Parameter estimation of spinning binary black-hole inspirals using MCMC

Marc van der Sluys[1]

Vivien Raymond[1], Ilya Mandel[1],
Christian Röver[2,3], Alexander Stroeer[4], Nelson Christensen[5],
Vicky Kalogera[1], Alberto Vecchio[1,6], John Veitch[6], Renate Meyer[2]
[1]Northwestern University, [2]University of Auckland, [3]AEI Hannover, [4]Goddard SFC, [5]Carleton College,
[6]University of Birmingham

Penn State, March 27, 2008

NORTHWESTERN UNIVERSITY

Outline

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:

	Rates (yr^{-1})			Horizon (Mpc)		
	NS-NS	BH-NS	BH-BH	NS-NS	BH-NS	BH-BH
Initial	0.015	0.004	0.01	32	67	160
Enhanced	0.15	0.04	0.11	71	149	349
Advanced	20	5.7	16	364	767	1850

Plausible, optimistic estimate:

	Rates (yr^{-1})			Horizon (Mpc)		
	NS-NS	BH-NS	BH-BH	NS-NS	BH-NS	BH-BH
Initial	0.15	0.13	1.7	32	67	160
Enhanced	1.5	1.4	18	71	149	349
Advanced	200	190	2700	364	767	1850

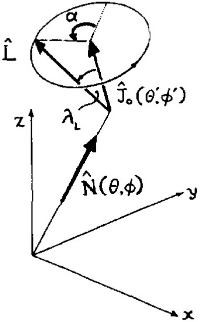
Estimates assume $M_{\mathrm{NS}} = 1.4\,M_{\odot}$ and $M_{\mathrm{BH}} = 10\,M_{\odot}$

CBC group, rates document

NORTHWESTERN UNIVERSITY

Spinning BH binaries: Simple waveform

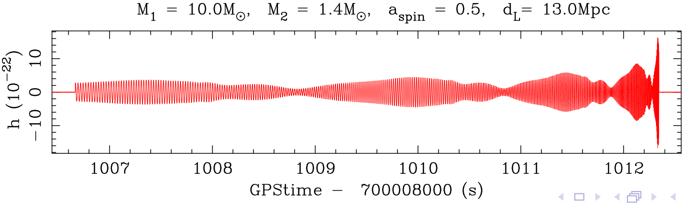
- Röver non-spinning code
- Waveform template:
  - Analytic waveform
  - Restricted 1.5 PN
  - Simple precession
  - 12-parameter set: $\vec{\lambda}$



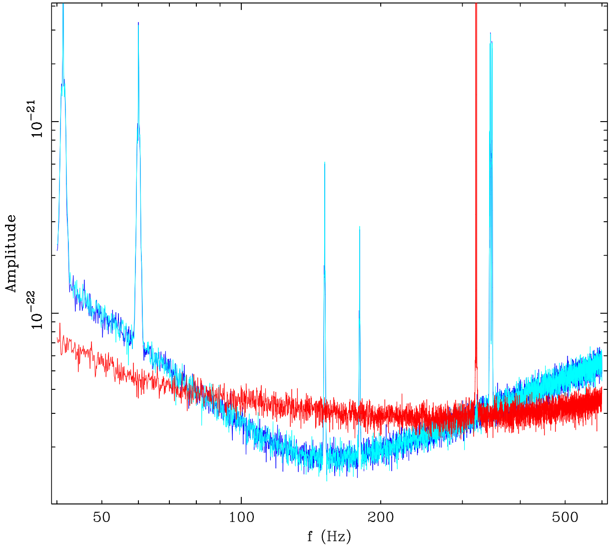
Apostolatos et al., 1994

Typical data stretch (f_{low} – coalescence):

