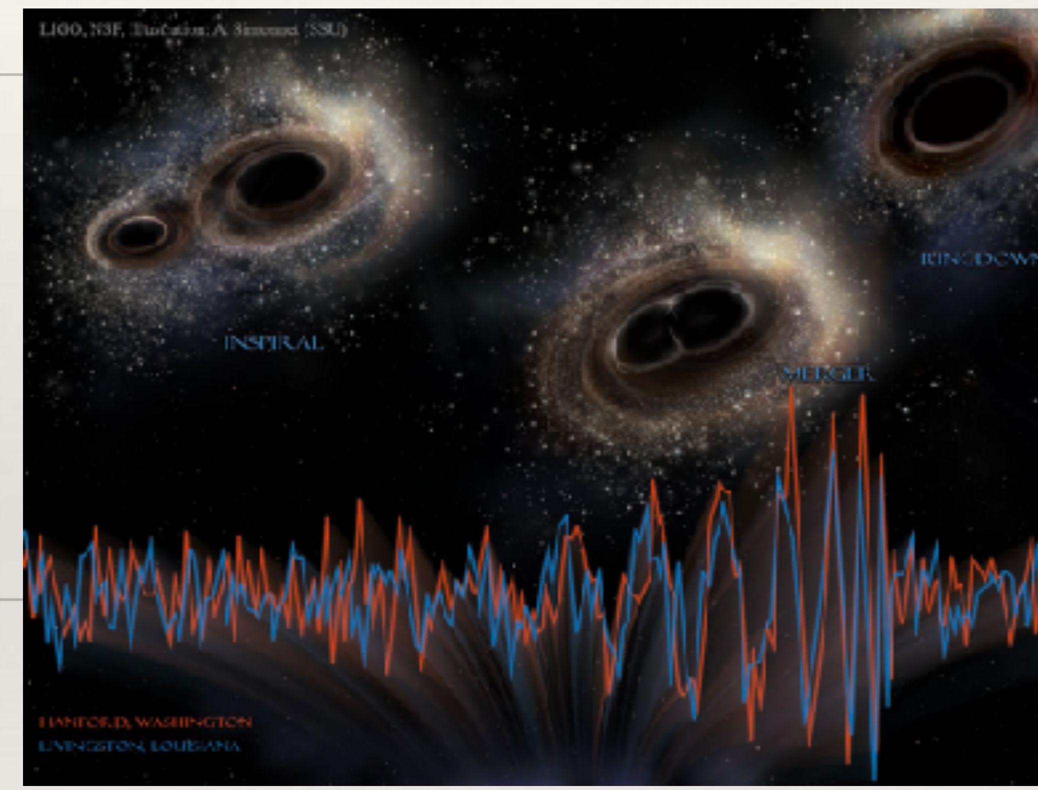
Stanislav (Stas) Babak.

AstroParticule et Cosmologie, CNRS (Paris)

université PARIS DIDEROT PARIS 7

Gravitational waves Part I



ICTP, 18-22 June 2018

Lecture 4

- Data analysis: Part II
- Testing GR with GWs
- Pulsar Timing Array: detecting GWs in the nano-Hz band

Likelihood

Let us assume that the data contains the signal.: hypothesis (model) H_1

$d(t) = n(t) + s(t, \theta_i)$ signal "s" depends on parameters θ_i

data = noise + signal

If the template matches the GW signal exactly $h(t, \theta_i) = s(t, \theta_i) \longrightarrow d(t) - h(t, \theta_i) = n(t)$

$$p(d(t)|H_1, \bar{s}(t, \lambda)) = p(d(t) - s(t, \vec{\lambda})) = p_n$$

- Assume that the noise is Gaussian (but not necessarily white): non white noise has different variance at different frequencies. The the likelihood can be written as

Likelihood: $p(d|H_1, \theta_i) \propto e^{-\frac{1}{2}(d-h(\theta_i)|d-h(\theta_i))}$

The inner product: matched filtering $(a|b) \equiv 4\Re \int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}\, df$

- We search for parameters which maximize the likelihood: making the residuals most noise-like — maximum likelihood estimators for parameters $\hat{\theta}_i$

Likelihood

- For a given noise realization the maximum likelihood estimators. $\hat{\theta} \neq \theta_{true}$
How close the estimated parameters to the true depends on the noise realization and on the strength of the signal: stronger the signal (high signal-to-noise ratio, SNR) closer estimators to the true values — less influence of the noise. Unbiased estimator: equal to true if averaged over the noise realizations. $< \hat{\theta} >= \theta_{true}$

- If $s \neq h(\theta_{true})$ — lack of accuracy in the signal modelling: systematic bias in parameter estimation.

$$min_{\theta_i}(s - h(\theta_i)|s - h(\theta_i)) \quad \rightarrow \tilde{\theta}_i \qquad |\theta_{true} - \tilde{\theta}_i| = \delta\theta_i \quad \text{bias}$$

If there is a bias, we still can detect the GW signal on expense of making error in parameters charactrizing the system (binary): *effectualness*

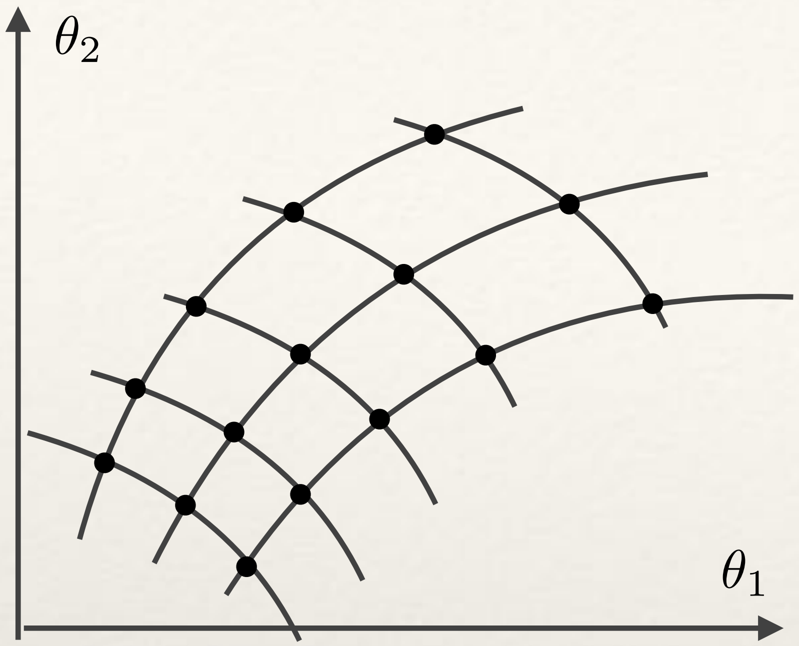
Faithfulness $$\frac{(s|h(\theta_{true})}{\sqrt{(s|s)(h(\theta_{true})|h(\theta_{true}))}}$$

Overlap $$(\hat{s}|\hat{h}), \quad \hat{h} = \frac{h}{\sqrt{(h|h)}}$$

$\hat{h}$ → normalized

Overlap varies [-1, 1]: 1 is a pefect match, related to the loss in SNR

Likelihood maximization

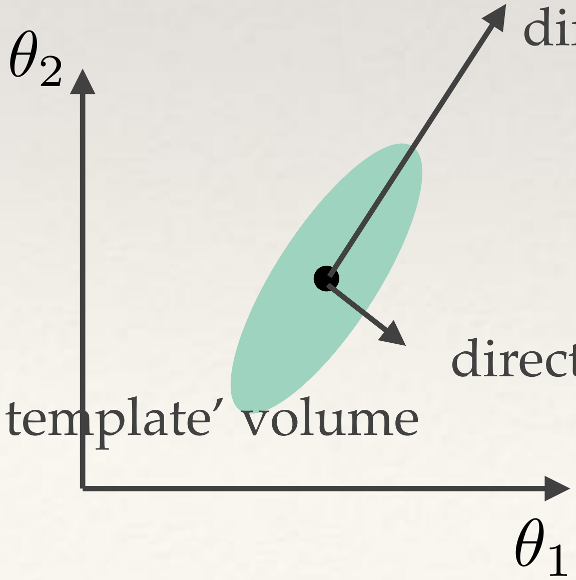


- We want to cover the parameter space (N-dim) by grid of points at equal distance from each other.
- Grid: not too coarse, not too fine
- The distance is determined **not** by a coordinate distance but by “proper” distance — correlation between nearby templates: introduce interval and metric

$$ds^2 = |\hat{h}(\theta_i + \delta\theta_i) - \hat{h}(\theta_i)| \approx (\hat{h}(\theta_i + \delta\theta_i) - \hat{h}(\theta_i)|\hat{h}(\theta_i + \delta\theta_i) - \hat{h}(\theta_i)) \approx \left(\frac{\partial \hat{h}}{\partial \theta_i}|\frac{\partial \hat{h}}{\partial \theta_j}\right)\delta\theta_i\delta\theta_j$$

metric on the parameter maniold

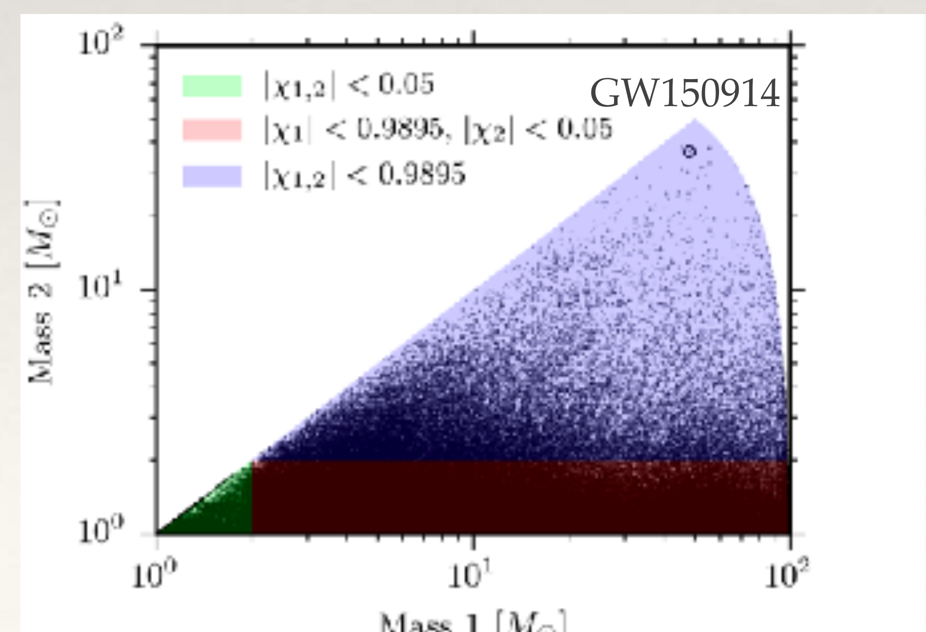
Consider 2-D parameter space and fix ds = 0.01



direction of strong correlation

the size and orientation of the ellipse is a function of a central pont

direction of weak correlation (chirp mass)



Bayesian approach

- Expensive computationally: often used when the signal is detected using the grid-based method (or something else) : H_1 is true. Allows to test several models (different signal' models, non-GR theories).
- We have to assign the ***prior*** probability to our models and parameters of each model. We treat parameters describing a signal as random variables and trying to estimate probability distribution function(s) for each parameter based on the observed data (***posterior***)
- Consider several models M_i each parametrized by set of parameters $\vec{\theta}_i$

Bayes' theorem
$$P(M_i|d) = \frac{P(d|M_i)\pi(M_i)}{p(d)}$$

For a given model M_i

likelihood

prior

$$P(\vec{\theta}_i|M_i, d) = \frac{P(d|\vec{\theta}_i, M_i)\,\pi(\vec{\theta}_i)}{p(d|M_i)}$$

posterior

Evidence of model M_i

Bayesian approach

$$p(d|M_i) = \int d\vec{\theta}_i \; p(d|\vec{\theta}_i, M_i)\pi(\vec{\theta}_i)$$

— Evidence of model M_i: important for the model selection

$$P(M_i|d) = \left[\int d\vec{\theta}_i \; p(d|\vec{\theta}_i, M_i)\pi(\vec{\theta}_i)\right] \frac{\pi(M_i)}{P(d)}$$

— probability of the model M_i given a data

- **Odd ratio**:The problem to evaluate the normalization P(d) - requires full set of models which are mutually exclusive. We can evaluate the ratio of probabilities:

$$O_{a,b} = \frac{p(M_a|d)}{p(M_b|d)} = \frac{p(d|M_a)}{p(d|M_b)} \frac{\pi(M_a)}{\pi(M_b)}$$

Bayes factor prior odds

Markov Chain Monte Carlo (MCMC)

- So for a given model M_i we need to evaluate postrior pdf for all parameters and the evidence: posterior pdf tells us about parameters of the GW signal (system) and evidence tells us how good this models fits the observations.

$$p(\vec{\theta}_i|M_i, d) = \frac{p(d|\vec{\theta}_i, M_i)\ \pi(\vec{\theta}_i)}{p(d|M_i)}$$ — posterior pdf

$$p(d|M_i) = \int d\vec{\theta}_i\ p(d|\vec{\theta}_i, M_i)\pi(\vec{\theta}_i)$$ — evidence of model M_i

- ***Markov Chain Monte Carlo (MCMC)*** approach:

We construct Markov chain: stochastic process where the next point in the chain depends only on the previous one. And:

- we want chain to move towards the region of parameter space with high likelihood
- we need to introduce the transitional probability: way to move from one point to another
- we want a transitional probability to satisfy the ballance equation

ballance eqn.: $$P(\vec{\theta}_{(k)})P(\vec{\theta}_{k+1}|\vec{\theta}_{(k)}) = P(\vec{\theta}_{(k+1)})P(\vec{\theta}_{k}|\vec{\theta}_{(k+1)})$$

distribution we want to sample (posterior)

transitional probability

MCMC

- Consider a particular implementation: Metropolis-Hastings
  - particular way of building transitional probability which satisfies the ballance equation
  - we start with introducing a proposal distribution (arbitrary*) $q(\vec{\theta}_{(k+1)}|\vec{\theta}_{(k)})$
  - then we build the chain by introducing the acceptance probability

$$\alpha(\vec{\theta}_{(k+1)}|\vec{\theta}_{(k)}) = min\left\{1, \frac{p(d|\vec{\theta}_{(k+1)})}{p(d|\vec{\theta}_{(k)})}\frac{q(\vec{\theta}_{(k)}|\vec{\theta}_{(k+1)})}{q(\vec{\theta}_{(k+1)}|\vec{\theta}_{(k)})}\frac{\pi(\vec{\theta}_{(k+1)})}{\pi(\vec{\theta}_{(k)})}\right\}$$

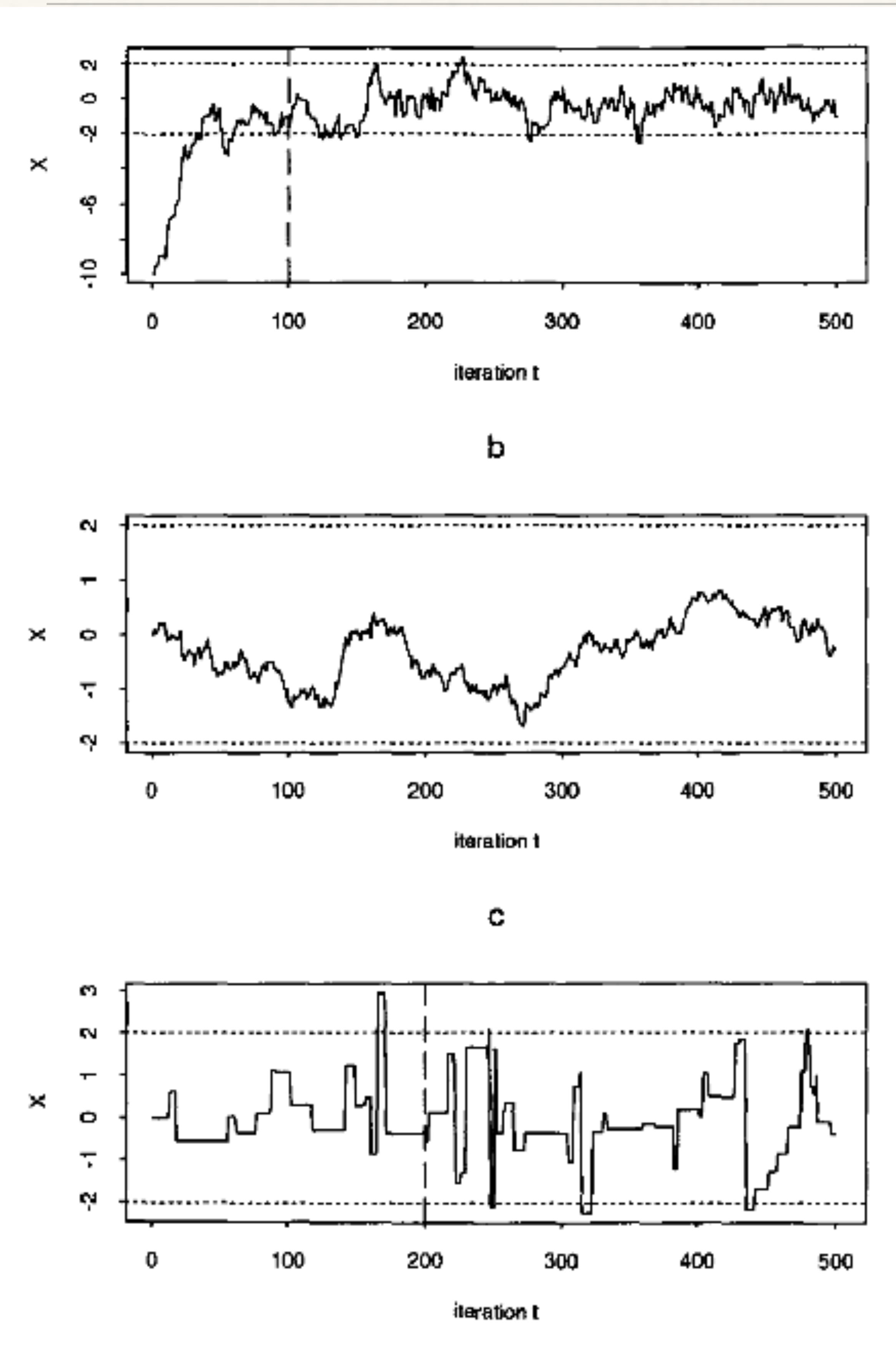
likelihood ratio **ratio of proposals** **ratio of priors**

It is easier to understand if we use symmetric proposal and uniform priors: likelihood ratio

α — probability of accepting new point $\vec{\theta}_{(k+1)}$

The chain moves predominantly in the direction of high likelihood.

