Building flexible models of gravitational wave data

…but not *too* flexible…

Tyson B. Littenberg

$$p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)}$$

Probability (density) of *A* given *B*

$$p(\theta \mid d, M) = \frac{p(d \mid \theta, M)\, p(\theta \mid M)}{p(d \mid M)}$$

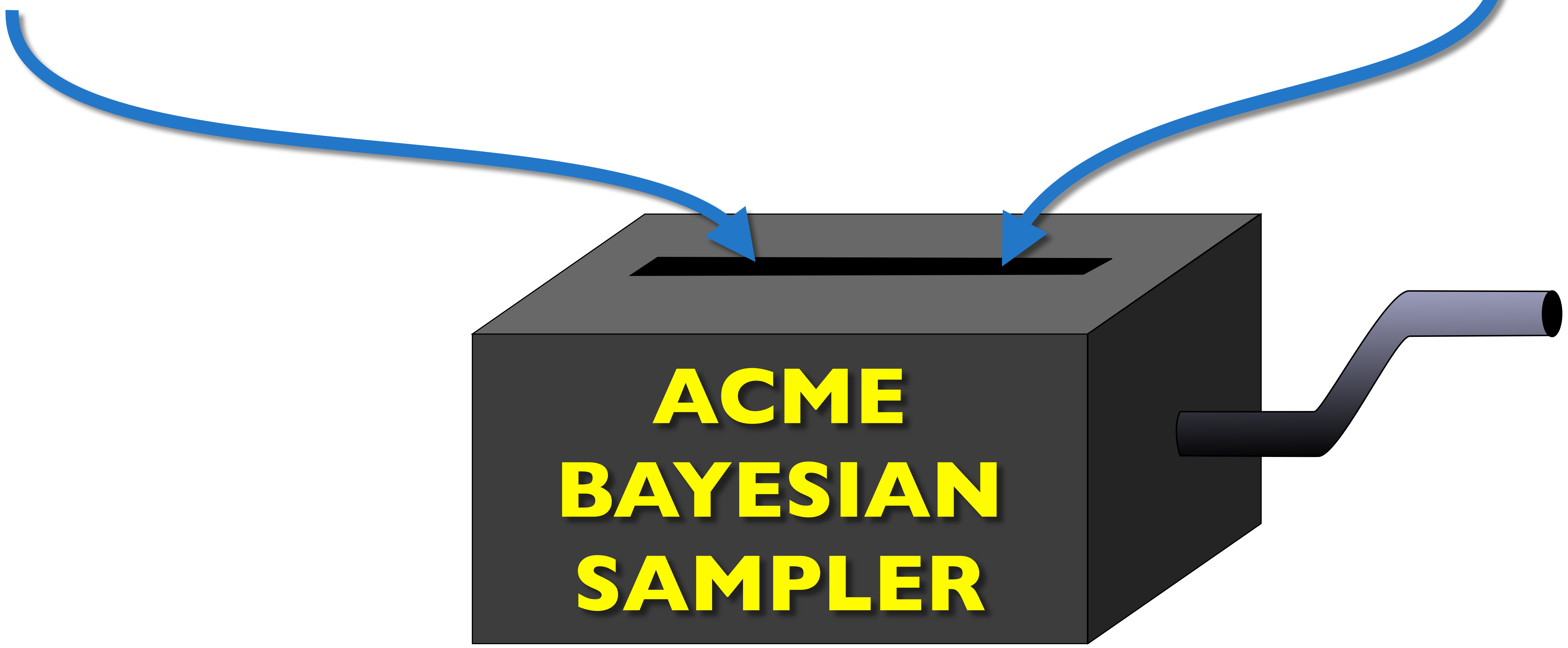
Probability (density) of *parameters*
given *data && model*