5.5s, 400 wave cycles, 5 precession cycles



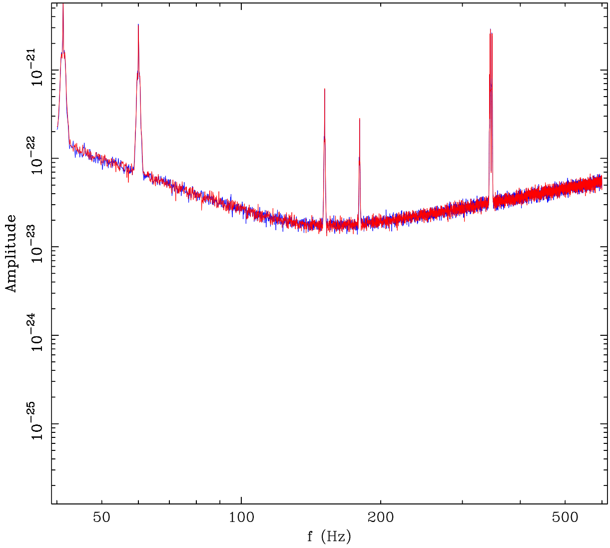
Detector noise



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

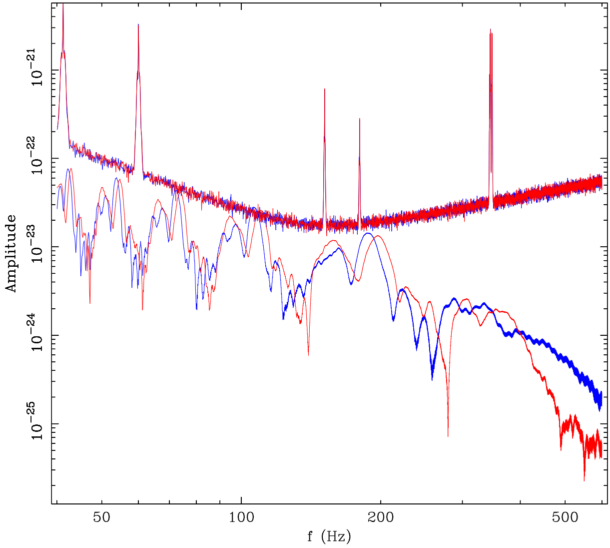
NORTHWESTERN UNIVERSITY

Detector noise



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

Detector noise



The game:

- Do software injections
- Retrieve physical parameters

Here, ΣSNR = 17

NORTHWESTERN UNIVERSITY

Compute posterior distribution

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

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

- Use Markov-Chain Monte Carlo to sample the posterior

Markov Chains

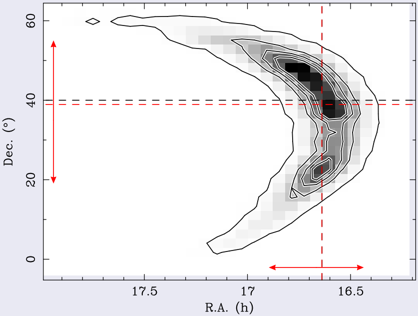
- Choose starting point for chain: $\vec{\lambda}_1$
- Calculate its likelihood: $L_j \equiv L(d|\vec{\lambda}_j)$
- do $j = 1, N$
    - draw random jump size $\Delta\vec{\lambda}_j$ from Gaussian with $\vec{\sigma}$
    - consider new state $\vec{\lambda}_{j+1} = \vec{\lambda}_j + \Delta\vec{\lambda}_j$
    - calculate $L_{j+1} \equiv L(d|\vec{\lambda}_{j+1})$
    - if($\frac{L_{j+1}}{L_j} >$ ran_unif[0,1]) then
        - Accept new state $\vec{\lambda}_{j+1}$
        - Increase jump size $\vec{\sigma}$
    - else
        - Reject new state; $\vec{\lambda}_{j+1} = \vec{\lambda}_j$
        - Decrease jump size $\vec{\sigma}$
    - end if
    - save state $\vec{\lambda}_{j+1}$
- end do (j)

NORTHWESTERN UNIVERSITY

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: $\mathcal{M}$ & η, distance: $\log d_{\rm L}$, time and phase at coalescence: $\varphi_{\rm c}$ & $t_{\rm c}$, position: R.A. & Dec, spin magnitude: $a_{\rm spin}$, angle between $\vec{S}$ and $\vec{L}$: $\theta_{\rm SL}$, precession phase: $\alpha_{\rm c}$, orientation of J_0: $\sin\theta_{\rm J_0}$ & $\varphi_{\rm J_0}$
- 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

