Gravitational waves: A statistical autopsy of a black hole merger

Renate Meyer and **Nelson Christensen** explain how statistics and statisticians helped unravel a story 1 billion years in the making

The detection of gravitational waves was a landmark moment for science, and a testament to both theoretical physics and experimental physics.[1] These "ripples" in the curvature of spacetime were first predicted by Albert Einstein in 1916, but it took the technological heft and spectacular performance of the Laser Interferometer Gravitational-Wave Observatory (LIGO) to confirm their existence.

Physicists are rightly celebrating their achievement. But, behind the scenes, statisticians had a part to play. Bayesian parameter estimation methods, specifically Markov chain Monte Carlo (MCMC) techniques, had a crucial role in extracting the parameters from the gravitational wave signal that could tell the story how, 1.3 billion years ago, two black holes (with masses estimated to be 29 and 36 times the mass of our sun, which we measure as 1 solar mass, $M_\odot$) spiralled into one another. These two objects collided with a relative speed of 0.6 times the speed of light, and in the process formed a new back hole with a mass of $62M_\odot$, which radiated away gravitational waves with a total energy equivalent to $3M_\odot$.

It was these gravitational waves that were detected by LIGO, and its two L-shaped laser-based detectors – one in Hanford, Washington, the other in Livingston, Louisiana.[2] As the wave passed through the detectors, it subtly altered the length of the two arms of the detectors, causing the laser beams to travel different distances – a difference that could be detected and measured.

LIGO was designed to detect a gravitational-wave strain, h, the difference in lengths relative to the total arm length, of amplitude $h \sim 10^{-21}$; a wave of this amplitude would create a difference in length between the two arms of order 10^{-18} m. This is the equivalent of a ruler so precise that it can measure a 1 000 000 000 000 000-metre-long stick to within 1 millimetre. How is it possible to measure such a small distance change? Every photon from the laser acts like a metre stick, and through the repeated measurements of all the photons it is possible to get an ensemble average and a value for that small distance displacement.

Renate Meyer is associate professor of statistics at the University of Auckland, New Zealand. Her research interests comprise applied Bayesian inference, MCMC techniques, time series and survival analysis. She has collaborated with Nelson Christensen on parameter estimation for gravitational waves

Nelson Christensen is the George H. and Marjorie F. Dixon Professor of Physics at Carleton College, Minnesota. He is a member of the LIGO Scientific Collaboration and conducts searches for gravitational waves from the Big Bang, detector noise studies, and parameter estimation

Credit: SXS

A data analysis task

Nowadays, the work of astrophysicists and cosmologists is not just about detecting a signal, but using that signal to describe the physical processes that occurred in our universe billions of years ago. Over the last 20 years, Bayesian parameter estimation methods have become the norm in helping to explain those processes. With the LIGO project, the ultimate goal is to extract a potential signal embedded in noise, and to estimate the parameters of the signal waveform that characterises the astrophysical event that created the gravitational wave. This is no doubt a data analysis task, and parameter estimation has been the focus of a dedicated group among the thousand-plus physicists, astronomers, and engineers who worked on the project. The combined efforts of all members of the parameter estimation group within the LIGO Scientific Collaboration and the Virgo Collaboration (LVC) made it possible to achieve this goal.

Before we go further, however, it is helpful to clearly distinguish between the detection of gravitational waves and the subsequent analysis of their measurements. Whereas frequentist chi-square test statistics were used to confirm the detection of a gravitational-wave signal and to calculate an (exceedingly small) false alarm probability of less than 2×10^{-7}, Bayesian inference played, and is playing, an important role in the follow-up analysis.[3] The independent measurement of the same signal at both LIGO detectors, with a difference in arrival time of 7 milliseconds, is important not only to confirm the detection with great confidence but also to estimate the sky location of the event.

Once a gravitational wave has been detected, it is important to estimate the parameters of the signal waveform as accurately as possible and thereby home in on the astrophysical characteristics of the cosmic event that caused it. Parameter estimation made it possible to infer that what caused the detected gravitational wave was indeed the merger of two black holes.

