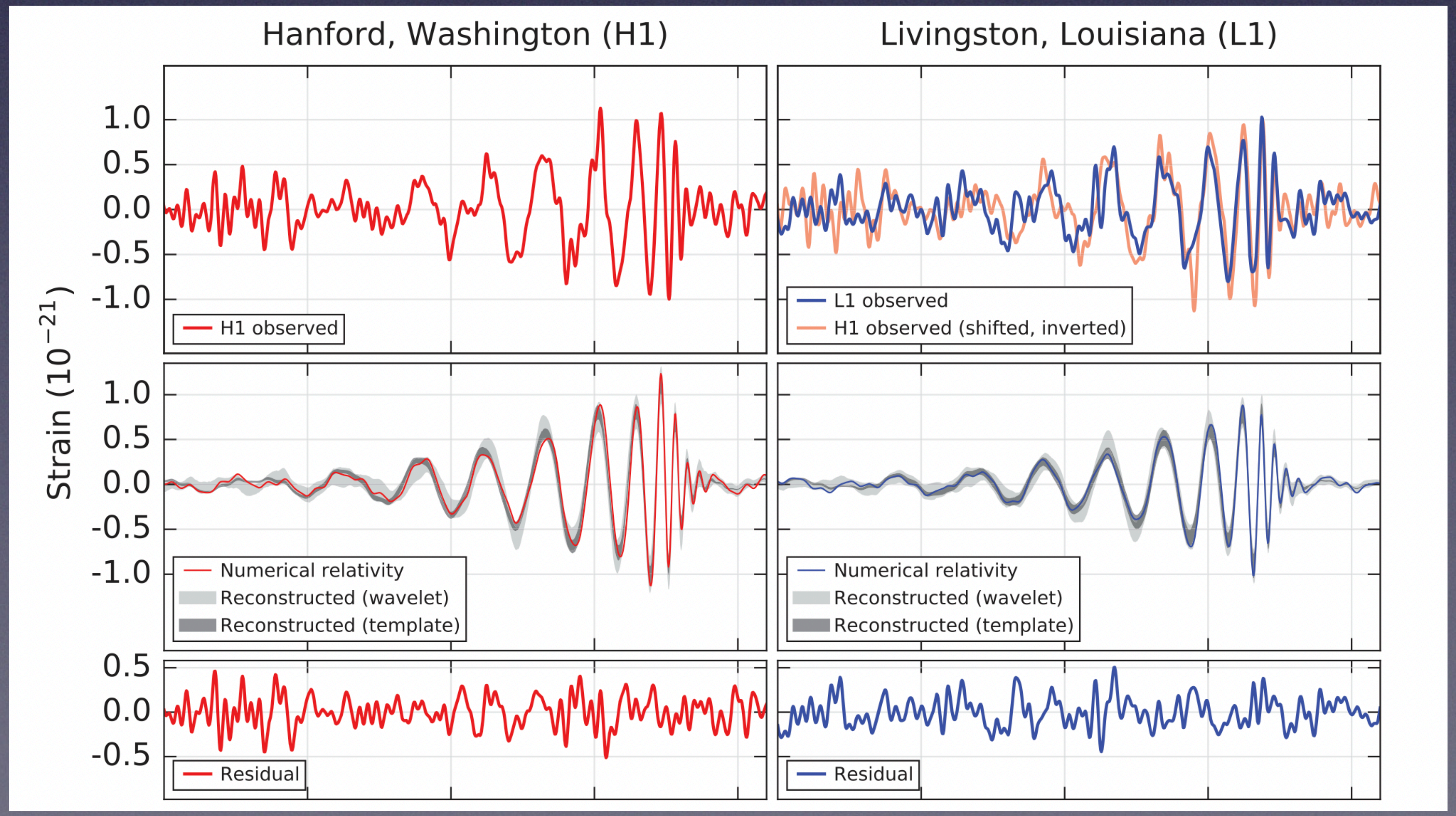
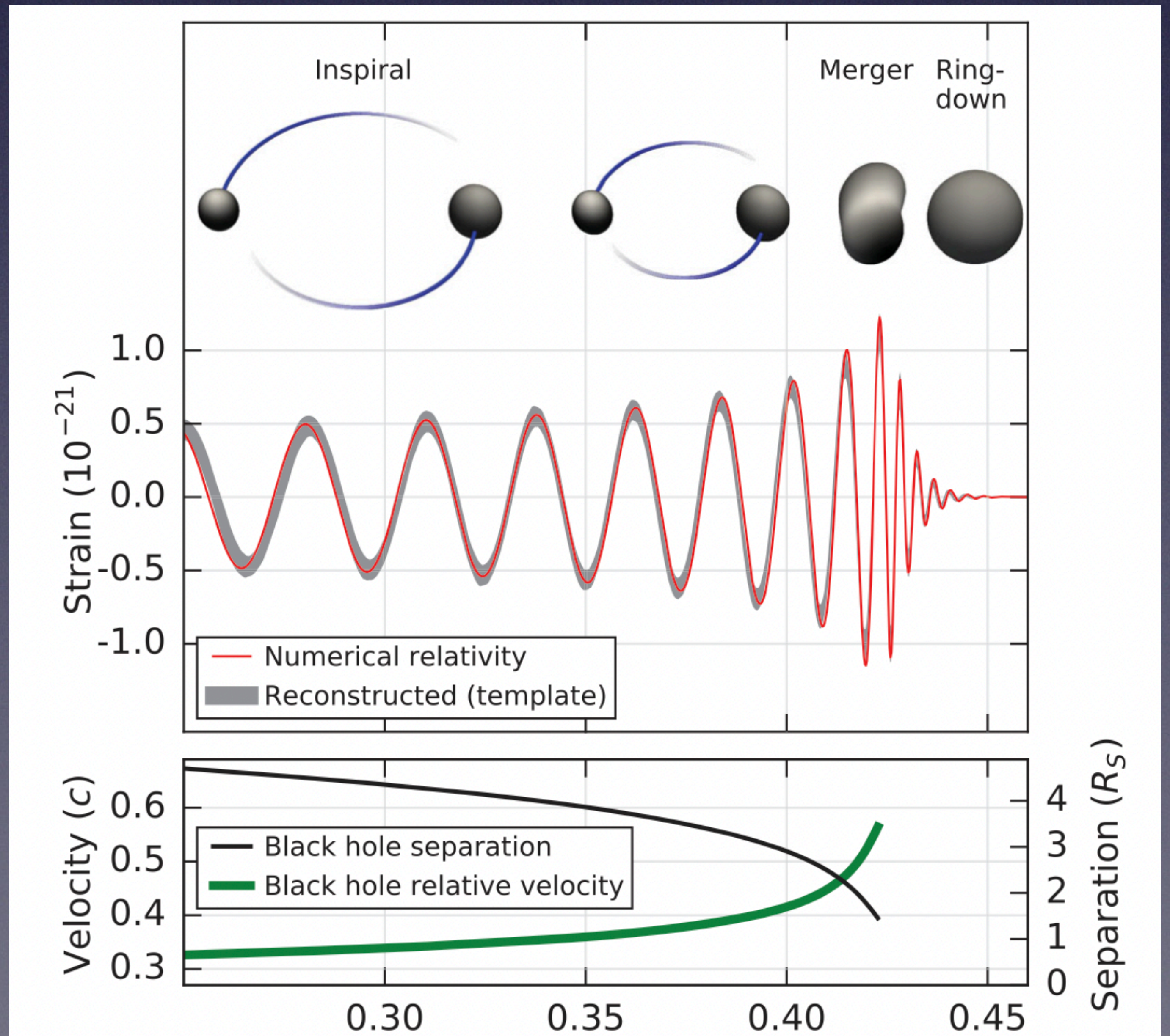
Bayesian statistics applied to gravitational waves

Ollie Burke
ollie.burke@l2it.in2p3.fr



(PRL 116, 061102)



(PRL 116, 061102)

The structure

1. **Review**: Gravitational Waves
2. **Crash course**: Bayesian Statistics + **simple application**
3. **Discussion**: Current challenges + prospects for future

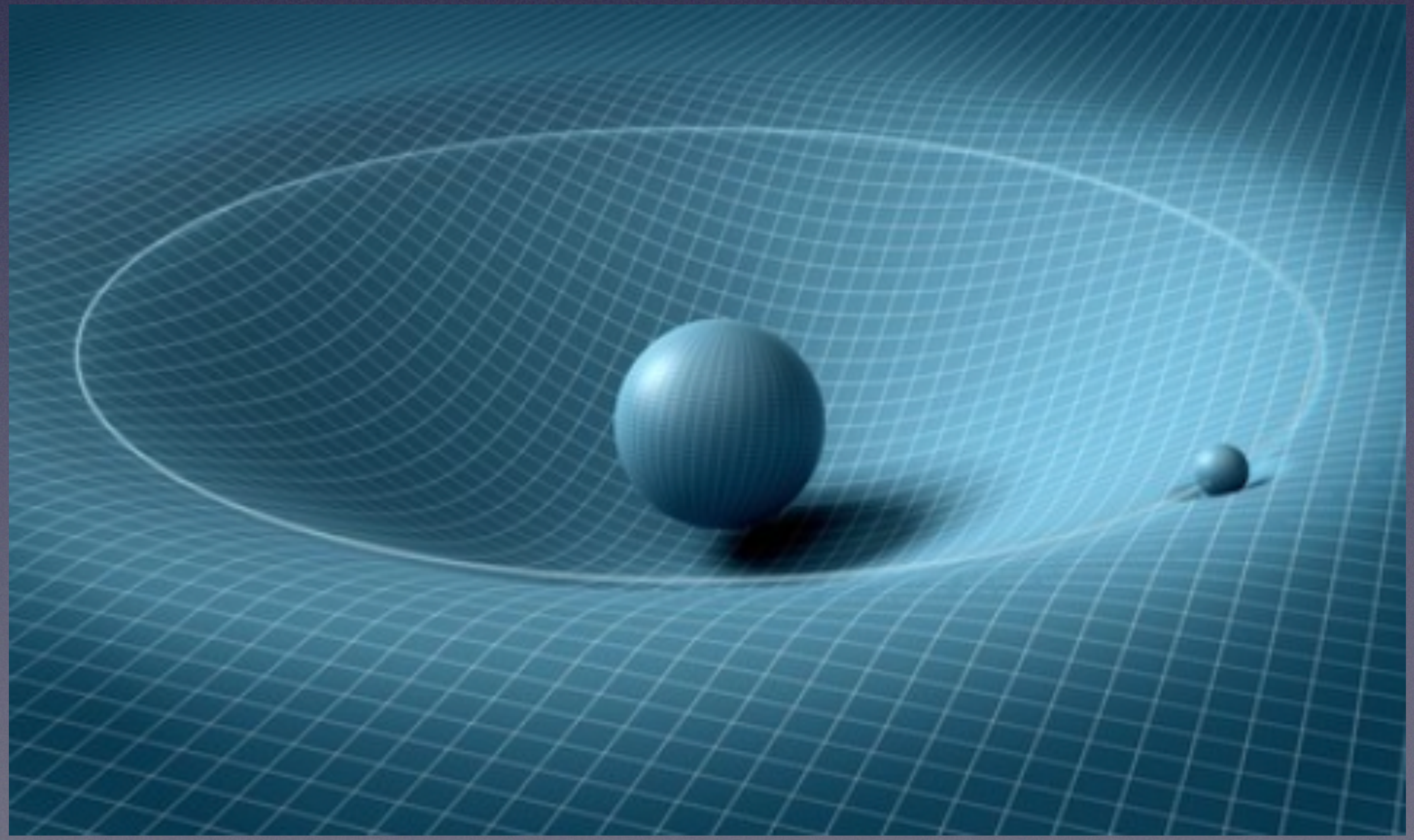
Part 1: **Gravitational Waves**

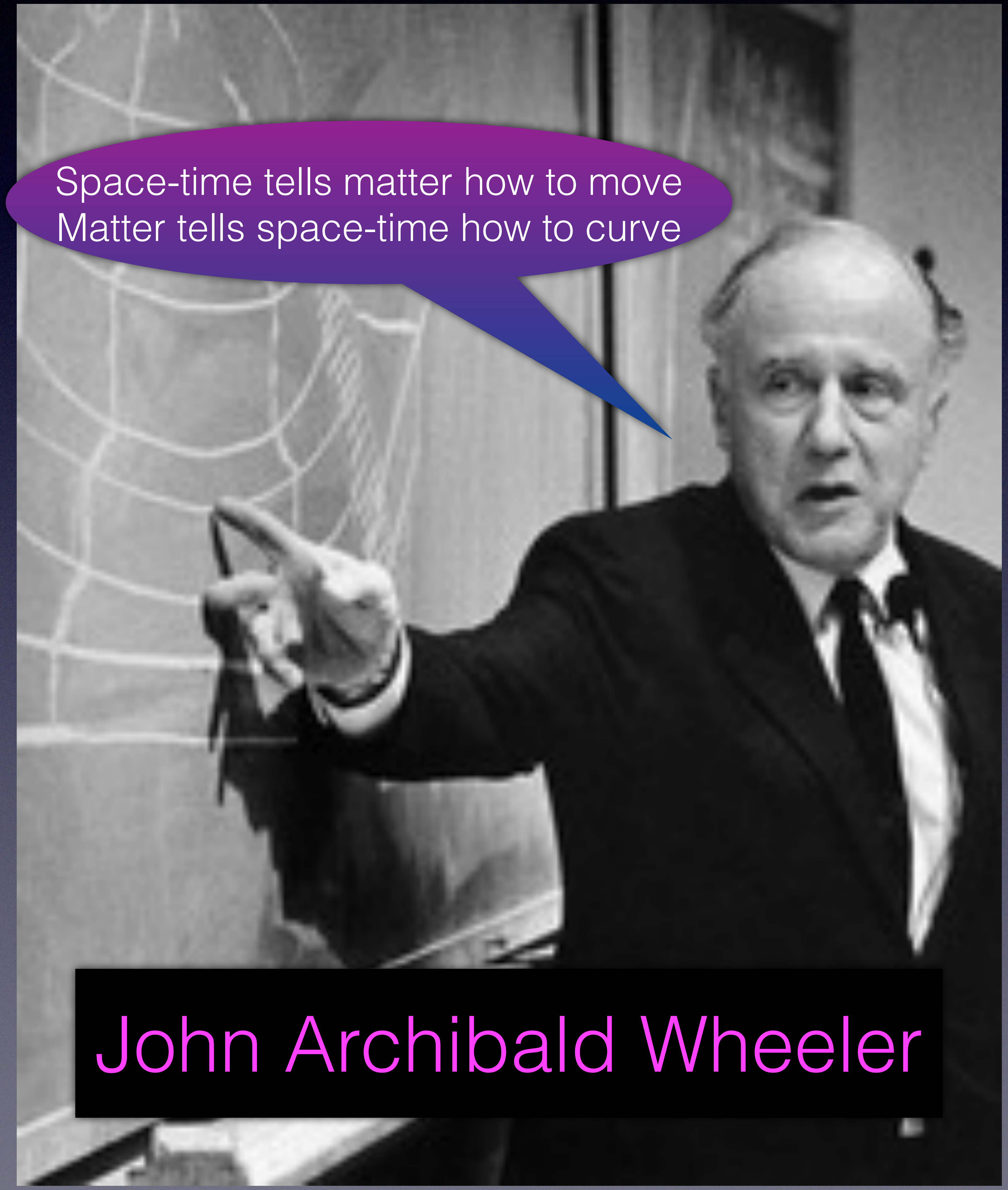
Einstein's Universe

$$\underbrace{G_{\mu\nu}}_{\text{Curvature}} \propto \underbrace{T_{\mu\nu}}_{\text{Stuff}}$$

Einstein's Universe

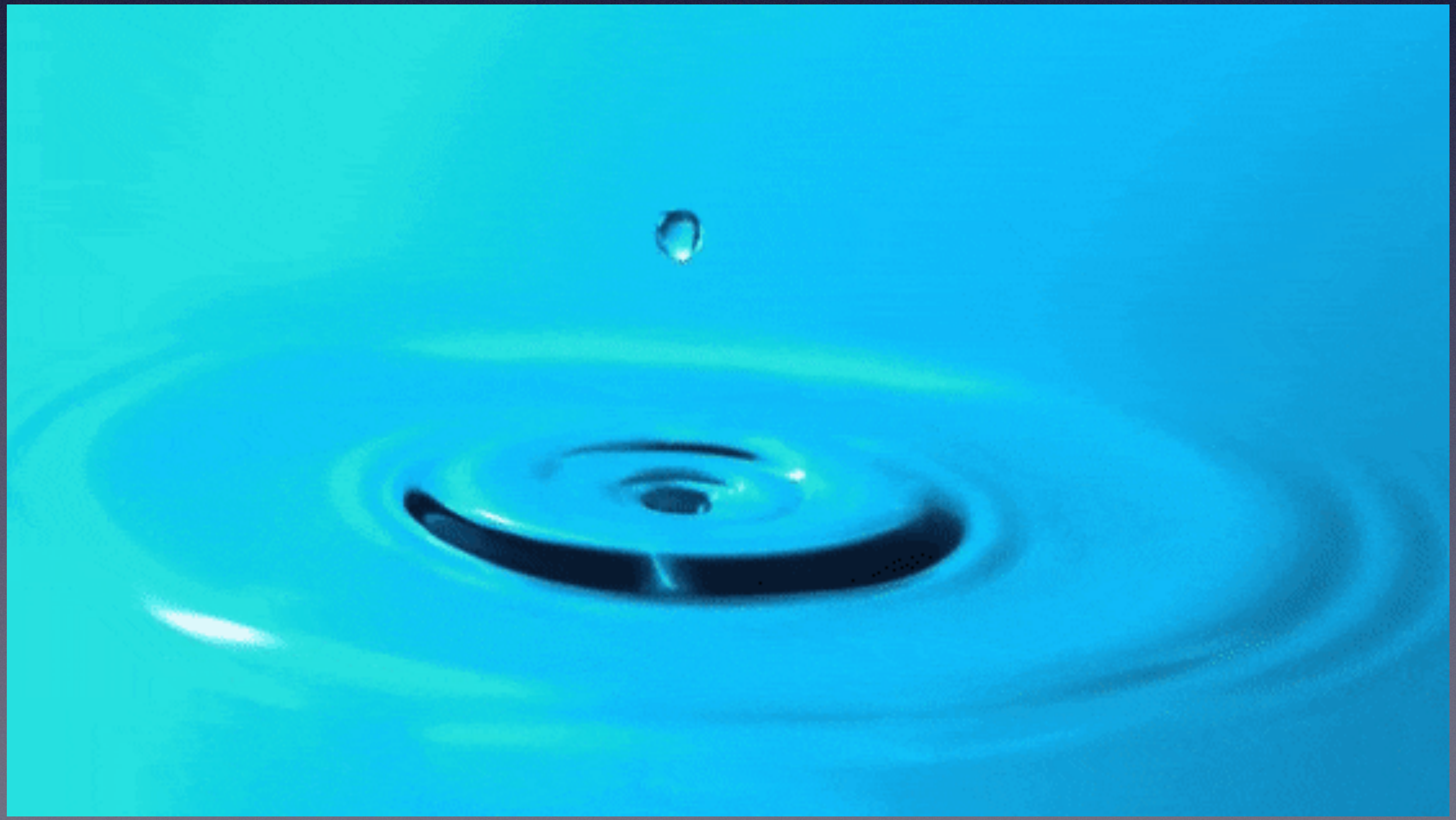
$$\underbrace{G_{\mu\nu}}_{\text{Curvature}} \propto \underbrace{T_{\mu\nu}}_{\text{Stuff}}$$





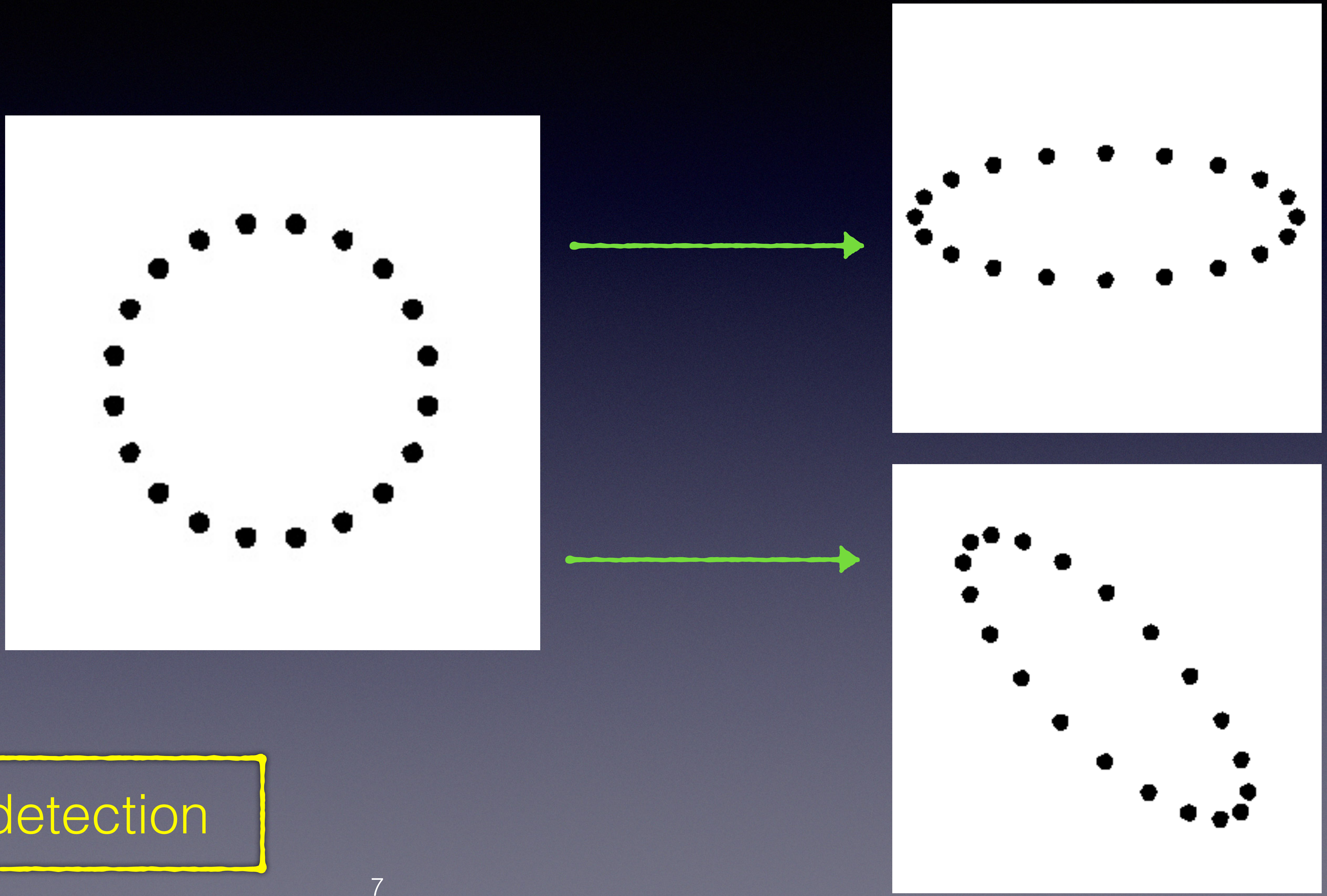
Small disturbance to this “fabric”

- When space-time is disturbed, those “disturbances” must go somewhere
- **Gravitational Waves**: Propagating, oscillating gravitational fields
- **Sourced by:** Acceleration of objects in binary system



The effect on particles

In General Relativity
Two polarisations



Fundamental to GW detection

How small is small?

$$\text{Strain} \approx 10^{-42} \left(\frac{\text{Separation}}{50\text{ cm}}\right)^2 \cdot \left(\frac{\text{Frequency}}{1\text{ Hz}}\right)^2 \cdot \left(\frac{\text{Mass}}{500\text{ grams}}\right) \cdot \left(\frac{\text{Distance}}{1000\text{ km}}\right)^{-1}$$