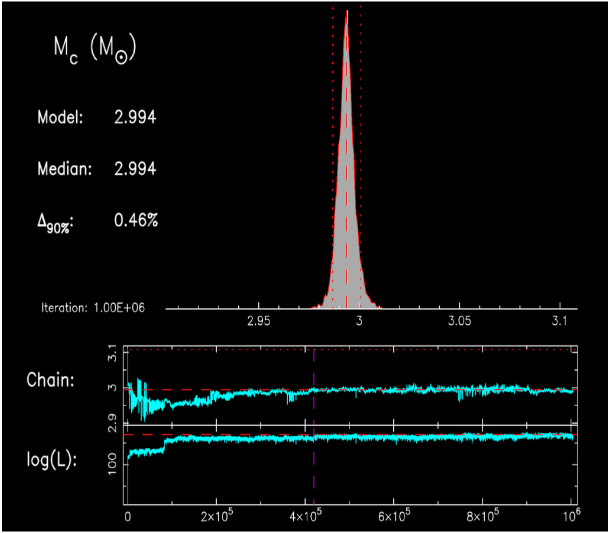
Set-up

- Use 80% correlated update proposals – more efficient
- Chains presented here, for 1 & 2 LIGO detectors:
  - Length: 7; 3×10^6 states
  - Burn-in 10^6; 5×10^5 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: $M_{1,2} = 10 + 1.4\,M_\odot$, $d_\mathrm{L} = 16 - 21$ Mpc
- Spin: $a_\mathrm{spin} = 0.0, 0.1, 0.5, 0.8$, $\theta_\mathrm{SL} = 20^\circ, 55^\circ$
- Using H1 @ SNR $\approx$12.7, H1L1 @ SNR $\approx$17.0
- Signals injected in simulated Gaussian noise

Example MCMC run

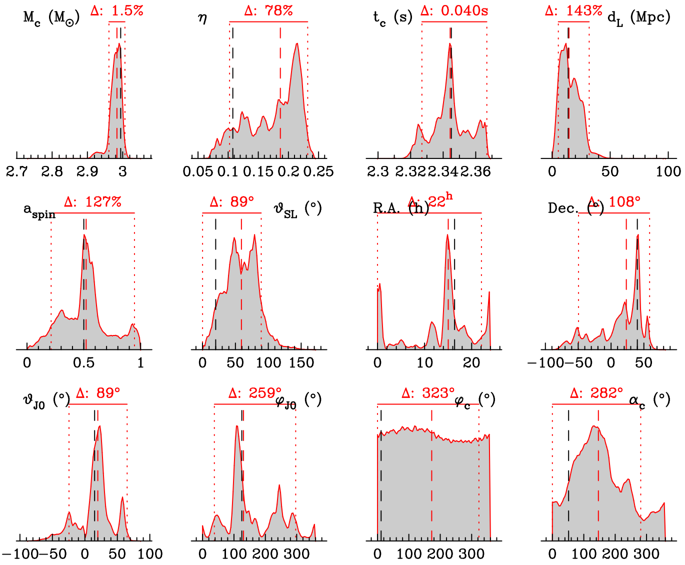


Parameters:

- H1 & L1
- M : $10, 1.4\, M_\odot$
- $a_{\rm spin} = 0.5, \theta_{\rm SL} = 20^\circ$
- ΣSNR ≈ 17.7

NORTHWESTERN UNIVERSITY

Results: 1 detector

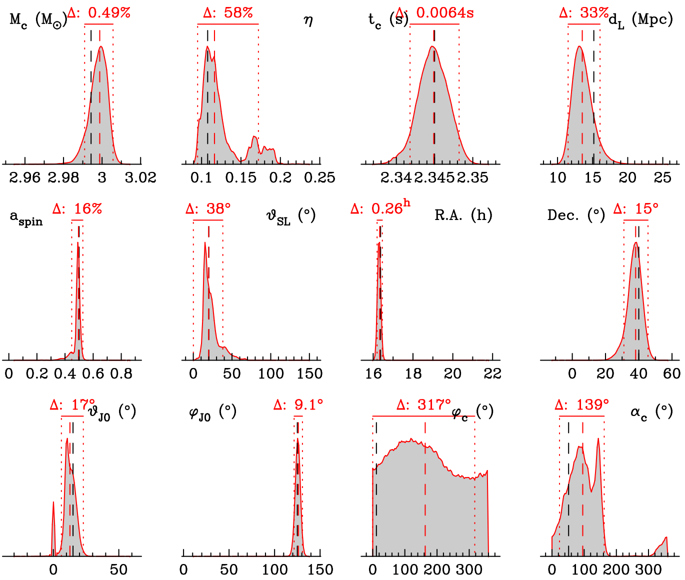


Parameters:

- H1 only
- $M = 10, 1.4\, M_\odot$
- $d_L = 18.7\,\mathrm{Mpc}$
- $a_{\mathrm{spin}} = 0.5$
- $\theta_{\mathrm{SL}} = 20^\circ$
- Network SNR ≈ 12.7
- Δ's are 90% probability
- Dashed lines show true values

NORTHWESTERN UNIVERSITY

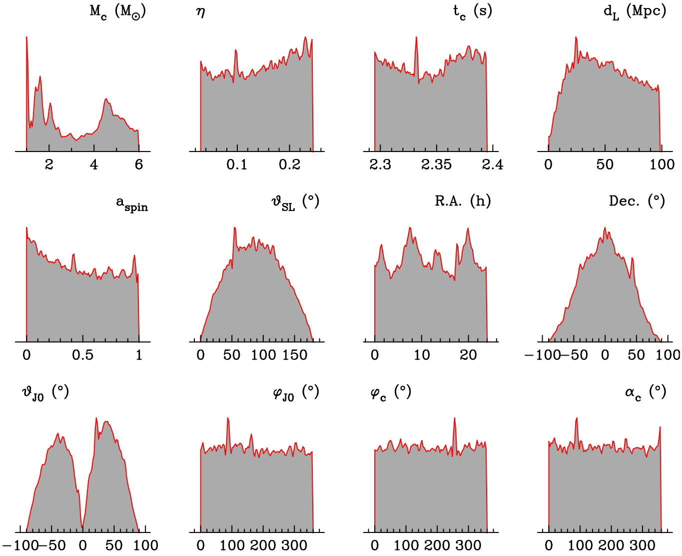
Results: 2 detectors



Parameters:

- H1 & L1
- $M = 10, 1.4\, M_\odot$
- $d_L = 18.7\,\mathrm{Mpc}$
- $a_{\mathrm{spin}} = 0.5$
- $\theta_{\mathrm{SL}} = 20^\circ$
- Network SNR ≈ 17.0
- Δ's are 90% probability
- Dashed lines show true values

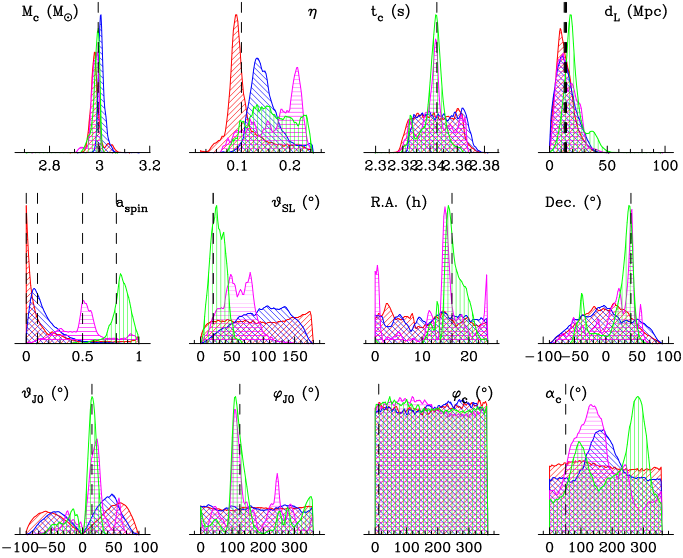
Run without signal



Parameters:

- H1 only
- Gaussian noise was used
- MCMC run was started as usual, but no signal was injected

Changing spin: 1 detector

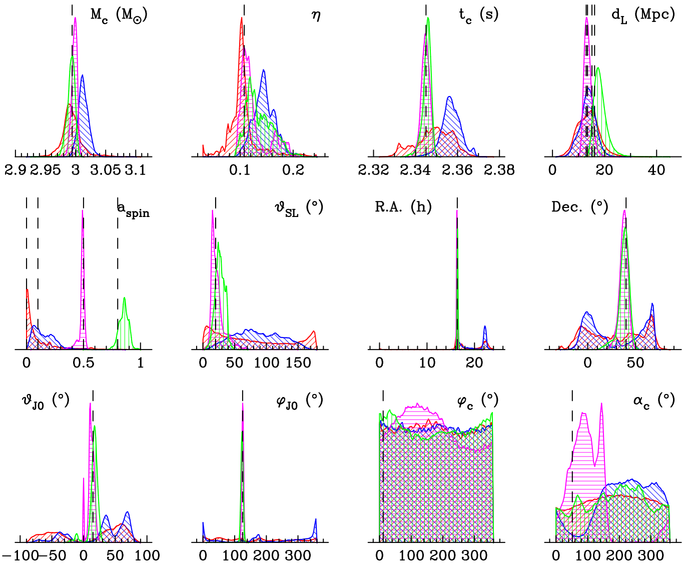


Parameters:

- H1 only
- $M = 10, 1.4\, M_\odot$
- $d_L \approx 16 - 21$ Mpc
- $a_{\rm spin} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\rm SL} = 20°$
- SNR $\approx$ 12.7
- Dashed lines show true values

NORTHWESTERN UNIVERSITY

Changing spin: 2 detectors

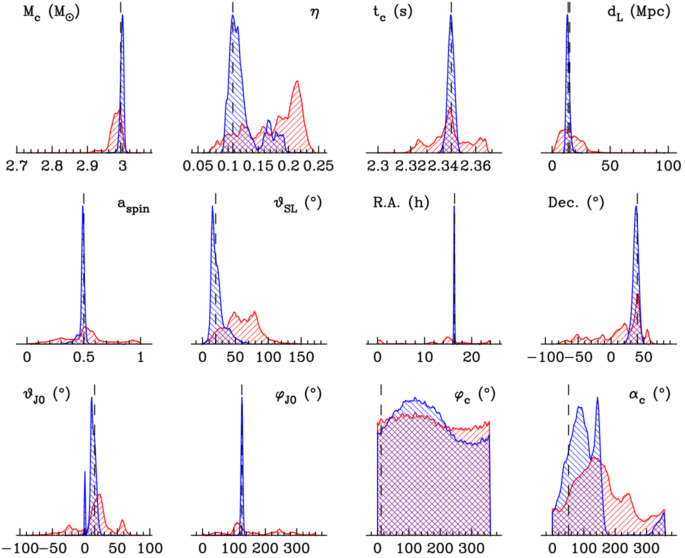


Parameters:

- H1 & L1
- $M = 10, 1.4\, M_\odot$
- $d_L \approx 16 - 21$ Mpc
- $a_{\rm spin} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\rm SL} = 20°$
- Network SNR $\approx$ 17.0
- Dashed lines show true values

NORTHWESTERN UNIVERSITY

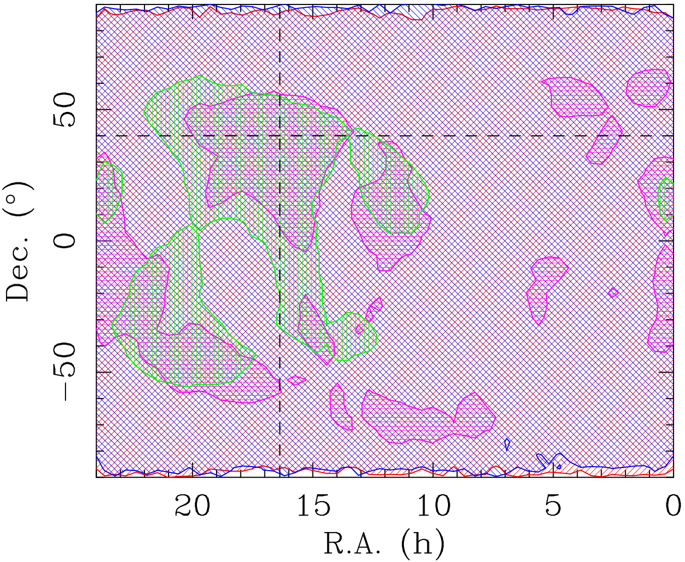
Changing the number of detectors



Parameters:

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

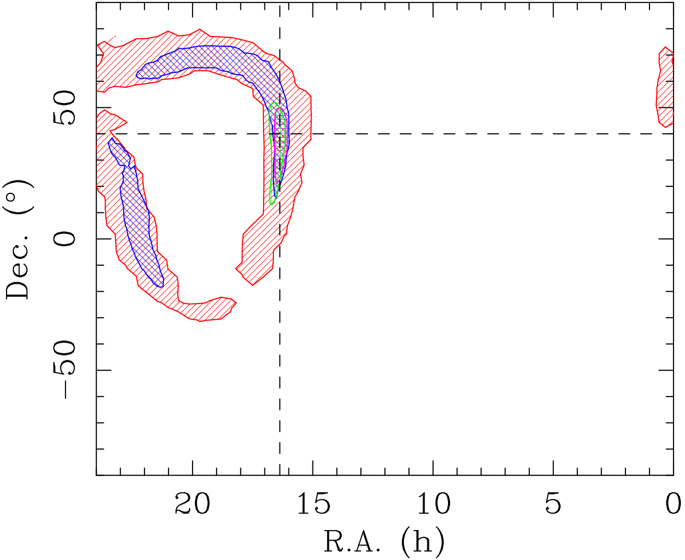
Sky map: 1 detector



Parameters:

- H1 only
- $M = 10, 1.4\,M_\odot$
- $d_L \approx 16 - 21$ Mpc
- $a_{\rm spin} = 0.0, 0.1, 0.5, 0.8$
- $\theta_{\rm SL} = 20°$
- SNR ≈ 12.7
- Dashed lines show true position

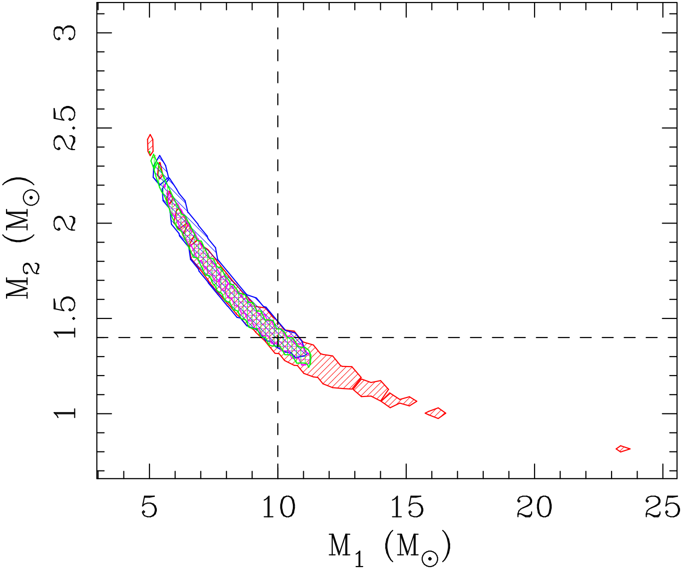
Sky map: 2 detectors



Parameters:

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

2D PDF: masses



Parameters:

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

Results

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

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

$*$: 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 $\mathcal{M}$ and η, 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

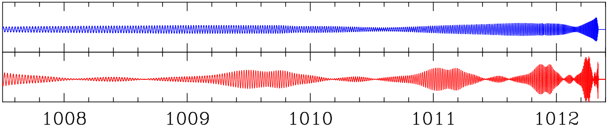
Offset start

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

Efficiency

- True modes will *eventually* be found by the chains
- Keyword: efficiency of sampling: how to we find the modes within *e.g.* a Hubble time?
- This becomes a more important issue for higher spin

Correlations increase with spin



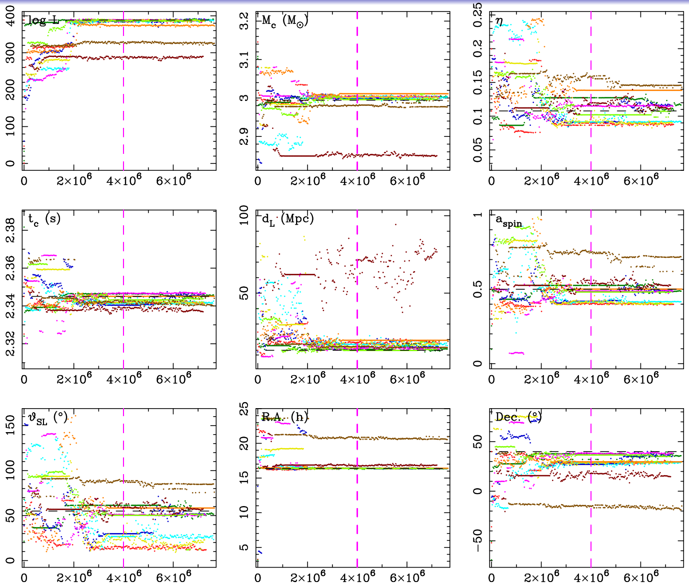
	M_c	η	a_{spin}	ϑ_{SL}	R.A.	Dec.
M_c		0.22	0.42	0.17	−0.40	0.19
η	−0.27		−0.34	−0.53	−0.07	−0.04
a_{spin}	−0.61	0.89		−0.04	0.11	0.62
ϑ_{SL}	0.66	−0.87	−0.99		0.02	−0.34
R.A.	−0.36	0.01	0.02	−0.02		0.12
Dec.	−0.23	0.08	0.18	−0.20	−0.05	