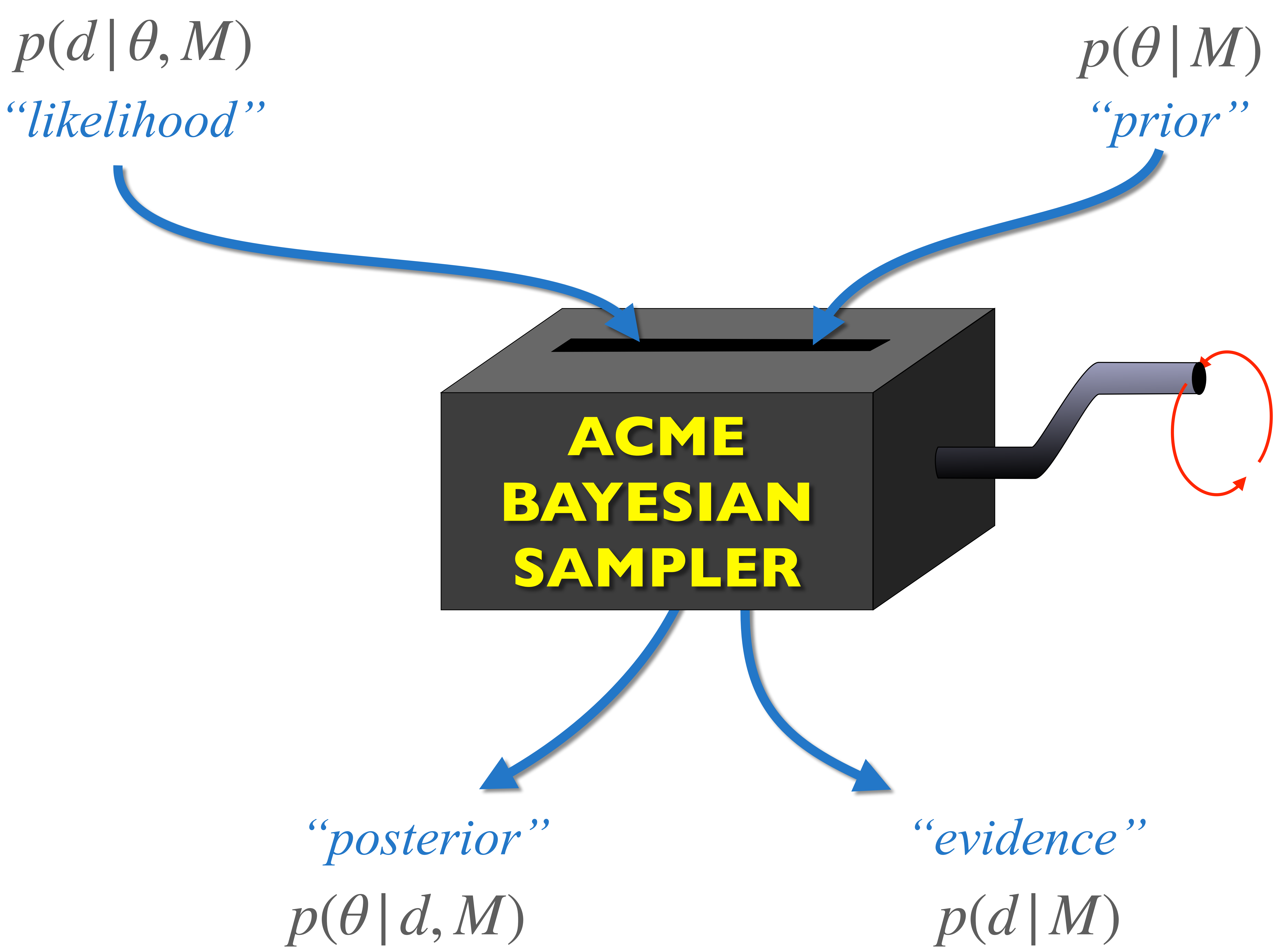
$p(d \mid \theta, M)$

"likelihood"

