The dangers of systematic errors in parameter estimation
What waveforms do data analysts want?

Ilya Mandel
University of Birmingham
GWPAW, Hannover, June 4, 2012

Why do we care about parameter estimation?

- If we want to do physics with GW detections -- astrophysics or studies of strong-field gravity -- we must know how to evaluate model fits and estimate parameters

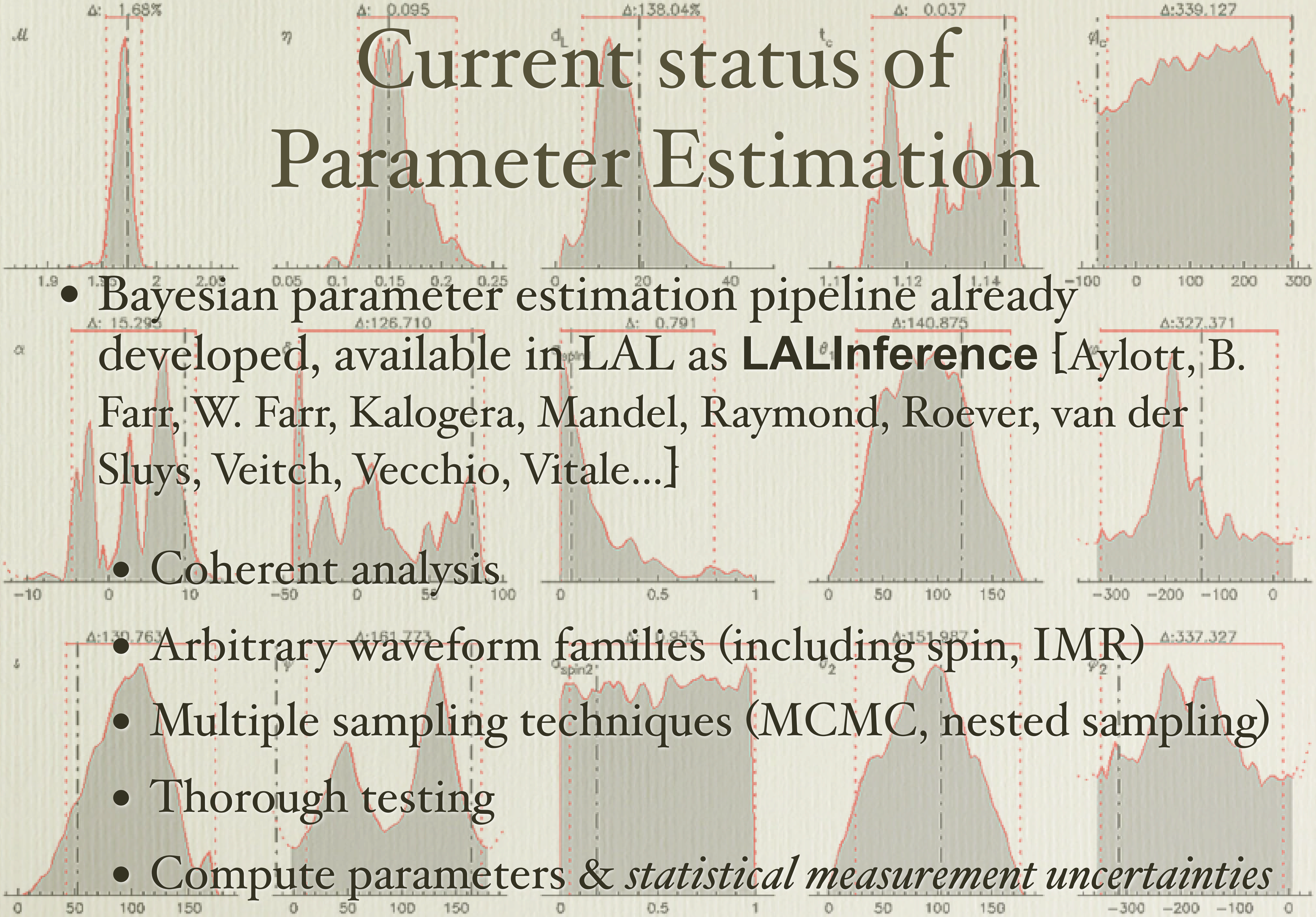
- See a number of PE posters at this meeting:
  - T. Sidery, Sky Localization
  - S. Vitale and R. Sturani, Spins
  - W. Vousden, Astrophysical Priors
  - R. Smith, IMRI waveforms
  - and talks by H. Tagoshi, T. Li, H. Pfeiffer, M. Vallisneri ...

Controversial Claims

- Parameter estimation problem is fundamentally solved when the model is known
- Systematic biases, particularly from imperfectly known waveforms, are the biggest challenge going forward
- We have been using inadequate techniques for estimating impact of bias [*A modest proposal for how to do this better*]
- Are BNSs clean systems from a PE perspective? [*Not quite*]
- What are the most urgent challenges for a data analyst? [*Personal view*]

