-1.0} M_\odot\) and \(m_2 = 2.59^{+0.08}_{-0.09} M_\odot\).

Like GW190412, the constraint on the mass ratio of GW190814 was also improved by the inclusion of higher order multipoles, though to a lesser extent. As can be seen in figure 5.3.1, it is the combined effects of both higher multipoles and precession that improves the constraint on the secondary mass. The constraint was obtained with both EOBNR and Phenom waveform families, and is robust against waveform systematics. This increases the confidence that the secondary object is in the mass gap: lighter than any black hole, but heavier than any neutron star known to exist in a compact binary system [12].

GW190814’s mass ratio and secondary component mass is most tightly constrained by waveform models that include spin precession effects. We understand this as follows. The large mass ratio of GW190814 implies that the spin of the
larger black hole dominates the total angular momentum, and spin phenomenology of the binary. Furthermore, a large component of in plane spin on the larger black hole would have measurable precession effects. As precession is not observed in GW190814, we tightly constrain the in plane components of the larger black hole as can be seen in figure 5.3.3. A constraint on the in-plane spin will also constrain the aligned spin, as they are correlated with each other. We can thus bound the dimensionless spin magnitude of the primary black hole to $\chi_1$. This is the tightest constraint to date. It is this constraint from precession which impinges on the $\chi_{eff} - q$ space consistent with the evolution of the binary phase [291], and narrows the credible bounds on the mass of the lighter object.

Inclusion of higher order multipoles can also serve to also improve the constraints on the binary orientation angles. Figure 5.3.4 shows how the improved constraint on the binary inclination angle (between the line of sight and total angular momentum) partially breaks the degeneracy with the luminosity distance [173] for GW190814.

It is the power in the (3,3) harmonic which drives the measurement of the inclination, as the two parameters are related by $\rho_{33} \propto \sin \iota$ and are highly correlated

<table>
<thead>
<tr>
<th>$m_1/M_\odot$</th>
<th>EOBNR PHM</th>
<th>Phenom PHM</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_2/M_\odot$</td>
<td>23.2$^{+1.0}_{-1.0}$</td>
<td>23.2$^{+1.3}_{-1.3}$</td>
<td>23.2$^{+1.1}_{-1.1}$</td>
</tr>
<tr>
<td>$m_2/m_1$</td>
<td>0.112$^{+0.008}_{-0.008}$</td>
<td>0.112$^{+0.009}_{-0.009}$</td>
<td>0.112$^{+0.008}_{-0.008}$</td>
</tr>
<tr>
<td>$M/L_\odot$</td>
<td>6.10$^{+0.05}_{-0.05}$</td>
<td>6.08$^{+0.05}_{-0.06}$</td>
<td>6.09$^{+0.06}_{-0.06}$</td>
</tr>
<tr>
<td>$L/L_\odot$</td>
<td>25.8$^{+1.0}_{-1.0}$</td>
<td>25.8$^{+1.2}_{-1.2}$</td>
<td>25.8$^{+1.0}_{-1.0}$</td>
</tr>
<tr>
<td>$\Theta_{rad}$</td>
<td>0.9$^{+0.3}_{-0.2}$</td>
<td>0.8$^{+0.2}_{-0.2}$</td>
<td>0.8$^{+0.3}_{-0.3}$</td>
</tr>
<tr>
<td>$\rho_{PH}$</td>
<td>10.6$^{+0.1}_{-0.1}$</td>
<td>10.7$^{+0.1}_{-0.2}$</td>
<td>10.7$^{+0.1}_{-0.2}$</td>
</tr>
<tr>
<td>$\rho_{PL}$</td>
<td>22.21$^{+0.09}_{-0.15}$</td>
<td>22.16$^{+0.17}_{-0.09}$</td>
<td>22.16$^{+0.17}_{-0.10}$</td>
</tr>
<tr>
<td>$\rho_{PV}$</td>
<td>4.3$^{+0.2}_{-0.2}$</td>
<td>4.1$^{+0.6}_{-0.2}$</td>
<td>4.2$^{+0.2}_{-0.6}$</td>
</tr>
<tr>
<td>$\rho_{HLV}$</td>
<td>25.0$^{+0.1}_{-0.2}$</td>
<td>24.9$^{+0.1}_{-0.2}$</td>
<td>25.0$^{+0.1}_{-0.2}$</td>
</tr>
</tbody>
</table>

Table 5.3.1: Source properties of GW190814: We report the median values along with the symmetric 90% credible intervals for the SEOBNRv4PHM (EOBNR PHM) and IMRPhenomPv3HM (Phenom PHM) waveform models. The primary spin magnitude and the effective precession is given as the 90% upper limit. The inclination angle is folded to $[0, \pi/2]$. The last column is the result of combining the posteriors of each model with equal weight. The sky location of GW190814 is shown in Figure 2 of Ref. [12].
5.3. GW190814

Figure 5.3.1: GW190814: The marginalized posterior distribution for the secondary mass obtained using a suite of waveform models. The vertical lines indicate the 90% credible bounds for each waveform model. The labels Phenom/EOBNR PHM (generic spin directions + higher multipoles), Phenom/EOBNR HM (aligned-spin + higher multipoles) and Phenom/EOBNR (aligned-spin, quadrupole only) indicate the different physical content in each of the waveform models. This plot is reproduced here from Ref. [12].

as can be seen in Fig. 5.3.2. The degeneracy between mass ratio and \( \cos \iota \) is less obvious for this event, partly because the mass ratio measurement appears to be driven by the non-observation of spin.

5.3.2 Evidence for higher multipoles

GW190814 exhibits stronger evidence for higher-order multipoles than GW190412, with \( \log_{10} B \approx 9.6 \) in favor of a higher-multipole vs. a pure quadrupole model. The \((\ell, m) = (3, 3)\) is the strongest subdominant multipole, with \( \log_{10} B \approx 9.1 \) in favor of a signal model including both the \((\ell, m) = (2, 2)\) and \((3, 3)\) multipole moments. GW190814’s stronger evidence for higher multipoles is expected given its more asymmetric masses and the larger network SNR.

We infer that the orthogonal optimal SNR of the \((\ell, m) = (3, 3)\) multipole is \(6.6^{+1.3}_{-1.4}\), as shown in Figure 5.3.5. This is the strongest evidence for measuring a subdominant multipole to date [9].
Figure 5.3.2: GW190814: Posterior distributions for the orthogonal optimal SNR in the $(\ell, m) = (3, 3)$ multipole, $\rho_{33}$ (see section 5.2.2 and Ref. [10], the inclination angle between the line-of-sight and orbital angular momentum, $i$, mass ratio $q = m_2/m_1 < 1$ and total mass in the detector frame under different waveform model assumptions. Results for the precessing higher multipole waveform IMRPhenomPv3HM model are shown as green contours. Red contours denote results for the non-precessing higher mode model IMRPhenomHM. Grey contours denote the precessing waveform IMRPhenomPv2 which contains only the $(2, 2)$ multipole in the co-precessing frame. Here we see that including precession has the largest effect on reducing the uncertainty in mass ratio. Shading shows the $1\sigma$, $3\sigma$ and $5\sigma$ credible regions. This plot was made with PESummary [11].
Figure 5.3.3: GW190814: Two-dimensional posterior probability for the tilt-angle and spin-magnitude for the primary object (left) and secondary object (right) based on the Combined samples. The tilt angles are 0° for spins aligned and 180° for spins anti-aligned with the orbital angular momentum. The tiles are constructed linearly in spin magnitude and the cosine of the tilt angles such that each tile contains identical prior probability. The color indicates the posterior probability per pixel. The probabilities are marginalized over the azimuthal angles. This plot is reproduced here from Ref. [12].
Figure 5.3.4: GW190814: The posterior distribution for the luminosity distance, $D_L$, and inclination, $\theta_{JN}$ (angle between the line-of-sight and total angular momentum), of GW190814. We illustrate the 90% credible regions. By comparing models that include either the dominant multipole (and no precession), higher multipoles and no precession, or higher multipoles and precession, we can see the great impact higher multipoles have on constraining the inclination and distance. This plot is reproduced here from Ref. [12].

Figure 5.3.5: GW190814: Posterior distributions for the precessing SNR, $\rho_p$ (green) and the optimal SNR in the (3,3) sub-dominant multipole moment, $\rho$ (orange). The grey dotted line shows the expected distribution for Gaussian noise. This plot is reproduced here from Ref. [12].
Appendix

5.A Waveform models used

Table 5.A.1 provides details and references for the two waveform families used in this chapter: Phenom (including IMRPhenomD, IMRPhenomHM, IMRPhenomPv2, IMRPhenomPv3HM), and EOBNR (including SEOBNRv4_ROM, SEOBNRv4HM_ROM, SEOBNRv4P and SEOBNRv4PHM). Occasionally we refer to the models by a shortened name. Usually this is their family name followed by the letters P if they include precession effects and HM if they include higher order multipoles.

