Neural Posterior Estimation with guaranteed exact coverage: the ringdown of GW150914

Marco Crisostomi

Università di Pisa

XXV SIGRAV Conference
07/09/2023

Based on arXiv:2305.18528

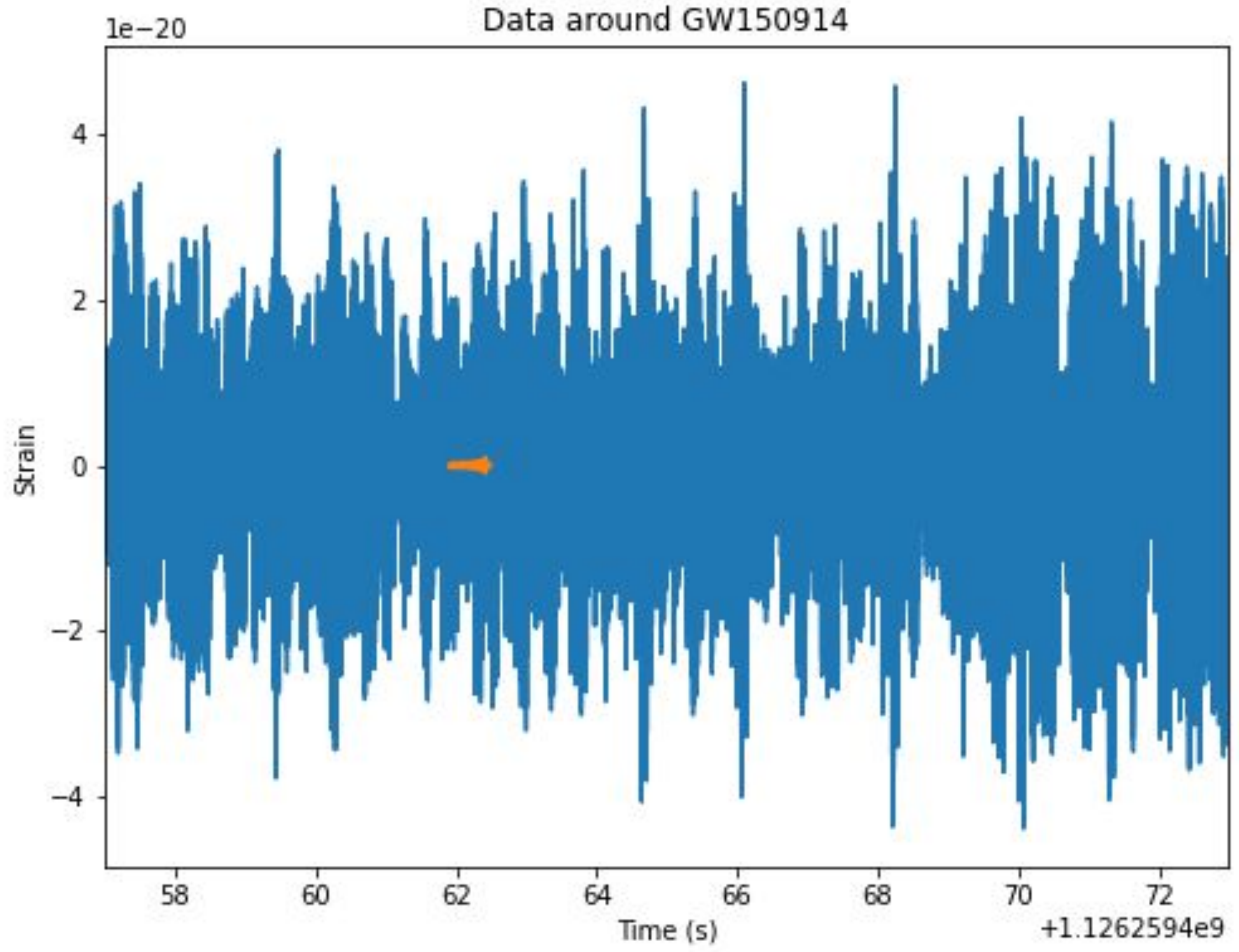
In collaboration with Kallol Dey, Enrico Barausse and Roberto Trotta

Outline

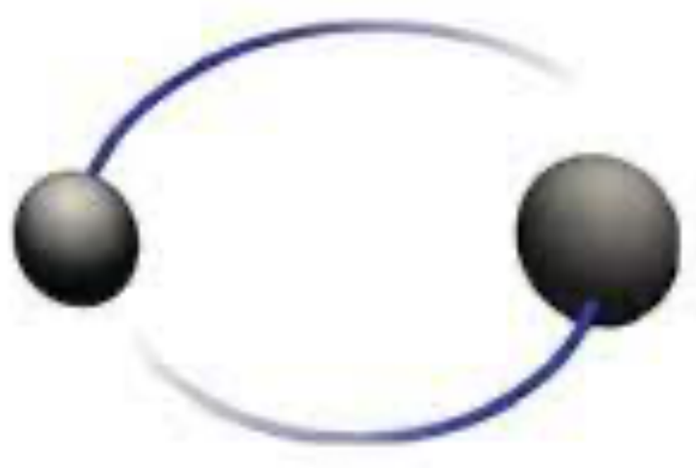
- Introduction to parameter estimation for GWs
- Simulation based inference (SBI) and calibration
- Ringdown parameter estimation (GW150914)
- SBI applied to the ringdown of GW150914
- Conclusions

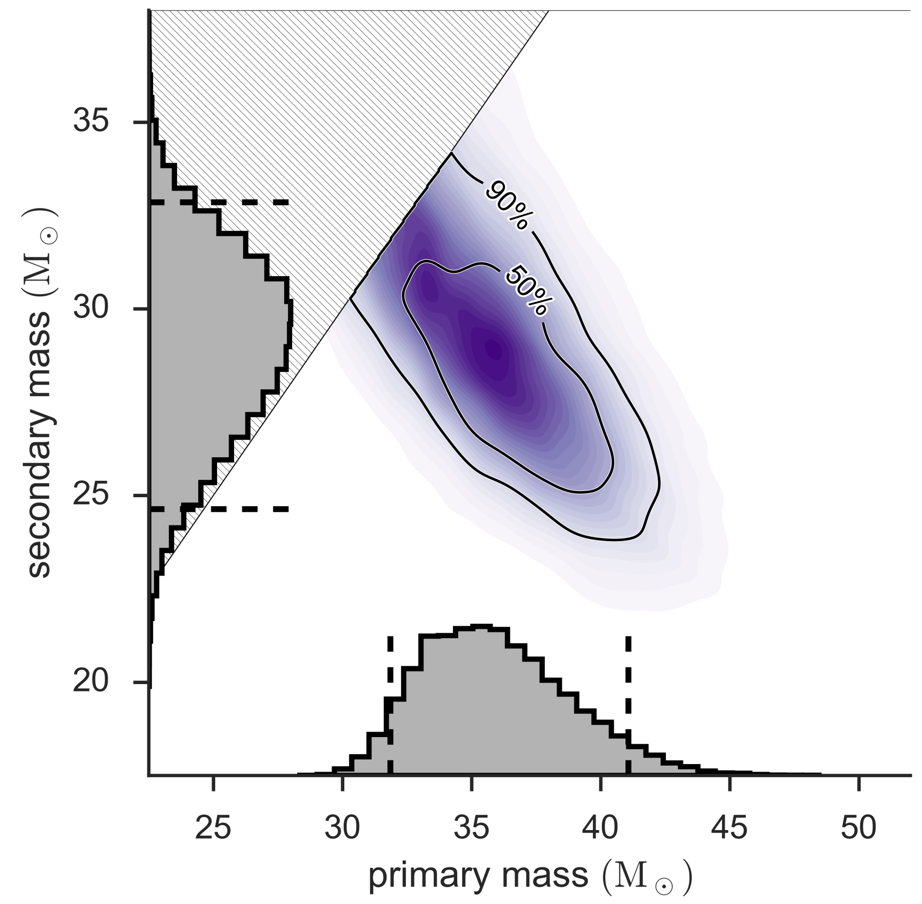
Introduction to PE for GWs

Extracting physics from data!



