Bayesian model comparison in gravitational wave data analysis

Massimo Tinto\(^1\) and Richard Umstätter\(^1\)

\(^1\) Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109
E-mail: Massimo.Tinto@jpl.nasa.gov

Abstract. We estimate the probability of detecting a gravitational wave signal from coalescing compact binaries in simulated data from a ground-based interferometer detector of gravitational radiation using Bayesian model selection. The simulated waveform of the chirp signal is assumed to be a spin-less Post-Newtonian (PN) waveform of a given expansion order, while the searching template is taken to be either of the same PN family as the simulated signal or one level below its PN expansion order. Within the Bayesian framework we estimate the detection probabilities and the statistical uncertainties due to noise as a function of the signal-to-noise ratio (SNR), and the posterior distributions of the parameters characterizing the chirp signal. Our analysis indicates that the detection probabilities are not compromised when simplified models are used, while the accuracies in the determination of the parameters can be significantly worsened.

1. Introduction

Ground-based interferometer detectors of gravitational waves have reached their design sensitivity, and have started to search \([1, 2, 3, 4]\), in the kilohertz frequency band, for gravitational waves emitted by astrophysical sources such as spinning neutron stars, supernovae, and coalescing binary systems. Coalescing binary systems in particular are expected to be the first to be detected and studied. These signals have a unique signature that enables them to be extracted from wide-band data by matched filtering \([5]\). This signature is their accelerating sweep upwards in frequency as the binary orbit decays because of energy loss due to the emission of gravitational radiation. Coalescing binaries have a potential advantage over other sources in signal-to-noise ratio (SNR) by a factor that depends on the square-root of the ratio between the corresponding number of cycles in the wave trains \([6]\).

The effectiveness of the matched filtering procedure relies on the assumption of exactly knowing the analytic form of the signal (possibly) present in the data. Recent breakthroughs in numerical relativity \([7, 8, 9]\) have started to provide a complete description of the radiation emitted during the inspiral, merger, and ring-down phases of generic binary black hole merger scenarios. Although the ability of obtaining numerically all the templates needed in a data analysis search (perhaps hundred of thousands of them) might be practically impossible, in principle we should be able to compare these numerically derived waveforms against various analytic templates obtained under different approximating assumptions. Work in this direction has already started to appear in the literature \([10, 11, 12]\) within the so called “frequentist framework”, in which estimates in the reduction in SNRs and inaccuracies of the determination of the parameters characterizing the signal, due to the use of approximated waveforms, have
been derived. Depending on the magnitude of these degradations one can decide whether to use these approximated waveforms as templates in a data analysis search.

Since it can be argued that contiguous PN approximations should well characterize the differences between the “true signal” present in the data and the highest-order PN approximation, we have performed such a comparison within the Bayesian framework. An analogous, frequentist analysis has recently been performed by Cutler and Vallisneri [13] for the case of super massive black holes binaries observed by LISA (the Laser Interferometer Space Antenna). Their approach relied on the use of the Fisher-Information matrix, which is known to give good results in the case of large (hundreds to thousands) SNRs. In the case of ground-based interferometers instead, since the expected SNRs will be probably smaller than 10, a parameter estimation error analysis based on the Fisher-Information matrix would lead to erroneous results [14].

In a recent paper [15] we have estimated the loss in probability of detection (i.e. loss of evidence of a signal to be present in the data) as a function of the SNR when the true signal present in the data is a spin-less waveform represented by a given PN approximating order and the search model is one PN order level below it. Our analysis was purposely limited to only two separate cases (i.e. PN3.5 vs. PN3.0 and PN2.0 vs. PN1.5) in order to cover the region of the PN approximations that have already, or are in the process of, being used in the analysis of the data collected by presently operated ground-based interferometers.

Our model comparison relies on a Bayesian Markov Chain Monte Carlo (MCMC) technique, as MCMC methods have successfully been applied to a large number of problems involving parameter estimations [16] in experimental data sets. In our analysis the chirp signals (the one present in the data and the one used as the model) are characterized by five parameters: the two masses of the system, \(m_1\) and \(m_2\), their time to coalescence \(t_C\), the coalescence phase \(\phi_C\), and their distance \(r\) from Earth.

