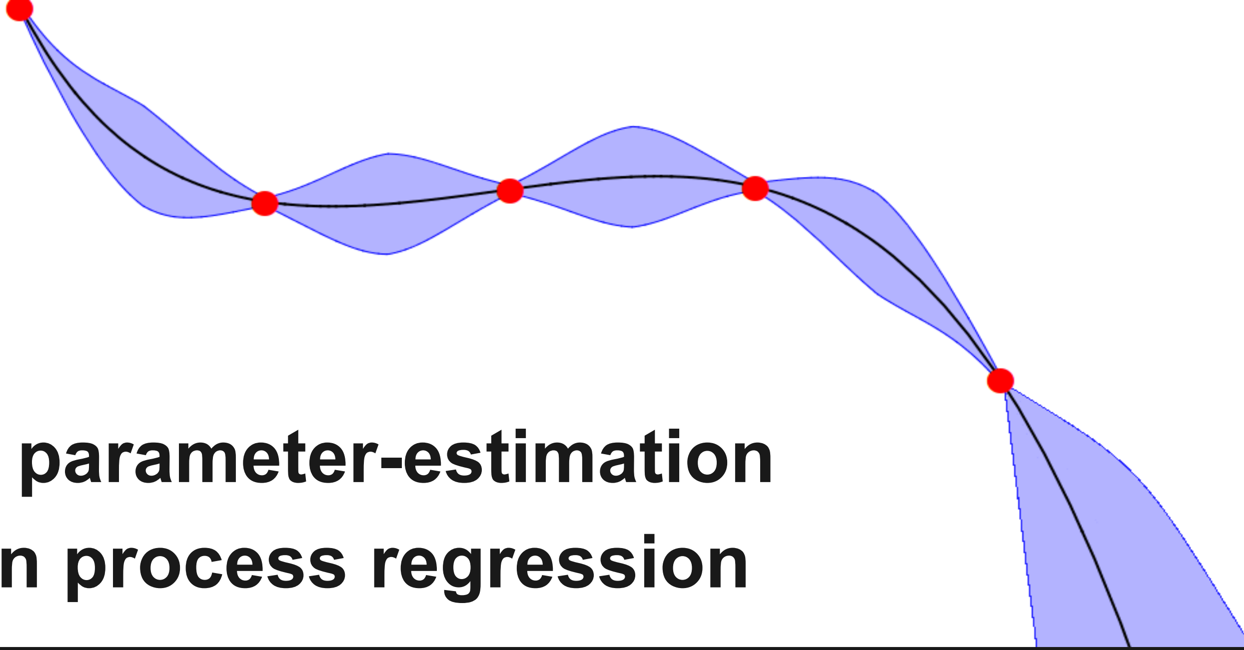


Improving GW parameter-estimation using Gaussian process regression

Christopher Moore
Institute of Astronomy, Cambridge, UK

20/08/2015

Work done in collaboration with Jonathan Gair, Christopher Berry, and Alvin Chua

Outline

- The problem with models
- The marginalised likelihood
- Implementation and results
- Summary

Data analysis preliminaries

GW data assumed to consist of a signal and noise.

$$s = h(\vec{\lambda}) + n$$

The key ingredient in any Bayesian detection or parameter estimation study is the **likelihood**,

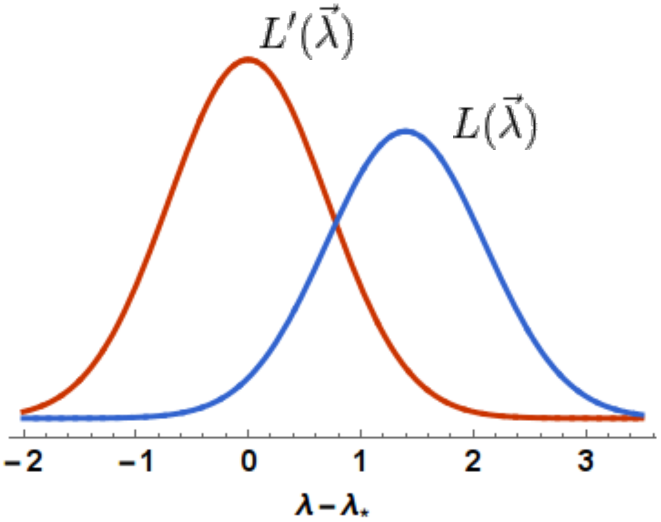
$$L'(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\langle s - h(\vec{\lambda}) | s - h(\vec{\lambda})\right\rangle\right) ,$$

$$\text{where} \quad \langle a|b\rangle = 4\Re\left\{\int_0^\infty \mathrm{d}f \, \frac{\tilde{a}(f)\tilde{b}(f)^*}{S_n(f)}\right\} .$$