(LIGO archives)

$$\text{Strain} \approx 10^{-21} \left(\frac{\text{Separation}}{100\text{ km}}\right)^2 \cdot \left(\frac{\text{Frequency}}{400\text{ Hz}}\right)^2 \cdot \left(\frac{\text{Mass}}{1.4 M_\odot}\right) \cdot \left(\frac{\text{Distance}}{15\text{ Mpc}}\right)^{-1}$$

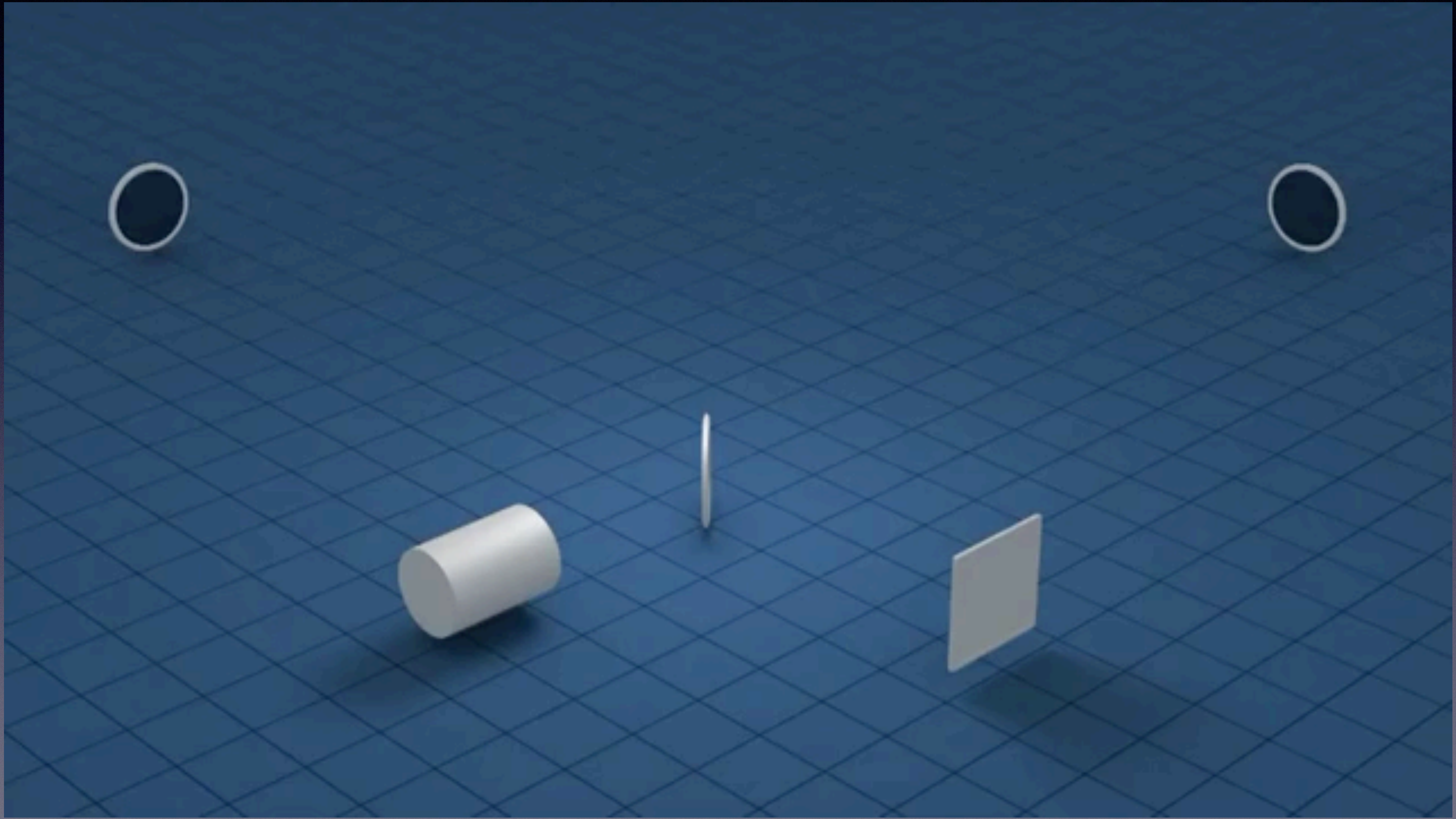
How do we measure:

$$\text{Strain} = \frac{\Delta L}{L} \sim 10^{-21} \text{ ?}$$

Laser Interferometry!

(Rainer Weiss, 1970s)

Laser Interferometry



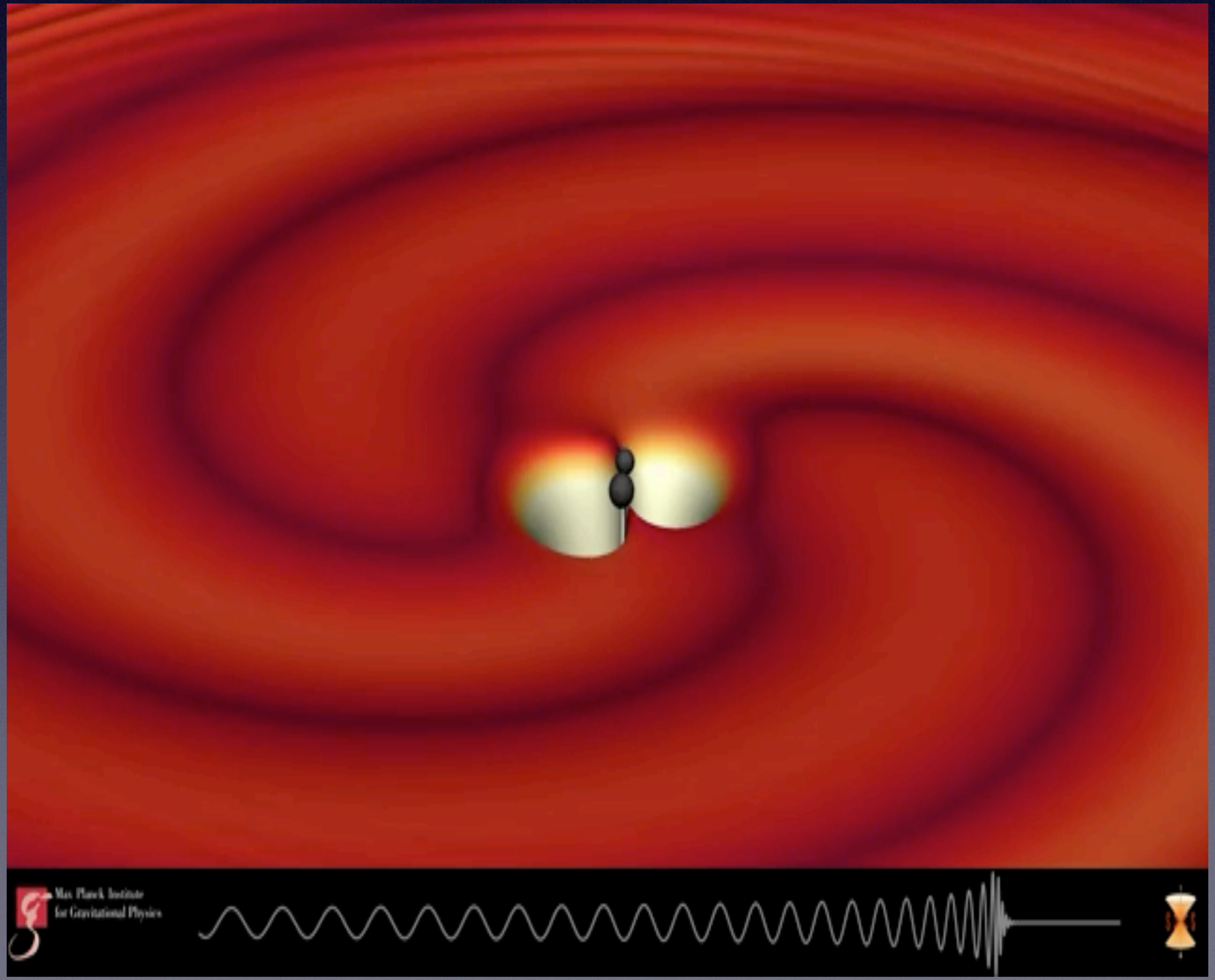
(LIGO archives)

Laser Interferometer Gravitational-Wave Observatory (LIGO)

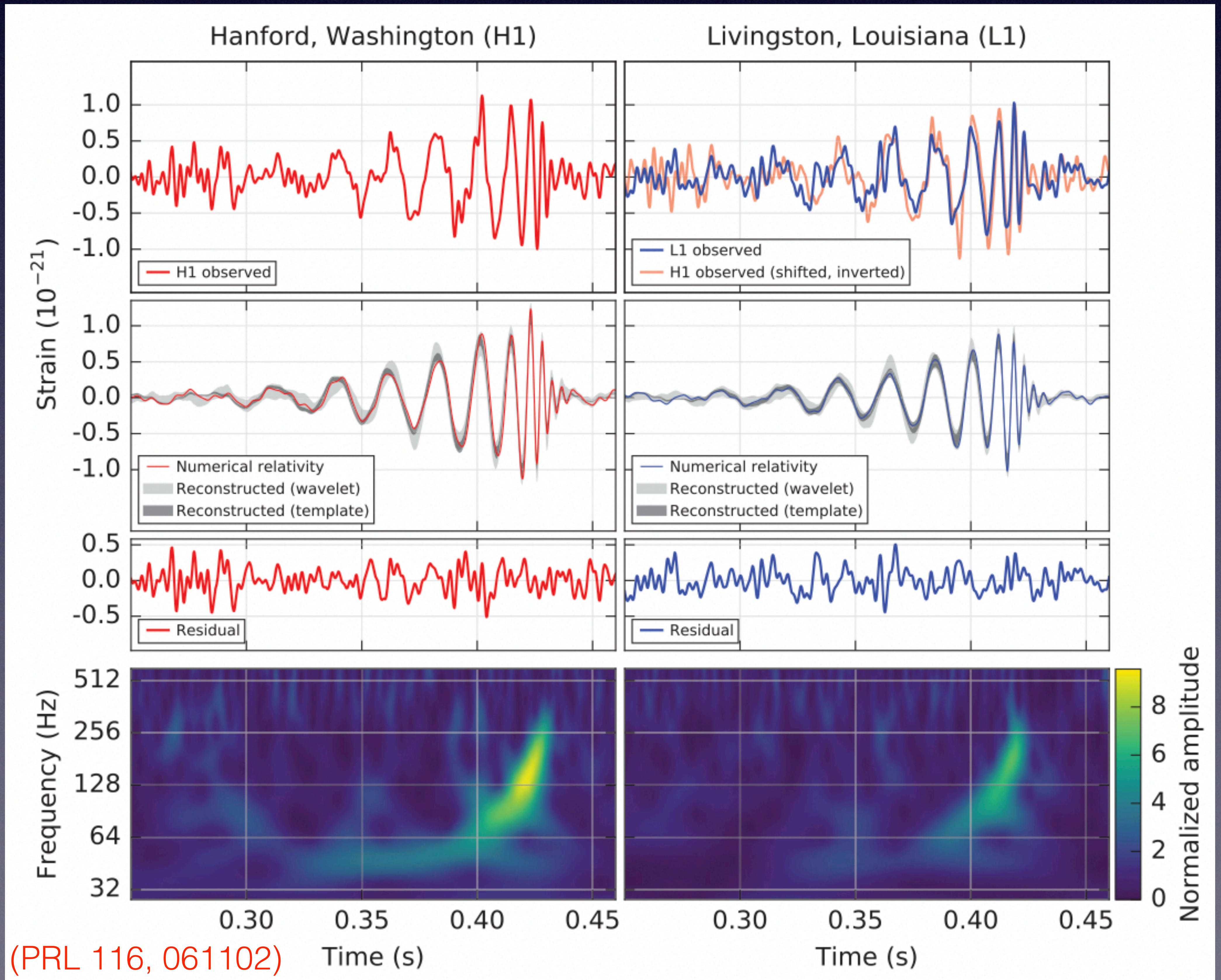
These detectors are the most sensitive instruments on earth

A new era in Observational Astronomy

(https://www.black-holes.org/)



LIGO-Virgo collaboration



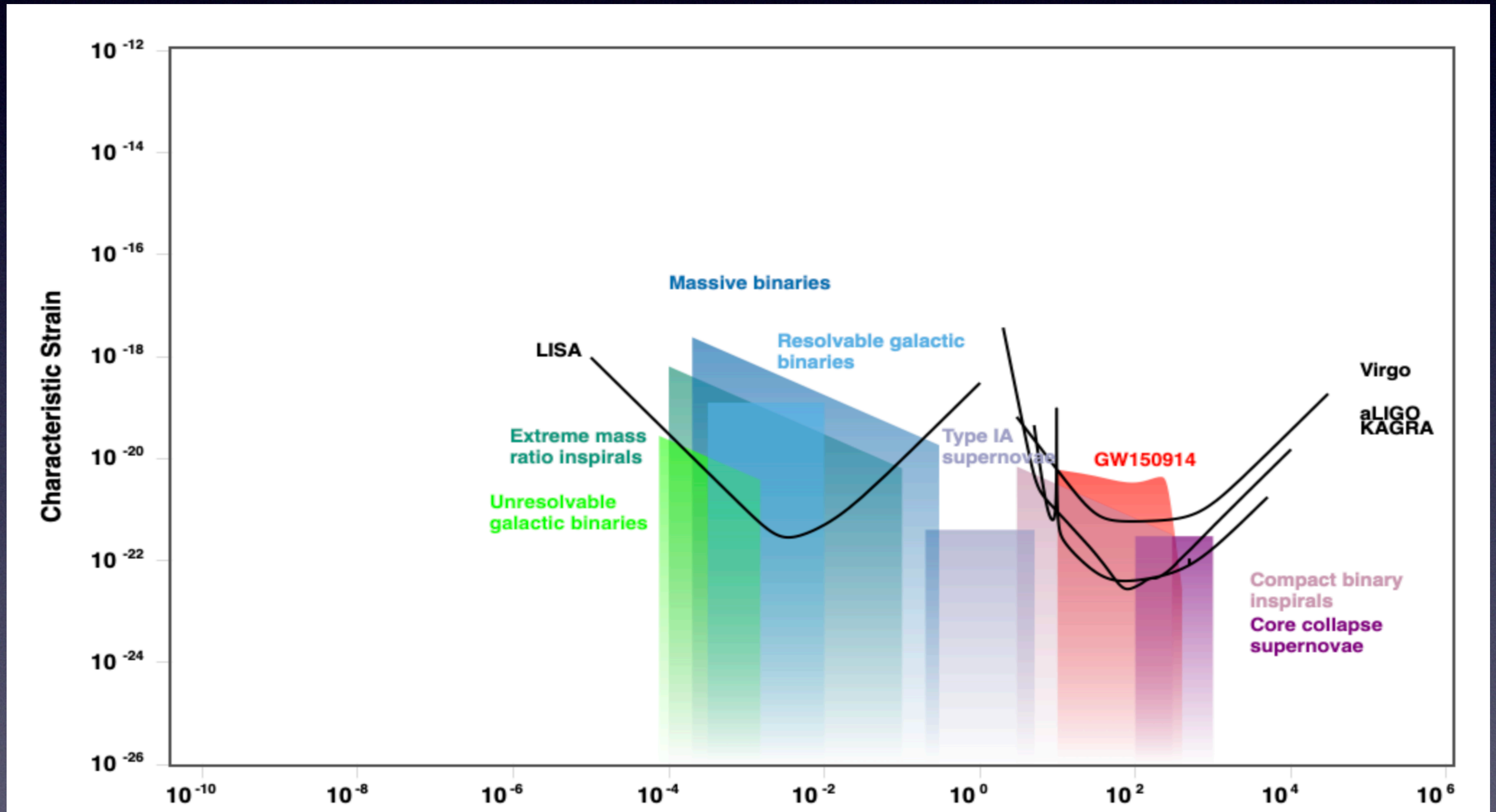
(PRL 116, 061102)

Masses in the Stellar Graveyard

(LIGO-Virgo-KAGRA | Aaron Geller | Northwestern)

The Laser Interferometer Space Antennae (LISA)

(Credit: AEI/Milde Marketing)

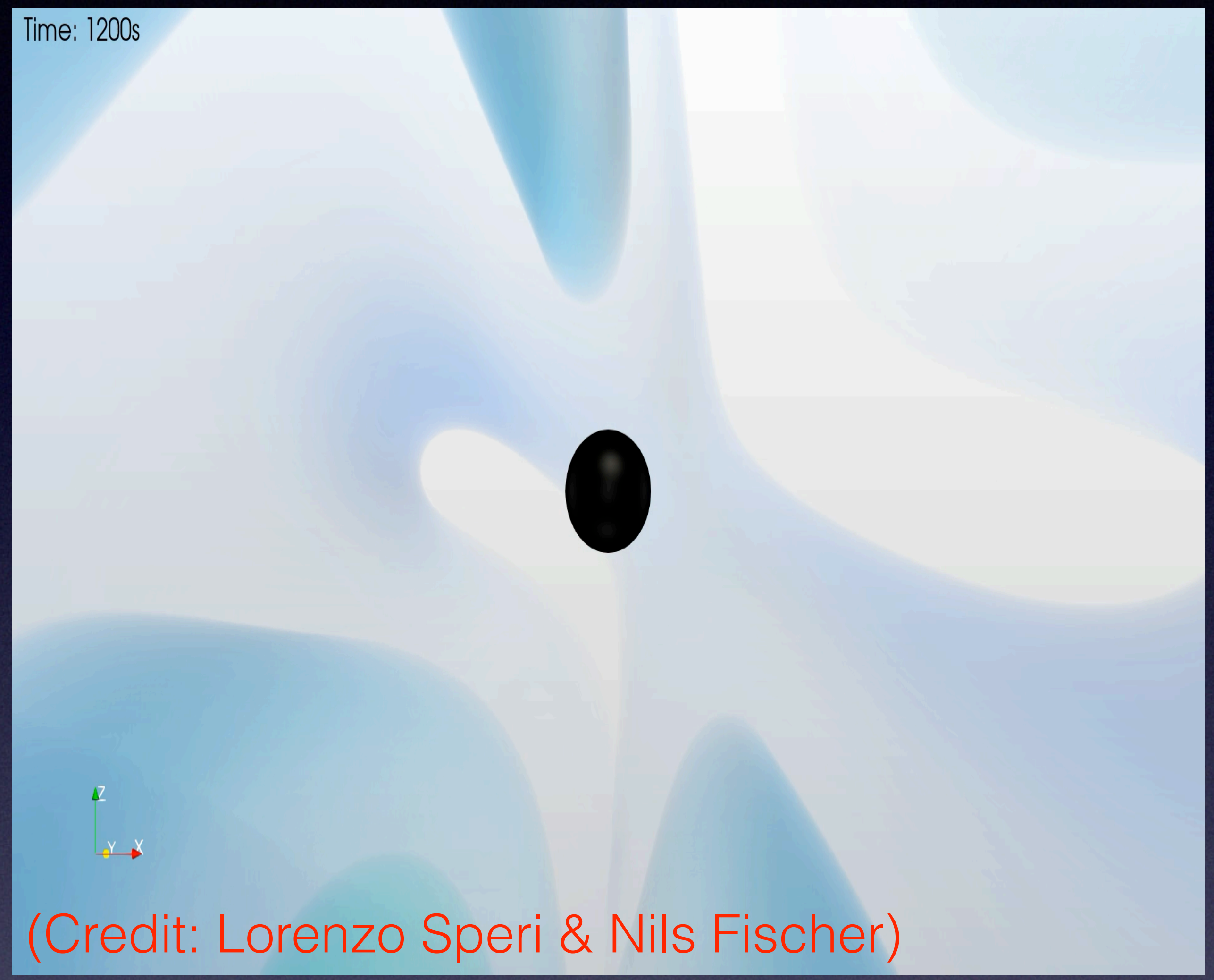


(http://gwplotter.com/)

