Towards a unified treatment of gravitational-wave data analysis

Neil J. Cornish and Joseph D. Romano

Phys. Rev. D **87**, 122003 — Published 3 June 2013

DOI: 10.1103/PhysRevD.87.122003

Towards a unified treatment of gravitational-wave data analysis

Neil J. Cornish
Department of Physics, Montana State University, Bozeman, MT 59718, USA.

Joseph D. Romano
Department of Physics and Astronomy and Center for Gravitational-Wave Astronomy, University of Texas at Brownsville, Brownsville, TX 78520, USA.

We present a unified description of gravitational-wave data analysis that unites the template-based analysis used to detect deterministic signals from well-modeled sources, such as binary-black-hole mergers, with the cross-correlation analysis used to detect stochastic gravitational-wave backgrounds. We also discuss the connection between template-based analyses and those that target poorly-modeled bursts of gravitational waves, and suggest a new approach for detecting burst signals.

PACS numbers: 04.80.Nn, 04.30.Db, 07.05.Kf, 95.55.Ym

Gravitational-wave data analysis is conventionally divided into three classes that depend on the nature of the signal: (i) well-modeled deterministic signals, such as those from compact binary inspirals; (ii) poorly-modeled deterministic signals, such as those from core-collapse supernovae; and (iii) stochastic signals, such as those from a phase transition in the early Universe. Here we will argue that this division is rather artificial and unnecessary, and suggest that a unified treatment can yield deeper insights. The elements needed to unify cases (i) and (ii) can be found in Refs. [1–4]. Here we provide a unification of cases (i) and (iii) using hierarchical Bayesian modeling [5].

The motivation for developing a unified description of gravitational-wave data analysis is two-fold. First there is the pedagogical value of a coherent picture that emphasizes the common foundation of the disparate analysis techniques found in the literature, and second, the unified picture can provide a deeper understanding that may suggest new approaches. To illustrate the latter point we conclude our discussion by proposing a novel technique for detecting un-modeled "bursts" of gravitational waves.

In the conventional picture, signals for which we have waveform templates are analyzed using a matched-filter statistic [6], un-modeled signals are characterized in terms of an excess power statistic [7], and stochastic signals are analyzed using a cross-correlation statistic between pairs of detectors [8]. The connection between the various forms of analysis is not immediately apparent, especially when described in a frequentist framework.

In the case of un-modeled signals, the analyses usually focus on short duration "bursts" of gravitational-wave energy that are localized in a time-frequency representation of the data. The connection between a waveform template-based search and a burst search becomes apparent in the case of fully-coherent network analyses, where it becomes possible to solve for the gravitational-wave signal by either maximizing the likelihood [2] or locating regions of high posterior density [3, 4]. To obtain meaningful results these analyses require constraints or priors on the signal models. The waveform template-based analyses are recovered in the limit that the signal priors become highly informative, ultimately mapping the individual signal samples h_i to a small number of physical parameters $\vec{\lambda}$ that describe the signal: $h_i(\vec{\lambda})$.

To develop the connection between the cases (i) and (iii) we found it advantageous to adopt a hierarchical Bayesian analysis framework [9, 10] with parametrized likelihood functions and a parameterized signal prior. We begin with the simplest case imaginable: that of two co-located and co-aligned detectors, which each provide a single datum $s_1 = n_1 + h$, $s_2 = n_2 + h$, respectively. The noise in the detectors is assumed to be Gaussian random and independent, with zero mean and variance σ_i^2. The signal, h, which is common to both detectors, is also assumed to come from a Gaussian distribution with zero mean and variance σ_h^2. The quantities σ_i, σ_h are our model hyper-parameters, which are to be determined from the data.