Table 5.A.1: Waveform models used for GW190814 and GW190412. We indicate which multipoles are included for each model. For precessing models, the multipoles correspond to those in the co-precessing frame.

| Name                  | precession | multipoles ($\ell, |m|$)                     | Ref. |
|-----------------------|------------|----------------------------|------|
| SEOBNRv4_ROM          | ×          | (2, 2)                     | [304]|
| SEOBNRv4HM_ROM        | ×          | (2, 2), (2,1), (3, 3), (4, 4), (5, 5) | [256, 261]|
| SEOBNRv4P             | ✓          | (2, 2), (2, 1)            | [305, 306, 307]|
| SEOBNRv4PHM           | ✓          | (2, 2), (2, 1), (3, 3), (4, 4), (5, 5) | [305, 306, 307]|
| IMRPhenomD            | ×          | (2, 2)                     | [308, 309]|
| IMRPhenomHM           | ×          | (2, 2), (2, 1), (3, 3), (3, 2), (4, 4), (4, 3) | [252]|
| IMRPhenomPv2          | ✓          | (2, 2)                     | [274]|
| IMRPhenomPv3HM        | ✓          | (2, 2), (2, 1), (3, 3), (3, 2), (4, 4), (4, 3) | [254]|
Chapter 6

Precession and higher order multipoles in the first half of the third gravitational wave observing run

Code: The code that produced the main results in this chapter will be made publicly available here: https://github.com/pesummary/pesummary/

Collaboration: This chapter was written by Charlie Hoy and CM, with the former (latter) paying most attention to the precessing (higher) harmonic measurements.

6.1 Introduction

Between 2015 and 2017, the Advanced LIGO [310] (aLIGO) and Advanced Virgo [311] (AdV) gravitational-wave (GW) detectors performed their first and second GW observing runs (O1 and O2). During this time, the LIGO Scientific and Virgo collaboration (LVC) announced GWs originating from 10 binary black hole mergers [312] and a single binary neutron star [313]. Additional compact binary candidates have also been reported by other groups [314, 315, 316, 317].

Two important General Relativistic effects that were not clearly observed during O1 and O2 are higher order multipoles [318, 264] and spin-induced orbital precession [312, 319]. Higher order multipoles are terms beyond the dominant quadrupole when a GW is expanded into multipole moments with spherical polar coordinates defined in the source frame [225]. Spin-induced orbital precession arises when there is a misalignment between the orbital angular momentum and the spins of each compact object [270]. Including these effects in waveform models used to infer source properties can improve parameter measurement accuracy and provides additional constraints on the in-plane spin components of the binary (see e.g.
The importance of both of these effects increase as the binary’s mass ratio \( q = m_1/m_2 \) increases. Clear evidence for asymmetric masses was absent in the binaries detected during O1 and O2, making the observation of either precession and higher order multipoles challenging.

Analysis of the first 6 months of data from the third GW observing run (O3a) revealed a further 39 GW candidates in the second gravitational-wave catalog (GWTC-2) [23]. In contrast to O1 and O2, several events in O3a had unequivocally unequal masses. First among these is GW190412 [9], with a mass ratio of \( \sim 4:1 \). The large mass ratio resulted in more significant higher order multipoles, and for the first time, imprints of subdominant multipole radiation oscillating at three times the orbital frequency were visible. Similarly, it was the first time that the amount of precession in the system was constrained away from the prior. Several months later GW190814 was detected with highly asymmetric component masses (\( \sim 9:1 \)) and a secondary component with a mass larger than any previously discovered neutron star and lighter than any black hole [12]. GW190814 had the largest evidence of the (3,3) multipole and the most precise precession measurement of any event observed to date. It was demonstrated that a combination of the higher order multipoles and the precession measurement improved parameter estimates, and in particular reduced the uncertainty on the mass of the smaller object. Other events in O3a have hinted at further evidence of higher order multipoles, though no single event in O3a unambiguously exhibits spin-induced orbital precession [9, 12, 23, 324].

In this chapter, we take advantage of recent developments in quantifying the presence of higher order multipoles [320] and precession [325, 319, 271] in GWs to build upon these statements. We calculate the signal-to-noise ratio in the \( (\ell,|m|) \in \{ (2, 1), (3, 3), (4, 4) \} \) multipoles\(^1\) and from precession for every event in O3a and compare it to the expected distribution from noise. Unlike Refs. [23, 324], we show that GW190412 exhibits evidence for spin-induced orbital precession. In addition to the previously reported evidence for higher multipoles in the signals of GW190814 and GW190412, we show that there is evidence for a (3,3) multipole in both the GW190519 and GW190929 signals\(^2\). Further, we demonstrate how a reanalysis of GW190521 by Nitz et al. [13] indicates a (3,3) multipole that is the most significant, as a fraction of total signal-to-noise ratio, detected to date.

\(^1\)\( (\ell,m) \) should everywhere be read as \( (\ell,|m|) \) unless otherwise indicated.
\(^2\)Throughout this chapter, we denote each gravitational-wave candidate by their shortened name. If there is only one gravitational wave on a given day, we identify it by its UTC date. If there are multiple, we add an underscore and as much of the time as needed to ensure the events can be distinguished.
6.2 Method

In General Relativity, GWs are fully described by two polarizations. These polarizations can be decomposed into multipole moments using the $-2$ spin-weighted spherical harmonic orthonormal basis (see Eq. (A1) in Ref. [10]). Coalescing compact binaries predominantly emit radiation at twice the orbital frequency in the leading order (2,2) multipole. The most important subdominant multipole for most compact binaries is the (3,3) multipole, though the (4,4) multipole can be more significant for binaries whose members have comparable masses [10].

For binaries where the dominant (2,2) multipole has been observed, it is natural to ask whether other multipoles can also be identified. Here, we quantify this by using parameter estimates inferred using a waveform model that includes these multipoles. These parameter estimates are used to generate the higher multipole waveform and calculate the noise-weighted inner product (also known as the optimal SNR) for each multipole using IMRPhenomXHM [230]. In general the $(l,m)$ multipole may overlap with the (2,2) signal, though for the (3,3) and (4,4) multipoles the overlap is usually negligible [10]. To quantify the significance of the higher multipole content it is necessary to project away the $(l,m)$ components that are parallel to the (2,2) multipole. Doing this we calculate the optimal orthogonal signal-to-noise ratio (SNR) $\rho_{tm}$ in the subdominant $(\ell,m)$ multipoles [10, 9, 12].

It is also natural to ask if precession has been observed. A binary on a quasicircular orbit is described by 8 intrinsic parameters: the individual component masses $m_1$ and $m_2$ and spin angular momenta $S_1$ and $S_2$. For a given binary, we can calculate the Newtonian orbital angular momentum $L$ and the total angular momentum $J = L + S = L + S_1 + S_2$. For the case where the total spin $S$ is misaligned with $L$, the system undergoes the General Relativistic phenomenon of spin-induced orbital precession [270]. In most cases, this phenomenon leads to $L$ precessing around the approximately constant $J$. This leads to characteristic modulations in the emitted gravitational wave [270, 326]. The strength of precession is characterised by the tilt angle of the binary’s orbit, $\beta$, defined as the polar angle between orbital angular momentum and total angular momentum.

Precession is often parameterised by the scalar quantity $\chi_p$ [327] (although alternative metrics have also been proposed [328, 329]). $\chi_p$ takes values between 0 (non-precessing) and +1 (maximal precession) and is widely used for inferring the occurrence of precession in gravitational-wave data (see e.g. [312, 23]).

Recently, Refs. [325, 319] introduced an alternative description for spin-induced orbital precession. This formalisation decomposes a gravitational-wave signal into 5 “precession harmonics” where the amplitudes form a power series in $b = \tan(\beta/2)$. It was demonstrated that each harmonic is equivalent to a gravitational wave emitted from non-precessing binary. The characteristic modulations in a precessing gravitational wave can then interpreted as the constructive and destructive interference of
these harmonics. In most regions of the parameter space, $b$ is small and therefore the leading two harmonics are sufficient to capture significant precession effects [325]. By decomposing a precessing GW signal in this way, it was shown that the power in the second most significant precession harmonic provides a good estimate of the contribution from precession to the SNR of the system. If this precession SNR $\rho_p$ is small, the second harmonic is insignificant and the observed GW looks like the dominant non-precessing signal. Analogously to the method above for higher order multipoles, we can use the results of parameter estimation obtained with a waveform model that includes precession to calculate the $\rho_p$ that are consistent with the data.

Both $\rho_p$ and $\rho_{33}$ are positively correlated with more unequal masses and larger inclination angles. In the case of precession, more unequal masses lead to a larger tilt angle of the binary’s orbit, and in general, the larger the tilt angle the more prominent the precession effects [270]. Likewise precession is easier to observe for binaries close to edge on, where a modulating inclination angle introduces larger fluctuations in the overall and relative amplitude of the two GW polarizations. The amplitude of the (3,3) multipole is identically zero at equal mass as only even $m$ multipoles respect the orbital symmetry of this arrangement [330]. As more asymmetry is introduced to the system in the form of unequal mass ratio, or misaligned spins the relative power of the (3,3) multipole grows [330, 232]. Relative importance of the (3,3) multipole also increases with inclination angle $\iota$: the ratio between the (3,3) and (2,2) amplitudes goes as $\sin \iota$ [10].