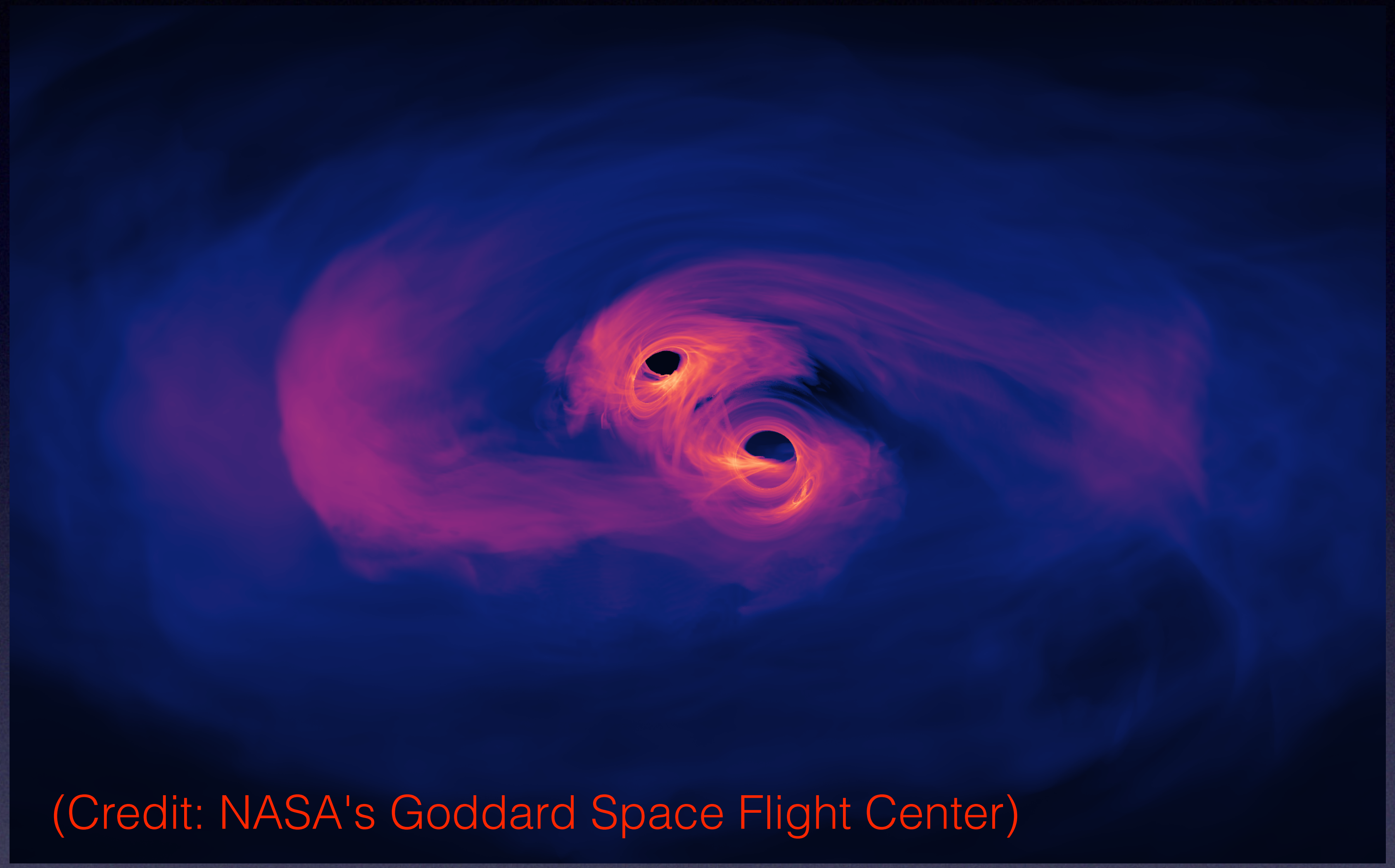
LISA: More sensitive to heavier systems

Example sources for LISA

Galactic binaries



Extreme mass-ratio Inspirals

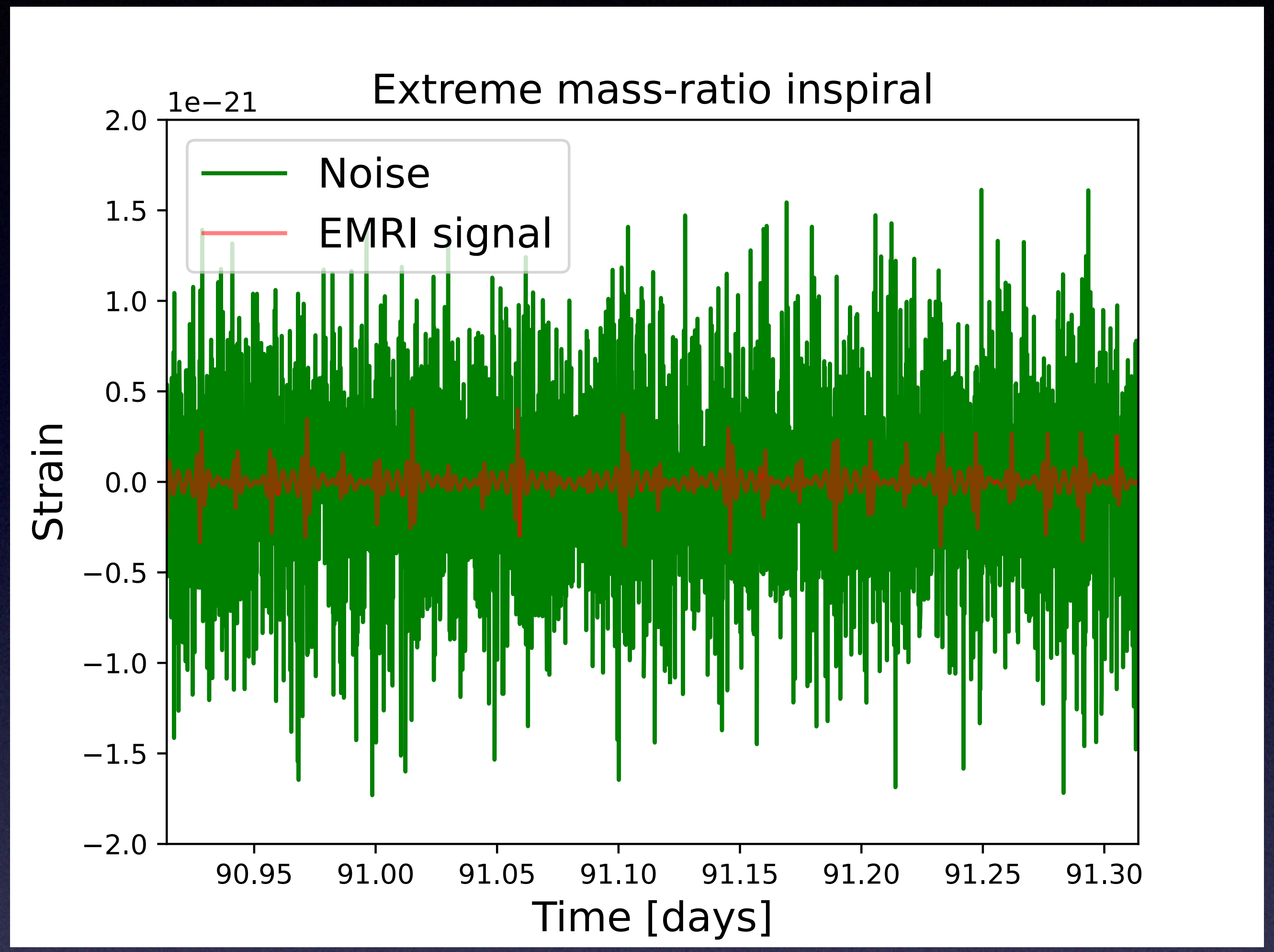


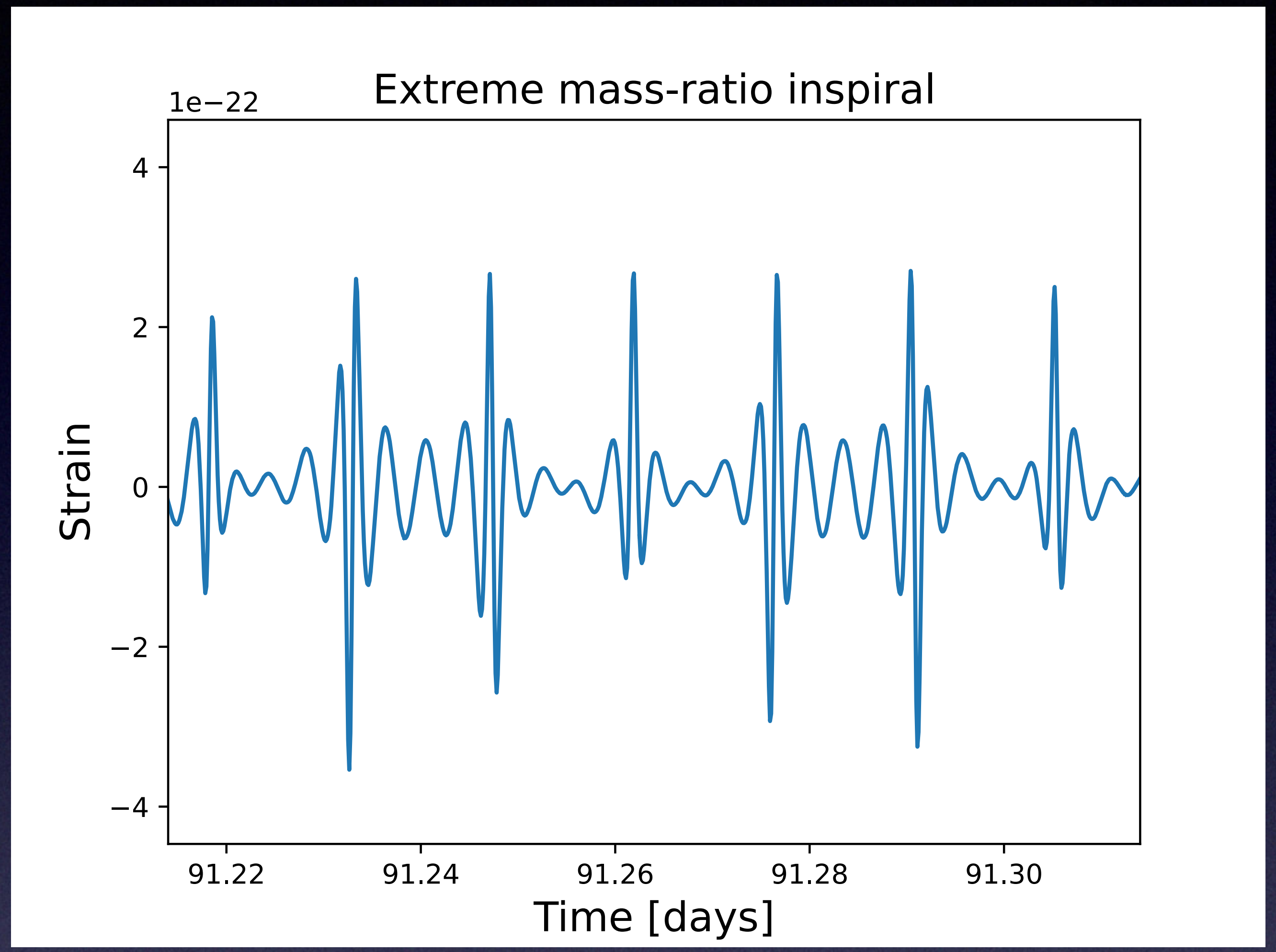
Massive black hole Binaries

All these sources emit gravitational waves in mHz band!
Prime targets for LISA!

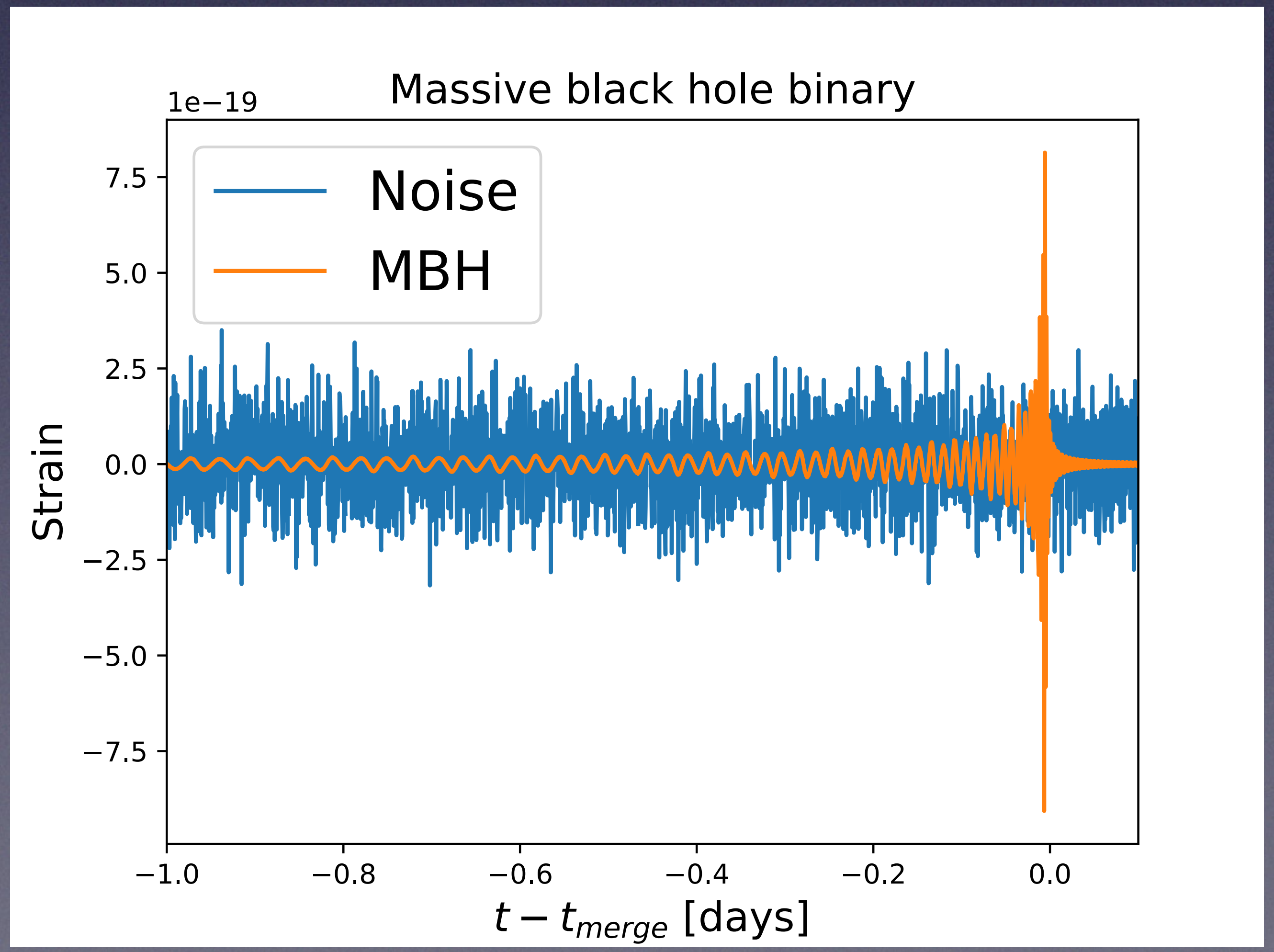
Part 2: **Bayesian Statistics**

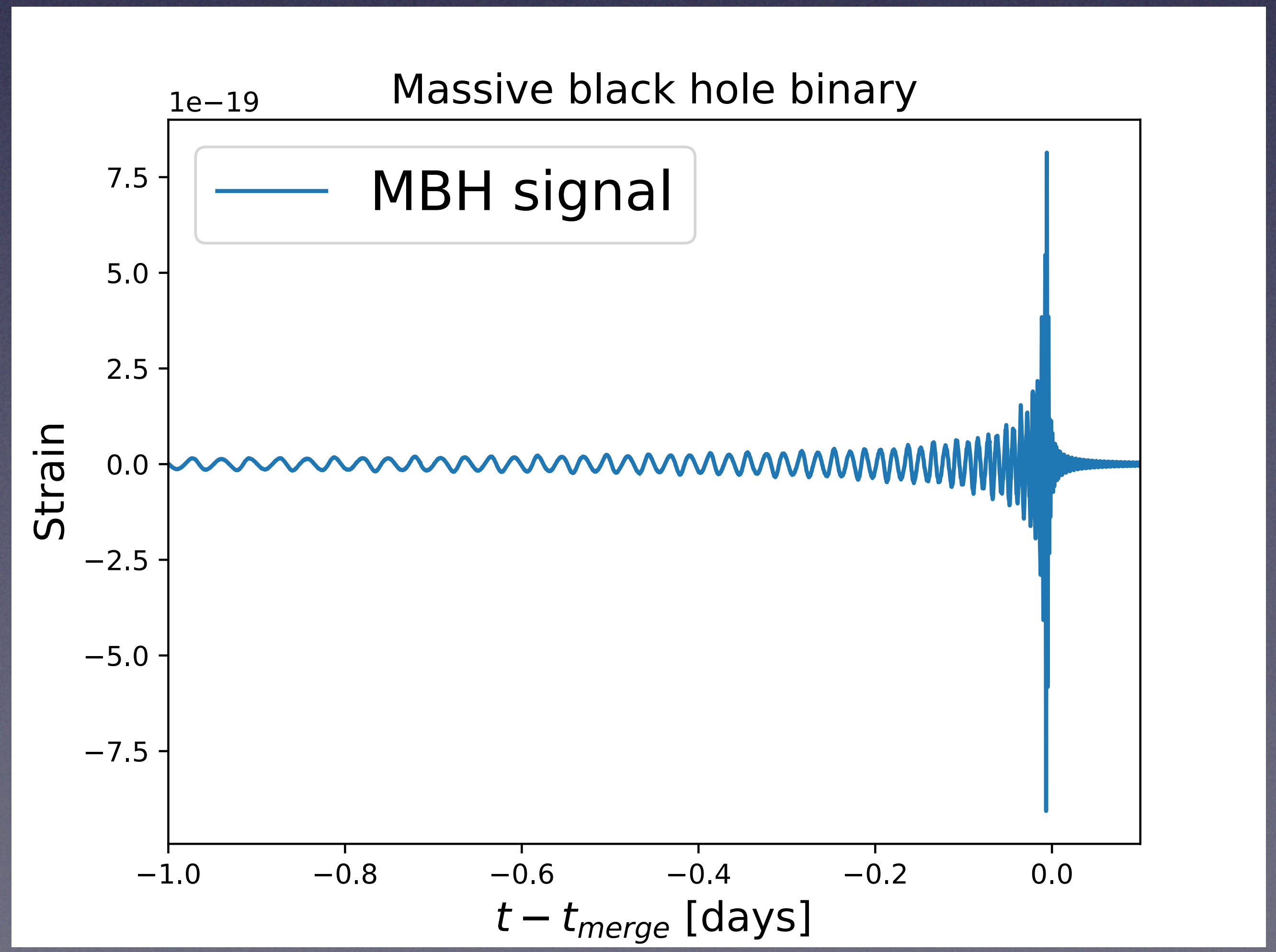
Motivation: Bayesian statistics





Parameters?

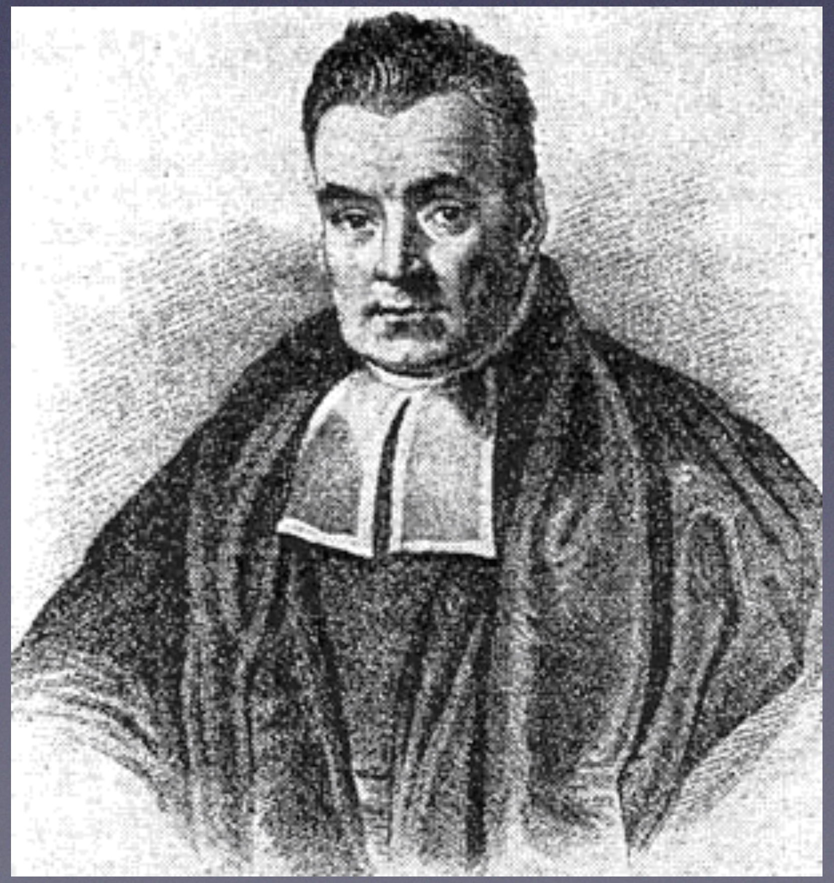




Bayes' Theorem

- $p(d|\boldsymbol{\theta})$: Likelihood
- $p(\boldsymbol{\theta})$: Prior distribution
- $p(\boldsymbol{\theta}|d)$: Posterior distribution

$$p(\boldsymbol{\theta}|d) \propto p(d|\boldsymbol{\theta})\, p(\boldsymbol{\theta})$$



(Reverend Thomas Bayes [1701-61])

Goal: Obtain samples

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}|d)$$

The data stream

data = Signal + Noise

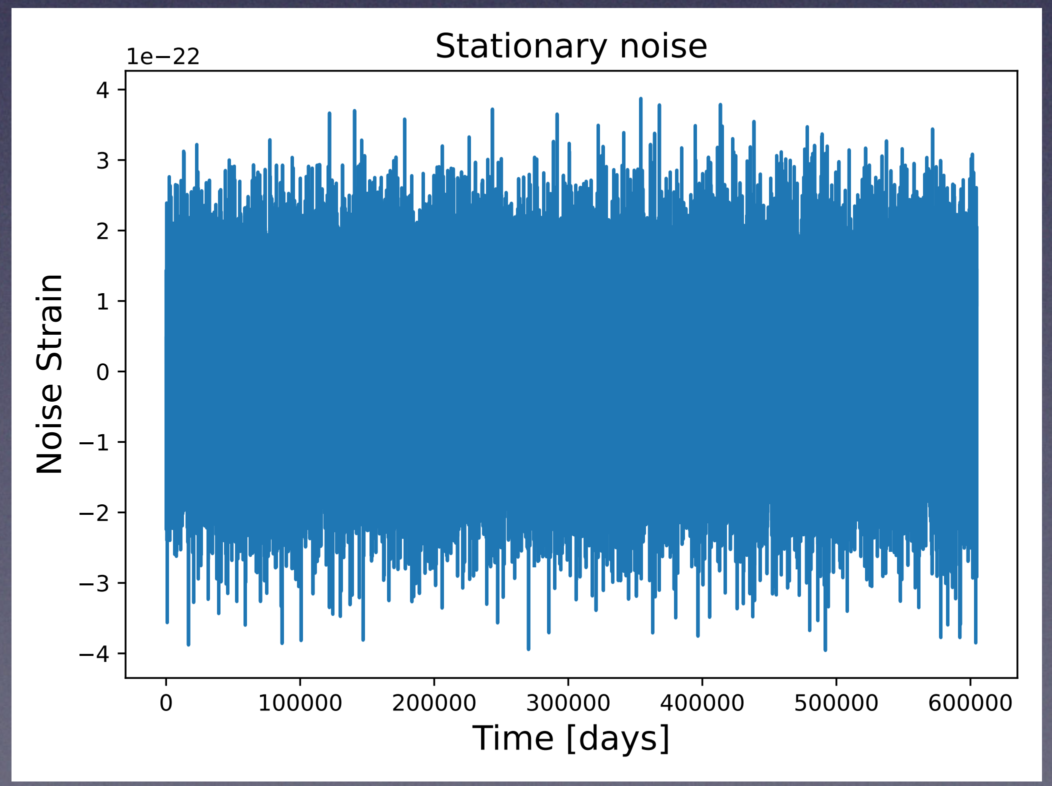
Signal: Deterministic

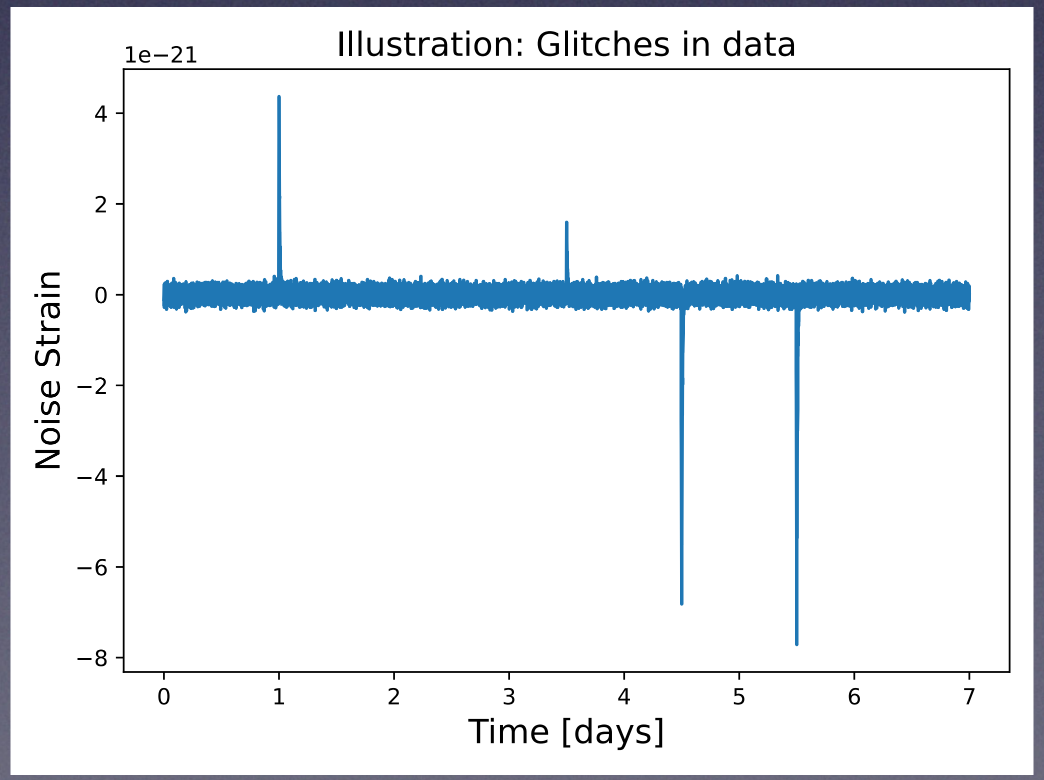
Noise: Probabilistic

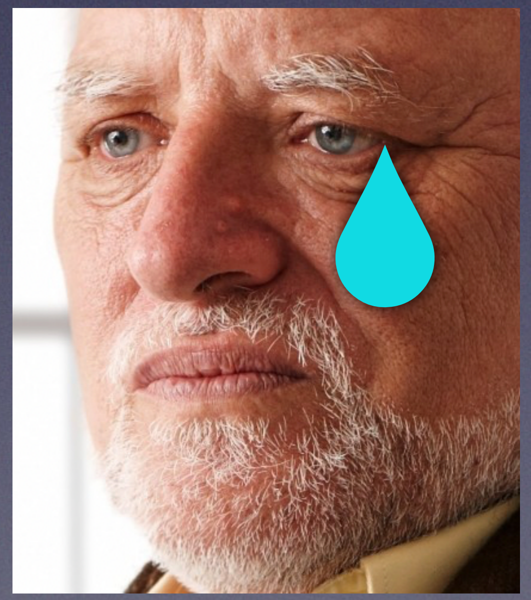
Probabilistic quantity ⟺ Probabilistic Models

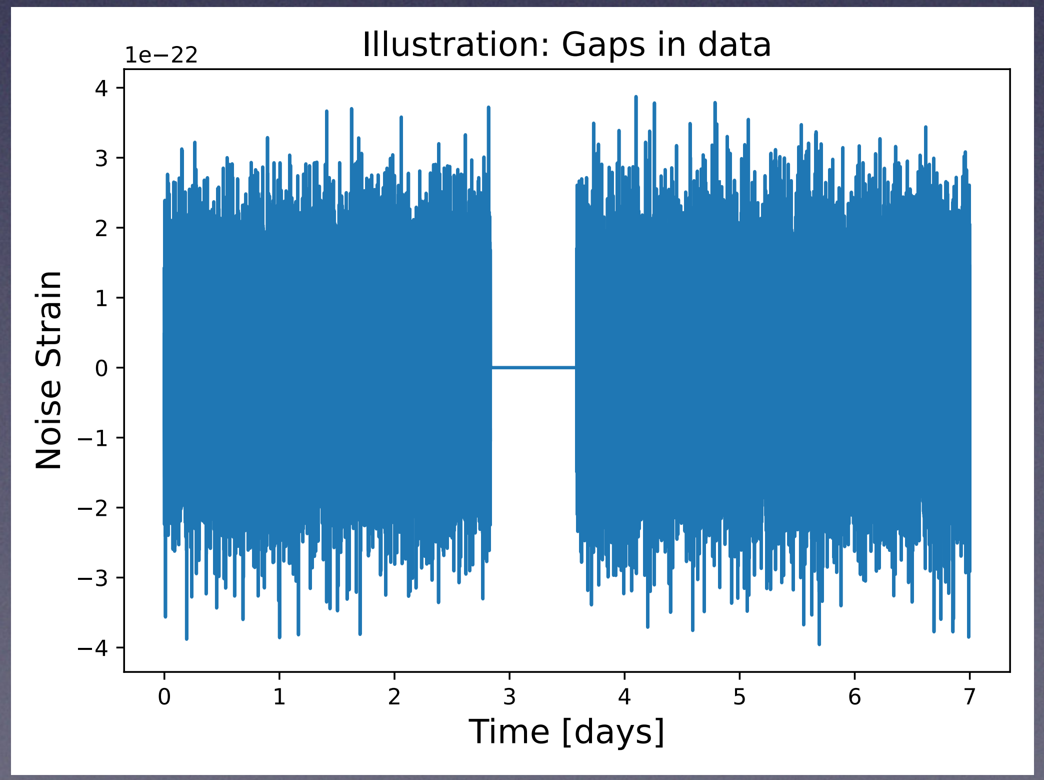
Noise

- Assume $n(t)$ is stochastic, Gaussian with finite mean and variance that **do not** depend on time.
- Further assume that covariance only depends on the lag.









