Parameter estimation of spinning binary black-hole inspirals using MCMC

Marc van der Sluys¹

Vivien Raymond¹, Ilya Mandel¹,
Christian Röver²,³, Alexander Stroeer⁴, Nelson Christensen⁵,
Vicky Kalogera¹, Alberto Vecchio¹,⁶, John Veitch⁶, Renate Meyer²

¹Northwestern University, ²University of Auckland, ³AEI Hannover, ⁴Goddard SFC, ⁵Carleton College,
⁶University of Birmingham

Penn State, March 27, 2008
Outline

1. Introduction
   - Goals
   - Waveform and noise

2. MCMC
   - Likelihood calculation
   - Markov chains
   - MCMC setup

3. Sampling modes
   - Results
   - Dependence on number of detectors and spin
   - Accuracy of parameter estimation

4. Finding modes
   - Offset runs
   - Spins, correlations and structure
   - Improving sampling

5. Future work
Goals of this project

Intermediate goals

- Show that Markov-Chain Monte Carlo (MCMC) with a large number of parameters (> 10) on LIGO data can be done
- Test MCMC code on software and hardware injections

Final goals

- Do parameter estimation on LIGO detection of inspiral signal
- Use as a follow-up for template-based search to:
  - Confirm spinning inspiral nature of signal
  - Determine physical parameters (masses, spin, position, ...)
- Provide final stage in automated CBC pipeline
Astrophysical goals

Populations of compact binaries

- Mass distributions
- Spins of BHs; alignment of spins
- Association of GW and EM events, e.g. GRB
- Empirical merger rates
- NS-NS/BH-NS/BH-BH merger ratios

Evolution of massive binaries

- Evolution of massive stars (in binaries)
- Constraints on CE evolution
- Initial-mass range for BH progenitors
## Predicted detection rates

### Realistic estimate:

<table>
<thead>
<tr>
<th></th>
<th>Rates (yr(^{-1}))</th>
<th>Horizon (Mpc)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS-NS</td>
<td>BH-NS</td>
</tr>
<tr>
<td>Initial</td>
<td>0.015</td>
<td>0.004</td>
</tr>
<tr>
<td>Enhanced</td>
<td>0.15</td>
<td>0.04</td>
</tr>
<tr>
<td>Advanced</td>
<td>20</td>
<td>5.7</td>
</tr>
</tbody>
</table>

### Plausible, optimistic estimate:

<table>
<thead>
<tr>
<th></th>
<th>Rates (yr(^{-1}))</th>
<th>Horizon (Mpc)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS-NS</td>
<td>BH-NS</td>
</tr>
<tr>
<td>Initial</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>Enhanced</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Advanced</td>
<td>200</td>
<td>190</td>
</tr>
</tbody>
</table>

Estimates assume \(M_{\text{NS}} = 1.4 M_\odot\) and \(M_{\text{BH}} = 10 M_\odot\)

CBC group, rates document
Spinning BH binaries: Simple waveform

- Röver non-spinning code
- Waveform template:
  - Analytic waveform
  - Restricted 1.5 PN
  - Simple precession
  - 12-parameter set: $\vec{\lambda}$

Typical data stretch ($f_{\text{low}}$ – coalescence):
5.5s, 400 wave cycles, 5 precession cycles

$$M_1 = 10.0M_\odot, \quad M_2 = 1.4M_\odot, \quad a_{\text{spin}} = 0.5, \quad d_L = 13.0\text{Mpc}$$
Detector noise

- Using 1–3 detectors from L1, H1, and Virgo
- Gaussian, stationary noise, at designed sensitivity level
- Noise is uncorrelated between detectors
Detector noise

Using 1–2 4-km detectors L1, H1:
- Gaussian, stationary noise
- LIGO S5 playground data
The game:

- Do software injections
- Retrieve physical parameters

Here, $\Sigma \text{SNR} = 17$
Compute posterior distribution

- Find posterior density of the model parameters
- Bayesian approach
- Coherent network of detectors:
  - PDF$(\vec{\lambda}) \propto \text{prior}(\vec{\lambda}) \times \prod_i L_i(d|\vec{\lambda})$
- The likelihood for each detector $i$ is:

  $$L_i(d|\vec{\lambda}) \propto \exp \left( -2 \int_0^\infty \frac{|\tilde{d}(f) - \tilde{m}(\vec{\lambda}, f)|^2}{S_n(f)} df \right)$$

