Inference on inspiral signals using LISA MLDC data

The Global LISA Inference Group (GLIG):

Christian Röver[1], Ed Bloomer[2], James Clark[2], Nelson Christensen[3], Martin Hendry[2], Chris Messenger[2], Renate Meyer[1], Matt Pitkin[2], Alexander Stroeer[4,5], Jennifer Toher[2], Richard Umstätter[1], John Veitch[2], Alberto Vecchio[4,5], Graham Woan[2]

[1]Department of Statistics, The University of Auckland, Auckland, New Zealand
[2]Department of Physics and Astronomy, University of Glasgow, Glasgow, UK
[3]Physics and Astronomy, Carleton College, Northfield, MN, USA
[4]School of Physics and Astronomy, University of Birmingham, Birmingham, UK
[5]Department of Physics and Astronomy, Northwestern University, Evanston, IL, USA

Overview:

1. Long-term context

2. Inference framework

3. The `TDI generator`

4. Preliminary results

5. Summary, outlook, to-do

Long-term approach

- LISA data analysis within Bayesian framework
- Markov Chain Monte Carlo (MCMC) methods for inference
- detection pipeline

what we can build upon

- proof of principle on 'toy problems':
  - unknown number of sinusoids[1]
  - unknown number of heterogeneous signals (WD, MBH, EMRI)[2,3]

---

[1] R. Umstätter et al.: *Bayesian modeling of source confusion in LISA data*. Phys. Rev. D 72, (2005).

[2] E. Wickham, A. Stroeer, A. Vecchio: *A Markov Chain Monte Carlo approach to the study of massive black hole binary systems with LISA*. Class. Quantum Grav. 23, S819 (2006).

[3] A. Stroeer, J. Gair, A. Vecchio: *Automatic Bayesian inference for LISA data analysis strategies*. `gr-qc/0609010` (2006).

what we can build upon

- realistic applications for ground-based interferometry
  - pulsars[4]
  - binary inspirals - single interferometer[5] and network[6]
- TDI observables within Bayesian approach[7]

---

[4] R.J. Dupuis, G. Woan: *Bayesian estimation of pulsar parameters from gravitational wave data*. Phys. Rev. D 72, 1022002 (2005).

[5] C. Röver, R. Meyer, N. Christensen: *Bayesian inference on compact binary inspiral gravitational radiation signals in interferometric data*. Class. Quantum Grav. 23, 4895 (2006).

[6] C. Röver, R. Meyer, N. Christensen: *Coherent Bayesian inference on compact binary inspirals using a network of interferometric gravitational wave detectors*. `gr-qc/0609131` (2006).

[7] J.D. Romano, G. Woan: *Principal component analysis for LISA*. Phys. Rev. D 73, 102001 (2006).

where we are now - MLDC1

- TDI generation
- waveform generation
- code framework
- basic MCMC running

Observing a binary inspiral through LISA

- **modeled**: waveform in radiation frame—
  $\times/+$ waveforms (functions of parameters θ), plus direction & polarisation

- **measured**: LISA time delay interferometry (TDI) response—
  (X/Y/Z or A/E/T variables)

- in between: **LISA** . . . or **LISA Simulator**

- (*contrast*: earth-bound interferometry
  $\rightarrow$ measurement $\sim$ linear combination)

Inference framework

- **Bayesian** framework:
  derive **posterior distribution** $p(\theta|y)$ of parameters θ given observations y

- use of **MCMC methods** for inference

- problem similar to previous work[8]
  (difference: LISA TDI measurements instead of ground-based)

- crucial: evaluation of **likelihood** function

---

[8] C. Röver, R. Meyer, N. Christensen: *Coherent Bayesian inference on compact binary inspirals using a network of interferometric gravitational wave detectors*. `gr-qc/0609131` (2006).