The likelihood

Data stream $d(t) = h_e(t;\boldsymbol{\theta}) + n(t)$, with $n(t) \sim N(0,\Sigma_n)$

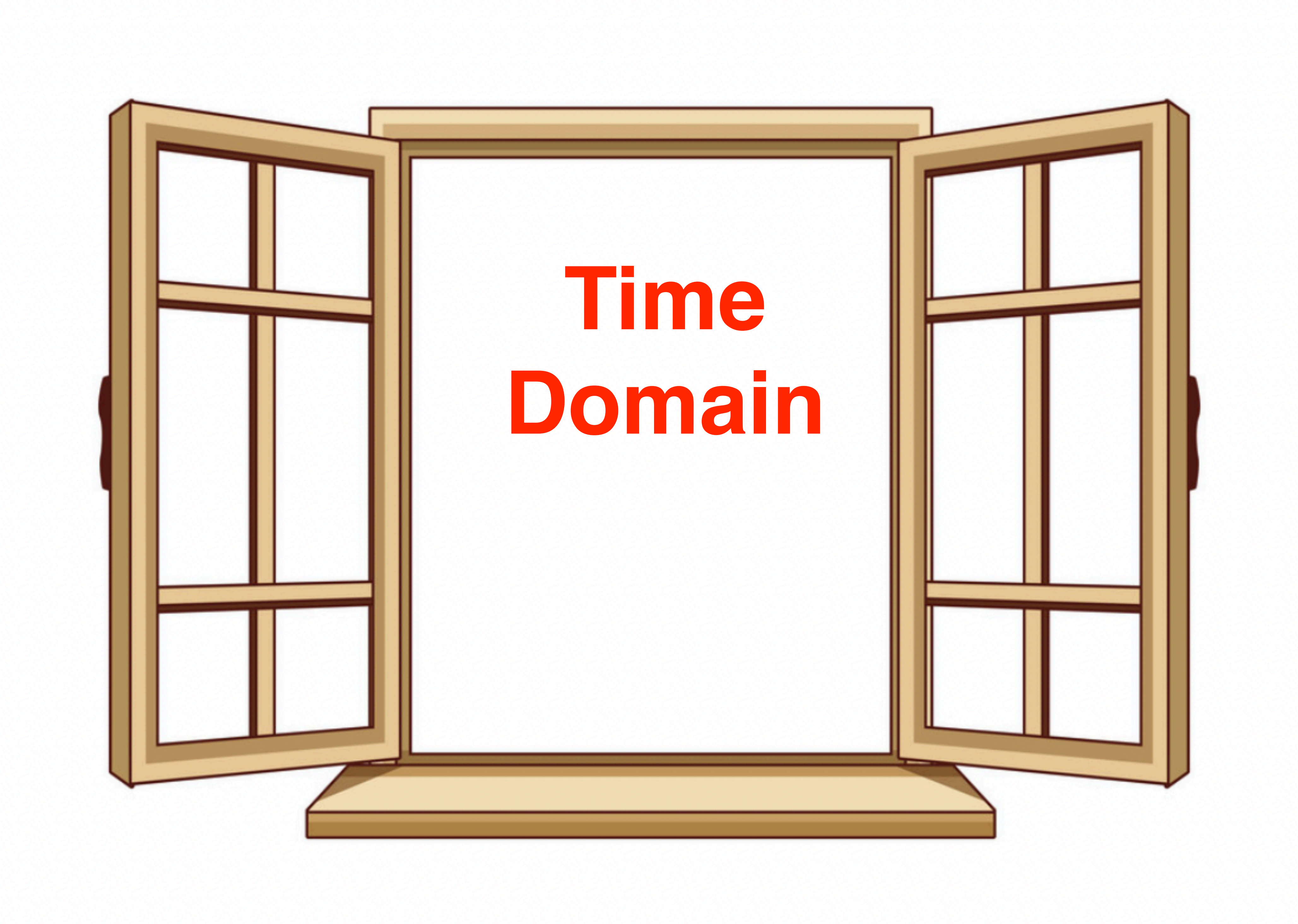
(Log) Likelihood: $\log \mathscr{L}(d\,|\,\boldsymbol{\theta}) \propto (\boldsymbol{d}(t) - \boldsymbol{h}_m(t;\boldsymbol{\theta}))^T \Sigma_n^{-1} (\boldsymbol{d}(t) - \boldsymbol{h}_m(t;\boldsymbol{\theta}))$

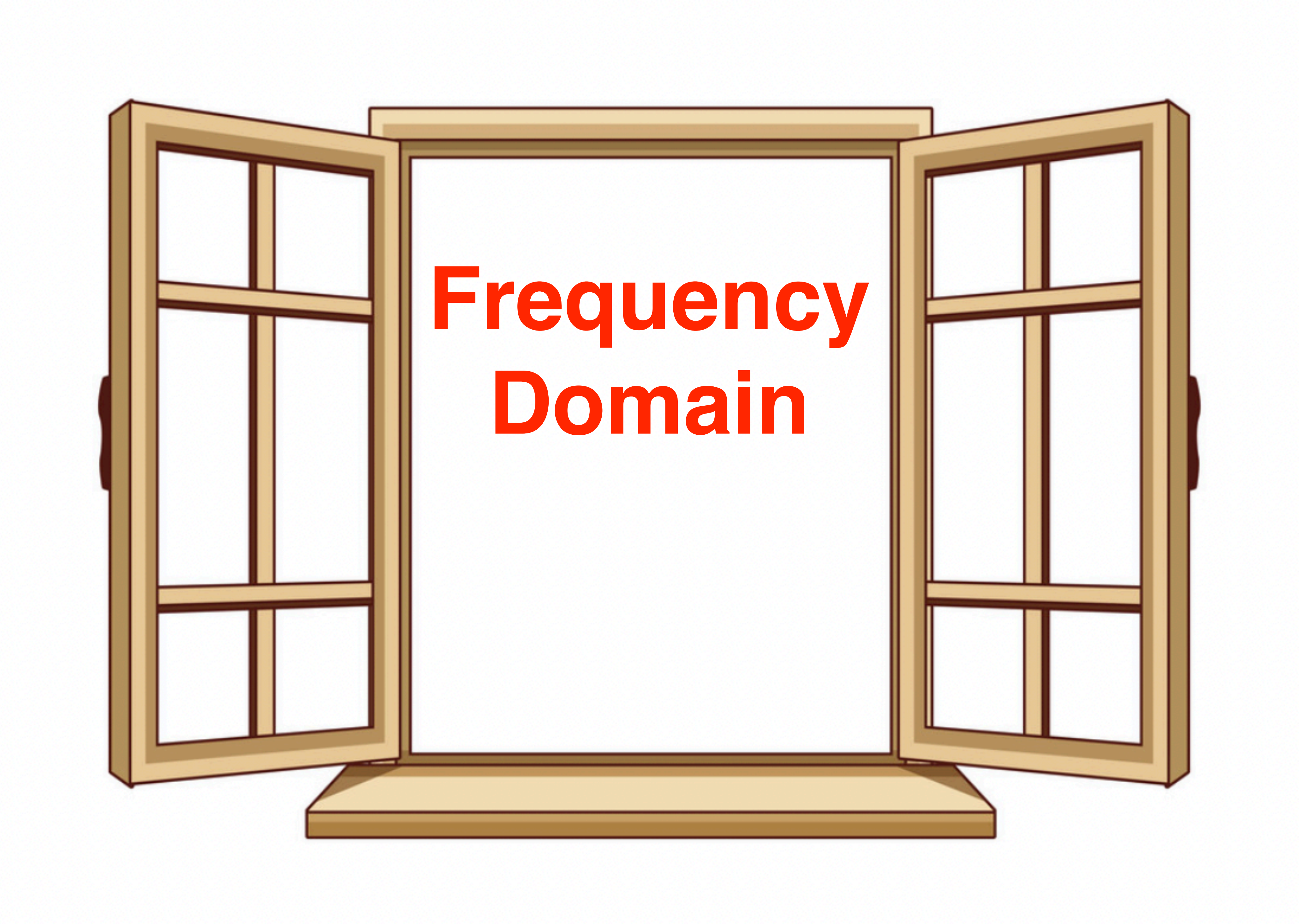
Evil noise co-variance Matrix

Model templates

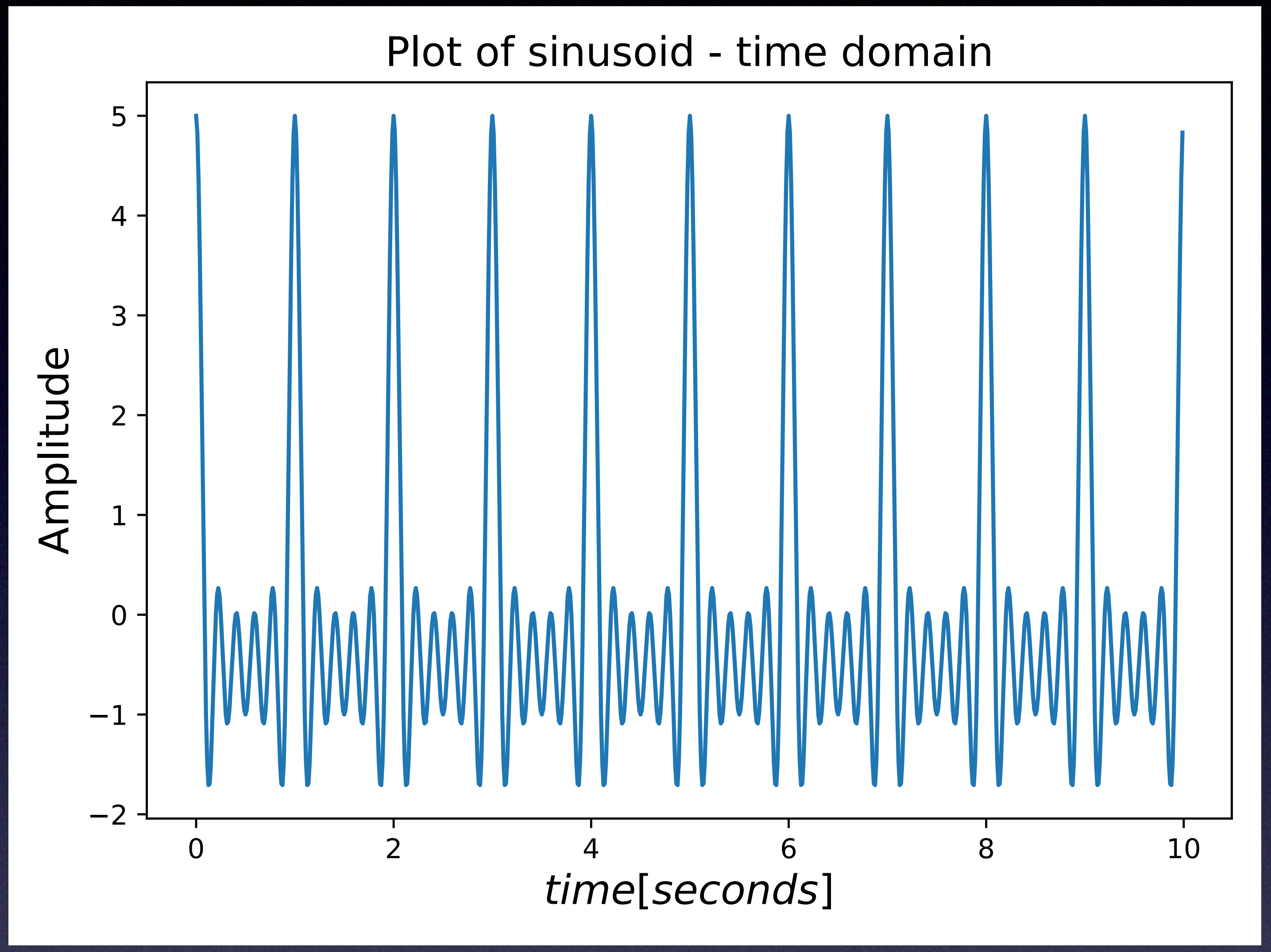
Goal: Identify "best" **template** $h_m(t;\boldsymbol{\theta})$ that **matches** $h_e(t;\boldsymbol{\theta})$ in data $d(t)$

The joy of the frequency domain

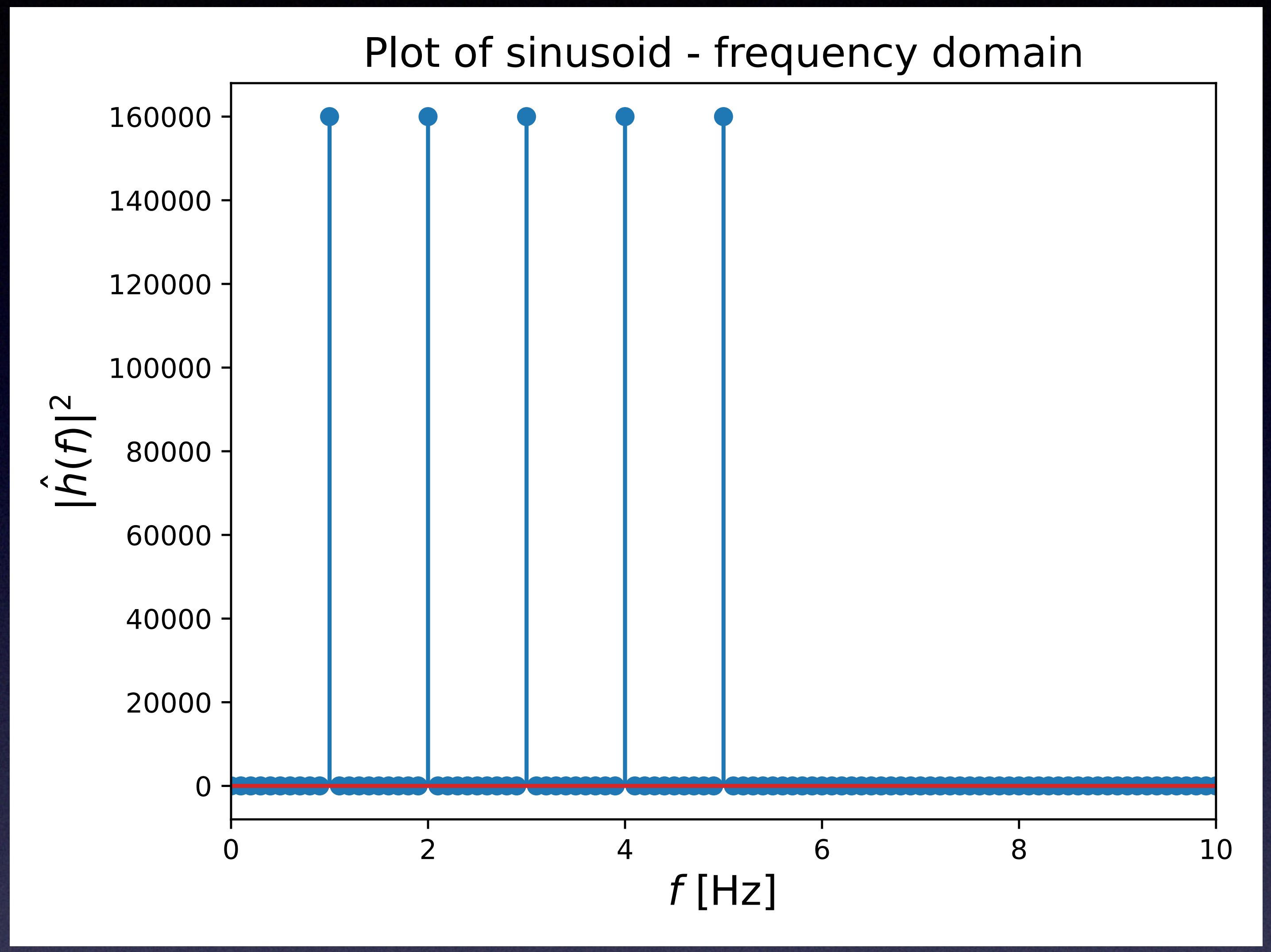


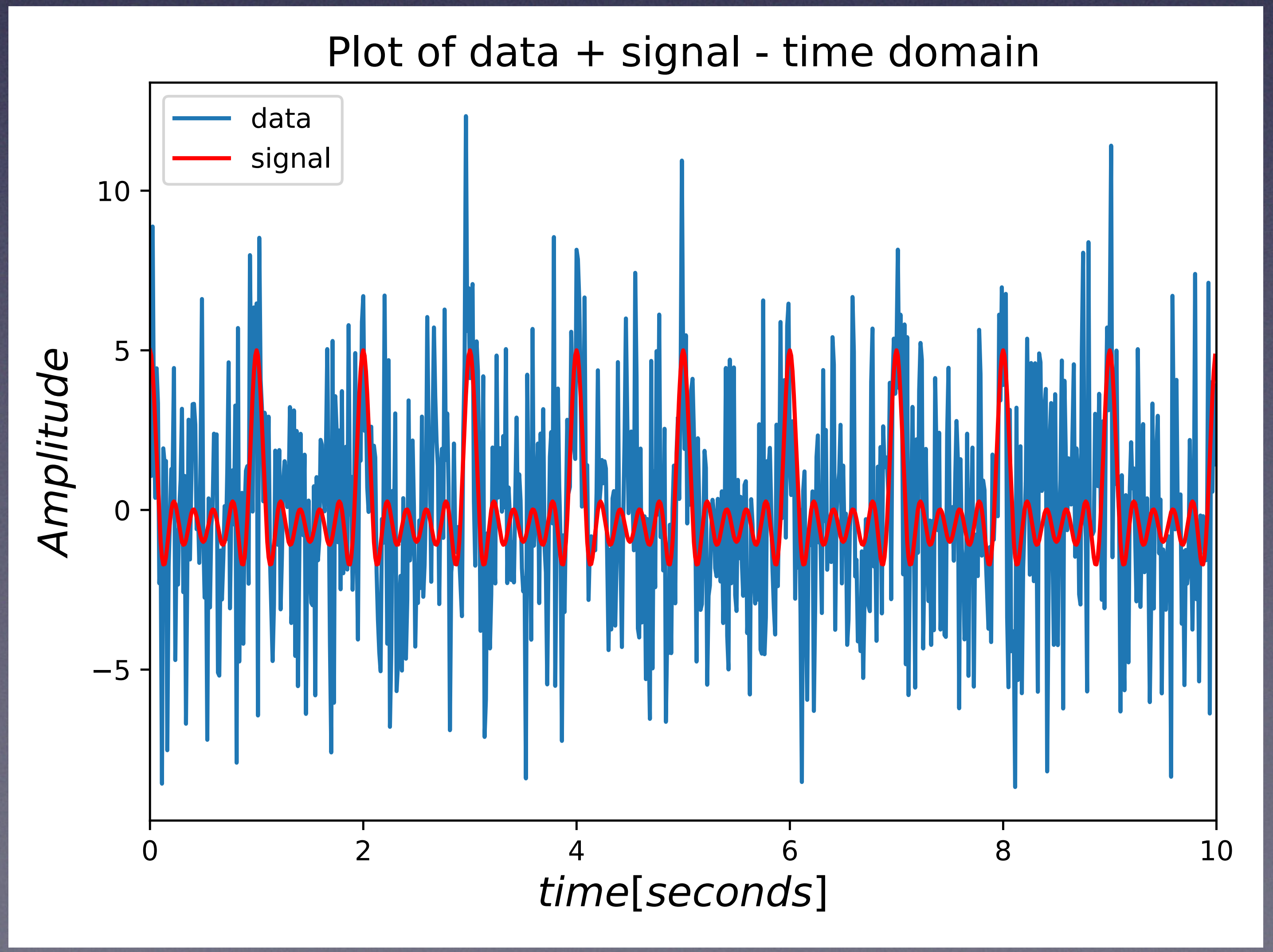


Signals from a different perspective

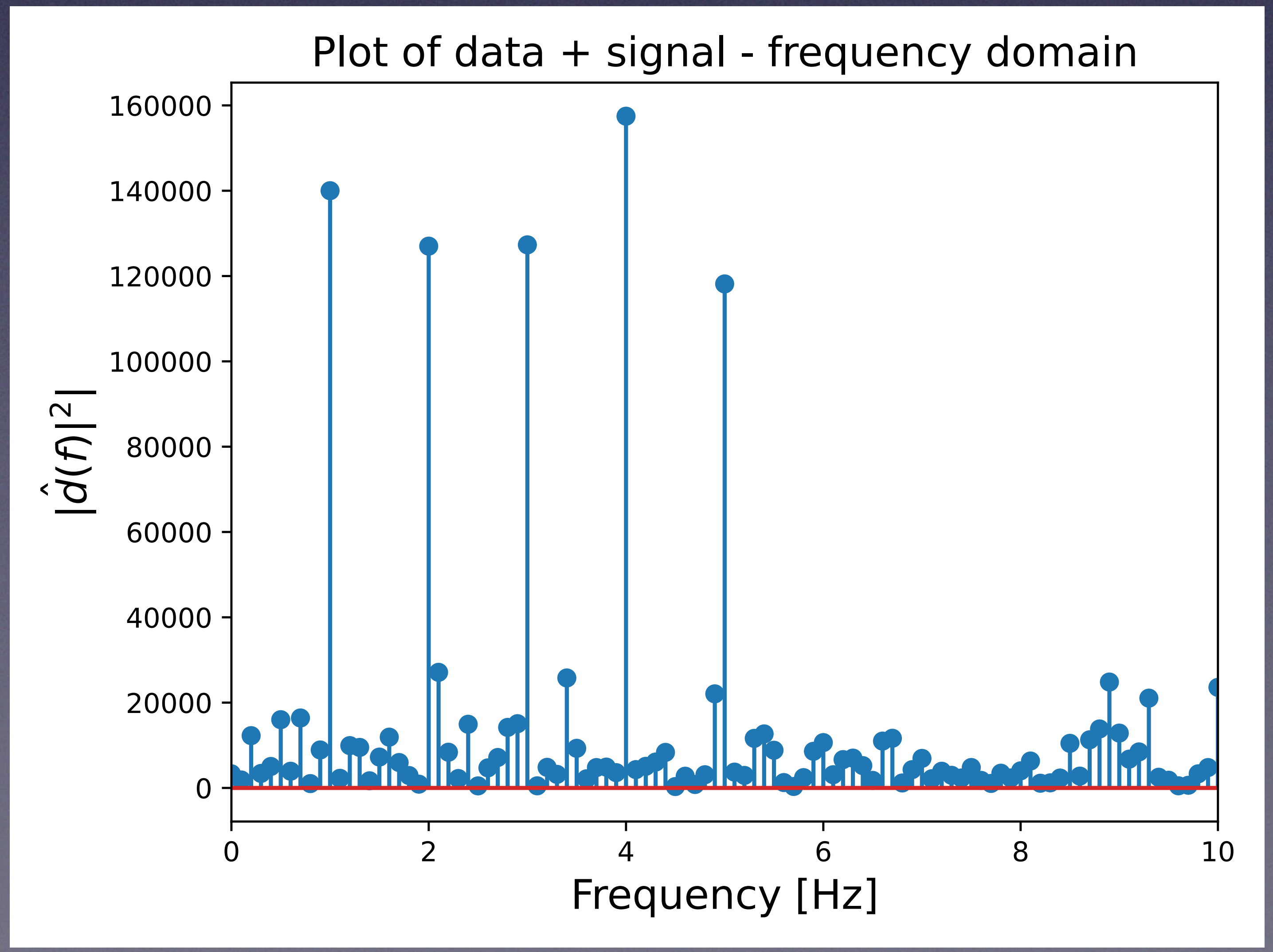


Fourier transform





Fourier transform



Likelihood time/frequency domain

Time domain

Stationary noise $\iff$ Dense Σ_n

$$\log \mathcal{L}(d|\boldsymbol{\theta}) \propto (\boldsymbol{d}(t) - \boldsymbol{h}_m(t;\boldsymbol{\theta}))^T \Sigma_n^{-1} (\boldsymbol{d}(t) - \boldsymbol{h}_m(t;\boldsymbol{\theta}))$$