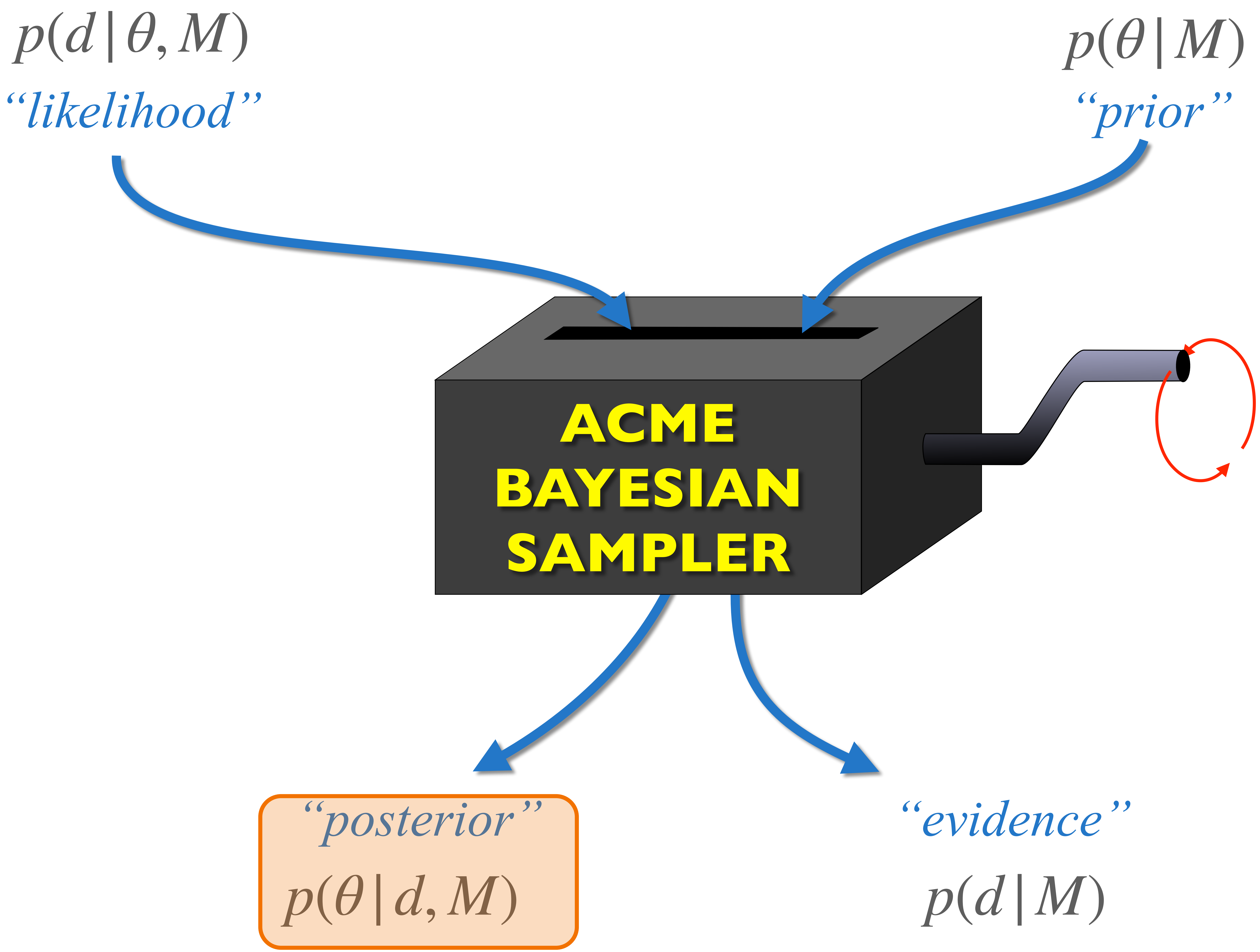
$p(\theta \mid M)$

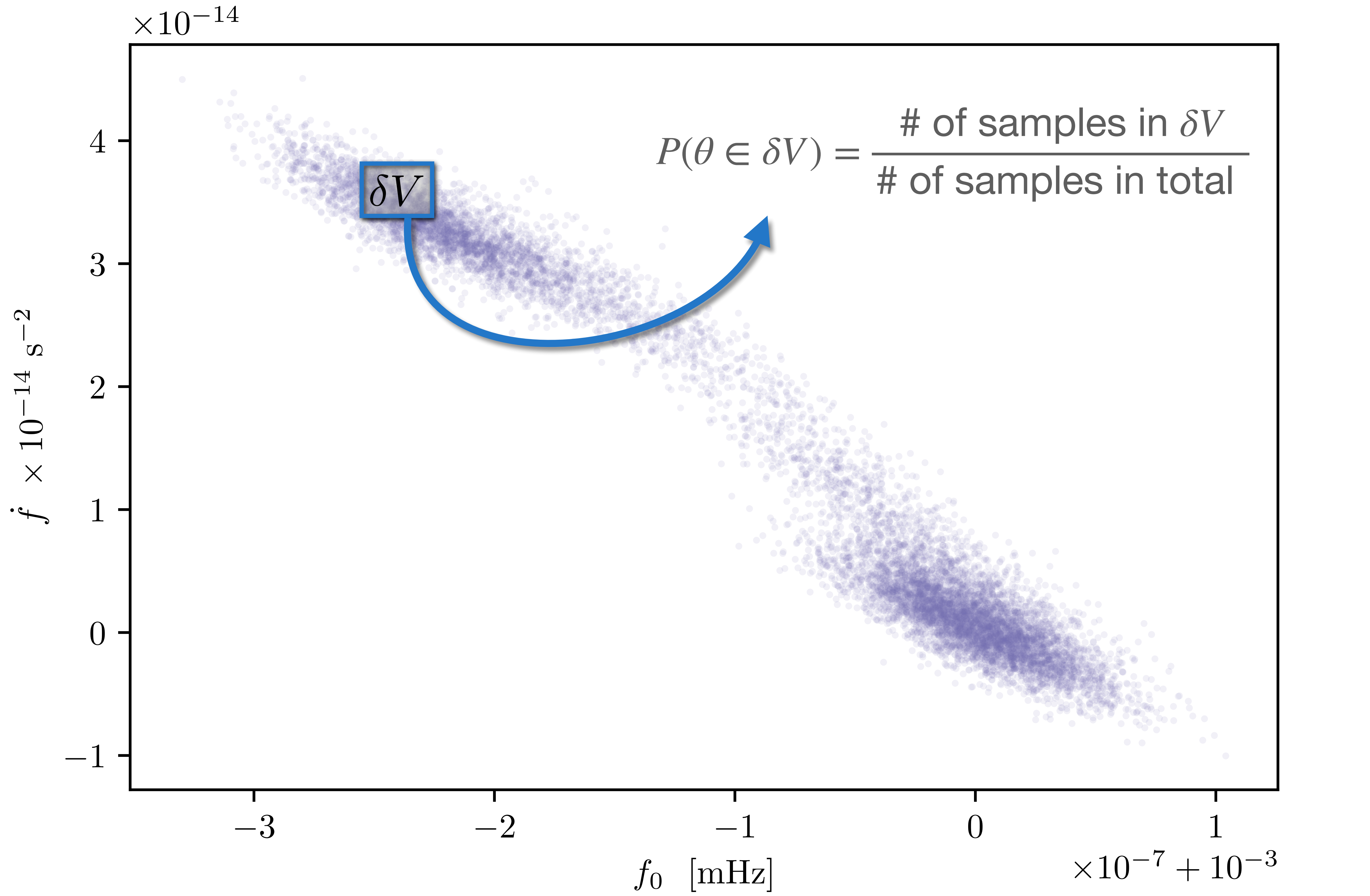
"prior"





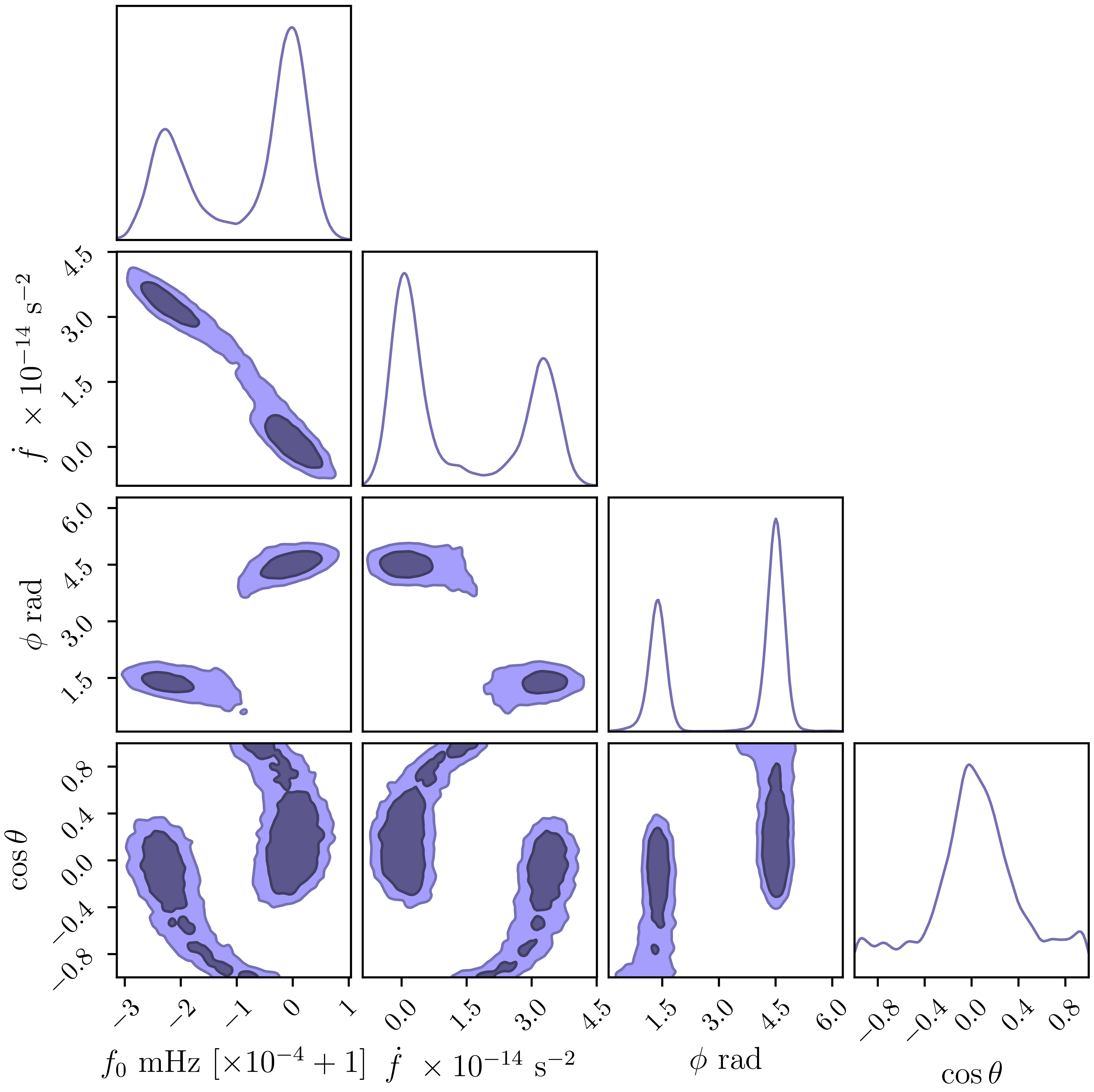
$p(d|\theta, M)$
"likelihood"
$p(\theta|M)$
"prior"
ACME
BAYESIAN
SAMPLER
"posterior"
$p(\theta|d, M)$
"evidence"
$p(d|M)$


$p(d|\theta, M)$
"likelihood"
$p(\theta|M)$
"prior"
ACME
BAYESIAN
SAMPLER
"posterior"
$p(\theta|d, M)$
"evidence"
$p(d|M)$