Bayes & MCMC

- **posterior distribution**: $p(\theta|y) \propto p(\theta)\,p(y|\theta)$
  (distribution of parameters θ *given* data y)

- $p(\theta)$: **prior distribution**
  (pre-experimental knowledge about parameters)

- $p(y|\theta)$: **likelihood**
  (distribution of data y *given* parameters θ)

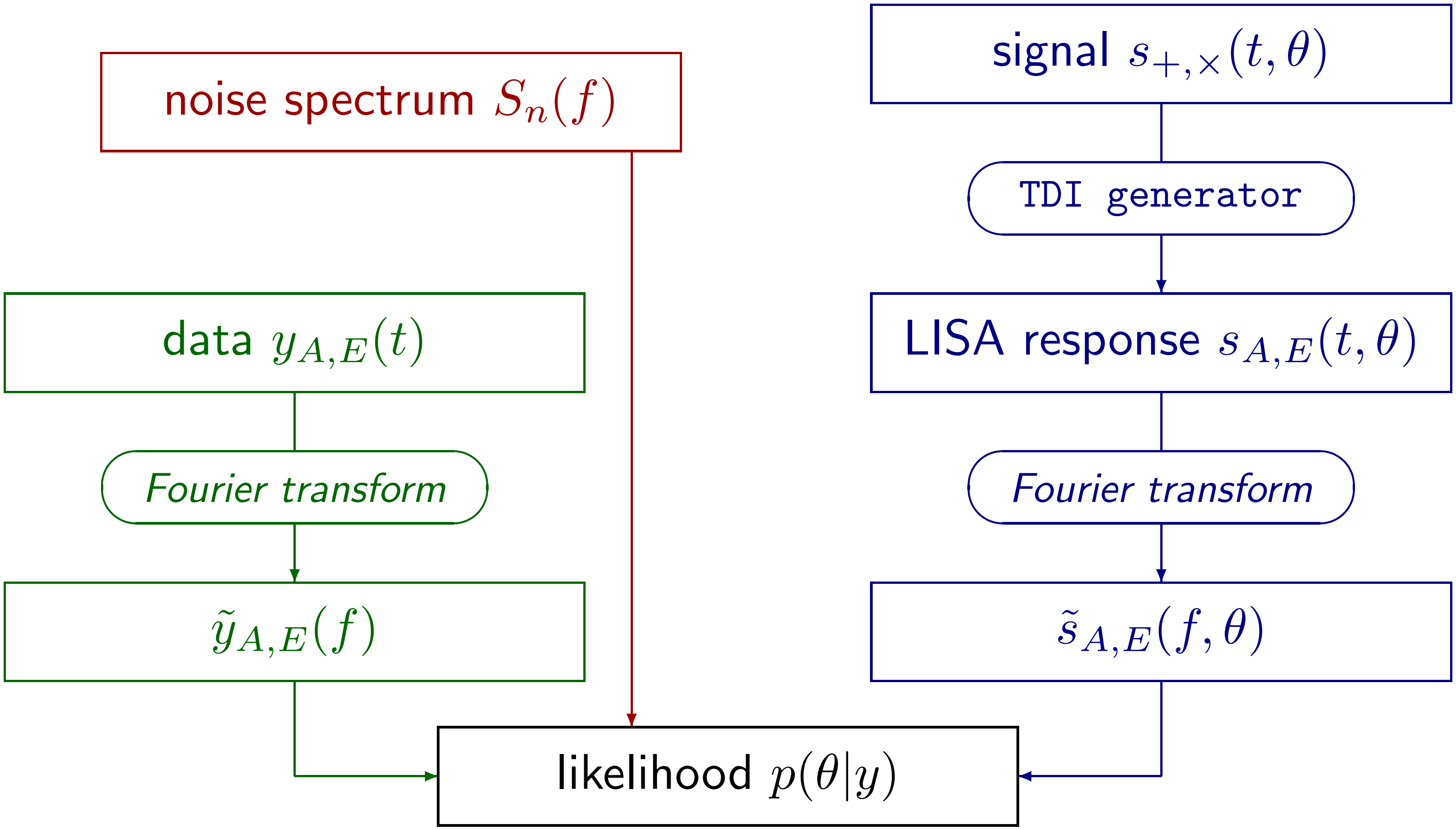
- use MCMC for
  - **optimisation** and
  - **integration**