Now suppose that we adopt a waveform template h to describe the gravitational wave signal, and form the residuals $r_i = s_i - h$. We demand that the residuals be consistent with the probability distribution for the noise [11], which in this case results in a multi-variate Gaussian likelihood function:

$$p(s|\sigma_i, h) = \frac{1}{\sqrt{(2\pi)^2 \det C'}}\, e^{-\frac{1}{2}\, r_i (C'^{-1})_{ij} r_j}, \tag{1}$$

with

$$C'_{ij} = \delta_{ij}\sigma_i^2. \tag{2}$$

Since our template is stochastic, we are not interested in the particular value of h, but rather, in the overall amplitude σ_h. Thus, using the parameterized signal prior

$$p(h|\sigma_h) = \frac{1}{\sqrt{2\pi\sigma_h^2}} e^{-h^2/2\sigma_h^2} \tag{3}$$

we marginalize over h:

$$\begin{aligned} p(s|\sigma_i,\sigma_h) &\equiv \int p(s|\sigma_i,h)\, p(h|\sigma_h)\, dh \\ &= \frac{1}{2\pi\,\sigma_1\sigma_2}\int e^{-(s_1-h)^2/2\sigma_1^2} e^{-(s_2-h)^2/2\sigma_2^2} \\ &\quad \times \frac{1}{\sqrt{2\pi}\,\sigma_h} e^{-h^2/2\sigma_h^2}\, dh \\ &= \frac{1}{\sqrt{(2\pi)^2\,\mathrm{det}C}}\, e^{-\frac{1}{2}\, s_i (C^{-1})_{ij} s_j}, \end{aligned} \tag{4}$$

where the matrix C has components

$$C_{ij} = \delta_{ij}\sigma_i^2 + \sigma_h^2\,. \tag{5}$$

The likelihood function (4) has the standard form used in cross-correlation analyses for stochastic signals [12, 13]. Thus we see that a template-based analysis (1) using stochastic templates is equivalent to a cross-correlation analysis (4) without templates.

The generalization to N co-aligned and co-located detectors is straightforward. Expression (1) becomes

$$p(s|\sigma_i,h) = \frac{1}{\sqrt{(2\pi)^N \mathrm{det}C'}}\, e^{-\frac{1}{2}\, r_i (C'^{-1})_{ij} r_j}, \tag{6}$$

and the marginalization proceeds as follows:

$$\begin{aligned} p(s|\sigma_i,\sigma_h) &= \int p(s|\sigma_i,h)\, p(h|\sigma_h)\, dh \\ &= \frac{e^{-\frac{1}{2}\, s_i (C^{-1})_{ij} s_j}}{\sqrt{(2\pi)^N \mathrm{det}C'}} \int \frac{dh}{\sqrt{2\pi}\,\sigma_h} \\ &\quad \times \exp\left(-\frac{\mathrm{det}C}{2\sigma_h^2 \mathrm{det}C'}\left(h - \frac{\sigma_h^2 \mathrm{det}C'}{\mathrm{det}C}\sum_i \frac{s_i}{\sigma_i^2}\right)^2\right) \\ &= \frac{e^{-\frac{1}{2}\, s_i (C^{-1})_{ij} s_j}}{\sqrt{(2\pi)^N \mathrm{det}C}} \int \frac{1}{\sqrt{2\pi}\,\sigma_h} e^{-h'^2/2\sigma_h^2}\, dh' \\ &= \frac{1}{\sqrt{(2\pi)^N \mathrm{det}C}}\, e^{-\frac{1}{2}\, s_i (C^{-1})_{ij} s_j}\,. \end{aligned} \tag{7}$$

The marginalization involves a completion of the square, followed by a change of variables from h to h' via a shift, then a rescaling. It is the rescaling that takes $\sqrt{\mathrm{det}C'}$ to $\sqrt{\mathrm{det}C}$ in the denominator.