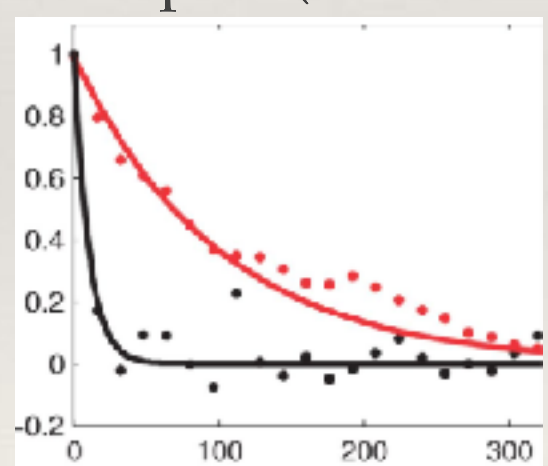
MCMC



- The theorem tells us that the chains will sample the posterior pdf (after some burn-in length) independent of the proposal distribution, BUT

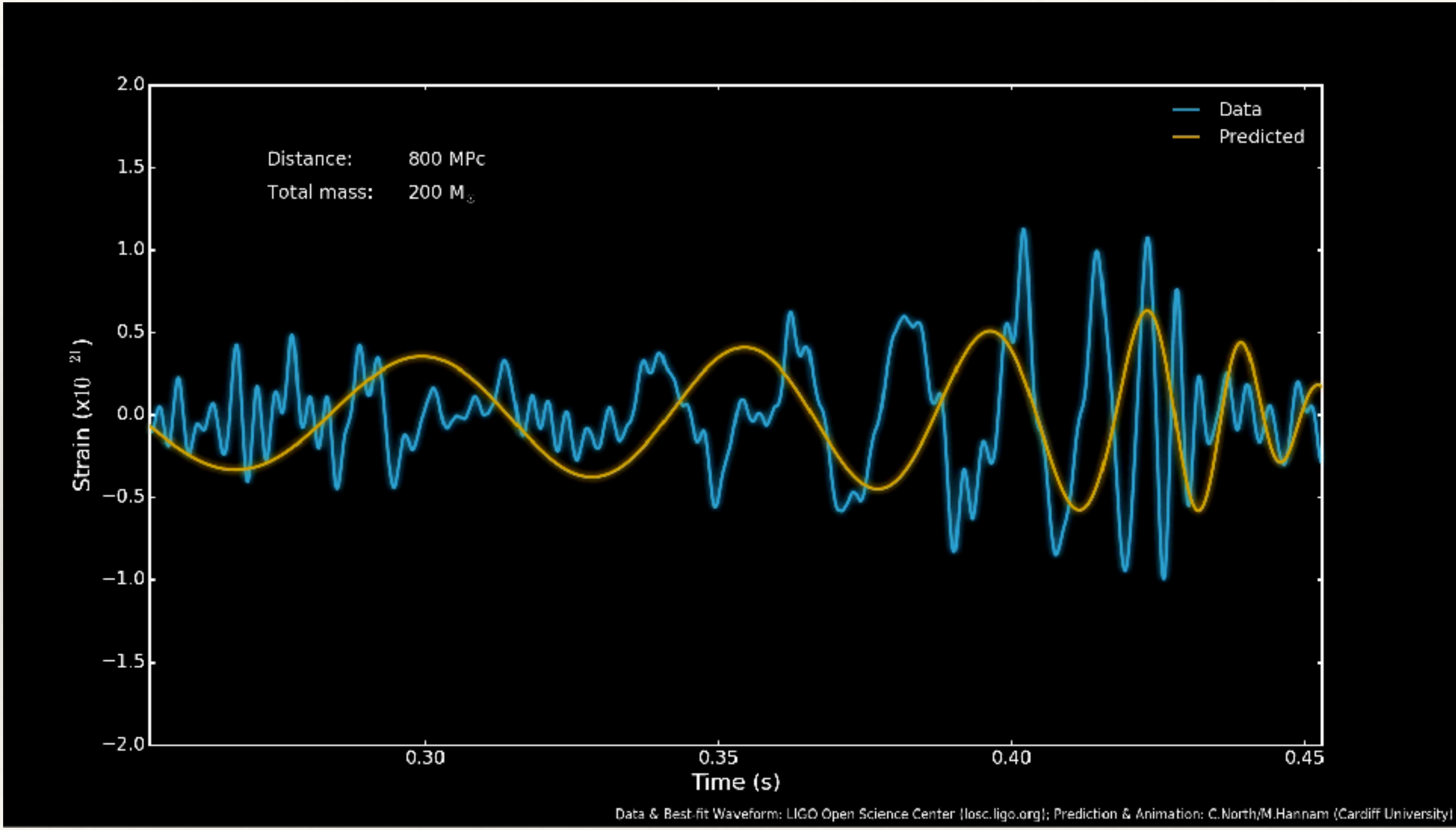
- The efficiency of the sampling strongly depends on the proposal (proposal should resemble the posterior)

- Number of samples vs. number of independent samples (defined by autocorrelation length)



- Multimodal posterior require special treatment! (simulated annealing, parallel tempering)

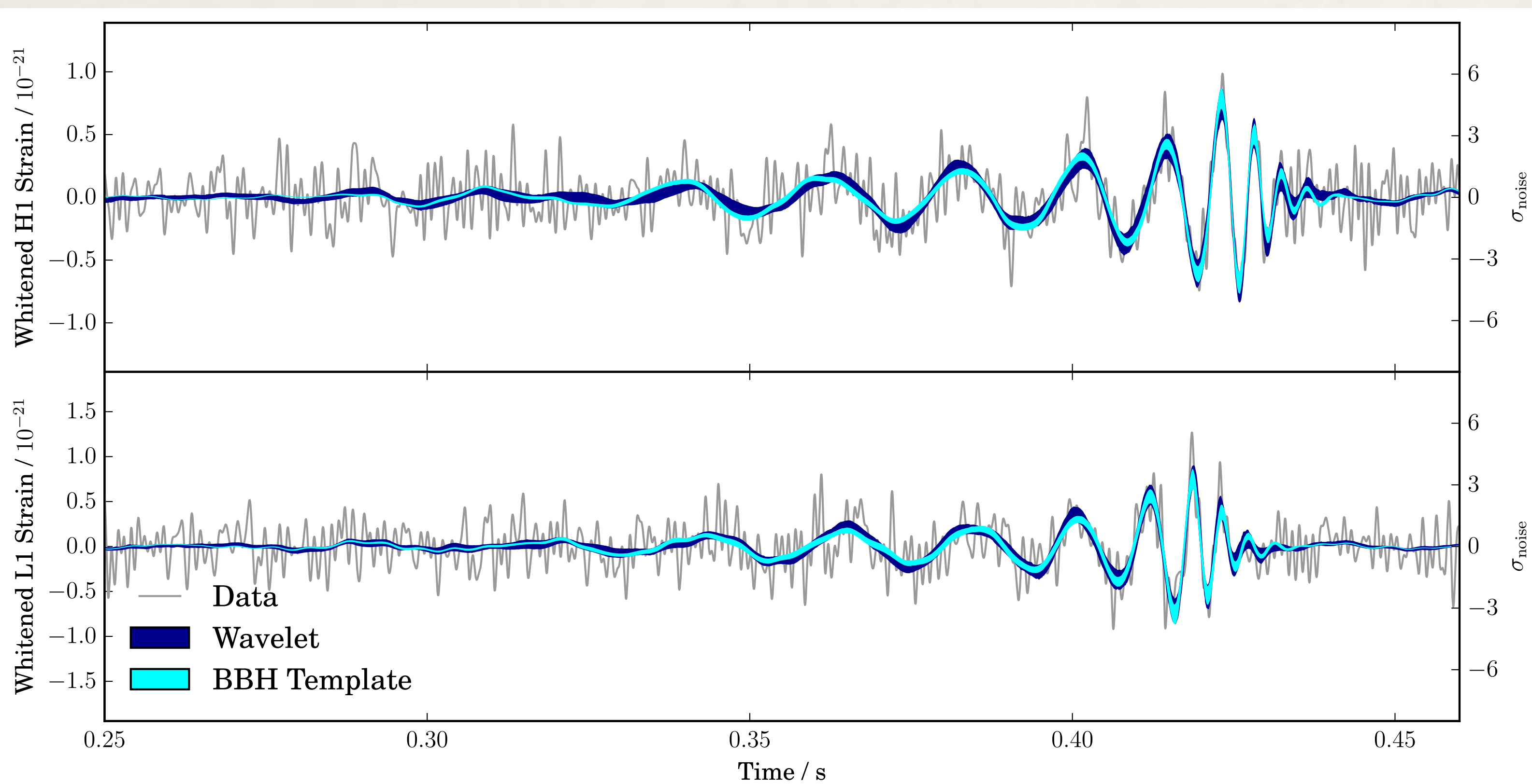
GW data analysis

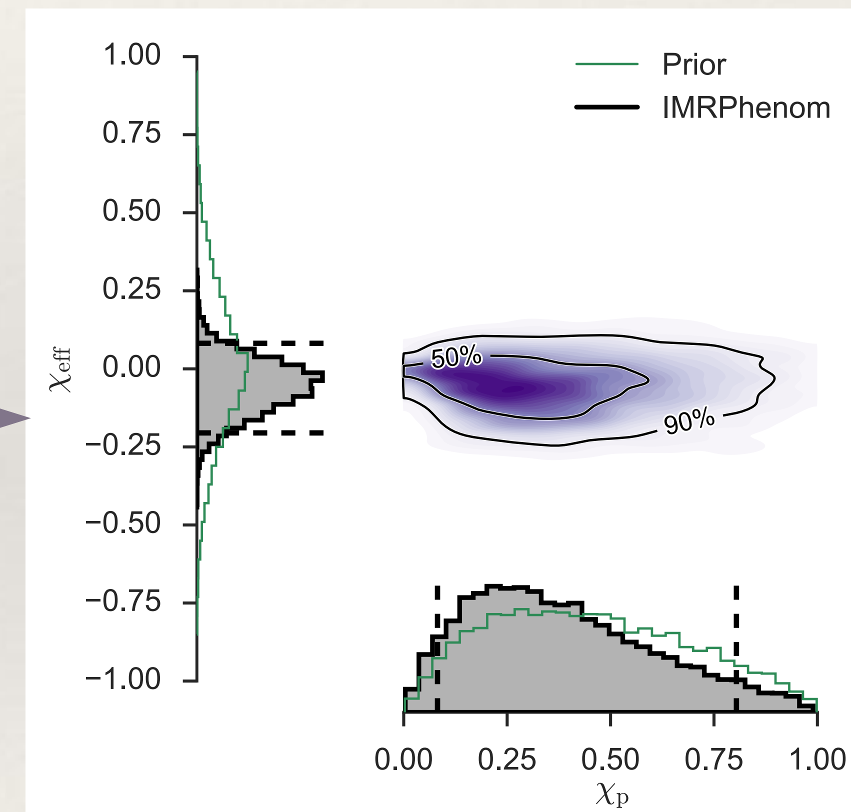


GW data analysis

$$p(\theta|d) = \frac{p(d|\theta)\ \ p(\theta)}{p(d)}$$

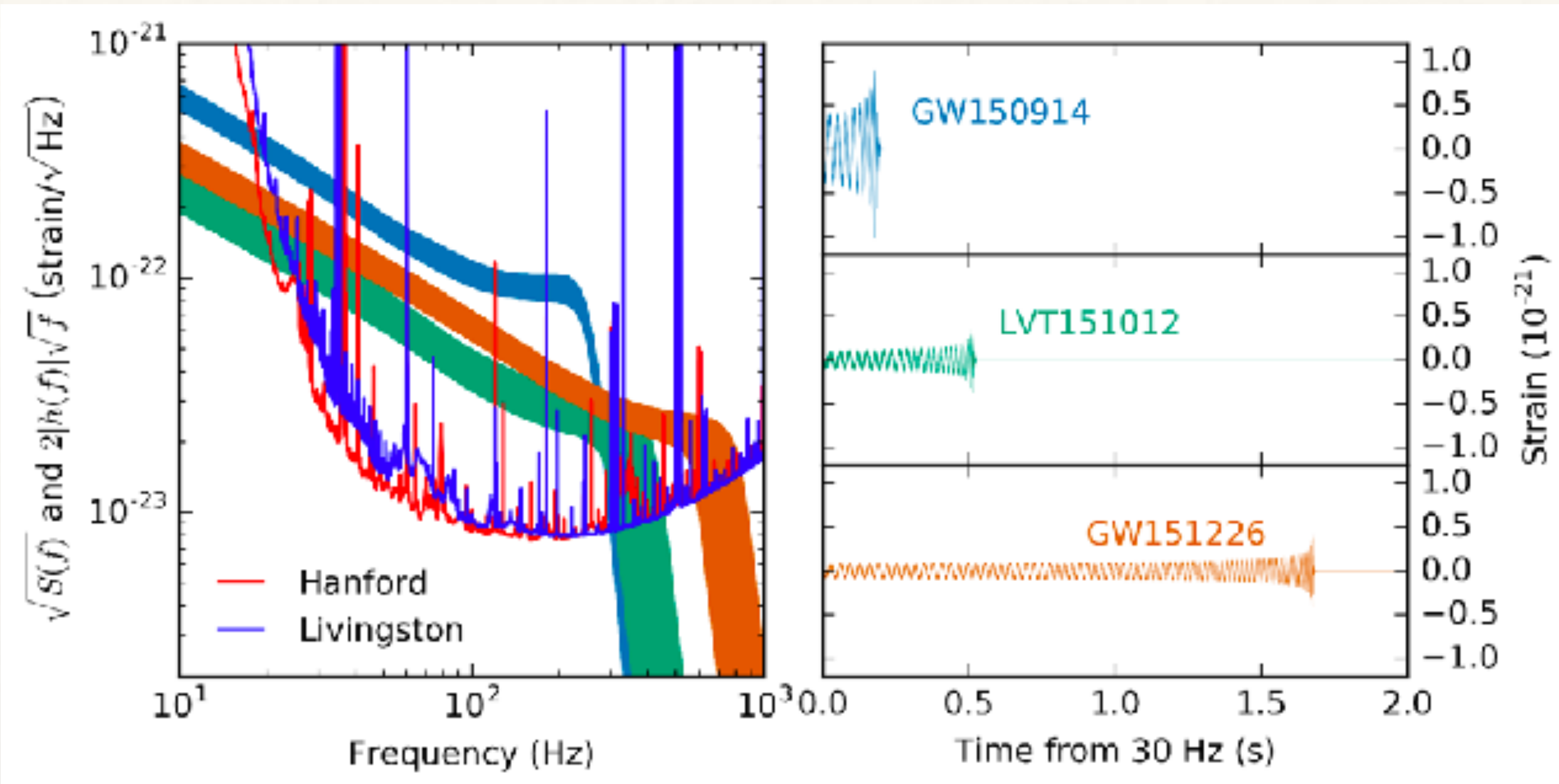
Fit to the data not a single line but multiple (for each parameter set in the posterioir)





[GW150914, LSC+VIRGO PRL (2016)]