Likelihood

- measurement $y(t)$: signal $s(t, \theta_0)$ plus noise
- signal: 2.0 post-Newtonian (PN) phase
- noise: stationary and Gaussian with spectral density $S_n(f)$
- likelihood:

$$p(y|\theta) \propto \exp\left(-\frac{2}{\delta_t} \sum_f \frac{|\tilde{y}(f) - \tilde{s}(f, \theta)|^2}{S_n(f)}\right)$$

where $\tilde{y}(f)$ and $\tilde{s}(f, \theta)$ are **Fourier-transformed** data & signal


noise spectrum $S_n(f)$
signal $s_{+,\times}(t,\theta)$
TDI generator
data $y_{A,E}(t)$
LISA response $s_{A,E}(t,\theta)$
Fourier transform
Fourier transform
$\tilde{y}_{A,E}(f)$
$\tilde{s}_{A,E}(f,\theta)$
likelihood $p(\theta|y)$

Likelihood implementation (general case)

- **input**: (arbitrary) waveform $s(t, \theta)$
  ($+/\times$ waveform, location, polarisation)

- **conversion**: `TDI generator`
  (X/Y/Z or A/E/T TDI response[9])

- . . . matching with data . . . (time- or frequency-domain)

- **output**: likelihood

[9] T.A. Prince et al.: *LISA optimal sensitivity*. Phys. Rev. D 66, 122002 (2002).

the `TDI generator`

- based on **LISA Simulator**[10],
  disassembled and wrapped into few 'black-box' `C` functions

- also:
  - interface to `lisaXML` files (data, metadata)
  - use of `FFTW`[11] library for Fourier transforms (+ noise spectrum)

- idea: supply **set of generic building blocks** for inference

[10] N.J. Cornish, L.J. Rubbo, O. Poujade: *The LISA Simulator*, `www.physics.montana.edu/LISA/`.
N.J. Cornish, L.J. Rubbo: *LISA response function*. Phys. Rev. D 67, 022001 (2003).

[11] M. Frigo, S.G. Johnson: *FFTW, a C subroutine library for computing the Discrete Fourier Transform (DFT)*, `http://www.fftw.org`.

applying the `TDI generator`

- problem: **speed**
  figures for *simplified* Challenge-1 model (constant arm lengths):

amount of data		seconds
days	# samples	per iteration*
364	2^{21}	146
182	2^{20}	75
91	2^{19}	38
46	2^{18}	19
23	2^{17}	10