A measurement of the (2, 2) harmonic alone typically does not result in tight constraints on the parameters that most determine both $\rho_p$ and $\rho_{lm}$. In large part this is because it is generally not possible to distinguish the second polarization of the (2,2) harmonic [173]. This means essentially all events are recovered as being consistent with a face-on orientation, where the two polarizations have identical amplitudes. Mass ratio is also typically poorly constrained by the phase evolution of the (2,2) harmonic, where it is degenerate with spin [291]. Finally, a measurement of the (2,2) harmonic alone does not constrain the precession parameters $\chi_p$ or the precession phase. These poorly (or un-)constrained parameters translate into a freedom in amplitude and phase of the loudest subdominant precessing and higher harmonics, which can be tuned by adjusting the properties of the system to better fit the data.

In order to assess the significance of precession and higher order multipoles, we compare the inferred posterior distributions for $\rho_p$ and $\rho_{lm}$ to the distribution expected from noise alone. Since the statistical properties of $\rho_p$ and $\rho_{lm}$ are similar, the expected distribution has the same form for both measurements. Below we summarize the derivation of this common distribution (parameterized by $\rho$ which denotes either $\rho_p$ or $\rho_{lm}$) but we refer the reader to Refs. [271, 10] for the details.

The observed gravitational wave including subdominant multipole moments, or the observed gravitational wave originating from a precessing system, can be ap-
proximately written as a two component harmonic decomposition \([325, 10]\) \( h = A_0(\lambda)h_0(\lambda) + A_1(\lambda)h_1(\lambda) \) where \( A_i(\lambda) \) are overall amplitudes, \( h_i(\lambda) \) evolve with time, and indices 0 and 1 respectively represent dominant and loudest subdominant (precessing or higher multipole) harmonics for a set of parameters \( \lambda \). The gravitational-wave likelihood may then be factorised into two components: one describing the contribution from the dominant harmonic, \( \Lambda_0(\lambda) \), and another describing the contribution from the subdominant harmonic, \( \Lambda_1(\lambda) \),

\[
p(d|\lambda) \propto \exp\left(-\frac{1}{2}\langle d - (A_0(\lambda)h_0(\lambda) + A_1(\lambda)h_1(\lambda))|d - (A_0(\lambda)h_0(\lambda) + A_1(\lambda)h_1(\lambda))\rangle\right) \\
\propto \exp\left(\langle d|A_0(\lambda)h_0(\lambda)\rangle - \frac{|A_0(\lambda)|^2}{2}\langle h_0(\lambda)|h_0(\lambda)\rangle\right) \\
\times \exp\left(\langle d|A_1(\lambda)h_1(\lambda)\rangle - \frac{|A_1(\lambda)|^2}{2}\langle h_1(\lambda)|h_1(\lambda)\rangle\right) \\
\propto \Lambda_0(\lambda) \times \Lambda_1(\lambda),
\]

where we have assumed that the dominant and subdominant harmonics are orthogonal \( \langle h_0|h_1\rangle = 0 \). In general, the phase evolution of the gravitational waveform is well constrained by the measurement of the dominant harmonic. For simplicity we assume therefore that the subdominant contribution to the phase evolution is negligible and that the squared matched filter SNR in the subdominant channel \( (\rho_{1}^{MF})^2 = [(s|h_1|^2 + |s|h_1|^2)/|h_1|^2 \) is fixed, leaving only the overall amplitude and phase of \( h_1 \) to vary. Analytically marginalizing over the phase constant, we can write the subdominant contribution to the marginalized likelihood as,

\[
\Lambda_1(\rho) \propto I_0(\rho_1^{MF} \rho) \exp\left(-\frac{(\rho_1^{MF} \rho)^2}{2} + \rho^2\right),
\]

where \( I_0 \) is the Bessel function of the first kind. In gaussian noise \( \rho_1^{MF} \) is chi-distributed with two degrees of freedom. The expected posterior distribution for \( \rho \) is therefore,

\[
p(\rho|d) \propto p(\rho) \Lambda_1(\rho)
\]

where \( p(\rho) \) is the prior distribution for \( \rho \).

For the case of uniform priors on the complex amplitude \( A_1 \), \( p(\rho|d) \) takes the form of a non-central \( \chi \) distribution with 2 degrees of freedom with non-centrality parameter equal to \( \rho_1^{MF} \). To a rough approximation, \( \rho_1^{MF} \) can be estimated as the peak of the inferred \( \rho \) distribution. For a better approximation, we can fit a chi distribution to the posterior to obtain the non-centrality parameter.

A simple maximization (equivalent to assuming uniform priors), tends to result in a predicted posterior distribution that overestimate the actual inferred poste-
6.2. Method

This is because the posterior probability inferred from the dominant harmonic alone tends to prefer face-on or face-off orientations [173], and comparable masses due to the prior that is used, where both the higher order multipole and precession harmonics are weaker. This results in a general preference for lower values of $A_1$ and $\rho$, before the actual likelihood associated with the loudest subdominant harmonic is probed. In order to obtain a better prediction for the posterior distribution $p(\rho|d)$, we can use the results from an analysis that includes only the dominant multipole during parameter inference. Since generally the phase constant that is marginalized over in Eq.6.2 is not well measured in dominant mode inference, it is safe to assume a uniform prior. For the other parameters, we can use the results of dominant harmonic as an informed prior in Eq.6.3 for $p(\rho)$. This informed prior is essentially what results from calculating $p(\rho|d)$ in Eq.6.3 while assuming $A_1(\rho) = 1$. For precession there are additional parameters that must be marginalized over that are not inferred with models using aligned spin templates: $\chi_p$ and the precession phase. However the inference of aligned spin and mass ratio does provide additional constraints on these parameters, and so rather than assuming the default prior on these parameters, we condition on the measured aligned spin and mass ratio. To obtain an estimate of the posterior including we can then reweight each sample in the informed prior using the likelihood in Eq.6.2 calculated for a particular value of $\rho_{1}^{MF}$. Refs. [10, 271] demonstrated that this procedure can result in predicted posterior distributions that more accurately resemble the inferred posterior distribution.

To calculate the informed prior for $\rho_p$ we use the ‘AlignedSpinIMR” dataset where available [24]. For $\rho_{lm}$’s informed prior we use parameter estimation results for an approximant from the same family as the higher harmonic results. To obtain an estimate of the expected posterior in noise alone, we randomly draw a sample from the $\chi$ distribution with 2 degrees of freedom for $\rho_{1}^{MF}$ – consistent with a specific realisation of the noise. Next, we combine the informed prior with the likelihood from Eq. 6.2 to obtain the distribution of $\rho$ for a specific realisation of the noise. We repeat this procedure 100 times to represent different realisations of the noise. We then take the median of these distributions as our estimate for $\rho_{NP}^{p}$ for precession, and the noise posterior for higher multipoles. We show the average distributions across all events, $\bar{\rho}_{NP}^{p}$ and $\bar{\rho}_{33}$, as black solid lines in Figure 6.3.1.

To calculate the $\rho_{lm}$ and $\rho_p$ for each GW candidate, we require samples obtained from a precessing higher-order multipole approximant. For the majority of events, this meant using the “PublicationSamples” dataset contained within the publically available data files obtained through the Gravitational Wave Open Science Center [24]. For cases where the “PublicationSamples” dataset was not obtained with a precessing higher-order multipole approximant (see Table VIII of Ref. [23]), we

---

3GW190521 data release only included parameter estimation results for precessing higher-order multipole waveform models, so we were not able to construct informed priors for this event.

---
Figure 6.3.1: Plot showing the Left: $\rho_{33}$ and Right: $\rho_p$ distributions for all observations in the second gravitational-wave catalogue (grey). In red we show the $\rho_{33}$ and $\rho_p$ distribution averaged across events. In black we show the average of the median expected noise distribution for Left: higher multipoles and Right: precession. Events which are discussed in the text are colored.

use the “C01:SEOBNRv4PHM” dataset for both analyses. For all cases we used the posterior samples re-weighted to a flat-in-comoving-volume prior to remain consistent with Ref [23].

### 6.3 Results

Table 6.3.1 presents a summary of the main results. All measurements are reported as symmetric 90% credible intervals around the median of the marginalized posterior distribution, unless otherwise specified. Figure 6.3.1 shows the inferred posteriors for $\rho_{33}$ and $\rho_p$ for all events in O3a.

The (3,3) multipole is the most significant subdominant multipole for almost every event in O3a. This is expected for the vast majority of binaries [10]. GW190910 is the sole exception, having $\rho_{33} = 0.66_{-0.58}^{+1.31}$ and $\rho_{44} = 1.04_{-0.47}^{+0.94}$, both of which are consistent with noise. This occurs because GW190910 is very close to equal mass, $m_2/m_1 = 1.22_{-0.20}^{+0.48}$, where the (3,3) multipole is zero, but has significant support for an edge-on orientation, where the (4,4) multipole is strongest. Other than GW190412 and GW190814, GW190519 is the only event with a median value for $\rho_{33} > 2.1$, suggesting possible evidence for a (3,3) multipole. Close behind is GW190929 with $\rho_{33} = 2.0_{-1.5}^{+1.6}$. We discuss both these events in section 6.3.2.