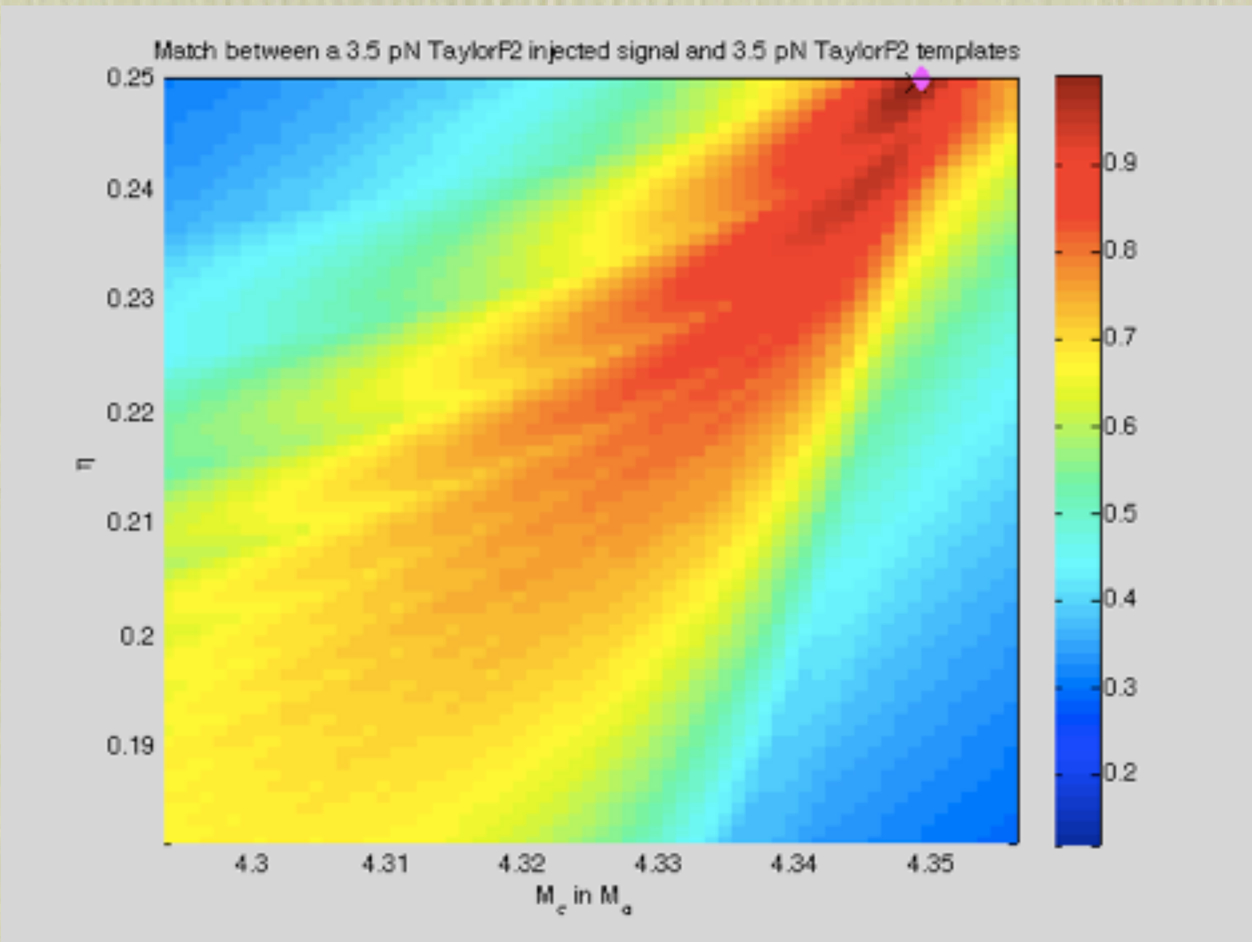
Current status of Parameter Estimation

- Bayesian parameter estimation pipeline already developed, available in LAL as **LALInference** [Aylott, B. Farr, W. Farr, Kalogera, Mandel, Raymond, Roever, van der Sluys, Veitch, Vecchio, Vitale...]
  - Coherent analysis
  - Arbitrary waveform families (including spin, IMR)
  - Multiple sampling techniques (MCMC, nested sampling)
  - Thorough testing
  - Compute parameters & *statistical measurement uncertainties*



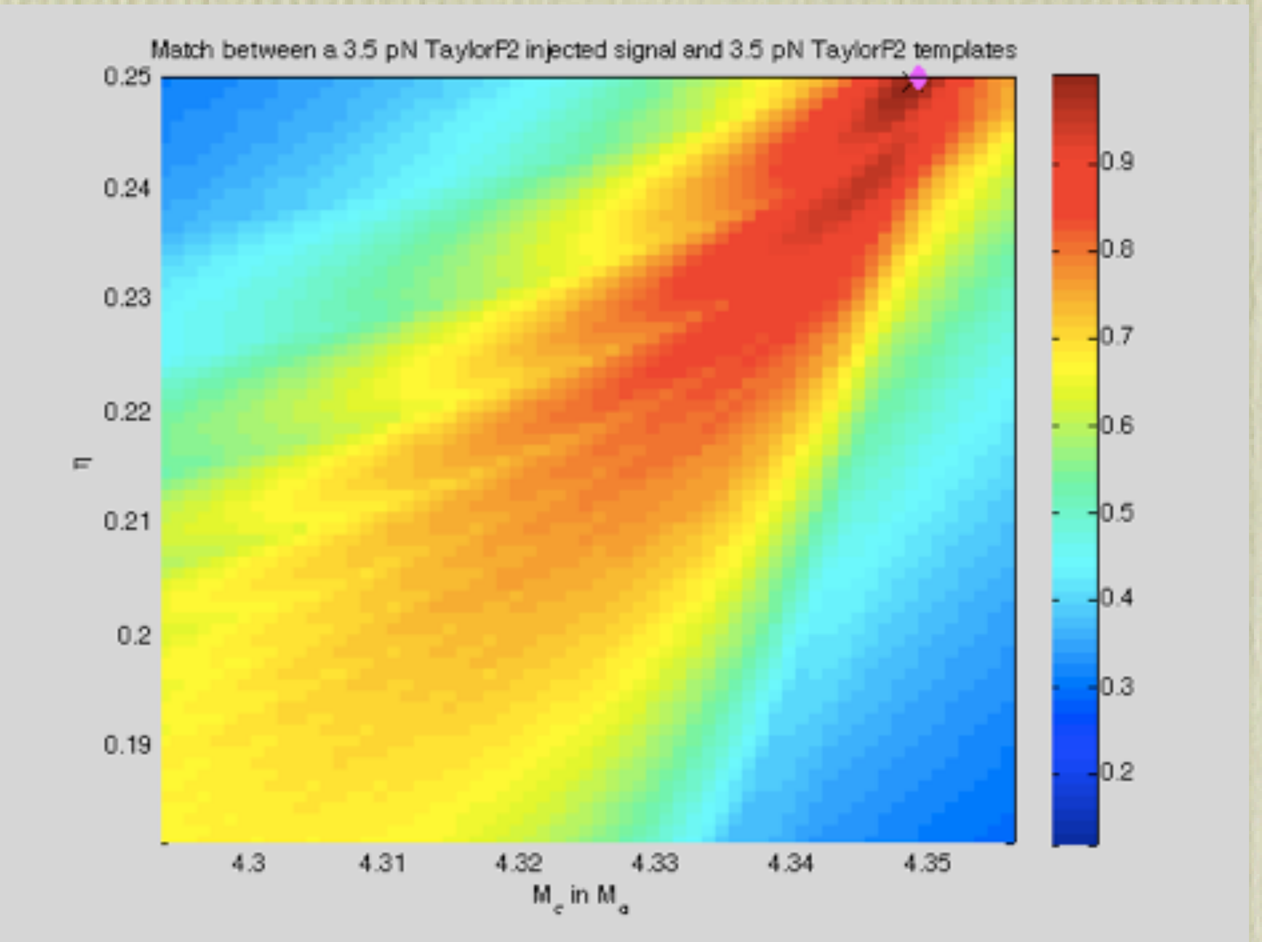
[van der Sluys, Mandel, Raymond, et al., 2009]