GW data analysis



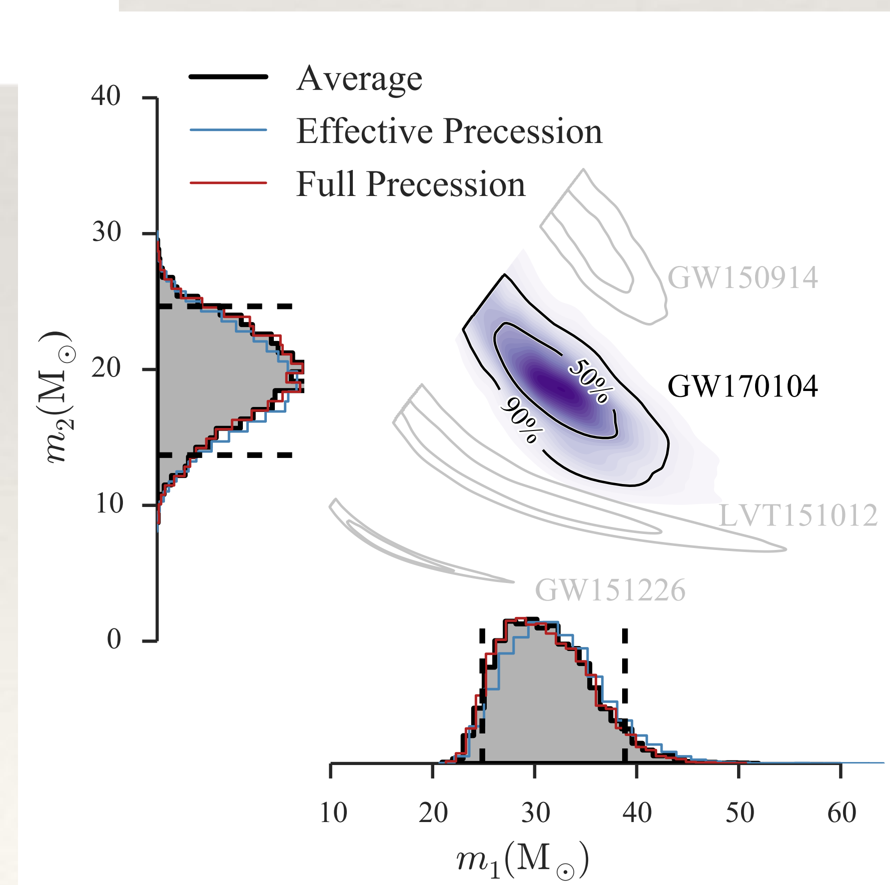
Frequency domain signal

Time domain signal

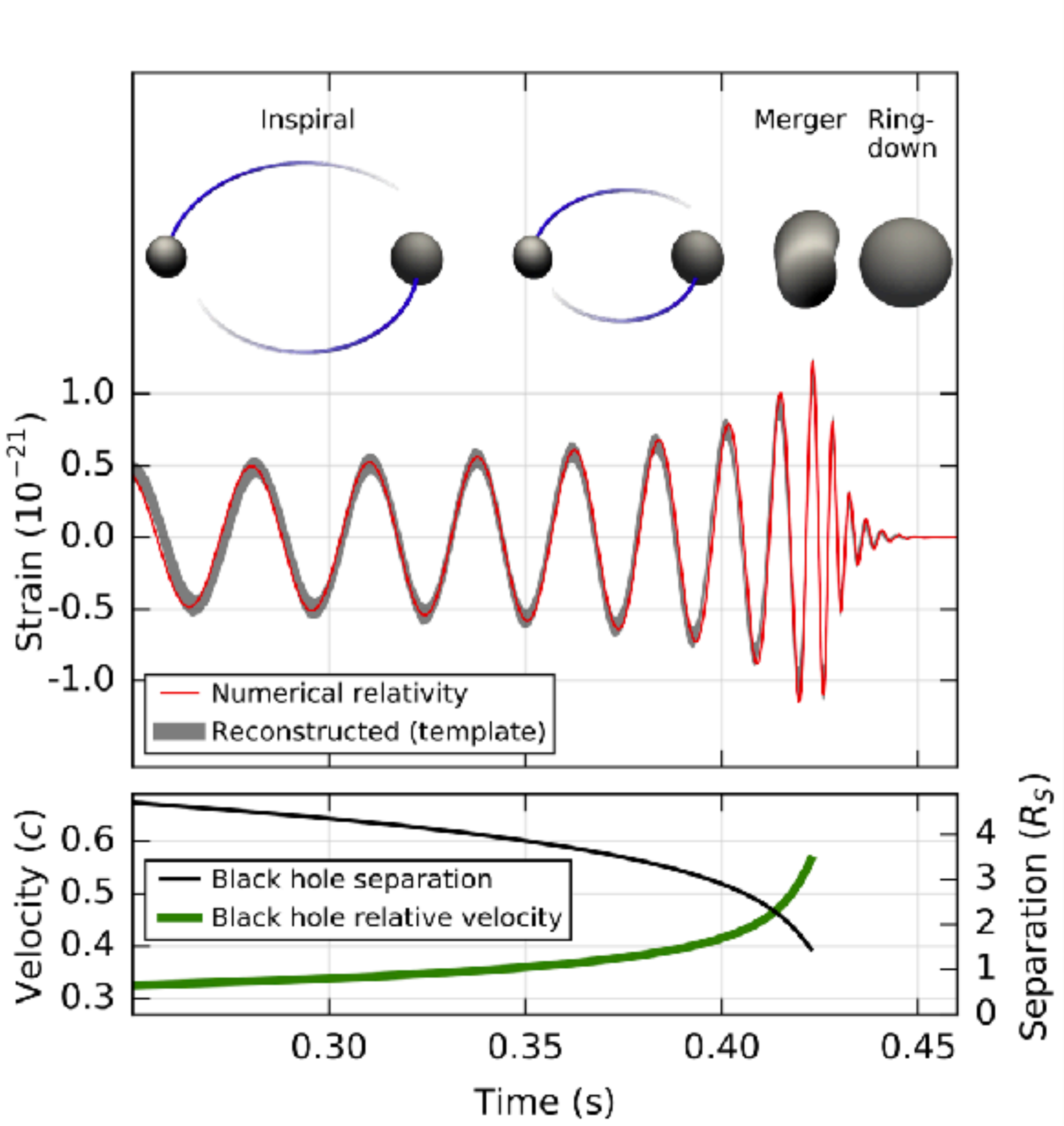
Chirp mass is the best measured parameter for the signal dominated by inspiral

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$

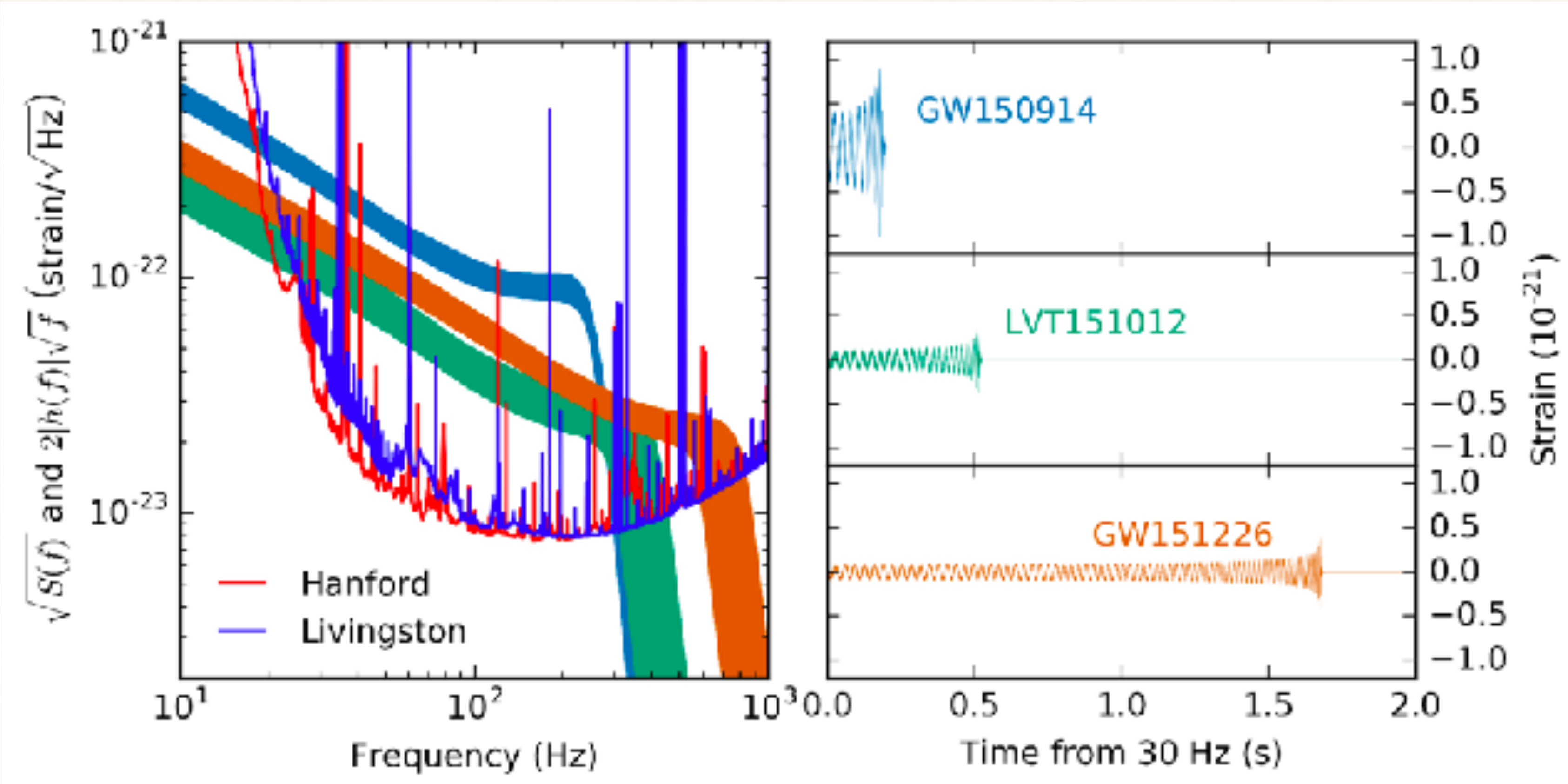
Estimation of BHs masses



GW data analysis



GW data analysis

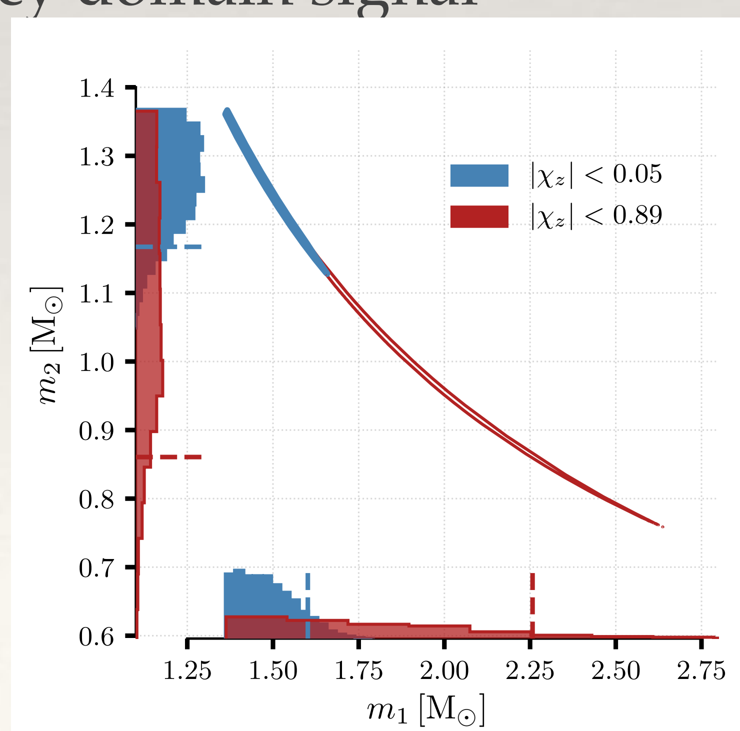


Frequency domain signal

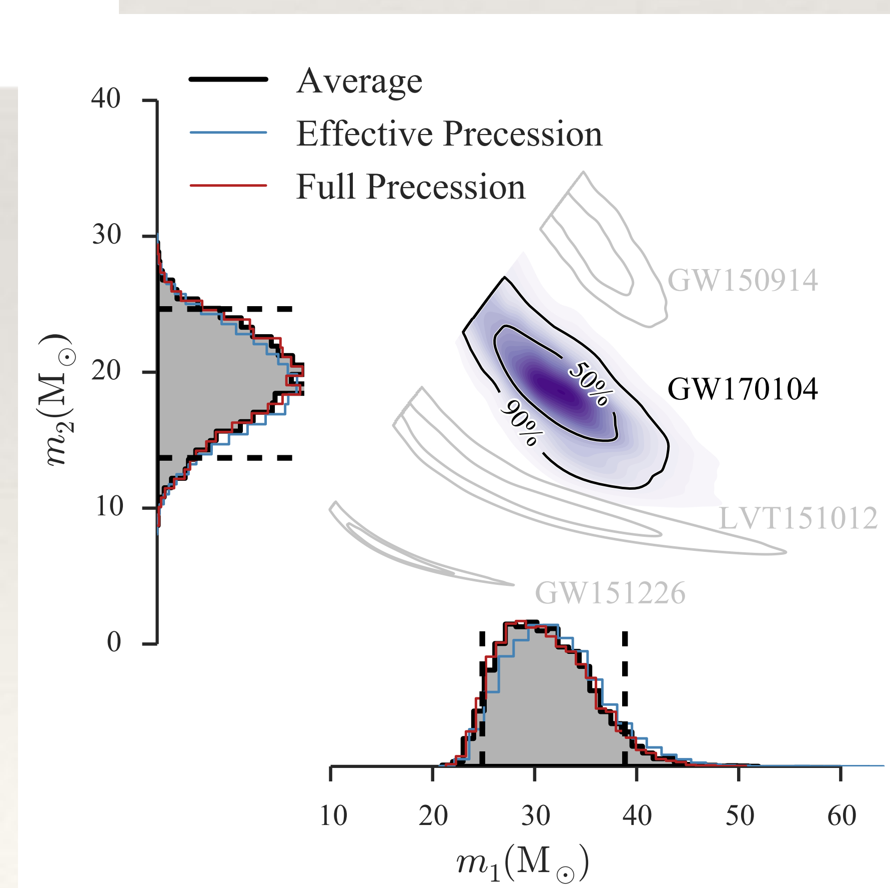
Time domain signal

Chirp mass is the best measured parameter for the signal dominated by inspiral

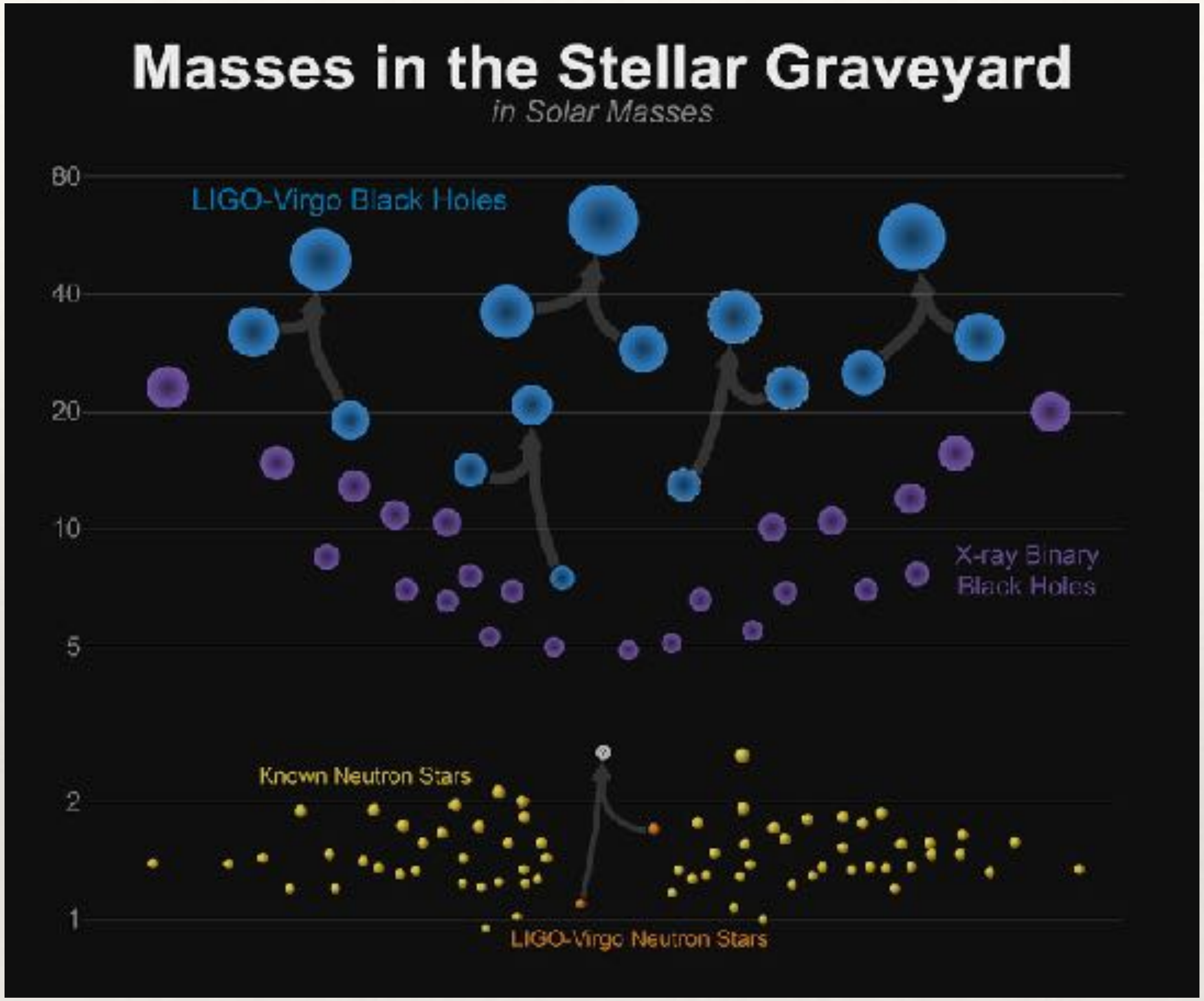
$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$