The right panel of Figure 6.3.1 shows that in general there is no strong evidence for precession in O3a as $\bar{\rho}_p$ is almost indistinguishable from $\rho_p^{NP}$. Of all of the events in O3a, GW190412 [9] and GW190814 [12] have the largest $\rho_p$ with $\rho_p = 2.99_{-1.51}^{+1.58}$ and $\rho_p = 1.75_{-1.20}^{+1.60}$ respectively. To identify which events show evidence for precession we compute the Jensen-Shannon divergence $D_{JS}$ [331] between $\rho_p$ and $\rho_p^{NP}$ (as was done in Ref [25] between $\chi_p$ and the conditioned $\chi_p$ prior). This statistic is designed to quantify the difference between probability distributions and is bounded between $D_{JS} \in [0, 1]$bits. $D_{JS} = 0$ suggests that both distributions are
precession. For events where the JSDs and the SNR from precession

distributions to the expected noise distributions respectively.

Table 6.3.1: Table showing the SNR in the ($\ell, m$) = (3, 3) multipole moment $\rho_{33}$ and the SNR from precession $\rho_p$ for all events in the second gravitational wave catalogue [23] plus the reanalysis of GW190521 by Nitz et al. [13]. For each event we also show two Jensen Shannon Divergences (JSDs); $D_{JS}$ and $D_{JS}^p$ compare the calculated $\rho_{33}$ and $\rho_p$ distributions to the expected noise distributions respectively. Events with a larger JSD show greater evidence for higher order multipoles and/or precession. For events where the JSD and $\rho_{33}/\rho_p$ could not be calculated, due to the lack of publically available posterior samples, we add a hyphen. Where applicable we report the median values along with the 90% symmetric credible intervals.
Figure 6.3.2: $\rho_p$ distributions for *First row*: GW190412, *Second row*: GW190915, *Third row*: GW190814. The blue line shows the expected distribution of $\rho_p$ in a stretch of noisy data under the assumption that the source is non-precessing, $\rho_p^{NP}$. The blue shaded region shows the 1$\sigma$ uncertainty of $\rho_p^{NP}$. The black line shows the average $\rho_p^{NP}$ across all events.

identical. We find that most events have $D_{JS} < 0.05$bits which indicates the $\rho_p$ and $\rho_p^{NP}$ agree well, see Table 6.3.1. This highlights that for most events, the observed signal is consistent with originating from a non-precessing binary. This is why $\bar{\rho}_p$ is almost indistinguishable from $\bar{\rho}_p^{NP}$. We discuss why our analysis demonstrates GW190412 shows the largest evidence for precession in detail below. Since Ref. [23] identified GW190521 and GW190814 as having two of the most informative $\chi_p$ distributions compared to the prior, we also describe why our analysis does not highlight these events.
6.3. Results

6.3.1 GW190412 and GW190814

GW190814 and GW190412 are found to be the most significant measurements of $\rho_{33}$ as previously reported [9, 12] and as expected as their individual black hole masses are the most unequal. The Jensen-Shannon divergences between the (2,2) multipole informed prior and posterior for these events are also the largest, $D_{JS} = 0.68$ bits and $D_{JS} = 0.28$ bits respectively. This suggests posteriors density estimates change significantly with the inclusion of higher order multipoles.

GW190412 likely originated from a precessing system as the inferred $\rho_p$ is significantly larger than $\rho_{pNP}$ with $D_{JS} = 0.36$ bits, see Figure 6.3.2. We understand this because a) GW190412 is consistent with an exceptionally large $\rho_p$ compared to the other events in O3a owing to the second largest mass ratio $q = 3.6^{+1.1}_{-1.1}$ and inclination angle constrained away from face-on $\theta_{JN} = 0.7^{+0.3}_{-0.2}$ rad (folded between $[0, \pi/2]$) and b) GW190412’s informed prior is broad $\rho_p = 3.3^{+5.6}_{-2.7}$, meaning that $\rho_{pNP}$ approximately peaks at the location of maximum $\rho_{1MF}^p$. Since $\rho_{1MF}^p$ peaks at $\rho_p = 1.2$, with $\sim 1\%$ probability of random drawing a value consistent with GW190412’s large $\rho_p = 2.99$, $\rho_{pNP}$ peaks at much smaller values than the inferred $\rho_p$. This results in a large Jensen-Shannon divergence. We find that GW190412 is consistent with a precessing system at $> 97\%$ probability. This result differs from the conclusions presented in Refs [9, 332] which are based upon the more commonly used Bayes factors between the precessing and non-precessing hypothesis.

Figure 6.3.2 also shows that $\rho_{pNP}$ is shifted to larger values than average. This follows from the fact that GW190412’s aligned spin analysis also confidently identifies GW190412 as an unequal mass ratio system ($q = 3.2^{+1.1}_{-1.1}$). This means that for a given in-plane spin sample drawn from the prior, the calculated value of $\rho_p$ will be larger for GW190412 than for an equal mass ratio binary.

We see that GW190814 is consistent with originating from a non-precessing system with $D_{JS} = 0.03$ bits. This is expected given the near-zero $\chi_p$ measurement ($\chi_p = 0.04^{+0.04}_{-0.03}$). Despite this, GW190814 has the third largest $\rho_p$ in O3a. This apparent contradiction is a result of GW190814’s large mass ratio ($q = 9.0^{+0.8}_{-0.6}$). As shown in Figure 6.3.2, we see that because of GW190814’s extraordinary mass ratio, $\rho_{pNP}$ peaks at significantly larger values than average, with $\rho_p$ also entirely contained within the $1\sigma$ uncertainty. This means that although $\rho_p$ is large for this system, it is still consistent with originating from a non-precessing binary. This explains why $D_{JS} = 0.03$ bits.

6.3.2 GW190519 and GW190929

GW190519 and GW190929 have $\rho_p$ consistent with noise. In order to assess the significance of the higher multipole content in GW190519 and GW190929, we are interested in comparing their $\rho_{33}$ posteriors to distributions that could be obtained from noise alone, i.e. in the absence of any true signal power in the (3,3) multipole.
Chapter 6. Precession and higher order multipoles in the first half of the third gravitational wave observing run

Figure 6.3.3: $\rho_{33}$ for Left: GW190519 Right: GW190929. The blue solid lines are the publicly available posteriors obtained with a waveform model that includes higher order multipoles. Solid orange lines indicate posteriors obtained with models that include only the dominant harmonic. Treating this as a prior, we multiply by the likelihood modelled as a non-central chi distribution with two degrees of freedom and non-centrality parameter $\rho_{33}^{MF} = 2.1$ to obtain an estimate of a posterior that could be obtained when the power in the (3,3) multipole is noise alone. These estimates are shown as the grey dashed lines.

As described in section 6.2 to evaluate the prior probability associated with $\rho_{33}$ we can inspect the posteriors obtained with waveform models that include only the dominant harmonic. These informed priors are shown in Fig. 6.3.3 as solid orange lines. Prior to any (3,3) measurement, each event has a general preference for face-on and equal mass parameters consistent with lower values of $\rho_{33}$. Treating the orange distribution as a prior, we multiply by a likelihood of the form in Eq. 6.2 and non-centrality parameter $\rho_{33}^{MF} = 2.1$ to obtain an estimate of the (3,3) posterior that would result from a draw from the noise with a p-value of 10%. This indicates an upper limit on a posterior that is consistent with noise. These are shown as the dotted lines in Fig. 6.3.3.

The measured posterior for GW190519 is shifted to slightly higher values compared to the noise distribution. In order to fit the noise distribution to the measured posterior results, we minimize $D_{JS}$ by varying $\rho_{33}^{MF}$. The best fit noise distribution is shown as a dashed line in Fig. 6.3.3, and corresponds to $\rho_{33}^{MF} \sim 2.9$, corresponding to a p-value of 1.5%. This suggests there is evidence for higher order multipoles in the GW190519 signal, though it is less significant than both GW190814 and GW190412. The extra likelihood from the (3,3) multipole results in narrower posteriors for mass ratio and inclination. The mass ratio is constrained more tightly between $q = 1.17 - 2.34$, compared with $q = 1.04 - 2.80$ from the (2,2) multipole alone. With the inclusion of higher harmonics, the inclination now peaks at edge-on, rather than face-on/face-off. What’s more, the improved constraints in inclination and mass ratio lead to better constraints on other properties of the system that are correlated such as $\chi_{eff}$, component and total masses, distance, redshift, source frame masses and polarization angle.

As can be seen in the right hand panel of Fig. 6.3.3, GW190929’s $\rho_{33}$ posterior
6.3. Results

Figure 6.3.4: Corner plot showing the inferred mass ratio and $\rho_{33}$ for the reanalysis of GW190521 by Nitz et al. [13] compared to the results in GWTC-2 [14, 15, 16]. Shading shows the 1σ, 3σ and 5σ confidence intervals. This plot was made with PESummary [11].

is slightly larger than the predicted posterior for noise with a p-value of 10%. The best fit $\rho_{33}^{MF} = 2.5$ is expected to arise from noise alone 4% of the time. Given that there are 38 events we repeat this procedure for, this p-value suggests inconclusive evidence for higher harmonics in the signal. Inclusion of higher harmonics does not improve the constraint on inclination, but does result in a slightly tighter constraint on mass ratio with $q = 1.16 - 6.77$ becoming $q = 1.34 - 5.28$.