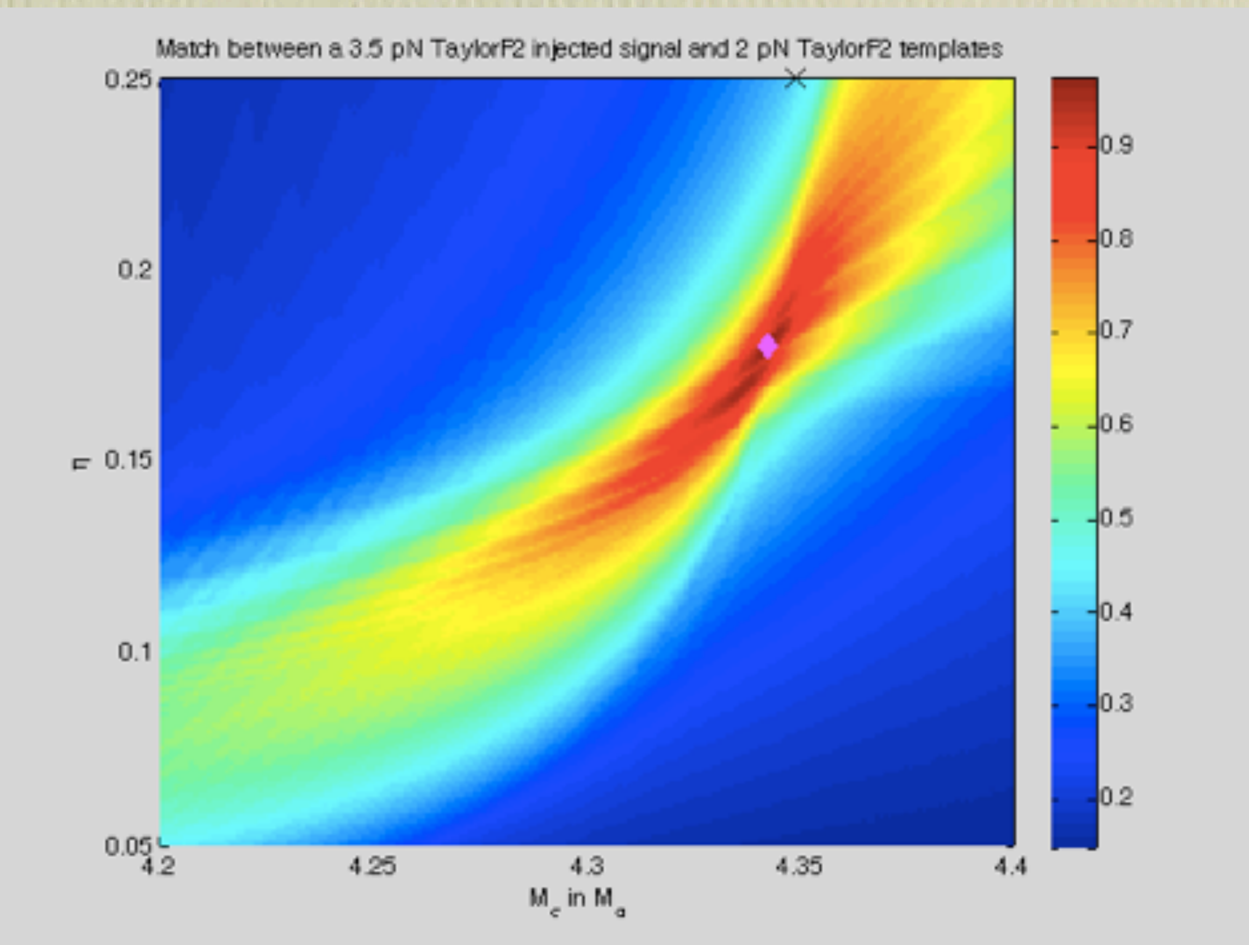
Statistical uncertainty



- 3.5pN injection and recovery
- 4.8+5.2 solar-mass BH system, overhead a single AdvLIGO detector
- Match of 0.97 contour corresponds to ~2-sigma confidence interval on masses (or does it?)

Systematic biases





- 3.5pN injection and recovery
- 4.8+5.2 solar-mass BH system, overhead a single AdvLIGO detector
- Match of 0.97 contour corresponds to ~2-sigma confidence interval on masses

- 3.5pN injection, 2pN recovery
- Best fit (magenta) is 9.2+2.8 system -- NS-BH???
- [Match with injection is above 0.97]

Waveform models

- Parameter estimation, not detection, is the threshold -- so "good to a few percent" is not enough.
- We need more confident waveform families that incorporate all of the relevant effects:
  - **Inspiral, merger, ringdown**
  - **Spinning, precessing waveforms**
  - Higher harmonics
  - Intermediate mass ratios
  - Matter effects
  - Eccentric binaries
  - Deviations from Kerr / GR?

How accurate do waveforms *really* need to be?

- Standard criterion: $\langle \delta h | \delta h \rangle \leq 1$ [Lindblom, Owen, Brown; 2008]

$$
\begin{aligned}
h(\lambda, f) &= (1-\lambda) h_e(f) + \lambda h_m(f), \\
&\equiv h_e(f) + \lambda \delta h(f),
\end{aligned}
$$

$$
\sigma_\lambda^{-2} = \left\langle \frac{\partial h}{\partial \lambda} \middle| \frac{\partial h}{\partial \lambda} \right\rangle = \langle \delta h | \delta h \rangle
$$

How accurate do waveforms *really* need to be?

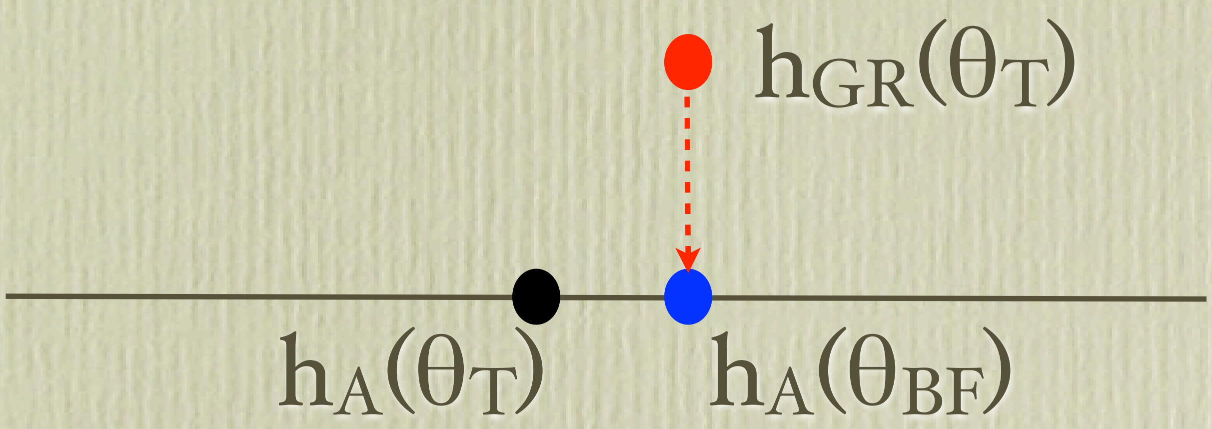
- Standard criterion: $\langle \delta h | \delta h \rangle \leq 1$ {Lindblom, Owen, Brown; 2008}

$$h(\lambda, f) = (1-\lambda) h_e(f) + \lambda h_m(f),$$
$$\equiv h_e(f) + \lambda \delta h(f),$$

$$\sigma_\lambda^{-2} = \left\langle \frac{\partial h}{\partial \lambda} \middle| \frac{\partial h}{\partial \lambda} \right\rangle = \langle \delta h | \delta h \rangle$$