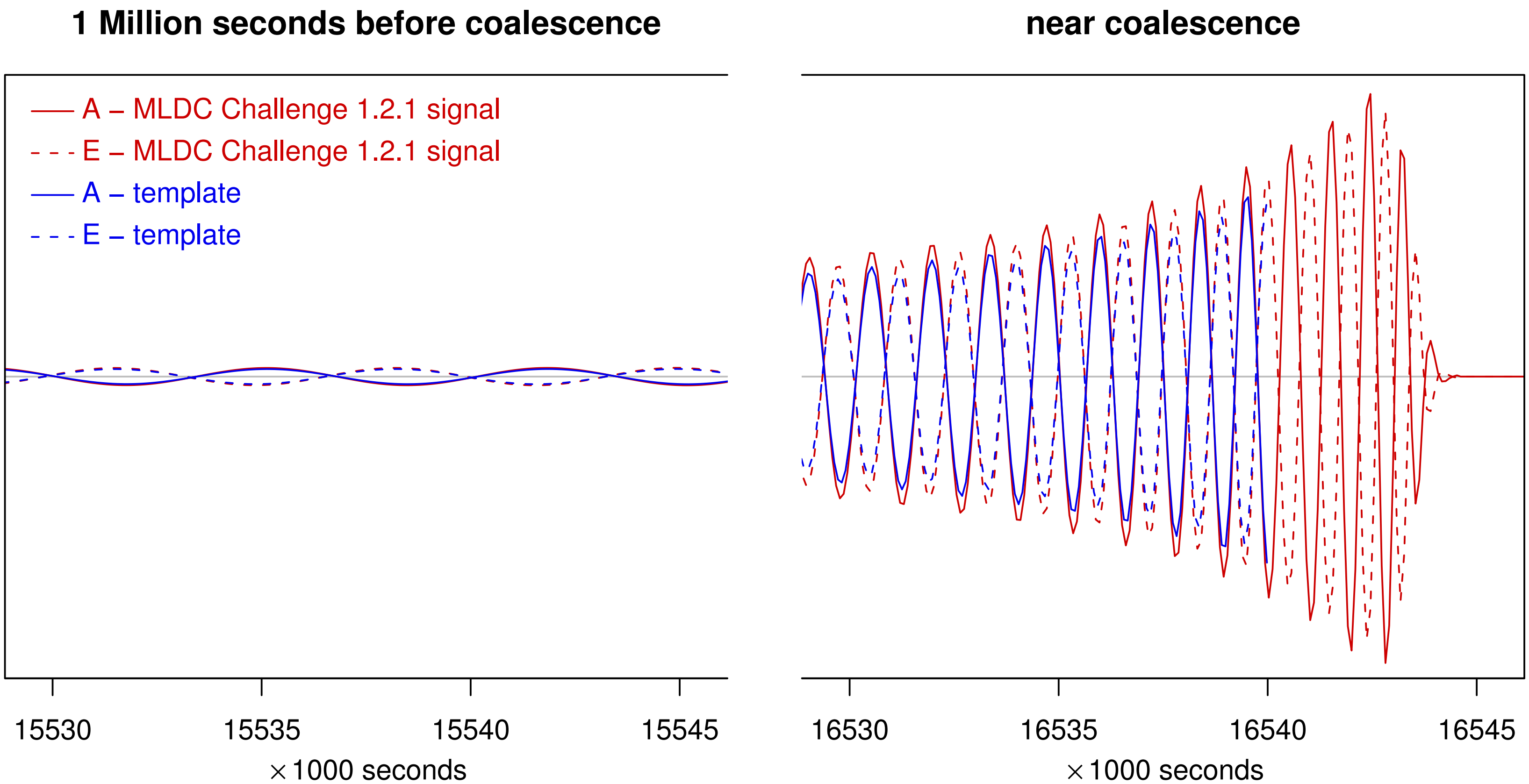
* *Intel Xeon 2.4 GHz, 4 GB memory*

$\approx$ 95% `TDI generator`,
$\approx$ 5% signal template, Fourier transform etc.

example application

- Challenge 1.2.1 (MBH binaries) training data set
- signal: 740 000 / 1 950 000 $M_\odot$ inspiral at 27 000 Mpc distance (2.0PN approximation)

matching noise-free templates (A/E TDI variables)