The generalization to multiple data points is trivial. The individual variances σ_i^2, σ_h^2 are replaced by variance-covariance *matrices* $\mathbf{\Sigma}_i$, $\mathbf{\Sigma}_h$, with corresponding replacements for C' and C:

$$\mathbf{C'}_{ij} = \delta_{ij}\,\mathbf{\Sigma}_i\,, \quad \mathbf{C}_{ij} = \delta_{ij}\,\mathbf{\Sigma}_i + \mathbf{\Sigma}_h\,. \tag{8}$$

Assuming time stationarity, these variance-covariance matrices can be written in the frequency domain as one-sided spectral density functions

$$C'_{ij}(f) = \delta_{ij} S_{n_i}(f)\,, \tag{9}$$

$$C_{ij}(f) = \delta_{ij} S_{n_i}(f) + S_h(f)\,. \tag{10}$$

In terms of these functions the likelihood and marginalized likelihood are given by

$$p(s|\sigma_i,h) = \prod_f \frac{1}{(2\pi)^N\,\mathrm{det}C'(f)}\, e^{-2\,\tilde{r}_i(f) C'_{ij}(f)^{-1} \tilde{r}^*_j(f)}\,, \tag{11}$$

$$p(s|\sigma_i,\sigma_h) = \prod_f \frac{1}{(2\pi)^N\,\mathrm{det}C(f)}\, e^{-2\,\tilde{s}_i(f) C_{ij}(f)^{-1} \tilde{s}^*_j(f)}\,. \tag{12}$$

The extension to multiple polarization states and signals with a spatial distribution is also relatively straightforward. To simplify the discussion, let us start by considering co-located, but no longer co-aligned detectors, and write

$$h_i = F_i^+ h_+ + F_i^\times h_\times\,, \tag{13}$$

where F_i^+ and $F_i^\times$ are the antenna beam patterns for the two polarization states for detector i in the long-wavelength limit. Then the signal prior for $h \equiv (h_+, h_\times)$ takes the form

$$p(h|\sigma_h) = \frac{1}{2\pi\sigma_h^2} e^{-(h_+^2 + h_\times^2)/2\sigma_h^2}\,. \tag{14}$$

Marginalization over h now involves integration over both h_+ and $h_\times$. For example, we start with the residual $r_i = s_i - F_i^+ h_+ - F_i^\times h_\times$ and a diagonal correlation matrix $C'_{ij} = \delta_{ij}\sigma_i^2$. After doing the integral over $h_\times$ we have a *new* residual $r'_i = s_i - F_i^+ h_+$, and the correlation matrix picks up the off-diagonal term $F_i^\times F_j^\times \sigma_h^2$. Then the h_+ integration takes us to $r''_i = s_i$, and the correlation matrix

$$C_{ij} = \delta_{ij}\sigma_i^2 + (F_i^+ F_j^+ + F_i^\times F_j^\times)\sigma_h^2\,. \tag{15}$$

To handle the case of sources that are distributed across the sky, we first write

$$h_i = \sum_{n=1}^{M} F_i^+(\hat{n}) h_+^n + F_i^\times(\hat{n}) h_\times^n\,, \tag{16}$$

where n labels the sky location in the direction $\hat{n}$. The continuum limit can be found by taking $M \to \infty$ and replacing the sum by an integral. For an isotropic background the signal prior for $h \equiv (h_+^n, h_\times^n)$ is given by

$$p(h|\sigma_h) = \prod_{n=1}^{M} \frac{1}{2\pi\sigma_h^2} e^{-((h_+^n)^2 + (h_\times^n)^2)/2\sigma_h^2}\,. \tag{17}$$

Repeated application of the marginalization as before yields the likelihood (7) with the correlation matrix

$$C_{ij} = \delta_{ij}\sigma_i^2 + \frac{\bar{\sigma}_h^2}{4\pi}\int d\Omega_{\hat{n}} (F_i^+(\hat{n}) F_j^+(\hat{n}) + F_i^\times(\hat{n}) F_j^\times(\hat{n}))\,, \tag{18}$$

where $\bar{\sigma}_h^2 = \sigma_h^2 M$. We have taken the continuum limit in writing the final expression for the correlation matrix, noting that as $M \to \infty$, the variance $\bar{\sigma}_h^2$ is held constant. The diagonal terms in the signal-dependent part of (18) are the sky (and polarization) averaged response functions for the detectors. The off-diagonal term in (18) is proportional to the *overlap reduction function* for co-located, but mis-aligned detectors. The derivation for anisotropic backgrounds is very similar, the only difference being that the σ_h are different for each sky location, leading to the correlation matrix

$$C_{ij} = \delta_{ij}\sigma_i^2 + \frac{1}{4\pi}\int d\Omega_{\hat{n}} (F_i^+(\hat{n})F_j^+(\hat{n}) + F_i^\times(\hat{n})F_j^\times(\hat{n}))\,\bar{\sigma}_h^2(\hat{n})\,. \tag{19}$$