In the detection phase, it is possible to ignore certain characteristics of the colliding black holes (generally referred

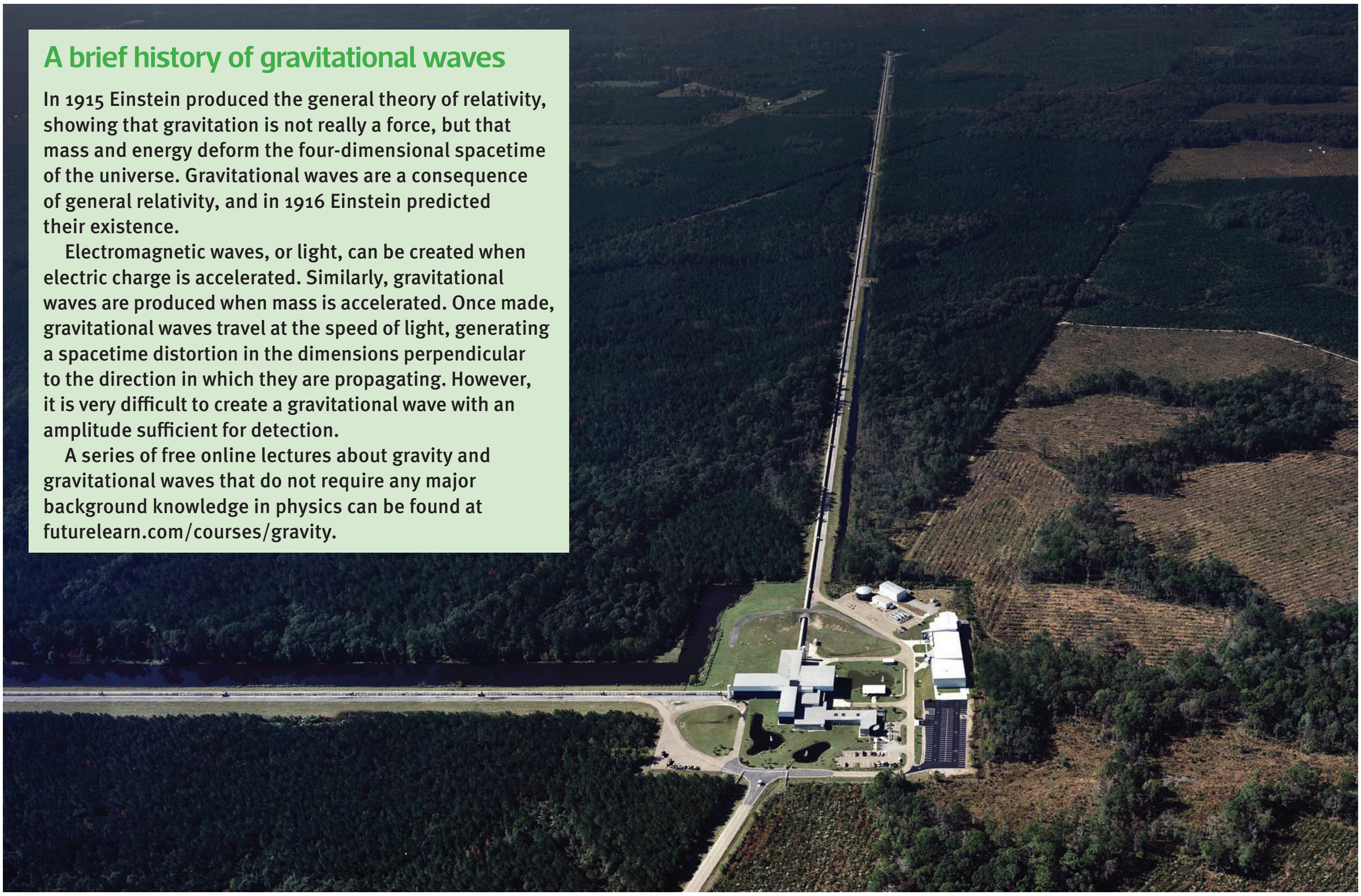
Physicists are rightly celebrating their achievement in detecting gravitational waves. But, behind the scenes, statisticians had a part to play

to as 'coalescing binaries') and take into account only the masses (in the range of 1–100$M_{\odot}$) of the binary system and the aligned spins, effectively discretising the parameter space into a grid of parameter combinations and trying all resulting 250 000 template waveforms in a matched-filter search procedure to check whether the observed waveform is close to any of the templates.