Parameters:

- H1 & L1
- $M_1 = 10\,M_\odot$
- $M_2 = 1.4\,M_\odot$
- $d_L = 13\,\mathrm{Mpc}$
- $a_{spin} = 0.1, 0.8$
- $\theta_{SL} = 55°$
- Network SNR $\approx 18.2, 30.5$

NORTHWESTERN UNIVERSITY

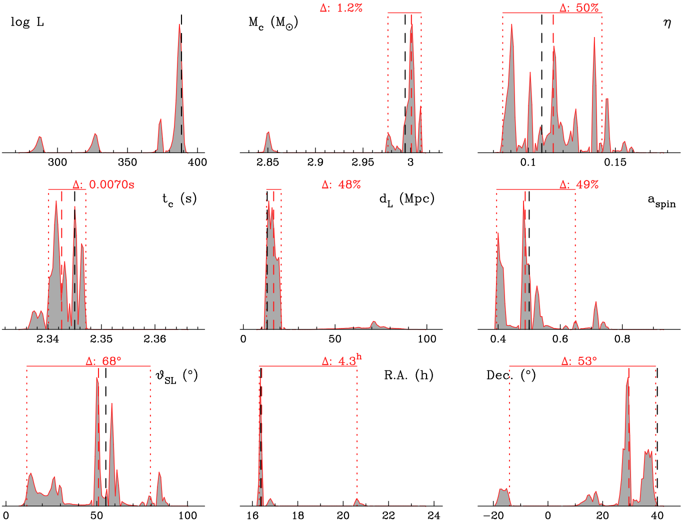
Structured parameter space



Parameters:

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

Structured parameter space



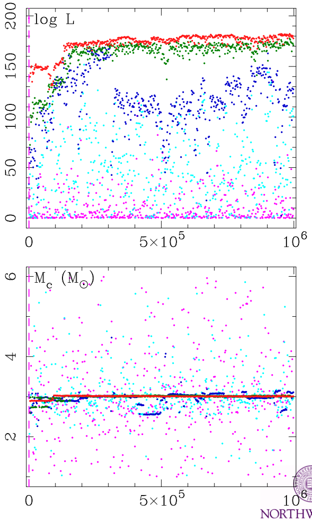
Parameters:

- H1 & L1
- $a_{\text{spin}} = 0.5$
- $\theta_{\text{SL}} = 20°$
- Network SNR ≈ 27.2
- 10 chains
- Offset start
- Black dashed lines are true values

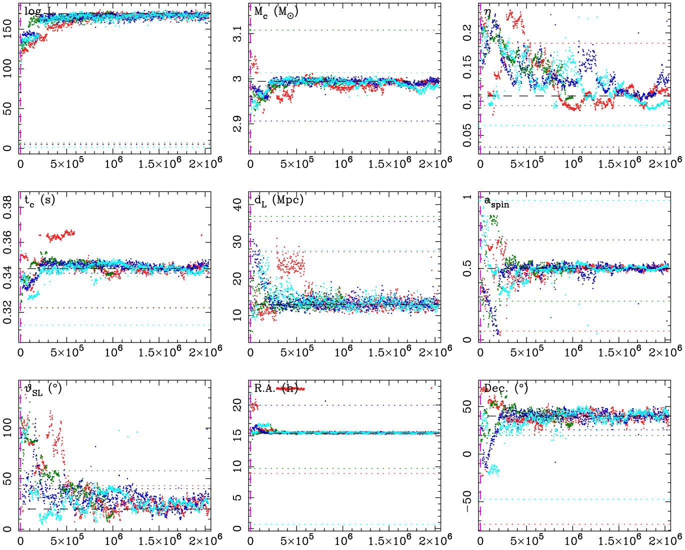
Parallel tempering

Parallel chains

- Use ∼5-10 parallel chains of temperatures $T = 1, \ldots, T_{\max}$
- 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: $\left(\frac{L_h}{L_c}\right)^{\frac{1}{T_c}-\frac{1}{T_h}}, \quad T_h > T_c$
- Hotter chains pass information to cooler chains



Converging chains



Parameters:

- H1 & L1
- $a_{\rm spin} = 0.5$
- $\theta_{\rm SL} = 20^\circ$
- Network SNR ≈ 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