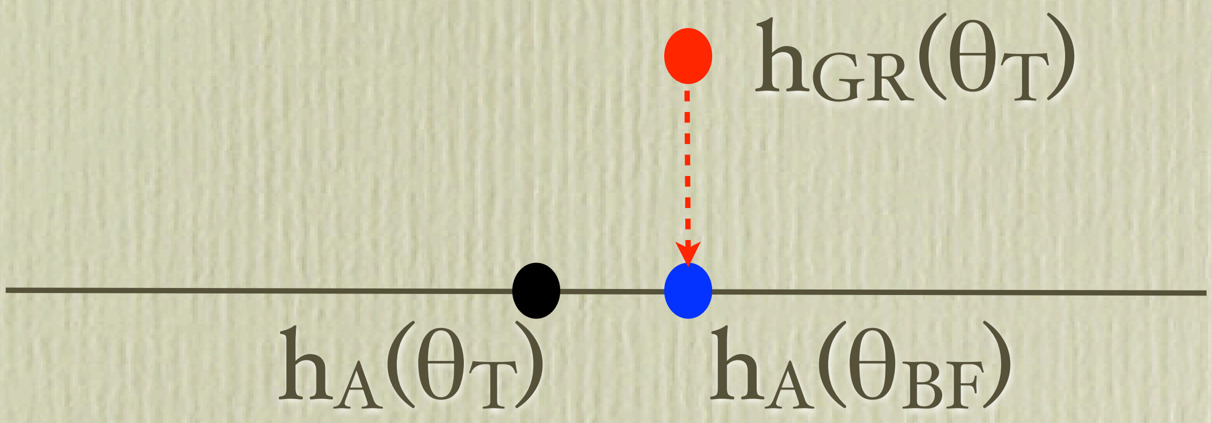
- Detectability of deviation between two alternative models with all other parameters fixed

How accurate do waveforms *really* need to be?



- Standard criterion: $\langle \delta h | \delta h \rangle \leq 1$ [Lindblom, Owen, Brown; 2008], $\delta h = h_{GR}(\theta_T) - h_A(\theta_T)$: Can we tell red and black apart?
- But the real question is, do we care about the difference between black and blue?
  - That difference can be much smaller because of projection; might also manifest in irrelevant parameters

How accurate do waveforms *really* need to be?



- Standard criterion: $\langle \delta h | \delta h \rangle \leq 1$ [Lindblom, Owen, Brown; 2008], $\delta h = h_{GR}(\theta_T) - h_A(\theta_T)$: Can we tell red and black apart?
- But the real question is, do we care about the difference between black and blue, for *parameters of interest*?
- Standard criterion sufficient, but not necessary -- can be far too demanding

How to estimate impact of systematic bias?

- Separately compute statistical uncertainty and systematic bias [e.g., Cutler & Vallisneri, 2007]
- Systematic bias: best-fit interesting parameters (*mass)* from grid-based search maximizing over uninteresting parameters (*time & phase)*
- Statistical uncertainty: ~~poor~~ computationally-limited man's version of Bayesian PE

Statistical uncertainty, I

- Likelihood: $L(h, \vec{\theta}) = p(s|\vec{\theta}) \propto \exp\left(-\frac{\langle s - h(\vec{\theta}) \mid s - h(\vec{\theta})\rangle}{2}\right)$

- Zero-noise realization:

$$L(\delta\vec{\theta}) \propto \exp\left(-\frac{\langle h(\vec{\theta_{\rm inj}}) - h(\vec{\theta_{\rm inj}} + \delta\vec{\theta}) \mid h(\vec{\theta_{\rm inj}}) - h(\vec{\theta_{\rm inj}} + \delta\vec{\theta})\rangle}{2}\right) \equiv \exp\left(-\frac{\langle \delta h \mid \delta h\rangle}{2}\right)$$

- Overlap / match:

$$M(s, h(\vec{\theta})) = \frac{\langle s \mid h(\vec{\theta})\rangle}{\sqrt{\langle s \mid s\rangle\langle h(\vec{\theta}) \mid h(\vec{\theta})\rangle}} \to \frac{\langle h \mid h + \delta h\rangle}{\sqrt{\langle h \mid h\rangle\langle h + \delta h \mid h + \delta h\rangle}}$$

where we maximize over “uninteresting” parameters

Statistical uncertainty, II

- If $\delta h << h$ and $\langle h|\delta h\rangle = 0$, then

$$\log L = -\frac{1}{2}\langle \delta h \mid \delta h\rangle \sim -(1-M)\langle h \mid h\rangle = -(1-M)\ \mathrm{SNR}^2$$

- Further, if $\langle \delta h \mid \delta h\rangle \approx \langle \partial h_{,\theta} \mid \partial h_{,\theta}\rangle (\delta\theta)^2$

$$\int_{-\Delta\theta}^{+\Delta\theta} L(\theta) d\theta = \int_{-\Delta\theta}^{+\Delta\theta} d\theta e^{-\frac{1}{2}\langle \delta h \mid \delta h\rangle} \sim erf\left(\frac{\Delta\theta}{\sqrt{2\langle \partial h_{,\theta} \mid \partial h_{,\theta}\rangle^{-1/2}}}\right)$$

- Under these assumptions, the boundary of the N-sigma confidence interval is given by $\sqrt{2(1-M)}\mathrm{SNR}$