In the follow-up phase, though, an accurate signal waveform model for the gravitational waves is essential. We will discuss these waveforms in more detail below. Taking all 17 parameters that describe a chirp waveform – like the one detected – into account, using a similar discretisation, would result in an astronomical number of templates, effectively rendering an exhaustive search unfeasible. Even if feasible with an ubercomputer, this procedure would give no idea of the uncertainty of the estimates – and a point estimate is useless without an uncertainty range. This is where MCMC provides a clever and efficient way of stepping through a high-dimensional parameter space, delivering parameter estimates and quantifying their uncertainty.

Markov chain Monte Carlo

Monte Carlo methods were developed in the 1950s by a team of physicists led by Nicholas Metropolis.[4] MCMC failed to thrive in an era of slow electronic computers but had an enormous growth spurt when computer scientists discovered its use for digital image restoration in 1985,[5] and when statisticians in the 1990s realised that MCMC could be harnessed for Bayesian posterior computation.[6] Since then, MCMC has revolutionised applied statistics. The late 1990s was when our team started to look at parameter



BELOW Aerial photo of the LIGO detector at Livingston, Louisiana. Credit: Caltech/MIT/LIGO Laboratory

A brief history of gravitational waves

In 1915 Einstein produced the general theory of relativity, showing that gravitation is not really a force, but that mass and energy deform the four-dimensional spacetime of the universe. Gravitational waves are a consequence of general relativity, and in 1916 Einstein predicted their existence.

Electromagnetic waves, or light, can be created when electric charge is accelerated. Similarly, gravitational waves are produced when mass is accelerated. Once made, gravitational waves travel at the speed of light, generating a spacetime distortion in the dimensions perpendicular to the direction in which they are propagating. However, it is very difficult to create a gravitational wave with an amplitude sufficient for detection.

A series of free online lectures about gravity and gravitational waves that do not require any major background knowledge in physics can be found at futurelearn.com/courses/gravity.

estimation for coalescing binaries and suggested a Bayesian approach using MCMC for posterior computation.[7] This was taken up by the LVC, as well as the cosmologists, who embraced the Bayesian philosophy.[8–10] MCMC quickly became the standard approach to estimating gravitational[11] and cosmological parameters.[12]

Before the upgrade to Advanced LIGO in 2010, the original LIGO detectors were operating for eight years or so. During this time, research efforts in the LVC were focused on the detection phase, by developing a robust and sensitive detection pipeline based on matched filtering designed to keep up with the large amount of data. The detection methods were tested through a "blind injection challenge" in which a synthetic coalescing binary signal was added without the knowledge of the data analysis teams and successfully detected.

However, many studies also had the objective of developing parameter estimation techniques for the follow-up analysis phase, once a candidate had been detected. To this end, starting with very simple waveform models with only five parameters[13] – namely the masses of the two black holes, the coalescence time, the phase of the signal at coalescence, and the effective distance to the source – increasingly accurate astrophysical waveform approximations have been employed, and MCMC methods developed, to take these into account.

Unpicking the signal

To give an idea of the statistical models and techniques involved, let us take a closer look at estimating the parameters of the very first gravitational-wave signal ever detected by LIGO on 14 September 2015.[3] The aim of the parameter estimation group of the LVC was to find out what caused the signal, where the astronomical event occurred, and to further characterise the source in detail.

The only information from the detection stage that was used to infer these parameters was the signal arrival time. Figure 1 (page 24) shows the observed signal at the two LIGO detectors, overlaid by the predicted shapes for the waveform. These inferred waveforms show what two merging black holes should look like, according to the equations of Albert Einstein's general theory of relativity.