Estimation of BHs masses

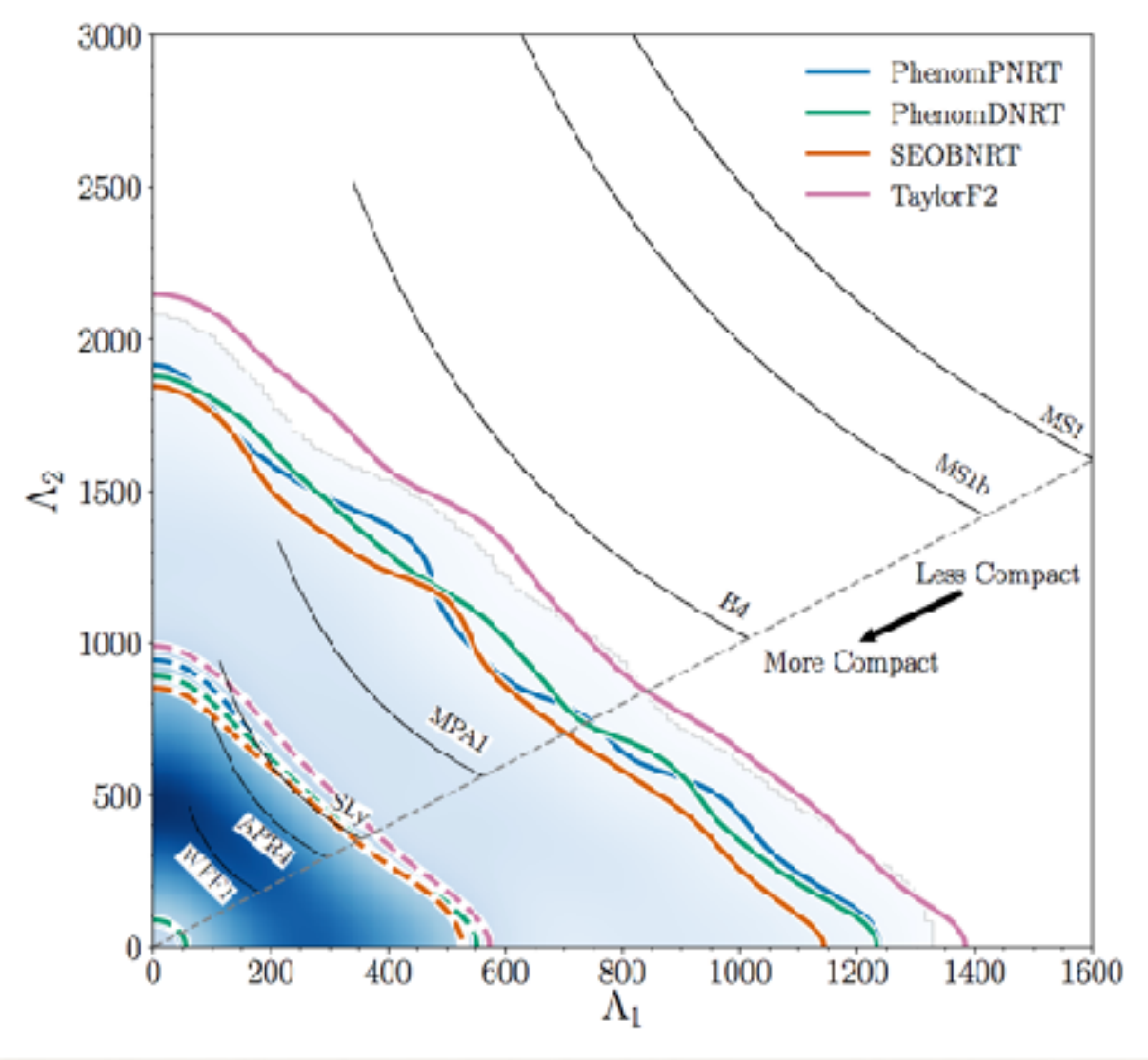


Detected GWs

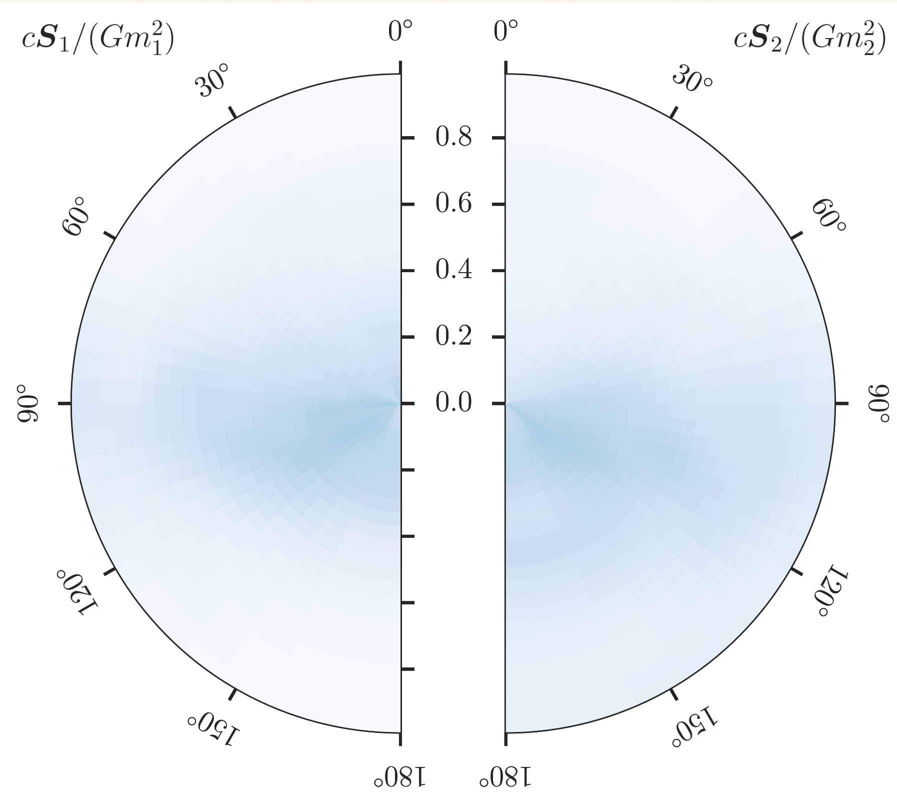


Mass distribution (LIGO)

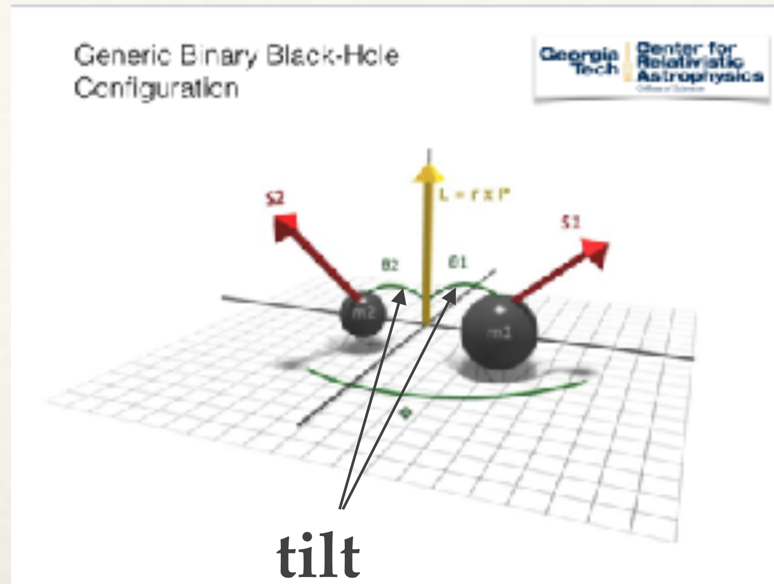
NS EOS



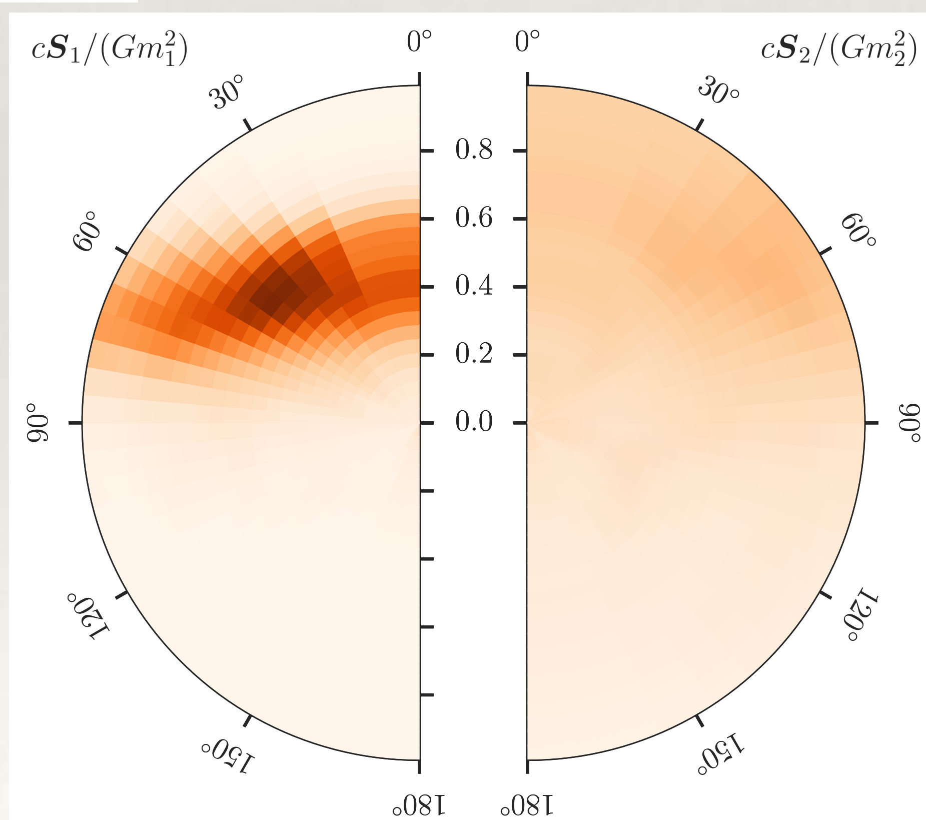
Black Hole spins

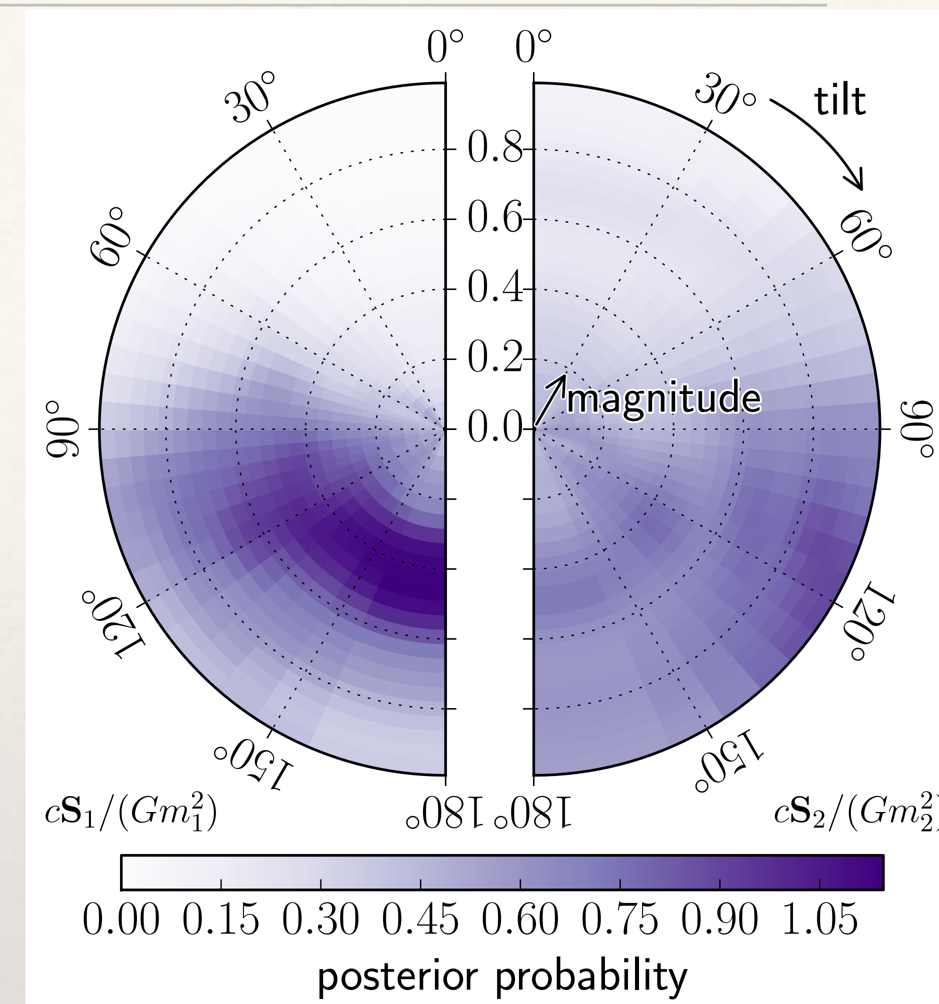


GW150914



GW151226

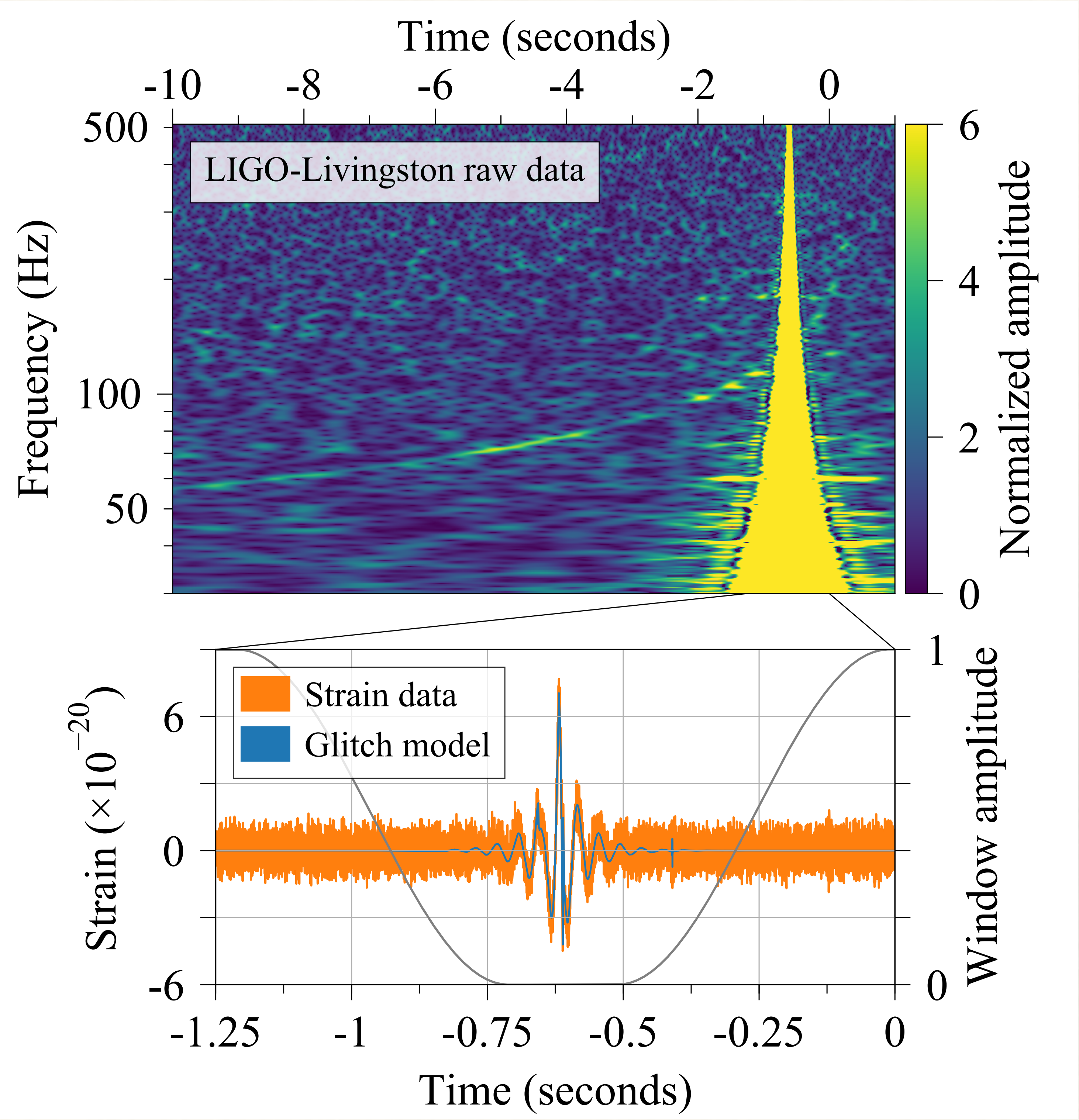




GW170104

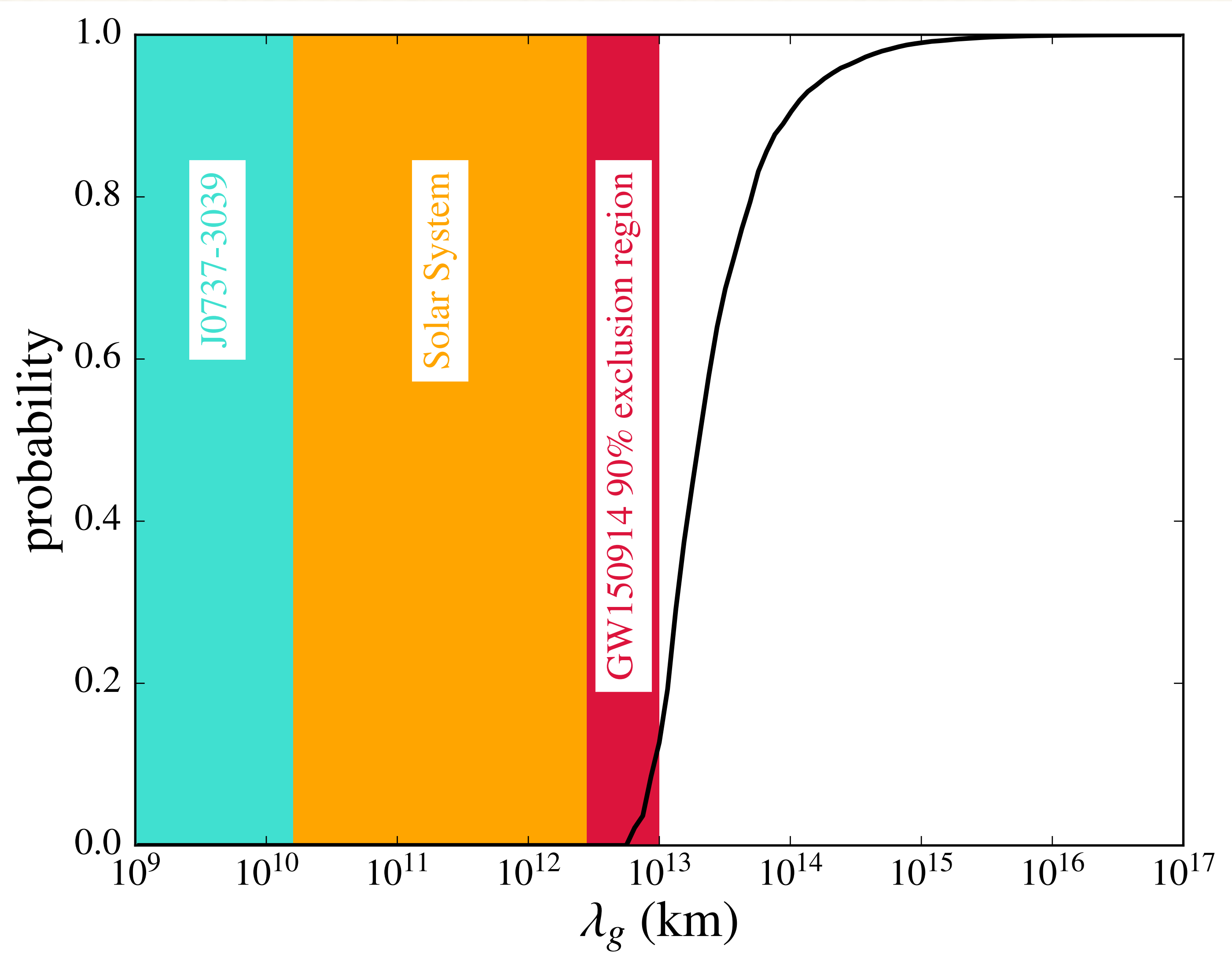
[LSC+VIRGO PRL (2016, 2017)]