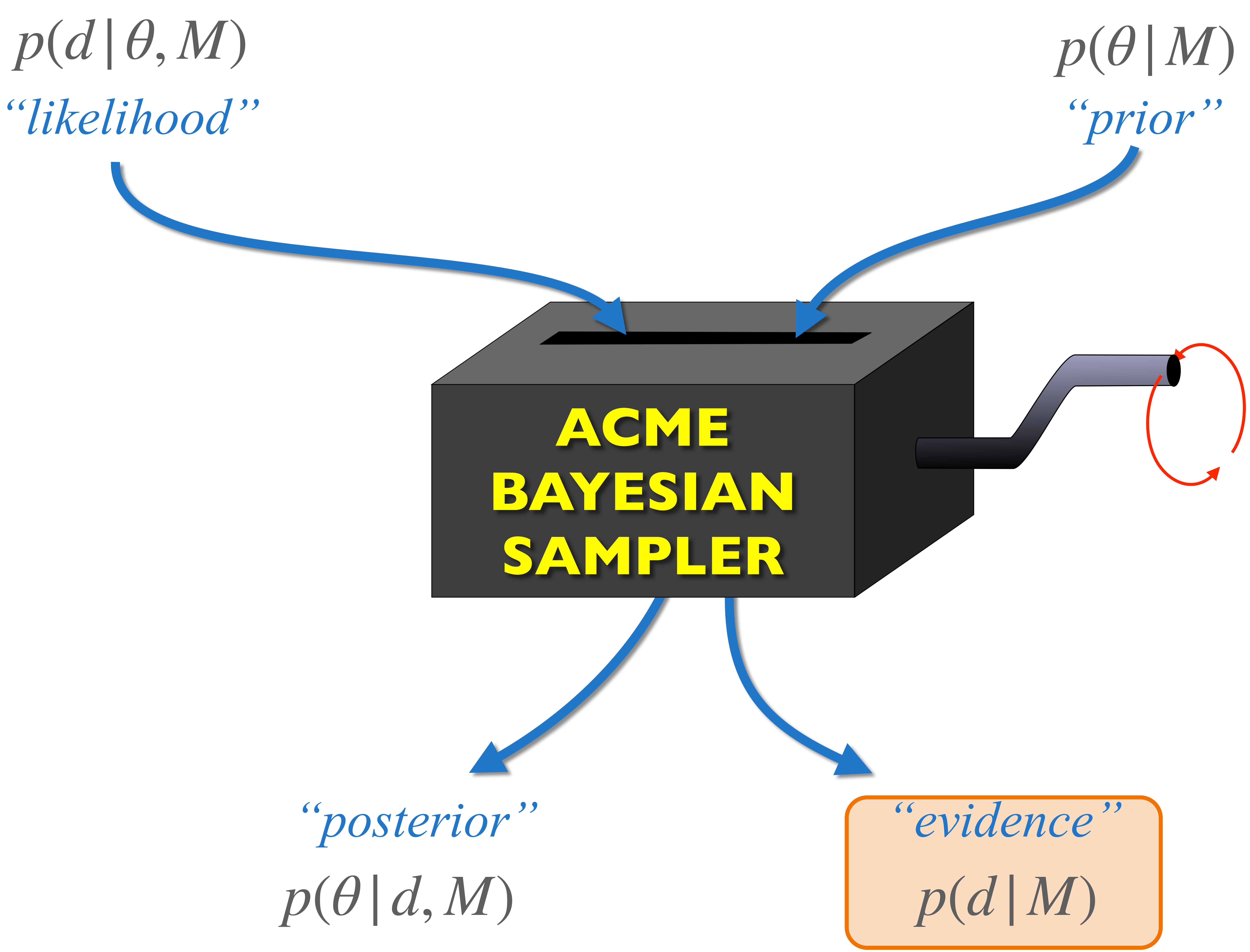


$$p(x\,|\,d, M) = \int p(x, y, z\,|\,d, M)\; dy\; dz$$

"nuisance parameters"



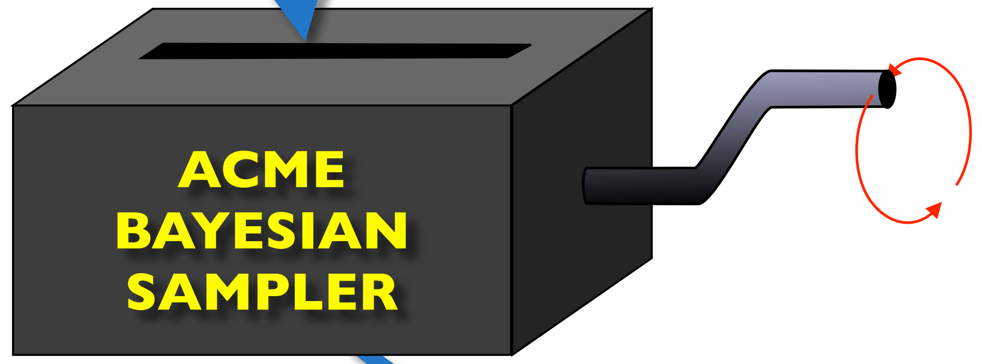
$\dot{f} \times 10^{-14}$ s^{-2}
ϕ rad
$\cos\theta$
f_0 mHz $[\times 10^{-4} + 1]$
$\dot{f} \times 10^{-14}$ s^{-2}
ϕ rad
$\cos\theta$


$p(d|\theta, M)$
"likelihood"
$p(\theta|M)$
"prior"
ACME BAYESIAN SAMPLER
"posterior"
$p(\theta|d, M)$
"evidence"
$p(d|M)$

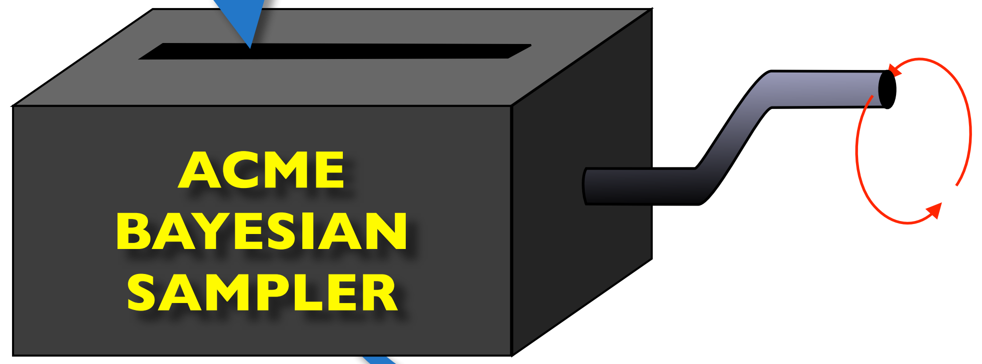
$$p(d\,|\,M) = \int d\theta\; p(d\,|\,\theta, M)\; p(\theta\,|\,M)$$

$$p(d\,|\,M) = \int d\theta\; p(d\,|\,\theta, M)\; p(\theta\,|\,M)$$

"I've detected Gravitational Waves!"