According to this theory,[14] two orbiting objects slowly spiral together because they lose energy from the emission of gravitational waves. As the two objects approach one another, the frequency and amplitude of the emitted gravitational waves increases. If the objects are black holes, they form a single perturbed black hole when they finally merge, which then emits gravitational waves at a constant frequency and exponentially damped amplitude.

The signal detected by LIGO matched nicely with the predictions of general relativity. Bayesian parameter estimation methods would then provide the means to know more about the black holes that caused these ripples.

Waveforms of two in-spiralling black holes are described by intrinsic parameters, the two masses (m_1, m_2) and spins (magnitude and orientation) of the individual black holes,

Technical notes: Estimating black hole parameters

Let us denote the vector of unknown waveform parameters by ϑ. The data $d(t)$ recorded at each detector is assumed to consist of the compact binary coalescence signal, the strain $h(t;\vartheta)$ plus additive noise:

$$d(t) = h(t;\vartheta) + n(t), \quad t = 1,\ldots,T$$

Strain represents the fractional amount by which distances are distorted. The strain at each detector consists of a linear combination of the two independent polarisation amplitudes h_+, $h_\times$ and the antenna response functions F_+, $F_\times$ that depend on the source location in the sky and the polarisation of the waves:

$$h(t;\vartheta) = F_+(t;\alpha,\delta,\psi)\, h_+(t;\vartheta) + F_\times(t;\alpha,\delta,\psi)\, h_\times(t;\vartheta)$$

There are a number of distinct waveform models $h(t;\vartheta)$ that differ in the post-Newtonian order of their approximation, their computational complexity, and their regime of applicability. Physicists are actively improving these models, none of which can describe all possible physical effects for all binary systems. In the case of the wave detected on 14 September 2015, a waveform model with 13 parameters was eventually used to characterise the coalescing black hole waveform.

The data $d(t)$ span the period $[t_c - T + 2,\ t_c + 2]$, i.e. a time $T (= 8\,\text{s})$ that ends 2 seconds after the trigger. The time series is usually down-sampled from its original sampling frequency (16 384 Hz) to a lower rate (typically 4096 Hz) and low-pass filtered to prevent aliasing.

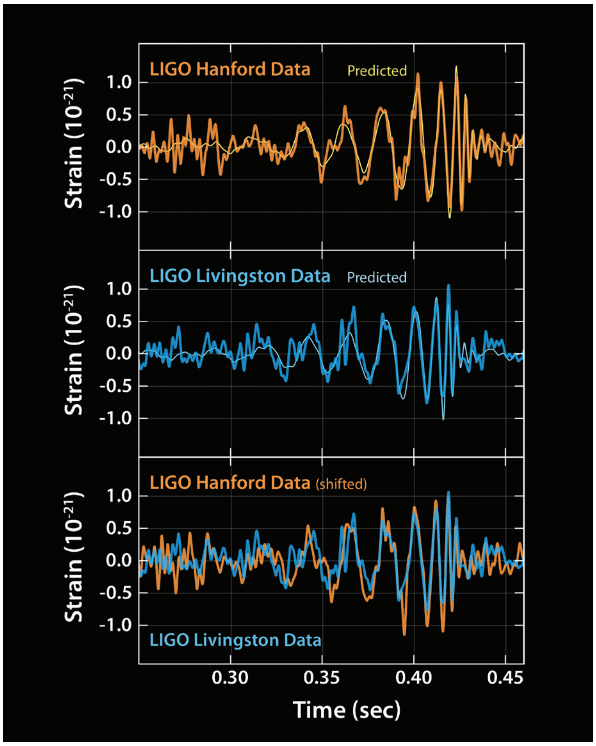
The statistical assumption on the noise time series $n(t)$ is that it is stationary and Gaussian. After Fourier-transforming the data, an approximate likelihood – the so-called Whittle likelihood[18,19] – is used. This has the big advantage that we no longer have to deal with dependent data because the discrete Fourier-transformed data are asymptotically *independent* Gaussian with variance equal to the power spectral density at the Fourier frequencies. So the (approximate) likelihood can be written as the product of Gaussians. The power spectral density is estimated beforehand from a stretch of data not containing the signal, so is effectively assumed to be known. Because of independence, the joint likelihood is the product of the individual likelihoods of the data at each detector.