- Use Markov-Chain Monte Carlo to sample the posterior
Markov Chains

- Choose starting point for chain: $\lambda_1$
- Calculate its likelihood: $L_j \equiv L(d|\lambda_j)$
- do $j = 1, N$
  - draw random jump size $\Delta \lambda_j$ from Gaussian with $\sigma$
  - consider new state $\lambda_{j+1} = \lambda_j + \Delta \lambda_j$
  - calculate $L_{j+1} \equiv L(d|\lambda_{j+1})$
  - if($\frac{L_{j+1}}{L_j} > \text{ran}_\text{unif}[0,1]$) then
    - Accept new state $\lambda_{j+1}$
    - Increase jump size $\sigma$
  - else
    - Reject new state; $\lambda_{j+1} = \lambda_j$
    - Decrease jump size $\sigma$
  - end if
- save state $\lambda_{j+1}$
- end do ($j$)
Correlated update proposals

Problem

- Often (strong) correlations exist
- Correlations make random jump proposals very inefficient

Solution

- Calculate covariance matrix from previous block of iterations
- Propose jumps according to these correlations
# MCMC runs – setup

## MCMC code
- Adaptive random-walk Metropolis sampler
- 12 parameters: masses: $M$ & $\eta$, distance: $\log d_L$, time and phase at coalescence: $\varphi_c$ & $t_c$, position: R.A. & Dec, spin magnitude: $a_{\text{spin}}$, angle between $\vec{S}$ and $\vec{L}$: $\theta_{SL}$, precession phase: $\alpha_c$, orientation of $J_0$: $\sin \theta_{J_0}$ & $\varphi_{J_0}$
- Software injections in simulated, Gaussian noise or (hopefully) clean S5 playground data

## MCMC runs
- Start chain from *true parameter values* (short burn-in) to assess efficiency of sampling the PDF
- Start chain from *offset values* to determine speed and quality of mode detection
Correlated MCMC

**Set-up**
- Use 80% correlated update proposals – more efficient
- Chains presented here, for 1 & 2 LIGO detectors:
  - Length: 7; $3 \times 10^6$ states
  - Burn-in $10^6$; $5 \times 10^5$ states
  - Run time: 10 days on a 2.8 GHz CPU
- 5 serial chains from the true values (one per CPU)

**Signal parameters**
- Fiducial binary: $M_{1,2} = 10 + 1.4 M_\odot$, $d_L = 16 - 21$ Mpc
- Spin: $a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8$, $\theta_{\text{SL}} = 20^\circ, 55^\circ$
- Using H1 @ SNR $\approx 12.7$, H1L1 @ SNR $\approx 17.0$
- Signals injected in simulated Gaussian noise
Example MCMC run

\( M_c (M_\odot) \)

Model: 2.994
Median: 2.994
\( \Delta_{90\%} \): 0.46%

Iteration: 1.00E+06

Parameters:
- H1 & L1
- \( M : 10, 1.4 M_\odot \)
- \( a_{\text{spin}} = 0.5, \theta_{\text{SL}} = 20^\circ \)
- \( \Sigma \text{SNR} \approx 17.7 \)
Results: 1 detector

Parameters:
- H1 only
- \( M = 10, 1.4 \, M_\odot \)
- \( d_L = 18.7 \, \text{Mpc} \)
- \( a_{\text{spin}} = 0.5 \)
- \( \theta_{\text{SL}} = 20^\circ \)
- Network SNR \( \approx 12.7 \)
- \( \Delta \)'s are 90% probability
- Dashed lines show true values
**Results: 2 detectors**

**Parameters:**
- H1 & L1
- $M = 10, 1.4 \, M_\odot$
- $d_L = 18.7 \, \text{Mpc}$
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 17.0$
- $\Delta$’s are 90% probability
- Dashed lines show true values
Run without signal

Parameters:
- H1 only
- Gaussian noise was used
- MCMC run was started as usual, but no signal was injected
Changing spin: 1 detector

Parameters:
- H1 only
- $M = 10, 1.4 \, M_\odot$
- $d_L \approx 16 - 21 \, \text{Mpc}$
- $a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\text{SL}} = 20^\circ$
- SNR $\approx 12.7$
- Dashed lines show true values
Changing spin: 2 detectors

Parameters:

- H1 & L1
- \( M = 10, 1.4 \, M_\odot \)
- \( d_L \approx 16 - 21 \, \text{Mpc} \)
- \( a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8 \)
- \( \theta_{\text{SL}} = 20^\circ \)
- Network SNR \( \approx 17.0 \)
- Dashed lines show true values
Changing the number of detectors

Parameters:
- H1, H1 & L1
- $M = 10, 1.4 \, M_\odot$
- $d_L = 18.7 \, \text{Mpc}$
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 12.7, 17.0$
- Dashed lines show true values
Sky map: 1 detector

Parameters:
- H1 only
- $M = 10, 1.4 M_\odot$
- $d_L \approx 16 - 21 \text{ Mpc}$
- $a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\text{SL}} = 20^\circ$
- SNR $\approx 12.7$
- Dashed lines show true position
Sky map: 2 detectors