- **Advantage**
  - Relatively clean - no artefacts!
- **Disadvantage**
  - Expensive - $\mathcal{O}(N^2)$

Frequency domain

Stationary noise $\iff$ Diagonal Σ_n

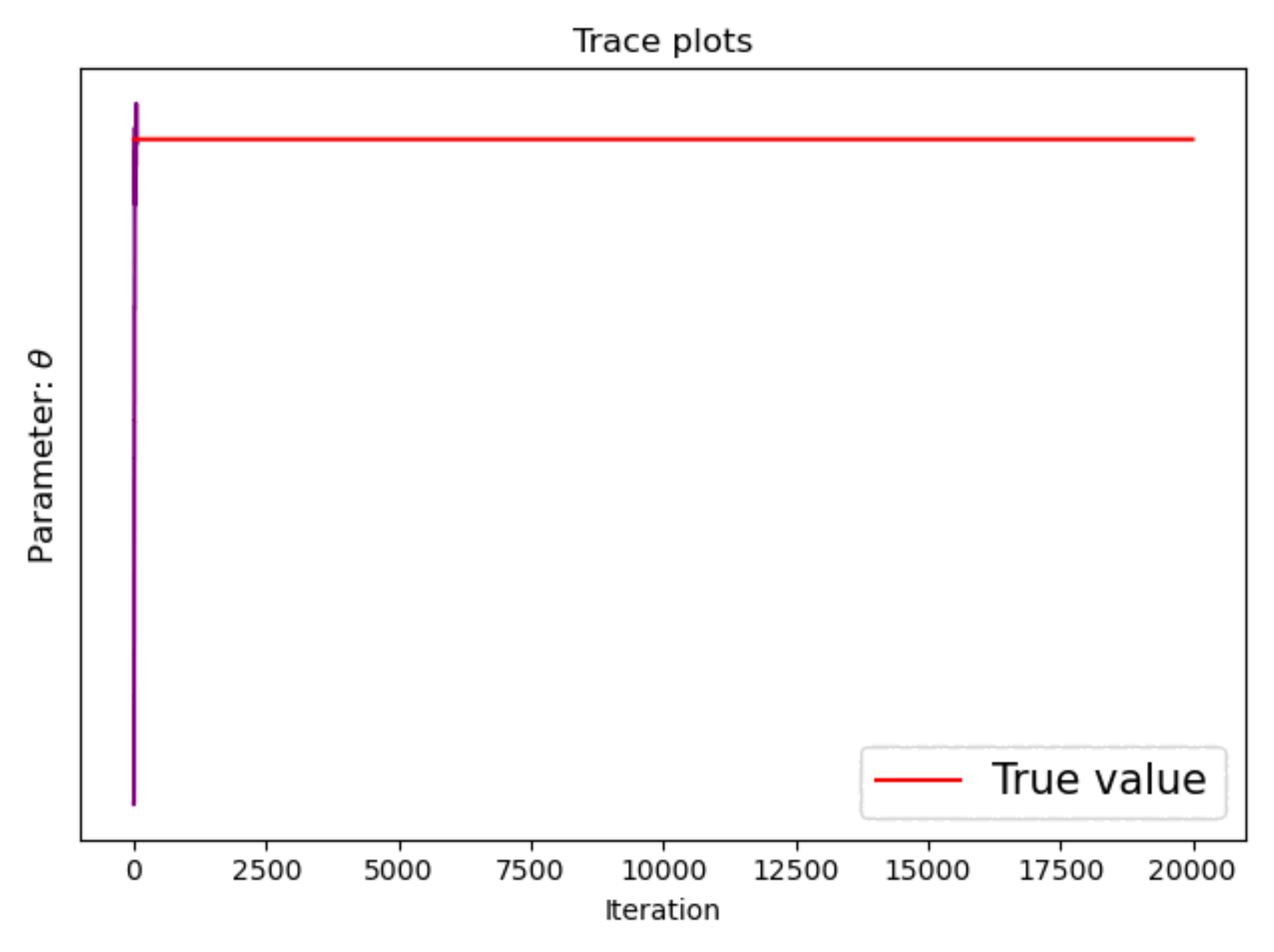
$$\log \mathcal{L}(d|\boldsymbol{\theta}) \propto -2 \sum_i \frac{|\hat{d}(f_i) - \hat{h}_m(f_i;\boldsymbol{\theta})|^2}{S_n(f_i)} \Delta f$$

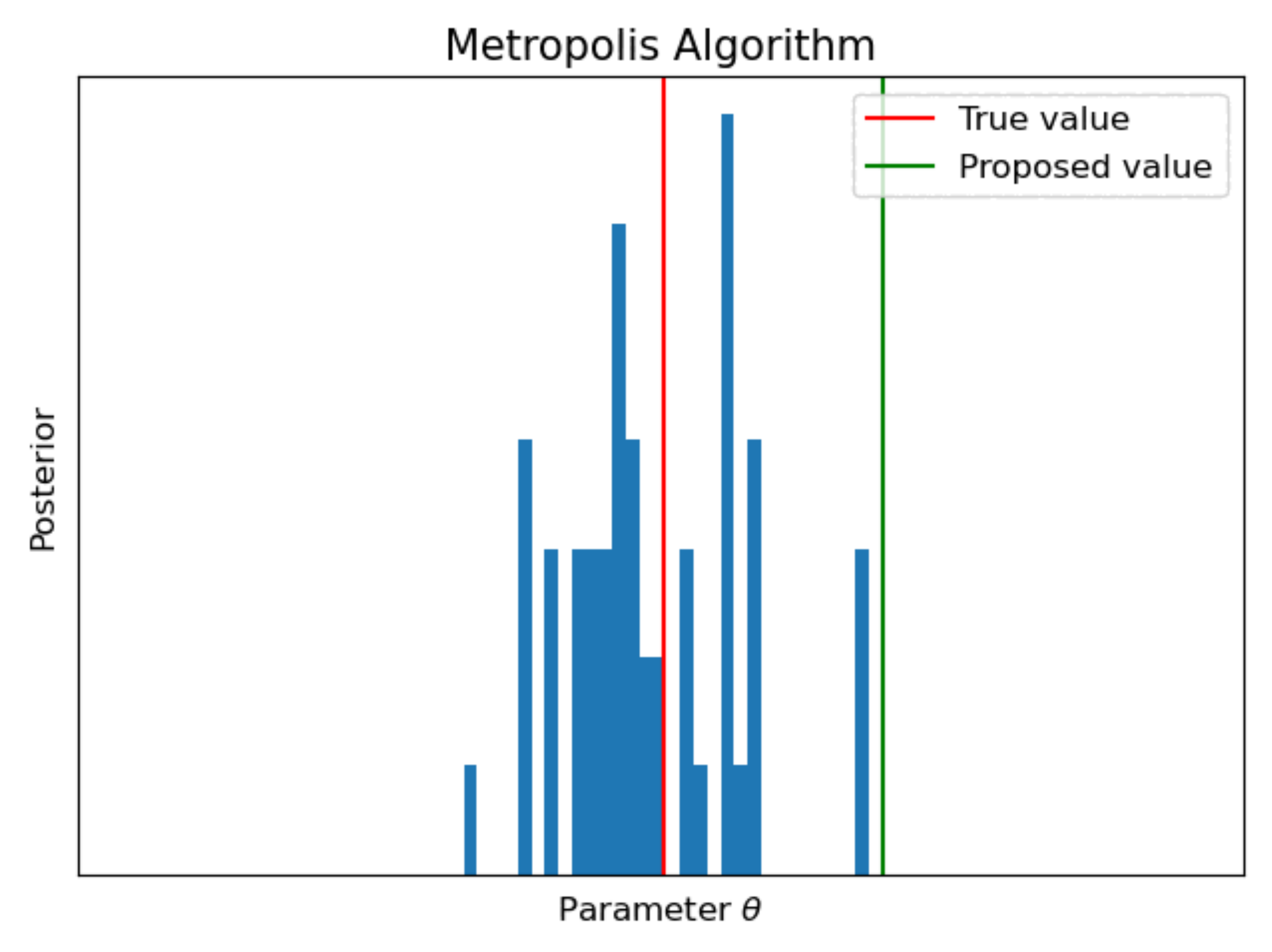
- **Advantage**
  - Cheap: $\mathcal{O}(N \log_2 N)$
- **Disadvantage**
  - Subject to sampling error.

(Whittle, 1953, IoMS)

Sampling from the posterior

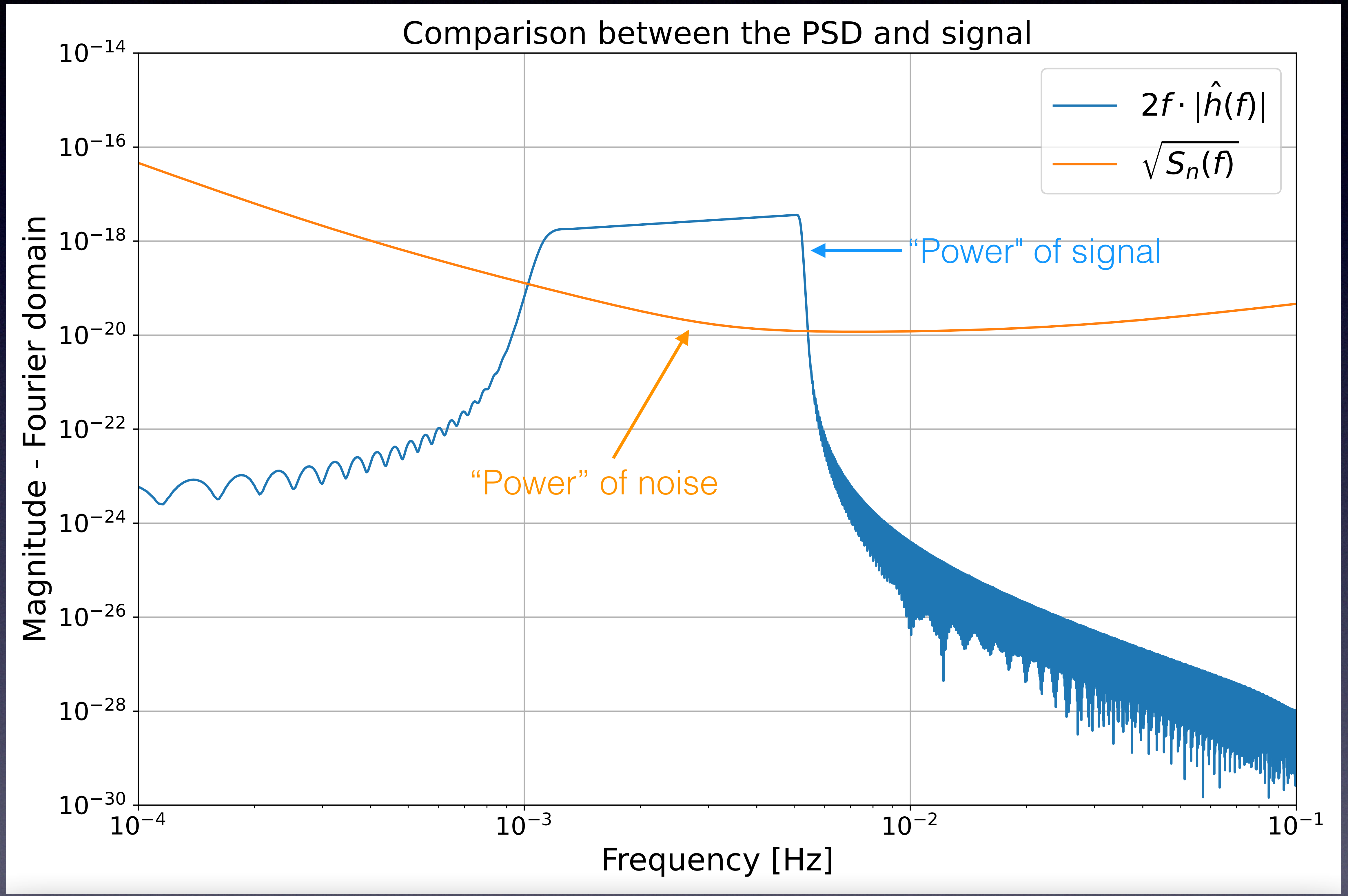
1. Initialise parameter values $\boldsymbol{\theta}^0$
2. For $i = 1, \ldots, N$
3. Draw new point $\boldsymbol{\phi} \sim q(\boldsymbol{\phi} \,|\, \boldsymbol{\theta}^{i-1})$
4. Compute $\alpha = \min\left(1, \frac{p(\boldsymbol{\phi} \,|\, d)}{p(\boldsymbol{\theta}^{i-1} \,|\, d)}\right)$
    1. Accept $\boldsymbol{\phi}$ with probability α. Set $\boldsymbol{\phi} = \boldsymbol{\theta}^i$.
    2. Reject $\boldsymbol{\phi}$ otherwise. Set $\boldsymbol{\theta}^{i-1} = \boldsymbol{\theta}^i$.
5. Increment i by one and return to step 3.





Instructive Example

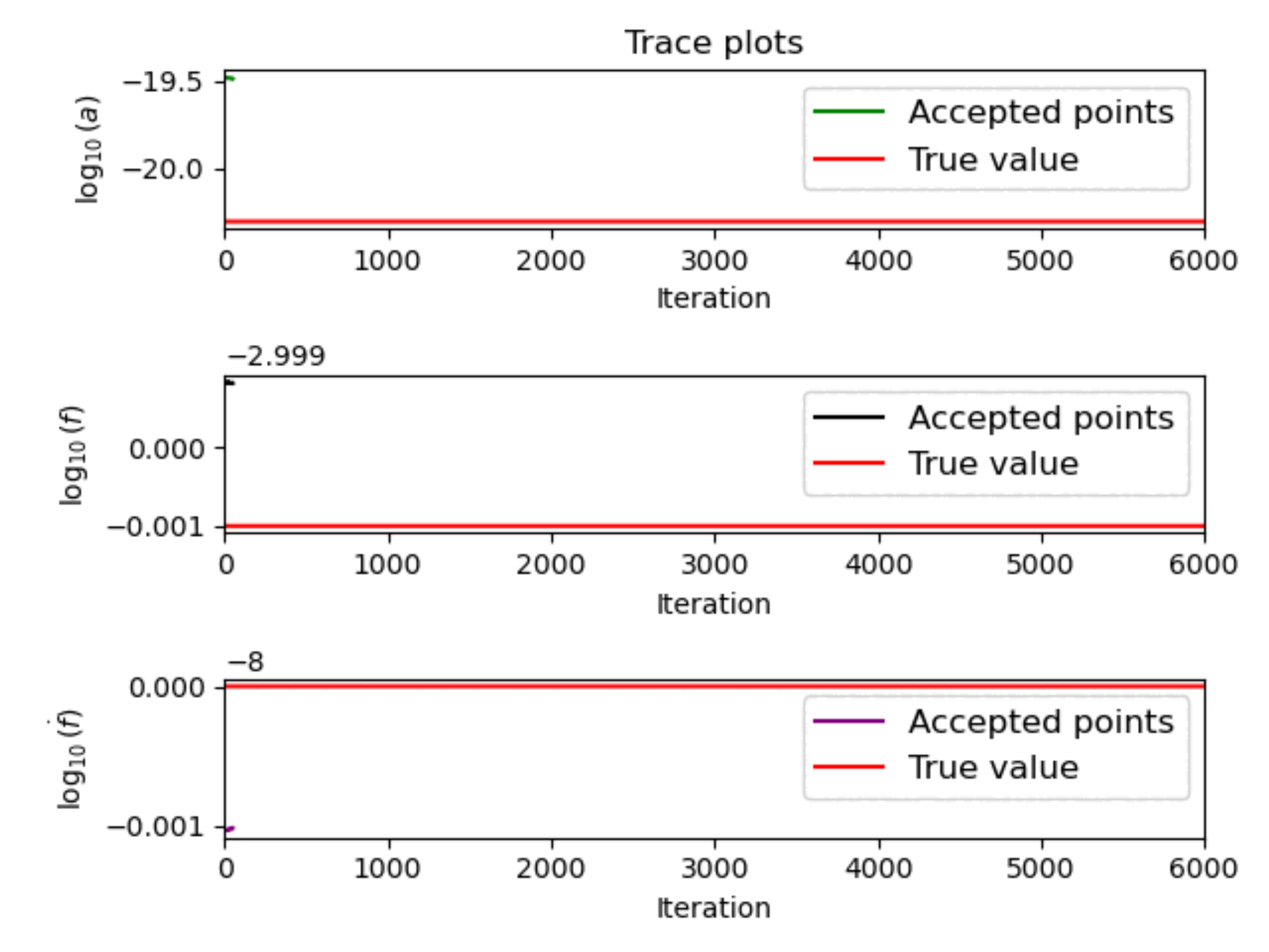
- $h_e(t;\boldsymbol{\theta}) = a\cos[2\pi t(f + \dot{f}t)]$, here $\boldsymbol{\theta} = \{a, f, \dot{f}\}$
- Estimate: $a = 5 \cdot 10^{-21}$, $f = 10^{-3}$ and $\dot{f} = 10^{-8}$
- Duration: 120 hours
- Data: $d(t) = h_e(t;\boldsymbol{\theta}) + n(t)$, noise gaussian + stationary
- Likelihood: $\log \mathscr{L}(d\,|\,\boldsymbol{\theta}) \propto -2\sum_i \frac{|\hat{d}(f_i) - \hat{h}_m(f_i;\boldsymbol{\theta})|^2}{S_n(f_i)}\Delta f$
- Uniform priors on parameters