6.3.3 GW190521

GW190521 is the first evidence of a new population of black holes that resist straightforward interpretation as supernovae remnants, with at least one black hole lying firmly in the pulsational pair-instability mass gap ($\sim 65 - 120 M_\odot$) [14, 15]. Nitz et al. [13] challenged this view, showing it is possible to obtain parameter estimates consistent with component masses that instead straddle the gap. They found that when using a uniform in mass-ratio prior, GW190521 is consistent with component masses $m_1 = 168^{+15}_{-61} M_\odot$ and $m_2 = 16^{+33}_{-3} M_\odot$ compared to $85^{+21}_{-14} M_\odot$ and $66^{+17}_{-18} M_\odot$ reported in Refs. [14, 15, 16]. Prior constraints on the mass ratio imposed by the original analysis [14, 15, 16] ruled out any possibility of sampling this region of the parameter space. It is possible, therefore, that GW190521 is consistent with largely
asymmetric masses – a region of the parameter space where the SNR in both the (3, 3) multipole \cite{10} and precession \cite{271} are expected to be large. We therefore include these alternative parameter estimates (hereafter denoted as GW190521 Nitz) alongside those from GWTC-2 and reported in Refs. \cite{14, 15, 16}.

The inferred value of $\rho_{33}$ using GWTC-2 parameter estimation results is $1.16^{+1.81}_{-1.17}$, consistent with gaussian noise. However the Nitz et al. analysis results in a much larger $\rho_{33} = 4.10^{+0.92}_{-1.27}$. We are unable to construct an informed prior for GW190521 because there are no publicly available samples that include only the (2,2) multipole. The median suggests a very significant measurement of the (3,3) multipole with a $p$-value of less than three parts in ten thousand. Figure 6.3.1 shows that GW190521 Nitz has the second largest $\rho_{33}$ for any candidate in O3a. As a fraction of the total SNR, this would be the most significant measurement of the (3,3) multipole to date. From Figure 6.3.4 we see that these large $\rho_{33}$ values are a direct consequence of the peak at high mass ratio. It is the extra likelihood from the measurement of the (3,3) multipole that is key to the Nitz et al. reinterpretation of GW190521 having component masses that instead straddle the mass gap. This motivates further investigation, and in particular a direct measurement of the matched filter SNR in the (3,3) multipole would provide useful insight on this event.

GW190521 had the largest inferred $\chi_p$ in GWTC-2: $0.68^{+0.26}_{-0.44}$. Surprisingly, GW190521 has a small $\rho_p$: $0.7^{+1.4}_{-0.6}$. We understand this because GW190521 is the largest mass event detected with LIGO/Virgo. This means that GW190521 is very short in duration: 4 cycles (2 orbits) within the detectors’ sensitive frequency band. Consequently, GW190521 is decomposed into two near parallel “precession harmonics” (with overlap $|O_{10}^{\text{prec}}| = 0.97^{+0.01}_{-0.03}$). This means that any power orthogonal to the dominant harmonic is small and $\rho_p$ is small by definition. Several explanations for the large $\chi_p$ have been suggested, including possible evidence for eccentricity \cite{333} and head-on collisions \cite{334}. Unfortunately we are unable to compare $\rho_p$ to $\rho_p^{\text{NP}}$ since there are no publicly available non-precessing samples for GW190521. We do not calculate $\rho_p$ for GW190521 Nitz, as most of the samples lie outside of the validity region of the two-harmonic approximation\footnote{Owing to the combination of a large mass ratio and the orientation of the primary spin, the opening angle for GW190521 Nitz is large: 94% of samples have $\tan(\beta/2) > 0.3$ compared to 3% for the GWTC-2 samples.}

### 6.4 Discussion

We have calculated the inferred precession SNR, $\rho_p$, and orthogonal optimal SNR in the higher multipoles, $\rho_{lm}$ (for $(\ell, |m|) \in \{(2, 1), (3, 3), (4, 4)\}$) for all binary black hole events in O3a. Comparing with predicted distributions expected from noise alone, we looked for significant excesses. In addition to the strong evidence for a (3,3) multipole previously reported for GW190412 and GW190814 \cite{9, 12}, we found evidence
for a (3,3) multipole in GW190519 and precession in GW190412. Observation of subdominant higher multipole and precessing harmonics in these signals improves the constraints on the properties of their source. The measurement of the (3,3) harmonic in GW190519 results in narrower constraints on mass ratio and breaks the inclination distance degeneracy to measure the binary orientation as peaked at edge-on. GW190412’s precessing harmonic results in the tightest constraint on $\chi_p$ away from zero to date. Finally, we have shown that a reanalysis of GW190521 by Nitz et al. [13] suggests a significant (3,3) harmonic. This additional signal content results in a second peak in the posterior at 10:1 mass ratio and calls into question GW190521’s astrophysical interpretation as having component masses in the pulsational pair-instability mass gap.

The method we have demonstrated here is straightforward, and clearly identifies the content of the observed signal that allow for improved parameter estimates. In the future we want to expand this method to do the following. First to calculate the subdominant higher multipole and precession SNR directly from the data. In principle this should enable the construction of predicted posteriors including the effects of precession and higher multipoles using posteriors computed with a simpler waveform model. A similar method to this has been suggested for higher multipoles [318]. There the authors demonstrated that reweighting posteriors inferred with a (2,2) only waveform model based on the full likelihood could result in posteriors that closely match those inferred with waveform models including higher order multipoles. Second, work is ongoing to calculate the power in the second polarization for all O3a events, as this can also be important in breaking parameter degeneracies [173].
Appendix

6.A Parameter estimation samples used for $\rho_{33}$

To estimate $\rho_{33}$ we have used publicly available [24, 23] parameter estimation results obtained assuming a waveform model that includes higher order multipoles. For most events we used SEOBNRv4PHM parameter estimation results. Table 6.A.1 lists the parameter estimation waveform model for the remaining events.
6.A. Parameter estimation samples used for $\rho_{33}$

<table>
<thead>
<tr>
<th>Event</th>
<th>Waveform model</th>
</tr>
</thead>
<tbody>
<tr>
<td>GW190413_13</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190514</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190412</td>
<td>IMRPhenomPv3HM</td>
</tr>
<tr>
<td>GW190503</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190521</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190527</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190727</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190731</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190803</td>
<td>NRSur7dq4</td>
</tr>
<tr>
<td>GW190814</td>
<td>IMRPhenomPv3HM</td>
</tr>
<tr>
<td>GW190521 Nitz</td>
<td>IMRPhenomXPHM</td>
</tr>
</tbody>
</table>

Table 6.A.1: Waveform model assumed for parameter estimation results used to calculate $\rho_{33}$ for each event in O3a. For brevity we omit events which used SEOBNRv4PHM, which was the most commonly used waveform model. All parameter estimation results are publicly available [24].
Chapter 7

Constraining seed black hole populations with third generation gravitational-wave detectors

7.1 Introduction

How do supermassive black holes form? With each discovery of a quasar at yet earlier times, this question gains renewed urgency [335]. There are now 3 detected quasars at redshift $z > 7.5$ [336, 337, 338]. The earliest of these at $z \sim 7.64$ is powered by an accreting supermassive black hole of mass $\sim 1.6 \times 10^9 M_\odot$ [338]. Accounting for the existence of such a massive object only $\sim 600 - 700$ million years after the Big Bang is a challenge. Eddington limited accretion would require a seed black hole of $\sim 10000 M_\odot$ at redshift 30 [338]. The origin of seed black holes is unknown and their existence is so far hypothetical (see Ref. [339] for a review). Gas is expected to primarily fuel their growth [340, 341, 342, 343], though mergers may also play a role. This raises the prospect of detecting their gravitational wave emission with future detectors.

Space-based LISA [45] will probe the merger of seed black holes with masses in the range $\sim 10^4 - 10^7 M_\odot$ across all cosmic ages. This may be complemented, at lower redshift, by observations from the most sensitive future X-ray missions such as Lynx [344] and Athena [345]. Lighter seeds, with masses in the range $\sim 10^2 - 10^4 M_\odot$ are targets for future ground based gravitational wave detectors the Einstein Telescope (ET) [346, 347] and Cosmic Explorer (CE) [348, 349]. Accretion of these systems produces electromagnetic radiation too faint to be detected by even the deepest future facilities [350]. Mergers, however, will emit gravitational waves detectable with next generation ground-based networks at a broad range of redshifts.
2 < z < 15 (see Ref. [17] and Fig. 7.3.1), providing a unique view of these objects at cosmic dawn.

In this chapter we review the prospects for detecting and measuring the properties of light black hole seeds with ET and CE. In section 7.2 we discuss results of simulations of supermassive black hole formation and the growth of seed black holes through accretion and mergers. We then explore their detectability with future detectors in section 7.3. A prominent feature of these signals is the observability of their higher multipole content. We discuss this and the potential consequences for their detection chances. Finally, we investigate the prospects for accurately measuring the properties of these systems in section 7.4.2.

7.2 Building a supermassive black hole: stars, seeds and binaries

In this section we summarise prescriptions made in the simulations of the formation of supermassive black holes (SMBHs) obtained with a semi-analytical code [351, 352, 353, 17], and presented in full in Ref. [17]. The history of three prototype SMBHs (∼ 10⁹M⊙) are reconstructed. Their properties correspond to the quasars J1148 [354], SDSS J2345+1104 (hereafter J2345) [355, 356], and PDS 456[357, 358] at redshift z = 6.4 near the epoch of reionization, z = 2 at the peak of star formation, and z = 0.2, respectively.