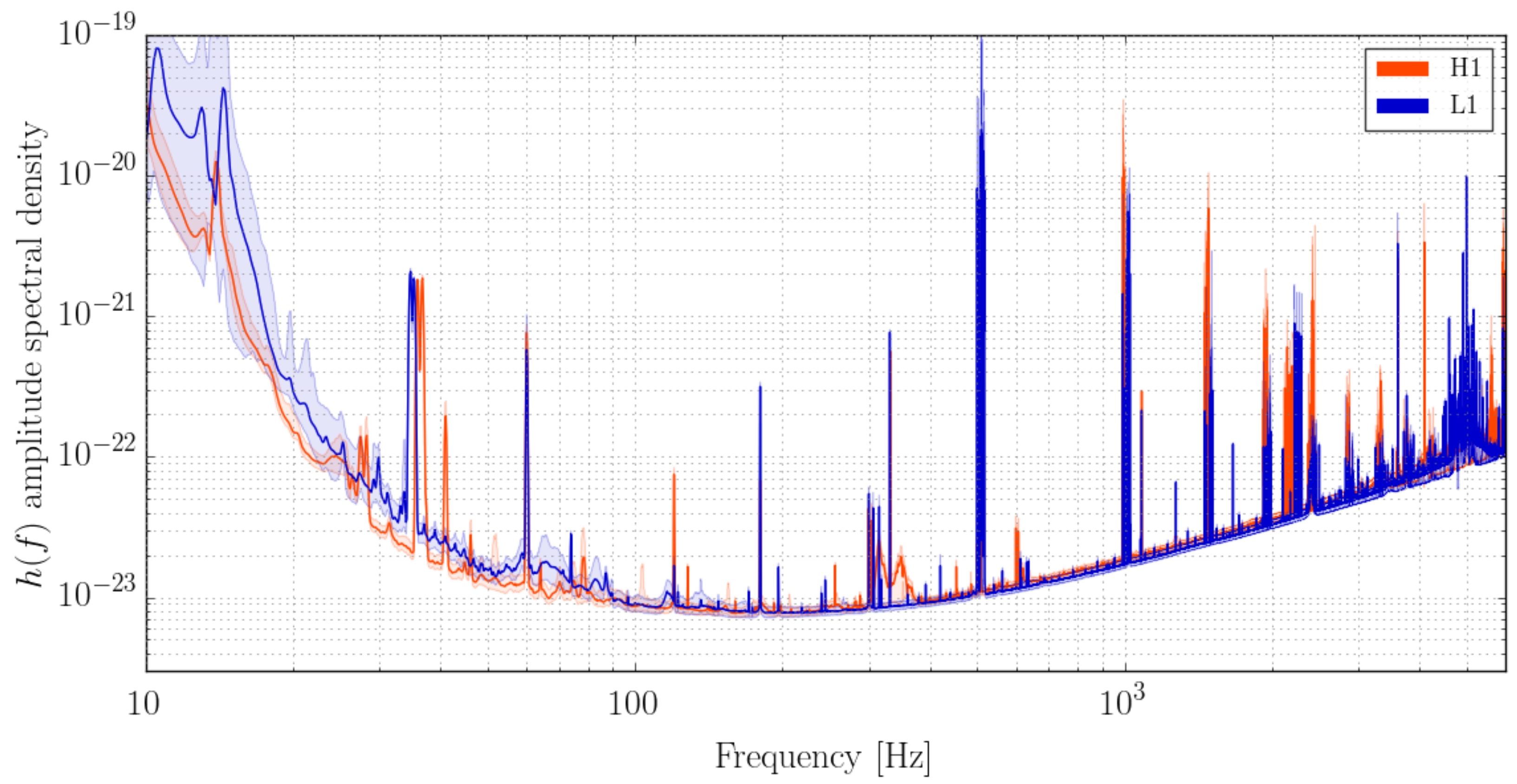
Bayesian inference



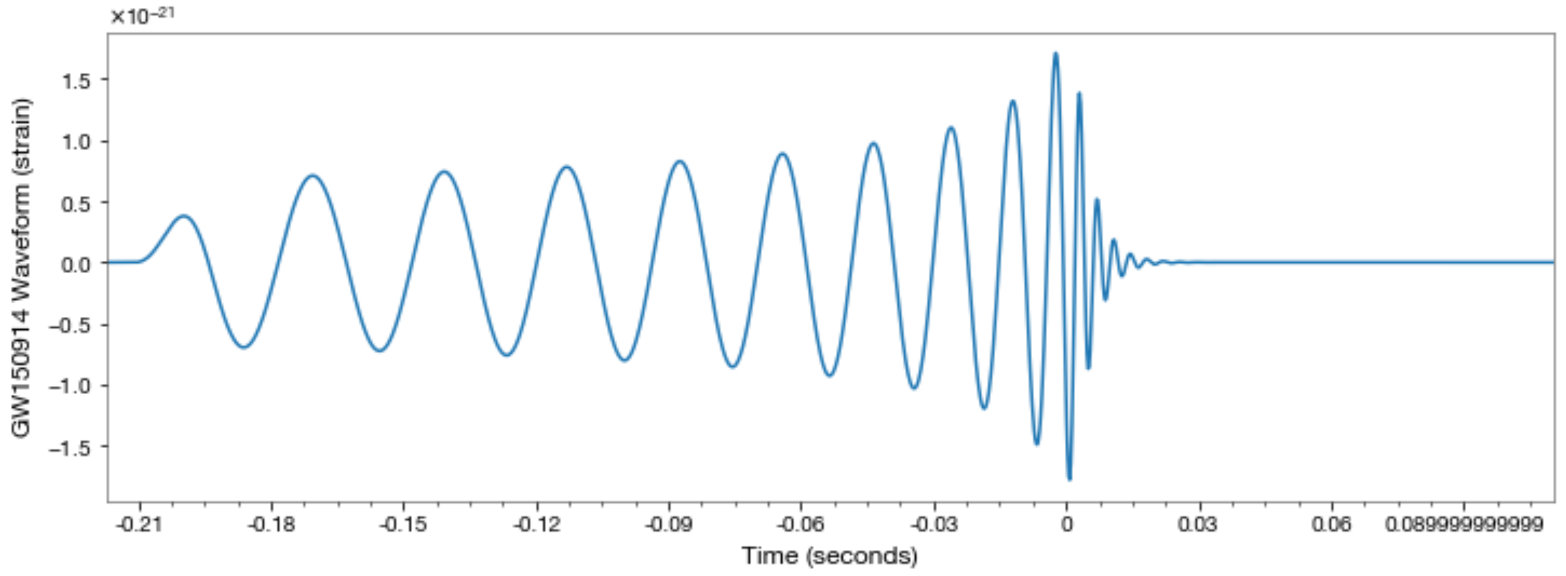


Modelling

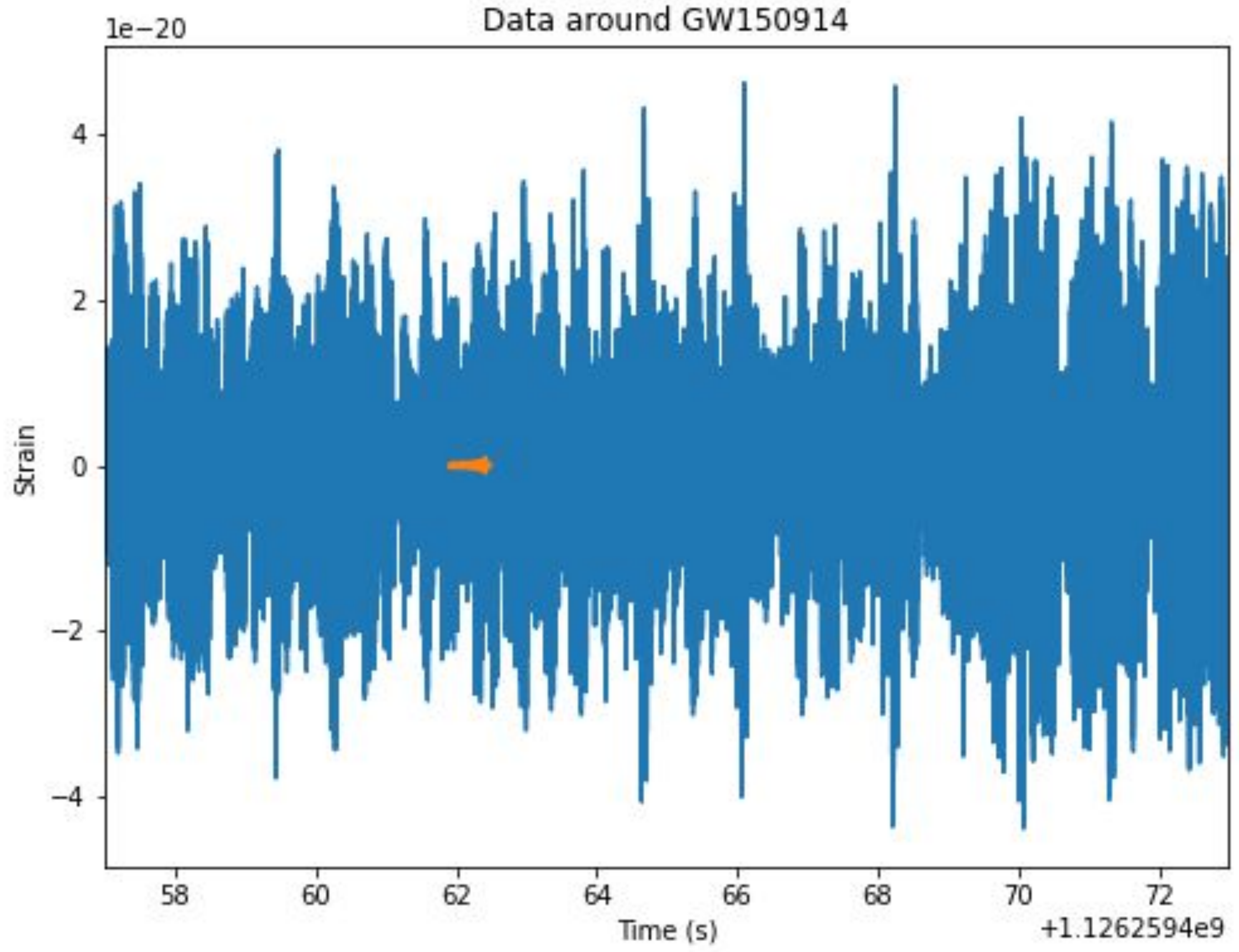
Noise



Physics



Introduction to PE for GWs



$$s(t) = h(t) + n(t)$$

Strain Signal Noise

$$\Lambda(s|\theta_t) = \mathcal{N} \exp\left\{-\frac{1}{2}(s - h(\theta_t)|s - h(\theta_t))\right\}$$

Likelihood

$$(A|B) = \mathrm{Re} \int_{-\infty}^{\infty} df \, \frac{\tilde{A}^*(f)\tilde{B}(f)}{(1/2)S_n(f)}$$

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \delta(f - f')\frac{1}{2}S_n(f)$$

Bayes th. $$p(\theta|s) \propto \Lambda(s|\theta)\pi(\theta)$$

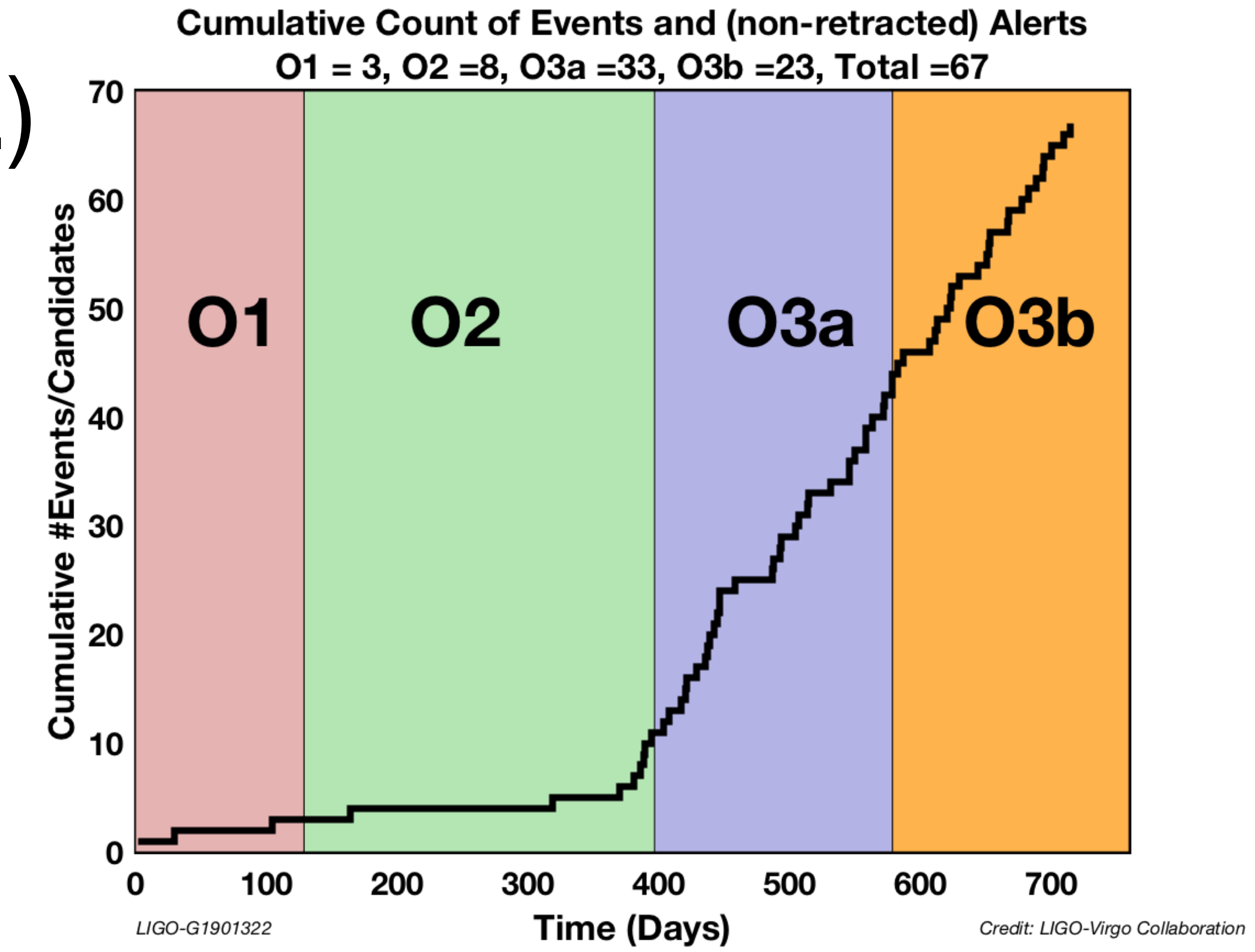
Posterior Prior

Introduction to PE for GWs

Likelihood-based inference

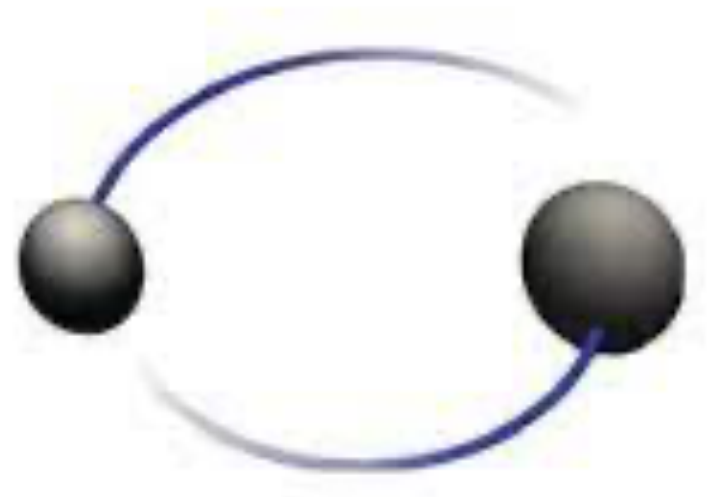
$$p(\theta|s) \propto \Lambda(s|\theta)\pi(\theta)$$

- Use of stochastic samplers as MCMC or Nested Sampling
- Many evaluations of the likelihood
- Drawbacks:
  - Noise is not stationary and gaussian (+glitches, gaps, etc.)
  - Very slow!