But, we have to rely on approximate models.

$$H(\vec{\lambda}) = h(\vec{\lambda}) + \delta h(\vec{\lambda})$$

$$L(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\langle s - H(\vec{\lambda}) | s - H(\vec{\lambda})\right\rangle\right) \approx L'(\vec{\lambda})$$



The problems with models

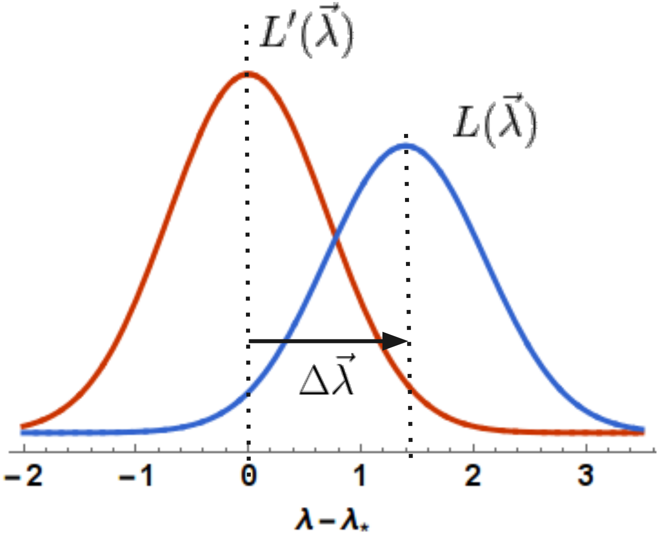
Two related problems with using approximate likelihood:

- **Detection**; reduced evidence
- **Parameter estimation**; shifted peak

Our focus is on the parameter estimation problem

Obvious solution is to develop better models!

Accurate (but not completely accurate) waveform models **do** exist, **but** very computationally expensive for exploring high dimensional parameter spaces.



$$\Delta\vec{\lambda}^a = -\left(\Gamma^{-1}\right)^{ab}\left\langle \delta h(\vec{\lambda}_*) \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle$$

Outline

- The problem with models
- The marginalised likelihood
- Implementation and results
- Summary

Marginalised likelihood

We propose the following alternative likelihood.

$$\mathcal{L}(\vec{\lambda}) \propto \int \mathrm{d}\left(\delta h(\vec{\lambda})\right)\; P(\delta h(\vec{\lambda})) \times \exp\left(-\frac{1}{2}\left\langle s - H(\vec{\lambda}) + \delta h(\vec{\lambda}) | s - H(\vec{\lambda}) + \delta h(\vec{\lambda})\right\rangle\right)$$

This likelihood uses the full waveform model, but has marginalised over the unknown part.

Two steps needed to evaluate this function: **(i) specify the prior (ii) perform the integral.**

If the final likelihood is to be useful in an MCMC-type search, it must not be any slower than standard techniques. In particular, the integration must not slow down the evaluation.

Moore & Gair (2014), PRL 113, 251101, arXiv:1412.3657

Specifying the prior: GPR

The prior is formed by interpolating a set of waveform differences precomputed

$$\mathcal{D} = \left\{ \left(\vec{\lambda}_i, \delta h(\vec{\lambda}_i) \right) \mid i = 1, 2, \ldots, n \right\}$$

GPR is used for the interpolation.

- Non-parametric
- Training to learn properties
- Allows for analytic marg

$$k(\vec{\lambda}_i, \vec{\lambda}_j) = \sigma_f \exp\left(-\frac{1}{2} g_{ab} (\vec{\lambda}_i - \vec{\lambda}_j)^a (\vec{\lambda}_i - \vec{\lambda}_j)^b \right)$$

At some new point in parameter space, $\vec{\lambda}$, GPR returns a Gaussian distribution for the waveform error at that point.