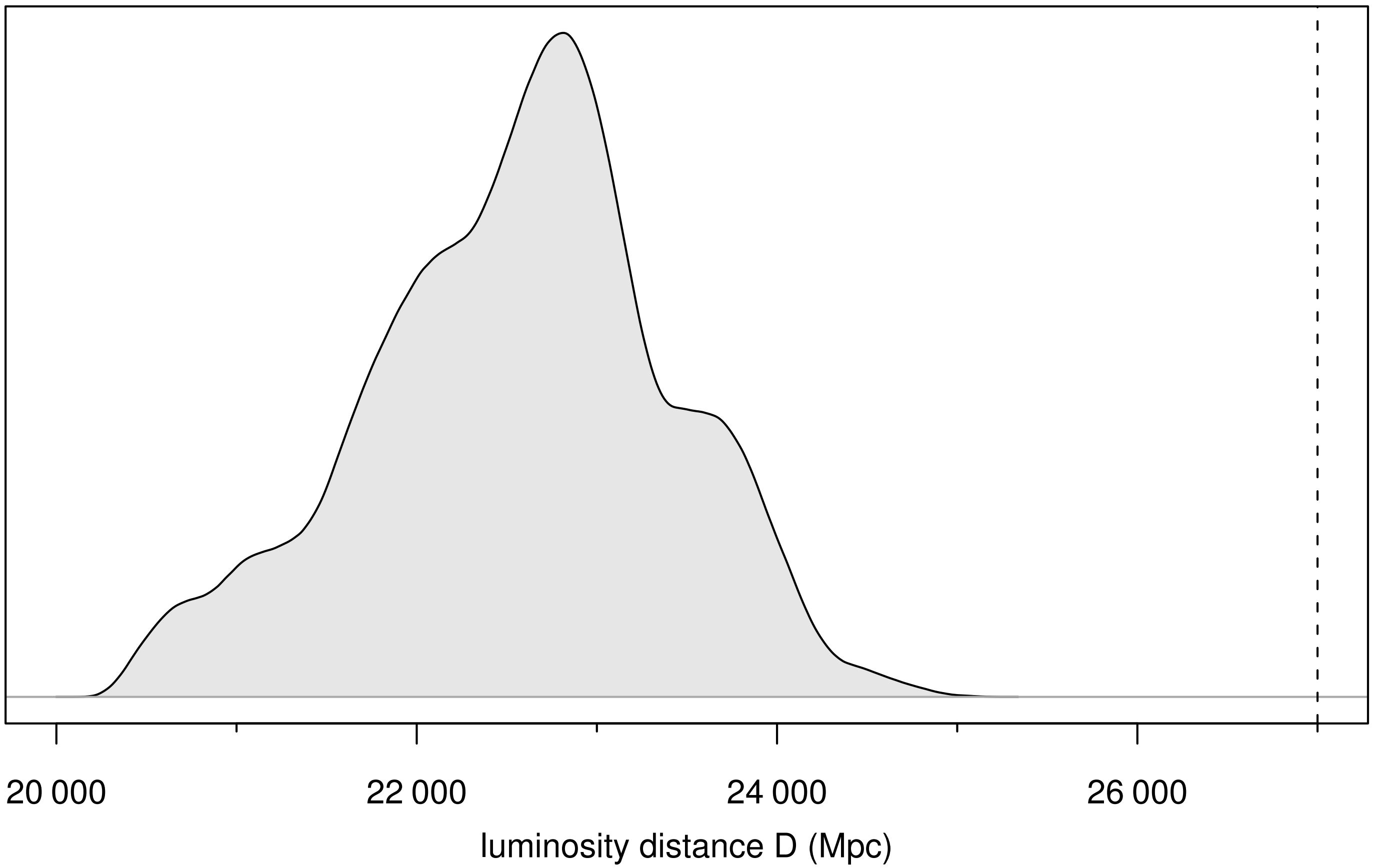
example application

- 'basic' MCMC (Metropolis algorithm)
- MCMC running on 2^{17} samples = 23 days of data