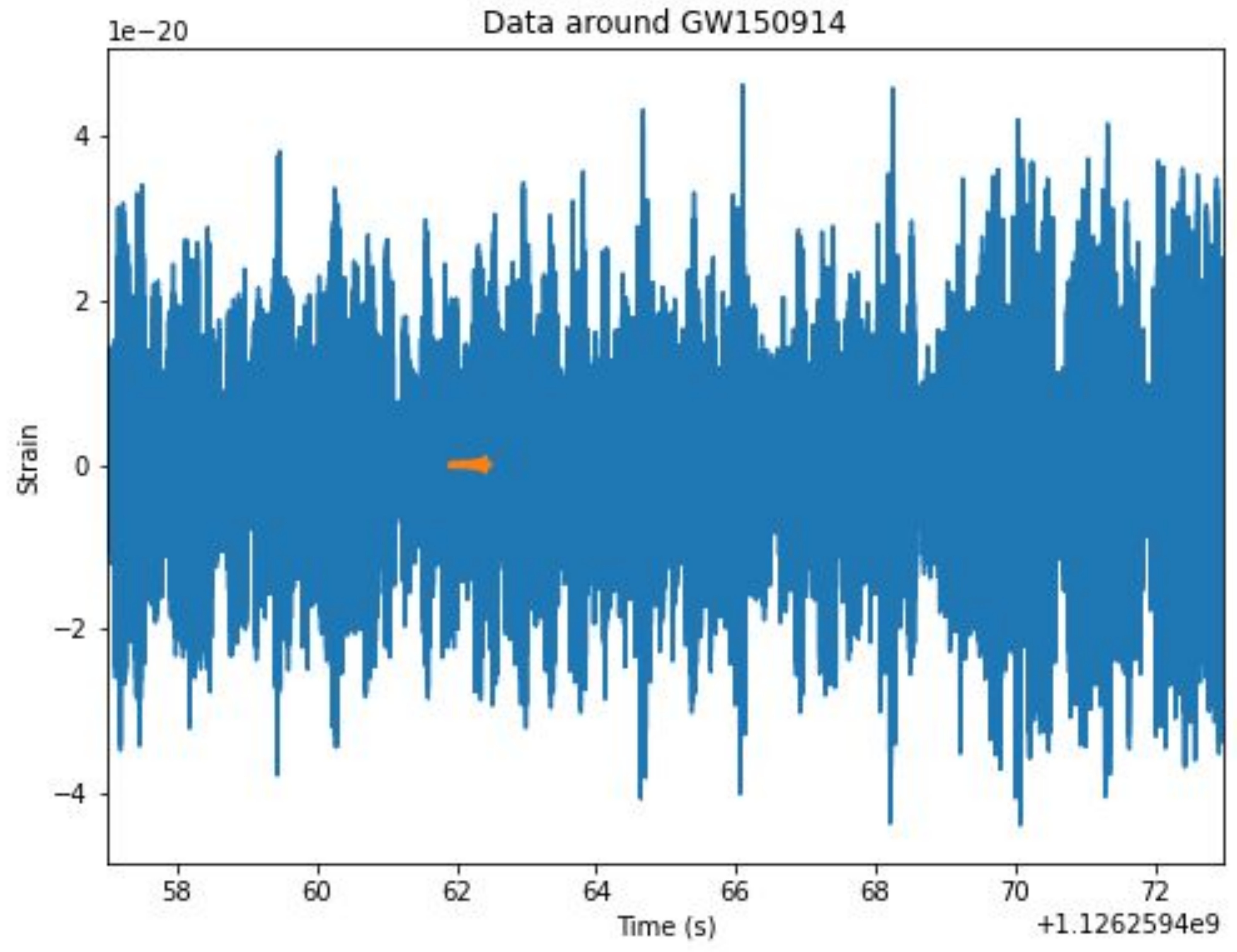


Simulation-based inference

Likelihood-free inference



Forward-simulator



Based on deep learning techniques

Once the neural network has been trained —> sampling is very fast!

Simulation-based inference

Neural density estimation

$$\{\boldsymbol{\theta}_n, \boldsymbol{x}_n\} \qquad n = 1, \ldots, N$$

Neural Network

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto q_\phi(\boldsymbol{\theta}|\boldsymbol{x}) \qquad \phi = F(\boldsymbol{x}, \psi)$$

Conditional density

$$\mathcal{L}(\psi) = -\sum_{n=1}^{N} \log \frac{q_{\phi(\boldsymbol{x}_n,\psi)}(\boldsymbol{\theta}_n)/\pi(\boldsymbol{\theta}_n)}{\sum_{m=1}^{M} q_{\phi(\boldsymbol{x}_n,\psi)}(\boldsymbol{\theta}_m)/\pi(\boldsymbol{\theta}_m)}$$

Cross-entropy loss

$$\boldsymbol{\theta}_n \in \{\boldsymbol{\theta}_m\}$$

Simulation-based inference

<u>Neural density estimation</u> <u>with Normalising Flows</u>

$$\boldsymbol{\theta} = f(\boldsymbol{\vartheta})$$

- Invertible
- Tractable Jacobian

$$q(\boldsymbol{\vartheta}) \longrightarrow p(\boldsymbol{\theta})$$

Base distribution

Target distribution

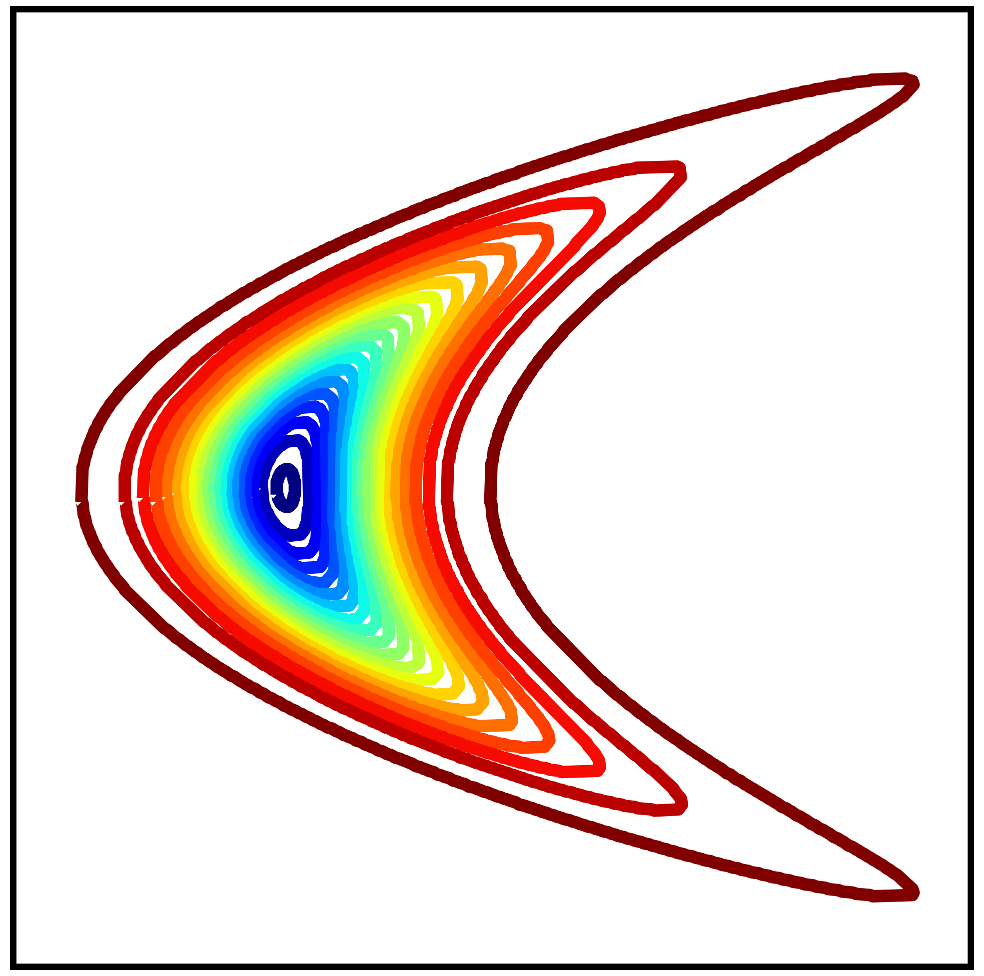
$$p(\boldsymbol{\theta}) = q\left(f^{-1}(\boldsymbol{\theta})\right) \left|\det\left(\frac{\partial f^{-1}}{\partial \boldsymbol{\theta}}\right)\right|$$

Simulation-based inference

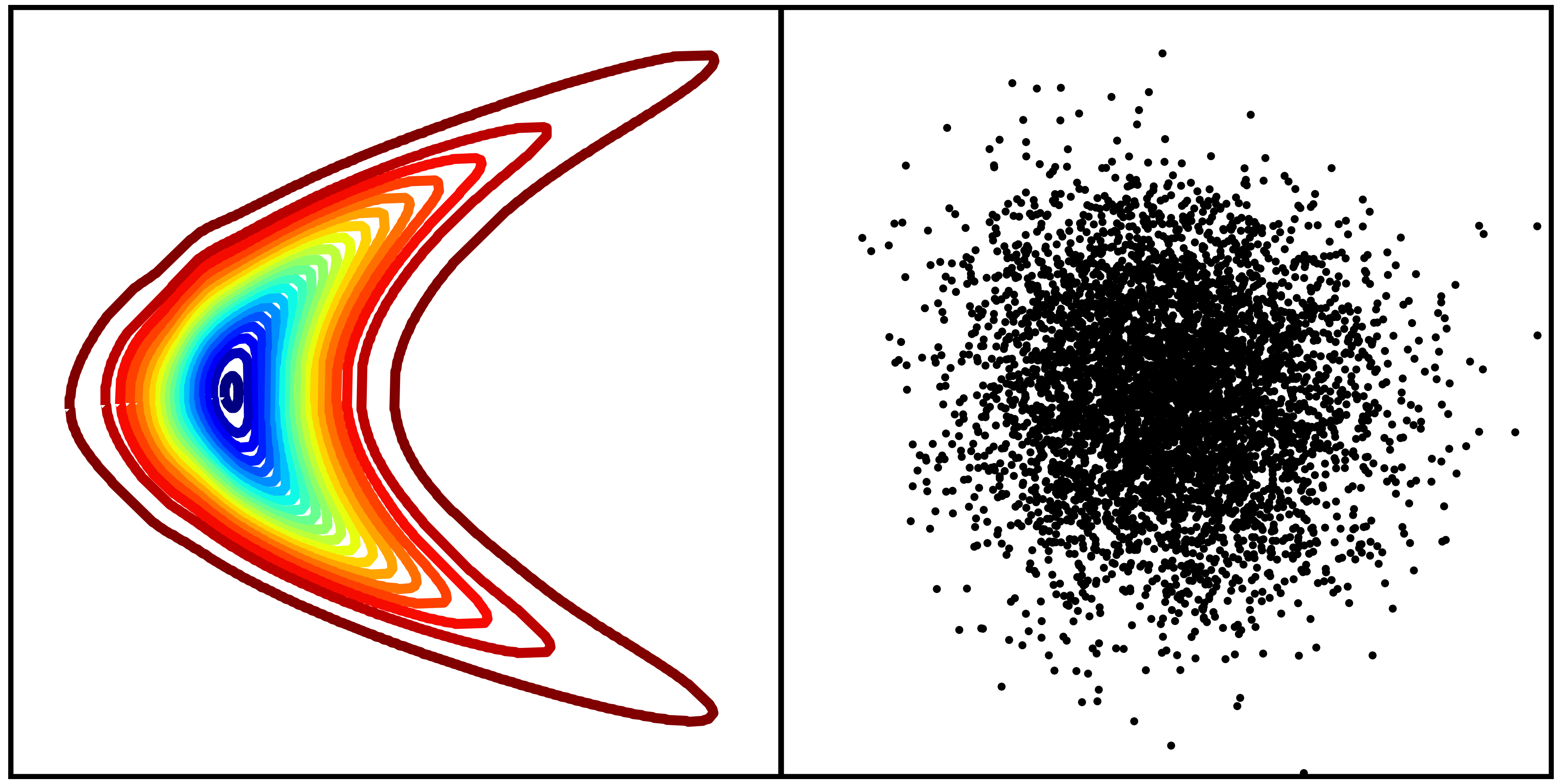
Neural density estimation with Normalising Flow

Masked Autoregressive Flow Papamakarios et al. (2018)

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i q_{\phi_i}(\theta_i|\boldsymbol{\theta}_{1:i-1}, \boldsymbol{x})$$



(a) Target density



MAF with 5 layers

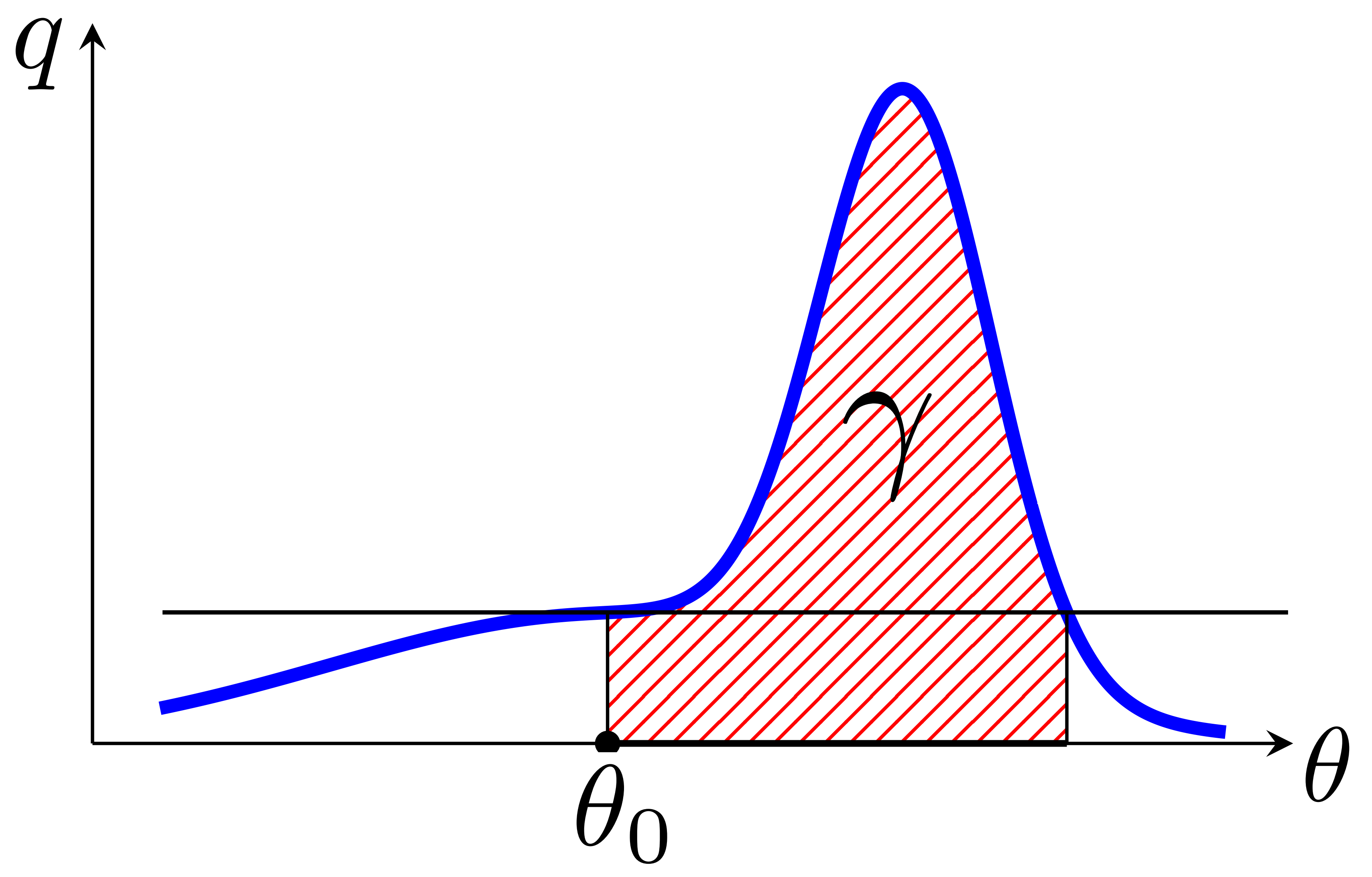
Validation and Calibration

How can we trust the posteriors to be right?