- So, for example, for SNR=8, M=0.97 is 2-sigma boundary

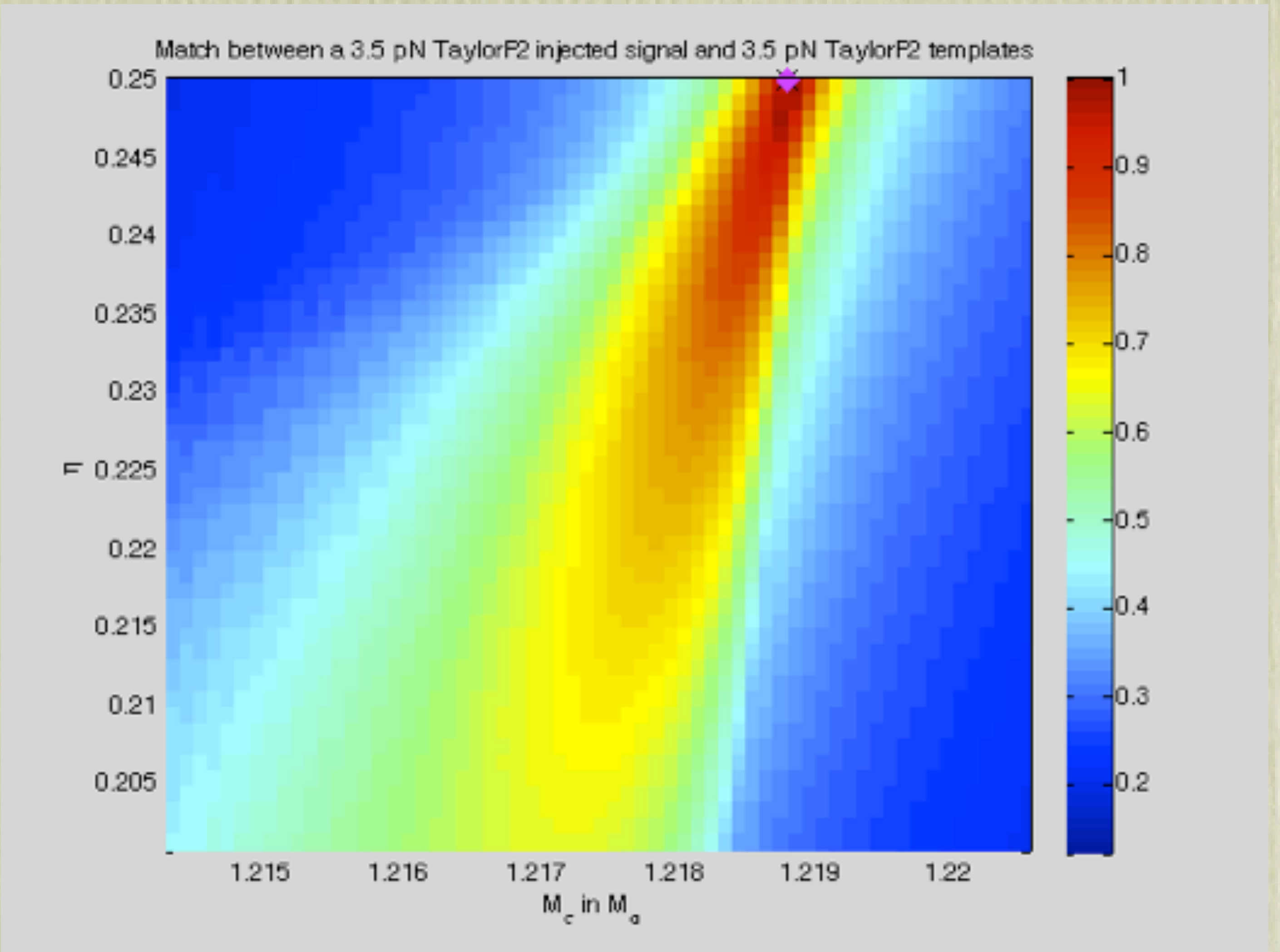
A Modest Proposal

- Can directly compute the desired confidence interval from a grid-based search:

$$\int_{-\Delta\theta}^{+\Delta\theta} L(\theta) d\theta = \int_{-\Delta\theta}^{+\Delta\theta} d\theta e^{-\frac{1}{2}\langle \delta h(\theta) \mid \delta h(\theta) \rangle}$$

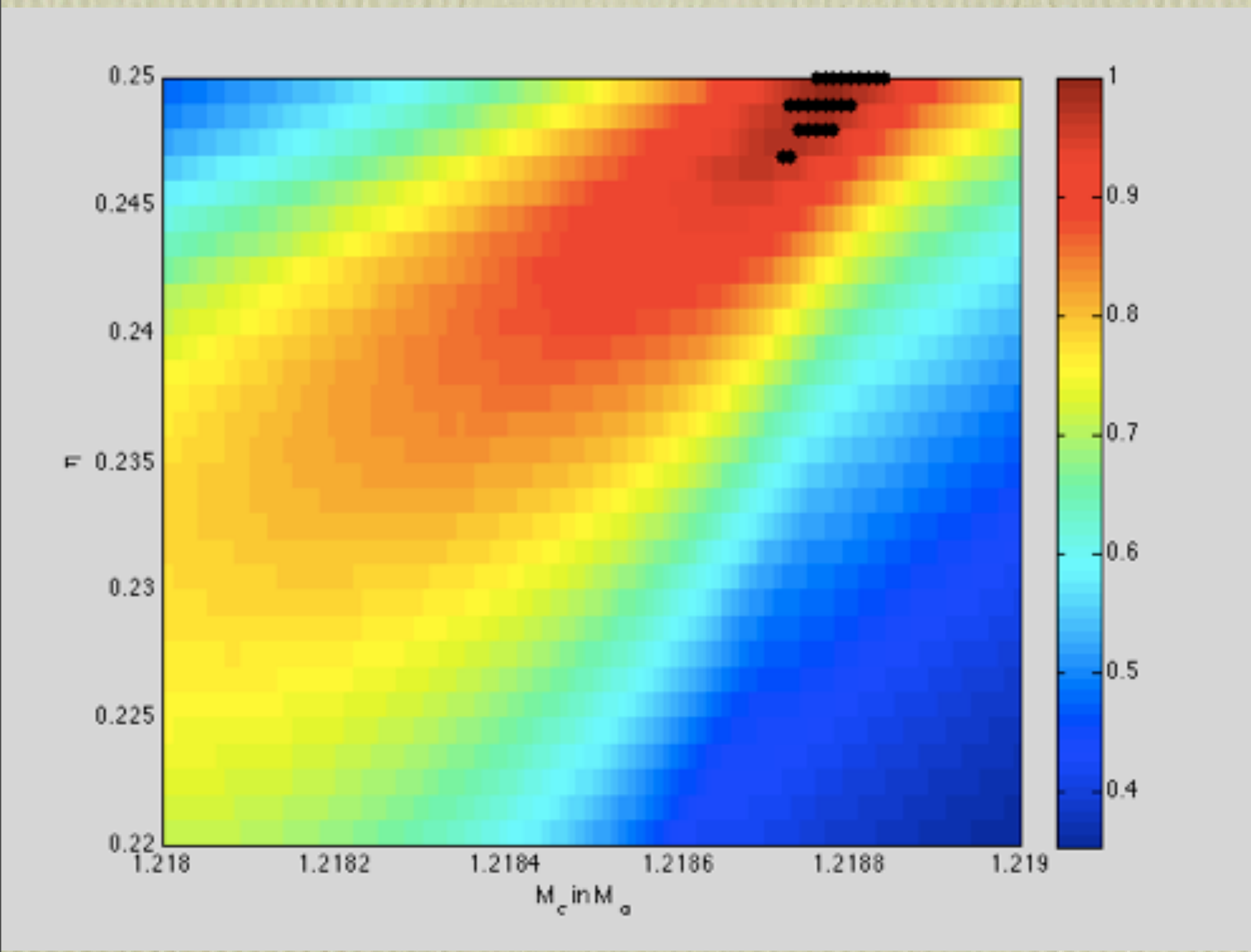
- Inner product maximized over *uninteresting* parameters
- Minimal over-head relative to standard overlap calculation
- Faster than Bayesian techniques in small dimensions [at the expense of not getting the priors right on parameters that are maximized over]