$$P(\delta h(\vec{\lambda})) = \exp\left(- \frac{\left\langle \delta h(\vec{\lambda}) - \mu(\lambda) | \delta h(\vec{\lambda}) - \mu(\lambda) \right\rangle}{2\sigma^2(\vec{\lambda})} \right)$$

$$\mu(\vec{\lambda}) = k(\vec{\lambda}, \vec{\lambda}_i) k^{-1}(\vec{\lambda}_i, \vec{\lambda}_j) \delta h(\vec{\lambda}_j)$$
$$\sigma^2(\vec{\lambda}) = k(\vec{\lambda}, \vec{\lambda}) - k(\vec{\lambda}, \vec{\lambda}_i) k^{-1}(\vec{\lambda}_i, \vec{\lambda}_j) k(\vec{\lambda}, \vec{\lambda}_i)$$

Performing the integral

GPR returns a probability distribution for the waveform difference, which is a Gaussian.

$$P(\delta h(\vec{\lambda})) = \exp\left(-\frac{\left\langle \delta h(\vec{\lambda}) - \mu(\lambda) | \delta h(\vec{\lambda}) - \mu(\lambda) \right\rangle}{2\sigma^2(\vec{\lambda})}\right)$$

The Marginalised likelihood was defined by the following Gaussian integral.

$$\mathcal{L}(\vec{\lambda}) \propto \int \mathrm{d}\left(\delta h(\vec{\lambda})\right) \; P(\delta h(\vec{\lambda})) \times \exp\left(-\frac{1}{2}\left\langle s - H(\vec{\lambda}) + \delta h(\vec{\lambda}) | s - H(\vec{\lambda}) + \delta h(\vec{\lambda}) \right\rangle\right)$$

This may be evaluated analytically to give the following expression

$$\mathcal{L}(\vec{\lambda}) \propto \frac{1}{\sqrt{1 + \sigma^2(\vec{\lambda})}} \exp\left(-\frac{1}{2}\frac{\langle s - H + \mu | s - H + \mu\rangle}{1 + \sigma^2(\vec{\lambda})}\right)$$

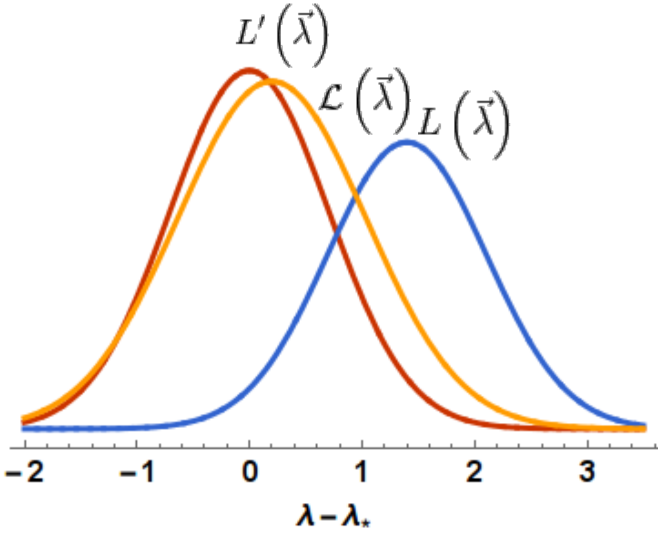
The marginalised likelihood

- Shifts the likelihood into better agreement with the true parameters

- Broadens the posterior to reflect the level of confidence we have in the results

- Even in the limit of large signal strength (when systematic model errors normally dominate over random error) posterior is never inconsistent with true parameters

- The broadening of the posterior reduces the bias in parameter estimation



$$\mathcal{L}(\vec{\lambda}) \propto \frac{1}{\sqrt{1+\sigma^2(\vec{\lambda})}} \exp\left(-\frac{1}{2} \frac{\langle s - H + \mu | s - H + \mu \rangle}{1+\sigma^2(\vec{\lambda})} \right)$$