This paper is organized as follows. In Section 2 we provide a brief summary of the Bayesian framework and its implementation in our problem. After deriving the expressions of the likelihood function and the priors for the parameters searched for, in Sec. 3 we give a brief description of the Markov-Chain Monte Carlo sampling technique adopted for calculating the posterior distributions, and summarize the results of our numerical analysis. We find that (i) the difference in detection probability when using a simplified model rather than the true one is negligibly small in comparison to the uncertainties associated with different noise realizations, while (ii) the posterior credibility regions of the parameters reveal offsets from the true parameter values that can be much larger than the statistical uncertainty.

2. The Bayesian formulation

The Bayesian full probability model for our problem, which involves the comparison between the two possibilities of having either a signal and noise or just noise in the data, requires the determination of the likelihoods, the prior distributions for the parameters associated with the models, and the resulting posterior distributions. Let us consider a data stream \(d(t) = s(t; \theta) + n(t)\) containing the instrumental noise \(n(t)\) and a chirp signal \(s(t; \theta)\) that we will regard as the “true” signal. Here, \(\theta\) is the vector representing all the parameters associated with the signal, and the noise is assumed to be a stationary Gaussian random process of zero mean. In the Fourier domain the observed data can equivalently be written as \(\tilde{d}(\theta; f) = \tilde{s}(\theta; f) + \tilde{n}(f)\) (where tilde denotes the Fourier transform operation) and we will refer to this expression as model \(\mathcal{M}_t\).

In what follows we will assume the “true” signal \(\tilde{s}(\theta; f)\) to be the gravitational wave emitted by a coalescing binary system and represented by a spin-less Post-Newtonian approximation in phase and Newtonian in amplitude for which \(\theta = (m_1, m_2, r, t_C, \phi_C)^T\). Here \(m_1\) and \(m_2\) are the masses of the rotating objects, \(t_C\) is the coalescence time, \(r\) the absolute distance to the binary
system, and $\phi_C$ the phase of the signal at coalescence.

We will then describe the detection and estimation of the parameters of the “true” signal by relying on a spin-less lower-order Post-Newtonian waveform, $\tilde{s}_t(\theta; f)$. This simpler model will be referred to as model $\mathcal{M}_s$.

The derivation of the detection probability implies a comparison between model $\mathcal{M}_s$ and the null-model, which postulates mere noise $\tilde{n}(f)$ within the data. This model will be referred to as model $\mathcal{M}_n$.

Since we have assumed the distribution of the random process associated with the noise of the detector to be Gaussian of zero-mean, it follows that the likelihood function is proportional to

$$M_d \propto \exp \left( -2 \int_{f_L}^{f_U} \frac{|\tilde{d}(f) - \tilde{s}_t(\theta; f)|^2}{S_n(f)} \, df \right),$$

where $S_n(f)$ is the one-sided power spectral density of the noise.

By substituting $\tilde{d}(f) = \tilde{s}_t(f; \theta_t) + \tilde{n}(f)$ into Eq. 1, the likelihood function becomes

$$p(\tilde{d}|\mathcal{M}_s, \theta) \propto \exp \left( -2 \int_{f_L}^{f_U} \frac{\tilde{s}_t(f; \theta_t) + \tilde{n}(f) - \tilde{s}_t(\theta; f)|^2}{S_n(f)} \, df \right).$$

In analogy to Eq. 2, under model $\mathcal{M}_n$ (with no parameters) the likelihood assumes the following form

$$p(\tilde{d}|\mathcal{M}_n) \propto \exp \left( -2 \int_{f_L}^{f_U} \frac{\tilde{s}_t(f; \theta_t) + \tilde{n}(f)|^2}{S_n(f)} \, df \right).$$

For comparison reasons, in the case of using the “true” model the likelihood function becomes

$$p(\tilde{d}|\mathcal{M}_t, \theta) \propto \exp \left( -2 \int_{f_L}^{f_U} \frac{\tilde{s}_t(f; \theta_t) + \tilde{n}(f) - \tilde{s}_t(\theta; f)|^2}{S_n(f)} \, df \right),$$

which will then give information about the impairment in detection when using model $\mathcal{M}_s$ instead of $\mathcal{M}_t$. Note, that the investigation of this question requires to do the two different model comparisons separately, i.e. $\mathcal{M}_n$ vs. $\mathcal{M}_s$ and $\mathcal{M}_n$ vs. $\mathcal{M}_t$ [21].