"I've measured lots of noise!"



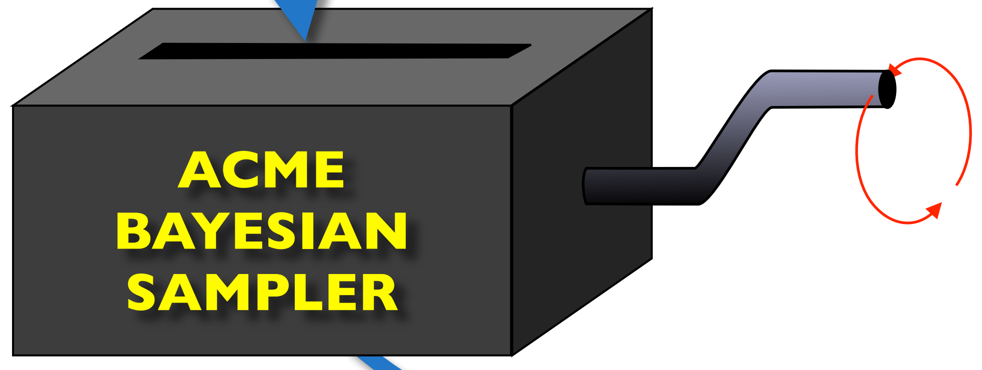
$$\mathcal{O}_{A,B} = \frac{p(M_A)}{p(M_B)} \times \frac{p(d\,|\,M_A)}{p(d\,|\,M_B)}$$

"odds ratio"

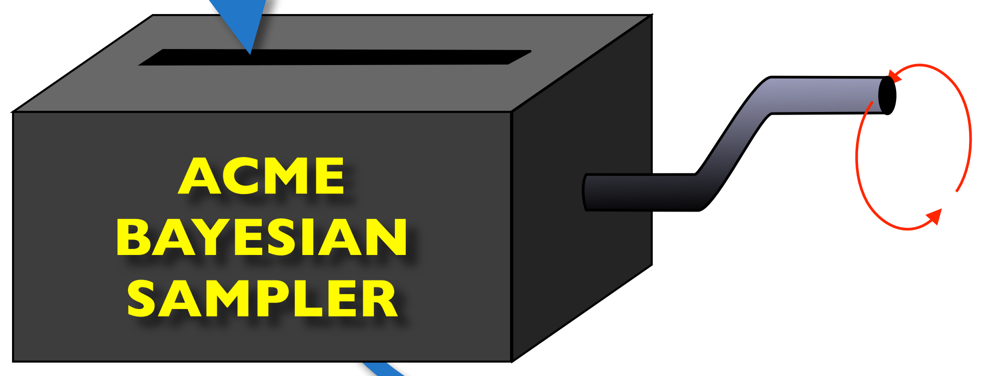
"Bayes factor"

$$p(d\,|\,M) = \int d\theta \; p(d\,|\,\theta, M)\; p(\theta\,|\,M)$$

"I've detected Gravitational Waves!"



"I've measured lots of noise!"



$$\mathcal{O}_{A,B} = \frac{p(M_A)}{p(M_B)} \times \frac{p(d\,|\,M_A)}{p(d\,|\,M_B)}$$

"odds ratio"

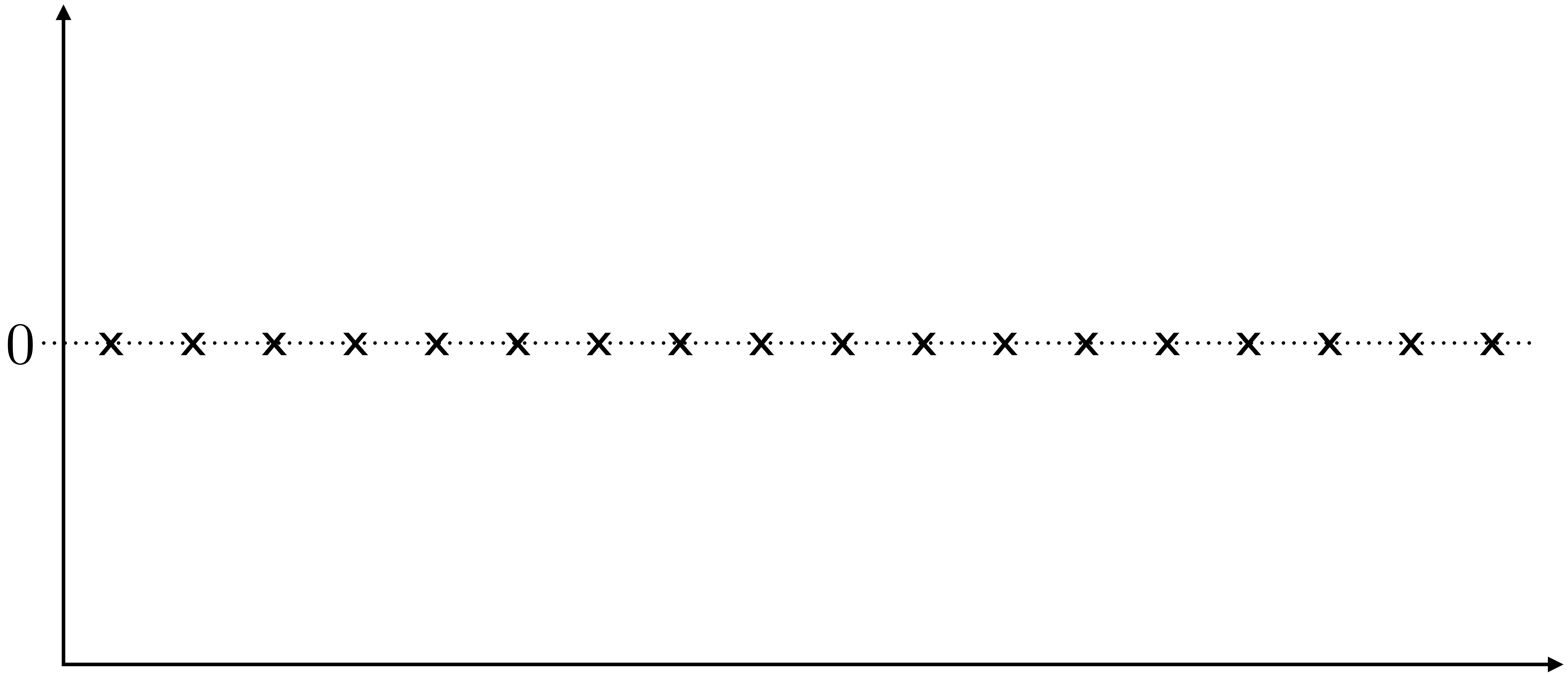
"Bayes factor"