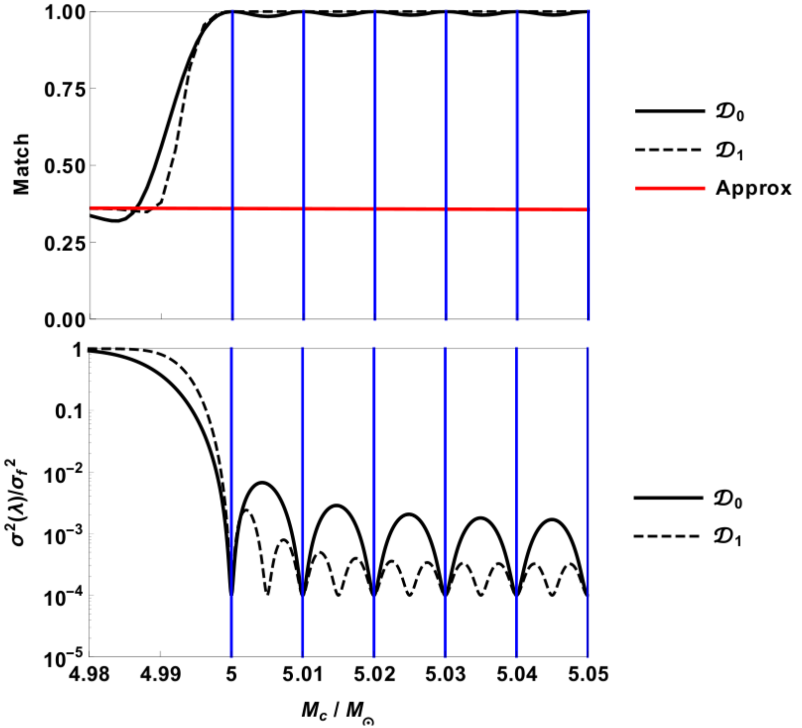
Moore & Gair (2014), PRD 91, 124062, arXiv:1504.02767

Outline

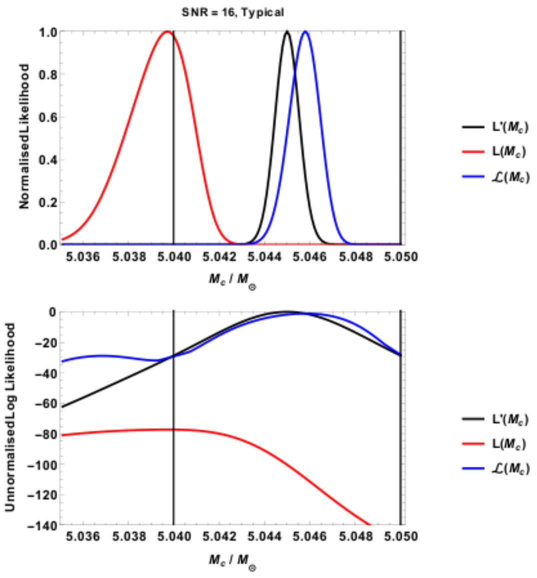
- The problem with models
- The marginalised likelihood
- Implementation and results
- Summary

Implementation

- Choice of model waveforms: accurate model IMRPhenomC, approximate model TaylorF2
- For simplicity, and to aid in developing new method, restrict to 1D interpolation in Chirp mass. (Symmetric mass ratio fixed to ~¼.)
- Two training sets were used with n=60 and n=120 points in range $M_c \in (5\text{-}5.6) M_\odot$
- A squared exponential covariance was found to perform best, with a typical length scale of $\sim 0.01 M_\odot$