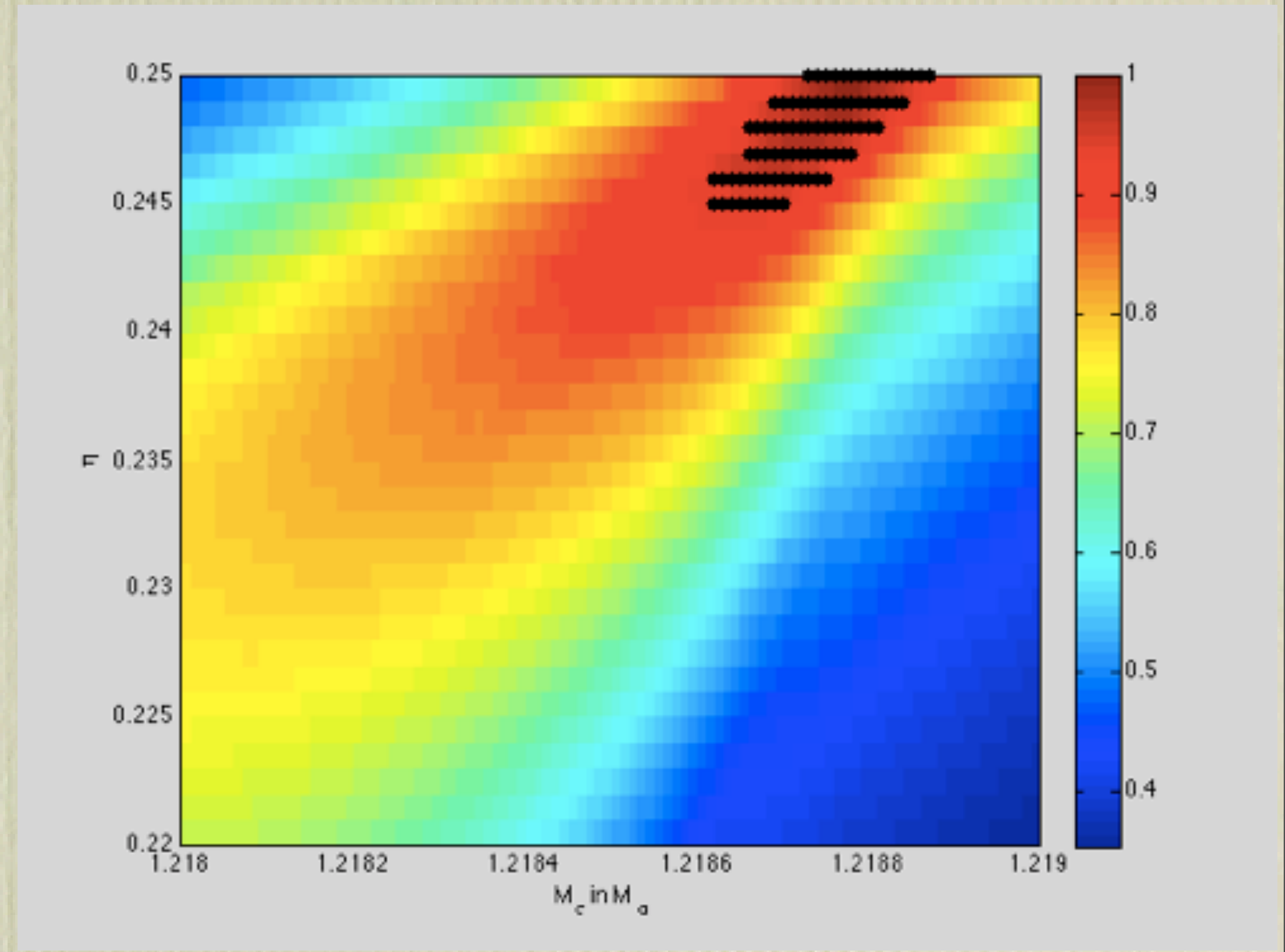
Are binary neutron stars clean systems?



- 3.5pN injection and recovery
- 1.4+1.4 solar-mass BH system, overhead a single AdvLIGO detector
- Match of 0.97 contour corresponds to ~2-sigma confidence interval on masses
- What about pN uncertainty? Matter effects? Spins?

