Parameter estimation on gravitational waves from multiple coalescing binaries *[arXiv:0912.5531]*

NORTHWESTERN UNIVERSITY

CIERA CENTER FOR INTERDISCIPLINARY EXPLORATION AND RESEARCH IN ASTROPHYSICS

Ilya Mandel (Northwestern, NSF AAPF)

Abstract

Future ground-based and space-borne interferometric gravitational-wave detectors may capture between tens and thousands of binary coalescence events per year. There is a significant and growing body of work on the estimation of astrophysically relevant parameters, such as masses and spins, from the gravitational-wave signature of a single event [see posters by Raymond and Petiteau, talks by Porter and Lang]. This poster introduces a robust Bayesian framework for combining the parameter estimates for multiple events into a parameter distribution of the underlying event population. The framework can be readily deployed as a rapid post-processing tool.

Introduction

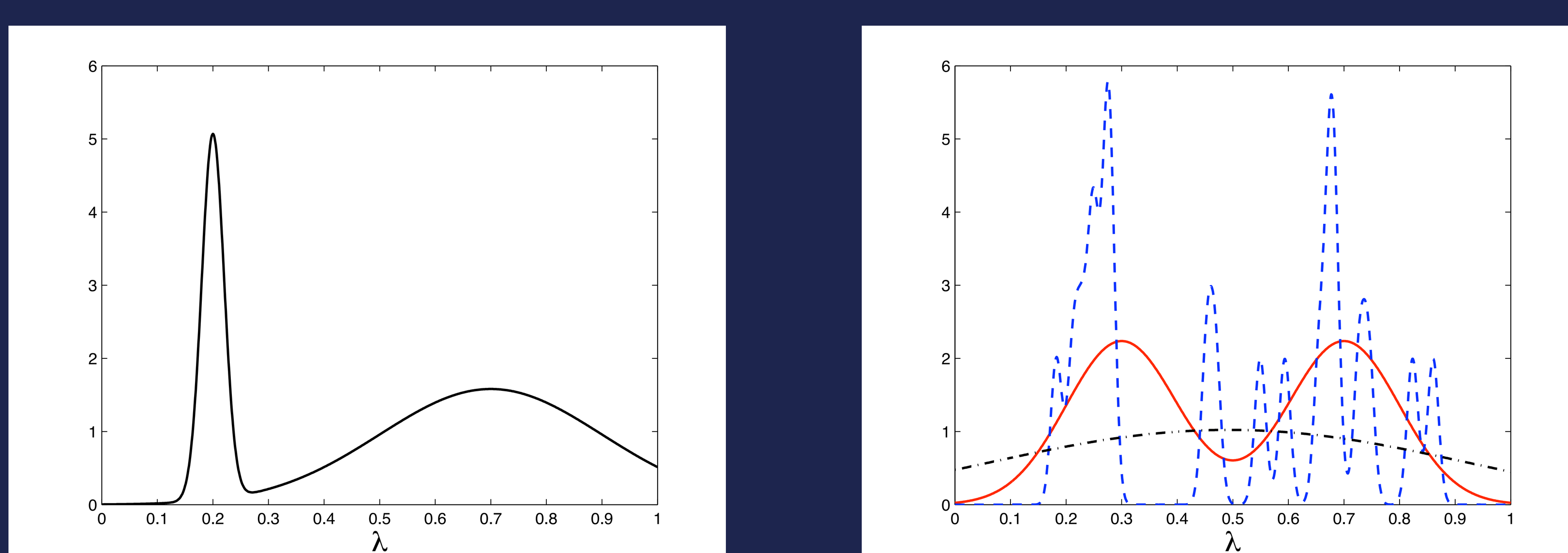

Figure 1 – The left panel (a) shows a possible multimodal marginalized posterior density function (PDF) for a single parameter of a single detected event. The right panel (b) shows the results of estimating the distribution from individual samples by means of a kernel density estimator (KDE). Twenty samples are taken from a bimodal distribution (solid red); the blue dashed curve shows a KDE where the smoothing parameter (bandwidth) was probably chosen to be too small, while the black dash-dotted curve shows the effect of a KDE with a large bandwidth parameter that over-smoothes the distribution.

The parameters of individual gravitational-wave events can be estimated with techniques like Markov Chain Monte Carlo [e.g., van der Sluys et al., 2008, ApJL 688 61], Nested Sampling [Veitch & Vechhio, 2008, CQG 25 184010], or MultiNest [Feroz et al., 2009, CQG 26 215003], allowing us to determine the posterior probability density functions (PDFs). These can be marginalized to produce PDFs for a single parameter, like the mass or spin of a binary member (e.g., Fig. 1a). We'd like to combine PDFs from multiple independent detections into a statement about the distribution of the population being sampled. However, smoothing methods like kernel density estimation (Fig. 1b) require ad hoc choices of smoothing parameters. Instead, we develop a theoretical framework based on Bayesian techniques to extract a population distribution from a set of individual detections.