Each SMBH resides in a dark matter (DM) halo of mass 10¹³M⊙ [359]. These halos are decomposed into progressively less massive fragments, with masses following the Extended Press-Schechter halo mass functions [360]. At redshift 24 the minimum mass of a resolved halo is 10⁶M⊙. While at the quasar final redshift, halos of mass 10¹⁰M⊙ are resolved. Structures with masses smaller form the inter galactic medium (IGM) from which progenitor halos accrete mass.

Starting at redshift 24, star formation proceeds at a rate regulated by the radiative and chemical properties of the halo and with periods of enhancement during major halo-halo mergers (see equation (2.2) in Ref. [17]). Population III stars form in these early metal poor environments, with a top heavy Larson initial mass distribution [361] with masses in the range 10 − 300M⊙. Light seed BHs (10²M⊙ < M < 10⁴M⊙) form by the collapse of these stars. Heavy seeds (10⁴M⊙ < M < 10⁶M⊙) form as the result of direct collapse of proto-stars of mass 10⁵M⊙. As heavy seeds form in more contrived environments, they are rarer with a relative fraction of ∼ 1% [17]. Within each halo the growth of the heaviest black hole is tracked. The seeds grow by accretion of surrounding material. Accretion is Eddington limited and follows the Bondi–Hoyle–Lyttleton formula (see Eq (6) of Ref. [362]). Both directions see a flow of energy and material: material is ejected into the inter galactic medium from accreting black holes and supernovae winds;
the external inter galactic medium flows into the inter stellar medium enriching the galaxy, and feeding the black hole accretion. Additional free parameters are included in both the star formation and accretion rate. These are used to tune the simulation to result in a system matching the observed properties of the final supermassive black hole and its surrounding galaxy (while respecting the Eddington limit).

During major halo-halo mergers (i.e. those with halo mass ratios greater than 4:1) it is assumed the nuclear seed black holes promptly form a binary. While in minor mergers, the evolution of the lighter black hole is no longer tracked. Black hole binaries can also form in situ within a halo. Here we do not consider these binaries. How black hole binaries shrink to milli parsecs to start to emit gravitationally and finally merge is an open problem. In these simulations, the role of taxing the binary orbit at this final stage, where other mechanisms are less efficient [363], is ascribed to triple interactions. These interactions have a limited efficiency in triggering binary mergers [363], and commonly result in stalled binaries with the ejection of the third black hole.

Thus each binary is assumed to form a Keplerian orbit and then stall until a third halo merges. The result of this interaction is evaluated statistically based on post-Newtonian simulations of triple interactions presented in Ref. [364]. If the merger happens at all, the simplifying assumption is made that it is immediate. 10 possible merger tree history realizations are followed of each supermassive black hole. In what follows we show results for the merger tree realization which best matches the simulation averaged predictions.

### 7.3 Observing seed black holes

During the simulated assembly of J1148 by $z = 6.4$, a total of 4228 light and 39 heavy seeds form. From these, 147 binaries are made and 70 are driven to merger by triple interactions. In our prescription, a binary promptly forms upon a halo-halo merger. This implicitly assumes that dynamical friction causes the black holes to pair within the halo-halo merger timescale of the simulations (around 500 million years) [365]. However, the timescale for dynamical friction can be much longer for black hole binaries with very unequal masses [366]. Further, the gravitational wave emission is not as strong for these binaries; we find that, averaged over orientation, only binaries with a mass ratio greater than 10:1 are detected in ET with an SNR greater than 10. For this reason we include only the binaries with mass ratio less than 10:1 in our analysis henceforth. 24 merging binaries meet this criteria for the J1148 quasar at $z = 6.4$.

A similar number of seed black holes form in the quasars J2345 and PDS456\(^1\), though a larger fraction pair and eventually merge due to the larger number of halos

\(^1\)5327 (5319) light and 31 (40) heavy seeds form respectively.
Observing seed black holes

7.3. Observing seed black holes

and halo-halo mergers: 45 and 84 mergers with mass ratios less than 10:1 for J2345 at $z = 2.0$ and PDS456 at $z = 0.2$ respectively.

Figure 7.3.1 shows the total mass and redshift of the mergers with mass ratios less than 10:1. Triangles, squares and circles denote mergers in the assembly histories of J1148, J2345 and PDS456 (represented by stars on the plot). The contours indicate the expected signal-to-noise ratio in ET, and LISA for a non-spinning system of mass ratio $q = 0.5$ (corresponding to the mean value from our simulations) averaged over orientation and sky angles. Mergers involving heavy seeds (shown as shapes with white edges) are loud sources for LISA, with SNRs in the range $10 - 1000$. Light seeds of several $10^2 M_\odot$ are possible sources for ET with SNRs in the range $10 - 50$ at $6 < z < 15$. A subset of light seeds that grow to have masses in the thousands will also be targets for LISA. For the longer assembly histories in the J2345 and PDS456, there is a large population of "starved" binaries in the more recent universe $2 < z < 8$, which are also accessible to ET.

In Figure 7.3.2 we plot the different flavours of seeds along with sensitivity limits for the future EM observatories Lynx and Athena (see Ref. [17] and the caption for details of how this was computed). Accretion onto heavy seeds are potential sources for EM observatories. Masses of $\sim 10^5 M_\odot$ will be observable at redshifts as high as 15 with Lynx. Light seeds, on the other hand, are likely too faint to be observed even with the deepest future facilities. ET therefore provides a unique opportunity to view this population of black holes forming at cosmic dawn.

The gravitational waveform observed at a detector is redshifted by the expansion of the universe, with its frequency evolution determined by the redshifted mass $M_{\text{total}} = M_{\text{source}} (1 + z)$. Consequently, the mass range observable by ET’s limited bandwidth depends on the redshift. In the low redshift universe ET is most sensitive to $M_{\text{source}} \sim 10^2 M_\odot$ binaries. Above redshift 10, only $M_{\text{source}} \sim 10^3 M_\odot$ binaries are detectable. As is clear from Figs. 7.3.1 and 7.3.2, light seeds densely populate the upper right hand side of the ET waterfall contours, where the redshifted masses are very high. This means that the majority of seeds that are not observable in ET are simply redshifted out of band. Those light seeds that are observable in ET have masses in the range $M_{\text{det}} \sim 1000 - 8000 M_\odot$, and waveforms characterized by just a few cycles.

The left panel of Figure 7.3.3 shows the waveform of a non-spinning $100 - 100 M_\odot$ binary at redshift 10, computed with SEOBNRv2 [19]. Due to redshifting, this binary is observed with a detector frame mass of $2200 M_\odot$, corresponding to the least massive quartile of the seeds in our simulations. We see that only a few cycles are visible, even with a detector with good sensitivity down to 3 Hz. A sharp cutoff in detector sensitivity at 7 Hz would mean only the final ringdown is above the detector noise. While above $\sim 8$ Hz the signal is no longer observable.

---

2Computed with the IMRPhenomC waveform [367] and the PSDs provided in Refs. [368] and [369].
Figure 7.3.1: Distribution of BBH coalescence events in the redshift $z$–$m_{BH,T}$ diagram. Data points describe cosmologically-driven BH mergers with mass ratio $q \geq 0.1$, triggered only by triple interactions among galaxy halos. Grey triangles, blue squares and red circles denotes the total mass and redshift of the coalescences extracted from the simulation of a $10^{13} M_\odot$ over-density, forming a $\sim 10^9 M_\odot$ SMBH at $z_{QSO} = 6.4$, 2 and 0.2 (represented with stars in the plot). Symbols with white edges indicate mergers involving at least one heavy seed. Color-coded areas represent lines of constant S/N ratios for ET (yellow/red) and LISA (azure/blue) computed for non spinning binaries assuming a mass ratio $q = 0.5$, which corresponds to the mean value of the merging binaries extracted from our samples. The ensemble of the color-coded areas for a given detector is often referred to ”waterfall” plot and provides averaged values of the S/N ratio at which a GW source is detected. Reproduced from Ref. [17].
7.3. Observing seed black holes

Figure 7.3.2: The GW and EM landscape. Color-coded areas give the average GW horizon computed for a detection threshold equal to $S/N = 10$: contour lines refer to binaries with mass ratios $q = 1, 0.5, 0.3, 0.1$ both in the ET and LISA bandwidth. Upper limits (shown as thick horizontal bars) indicate the sensitivity of the deepest pointing, in the $[0.5 - 2]$ keV observed band, by Athena (orange) and Lynx (black) given the limiting fluxes of $2.4 \times 10^{-17}$ and $10^{-19}$ erg s$^{-1}$ cm$^{-2}$, respectively. The upper limits are inferred assuming that BHs are emitting at the Eddington limit and adopting a bolometric correction ($L_X/L_{bol}$) of 10%. Ellipses highlight the islands in the $z - m_{BH,T}$ plane where light (blue) and heavy (white) seeds are expected to form as well as where light seeds are expected to grow via accretion and mergers (yellow). The transit to the SMBH domain covers the entire LISA area and EM observations are key to discover the high-mass tail of the SMBH distribution. The light-grey ellipse below $z \sim 5$ marks the population of long-living "starved" seeds. Note that in this island, coordinated multi-band observations are possible having LISA the capability to first follow the early inspiral in intermediate-mass black holes and ET the merger phase, enhancing the ability to carry on precise measurements of the source parameters also at $z \sim 5$ [18]. The islands have overlap with the GW horizon, but an empty inaccessible region is present between ET and LISA, corresponding to the Deci-Hz GW domain. The island corresponding to the stellar realm is included, on the left, for comparison. Reproduced from Ref. [17].
Chapter 7. Constraining seed black hole populations with third generation gravitational-wave detectors

Figure 7.3.3: 100-100 $M_\odot$ binary at redshift 10 (i.e. with detector frame total mass of 2200 $M_\odot$) Left: time domain at different frequencies. Computed with the model SEOBNRv2 [19]. Right: Frequency domain for (2,2) and (4,4) multipole inclined at 40 degrees, computed with IMRPhenomHM [20]. Normalized such that area between the waveform and noise is indicative of SNR [21]. A version of the time domain plot appeared in the science case document for third generation detectors [22].