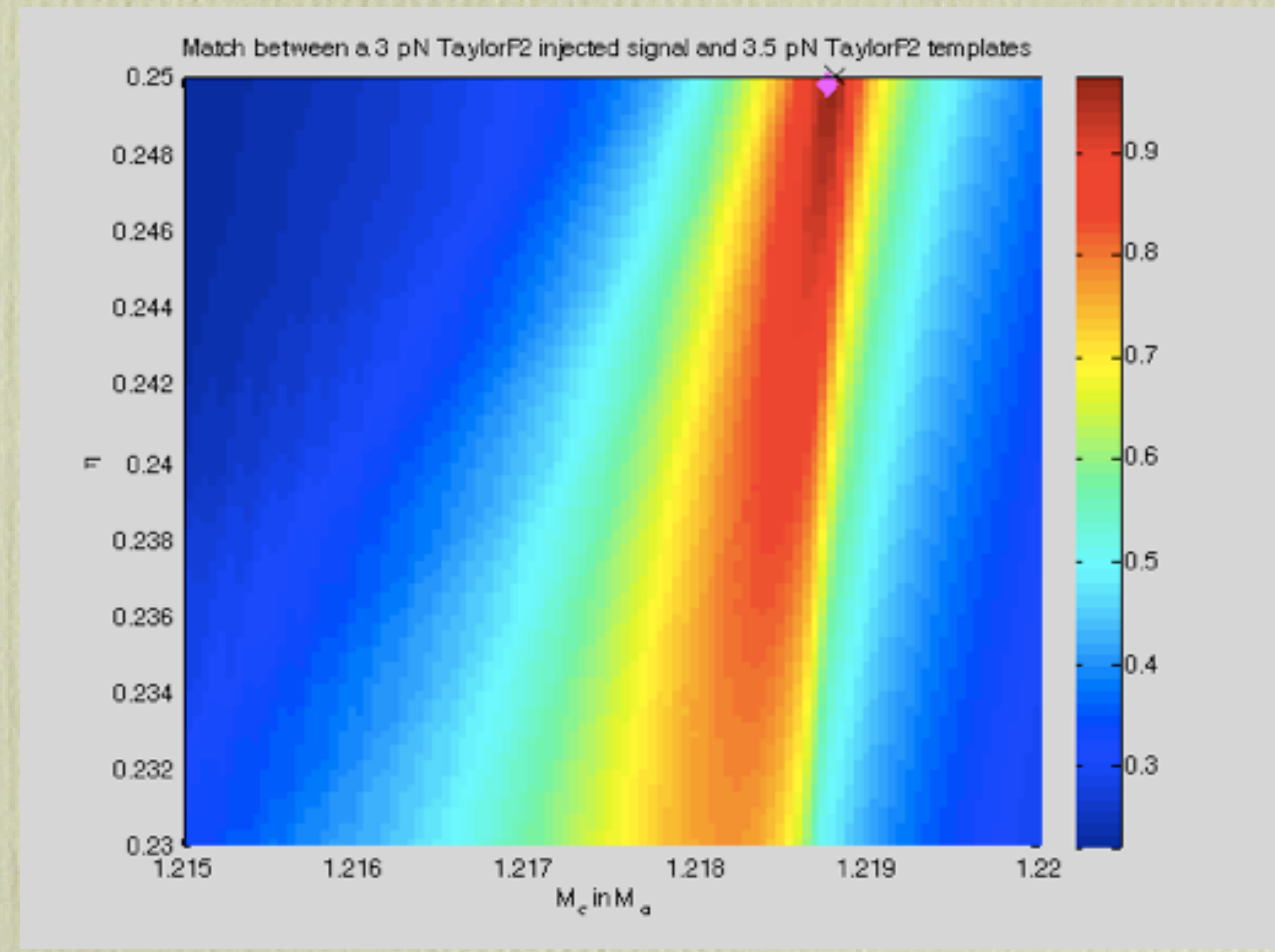
Are binary neutron stars clean systems?





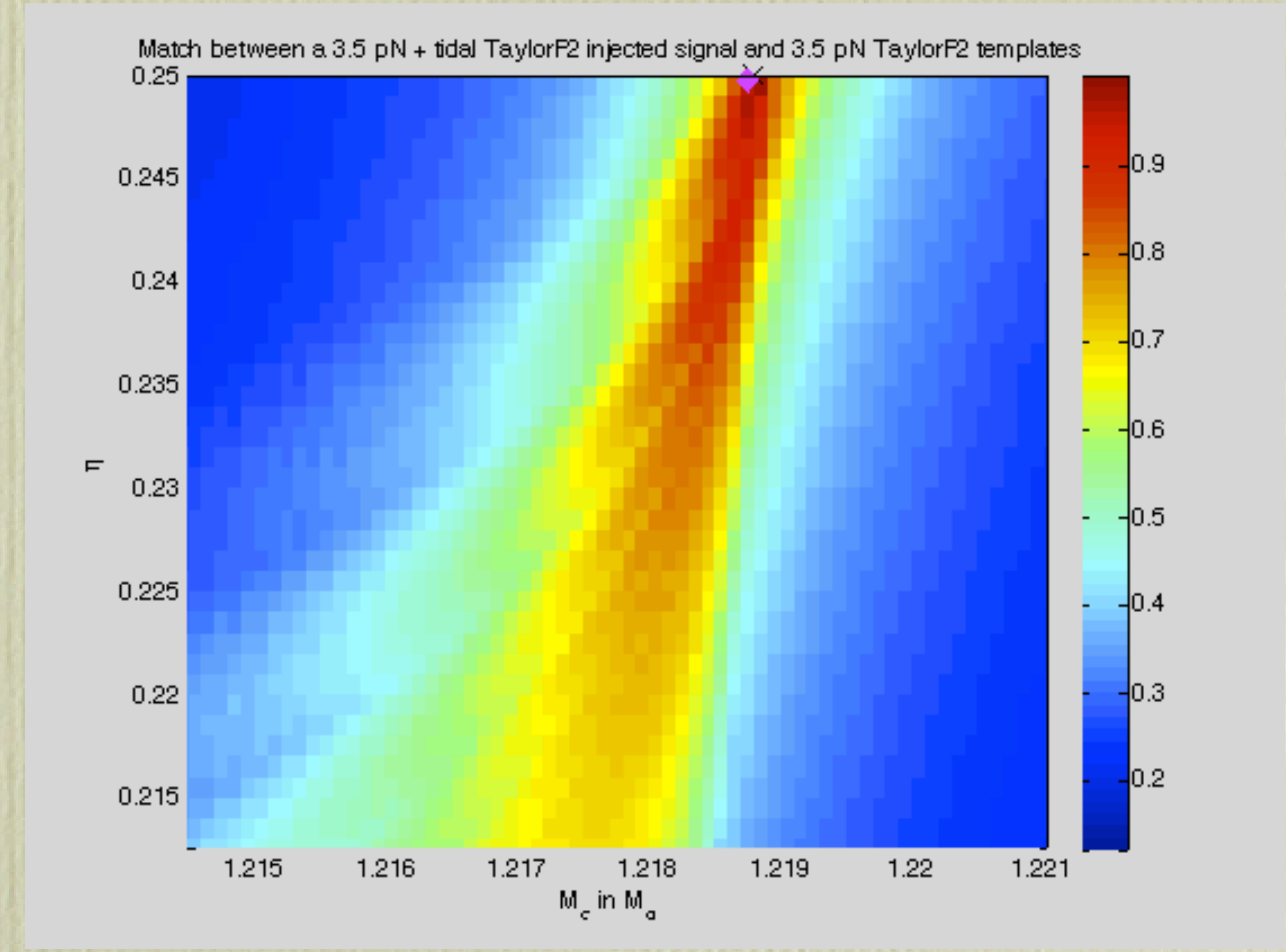
- 3.5pN injection and recovery, 1.4+1.4 solar-mass BH system, overhead a single AdvLIGO detector
- 95% confidence interval on masses includes matches ≳ 0.94 at SNR=8, $\langle \delta h | \delta h \rangle \lesssim 8$
- 68% confidence interval on masses includes matches ≳ 0.976 at SNR=8, $\langle \delta h | \delta h \rangle \lesssim 3$
- What about pN uncertainty? Matter effects? Spins?