Theoretical framework

We have k marginalized PDFs, $p_i(\lambda)$, and a parametrized model for the underlying population distribution, $f(\lambda)$. We'd like to know how consistent a given population distribution is with the observations; this is given by Bayes rule:

$$p\left(f(\lambda)|p_1(\lambda),...p_k(\lambda)\right) = \frac{p\left(p_1(\lambda),...p_k(\lambda)|f(\lambda)\right)\ \pi(f(\lambda))}{p\left(p_1(\lambda),...p_k(\lambda)\right)}.$$

If the k samples are independent, the likelihood of drawing those samples is:

$$p\left(p_1(\lambda),...p_k(\lambda)|f(\lambda)\right) = \prod_{i=1}^{k} p\left(p_i(\lambda)|f(\lambda)\right) = \prod_{i=1}^{k} \int_{\text{all }\lambda} p_i(\lambda) f(\lambda) d\lambda.$$

The model distribution can be discrete (e.g., a histogram) or analytical (e.g., a multi-modal Gaussian). In practice, we will have discrete PDFs, which could present difficulties in computing the overlap integral with an analytical model. However, we can use the fact that the PDFs are sampled according to the posterior, so the typical "volume" dλ corresponding to each value in the PDF chain is inversely proportional to the value of the posterior there, so

$$p\left(p_i(\lambda)|f(\lambda)\right) = \frac{1}{N_i}\sum_{j=1}^{N_i} f(\lambda_i^{(j)})$$

This allows us to compute the posterior probability for a distribution (thus giving us a *distribution of distributions*) as

$$p\left(f(\lambda)|p_1(\lambda),...p_k(\lambda)\right) \propto \pi(f(\lambda)) \prod_{i=1}^{k} \left| \frac{1}{N_i}\sum_{j=1}^{N_i} f(\lambda_i^{(j)}) \right|.$$

We can also carry out model selection, by computing the odds ratio between two competing model families M_1 and M_2:

$$\mathcal{O} = \frac{p(M_1)}{p(M_2)} \frac{\int p\left(p_1(\lambda),...p_k(\lambda)|f(\lambda);M_1\right)\ \pi(f(\lambda)|M_1)\ d\lambda}{\int p\left(p_1(\lambda),...p_k(\lambda)|f(\lambda);M_2\right)\ \pi(f(\lambda)|M_2)\ d\lambda}$$

Sample Calculation of Population Distribution

We demonstrate an example of the calculation of the population probability distribution from a set of imperfectly estimated samples using this methodology. We imagine that k points are selected from a Gaussian distribution with true mean 0.4 and standard deviation 0.1. For each detected event, the parameters are not known perfectly, but described by a discrete PDF of N points, with an unbiased mean centered on the chosen sample value and a standard deviation uniformly chosen in [a,b], intended to represent statistical errors in parameter estimation. We model the underlying probability distribution as a Gaussian with two parameters, mean μ and standard deviation σ. We construct a posterior on μ and σ with an Metropolis-Hastings MCMC, using uniform sampling from the prior for the jump proposal distribution. The result is shown in Fig. 2.

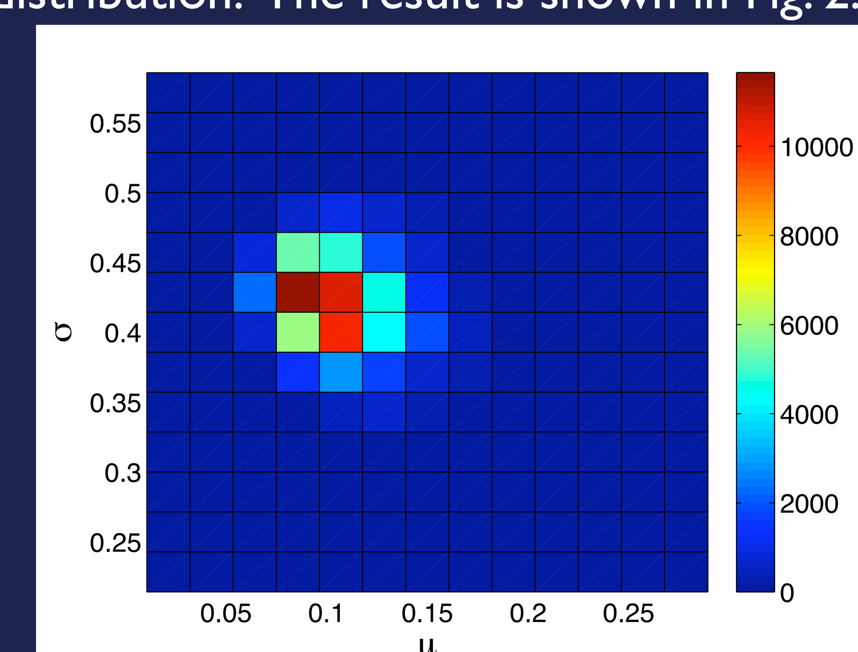

Figure 2 – Recovered joint distribution for μ and σ. The true probability distribution with mean 0.4 and standard deviation 0.1 was sampled with k=20 detected events. The PDF obtained for each event is a Gaussian centered on its true value with standard deviation taken uniformly from [a,b]=[0,0.2], sampled with N=100 points. The weighted averages of the recovered parameters over the posterior density of the reconstructed population distribution are 0.4055 for the mean and 0.0958 for the standard deviation.