Parameters:
- H1 & L1
- $M = 10, 1.4 \, M_\odot$
- $d_L \approx 16 - 21 \, \text{Mpc}$
- $a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 17.0$

Dashed lines show true position
2D PDF: masses

Parameters:
- H1 & L1
- $M = 10, 1.4 \ M_\odot$
- $d_L \approx 16 - 21 \ Mpc$
- $a_{\text{spin}} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{SL} = 20^\circ$
- Network SNR $\approx 17.0$
- Dashed lines show true position
## Results

Width of the 90%-probability ranges ($\Delta_{90\%}$):

<table>
<thead>
<tr>
<th>$n_{\text{det}}$</th>
<th>$a_{\text{spin}}$</th>
<th>$\theta_{\text{SL}}$ ($^\circ$)</th>
<th>$d_L$ (Mpc)</th>
<th>SNR</th>
<th>$M_1^a$ (%)</th>
<th>$M_2^a$ (%)</th>
<th>$t_c$ (s)</th>
<th>$d_L$ (%)</th>
<th>$a_{\text{spin}}$ (%)</th>
<th>$\theta_{\text{SL}}$ ($^\circ$)</th>
<th>RA$^b$ ($^\circ$)</th>
<th>Decl. ($^\circ$)</th>
<th>$\theta_{J_0}$ (°)</th>
<th>$\varphi_{J_0}$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>0</td>
<td>13.6</td>
<td>12.7</td>
<td>85</td>
<td>65</td>
<td>0.042</td>
<td>150</td>
<td>200</td>
<td>157**</td>
<td>241</td>
<td>119</td>
<td>158</td>
<td>326</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>20</td>
<td>12.7</td>
<td>12.7</td>
<td>52**</td>
<td>41*</td>
<td>0.041</td>
<td>156</td>
<td>194</td>
<td>133**</td>
<td>248</td>
<td>135</td>
<td>132</td>
<td>320</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>55</td>
<td>12.3</td>
<td>12.7</td>
<td>34*</td>
<td>25*</td>
<td>0.023</td>
<td>85</td>
<td>185</td>
<td>126</td>
<td>75</td>
<td>94</td>
<td>52</td>
<td>354</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>20</td>
<td>13.8</td>
<td>12.7</td>
<td>79</td>
<td>64</td>
<td>0.040</td>
<td>143</td>
<td>127</td>
<td>89</td>
<td>254</td>
<td>108</td>
<td>89</td>
<td>259</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>55</td>
<td>18.8</td>
<td>12.7</td>
<td>64</td>
<td>48</td>
<td>0.022</td>
<td>100</td>
<td>67</td>
<td>79</td>
<td>63</td>
<td>29</td>
<td>20</td>
<td>93</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>20</td>
<td>14.7</td>
<td>12.7</td>
<td>80*</td>
<td>62</td>
<td>0.027</td>
<td>117</td>
<td>29</td>
<td>39</td>
<td>94</td>
<td>88</td>
<td>60</td>
<td>271</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>55</td>
<td>20.9</td>
<td>12.7</td>
<td>102</td>
<td>83</td>
<td>0.024</td>
<td>113</td>
<td>58</td>
<td>75</td>
<td>150</td>
<td>93</td>
<td>43</td>
<td>255</td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>0</td>
<td>13.5</td>
<td>17.0</td>
<td>66</td>
<td>49</td>
<td>0.028</td>
<td>92</td>
<td>200</td>
<td>167**</td>
<td>80</td>
<td>83</td>
<td>154</td>
<td>323</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>20</td>
<td>13.0</td>
<td>17.0</td>
<td>41*</td>
<td>32*</td>
<td>0.015**</td>
<td>72</td>
<td>170</td>
<td>120*</td>
<td>72*</td>
<td>76</td>
<td>120</td>
<td>354</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>55</td>
<td>13.5</td>
<td>17.0</td>
<td>35**</td>
<td>27</td>
<td>0.008</td>
<td>40</td>
<td>189</td>
<td>115*</td>
<td>3.6</td>
<td>23</td>
<td>23*</td>
<td>8.2</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>20</td>
<td>15.2</td>
<td>17.0</td>
<td>48</td>
<td>37</td>
<td>0.006</td>
<td>33</td>
<td>16</td>
<td>38</td>
<td>3.0</td>
<td>15</td>
<td>17</td>
<td>9.1</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>55</td>
<td>20.8</td>
<td>17.0</td>
<td>43</td>
<td>32</td>
<td>0.006</td>
<td>54</td>
<td>51</td>
<td>65</td>
<td>3.0</td>
<td>12*</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>0.8</td>
<td>20</td>
<td>16.2</td>
<td>17.0</td>
<td>49</td>
<td>37</td>
<td>0.006</td>
<td>40</td>
<td>15</td>
<td>24</td>
<td>3.8</td>
<td>18</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>0.8</td>
<td>55</td>
<td>23.2</td>
<td>17.0</td>
<td>33</td>
<td>25</td>
<td>0.006</td>
<td>57</td>
<td>29</td>
<td>26</td>
<td>3.3</td>
<td>10</td>
<td>9.2</td>
<td>16</td>
</tr>
</tbody>
</table>

*: The true value lies outside the 90%-probability range, but inside 95%.