Why not address this with the usual frequentist approach by maximising the likelihood? Given the large number of data points, an asymptotic Gaussian approximation of the maximum likelihood estimator might be justified and the Hessian matrix easily computed to yield an approximate covariance matrix. This would give an asymptotic confidence interval for each parameter estimate and thus describe the variability of the estimator over repeated sampling. What it does not do, and neither do more sophisticated frequentist bootstrap techniques, is quantify the uncertainty in the parameter estimate for this particular data set. Bayesian inference does just that by quantifying the uncertainty of a parameter estimate via a probability distribution on the parameter space. Starting with a prior distribution of ϑ, this distribution is updated via the likelihood after observing the data to the posterior distribution.

Current research looks at Bayesian non-parametric estimates of the power spectral density for locally stationary time series so that estimates of the waveform parameters ϑ can be obtained by properly marginalising over the power spectral density[20–22] rather than assuming it to be known.

> The aim of the parameter estimation group was to find out what caused the signal and where the astronomical event occurred

FIGURE 1 Gravitational-wave signals detected by the twin LIGO observatories at Livingston, Louisiana, and Hanford, Washington. The observed waveforms are overlaid by the predicted waveforms of two merging black holes according to the equations of Albert Einstein's general theory of relativity. Credit: LIGO Laboratory



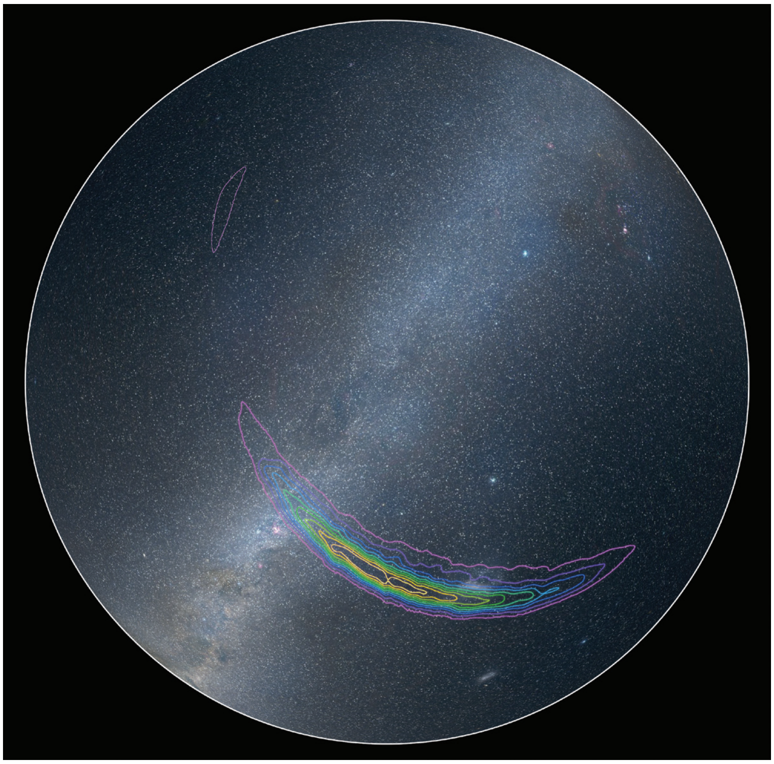


FIGURE 2 The approximate sky location in the southern hemisphere of the source of LIGO's gravitational-wave signals. The coloured lines represent different posterior probability regions for where the signal originated: the purple line defines the 90% level; the inner yellow line defines the 10% level. Credit: LIGO Laboratory/Axel Mellinger

as well as extrinsic parameters comprising the luminosity distance, right ascension α and declination δ (which describe the sky location), the binary's orbital inclination between the system's orbital angular momentum and the line of sight and its polarisation ψ, the time t_c and phase φ_c of coalescence, and the two eccentricity parameters of the system. Some of these are of particular interest, such as the masses of the black holes, the distance and sky location; others can be considered as "nuisance" parameters.