An important omission from the above discussion are higher order multipoles, which are not included in the SEOBNRv2 waveform model. Figure 7.3.3’s right panel shows the frequency domain plot of the same system inclined at an angle of 40 degrees. The (2,2) and (4,4) multipoles are plotted separately using IMRPhenomHM [20]. As this is an equal mass system, the (3,3) multipole is zero. Also shown are the sensitivity curves for ET and CE [1]. This is a visual representation of the contribution of each harmonic to the total SNR, where the area between the noise and the waveform is proportional to signal-to-noise ratio. The (4,4) harmonic is above the noise in both ET and CE, and merger occurs at roughly twice the frequency of the (2,2) multipole. This highlights the prospect that the (3,3) and (4,4) multipoles, with merger frequencies approximately at 3 and 4 times the orbital, may still be observed even if the (2,2) is out of band.

To test this, we calculate the SNR in the (3,3) and (4,4) harmonics for each of the light seed black hole binaries in ET, averaged over orientation and sky position. For a binary to be considered detected we require an SNR greater than 8. An additional harmonic is detected when the SNR in that harmonic is greater than 2.1 [10]. We find that, due to the heavy detector frame masses, 98% of detected binaries also have a detected subdominant harmonic. Only systems viewed very close to face-on or face-off had higher multipole content too small to be clearly detected. This is in stark contrast to stellar mass binaries observed in current detectors where the fraction of detected binaries with measurable higher order multipoles is about 5% [10]. Current matched-filter gravitational-wave searches only search for the (2,2) harmonic [284,
A search also incorporating the (3,3) and (4,4) harmonics would increase the fraction of binaries having a total SNR above the detection threshold by a factor of 1.8. In fact, of the seeds detected with the (2,2) mode only, about 2 in 5 have a louder signal in either the (3,3) or the (4,4) harmonic, while almost 1 in 5 have a louder signal in both "subdominant" harmonics. Among the set of binaries that are detected with SNR greater than 8 in a higher multipole search, the fraction with a larger SNR in either the (3,3) or (4,4) multipole increases to two-thirds, with 2 in 5 having both harmonics louder than the (2,2).

Figure 7.3.4 shows the relative power in the (3,3) and (4,4) harmonics ($a_{lm}$, see Chapter 4) as a function of mass ratio and detector-frame total mass in ET. The dashed contour shows the point at which the harmonic is detected with SNR greater than 2.1 (approximately equivalent to a p-value of 10%), assuming SNR 20 in the (2,2) harmonic. Solid lines denote approximate 5-sigma significance for an additional harmonic after observing the (2,2) multipole with SNR 20. We see that for binaries with detector frame mass greater than 2000, the (4,4) harmonic is always observable, and often with high confidence. The dotted contours show the corner of parameter space for which the higher harmonics are equally significant to the (2,2) harmonic: at large, unequal masses. The light seed black holes observable in ET have detector frame masses in the range $1000 - 8000 M_\odot$, extending to masses beyond what is shown in this figure. The abundance of observable signal power in higher harmonics for light seed black holes has important implications not just for their detection, but also for parameter estimation. We discuss this in the next section.

## 7.4 Parameter estimation

Measuring the properties of light seed black holes, and in particular their masses and redshift, is important if we are to accurately determine their formation history. As discussed above, the large detector frame masses of these binaries means they are only in band for just a few cycles, making precise estimation of parameters challenging. Observation of the dominant harmonic oscillating at twice the orbital frequency typically results in broad constraints on both mass ratio and the inclination angle between the orbital angular momentum and line of sight. This latter parameter is particularly important as it is degenerate with redshift (see Chapter 3) and therefore also source frame masses. Usefully, higher harmonics have a different dependence on the inclination angle and can break the inclination-distance degeneracy, reducing uncertainties on both redshift and source-frame masses. Precession can in principle further improve parameter estimates. However given that so few cycles of the wave-
Figure 7.3.4: Ratio of the intrinsic amplitude, $\alpha_{\ell m}$, of signal harmonics to the (2,2) harmonic as a function of the total (detector frame) mass and mass ratio of the system, in the Einstein telescope. Top: the (3,3) harmonic; Bottom: the (4,4) harmonic. In all cases, the spins of the black holes are set to zero. The solid white line corresponds to $\alpha_{\ell m} = 5.3/20$ and the dashed line to $\alpha_{\ell m} = 2.1/20$, which correspond, approximately, to the threshold for the higher harmonics being confidently/marginally observable for a signal with SNR=20 in the (2,2) multipole. Dotted lines are at $\alpha_{\ell m} = 1$. 

---

- 117 -
7.4. Parameter estimation

Figure 7.4.1: Progression of the cos \( \iota \)-distance probability distribution for a \( 100M_\odot - 100M_\odot \) binary black hole at redshift \( z = 10 \) detected overhead the Einstein Telescope. The white star represents the injected signal. The top row shows the inference with a waveform model that includes only the dominant (2,2) multipole. The bottom panel shows the distribution inferred when the (4,4) multipole is also included. The leftmost column show the distribution for the un-marginalized likelihood, generated by calculating the SNR fall-off over the parameter space. Since we have not yet marginalized over the phase \( \phi_0 \) and polarization \( \psi \), these angles are taken to be equal to the values used in the injection, in this case zero for both \( \phi_0 \) and \( \psi \). The middle column show how these distributions change when marginalizing over \( \psi \) and \( \phi_0 \). Lastly, the rightmost column show the posterior distribution, calculated by multiplying the marginalized likelihood by the prior. We use a uniform in cos \( \iota \) and comoving volume prior. The white contours represent the 50% and 90% confidence intervals obtained from our simplified model. From these plots, we can see that including the (4,4) multipole in the waveform model breaks the degeneracy between distance and inclination, resulting in a posterior peaked at the correct value and excluding a face-off orientation.

form are visible in ground based detectors, the prospects for observing precession are slim [271], and in what follows we neglect precession effects.

To explore the general features of parameter estimation for light seed binaries, we perform parameter estimation on our fiducial non-spinning seed black hole binary with component masses of 100-100 \( M_\odot \) at redshift 10. In section 7.4.1 we will investigate the impact of higher order multipoles on the inclination distance degeneracy. We then present results of full parameter estimation in section 7.4.2.

7.4.1 Impact of higher order multipoles on inclination distance degeneracy

Here we are interested in the impact of higher-order multipoles on the inclination distance degeneracy. In Ref. [173] and Chapter 3 we demonstrated that the main features of this degeneracy can be captured by allowing just four parameters to vary (luminosity distance \( d_L \), inclination \( \iota \), polarization angle \( \psi \), and coalescence phase
φ₀) and fixing the rest. It is straightforward to extend this method to include the next most significant higher order multipoles after the dominant (2,2) multipole: (3,3) and (4,4). Apart from the additional terms in the gravitational wave polarizations, the method is identical to that outlined in Ref. [173] and Chapter 3. In this case, the gravitational wave polarizations are given by

\[ h_\times(t) = A_{22}^2 h_{22}^0(t) + A_{33}^3 h_{33}^0(t) + A_{44}^4 h_{44}^0(t) + A_{22}^2 h_{22}^0(t) + A_{33}^3 h_{33}^0(t) + A_{44}^4 h_{44}^0(t). \] (7.1)

\[ h_\times(t) = A_{22}^2 h_{22}^0(t) + A_{33}^3 h_{33}^0(t) + A_{44}^4 h_{44}^0(t). \] (7.2)