<u>Bayesian P-P Plot</u>

$$\gamma(\boldsymbol{\theta}_0, \boldsymbol{x}) \equiv \int_{q(\boldsymbol{\theta}|\boldsymbol{x}) > q(\boldsymbol{\theta}_0|\boldsymbol{x})} q(\boldsymbol{\theta}|\boldsymbol{x})\, \mathrm{d}\boldsymbol{\theta}$$

Credibility



$$\boldsymbol{\theta}_0 \sim \pi(\boldsymbol{\theta}) \qquad \gamma(\boldsymbol{\theta}_0, \boldsymbol{x}_0) \qquad \mathcal{C}(\gamma) \equiv \int_0^{\gamma} \mathcal{P}(\gamma')\mathrm{d}\gamma'$$

Empirical coverage

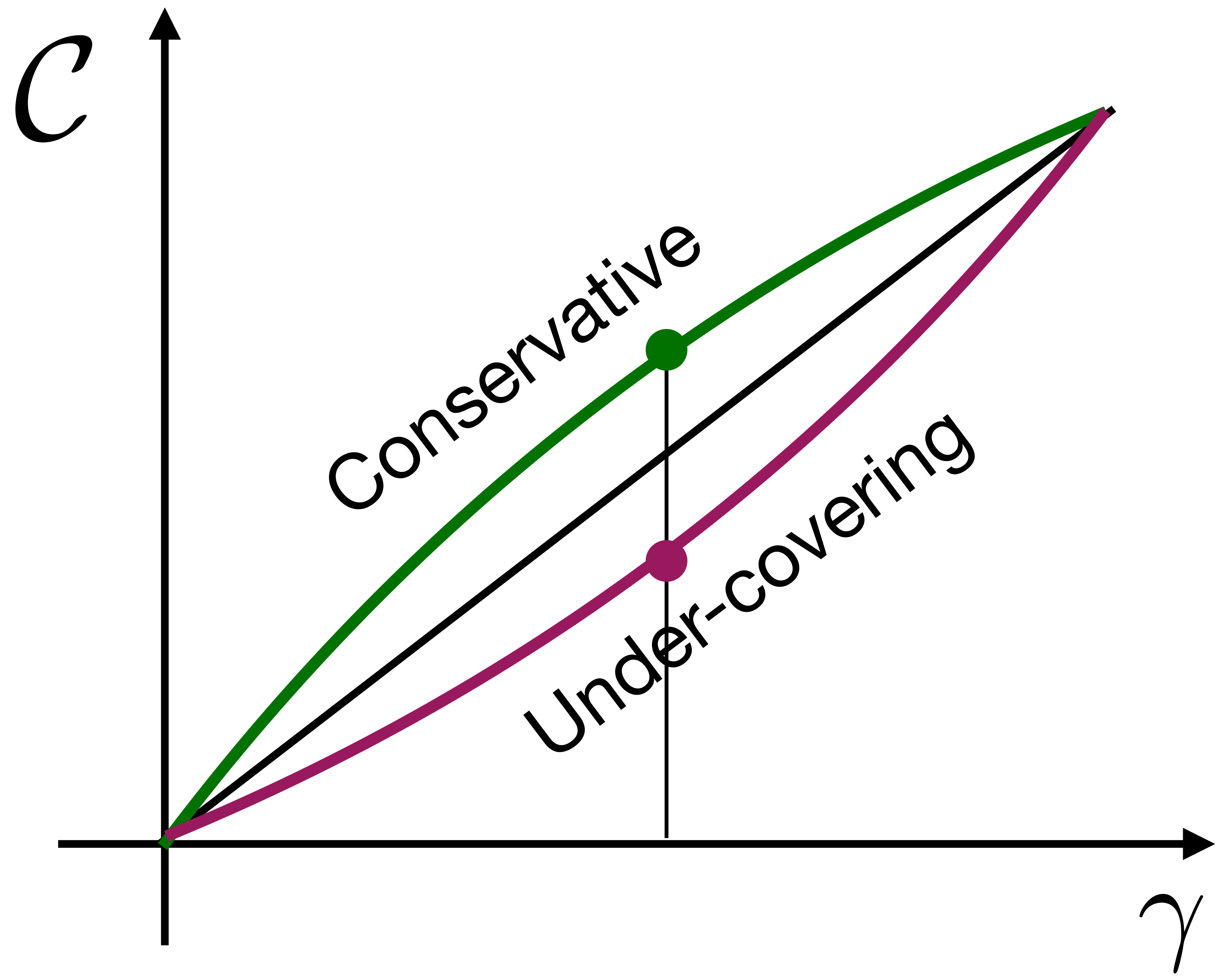
Validation and Calibration

How can we trust the posteriors to be right?

<u>Bayesian P-P Plot</u>

$\mathcal{C}(\gamma) \longrightarrow$ Empirical coverage

the frequency with which regions of different credibility include (cover) $\boldsymbol{\theta}_0$



Averaging across the prior space!

Validation and Calibration

How can we trust the posteriors to be right?