The next step needed for completing our Bayesian full probability model is the identification of suitable prior distributions for the five parameters characterizing the chirp signal. The derivation of suitable priors has been discussed in [15] and more extensively in [17, 18], and we refer the reader to those references for details. Note that the selection of the appropriate priors $p(\theta)$ bears significant influence on the evidence of a signal presence within noise, since the prior identifies the size of the parameter space which the evidence is based on. In short, the masses are assumed to be uniformly distributed over a specified range, $[m_{\text{min}}, m_{\text{max}}]$, and the prior distribution for the distance is chosen to be a cumulative distribution of having systems out to a distance $x$ smaller than $r$, $P(x < r)$, proportional to the cube of the distance, $x^3$. In order to obtain a proper prior distribution that does not diverge once integrated to infinity, it is down-weighted by including an exponential decaying and includes the assumption of a uniform distribution for the masses. The sigmoid function that serves as the down-weighing term, is chosen in such way that a $(2 - 2)\mathcal{M}_C$ inspiral system is assumed to be detectable with probabilities 0.1 and 0.9 out to distances 95 Mpc and 90 Mpc respectively [17, 18]. This choice depends on the assumed noise model which roughly coincides with the noise sensitivity curve of the LIGO-I detector [23].
As far as the time to coalescence is concerned, we have assumed it to be uniformly distributed over a time interval of 1 second centered around the value identified by the masses of the binary and the lower frequency cut-off of the detector [24]. This search range for the time to coalescence $t_C$ is larger than that used in [17] because when using the simplified model the posterior peak can be offset from the true value by more than the posterior width. Finally, we have chosen the phase of the signal at coalescence, $\phi_C$, to be uniformly distributed over the interval $[0, 2\pi]$, i.e. $p(\phi_C) = \frac{1}{2\pi}$. The choice of priors is different when it comes to the analysis of model $M_n$. Since this model postulates mere noise, there are no parameters entering the likelihood, which is therefore a constant.

The final remaining step in defining the Bayesian procedure is to assign prior probabilities to the models themselves. Since we have no a priori knowledge, the unbiased choice is to assume equal probability for each. However, one can show that for SNRs equal to about 7 or larger the effects of assigning different prior detection probabilities are negligible. See the appendix in [15] for a detailed analysis of this point.

By applying Bayes’ theorem using the likelihoods and priors defined above, we then derive the multidimensional posterior probability distribution for the model and its parameters

$$p(i, \theta, \tilde{d}) = \begin{cases} p(M_n) \cdot p(\tilde{d} | M_n) & \text{if } i = 0 \\ p(M_s, \theta) \cdot p(\tilde{d} | M_s, \theta) & \text{if } i = 1 \end{cases} \int p(M_s, \theta) \cdot p(\tilde{d} | M_s, \theta) d\theta + p(M_n) \cdot p(\tilde{d} | M_n) ,$$

where $i \in \{0, 1\}$ corresponds to the two models $\{M_n, M_s\}$. In the same way, it is possible to derive the posterior for the comparison of $M_n$ vs. $M_t$. A Bayesian analysis naturally justifies Occam’s Razor [19, 20] due to the penalization of unreasonably complex models by integrating over the parameter space resulting in the preference for a simpler model.

3. Bayesian model selection and posterior probabilities

There exist various techniques for tackling the multi-dimensional problem of estimating the posterior distributions. We have implemented a relatively new procedure, called Reversible Jump MCMC (RJMCMC) technique [25, 26], which simultaneously addresses the problems of model selection and parameter estimation. The RJMCMC is combined with traditional fixed dimension MCMC techniques that sample from the parameters of the current model. See [15] for details.