The \( h_{lm}^{0/\pi} \) are the waveforms for each of the harmonics. In general they are close to orthogonal. However, their relative amplitudes \( \alpha_{lm} \) (see Chapter 4 or Ref. [10] for definition) vary with mass ratio and total mass, as shown in Fig. 7.3.4. The amplitude parameters \( A_{lm}^i \) are [198, 199, 173]:

\[ A_{lm}^1 = A_{lm} \cos m\phi_0 \cos 2\psi - A_{lm}^\times \sin m\phi_0 \sin 2\psi \] (7.3)

\[ A_{lm}^2 = -A_{lm}^\times \cos m\phi_0 \sin 2\psi - A_{lm}^\times \sin m\phi_0 \cos 2\psi \] (7.4)

\[ A_{lm}^3 = A_{lm}^\times \sin m\phi_0 \cos 2\psi + A_{lm} \sin m\phi_0 \sin 2\psi \] (7.5)

\[ A_{lm}^4 = -A_{lm}^\times \sin m\phi_0 \sin 2\psi + A_{lm}^\times \cos m\phi_0 \cos 2\psi, \] (7.6)

with

\[ A_{lm}^{22} = \frac{d_0}{d_L} \frac{1 + \cos^2 \iota}{2} \] (7.7)

\[ A_{lm}^{22} = \frac{d_0}{d_L} \cos \iota. \] (7.8)

\[ A_{lm}^{33} = 2 \sin \iota A_{lm}^{22} \] (7.9)

\[ A_{lm}^{33} = 2 \sin \iota A_{lm}^{22} \] (7.10)

\[ A_{lm}^{44} = 2 \sin^2 \iota A_{lm}^{22}. \] (7.11)

\[ A_{lm}^{44} = 2 \sin^2 \iota A_{lm}^{22}. \] (7.12)

In Figure 7.4.1 we see an example of parameter estimation for our fiducial system at an inclination of 40 degrees⁴ overhead ET, with (top row) and without (bottom row) higher multipoles. The injected value is denoted by a white star. The leftmost column shows the unmarginalized likelihood at the injected values. Marginalizing over phase and polarization angle produces the likelihood in the middle column. The dominant multipole measurement only constrains to a combination of \( \phi_0 \pm \psi \) when the inclination is close to face-on (see Chapter 4 Fig. 4.4.2), but by including the (4,4) multipole this degeneracy is broken. For this reason, rather than pushing.

⁴The inclination here is folded between 0 and \( \pi/2 \). Equivalent to 140 degrees, or \( \cos \iota = -0.77 \)
7.4. Parameter estimation

Figure 7.4.2: Inferred posteriors for the (clockwise from top left) redshift $z$, mass ratio $q < 1$, total mass in source frame $M_{\text{total}}$, and inclination $\cos \theta_{JN}$ of a non-spinning $100 - 100M_\odot$ binary black hole at redshift 10. The signal is detected with SNR $\sim 67$ by a network including ET located at the site of Virgo, and two Cosmic Explorers located at LIGO India and LIGO Livingston. We compare posteriors inferred with waveform models containing the dominant multipole (IMRPhenomPv2) with those inferred with additional higher multipoles (IMRPhenomHM). Dashed vertical lines indicate the 90% credible bounds for each waveform model.

the likelihood towards face-on where there is a greater volume of parameter space, marginalizing over phase and polarization now results in a marginalized likelihood peaked at the correct value. The final column shows the posterior: the product of the prior and the marginalized likelihood. Priors are chosen to be uniform in $\cos \iota$ and comoving volume\(^5\). We see that when the (4,4) multipole is included, the system can now be recovered at the correct inclination with the posterior excluding a face-on orientation.

7.4.2 Results from full parameter estimation

Figure 7.4.2 shows the posteriors obtained with full parameter estimation\(^6\). We show the inferred posterior for a dominant multipole model, IMRPhenomPv2 [309], and IMRPhenomHM [20] which has higher multipoles including the (3,3) and (4,4). We

\(^5\)the prior on distance prefers nearby distances, since there is a greater comoving volume there at high redshifts

\(^6\)Obtained with the parameter estimation code LALInference [200], modified to have a uniform in comoving volume prior.
Figure 7.4.3: Posterior distribution for the component masses of the fiducial seed black hole binary described in Fig 7.4.2’s caption. We illustrate the 90% credible regions. The red cross indicates the injected values. Comparing IMRPhenomPv2, a model which only includes the dominant multipole with one including higher multipoles, we see the great impact higher harmonics have for constraining the source frame masses.

see a similar improvement in the inclination angle $\theta_{JN}$ due to the extra likelihood from the (4,4) multipole, with the posterior now tightly peaked at the correct value\(^7\). This has a domino effect across the parameter space, resulting in a large reduction in uncertainty on the redshift, and therefore source-frame masses. The span of the 90% credible bounds for source frame total mass reduces from $\sim 60\%$ of the value, to $\sim 10\%$ with the inclusion of the higher multipoles. The redshift 90% limits for the dominant multipole roughly span from $z = 6 - 12$, corresponding to either a binary observed merging when the universe was less than 400 million years old or more than 900 million years old by which point several supermassive black holes have formed. Higher harmonics reduce this ambiguity, constraining the merger time to within 100 million years of its injected value. A further improvement is seen in the measurement of the mass ratio, the other key parameter determining the strength of the higher harmonics. The absence of a (3,3) multipole in the data despite an inclination constrained away from face-off causes unequal masses to be highly disfavoured. The combined improvements in the measurements of the mass ratio and source frame total mass results in a great reduction in the range of black hole

\(^7\)For precessing models the angle between the total angular momentum and line-of-sight $\theta_{JN}$ is not necessarily the same as the angle between the orbital angular momentum and the line of sight $\iota$. For a non-precessing system, we do not expect this to make a substantial difference to the inferred posterior, and so neglect it here.
component masses consistent with the data. This is shown in Fig. 7.4.3. Here the primary mass 90% credible bound span reduces from $\sim 140\%$ of the injected value, to $\sim 10\%$. These vastly improved parameter estimates due to higher harmonics, though only demonstrated for a single binary, suggest it may be possible to get precise parameter estimates for other seed black holes with significant higher harmonic content. This increases the prospects for distinguishing and studying the history of this population with fewer detections.

7.5 Discussion

We have demonstrated that seed black holes will be targets for future ground and space gravitational wave networks. A subset of these seeds are only detectable with ground based detectors, and especially ET which has good low frequency sensitiv-
ity. We find that higher harmonics ringing at 1.5 and 2 times the frequency of the dominant will be observable for the majority of these systems, allowing a greater fraction to be detected. These additional harmonics also result in tighter constraints on both mass ratio and inclination angle. This improves the prospects for differentiating light seed black hole candidates from black hole mergers of stellar origin. In future work we will demonstrate the expected parameter uncertainties for a broader sample of seeds, and explore more completely the consequences this has on our ability to distinguish populations.
Chapter 8

Discussion

The first half of the third observing run of the LIGO-Virgo gravitational-wave network resulted in confident detections of $\sim 40$ compact binary coalescences \[16, 32\]. Soon results from the second half of the observing run will uncover further candidates. Data collection will resume with the onset of the fourth observing run circa 2022 \[370\]. KAGRA will join the LIGO-Virgo network, which will be approaching its design sensitivity. The expected binary neutron star surveyed volume will increase by a factor of $\sim 3.8$ \[370\], suggesting that we can expect almost daily detections of gravitational waves from colliding compact objects. Before the decade is out, Voyager may usher in the era of overlapping signals from compact binaries \[371\]. Finally, in the decades that follow, third generation networks will detect the majority of binary black hole mergers in the observable universe \[22\].

Part of this thesis has been about understanding what is measurable in a gravitational-wave signal, and how this impacts our ability to determine the properties of black holes and neutron stars. In chapter 3 we showed that measuring the luminosity distance requires distinguishing the two gravitational-wave polarizations, which will not be possible for the majority of gravitational waves detected in binary neutron star mergers. Chapter 4 described how higher harmonics can break the degeneracies in the parameter space, and demonstrated a simple criterion for their observation. Refs. \[275, 276, 271\] describes a similar criterion for the observation of precession. These viewpoints can be useful in the development of searches that include higher harmonics and precession. Upon the detection of a binary merger with parameters \textit{a priori} likely to have measurable higher or precessing harmonics, the data can be filtered for these additional harmonics. Another future application is to use the observation of the second polarization and precessing and higher harmonics to predict the posterior probability distribution for the properties of compact binaries. With ever increasing sensitivity, and detections, the need to produce posteriors rapidly and with minimal computational expense becomes more pressing. Considerable effort has been expended on these fronts, yielding a broad variety of techniques \[372, 373, 374, 375, 376, 377, 378, 132, 379, 380, 381, 382, 383, 384, 385, 386, 387\].
This work will complement these efforts, while also providing a simple way to interpret parameter estimation results.

In chapter 5 and 6 we quantified the evidence for higher and precessing harmonics and used this to interpret parameter estimation of events in O3a. This thesis also aims to inform the science case for future gravitational-wave observatories. This was done in chapter 3 where we demonstrated that uncertainty in inclination may be the dominant source of uncertainty in the source frame masses for observations of dominant harmonic emission with future gravitational-wave detectors. Chapter 7 showed this is unlikely to be the case for a population of black holes in the early universe that are hypothesized to grow to become supermassive black holes, due to their higher harmonic signal content. Finally, chapter 2 presented detailed calculations of the localization capabilities for various configurations of future gravitational-wave networks. In addition to providing scientific justification for future gravitational-wave detectors, these calculations will guide astronomers in the wider community seeking to fine tune their own scientific priorities.
Bibliography


[261] Roberto Cotesta, Sylvain Marsat, and Michael Purrer. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with higher-order modes. 3 2020.


[293] Sylvain Marsat and John G. Baker. Fourier-domain modulations and delays of gravitational-wave signals. 6 2018.