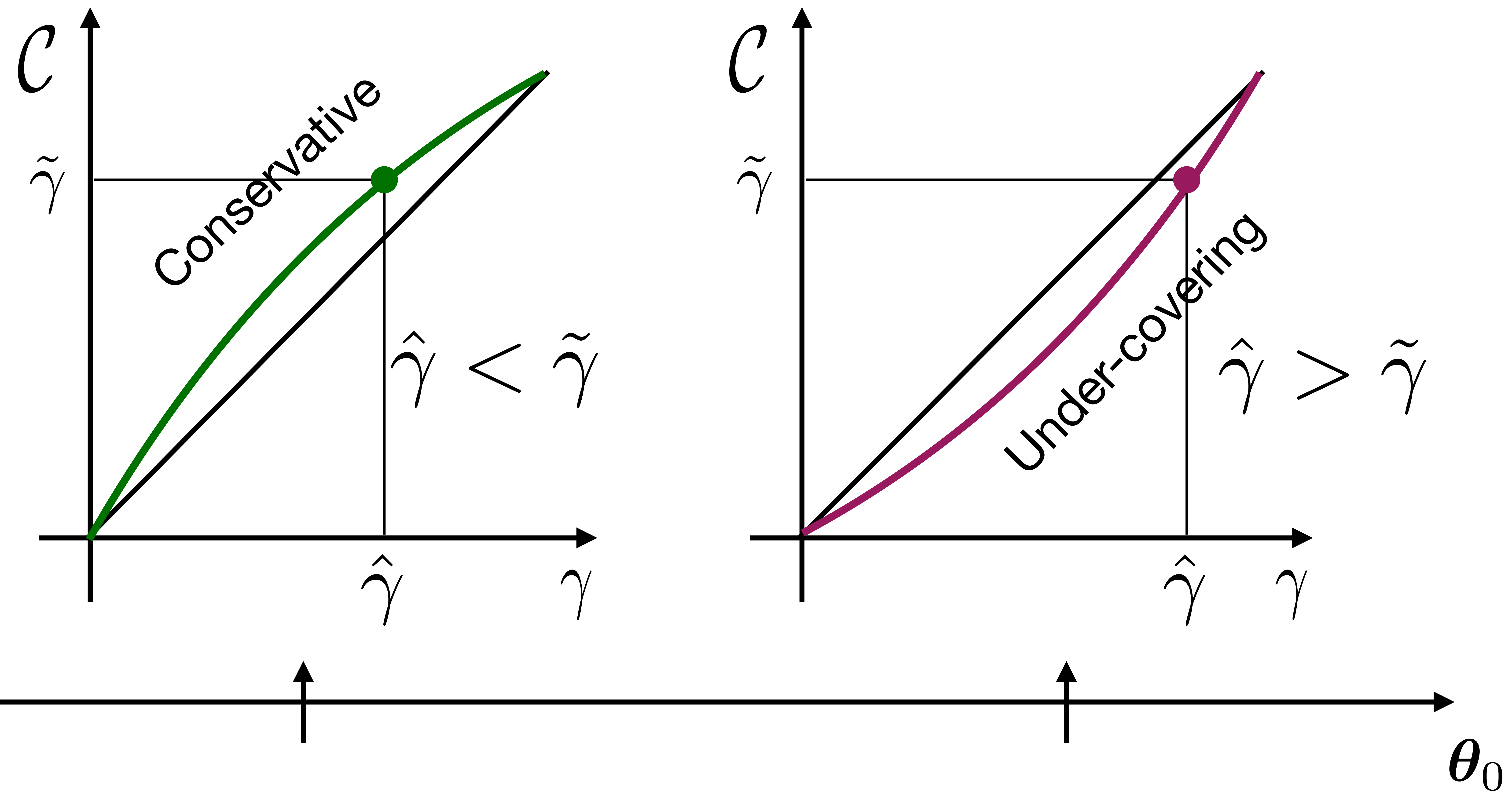
Frequentist P-P Plot

$$\hat{\gamma}$$

$$\downarrow$$

Required credibility

$$\mathcal{C}(\hat{\gamma}) = \tilde{\gamma}$$

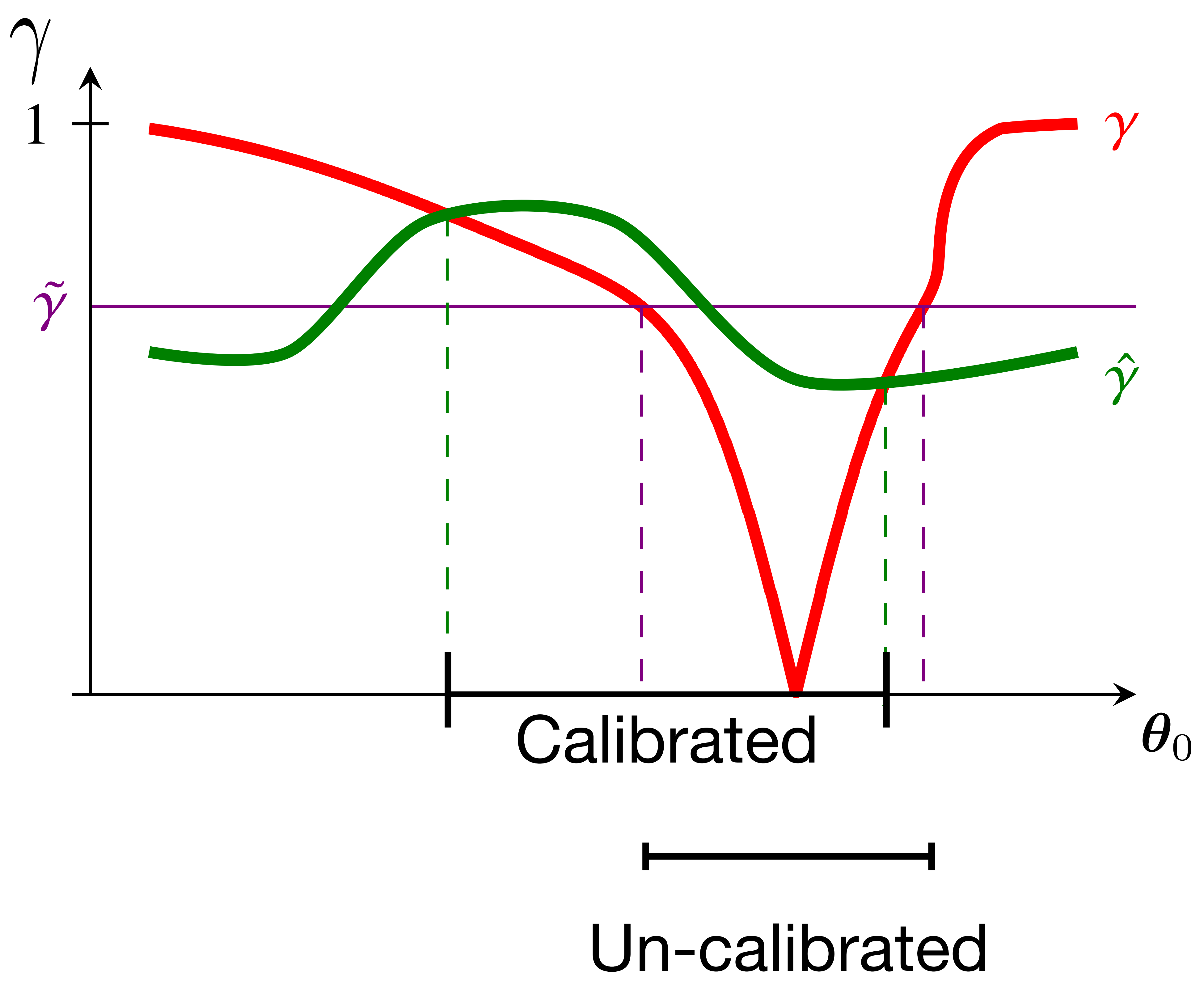


Make a map of $\hat{\gamma}$

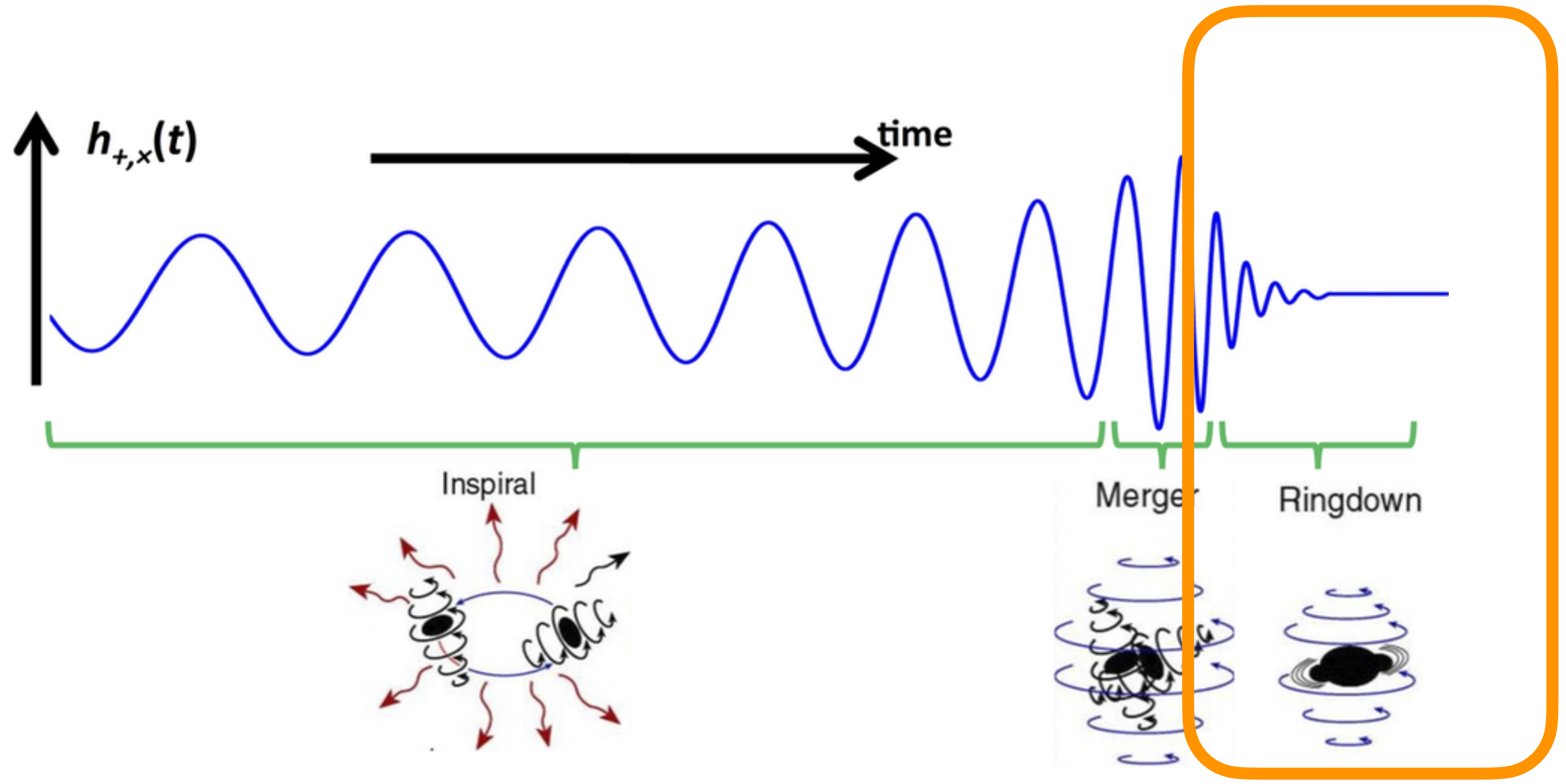
Validation and Calibration

From this map we can construct
frequentist confidence regions with guaranteed exact coverage!

For a given observation $\mathbf{x}$ we include in the region the parameters θ_0 for which $\gamma(\boldsymbol{\theta}_0, \boldsymbol{x}) < \hat{\gamma}(\boldsymbol{\theta}_0, \tilde{\gamma})$



Ringdown PE



$$h(t) = F_+ h_+ + F_\times h_\times$$

$$h_+ - i h_\times = \sum_{\ell m} h_{\ell m}(t)\, {}_{-2}Y_{\ell m}(\iota, \phi)$$

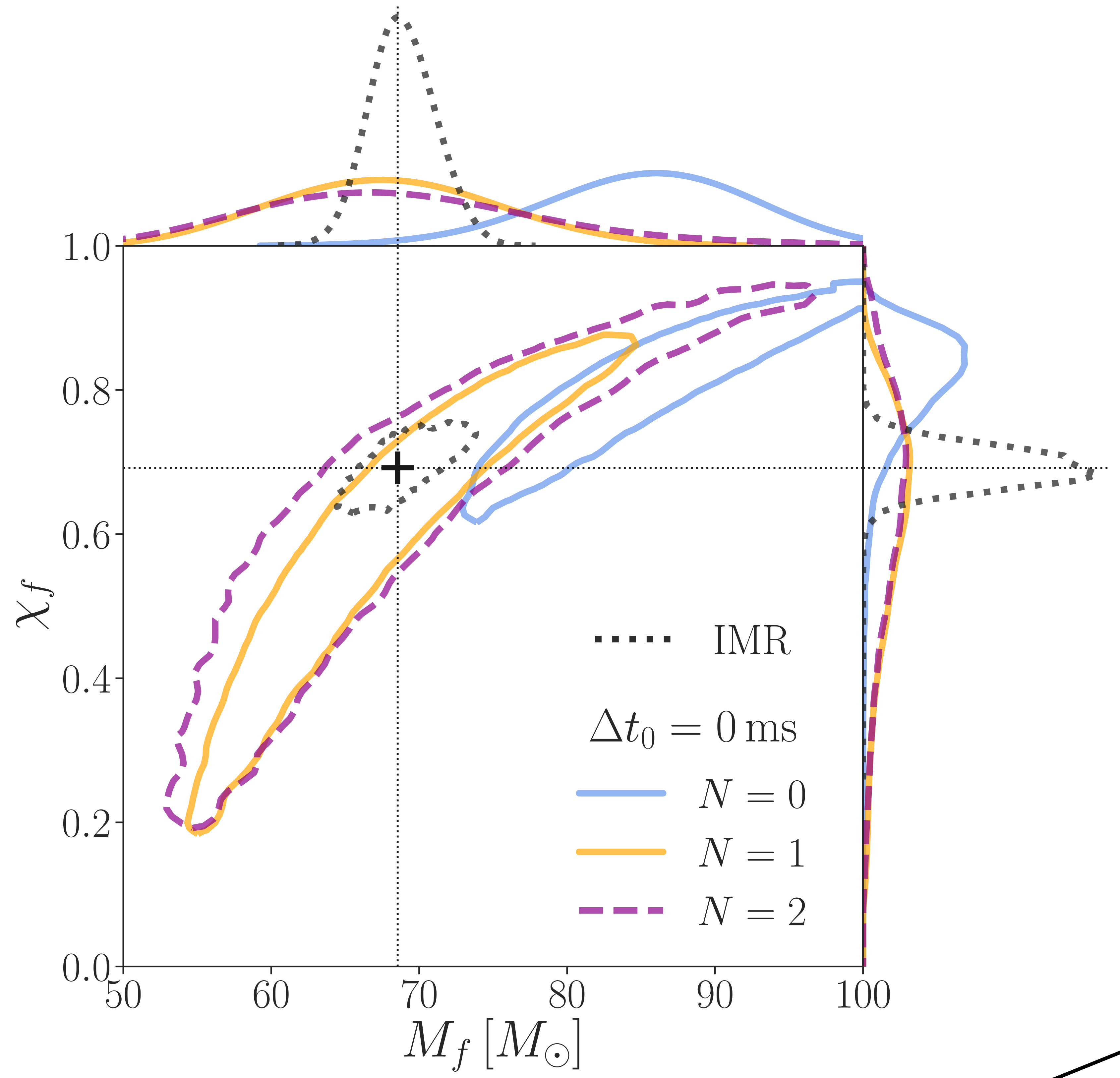
$$h_{\ell m}(t) = \sum_{n=0}^{\infty} \underbrace{A_{\ell m n} e^{i\,\phi_{\ell m n}}}_{\text{Estimated from data}} {}^{-\underbrace{\left(i\,\omega_{\ell m n} + \tau_{\ell m n}^{-1}\right)}_{\text{Kerr spectrum}}(t - t_0)}$$

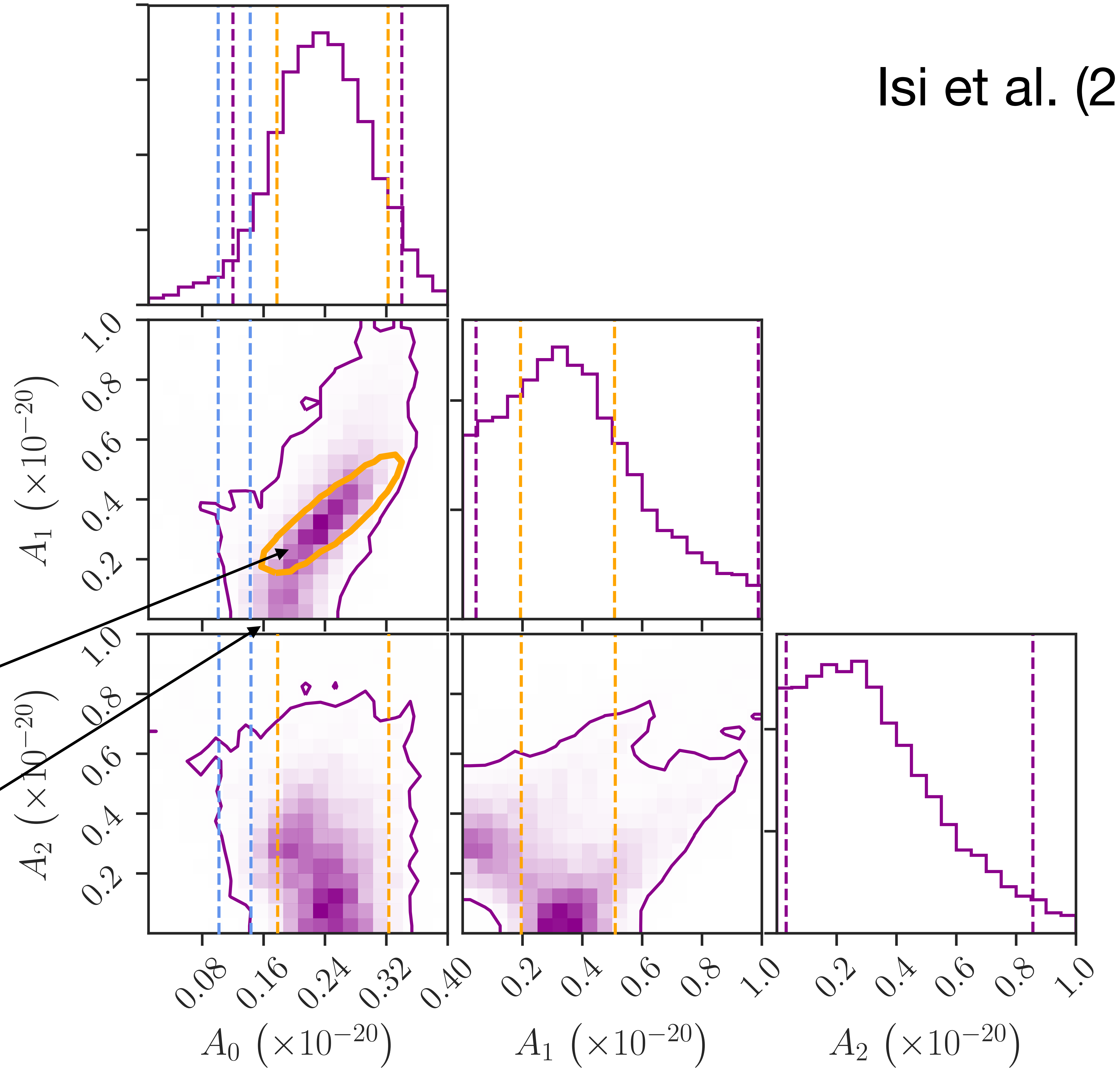
Kerr spectrum depends only on mass and spin (No-hair theorem) — perturbation theory