Since our simulations were computationally time-intensive, we have created data sets from “true” wave forms of three hypothetical binary inspiral systems whose masses were selected in order to get a fair representation of the entire mass-parameter space. We also considered two scenarios where the “true” wave form was either of PN 2.0 or PN 3.5 order. The detection of each scenario was attempted by either a PN 1.5 or a PN 3.0 wave form respectively [15].

The distance of each binary system was varied in order to obtain different signal-to-noise ratios. The noise realizations were drawn in such a way that they corresponded to the approximated expression of the one-sided power spectral density of initial LIGO [27]

$$S_n(f) = \frac{S_0}{5} \left[ \left( \frac{f_0}{f} \right)^4 + 2 \left[ 1 + \left( \frac{f}{f_0} \right)^2 \right] \right]$$

with $S_0 = 8.0 \times 10^{-46}$ Hz$^{-1}$ being the minimum noise of the detector and $f_0 = 175$ Hz the frequency at which the sensitivity of the detector reaches its maximum. The noise samples were generated directly in the Fourier domain by first simulating white noise samples, and then by scaling their amplitudes according to the required noise spectrum $S_n(f)$. The simulated data
were sampled at 4096 Hz for a duration of about 24 s, and they were produced by embedding the different signals into noise samples that were generated in the Fourier domain.

The integration bandwidth for the likelihood was chosen from 12 Hz up to the frequency of the last stable orbit or 600 Hz, whichever is the smaller. Since the SNR is negligible above 600 Hz, we fixed this to be the upper frequency cut-off. For the binary system \((m_1 = 1.8M_\odot, m_2 = 0.6M_\odot)\) the frequency at the last stable orbit is equal to 1832.2 Hz, which gives an integration limit of 600 Hz. This translates in 14355 complex samples that contribute to the posterior distribution.

In the case of the binary system with \((m_1 = 45.0M_\odot, m_2 = 0.52M_\odot)\) the last stable orbit is at 96.6 Hz, resulting in 2065 complex samples involved in the determination of the likelihood function. Finally, for the high mass binary system \((m_1 = 45.0M_\odot, m_2 = 30M_\odot)\) in our set of systems the last stable orbit is at 58.6 Hz, implying now only 1138 complex samples over which the likelihood is calculated.

After we conducted the MCMC simulations we derived the posterior detection probabilities for the competing models from the MCMC outputs regarding the three example binary systems, the four different model comparisons, and the different sets of SNRs. For each binary system we computed the posterior probabilities for the considered scenarios and contrasted the probability of detection based on a lower order PN expansion against the one using the true wave form.

The results are shown in figures (3,4,5) of reference [15]. The main conclusion of our analysis is that the probability of detection is not impeded by using a simplified model for detecting wave forms of higher PN order in the low-SNR regime. The Bayesian approach provides the means to gain insight into the variation of the detection probability over different noise realizations. The difference between the posterior detection probabilities corresponding to the true and the simplified model is very small as compared to its variance over different noise realizations. It is worth mentioning that analyses performed within the frequentist framework [27, 28, 12] and aimed at comparing the detectability of a signal by using a simplified wave form were focused entirely on estimating the resulting loss of SNR. The Bayesian model comparison presented here has the inherent ability to estimate probabilities and their uncertainties due to noise, providing much more insights into this issue.

The impact of the use of lower PN order wave forms on the bias of posterior distributions of the parameters was then estimated by calculating the posterior distributions of the parameters characterizing the analyzed binary systems. We found that when the model matches the signal present in the data the posteriors cover well the true parameter values. However, when the searching model is different from that of the signal present in the data the posterior distributions display a marked displacement of the mass parameters from their “true values” that can be as large as several of their standard deviations. Very striking is also the error in the time to coalescence. The masses and time to coalescence are obviously the parameters subject to biases when using a simplified model. This is physically understandable since these three parameters define the phase of the signal.