Again non-stationarity



Strong glitch during the GW signal from BNS system

Testing GR



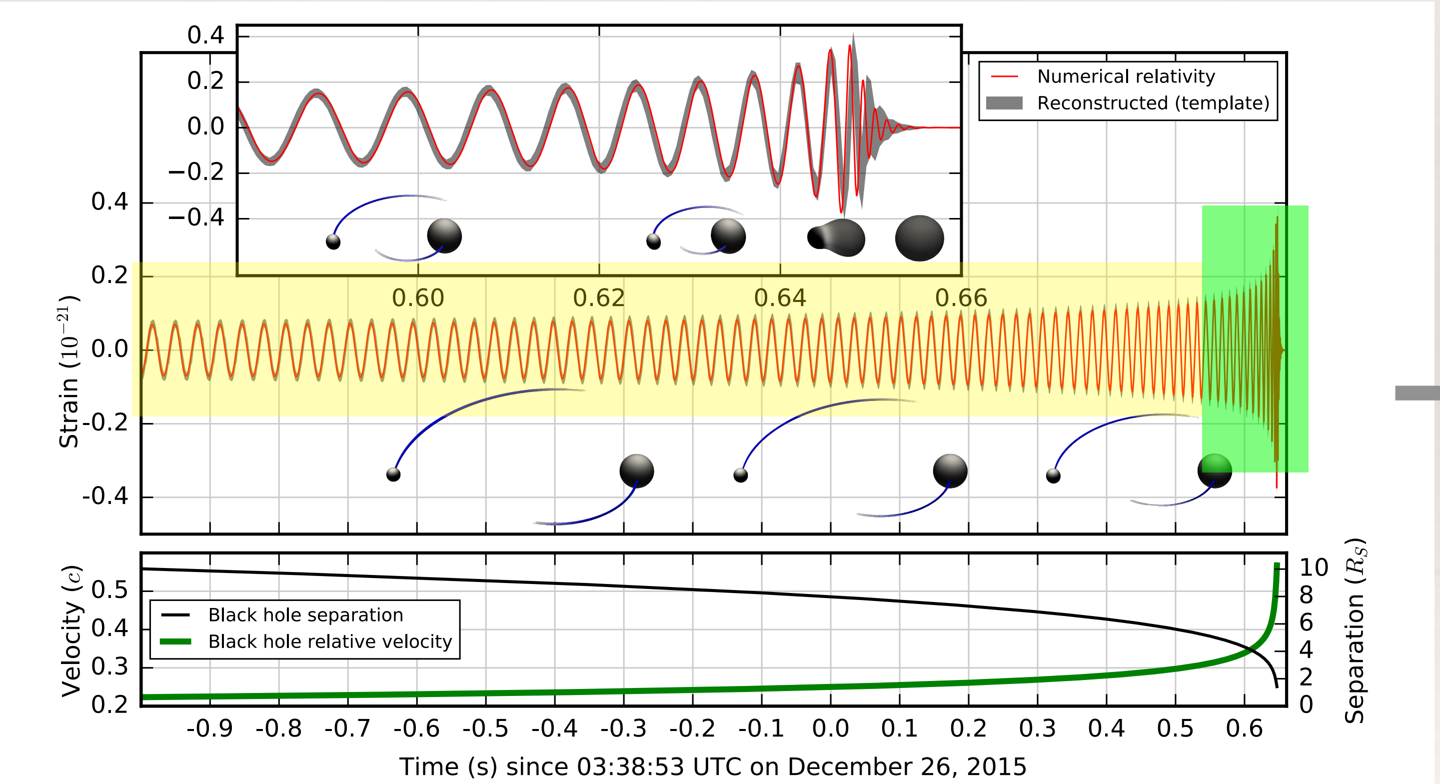
- In GR graviton is massless (propagates with the speed of light)
- If graviton has non-zero mass (m_g) then we should observe dispersion of GWs
- The observed data does not contradict GR: we can set upper limit on the mass of grviton (lower bound on the Compton wavelength

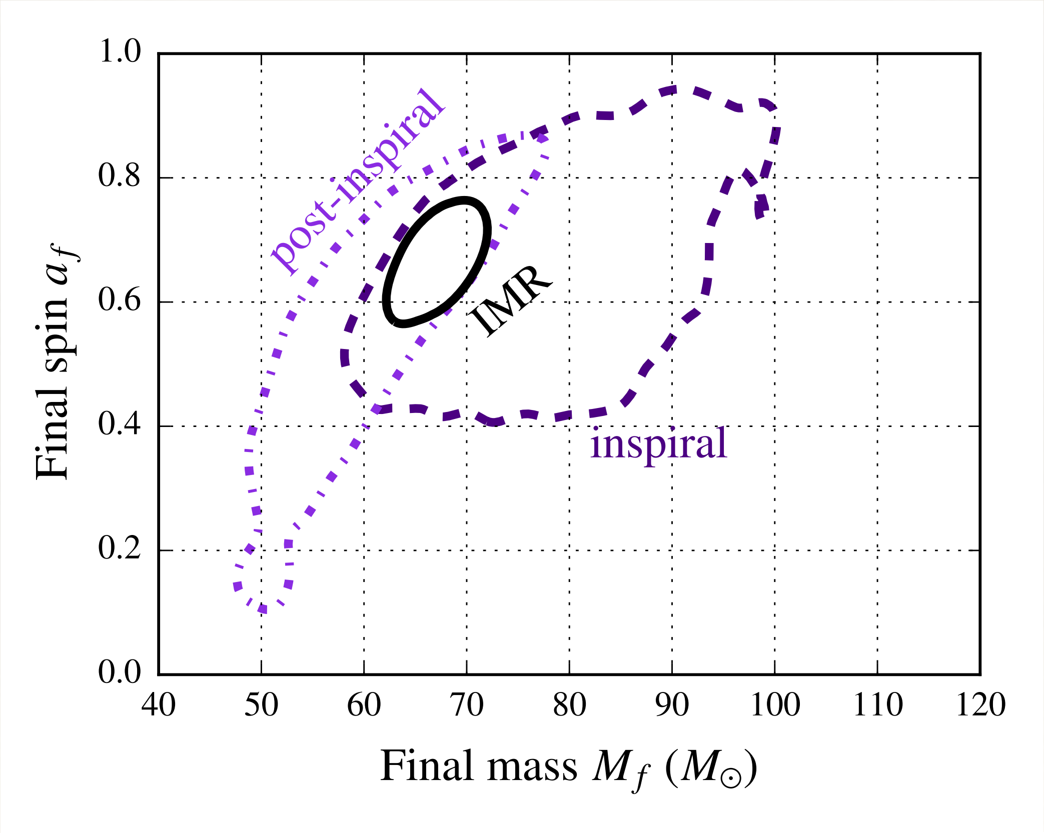
$$\lambda_g = h/(cm_g)$$

[LVC PRL (2016)]

Testing GR

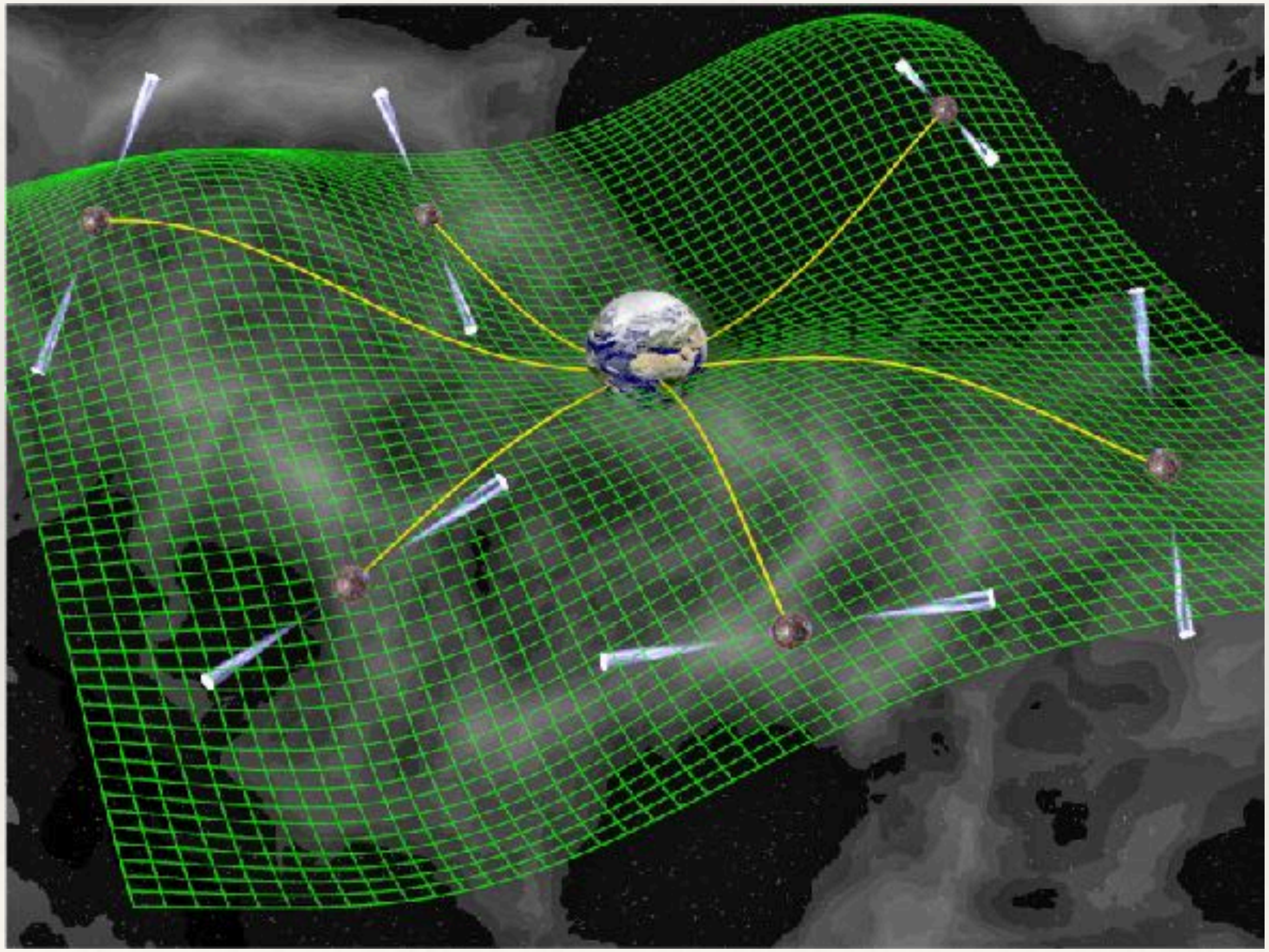
- We can test self-consistency of GR
- We split the signal in two parts and analyze both parts separately: estimate parameters
- We compare if parameter estimation from each part are consistent





Pulsar Timing Array

The main idea behind pulsar timing array (PTA) is to useultra-stable millisecond pulsars as beacons (clocks sending signals) for detecting GW in the nano-Hz range (10^{-9} - 10^{-7} Hz).

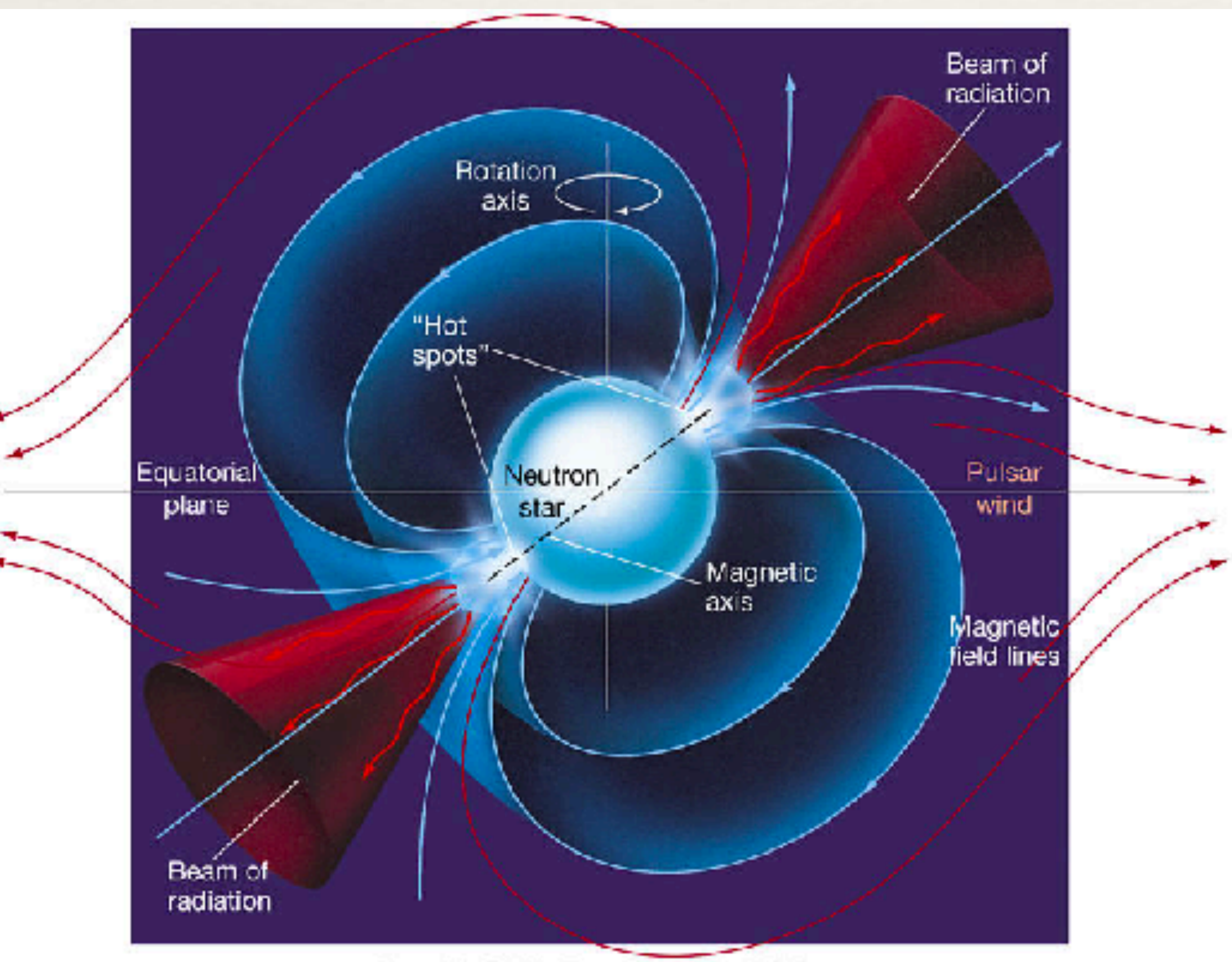


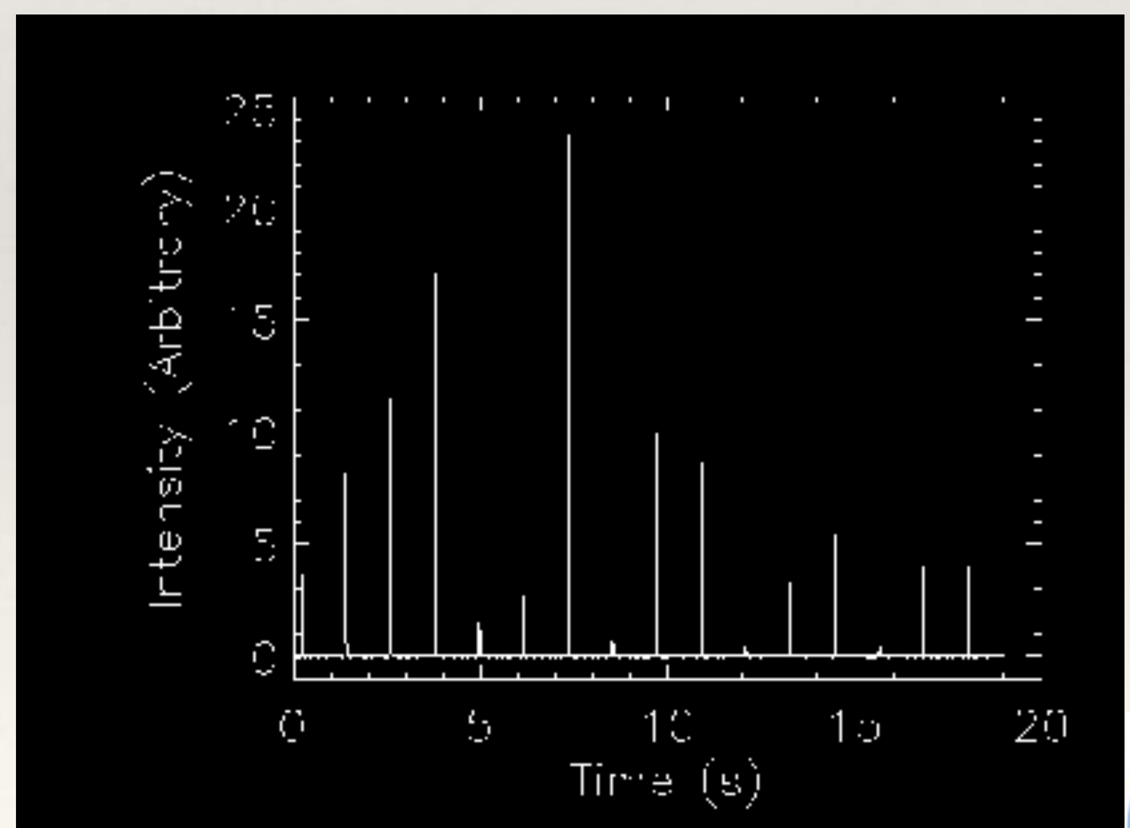
[Credits: D. Champion]