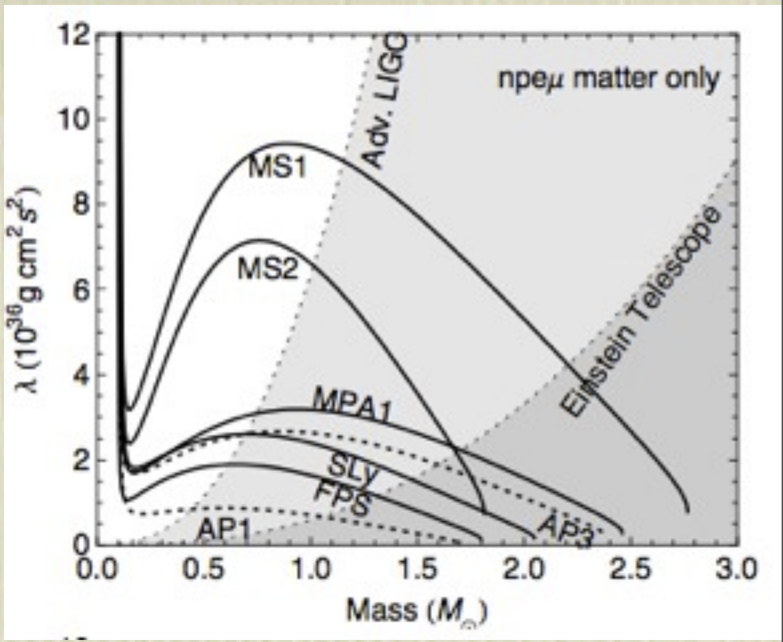
Effect of higher-order pN terms



- 3pN injection, 3.5 pN recovery
- Best-fit masses are 1.44+1.36

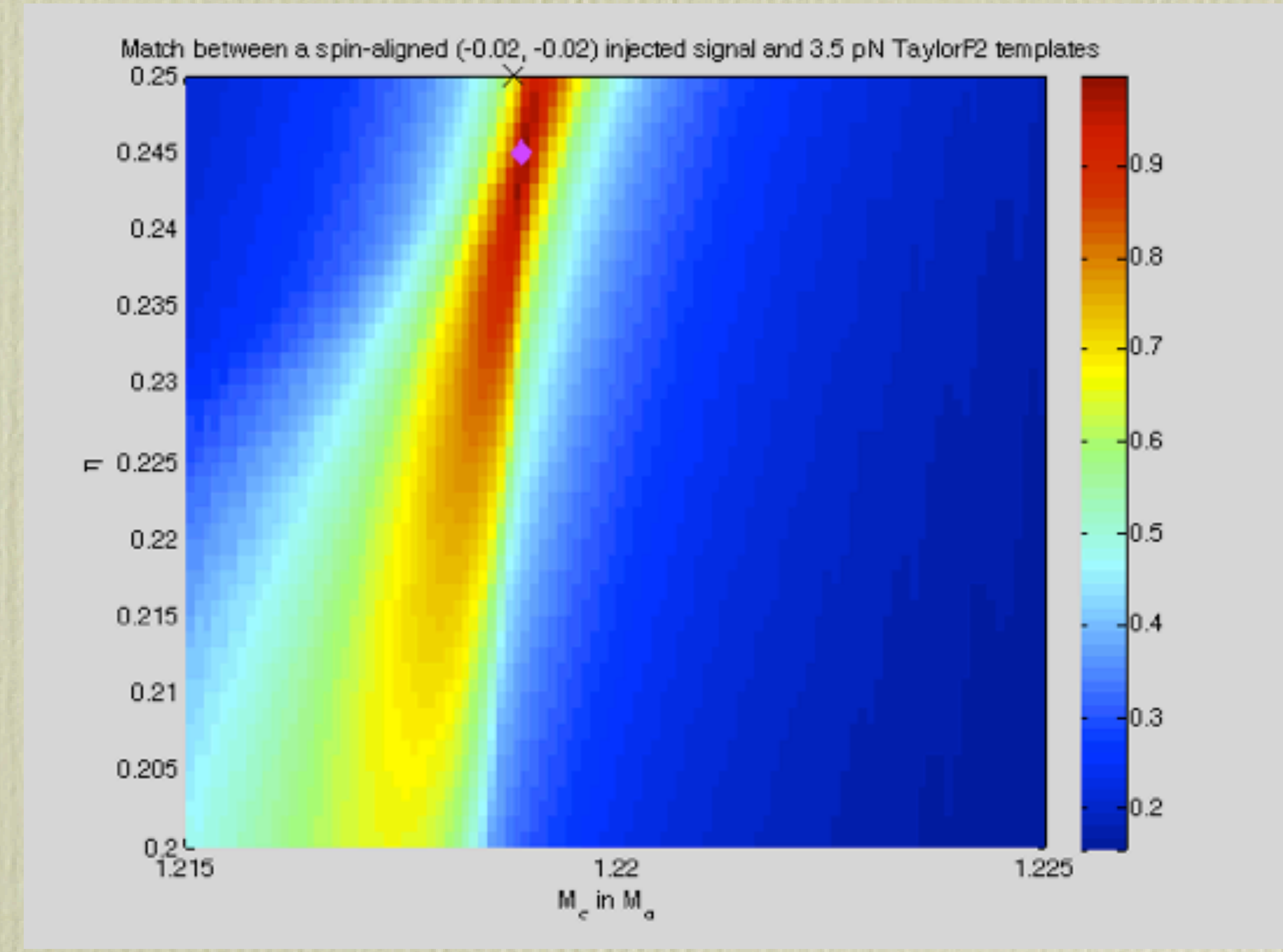
Tidal effects





- 3.5pN+tidal terms injection, 3.5 pN recovery
- Use single-parameter parametrization for tidal deformability [Hinderer et al., 2009], valid to ~500 Hz

Effect of spins



- 3.5pN injection with non-precessing aligned spins (-0.02,-0.02), 3.5 pN recovery
- Best-fit masses are 1.6+1.2 solar masses
- Preliminary results, need to carefully evaluate spin effects

What do ~~data analysts~~ I want?

- A family of IMR waveforms with two spinning, precessing components: hybridized with NR results
- Need practical confidence statements, not just match to NR in regime of matching
- Request: provide several approximate waveform families that are within systematic uncertainty in fits to NR
- *Direct use of NR waveforms for parameter estimation?*

What can data analysts contribute?

- Studies of systematic impacts of variations in waveform families (e.g., NINJA context)
- Improved (relative to the overly strict $|dh|<1$) accuracy requirements on numerical and approximate waveforms
- Accounting for waveform uncertainty directly in Bayesian parameter-estimation methods