Estimated from data

Ringdown PE for GW150914

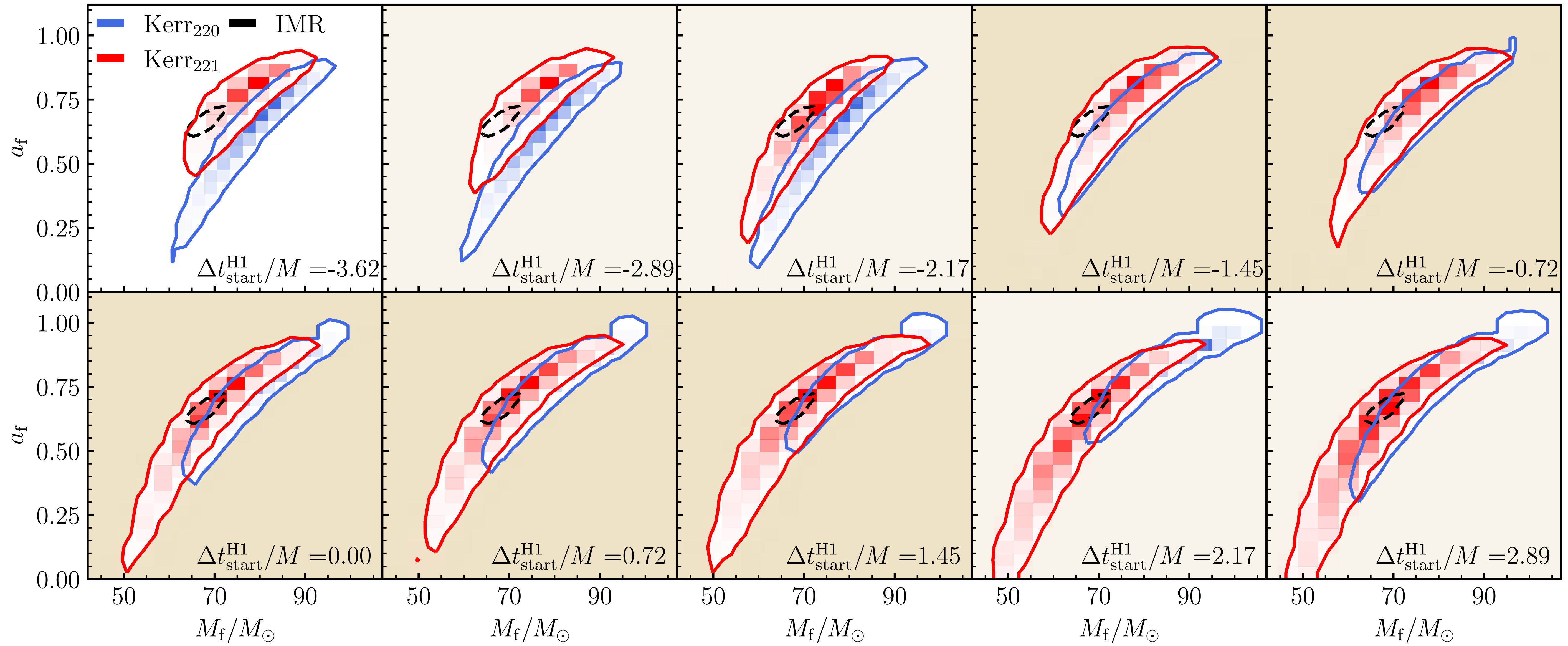
Isi et al. (2019)



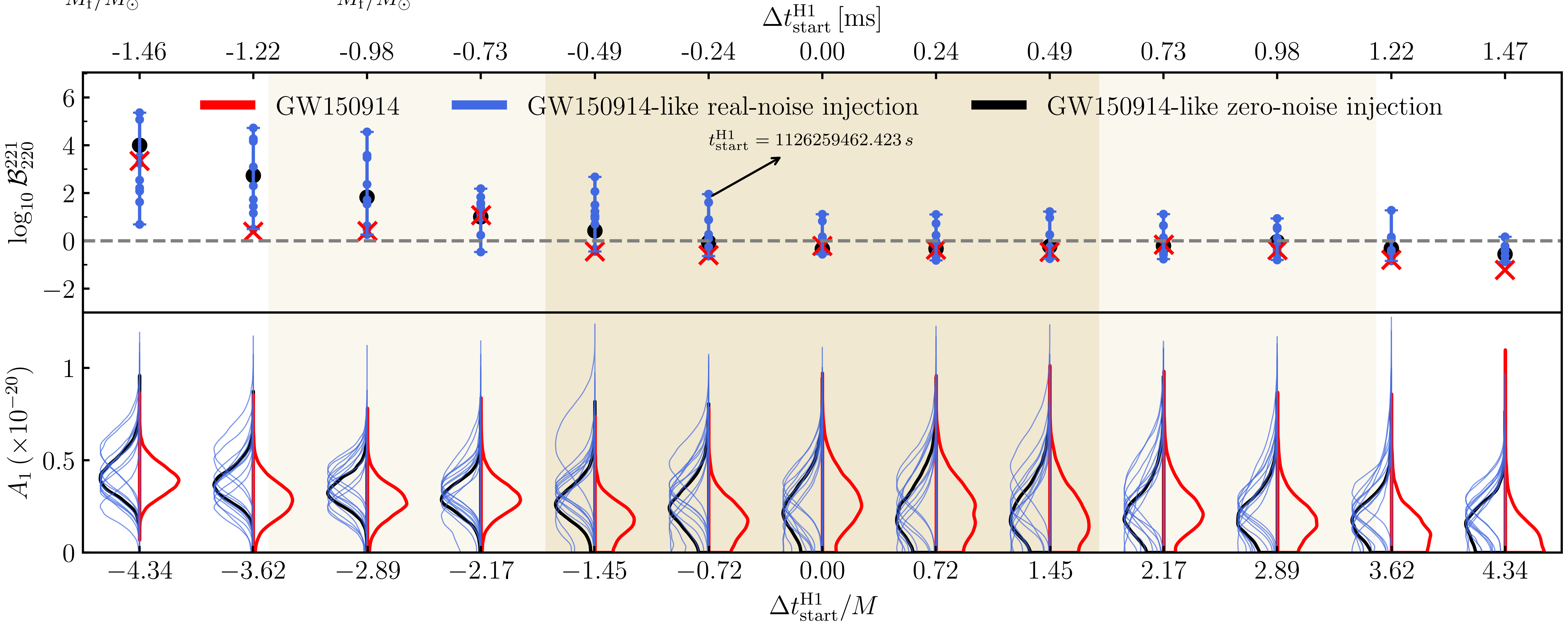


$$N = 1 \begin{cases} 90\% \\ 3.6\sigma \end{cases}$$

Ringdown PE for GW150914

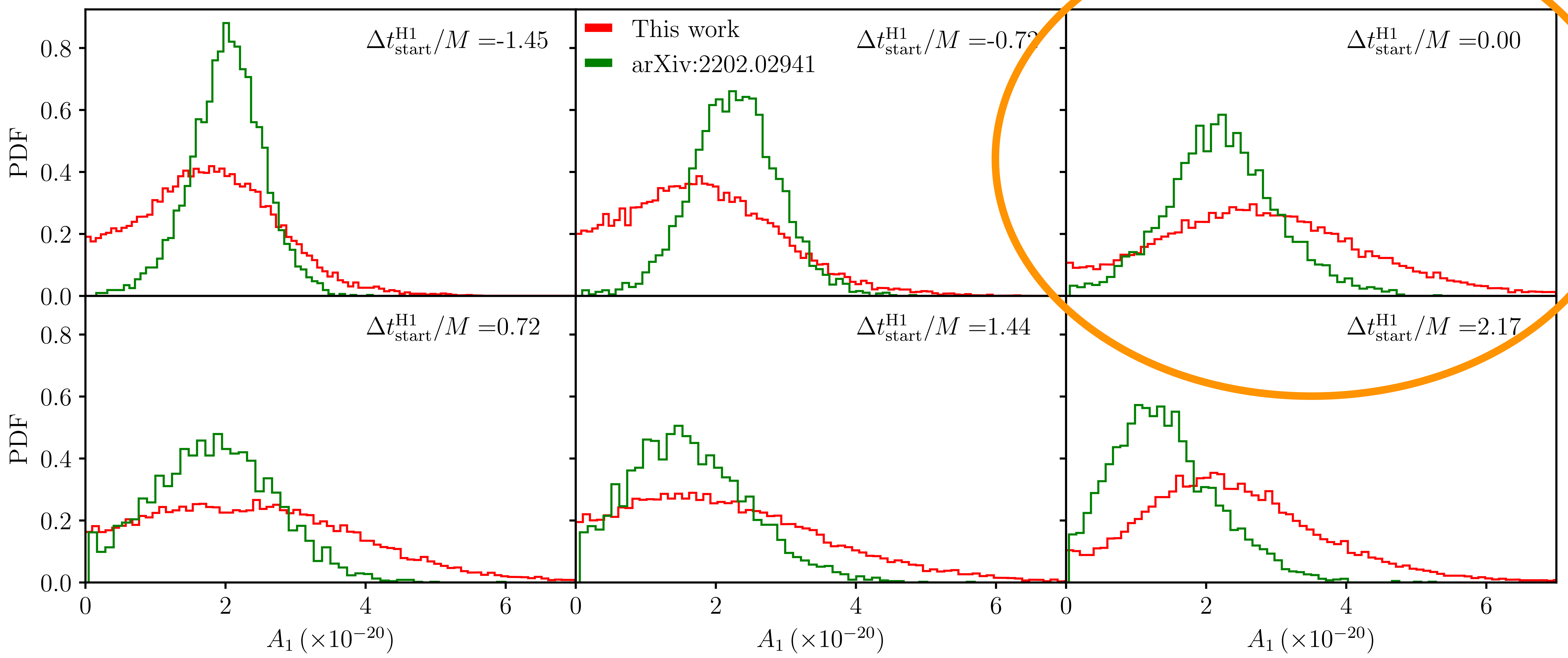


Cotesta et al. (2022)

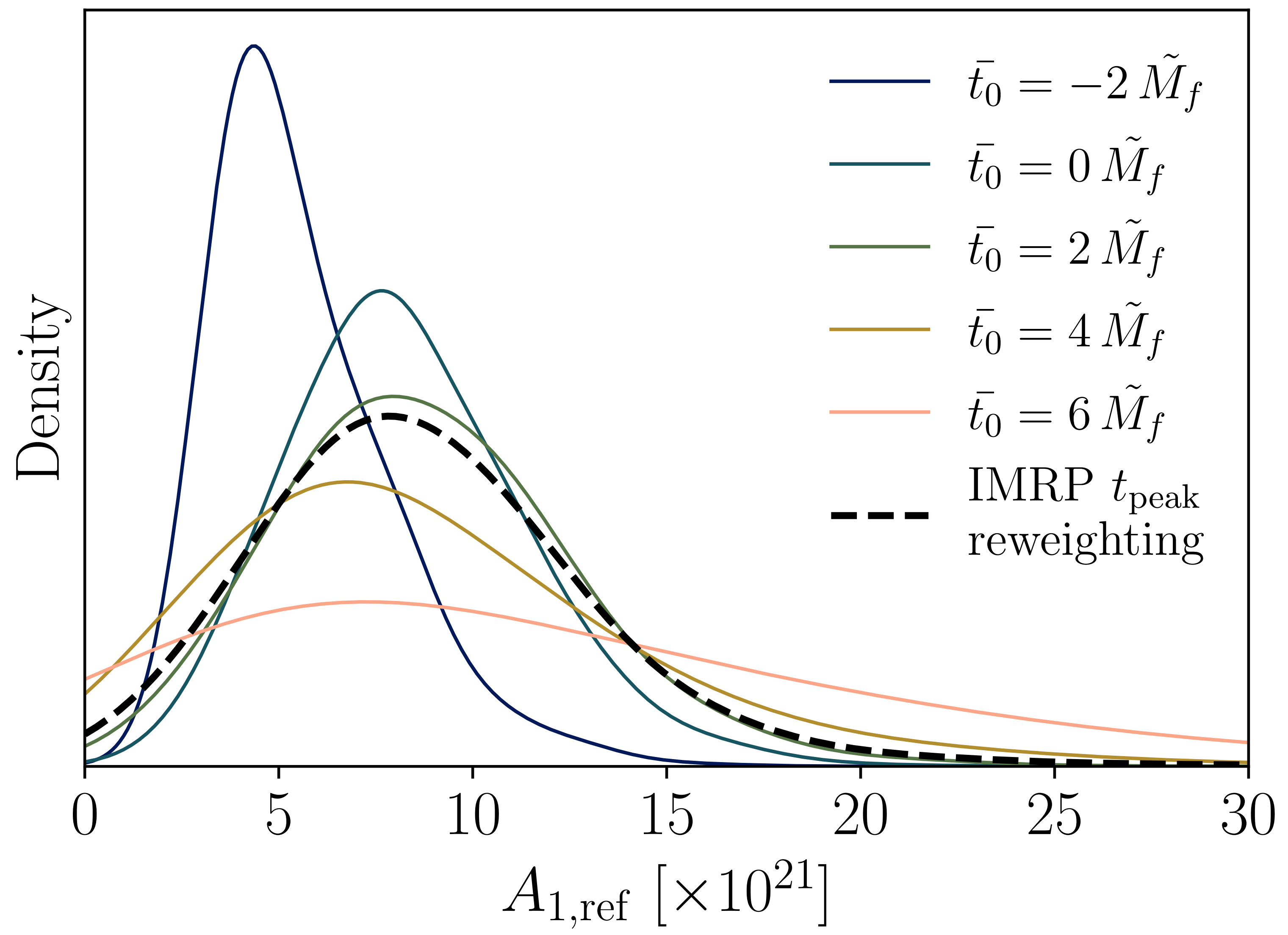


Ringdown PE for GW150914

Cotesta et al. (2022)