EPTA

Millisecond pulsars

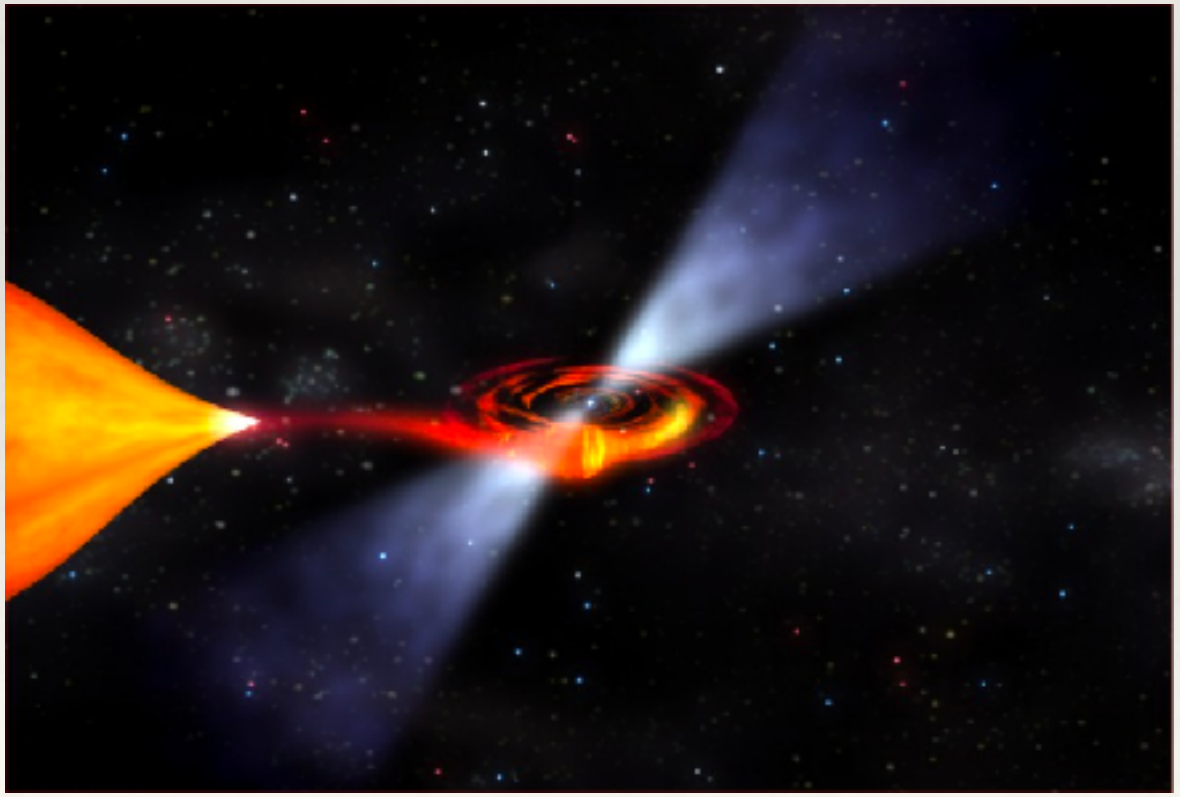
- Pulsars - neutron stars (end product of evolution of stars with the mass > 7 solar) with rapid rotation and strong magnetic field
- Emit beamed e/m radiation from the magnetic poles. Powered by rotation: spinning down.
- Beamed radio emission swaps across the line of sight — seen as pulses in observations (similar to the lighthouse)



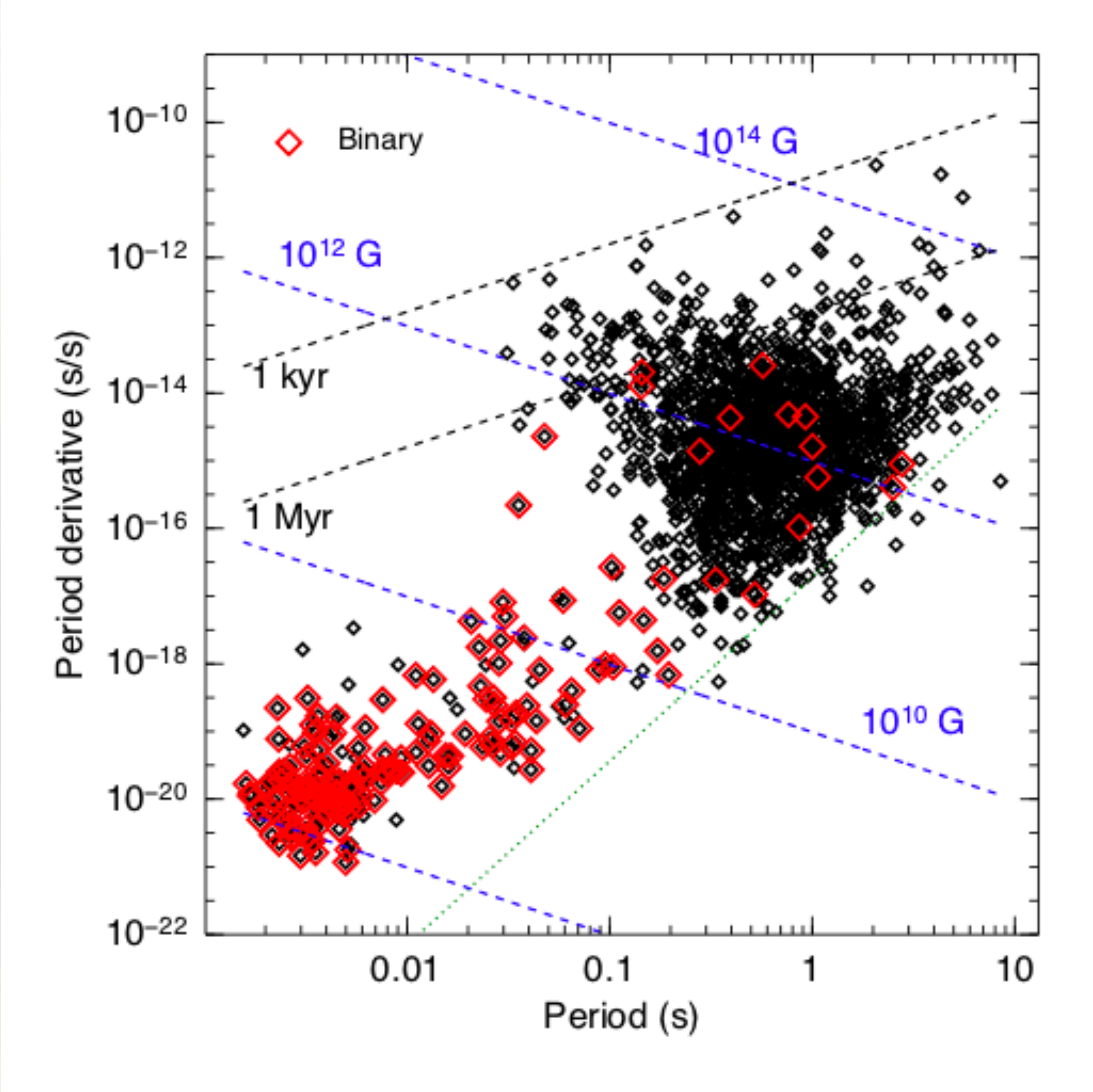


Millisecond pulsars

- Millisecond pulsars: period of rotation ~ millisec
- Often in binaries
- Very old NSs, very stable rotation
- The most accurate clock on the long time scale (decades)

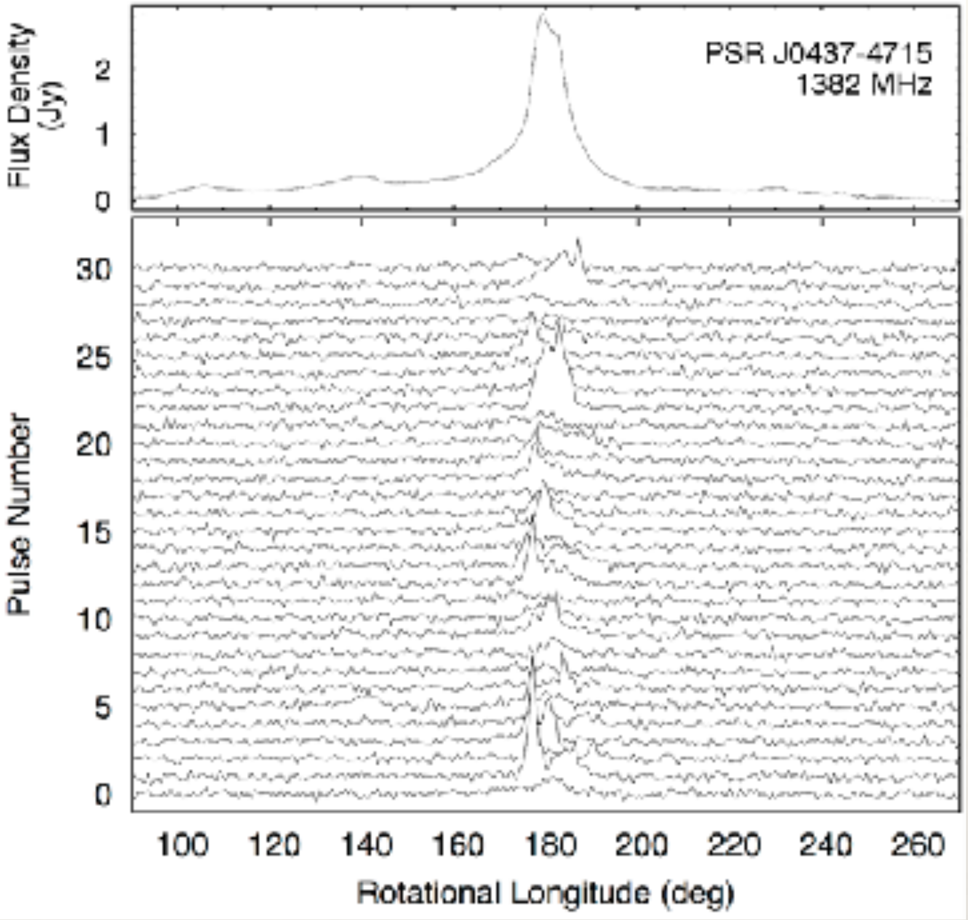


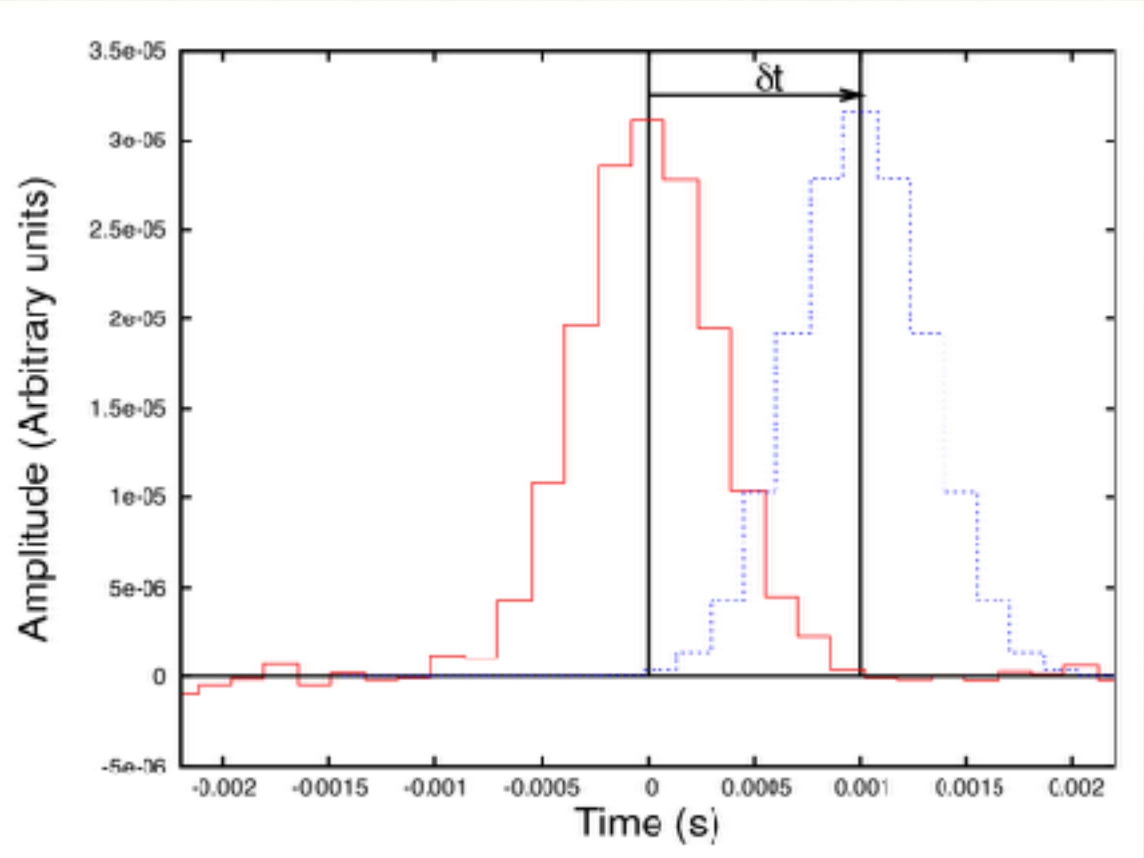
[Credits: NASA]



[Wikimedia]

Pulsar timing

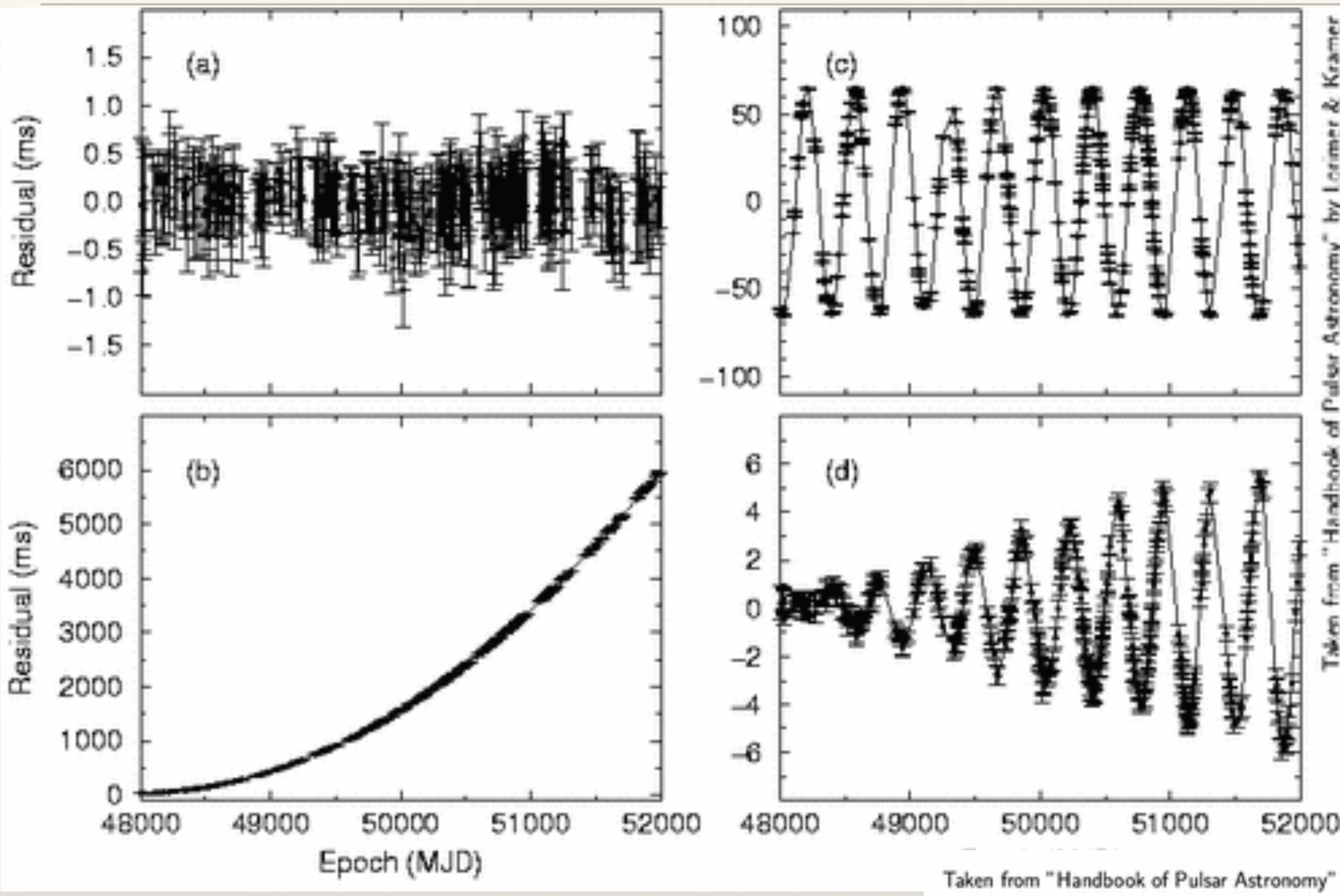




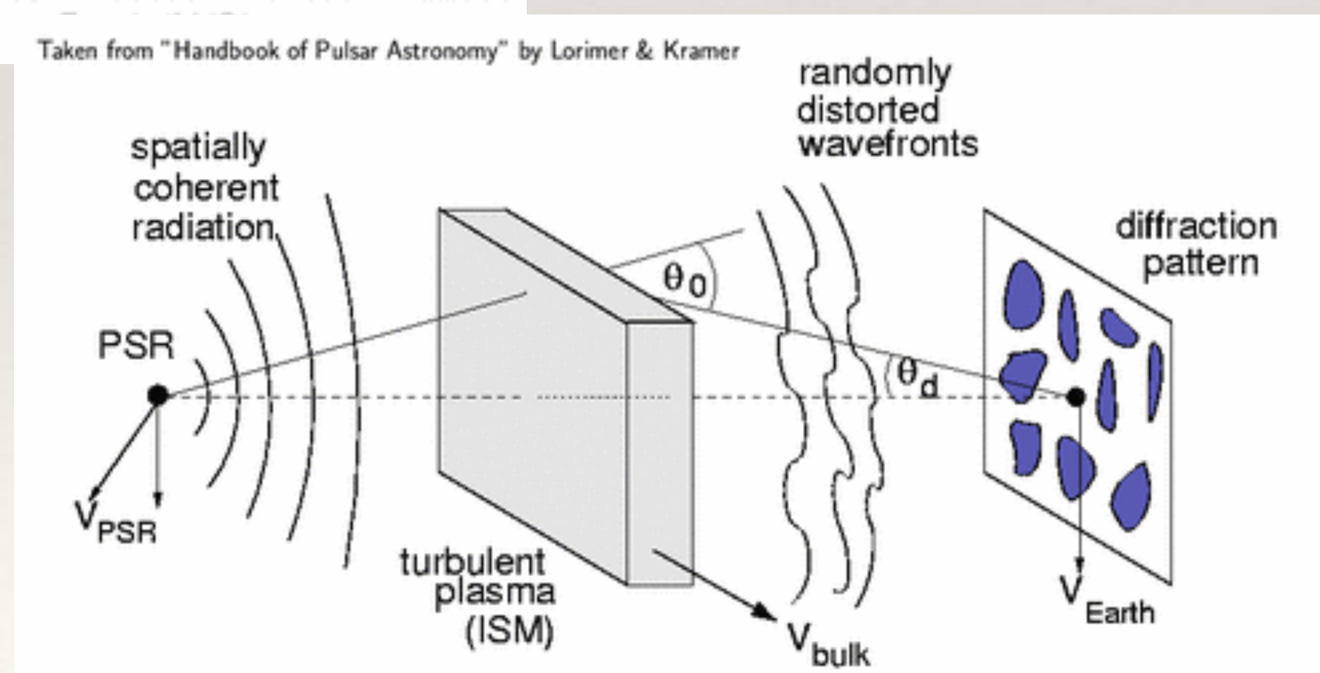
[Figs: credits
S. Burke-Spolar & L. Lentati]