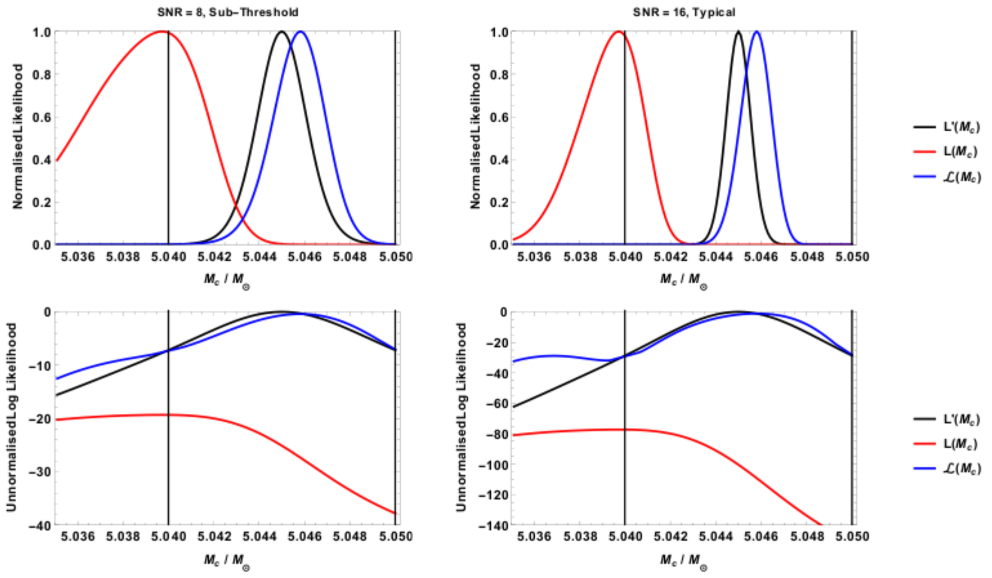


Results

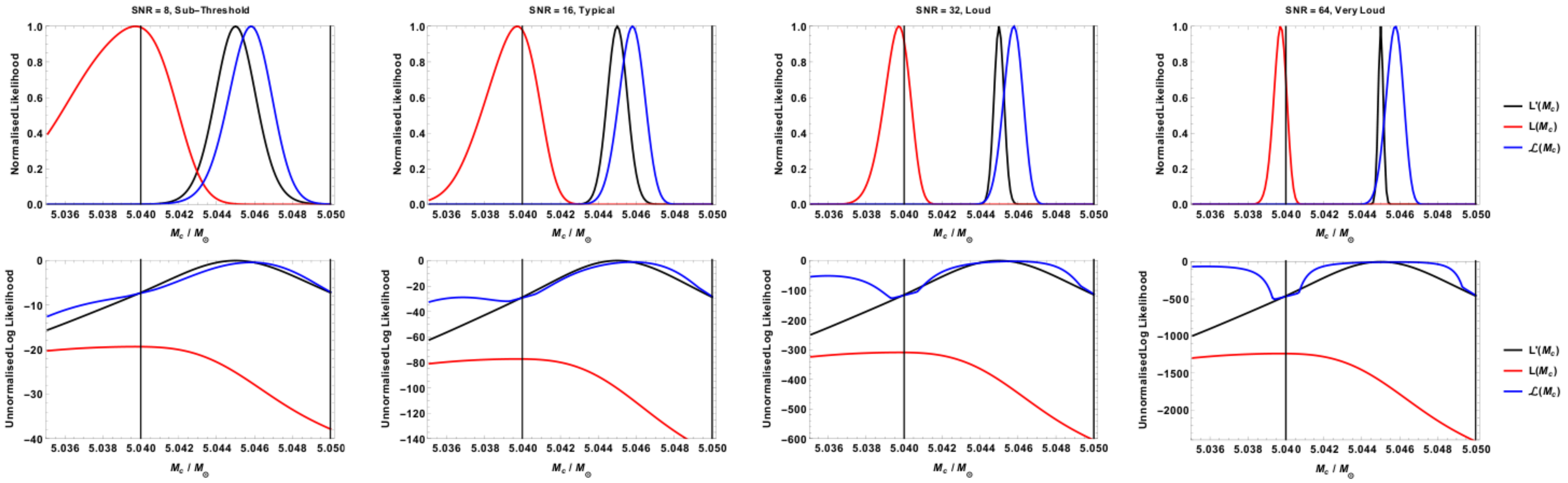


SNR = 16, Typical
Normalised Likelihood
1.0
0.8
0.6
0.4
0.2
0.0
5.036 5.038 5.040 5.042 5.044 5.046 5.048 5.050
M_c / $M_\odot$
L'(M_c)
L(M_c)
$\mathcal{L}$(M_c)
Unnormalised Log Likelihood
0
−20
−40
−60
−80
−100
−120
−140
5.036 5.038 5.040 5.042 5.044 5.046 5.048 5.050
M_c / $M_\odot$
L'(M_c)
L(M_c)
$\mathcal{L}$(M_c)

Results


SNR = 8, Sub-Threshold
SNR = 16, Typical
Normalised Likelihood
Unnormalised Log Likelihood
M_c / $M_\odot$
L'(M_c)
L(M_c)
$\mathcal{L}$(M_c)

Results


SNR = 8, Sub-Threshold
SNR = 16, Typical
SNR = 32, Loud
SNR = 64, Very Loud
Normalised Likelihood
Unnormalised Log Likelihood
M_c / $M_\odot$
L'(M_c)
L(M_c)
$\mathcal{L}$(M_c)

Outline

- The problem with models
- The marginalised likelihood
- Implementation and results
- Summary

Summary

- Model errors are a known problem for advanced LIGO, particularly for high mass binary black hole systems.

- The marginalised likelihood **(i)** reduces the size of the error and **(ii)** properly accounts for any remaining error.

- In this paper we...
  - Provided a detailed description of the marginalised likelihood for the first time
  - Implement the method using approximants from LAL
  - Explore the effects of different choices for the GPR training set and covariance function on the method
  - Demonstrate that the marginalised likelihood works for binary black holes at realistic signal amplitudes

Thank you for listening!