  (10 seconds/iteration, 360 iterations/hour)

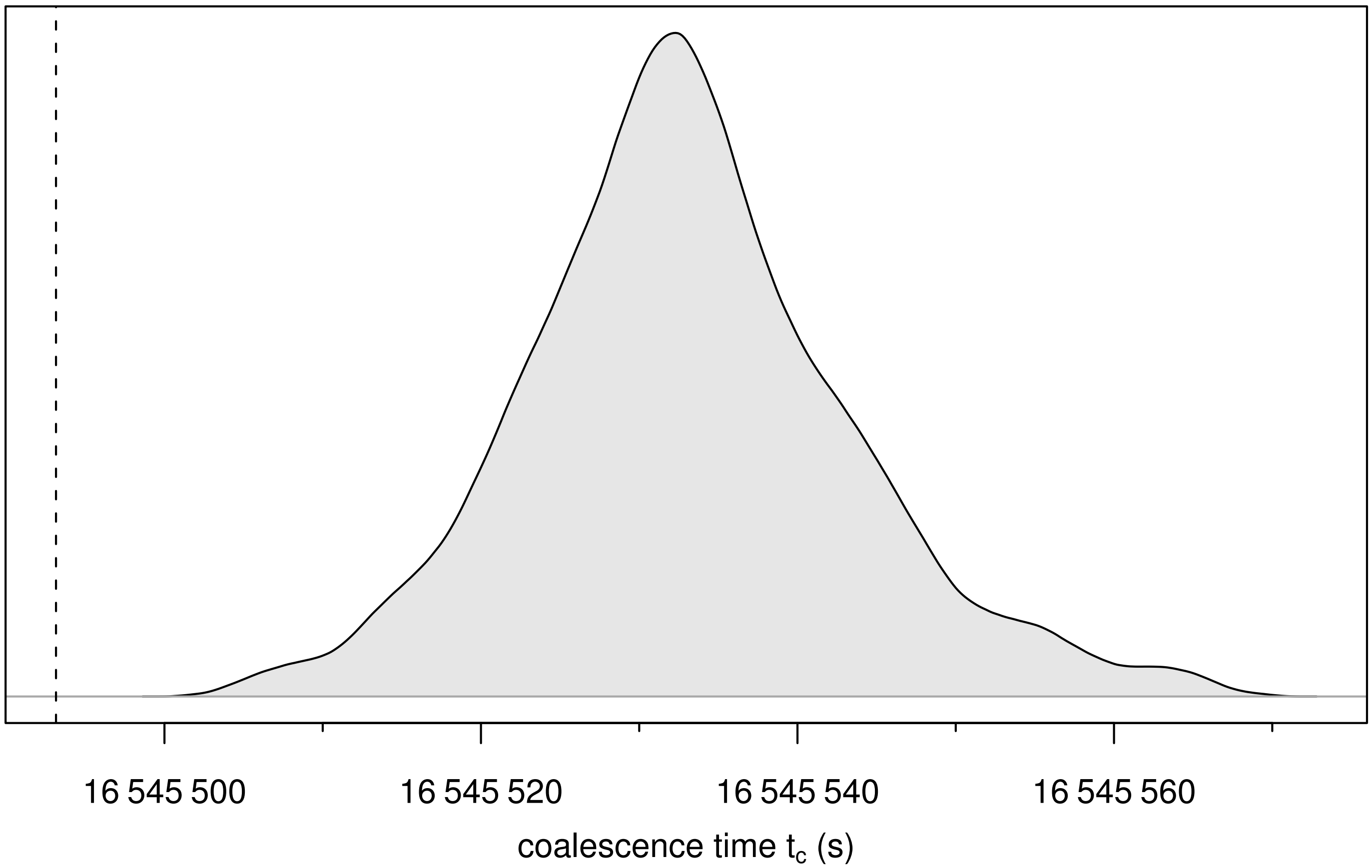
  + speed
  – resolution, esp. location parameters