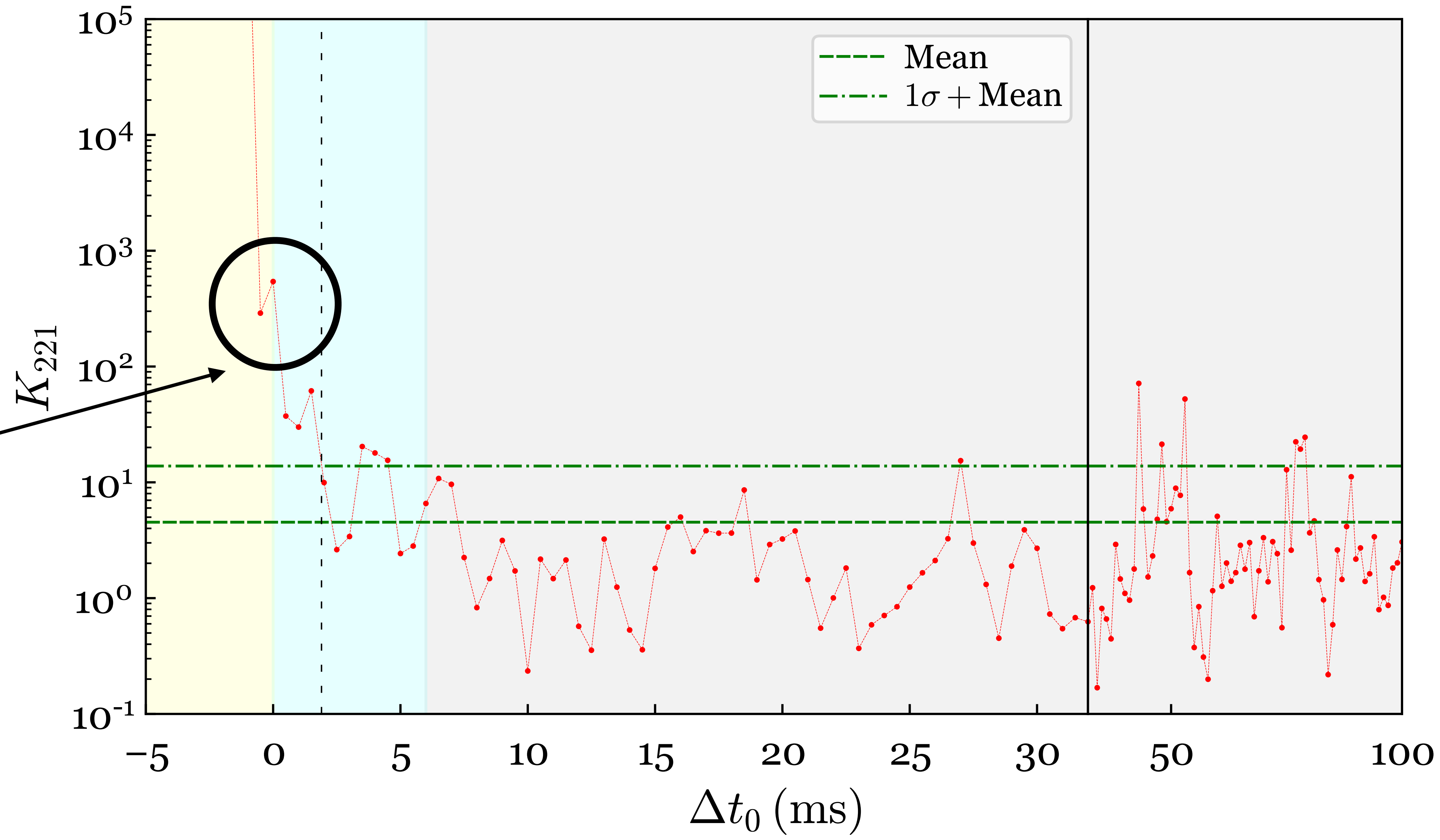
Ringdown PE for GW150914



Finch & Moore (2022) $\sim 1.8\sigma$

Frequency domain analysis

Ma et al. (2023)



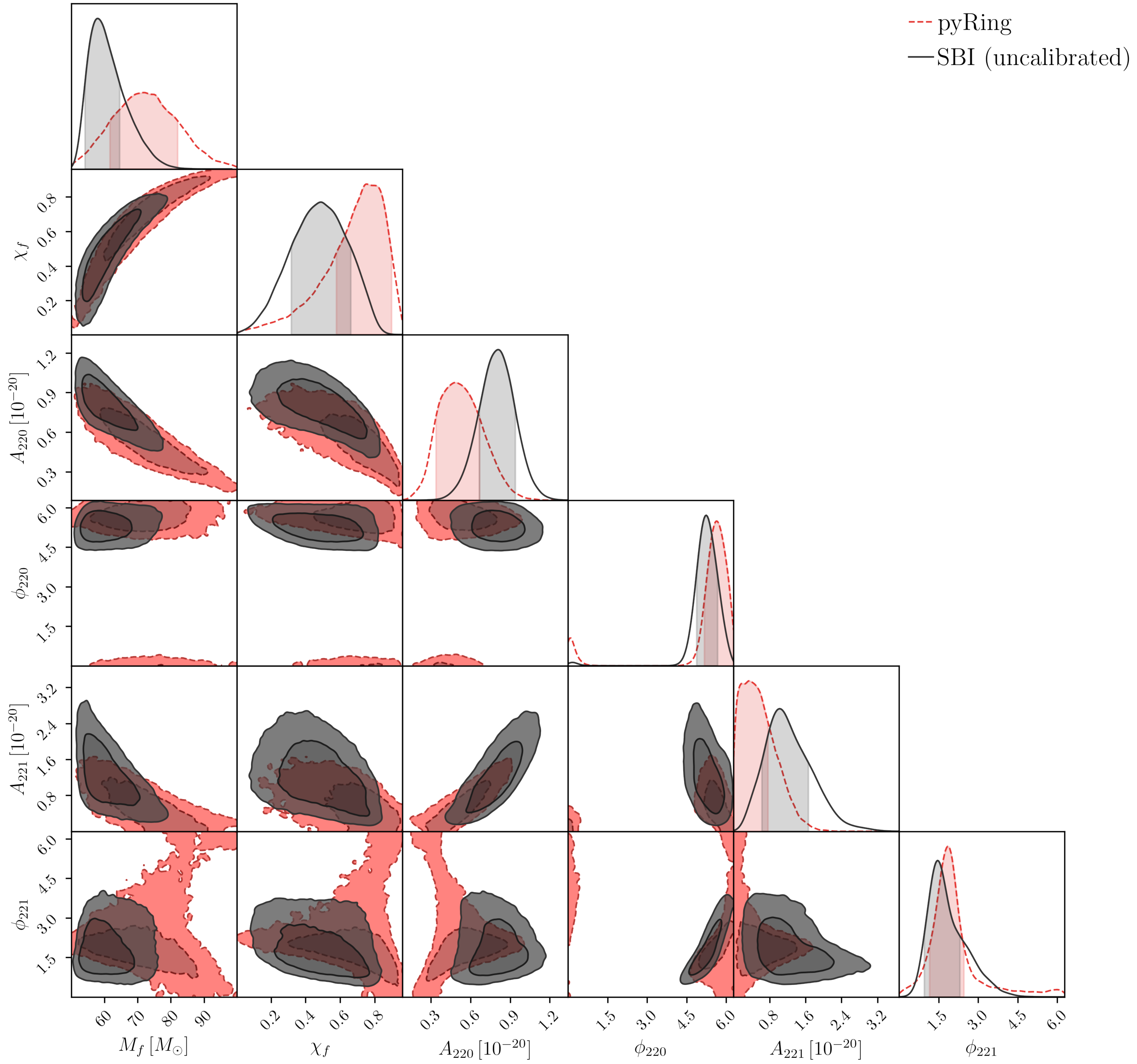
Bayes factor 600

SBI PE for ringdown GW150914

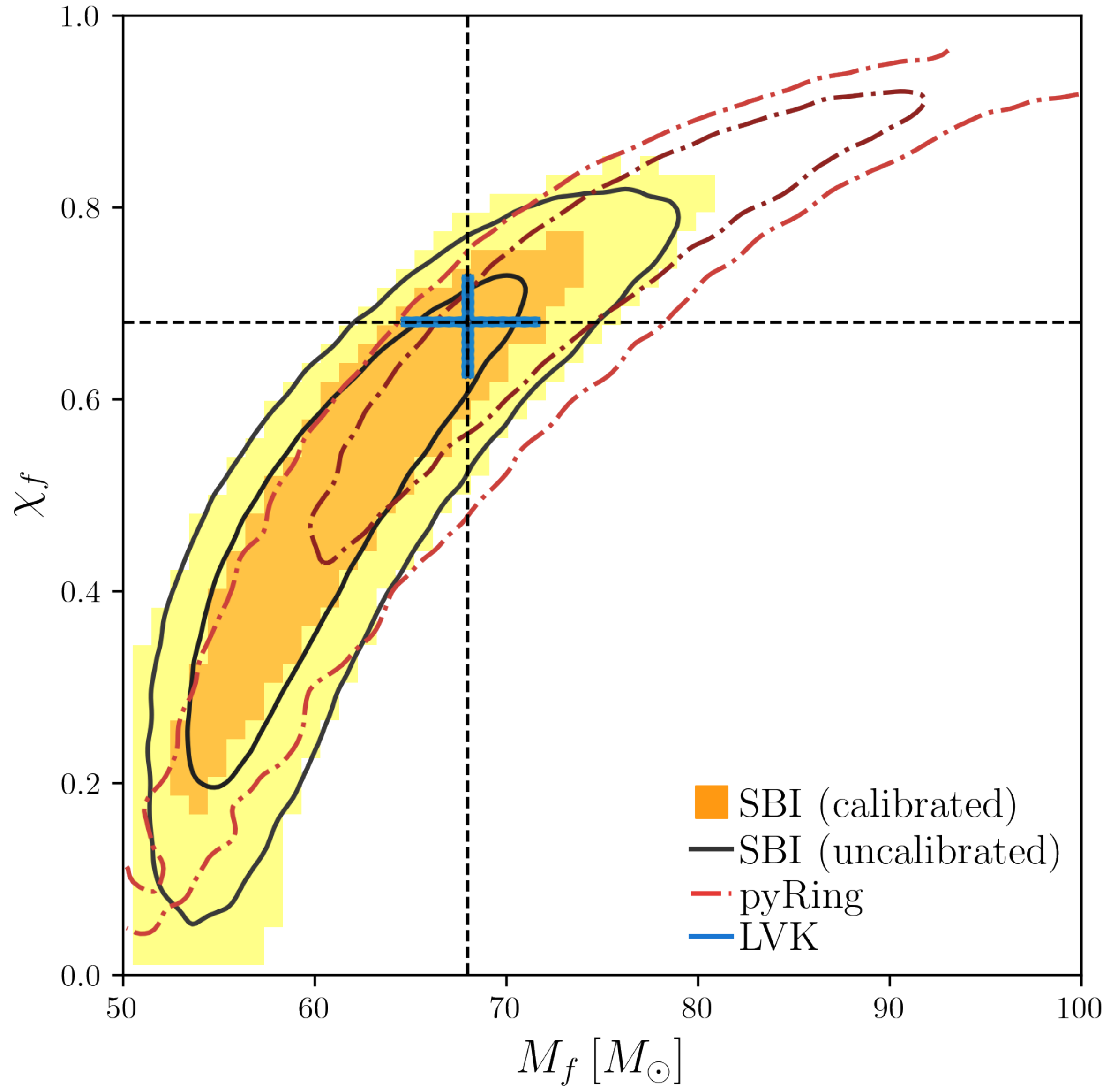
- 10^5 waveforms x 50 noise realisations each
- $\ell = m = 2 \quad , \quad n = 0, 1$
- T=0.1 s, sample rate 4096Hz, $t_0 = t_{\rm peak} = 1126259462.42323 \text{ GPS} \quad , \quad f \in [100, 650]\text{Hz}$
- Uniform priors $M_f \in [50, 100]\, M_\odot,\ \chi_f \in [0, 1],\ A_{22n} \in [0, 5] \cdot 10^{-20} \text{ and } \phi_{22n} \in [0, 2\pi]$
- Learning rate $5 \cdot 10^{-4}$, batches size 100, stop after 25 epochs of no loss decrees
- Priors truncation scheme at 4σ, factor 2 condition
- H1 (107+125) epochs, L1 (145) epochs
- ~10 minutes x epoch on NVIDIA A-100 GPU (total training time: 2 days)