The technical details of the parameter estimation process can be found in the "Technical notes" box on page 23, but, in essence, the first step in the process was to make some prior assumptions about the likely parameters of the two black holes and to input these into the MCMC algorithms in order to produce a configuration of parameter estimates that closely approximated the signal detected by LIGO.

Uniform priors on 10–80$M_\odot$ were chosen for the two masses, with the constraint that m_2 would be less than m_1. The prior on the coalescence time t_c was uniform of width ±0.1 s, centred on the trigger time, and the prior on the polarisation angle φ_c was uniform on $[0,2\pi]$. In the absence of any additional astrophysical constraints on the source, prior choices for the other parameters were non-informative.

Here, the posterior distribution lives in a 13-dimensional parameter space. We cannot analytically integrate this distribution to obtain the marginal posterior distribution of each parameter of interest, such as the final mass. But MCMC techniques enable an easy sampling-based approach to marginalisation. For that purpose, a suite of state-of-the art MCMC methods[15] – such as adaptive Metropolis–Hastings, simulated annealing, delayed rejection, parallel tempering and nested sampling – have been implemented in the LALInference software library.[16] Figure 2, for instance, shows the two-dimensional marginal posterior distribution of the sky location of the source. Similarly, parameter estimates in Table 1 were obtained by MCMC marginalisation.

The Bayesian approach to parameter estimation combined with MCMC methods made it possible to estimate the parameters of the coalescing black hole merger waveform after its detection by LIGO and quantify the inherent uncertainties. It provided the means to infer important characteristics such as the individual masses of the black holes and the final mass, the distance to the Earth and the energy radiated in gravitational waves. This analysis assumed that general relativity is correct, and the waveform reconstruction in Figure 1 under this assumption gives no evidence against its validity. But Bayes factors have also been used to compare the relative fit of different waveform approximations derived by solving Einstein's equations so that various aspects of general relativity could be put to the test.[17]

This is only the beginning of the story of gravitational waves. Soon, the Virgo detector (near Pisa, Italy) will be operating in concert with LIGO, and with more than two detectors, located in different parts of the world, it will be possible to estimate the sky location of future events even more accurately. Then,

TABLE 1 Important parameters of the coalescing black hole waveform and their estimates (posterior medians) along with 90% credible range.[1]

Mass of initial black hole no. 1	36 (+5, −4) $M_\odot$
Mass of initial black hole no. 2	29 (+4, −4) $M_\odot$
Mass of final black hole	62 (+4, −4) $M_\odot$
Distance to the event	1.34 (+0.52, −0.59) billion light-years
Energy radiated in gravitational waves	3 (+5, −5) $M_\odot c^2$

The Bayesian approach, combined with MCMC methods, provided the means to infer important characteristics of the black holes

in 2019, the Japanese KAGRA detector should come online, followed by a third LIGO detector (this one in India) in 2023. In addition, the LISA Pathfinder spacecraft (sci.esa.int/lisa-pathfinder), launched in December 2015, will provide an important test of the feasibility of putting a gravitational wave detector in space.