We also explore the effects of the choices of k, N, and [a,b] on the accuracy of the recovery of the population distribution. We find that N=100 is a sufficient number of points to sample each PDF, and higher values of N do not improve accuracy. Varying the number of detected samples, k, we observe a clear anti-correlation between k and the accuracy of determining μ and σ: as expected, it's harder to reconstruct the population with fewer events. Fig. 3a shows that the error in the recovered population mean scales approximately as the inverse square root of k, while the standard deviation is over-estimated at low k because we are sensitive to the variance in individual PDFs. Finally, we study the impact of the accuracy of individual PDFs by varying [a,b] and find that the accuracy deteriorates significantly once the PDF width exceeds the standard deviation of the underlying population.

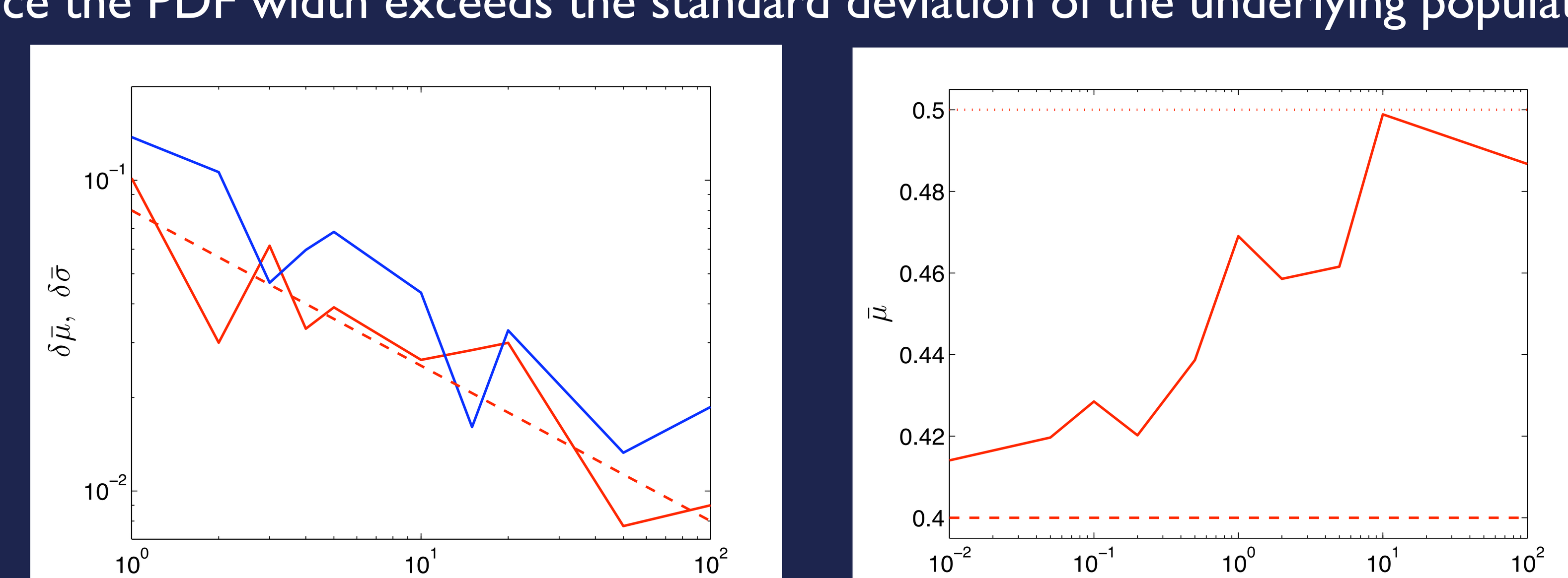

Figure 3 – The left panel (a) shows the error in the recovered mean (red) and standard deviation (blue) of the population distribution as a function of the number of detected events k. The dashed red curve is the approximate analytic fit to $k^{-1/2}$. The right panel (b) shows the estimated mean as a function of the statistical uncertainty in each event a=b. The dashed line is the true mean of the population distribution, 0.4, while the dotted line is the mean of the prior on μ.

Discussion

We introduced a robust Bayesian technique for estimating a modeled population distribution based on an arbitrary number of independent detected events, each with an associated posterior probability distribution function describing the degree of statistical uncertainty in parameter estimation. This technique can be readily deployed as an inexpensive post-processing tool, which can run in a tiny fraction of the time required to sample the PDFs of individual events.

To better apply this technique to GW astronomy, we should:

- Properly account for selection biases in the detected sample.
- Characterize systematic errors in parameter estimation for individual events.
- Choose model families for the population distribution based on prior astrophysical knowledge.
- Determine the number and accuracy of detections which would be necessary to constrain astrophysical models with the observed population distributions of specific parameters, e.g., masses and spins [see Mandel & O'Shaughnessy, 2009].