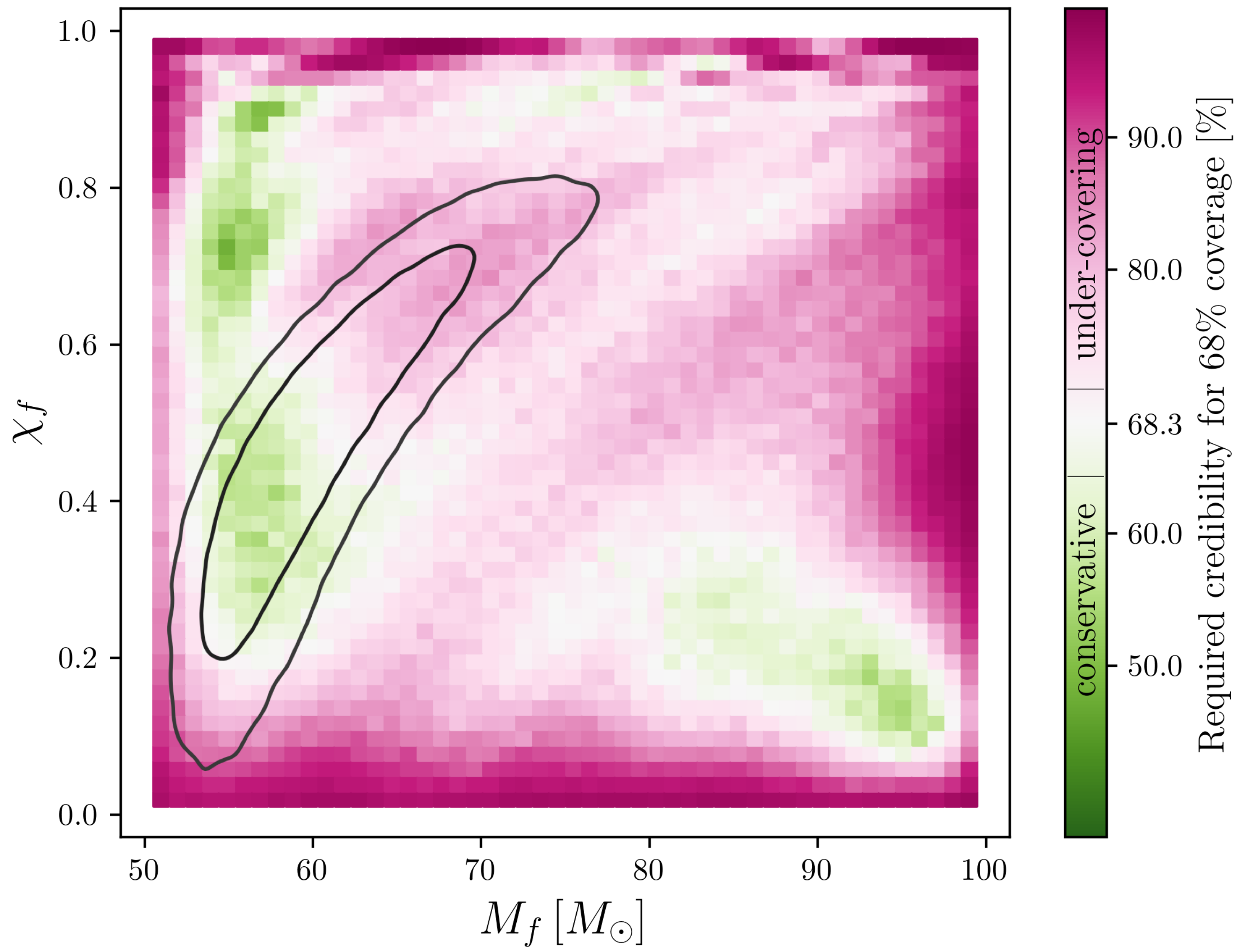
SBI PE for ringdown GW150914

<u>Un-calibrated inference</u>



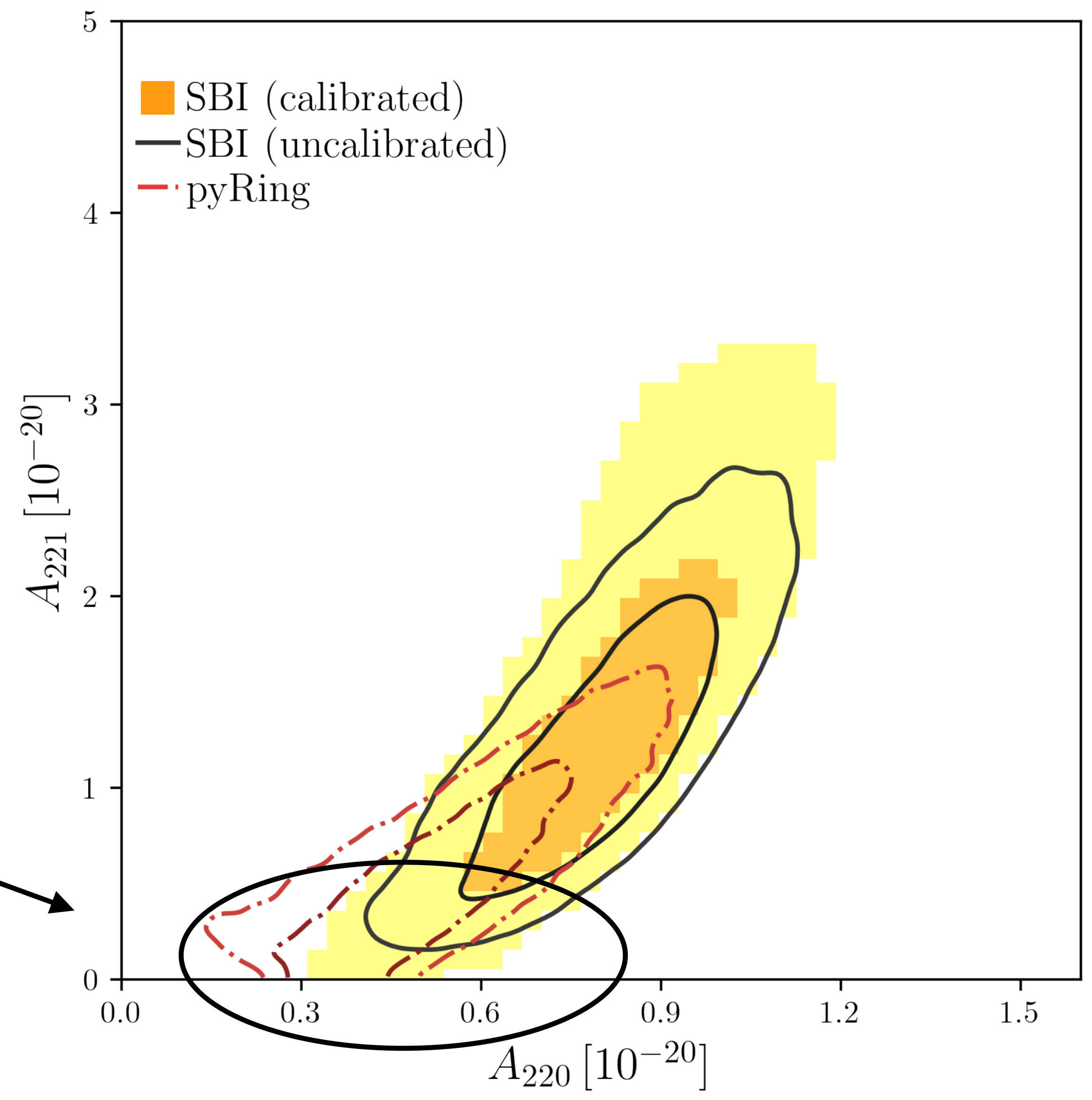
SBI PE for ringdown GW150914

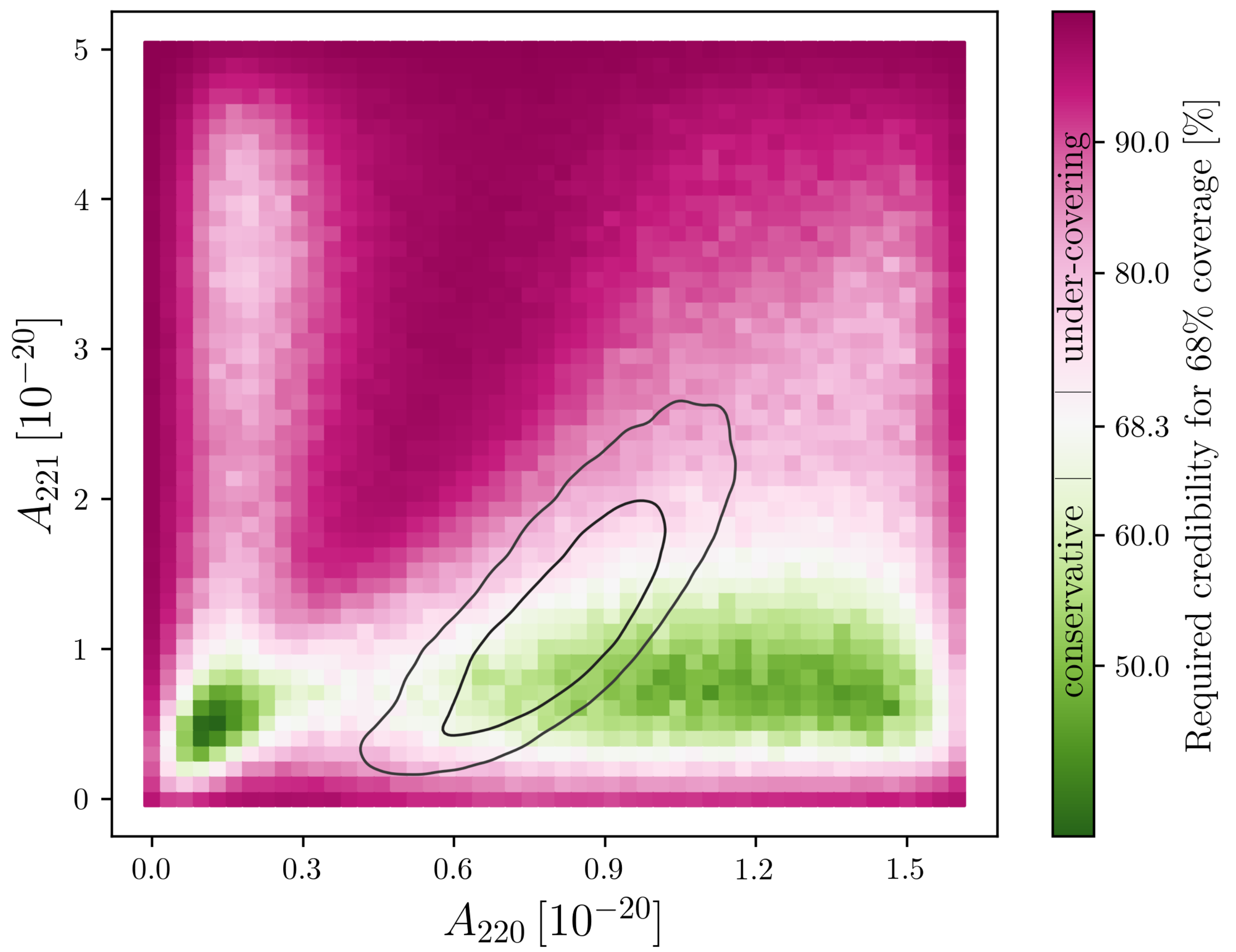




2500 pixels (P-P plot)
x
1000 inferences each

Calibration





$\sim 2\sigma$

Conclusions

- Fast inference allows frequentist validation and calibration of posteriors
- Extremely important for low SNR (~ 9) GW events
- Applied to ringdown of GW150914
- Calibration has an impact on the physical results, lowering the evidence of the first overtone
- Future extensions:
  - non-Gaussian/non-stationary noise
  - More complex waveforms [e.g. higher harmonics (GW190521), non-linear QNMs]
  - Tests of gravity
  - LISA / ET (higher SNR)

Thank you!