Finally, for spatially separated and mis-aligned detectors we can adopt a coordinate system where the detectors are located at $\vec{x}_1$ and $\vec{x}_2$, so that in the Fourier domain the signals can be written as

$$\tilde{h}_i(f) = \sum_{n=1}^{M} (F_i^+(\hat{n})\tilde{h}_+^n(f) + F_i^\times(\hat{n})\tilde{h}_\times^n(f))e^{2\pi i f \vec{x}_i\cdot\hat{n}}\,. \tag{20}$$

Marginalizing over the stochastic signal for an isotropic background yields the likelihood (12) with the Hermitian correlation matrix

$$C_{ij}(f) = \delta_{ij}S_{n_i}(f) + \gamma_{ij}(f)S_h(f)\,, \tag{21}$$

where $\gamma_{ij}(f)$ is the overlap reduction function [14, 15]

$$\gamma_{ij}(f) = \frac{1}{4\pi}\int (F_i^+(\hat{n})F_j^+(\hat{n}) + F_i^\times(\hat{n})F_j^\times(\hat{n})) \times e^{2\pi i f(\vec{x}_i - \vec{x}_j)\cdot\hat{n}}\, d\Omega_{\hat{n}}\,. \tag{22}$$

If the background were anisotropic, the correlation matrix would have the form

$$C_{ij}(f) = \delta_{ij}S_{n_i}(f) + \kappa_{ij}(f), \tag{23}$$

with

$$\kappa_{ij}(f) = \frac{1}{4\pi}\int (F_i^+(\hat{n})F_j^+(\hat{n}) + F_i^\times(\hat{n})F_j^\times(\hat{n})) \times S_h(\hat{n}, f)\, e^{2\pi i f(\vec{x}_i - \vec{x}_j)\cdot\hat{n}}\, d\Omega_{\hat{n}}\,. \tag{24}$$

Note that all of the above results are special cases of the general result for the convolution of two multi-variate Gaussian distributions:

$$\int \frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{F}\cdot\mathbf{y})^T\mathbf{D}^{-1}(\mathbf{x}-\mathbf{F}\cdot\mathbf{y})}}{\sqrt{(2\pi)^{N_x}\det\mathbf{D}}}\frac{e^{-\frac{1}{2}\mathbf{y}^T\mathbf{E}^{-1}\mathbf{y}}}{\sqrt{(2\pi)^{N_y}\det\mathbf{E}}}\,d\mathbf{y} = \frac{1}{\sqrt{(2\pi)^{N_x}\det\mathbf{C}}}e^{-\frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}} \tag{25}$$

where

$$\mathbf{C}^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{F}(\mathbf{E}^{-1} + \mathbf{F}^T\mathbf{D}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{D}^{-1} \tag{26}$$

and

$$\mathbf{C} = \mathbf{D} + \mathbf{F}\mathbf{E}\mathbf{F}^T\,. \tag{27}$$

The connection between a stochastic template-based analysis and a cross-correlation analysis was partially developed in Refs. [16, 17], where a connection was made between a template-based maximum-likelihood estimator and the cross-correlation statistic used to search for stochastic backgrounds. However, the analysis was limited to a pair of co-located and co-aligned detectors, and the possibility of a full unification was not developed.