- Each observed radio pulse profile has a lot micro-structure. If we average over ~hour the (average) profile is very stable
- We can use the average pulse profile to estimate the time-of-arrival (TOA) of the pulses.
- The idea is to measure the TOA, and compare to the expected TOA. We know the spin of the pulsars, so we can predict the TOA. The difference between measure and expected TOA: ***residuals***

Timing pulsars



- We need to build a timing model to make accurate prediction for TOAs - take into account various physical effects
- Dispersion of e/m wave and its time dependence
- Rate of change of rotation (b)
- Sky position of the pulsar (c)
- Proper motion of a pulsar (d)



- Timing model could be quite complex if pulsar is in the binary

Residuals

- Building the timing model: depends on many parameters

$$t_{toa} = t_{toa}(P, \dot{P}, \ddot{P}, \Delta_{clock}, \Delta_{DM}(L), \Delta_{\odot-\oplus}, \Delta_E, \Delta_S)$$

$P, \dot{P}, \ddot{P}$ period of pulsar' rotation and its derivatives: spin-down

Δ_{clock} difference in the local clock and terrestrial standrad

$\Delta_{DM}(L)$ delays caused by propagation in the interstellar medium

$\Delta_{\odot-\oplus}$ Transformation from the local frame to the solar system barycentre

Δ_E Accounts for relative motion (Doppler) + gravitational redshift caused by the Sun, plantes or binary companion.

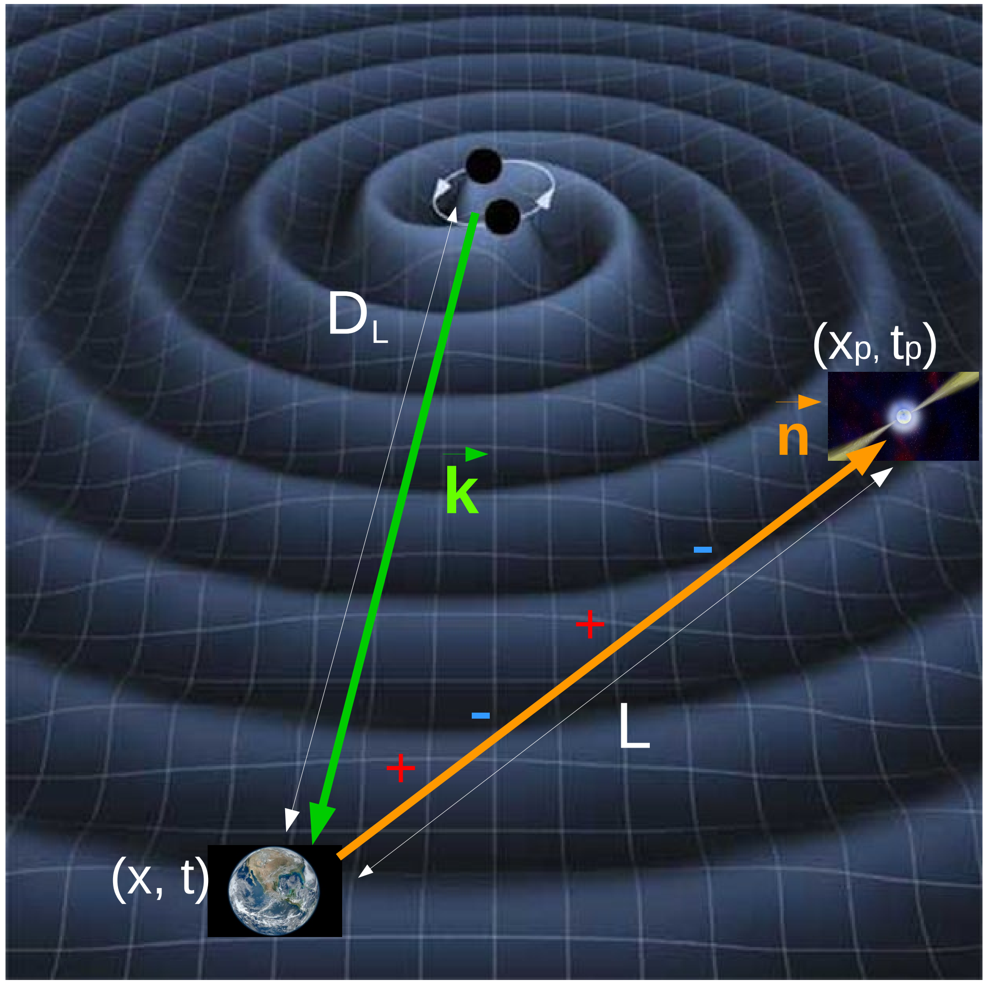
Δ_S Extra time required to trave in the curved spacetime containng Sun/companion (if in binary)

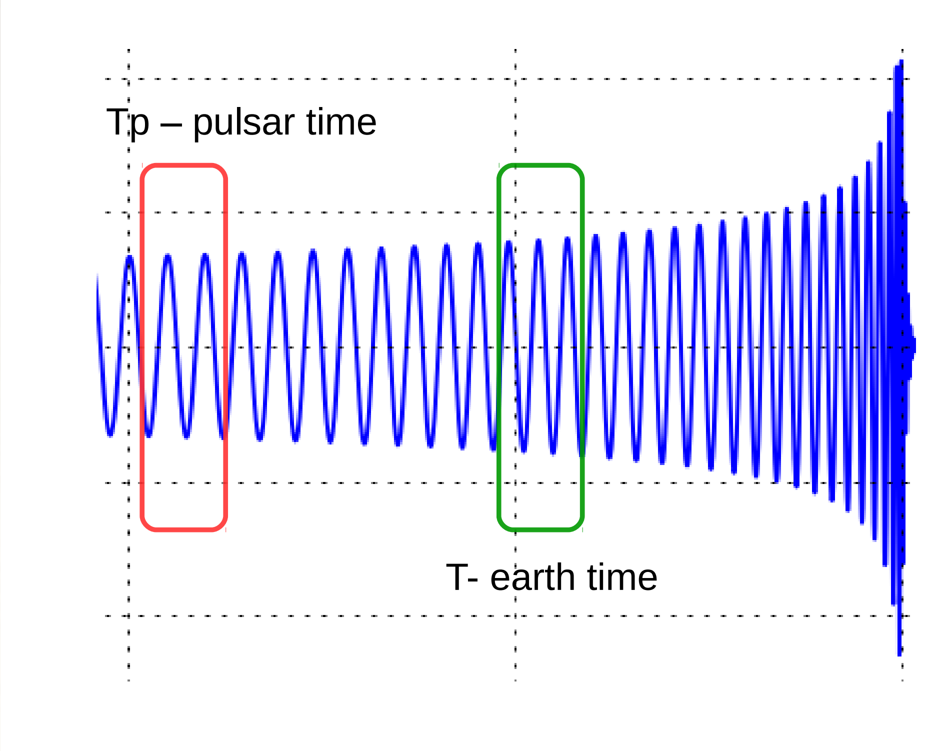
$$dt = t^p_{toa} - t^o_{toa} = dt_{errors} + \delta\tau_{GW} + noise$$

Errors in fitting the model → dt_{errors}

due to GWs → $\delta\tau_{GW}$

Response to GW signal





Important quantity which characterizes the response of any GW observatory is $\epsilon = (2\pi f_* L/c)$

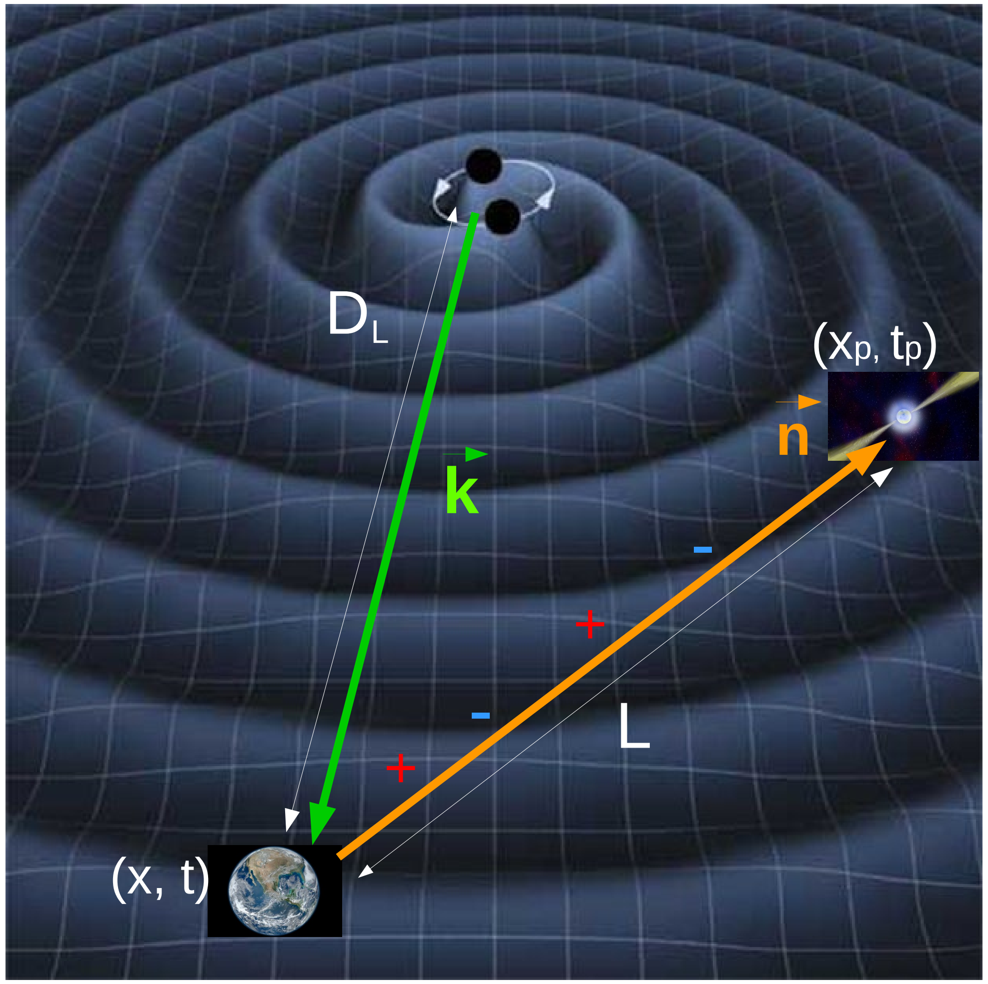
size of GW detector

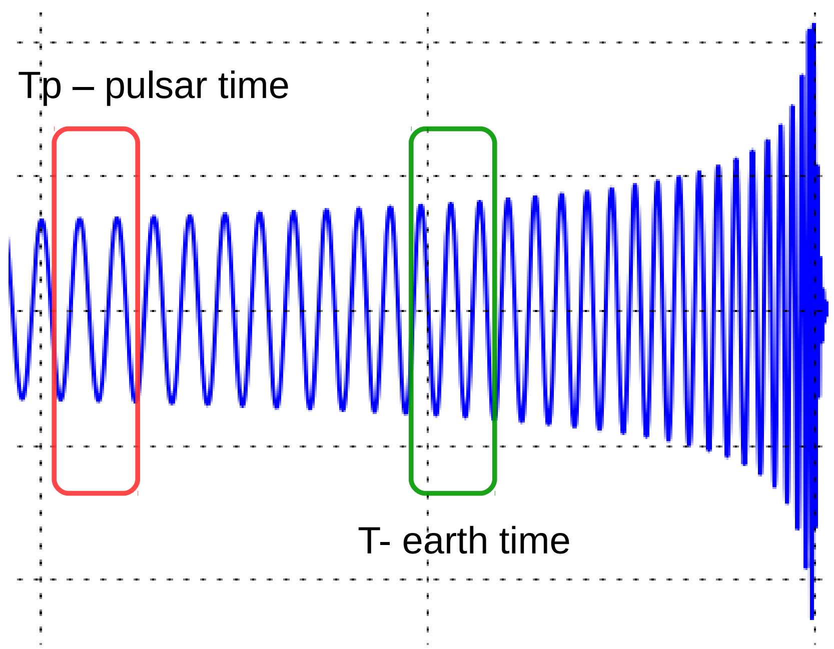
$$\epsilon \ll 1 \rightarrow \ R \propto h_{ij}n^i n^j$$

long wavelength approximation: already seen it with LIGO/Virgo

$\epsilon = 1 \quad \rightarrow$ LIGO: f* ~12 kHz, LISA: f*~ 0.05 Hz, PTA: f*~0.002 nHz

Response to GW signal





$$dt = t^p_{toa} - t^o_{toa} = dt_{errors} + \delta\tau_{GW} + noise$$

$$\delta\tau_{GW} = r(t) = \int_0^t \frac{\delta\nu}{\nu_0}(t')dt'; \quad \frac{\delta\nu}{\nu_0} = \frac{1}{2}\frac{\hat{n}^i\hat{n}^j\Delta h_{ij}}{1+\hat{n}.\hat{k}}$$

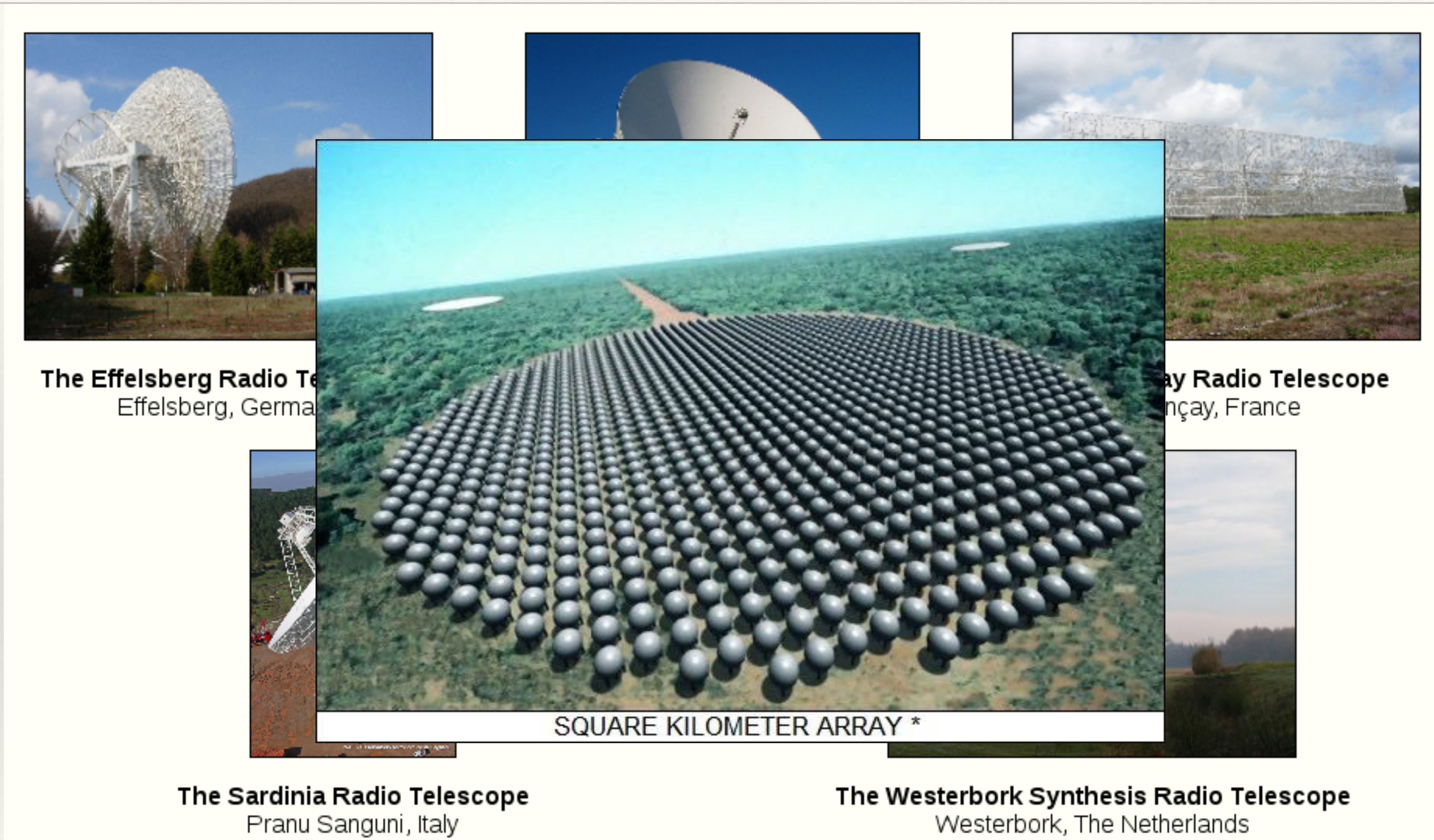
Familiar from LISA

$$\Delta h_{ij} = h_{ij}(t_p = t - L(1+\hat{n}.\hat{k})) - h_{ij}(t)$$

tp — pulsar time, ~ time of emission of the radio pulse:

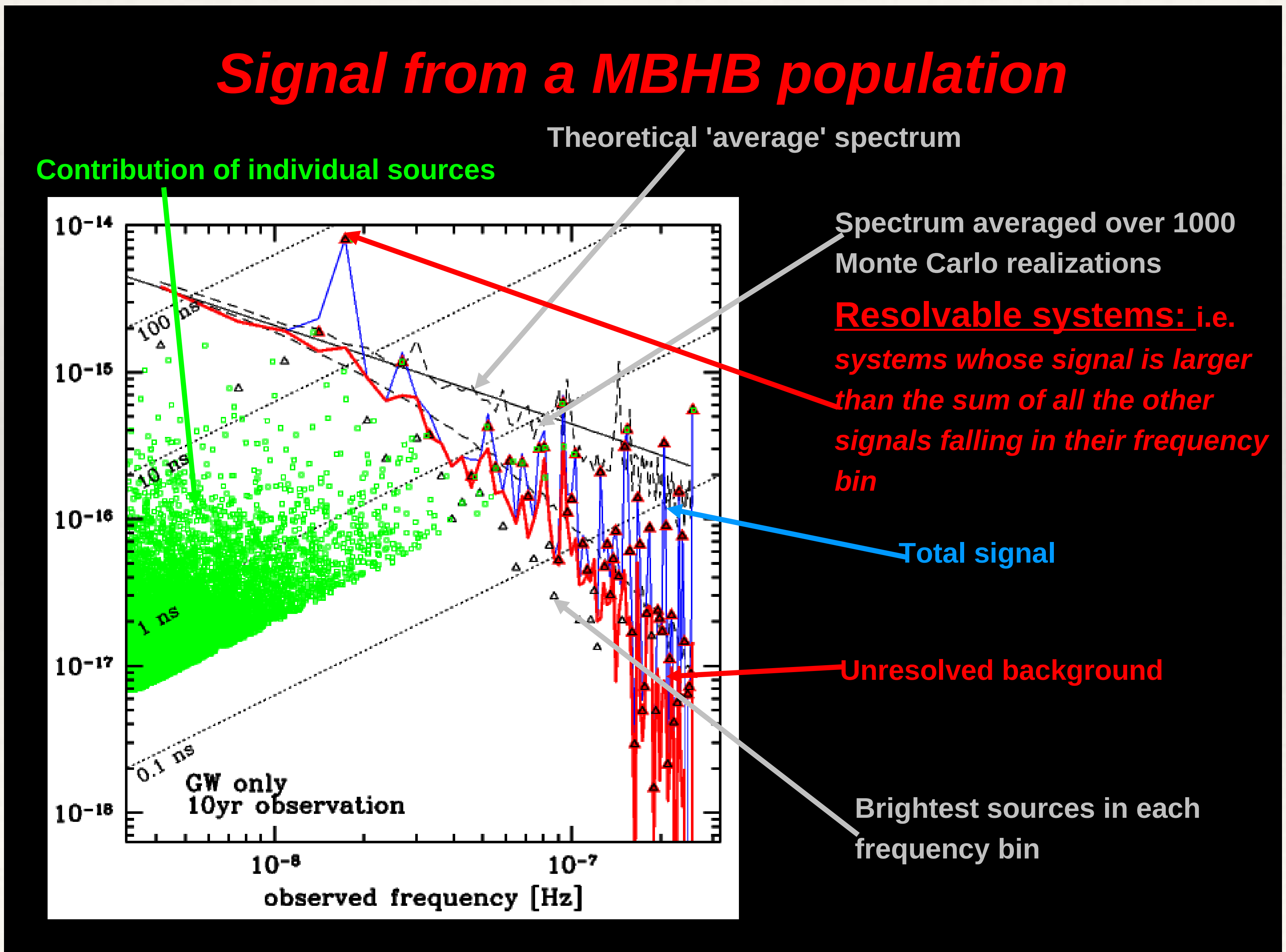
- depends on the relative position of a pulsar and GW source
- depends on the distance to the pulsar L
- L ~ few kpc ~10 000 years — “pulsar” term h(tp) contains info about the system 10^5 years in the past as compared to the “earth” term
- pulsar term depends on the pulsar.

Radiotelescops: EPTA



Supermassive black hole binaries

- Main sources are supermassive black hole binaries (mass 10^7 — 10^{10} solar) on very broad orbit (period ~ year(s))
- The orbital evolution due to GW emission is very slow: $\frac{dE}{dt} \propto \eta (M/r)^5$ signal is (almost) monochromatic over period of observations

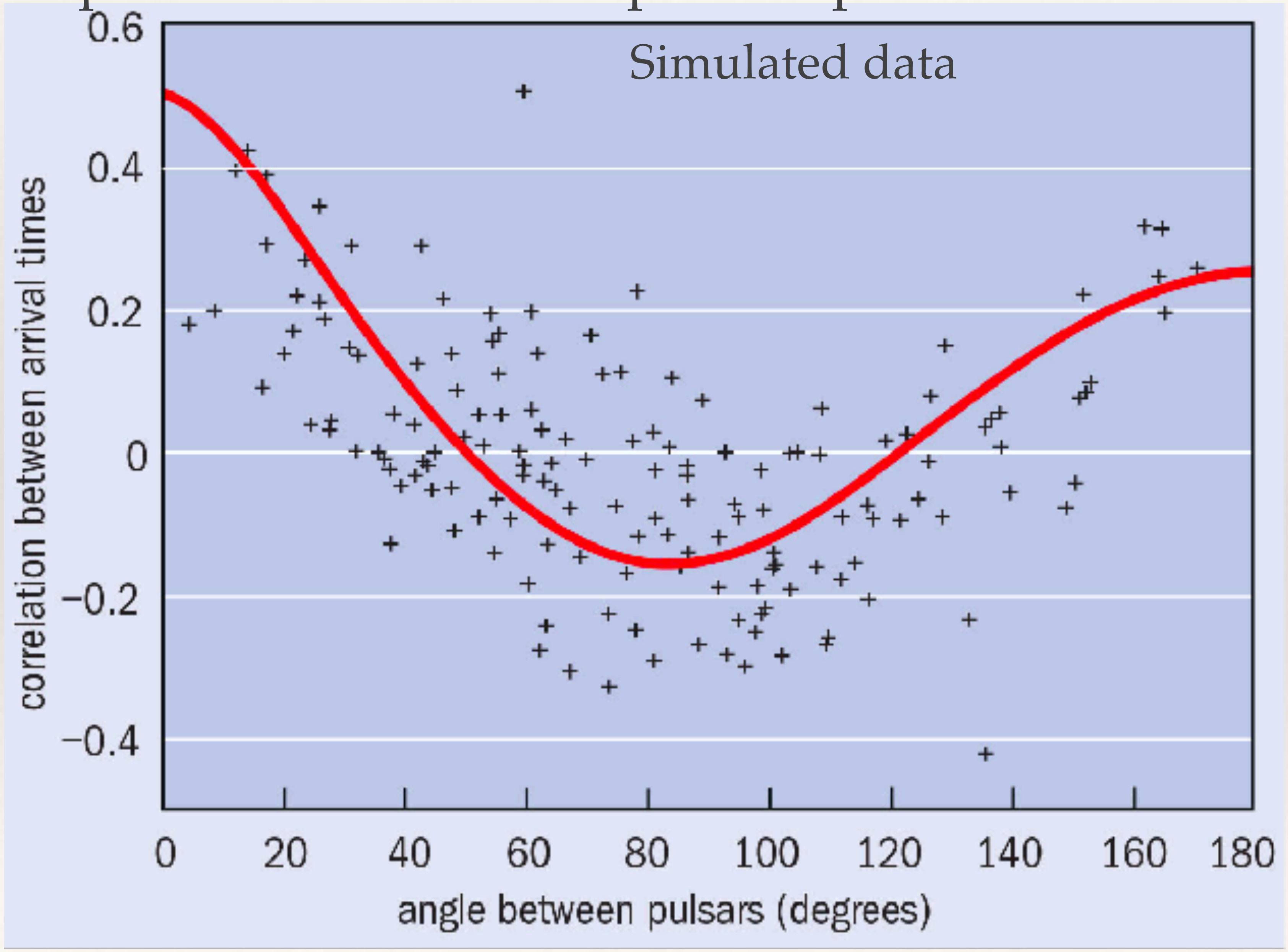


GW signal from the population of SMBH binaries: forms a stochastic signal at low freqs. (similar to Galactic binaries in LISA

[Credits: A. Sesana]

Correlation

- Stochastic GW signal — noise like signal which is correlated in observation of a pulsars. The correlation due to GW is very specific: Hellings-Downs curve.
- Correlation for the isotropic stochastic GW signal depends only on the angular separation between the pairs of pulsars.



[Fig. from IOP, Physics World]

Data analysis: likelihood function

- The likelihood function (likelihood of the signal with given parameters is present in the observed data) is given as

$$P(\vec{\delta t}|\vec{\theta}) = \frac{1}{\sqrt{(2\pi)^n det(C)}} \exp\left[-\frac{1}{2}(\vec{\delta t} - \vec{s})^T C^{-1} (\vec{\delta t} - \vec{s})\right]$$

data (TOA residuals) unevenly sampled

determenistic signal

Noise variance-covariance matrix

- We pack ***all*** observed data (from all telescopes and all pulsars) into a single array of size n

- The GW signal (template) is located in "s" vector (resolvable binaries) and in noise matrix C

greek: pulsar index,
latin: data index

$$C_{\alpha i,\beta j} = C^{wn}\delta_{\alpha\beta}\delta_{ij} + C^{rn}_{ij}\delta_{\alpha\beta} + C^{dm}_{ij}\delta_{\alpha\beta} + C^{GW}_{\alpha i,\beta j}$$

white noise

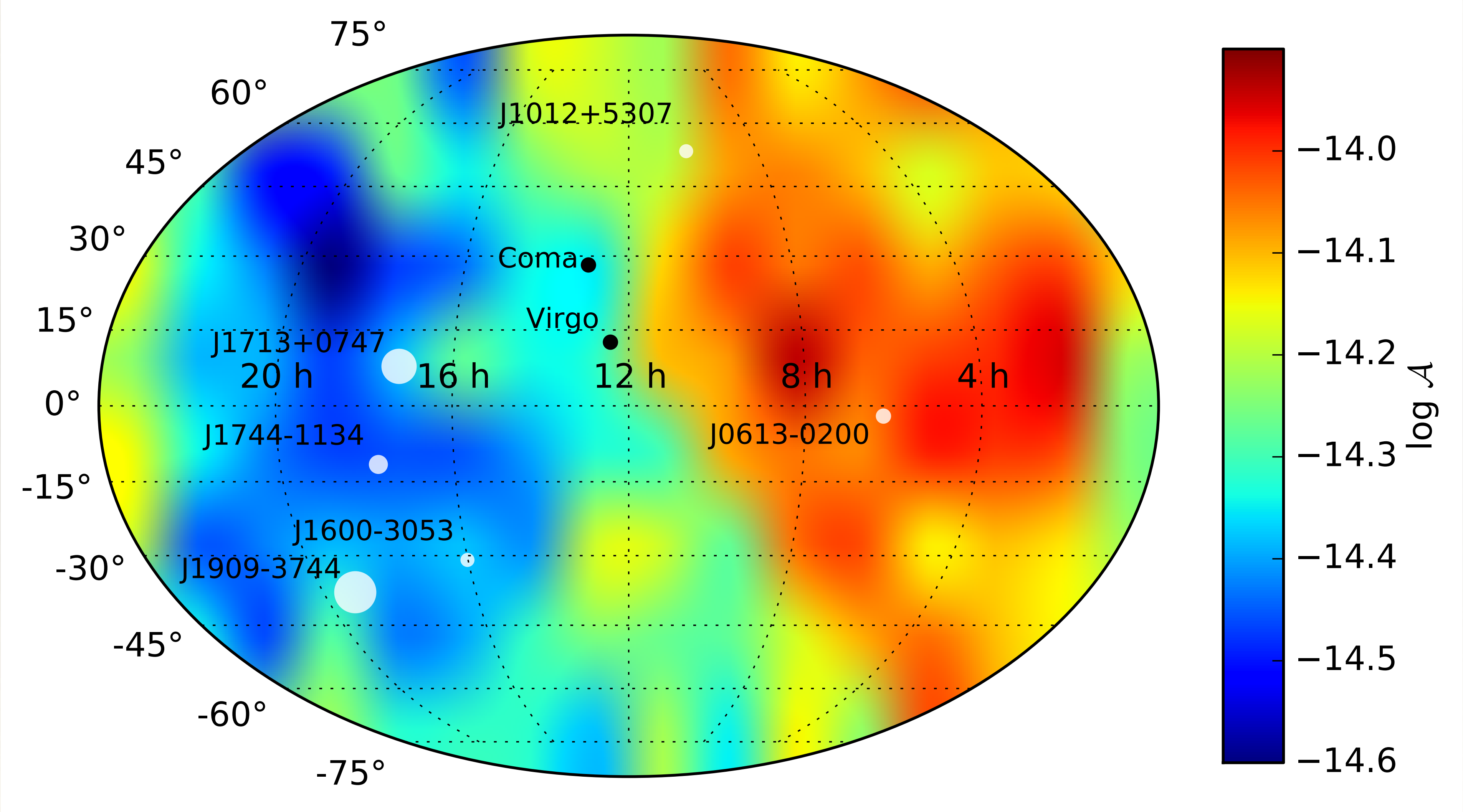
red noise

dispersion variation

stochastic GW signal

Upper limit on GW in the PTA band

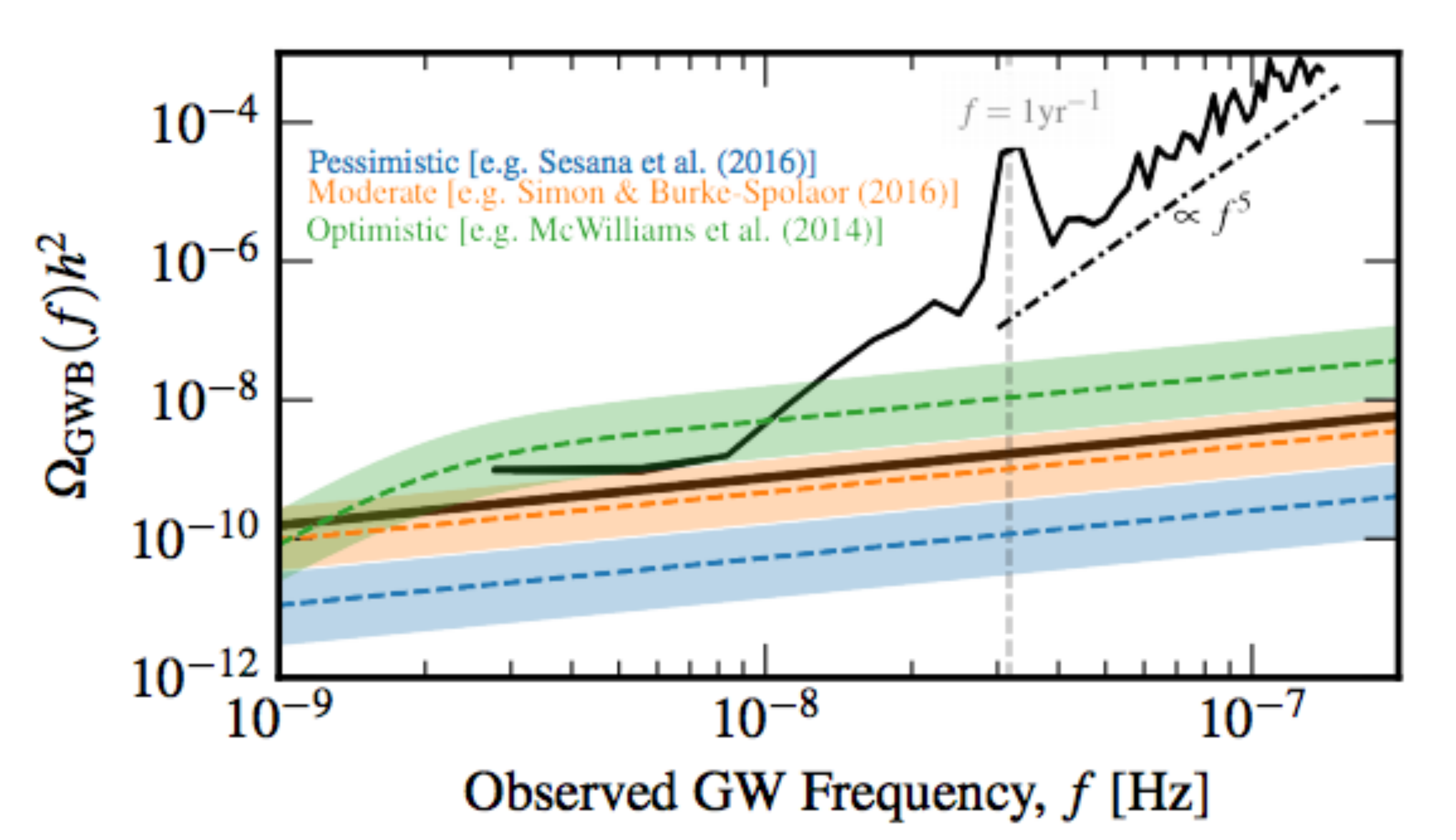
- GW not yet detected with PTA: GW are weak and need to be integrated out the noise (decades of observation time). We also need more good “timers”.
- Give a current sensitivity we can set upper limit on GW strain: for individua binaries and for the stochastic GW signal.



[Babak+ MNRAS (2015), EPTA]

Upper limit on GW in the PTA

Current observational data allows us close some (over)optimistic astrophysical models and constrain some parameters



[Nanograv, arXiv: 1801.02617]