4. Acknowledgments
This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References
[1] The LIGO project home page. http://www.ligo.caltech.edu
[2] The Virgo project home page. http://www.virgo.infn.it/
[3] The GEO-600 project home page. http://www.geo600.uni-hannover.de/
[4] The TAMA project home page. http://tamago.mtk.nao.ac.jp/
[5] Helström C W, 1968 Statistical Theory of Signal Detection, Pergamon Press, London (UK).
[6] Thorne K S 1987, Gravitational radiation. In Three Hundred Years of Gravitation, pages 330–458. Cambridge University Press. Hawking S W and Israel W, eds.
[7] Pretorius F 2005 Evolution of binary black-hole spacetimes. Physical Review Letters, 95, 121101
[8] Campanelli M, Lousto C O, Marronetti P and Zlochower Y. 2006, Accurate evolutions of orbiting black-hole binaries without excision. Physical Review Letters, 96, 111102.
[9] Baker J G, Centrella J, Choi D, Koppitz M and van Meter J, 2006. Gravitational-wave extraction from an inspiraling configuration of merging black holes. Physical Review Letters, 96, 111102.
[10] Buonanno A, Cook G B and Pretorius F, 2007. Inspirial, merger, and ring-down of equal-mass black-hole binaries. Physical Review D, 75, 124018.
[11] Baker J G, van Meter J R, McWilliams S T, Centrella J and Kelly B J, 2007. Consistency of post-Newtonian waveforms with numerical relativity. Physical Review Letters, 99, 181101.
[12] Baker J G, Buonanno A, Baker J G, Centrella J, Kelly B J, McWilliams S T, Pretorius F, and van Meter J R, 2008. A data-analysis driven comparison of analytic and numerical coalescing binary waveforms: nonspinning case. Phys. Rev. D, 77, 024014.
[13] Cutler C and Vallisneri M, 2007. Lisa detections of massive black hole inspirals: parameter extraction errors due to inaccurate template waveforms. Phys. Rev. D, 76, 104018.
[14] Vallisneri M, 2008. Use and abuse of the fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. Phys. Rev. D, 77, 042001
[15] Unstätter R and Tinto M, 2008. Bayesian comparison of Post-Newtonian approximations of gravitational wave chirp signals. Physical Review D, 77, 082002.
[16] Gilks W R, Richardson S, Spiegelhalter P J, and Green D J, editors, 1996. Markov chain Monte Carlo in practice. Chapman and Hall, London (UK).
[17] Röver C, Meyer R and Christensen N, 2006. Bayesian inference on compact binary inspiral gravitational radiation signals in interferometric data. Classical and Quantum Gravity, 23 (15), 4895–4906.
[18] Röver C, Meyer R and Christensen N, 2007. Coherent Bayesian inference on compact binary inspirals using a network of interferometric gravitational wave detectors. Physical Review D, 75(6), 062004
[19] Jaynes E T, 2003. Probability Theory: The Logic of Science. Cambridge University Press, Cambridge (UK).
[20] Loredo T J, 1992. Statistical challenges in modern astronomy, pages 275–297. Springer-Verlag, New York.
[21] Gregory P C, 2005. Bayesian Logical Data Analysis for the Physical Sciences. Cambridge University Press, Cambridge (UK).
[22] Sandage A, 2001. Encyclopedia of Astronomy and Astrophysics. Institute of Physics Publishing, Bristol (UK).
[23] Sigg D, 2004. Commissioning of LIGO detectors. Classical and Quantum Gravity, 21, S40915.
[24] Pai A, Dhurandhar S V and Bose S, 2001. Data-analysis strategy for detecting gravitational-wave signals from inspiraling compact binaries with a network of laser-interferometric detectors. Physical Review D, 64, 042004.
[25] Green P J, 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika, 82(4), 711–732
[26] Green P J, Hjort N L and Richardson S editors. 2003. Highly Structured Stochastic Systems. Oxford University Press
[27] Chronopoulos A E and Apostolatos T A, 2001. Less accurate but more efficient family of search templates for detection of gravitational waves from inspiraling compact binaries. Physical Review D, 64, 042003.
[28] Canitrot P, 2001. Systematic errors for matched filtering of gravitational waves from inspiraling compact binaries. Physical Review D, 63, 082005.