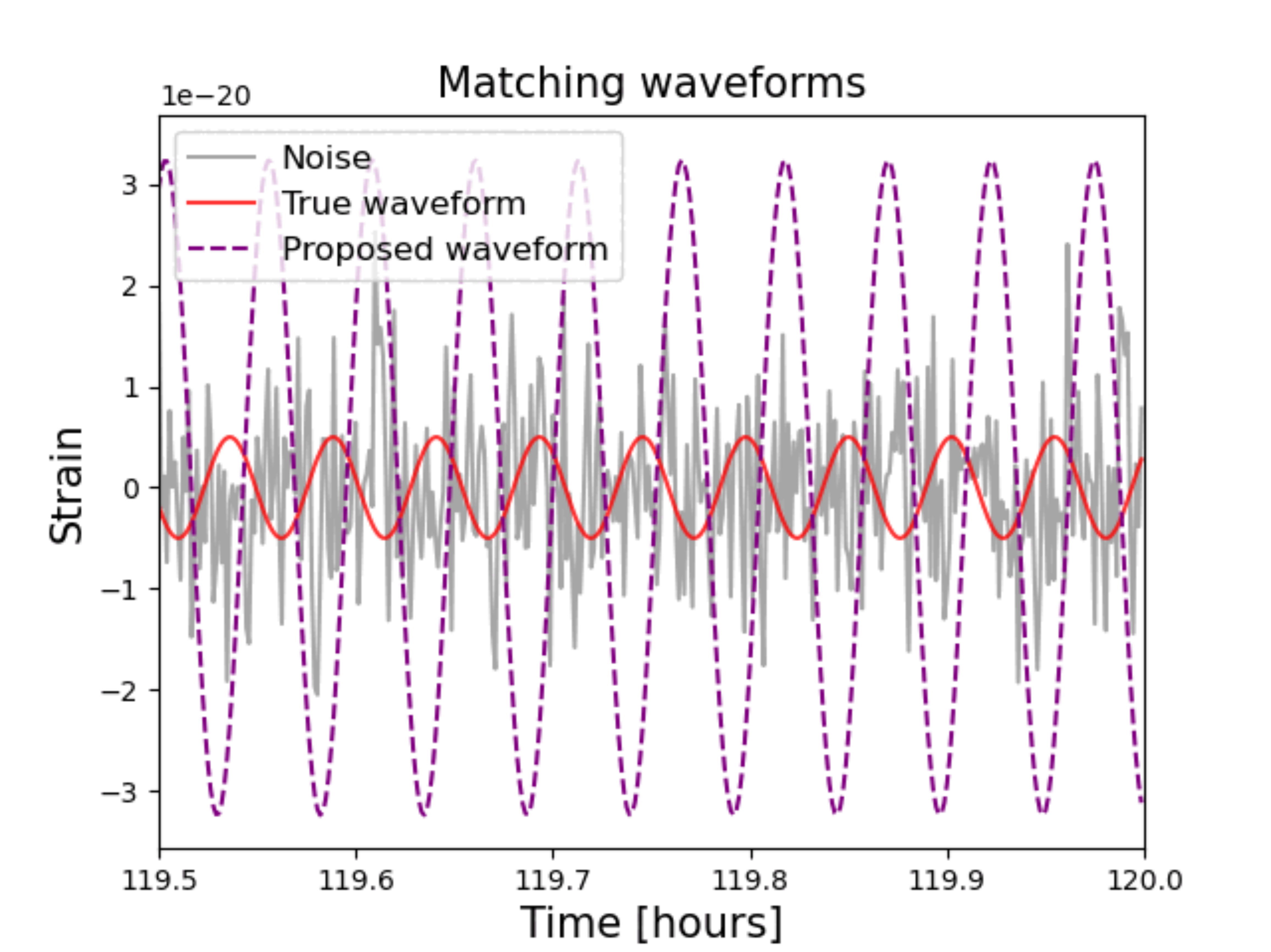


$$\mathrm{SNR}^2 = 4\int_0^\infty \frac{f\,|\hat{h}(f)|^2}{S_n(f)}\, d\log(f)$$

Finding the "best" signal

Goal: Identify parameters $\boldsymbol{\theta}$ that best match the waveform

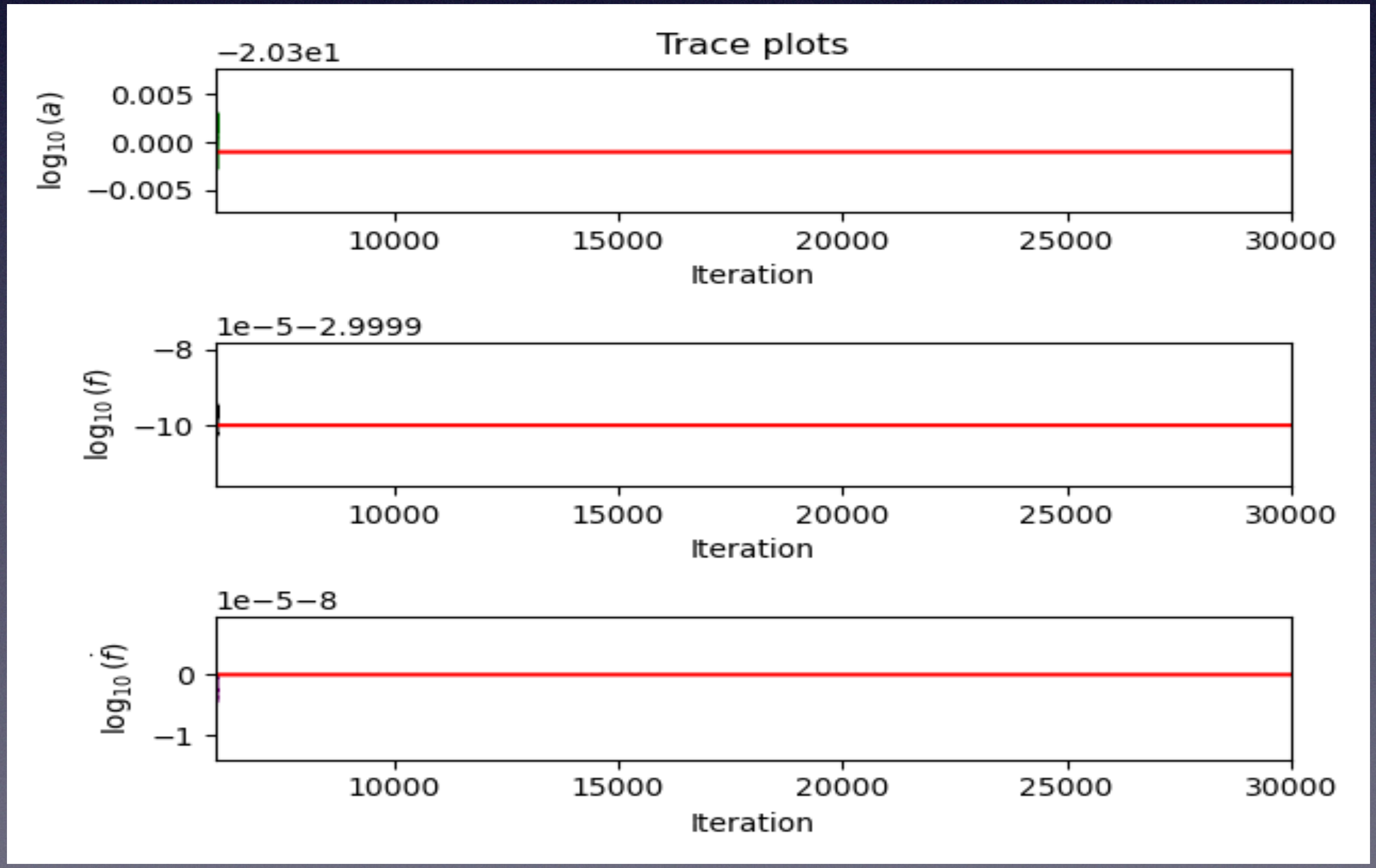


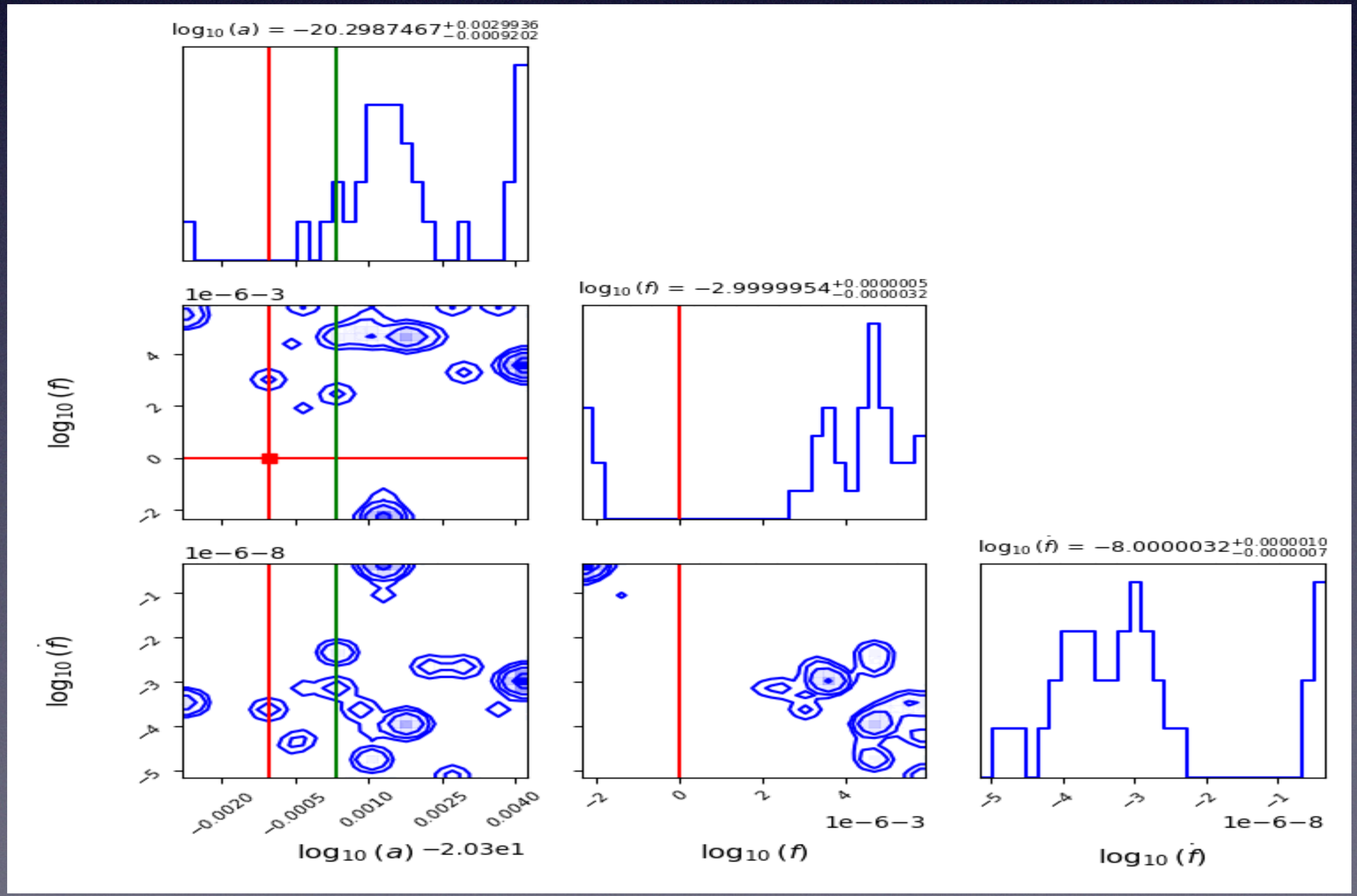


https://github.com/OllieBurke/Tutorials.git

Parameter Estimation

- Identified "best" location in parameter space
- Now sample $\boldsymbol{\theta} \sim p(\boldsymbol{\theta}\,|\,d)$ and explore posterior!
- Make statements on $\boldsymbol{\theta}$ given observed data d





https://github.com/OllieBurke/Tutorials.git

Part 3: **Real life situations**

Reality Strikes

- **Unknown** noise
- **Multimodal** likelihood surface
- **Expensive** and **imperfect** waveform models
- **Multiple unknown sources** in the data
- **Gaps**, **glitches** and **~~aliens~~** other gremlins in data

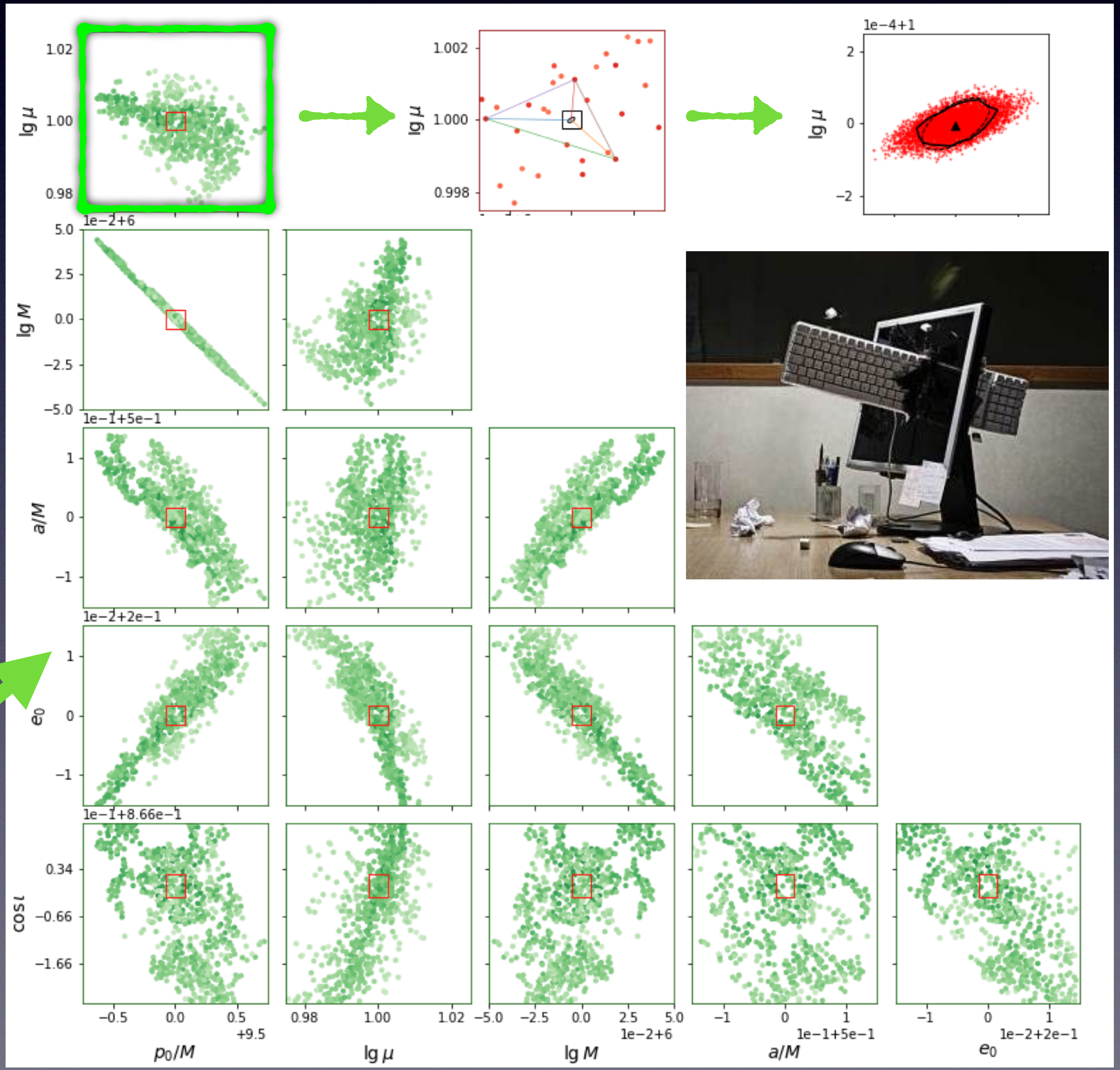
We require sophisticated samplers to account for each obstacle

Multimodal likelihood surface

Samplers **must** be able to tackle the issue of multimodality!

Multimodal likelihood surface

(Chua, Cutler, 2022)

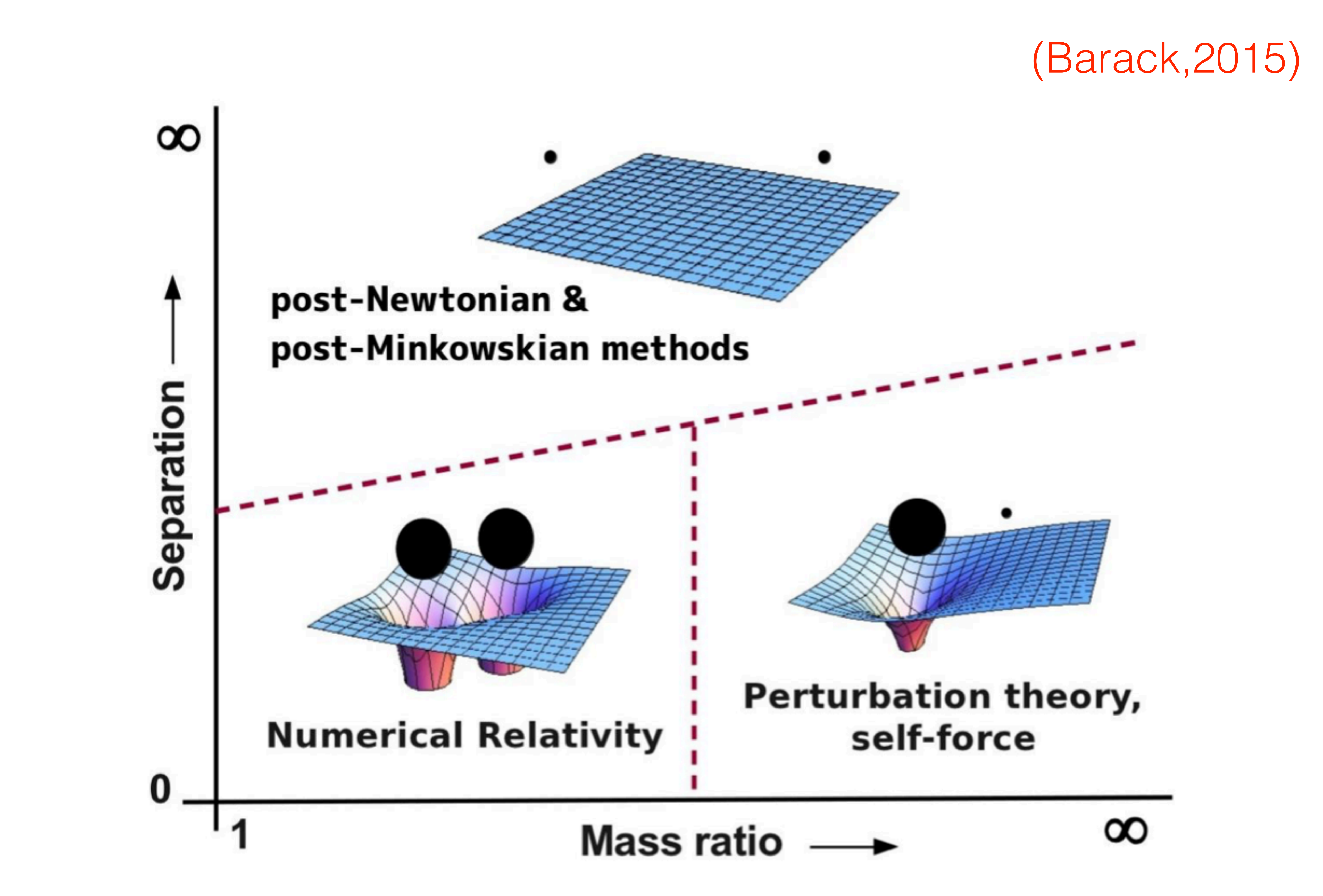


Special place in hell for people who enjoy this stuff

Example EMRI Posterior

Imperfect Waveforms

- Waveforms are approximate
- Waveform errors $\Longleftrightarrow$ biases
- Requirement:
  - Accuracy
  - Speed
  - Cover all parameter space

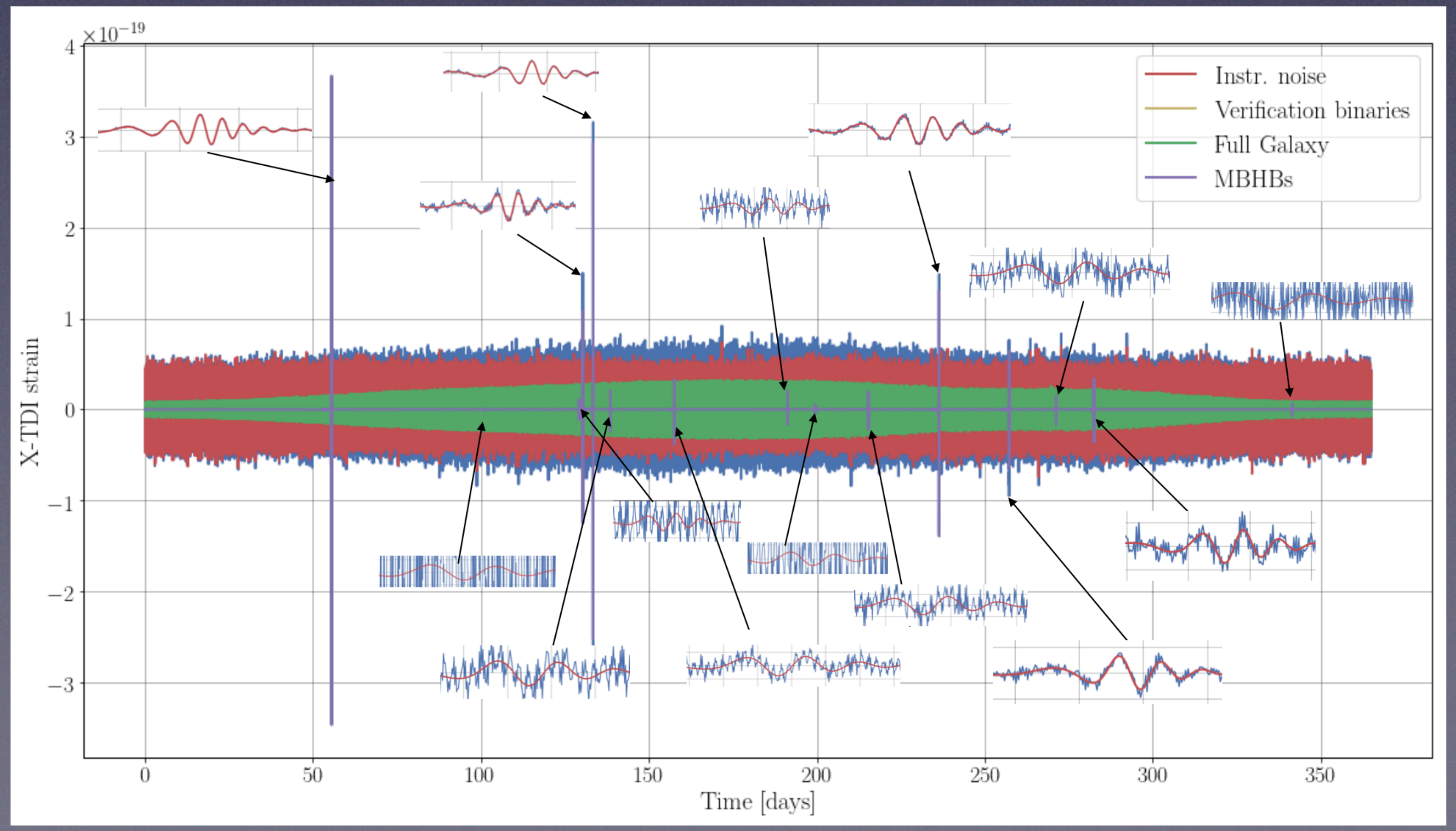


(LDC group)
(https://lisa-ldc.lal.in2p3.fr)

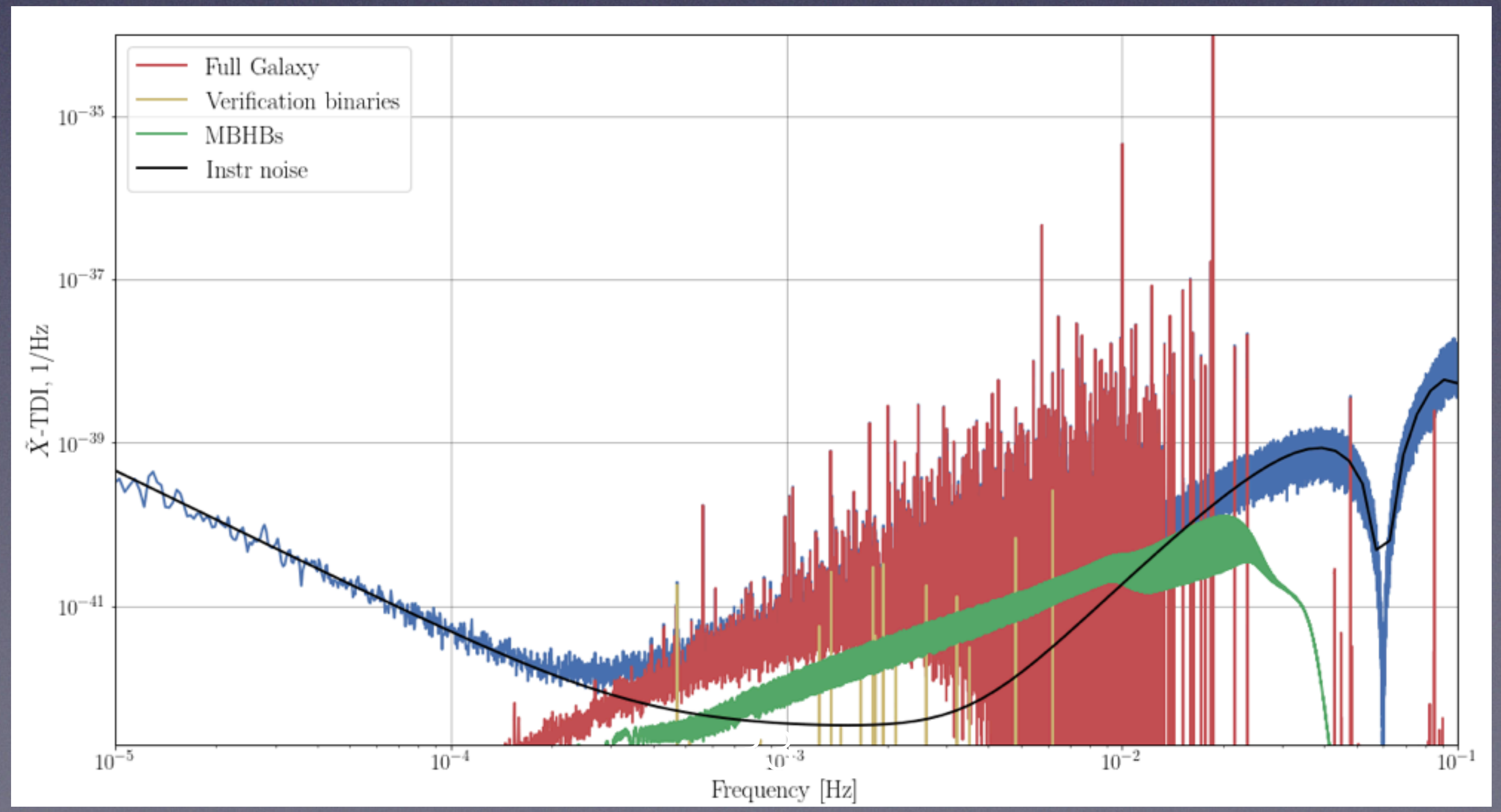
Multiple Waveforms

- Unlike ground-based detectors, LISA will be **dominated** by signals
- Must account for **all** resolvable signals in data