$\mathcal{O}_{A,B} = X \equiv$ Model A is preferred over model B with $X : 1$ odds

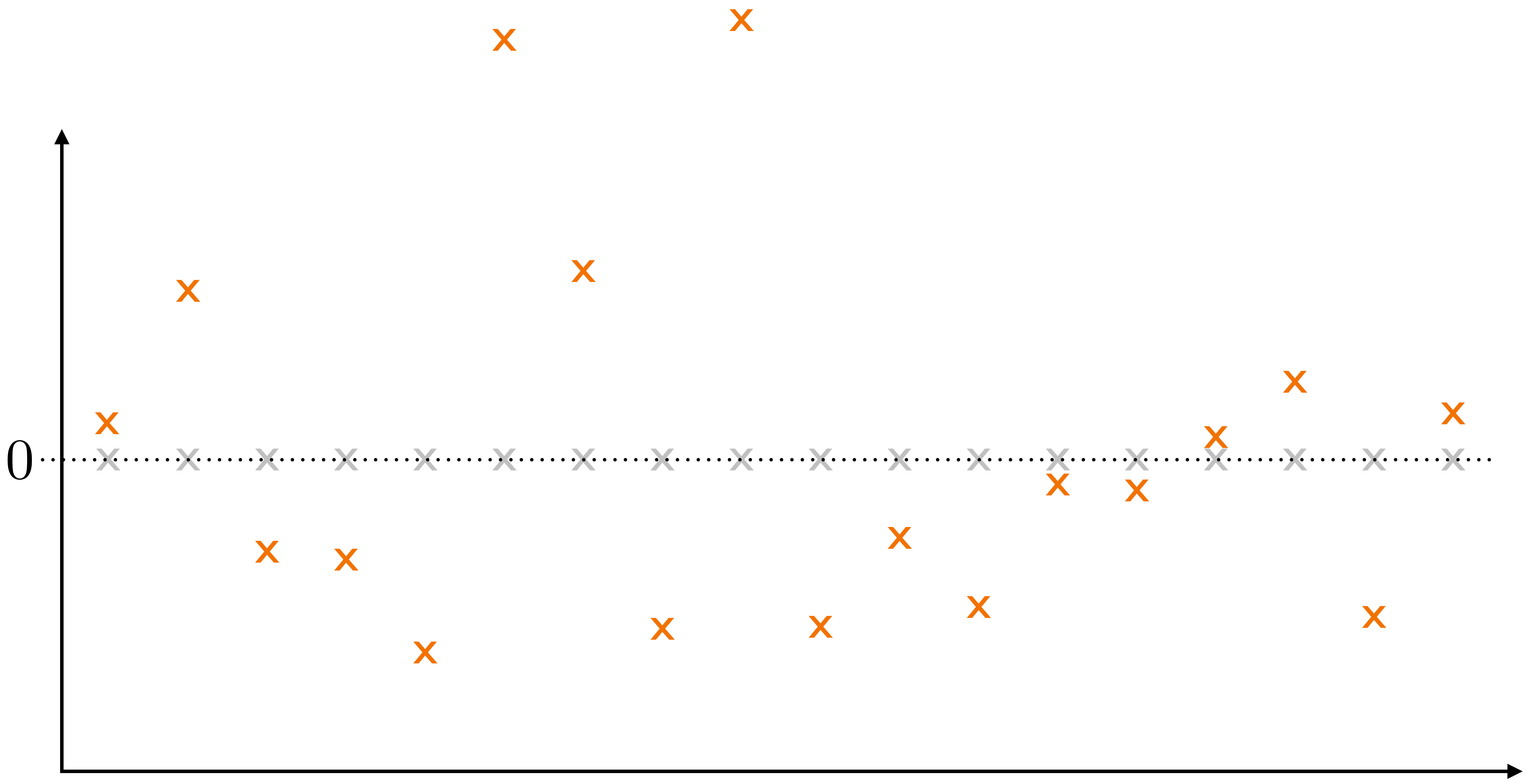
Bayesian Analyses: Not magic.