Of course, the analysis of all the data – from LIGO-Virgo and LISA – will provide ample opportunities for statisticians to make significant contributions within the field of astrostatistics. As the late John Tukey once said: "The best thing about being a statistician is that you get to play in everyone's backyard." The backyard of the astrophysicist is vast and expanding. ■

References

1. Abbott, B. P. *et al.* (LIGO Scientific Collaboration and Virgo Collaboration) (2016) Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, **116**, 061102.
2. Aasi, J. *et al.* (LIGO Scientific Collaboration) (2015) Advanced LIGO. *Classical and Quantum Gravity*, **32**, 074001.
3. Abbott, B. P. *et al.* (LIGO Scientific Collaboration and Virgo Collaboration) (2016) Properties of the binary black hole merger GW150914. arXiv:1602.03840 [gr-qc].
4. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. and Teller, E. (1953) Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**(6), 1087–1092.
5. Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.
6. Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (eds) (1996) *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall.
7. Christensen, N. L. and Meyer, R. (1998) Markov chain Monte Carlo methods for Bayesian gravitational radiation data analysis. *Physical Review D*, **58**, 082001.
8. Christensen, N. L., Meyer, R., Knox, L. and Luey, B. (2001) Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements. *Classical and Quantum Gravity*, **18**, 2677–2688.
9. Schafer, C. M. (2015) A framework for statistical inference in astrophysics. *Annual Review of Statistics and Its Application*, **2**, 141–162.
10. Knox, L., Christensen, N. and Skordis C. (2001) The age of the universe and the cosmological constant determined from cosmic microwave background anisotropy measurements. *Astrophysical Journal Letters*, **563**, L95–L98.
11. Van der Sluys, M., Röver, C., Stroeer, A., Raymond, V., Mandel, I., Christensen, N., Kalogera, V., Meyer, R. and Vecchio, A. (2008) Gravitational-wave astronomy with inspiral signals of spinning compact-object binaries. *Astrophysical Journal Letters*, **688**, L61–L64.
12. Verde, L., Peiris, H. V., Spergel, D. N., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Wollack, E. and Wright, E. L. (2003) First-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) observations: Parameter estimation methodology. *Astrophysical Journal Supplement Series*, **148,** 195–211.
13. Christensen, N. and Meyer, R. (2001) Using MCMC methods for estimating parameters with gravitational radiation data. *Physical Review D*, **64**, 022001.
14. Einstein, A. (1916) Näherungsweise Integration der Feldgleichungen der Gravitation. *Preussische Akademie der Wissenschaften Berlin, Sitzungsberichte der physikalisch-mathematischen Klasse*, **1916**, 688–686.
15. Brooks, S., Gelman, A., Jones, G. L. and Meng, X. L. (eds) (2011) *Handbook of Markov Chain Monte Carlo*. Boca Raton, FL: Chapman & Hall/CRC.
16. Veitch, J., Raymond, V., Farr, B., Farr, W., Graff, P., Vitale, S., Aylott, B., Blackburn, K., Christensen, N., Coughlin, M., Del Pozzo, W., Feroz, F., Gair, J., Haster, C.-J., Kalogera, V., Littenberg, T., Mandel, I., O'Shaughnessy, R., Pitkin, M., Rodriguez, C., Röver, C., Sidery, T., Smith, R., van der Sluys, M., Vecchio, A., Vousden, W., Wade, L. (2015) Parameter estimation for compact binaries with ground-based gravitational-wave observations using LALInference. *Physical Review D*, **91**, 042003.
17. Abbott, B. P. *et al.* (LIGO Scientific Collaboration and Virgo Collaboration) (2016) Tests of general relativity with GW150914. arXiv:1602.03841 [gr-qc].
18. Whittle, P. (1957) Curve and periodogram smoothing. *Journal of the Royal Statistical Society*, *Series B*, **19**, 38–63.
19. Shao, X. and Wu, B. W. (2007) Asymptotic spectral theory for nonlinear time series. *Annals of Statistics*, **35**, 1773–1801.
20. Littenberg, T. B. and Cornish, N. J. (2014) Bayesian inference for spectral estimation of gravitational wave detector noise. *Physical Review D*, **91**, 084034.
21. Edwards, M., Meyer, R. and Christensen, N. L. (2015) Bayesian semiparametric power spectral density estimation with applications in gravitational wave data analysis. *Physical Review D*, **92**, 064011.
22. Cornish, N. J. and Littenberg, T. B. (2015) Bayeswave: Bayesian inference for gravitational waves and instrument glitches. *Classical and Quantum Gravity*, **32**, 135012.