**: The true value lies outside the 95%-probability range, but inside 99%.

$^a$: The values of $M_1$ and $M_2$ are derived from $M_1$ and $\eta$, used in the MCMC code.

$^b$: The column RA shows the value $\Delta_{90\%} \cdot \cos 40^\circ$, ($40^\circ$ is the declination of the source) and converted to degrees to make the value comparable to that of the declination.
Conclusions

Accuracies:

- Detection with 1 detector: degeneracy in sky position and binary orientation:
  - no or low spin: whole sky/all directions
  - intermediate or high spin: multimodal distribution
- Detection with 2 detectors can produce astronomically relevant information:
  - individual masses and spin with $\sim 30 - 40\%$ accuracy
  - distance with $\sim 40\%$ accuracy
  - position and orientation down to typically $10 - 20^\circ$
  - timing better than 0.01s
- Combination of the above can lead to association with E&M detection (e.g. gamma-ray burst)
Finding the modes of the PDFs

Offset start

- Start chains with offset initial parameter values
- Choose initial values randomly from a range around the true values
- Typical offset: \( M \sim 0.1 M_\odot, \ t_c \sim 0.03s, \) rest: \( \sim \) random

Efficiency

- True modes will \textit{eventually} be found by the chains
- Keyword: \textbf{efficiency} of sampling: how to we find the modes within \textit{e.g.} a Hubble time?
- This becomes a more important issue for higher spin
Correlations increase with spin

Parameters:
- H1 & L1
- $M_1 = 10\, M_\odot$
- $M_2 = 1.4\, M_\odot$
- $d_L = 13\, \text{Mpc}$
- $a_{\text{spin}} = 0.1, 0.8$
- $\theta_{\text{SL}} = 55^\circ$
- Network SNR $\approx 18.2, 30.5$
Structured parameter space

Parameters:
- H1 & L1
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 27.2$
- 10 chains
- Offset start
- Black dashed lines are true values
Structured parameter space

Parameters:
- H1 & L1
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 27.2$
- 10 chains
- Offset start
- Black dashed lines are true values
Parallel tempering

### Parallel chains

- Use $\sim$5-10 parallel chains of temperatures $T = 1, \ldots, T_{\text{max}}$

- Acceptance probability for chain with temperature $T$: $\left(\frac{L_i}{L_{i-1}}\right)^{\frac{1}{T}}$

- Hotter chains explore wider ranges, at lower likelihood

- Probability for swap between chains: $\left(\frac{L_h}{L_c}\right)^{\frac{1}{T_c} - \frac{1}{T_h}}$, $T_h > T_c$

- Hotter chains pass information to cooler chains
Converging chains

Parameters:
- H1 & L1
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20^\circ$
- Network SNR $\approx 17.7$
- 4 chains
- Offset start
- Black dashed lines are true values
Improve sampling

Included techniques
- Parallel tempering
- Mix of uncorrelated and correlated updates
- Extra-large steps

Planned techniques
- Partial updates of only intrinsic/extrinsic parameters
- ‘Smart’ updates:
  - use knowledge of waveform to identify near-degenerate islands
  - take large steps top hop islands
  - beach-to-mountain-top routine
Conclusions

Sampling modes

- Our code samples PDFs fine, using one or multiple detectors, for no, small or high spin.
- We can give a good indication of the expected accuracies with which the astrophysical parameters of the binary can be determined.
- For two or more detectors, the accuracy of $t_c$, position and distance is good enough for association with E&M detection.

Finding modes

- For intermediate or high spin, parameter space is strongly structured.
- Strong correlations between parameters demand efficient, perhaps even ‘smart’ sampling.
Future work

MCMC wish list

- Keep improving sampling efficiency, find modes faster
- Explore wider range of parameters
- Improve signal:
  - more realistic inspiral (Vivien):
    - add second spin
    - higher PN
  - add ring-down and merger
  - use NR waveforms with physical parameters

CBC pipeline

- Add MCMC to data-analysis pipeline
- Map parameters of filter triggers into priors for MCMC
- Include noise as one of the unknown parameters