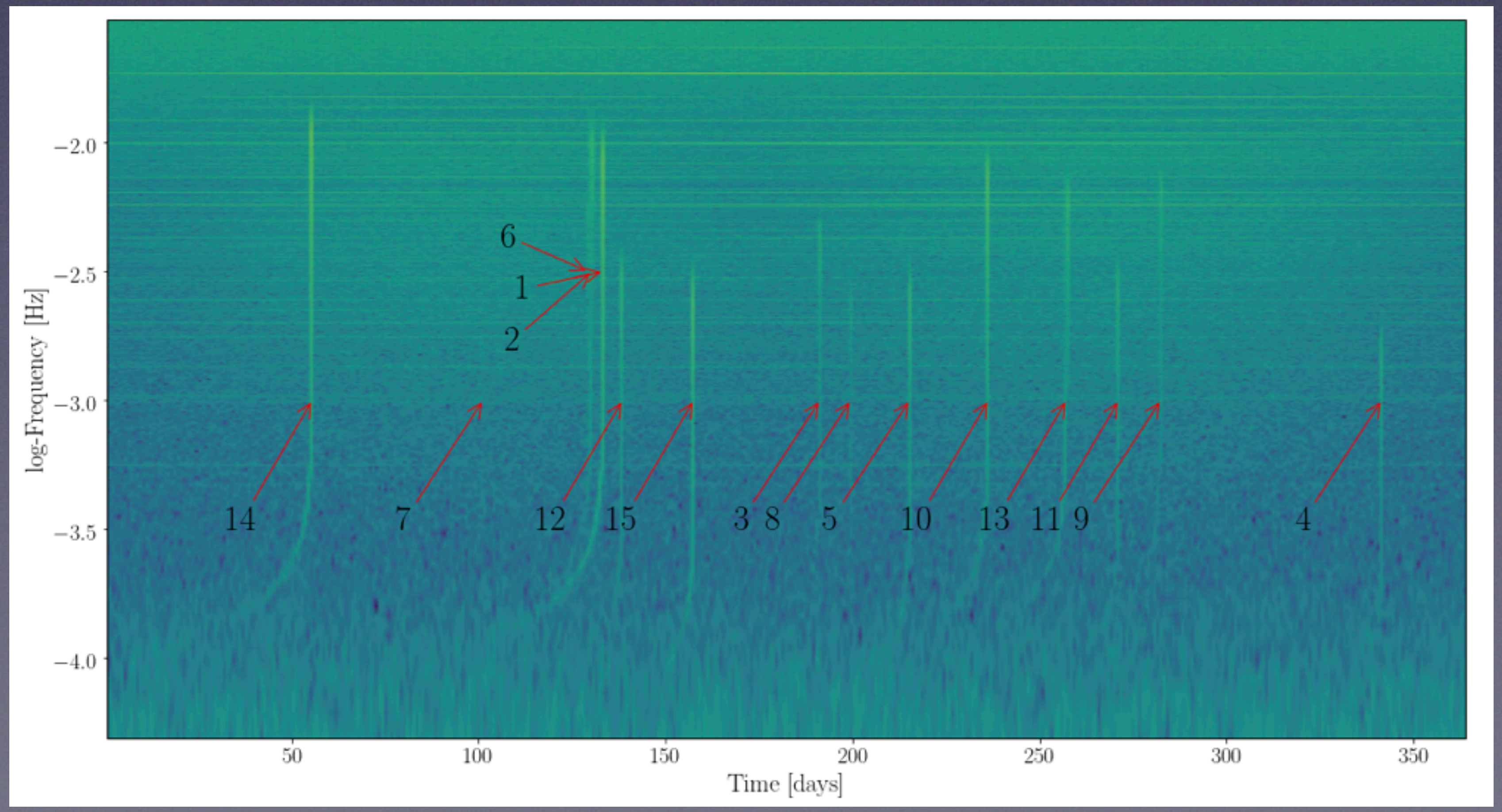
Time domain



Frequency domain



Time/Frequency domain

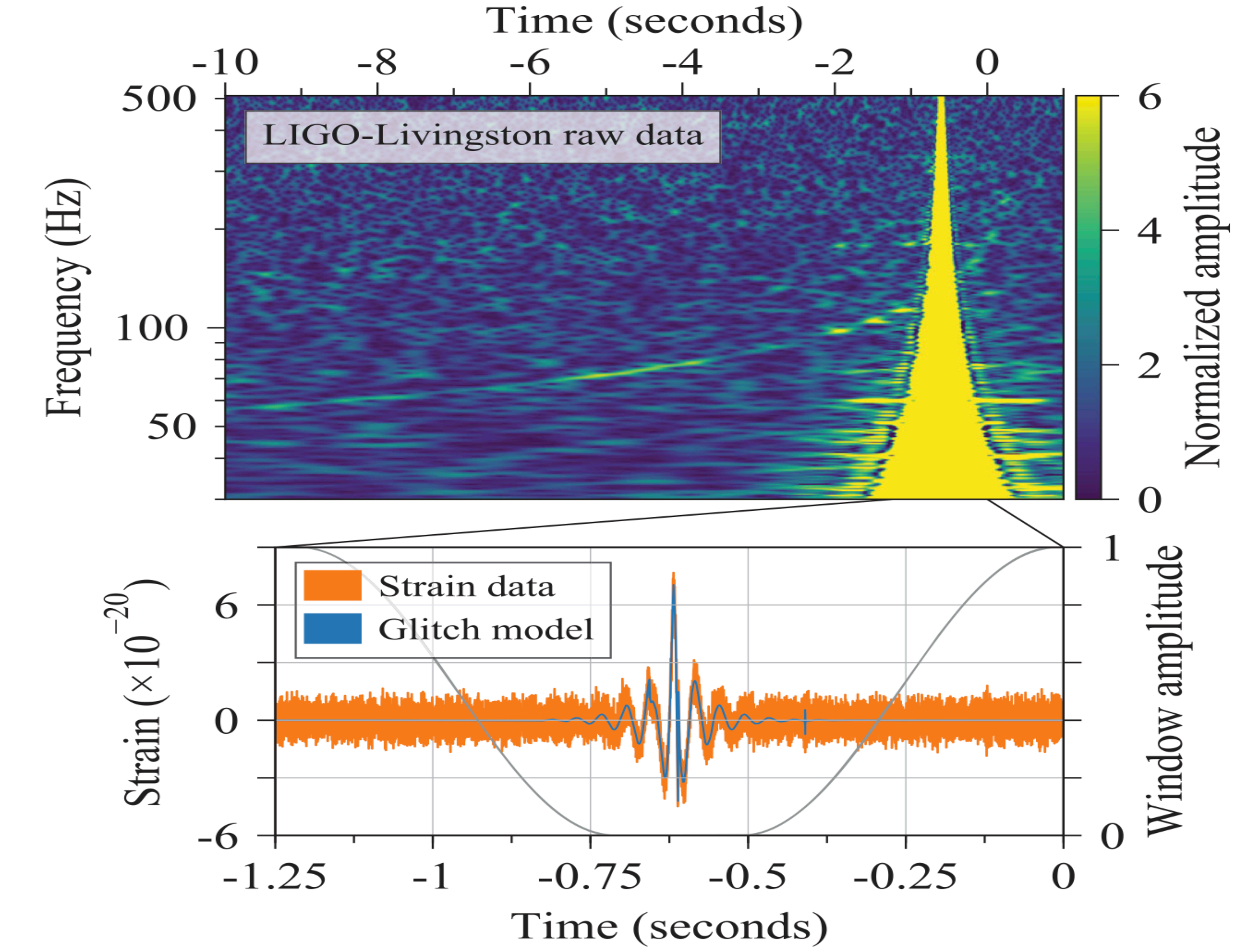


Understanding noise

Noise model $\Longleftrightarrow$ Probabilistic Models

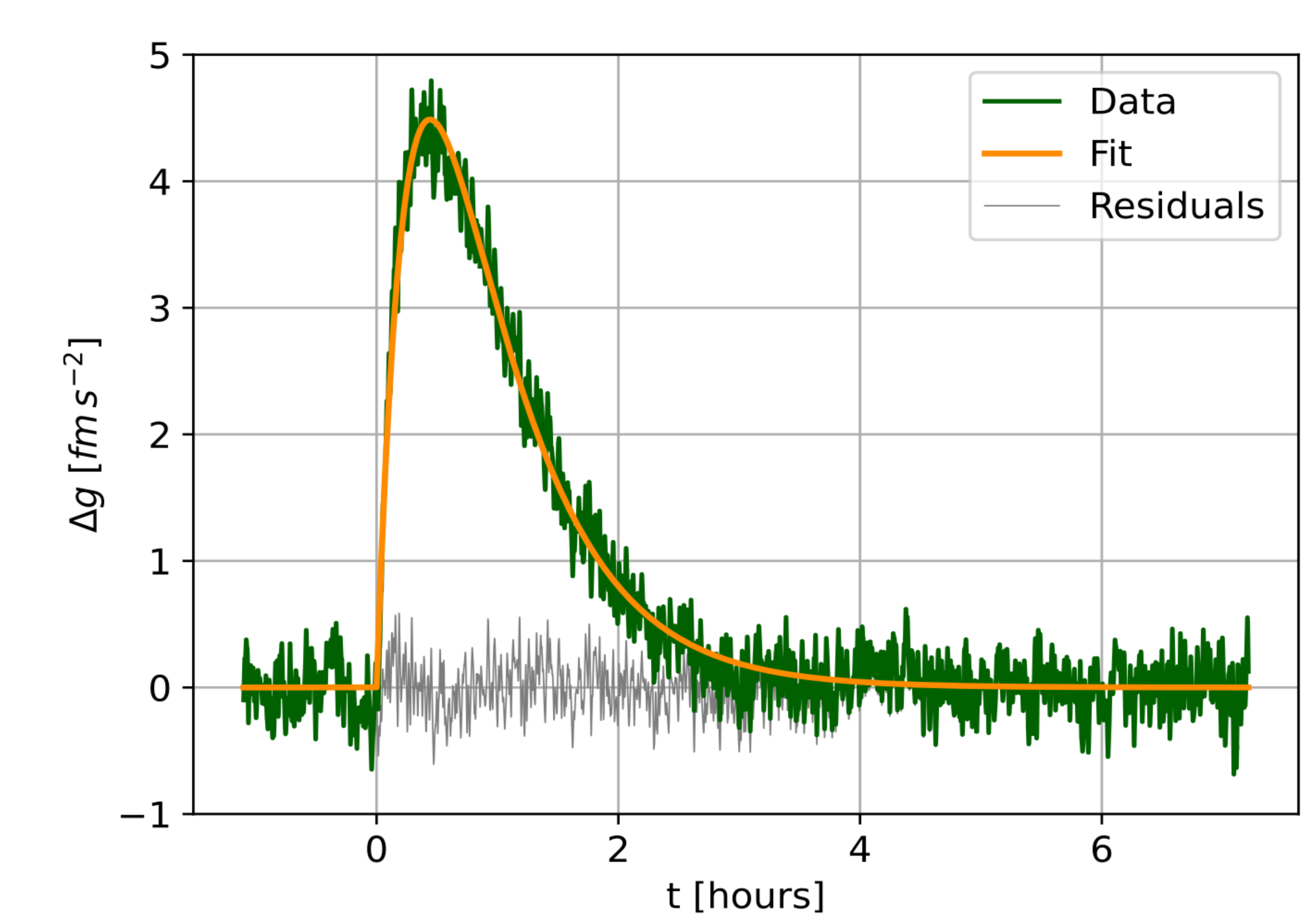
LIGO - glitch

(Credit: LIGO-Virgo Collaboration)



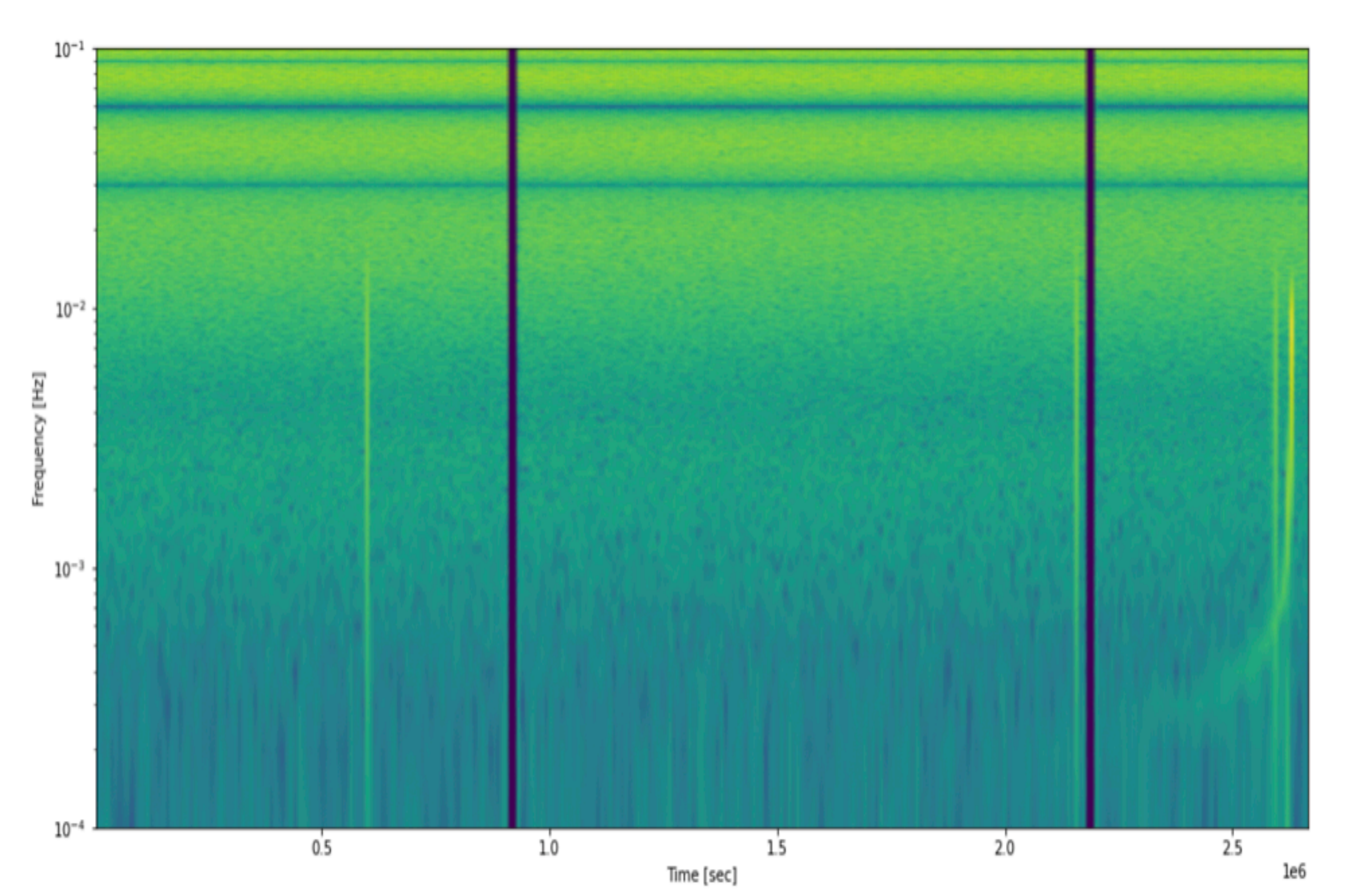
LISA pathfinder - glitch

(Credit: LISA Pathfinder Collaboration)



LISA - gaps and glitch

https://lisa-ldc.lal.in2p3.fr

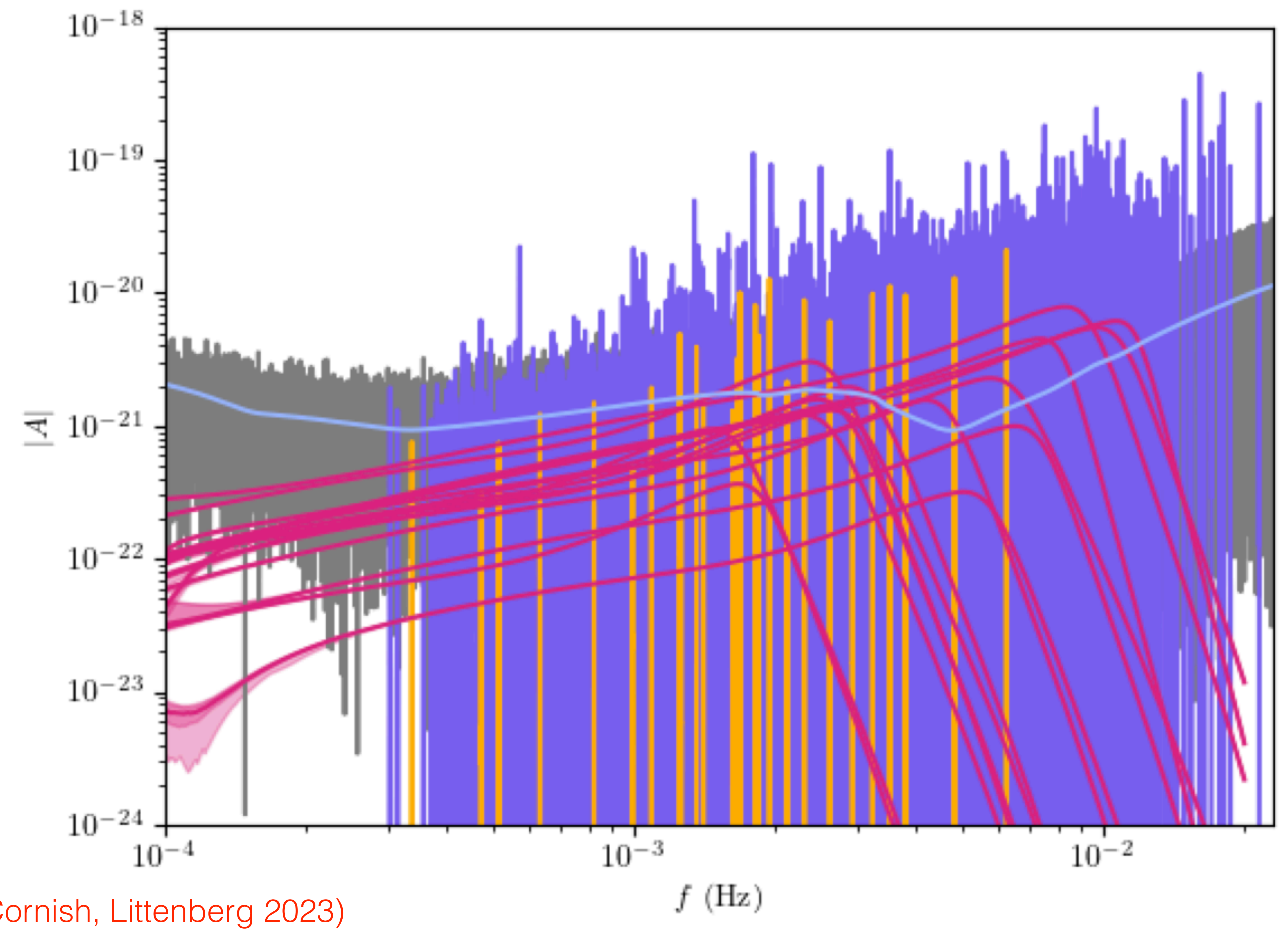


Combination — The global fit

- Global fit: Simultaneous characterisation of **all resolvable signals and noise**

- **Extremely difficult**

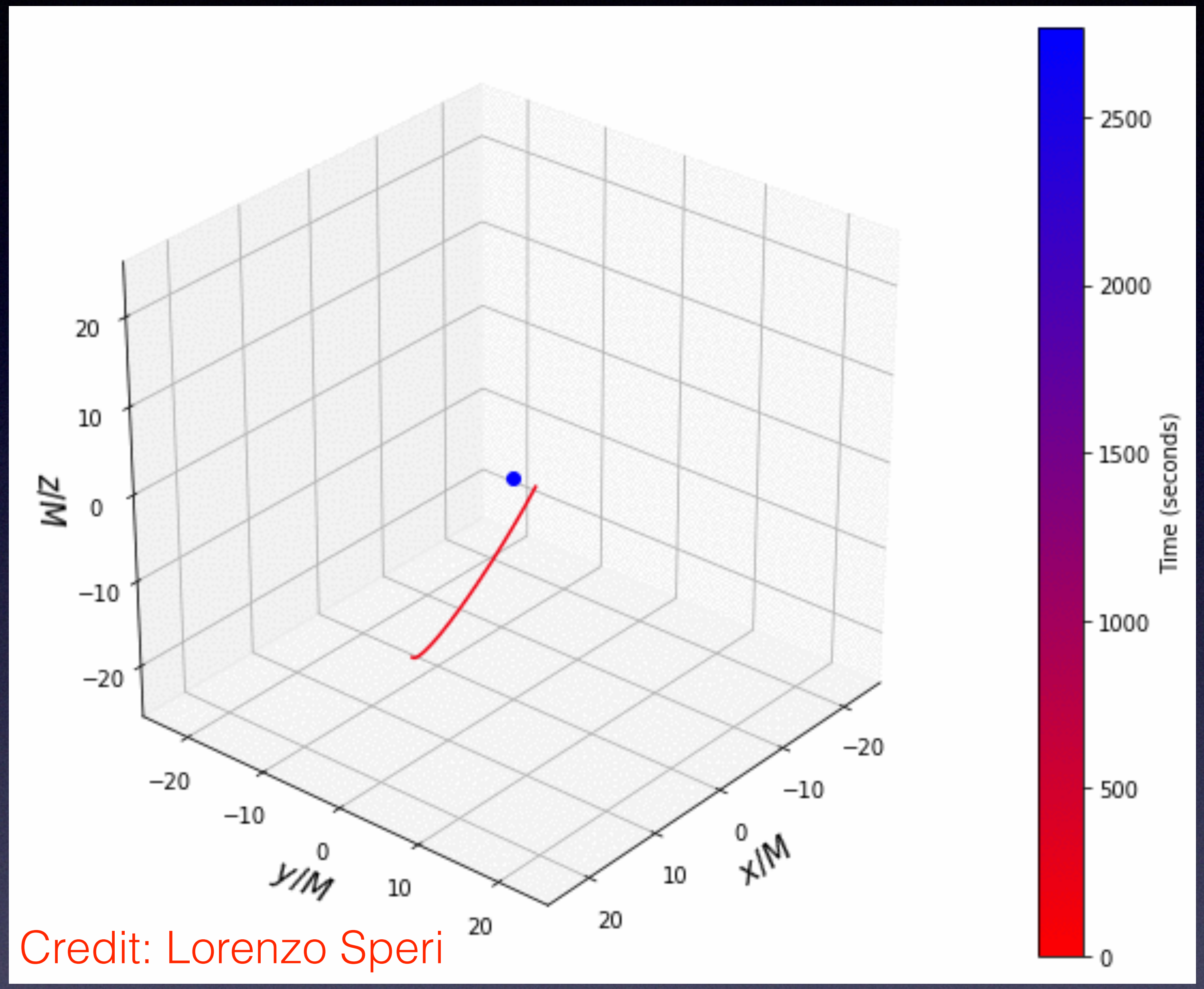
1. Massive black hole binaries
2. Galactic Binaries
3. Verification Galactic Binaries
4. Noise — instrumental + confusion



(Cornish, Littenberg 2023)