- started at true parameter values
  (blind search did not converge)
- $\longrightarrow$ other modes?

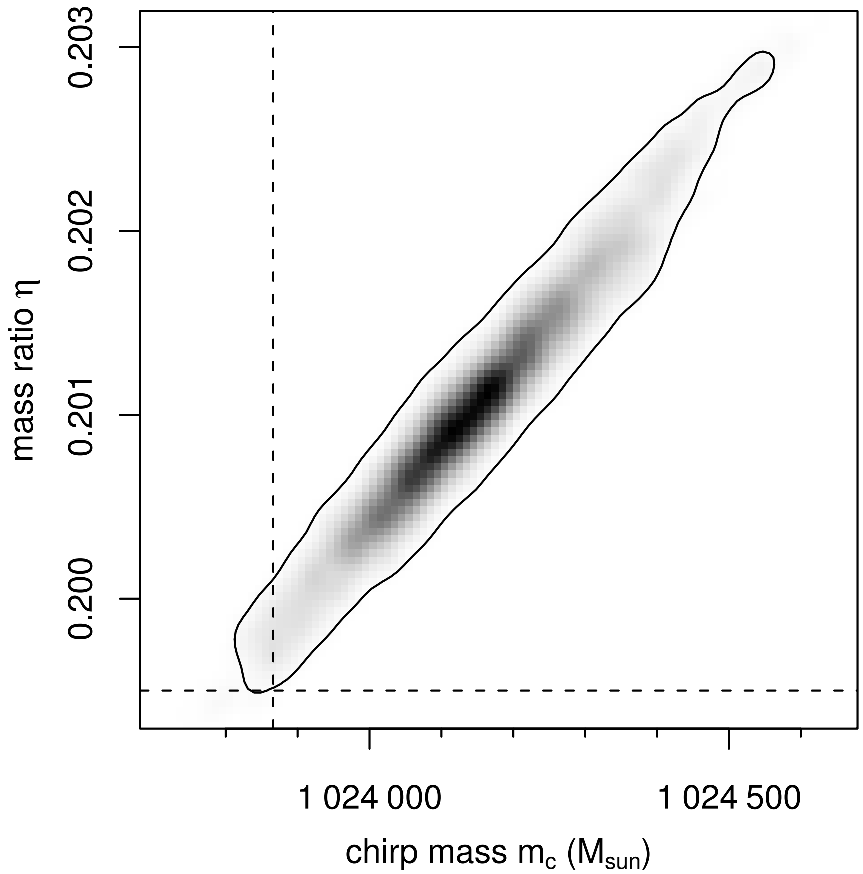
marginal posterior



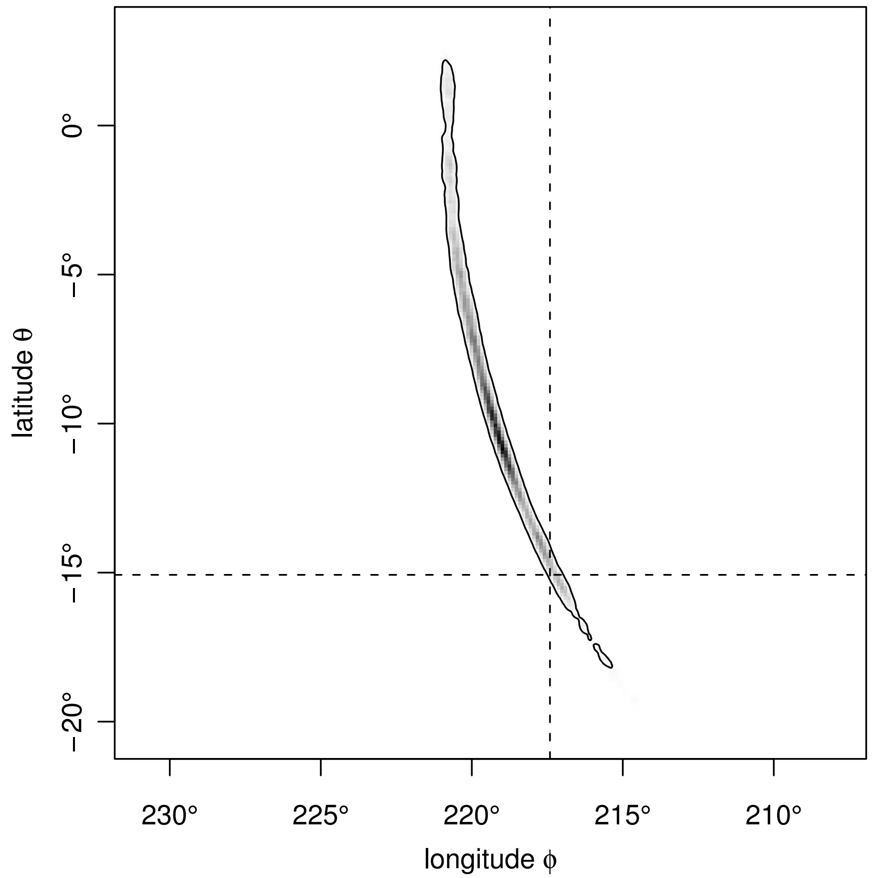
marginal posterior



marginal posterior



marginal posterior



correlations of parameters

	m_c	η	ϕ_0	t_c	θ	φ	D	ψ	ι
chirp mass (m_c)	1.000	**0.974**	**0.951**	**0.918**	0.368	0.312	-0.357	-0.355	-0.161
mass ratio (η)	**0.974**	1.000	**0.992**	**0.976**	0.301	0.245	-0.292	-0.304	-0.127
coalescence phase (ϕ_0)	**0.951**	**0.992**	1.000	**0.992**	0.345	0.285	-0.336	-0.326	-0.129
coalescence time (t_c)	**0.918**	**0.976**	**0.992**	1.000	0.322	0.266	-0.316	-0.324	-0.127
latitude (θ)	0.368	0.301	0.345	0.322	1.000	**0.945**	**-0.995**	-0.845	-0.340
longitude (φ)	0.312	0.245	0.285	0.266	**0.945**	1.000	**-0.955**	-0.699	-0.247
luminosity distance (D)	-0.357	-0.292	-0.336	-0.316	**-0.995**	**-0.955**	1.000	0.836	0.272
polarisation (ψ)	-0.355	-0.304	-0.326	-0.324	-0.845	-0.699	0.836	1.000	0.381
inclination (ι)	-0.161	-0.127	-0.129	-0.127	-0.340	-0.247	0.272	0.381	1.000

- (some) correlations should decrease when more data are considered

MLDC application

- template discrepancy (esp.: coalescence time t_c and luminosity distance D) turns out 'significant'
- note: *significance* $\not\approx$ *relevance*
- also: SNR$=$500
- other parameters still mostly match

- MCMC basically running (. . . slowly. . .)

what's next

- tidy, speed & wrap up `TDI generator`
- match templates
- implement **Parallel Tempering** MCMC for efficient sampling (convergence and mixing)
- use of parallel computing
- treat noise (-spectrum) as unknown