Establishing a unified description of gravitational-wave data analysis has value beyond pedagogy. Seeing the connection between the analyses can suggest new approaches. Of particular promise are novel approaches that fall between the conventional divisions. In the unified approach, the analysis begins with a model for the instrument noise, which then becomes the likelihood for the residuals $r = s - h$. The specification of the prior on the signal model then completes the model, giving a continuum of analysis techniques that range from the highly informative waveform priors $p(h|\vec{\lambda}) = \delta(h - h(\vec{\lambda}))$ of a standard matched filter analysis to the stochastic prior $p(h|\sigma_h)$ that yields the standard cross-correlation analysis. While it may be difficult to reverse engineer the form of the signal prior that yields some of the existing hybrid analysis techniques [18–24], it is likely that such a mapping will always exist as we have almost infinite freedom in the choice of the signal prior. Specific examples of such reverse engineering of the signal priors can be found in Ref. [4]. Rather than trying to recover existing approaches, a more promising avenue for future research is to explore alternative choices for the signal prior that may yield useful analysis techniques.

To give a concrete example of a new analysis approach that is suggested by the unified picture, we propose a simple signal prior for un-modeled bursts of gravitational wave radiation. We picture these signals as occupying a relatively small area in time-frequency space, so it is natural to work in a wavelet basis where the signal in each detector can be written as $s_{\mu ij}$, where the Greek index μ labels the detector and the Roman indices i, j denote the location in time and frequency respectively. For a co-located and co-aligned two detector network the likelihood is then

$$p(s|\sigma_{\mu j}, \sigma_{\nu l}, h) = \frac{e^{-\frac{1}{2}r_{(\mu ij)}(C'^{-1})_{(\mu ij)(\nu kl)}r_{(\nu kl)}}}{\sqrt{(2\pi)^{2N}\det C'}}\,, \tag{28}$$

with

$$C'_{(\mu ij)(\nu kl)} = \sigma_{\mu j}^2\,\delta_{\mu\nu}\delta_{ik}\delta_{jl}. \tag{29}$$

Here N is the number of wavelet components. The expression for the correlation matrix is only approximate as there will be a some noise correlation between frequency layers j, l, but in a well-chosen wavelet basis the correlation is negligible. For the signal model we assume that the wavelet amplitudes are Gaussian random distributed:

$$p(h_{(ij)}|\sigma_{h(ij)}(f_c, t_c, \Delta f, \Delta t)) = \frac{1}{\sqrt{2\pi}\,\sigma_{h(ij)}} e^{-h_{(ij)}^2/2\sigma_{h(ij)}^2} \tag{30}$$

with amplitudes $\sigma_{h(ij)}$ that depend on a central frequency f_c, central time t_c, and widths Δf and Δt:

$$\sigma_{h(ij)}(f_c, t_c, \Delta f, \Delta t) = \sigma e^{-((t_i-t_c)^2/2\Delta t^2 + (f_j-f_c)^2/2\Delta f^2)}\,. \tag{31}$$

Marginalizing over h yields

$$p(s|\sigma_{\mu j}, \sigma_{\nu l}, h) = \frac{e^{-\frac{1}{2}\, s_{(\mu ij)}(C^{-1})_{(\mu ij)(\nu kl)} s_{(\nu kl)}}}{\sqrt{(2\pi)^N \det C}}\,, \tag{32}$$

with

$$C_{(\mu ij)(\nu kl)} = (\sigma_{\mu j}^2\, \delta_{\mu\nu} + \sigma_{h(ij)}^2)\delta_{ik}\delta_{jl}\,. \tag{33}$$

For frequencies and times f_j, t_i that are far from the central values f_c, t_c the correlation matrix reduces to that of uncorrelated noise. In effect we end up with a cross-correlation search that targets a small region in time-frequency. The analysis is easily extended to M mis-aligned and spatially distributed detectors by linearly transforming the data to form $M-2$ null streams and 2 signal streams [2]. The cross-correlation statistic then depends on 8 parameters: the sky location θ, ϕ and polarization angle ψ; the central values f_c, t_c; the widths $\Delta f, \Delta t$ and the overall amplitude σ. Our example burst search is essentially equivalent to applying a radiometer style search [21] to data that has been convolved with a Gaussian time-frequency window function, and is also similar to the *STAMP* algorithm [23].