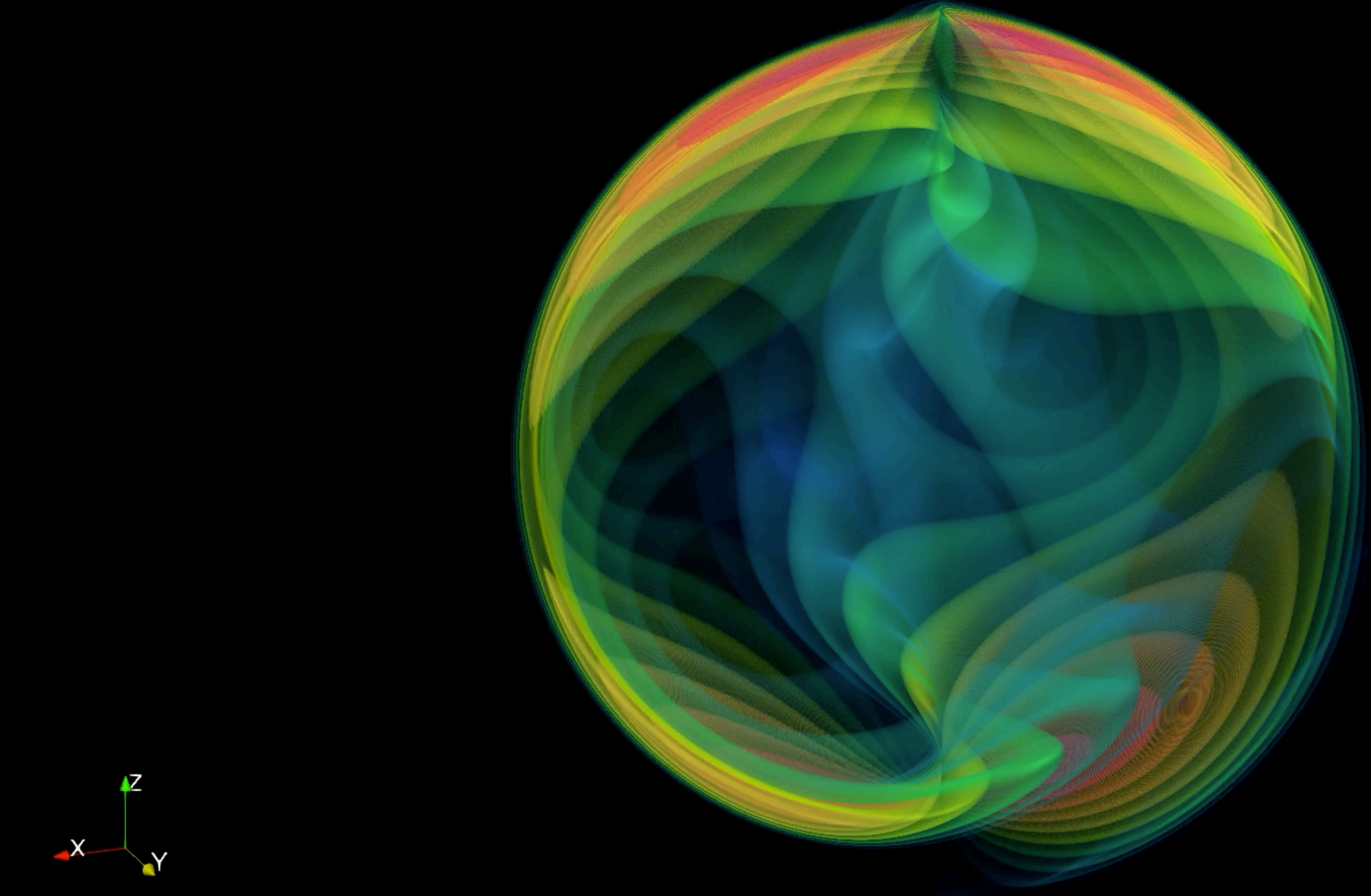
… and the rest?

- Need to include EMRIs in search and characterisation pipelines.
  - **BRUTAL task**. **We can't even find one!**
  - Event rates: **Anywhere between 1 and 10,000…**



- Need to include non-stationary features of noise.
  - **Gaps/glitches and other demons could change entire search strategies**
  - Completely unavoidable and yet unsolved.
- Multi-band detections? Stellar origin sweeping through LISA & LIGO band? **Horrendous.**

Questions



(Credit: Nils Fischer)

ollie.burke@l2it.in2p3.fr
https://github.com/OllieBurke/Tutorials.git

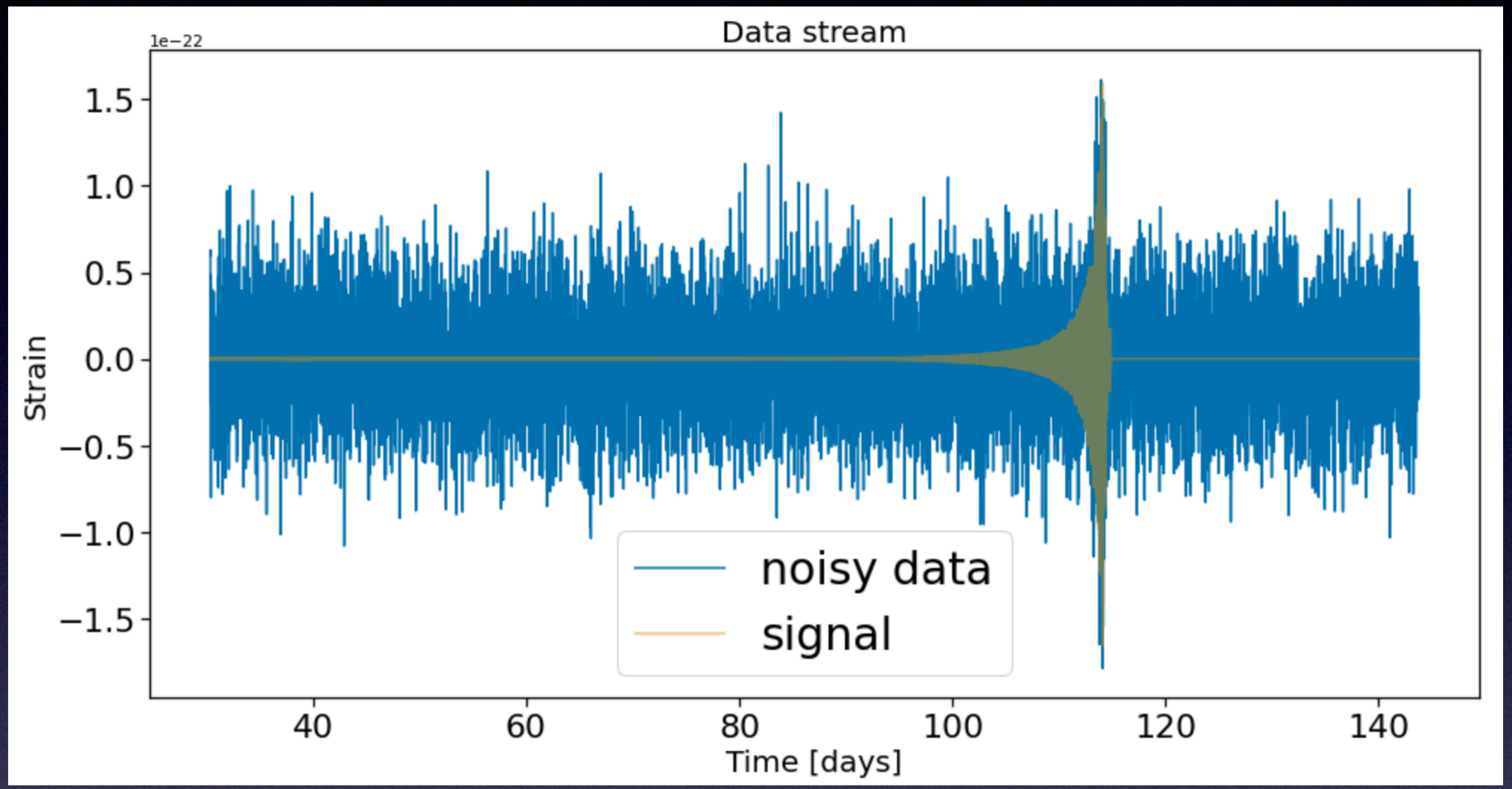
Signal-to-noise Ratio

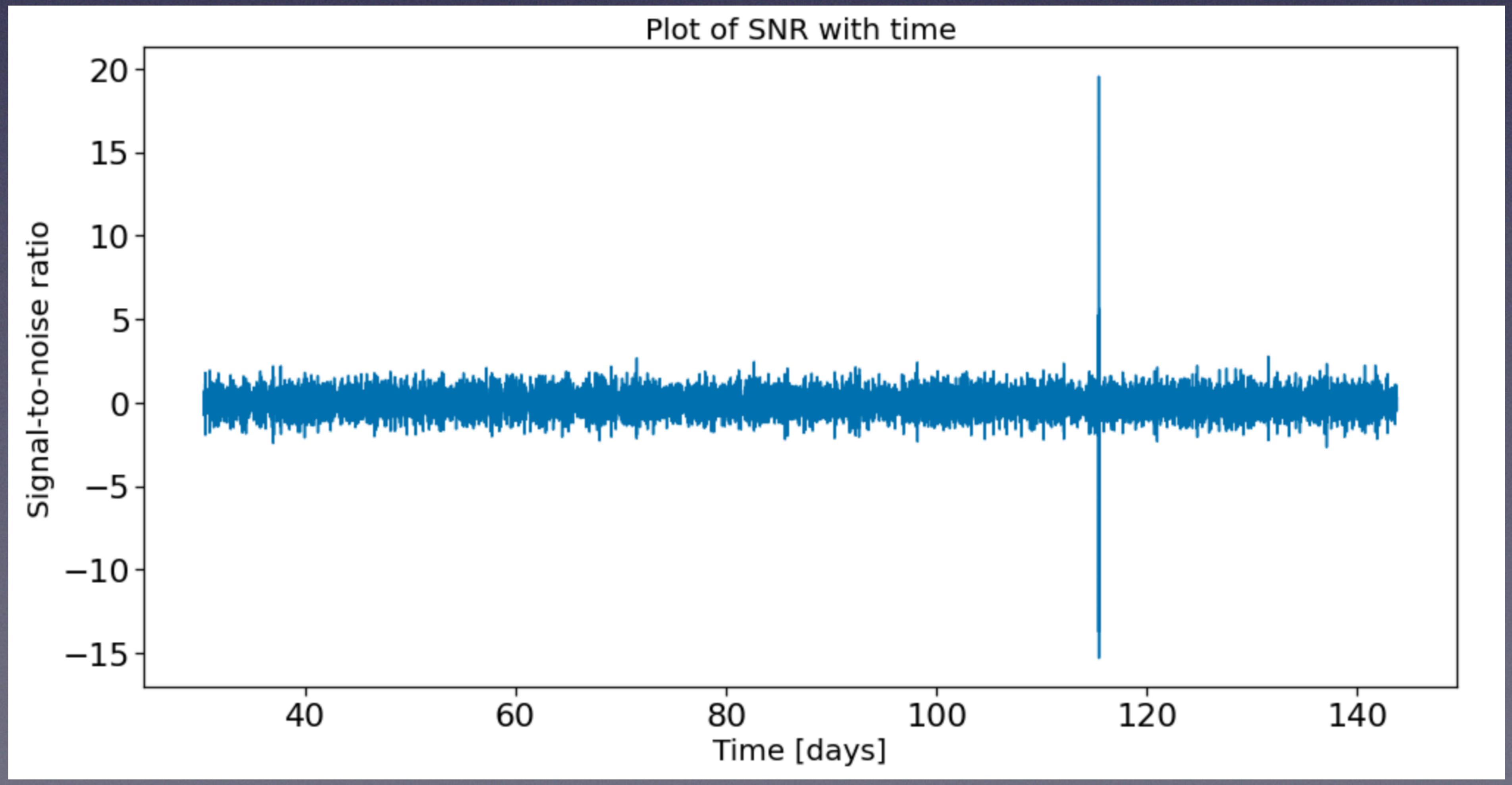
Data stream: $d(t) = h_e(t;\boldsymbol{\theta}) + n(t)$

Question: How bright is $h_e(t;\boldsymbol{\theta})$ compared to background noise $n(t)$

$$\mathrm{SNR}^2 = (h_e | h_e) = 4\int_0^\infty \frac{|\hat{h}_e(f)|^2}{S_n(f)} df = \frac{\text{Power of noise}}{\text{Variance of Noise}}$$

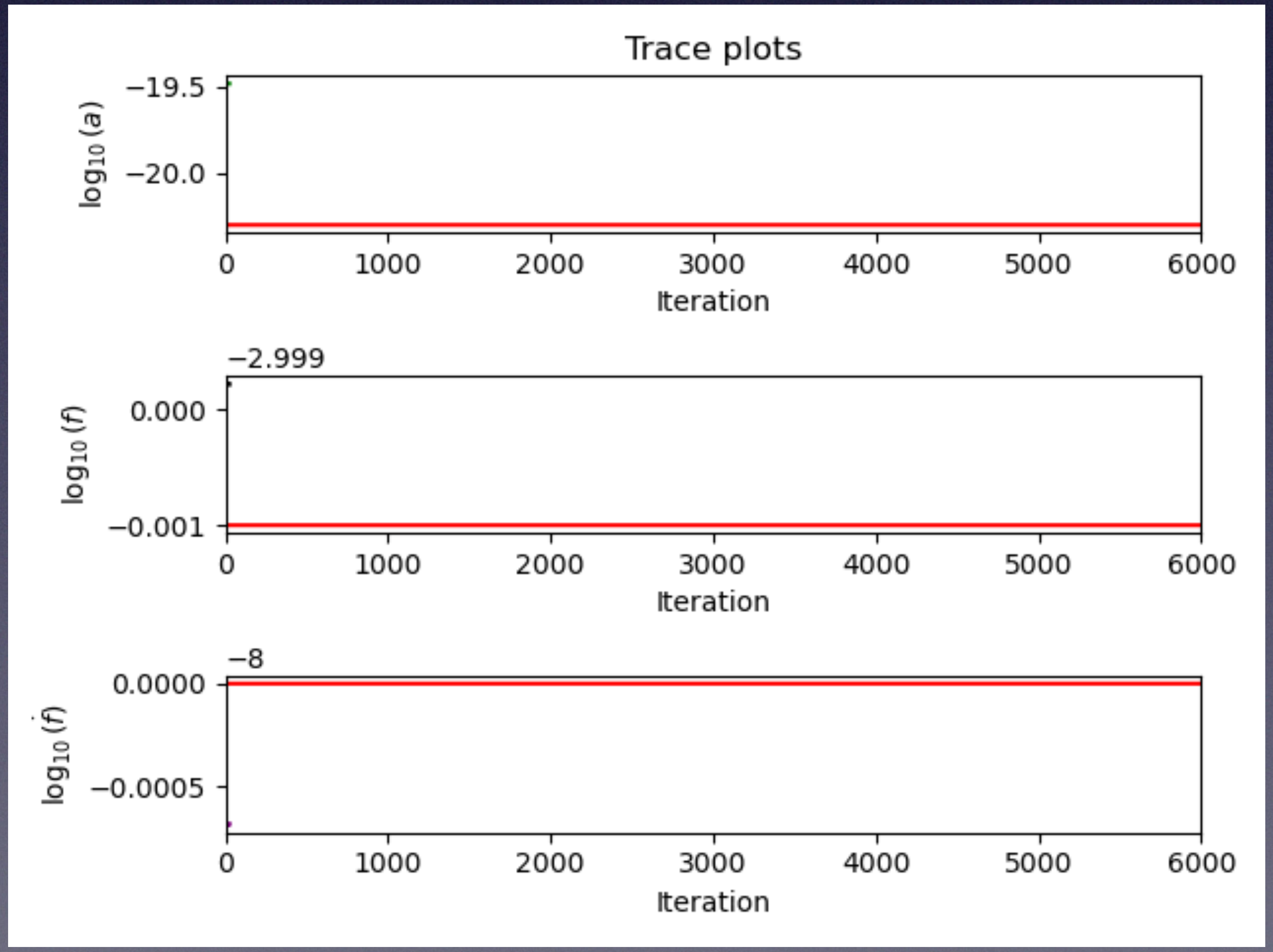
Define statistic: $\rho_d = \dfrac{(h_m|d)}{\sqrt{(h_m|h_m)}} \implies N\left[\dfrac{(h_m|h)}{\sqrt{(h_m|h_m)}}, 1\right]$

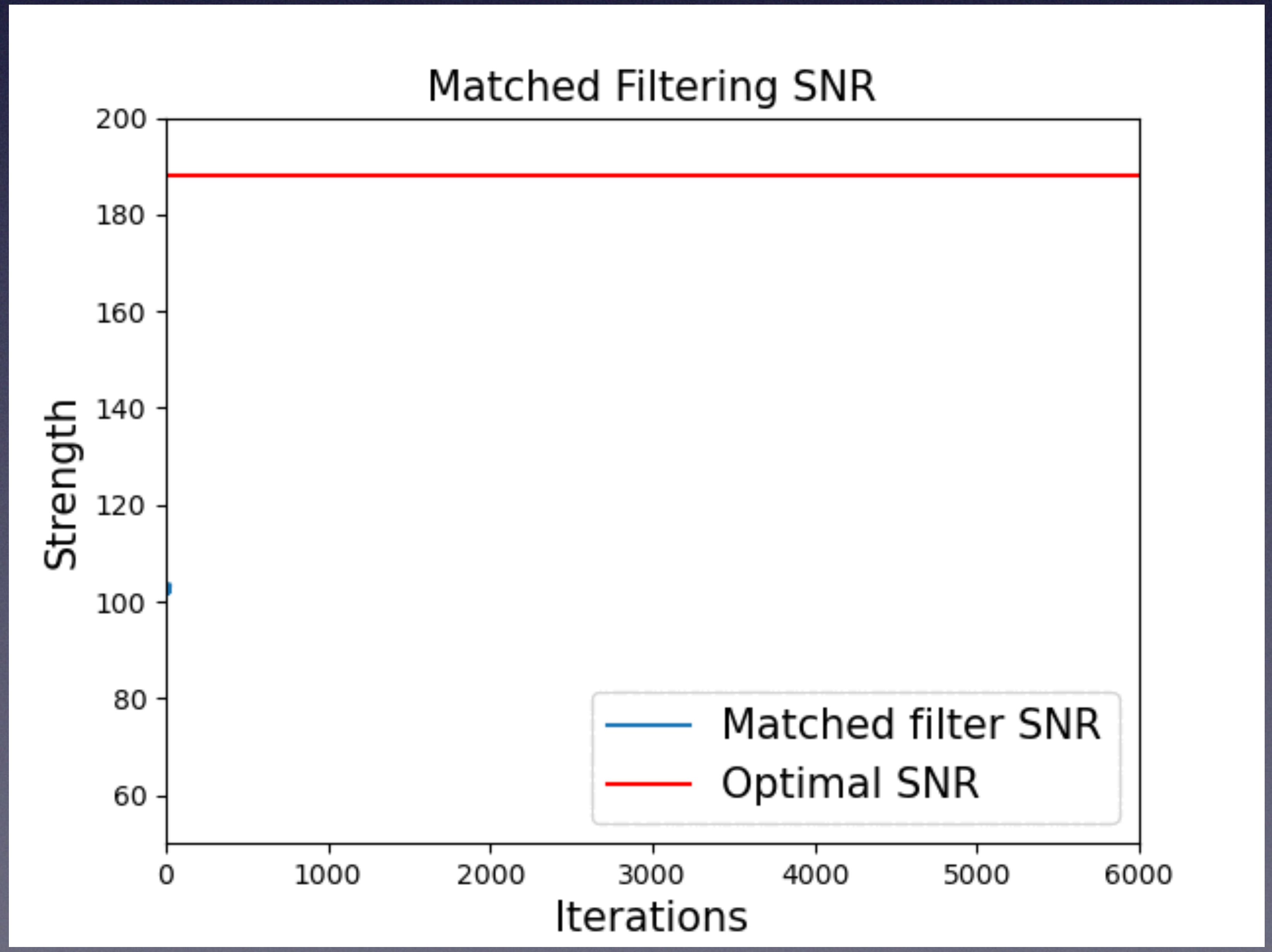


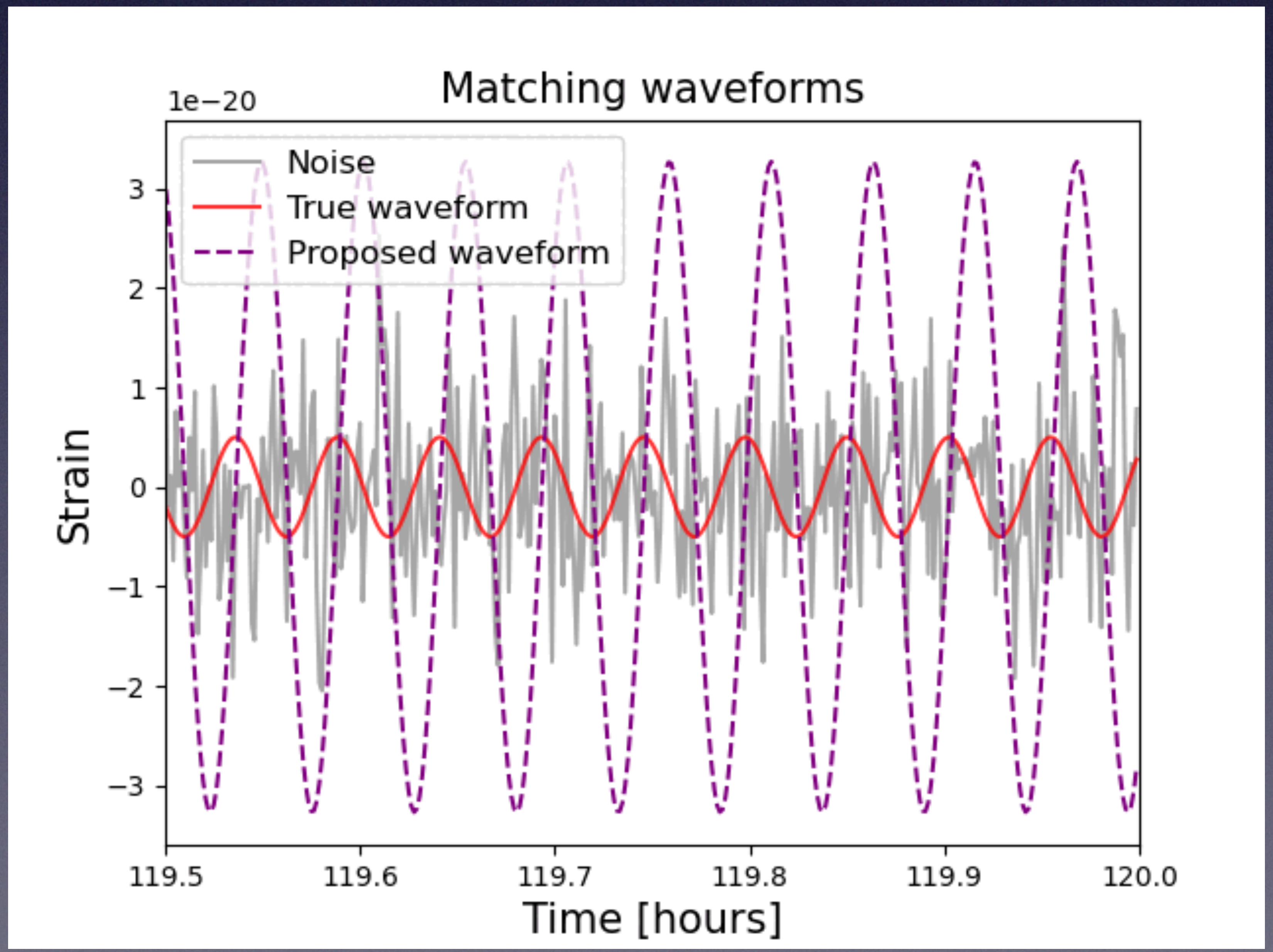


Finding the "best" signal

Goal: Identify parameters $\boldsymbol{\theta}$ that best match the waveform







https://github.com/OllieBurke/Tutorials.git