Acknowledgments

NJC acknowledges support from NSF Award PHY-1205993 and NASA grant NNX10AH15G. JDR acknowledges support from NSF Awards PHY-1205585 and CREST HRD-1242090. We would also like to thank Nelson Christensen and Eric Thrane for comments on an earlier draft of this paper.

[1] E. E. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4566 (1998)
[2] S. Klimenko, S. Mohanty, M. Rakhmanov and G. Mitselmakher, Phys. Rev. D **72**, 122002 (2005) [gr-qc/0508068].
[3] S. Chatterji, A. Lazzarini, L. Stein, P. J. Sutton, A. Searle & M. Tinto, Phys. Rev. D**74**, 082005 (2006).
[4] A. C. Searle, P. J. Sutton and M. Tinto, Class. Quant. Grav. **26**, 155017 (2009) [arXiv:0809.2809 [gr-qc]].
[5] M. R. Adams, N. J. Cornish and T. B. Littenberg, Phys. Rev. D **86**, 124032 (2012) [arXiv:1209.6286 [gr-qc]].
[6] K. S. Thorne, “Gravitational Radiation.” *In 300 Years of Gravitation*, ed. Stephen Hawking and Werner Israel, pp. 350458. (Cambridge, Cambridge University Press 1987).
[7] W. G. Anderson, P. R. Brady, J. D. E. Creighton and E. E. Flanagan, Phys. Rev. D **63**, 042003 (2001) [gr-qc/0008066].
[8] P. F. Michelson, MNRAS **227**, 933 (1987).
[9] I. J. Good, *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*, (MIT Press, Cambridge, Mass. 1965)
[10] C. N. Morris and S. L. Normand, *in Bayesian Statistics 4*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, pages 321-334, (Oxford Univeristy Press, Oxford, 1992)
[11] L. S. Finn, Phys. Rev. D **46**, 5236 (1992)
[12] R. van Haasteren, Y. Levin, P. McDonald and T. Lu, arXiv:0809.0791 [astro-ph].
[13] M. R. Adams and N. J. Cornish, Phys. Rev. D **82**, 022002 (2010) [arXiv:1002.1291 [gr-qc]].
[14] N. Christensen, Phys. Rev. D **46**, 5250 (1992).
[15] E. E. Flanagan, Phys. Rev. D **48**, 2389 (1993) [astro-ph/9305029].
[16] S. Drasco and E. E. Flanagan, Phys. Rev. D **67**, 082003 (2003) [gr-qc/0210032].
[17] B. Allen, J. D. E. Creighton, E. E. Flanagan, and J. D. Romano, Phys. Rev. **D67**, 122002 (2003).
[18] P. R. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000) [gr-qc/9812014].
[19] B. Krishnan, A. M. Sintes, M. A. Papa, B. F. Schutz, S. Frasca and C. Palomba, Phys. Rev. D **70**, 082001 (2004) [gr-qc/0407001].
[20] C. Cutler, I. Gholami and B. Krishnan, Phys. Rev. D **72**, 042004 (2005) [gr-qc/0505082].
[21] S. W. Ballmer, Class. Quant. Grav. **23**, S179 (2006) [gr-qc/0510096].
[22] S. Dhurandhar, B. Krishnan, H. Mukhopadhyay and J. T. Whelan, Phys. Rev. D **77**, 082001 (2008) [arXiv:0712.1578 [gr-qc]].
[23] E. Thrane, S. Kandhasamy, C. D. Ott, W. G. Anderson, N. L. Christensen, M. W. Coughlin, S. Dorsher and S. Giampanis *et al.*, Phys. Rev. D **83**, 083004 (2011) [arXiv:1012.2150 [astro-ph.IM]].
[24] R. Prix, S. Giampanis and C. Messenger, Phys. Rev. D **84**, 023007 (2011) [arXiv:1104.1704 [gr-qc]].