Let’s Build a Likelihood Function

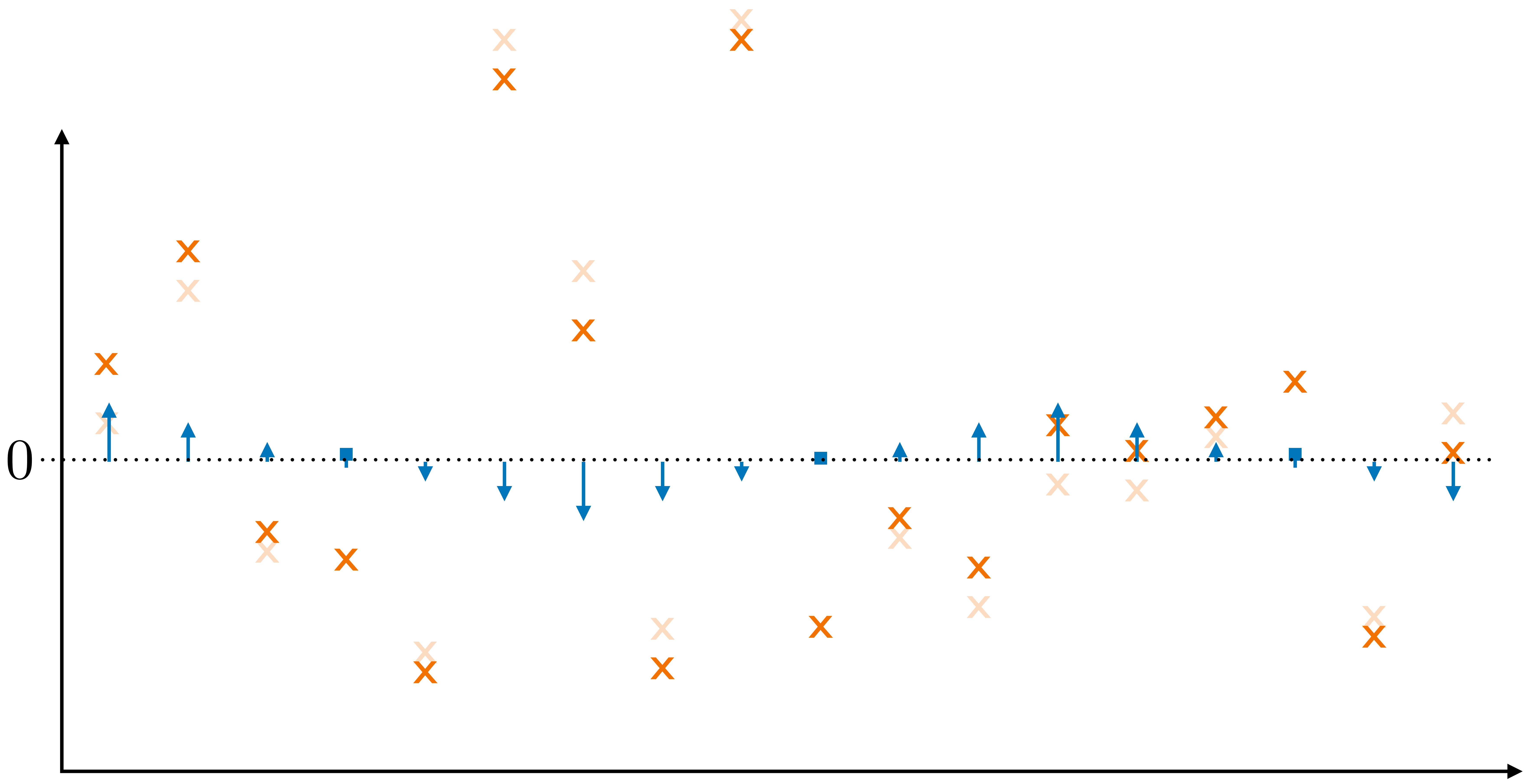
d



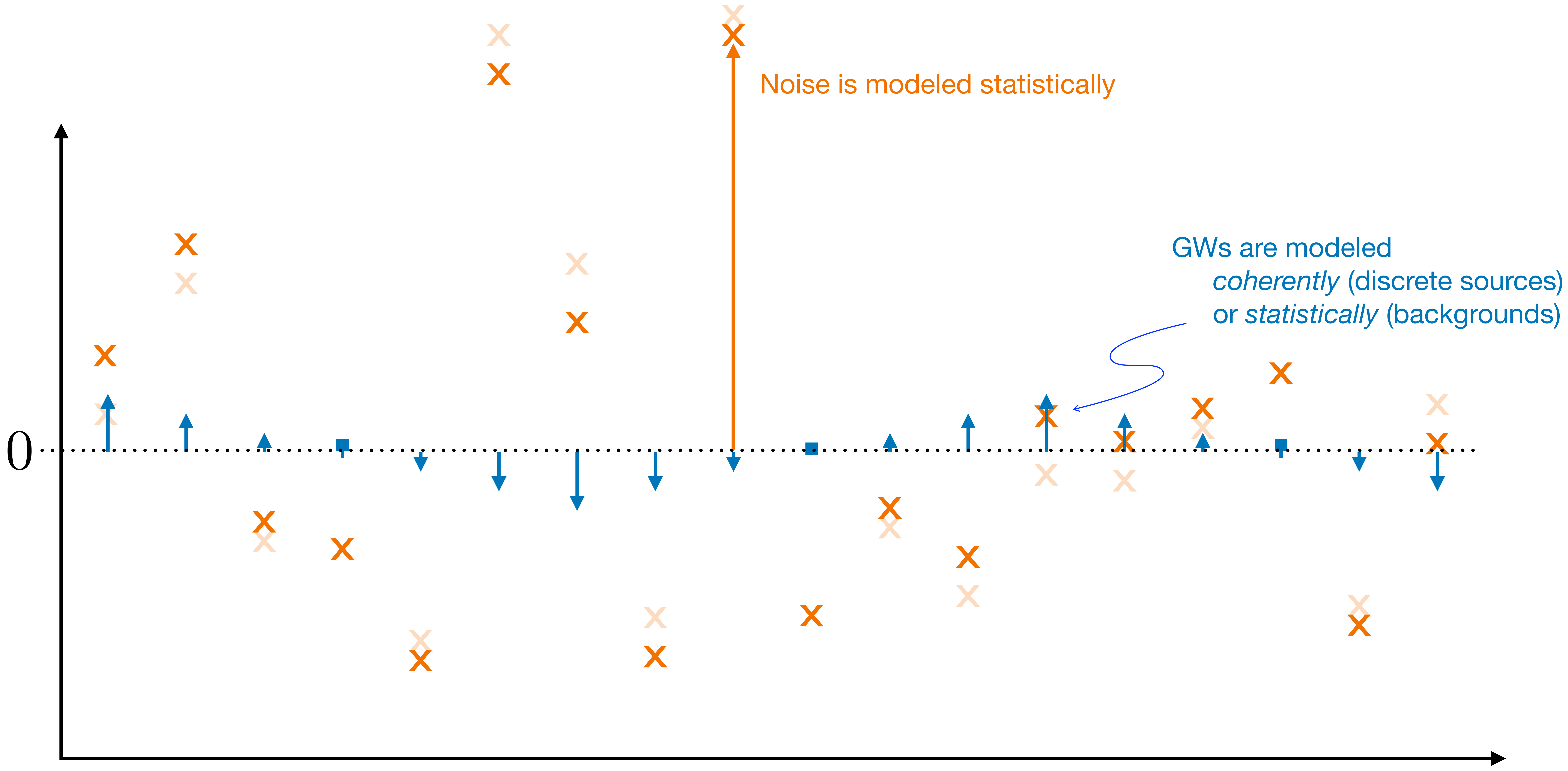
$d = n$



$$d = n + h$$



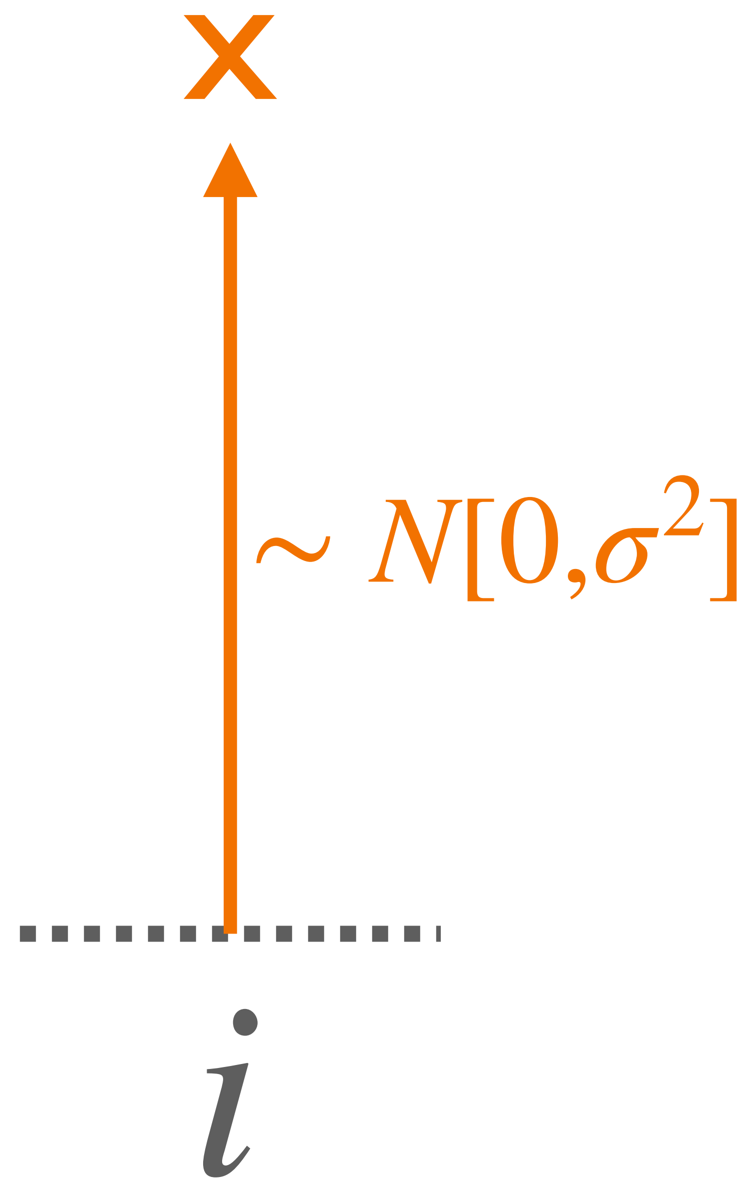
$$d = n + h$$


Noise is modeled statistically
GWs are modeled
coherently (discrete sources)
or *statistically* (backgrounds)
0

$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance



Probability of measuring noise n_i

$$p(n_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|n_i|^2}{2\sigma^2}}$$

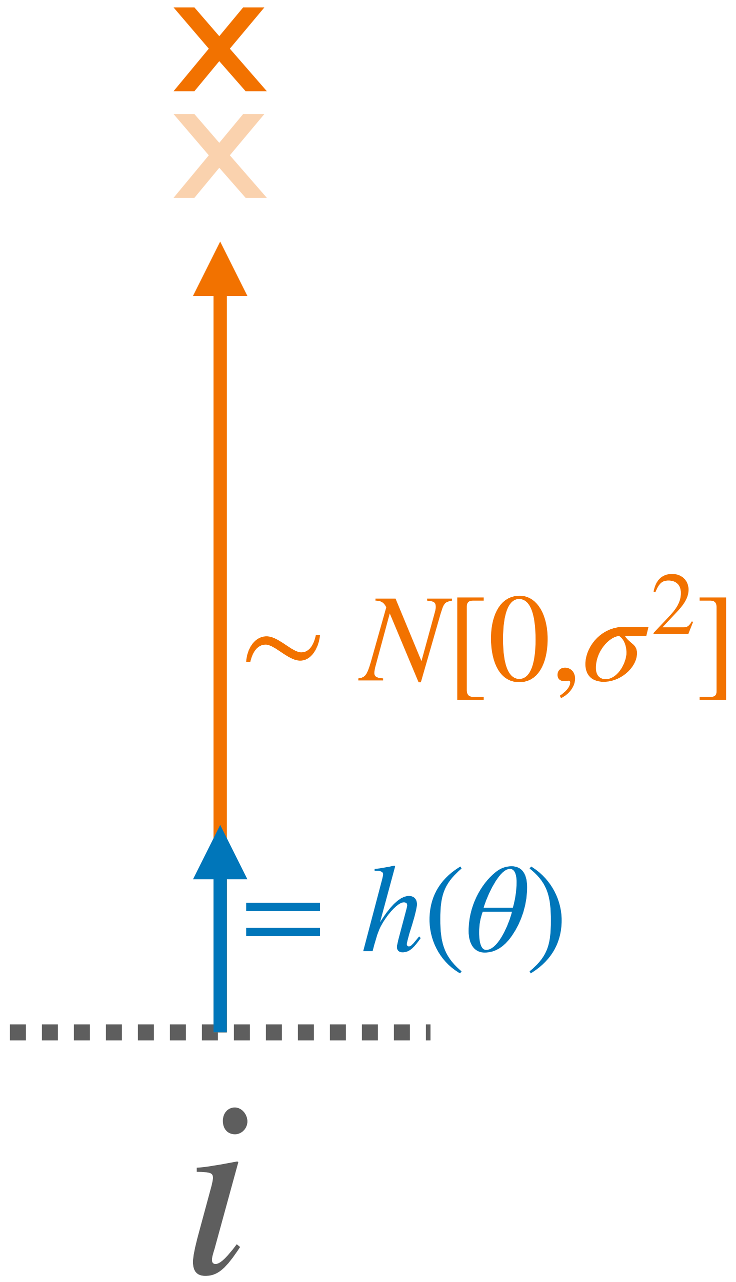
$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect



Probability of measuring data d_i

$$p(d_i \,|\, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|d_i - h_i|^2}{2\sigma^2}}$$

$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect

Probability of measuring set of data **d** with *k* samples:

$$p(\mathbf{d}\,|\,\theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^T C^{-1}(\mathbf{d}-\mathbf{h})}$$

$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect

Noise variance is *stationary*

$$\langle \tilde{n}_i \tilde{n}_j \rangle = \sigma_i^2 \delta_{i,j} \equiv \frac{T}{2} S_n(f_i)$$

Probability of measuring set of data **d** with *k* samples:

$$p(\mathbf{d} \,|\, \theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^T C^{-1} (\mathbf{d}-\mathbf{h})}$$

$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect

Noise variance is *stationary*

$$\langle \tilde{n}_i \tilde{n}_j \rangle = \sigma_i^2 \delta_{i,j} \equiv \frac{T}{2} S_n(f_i)$$

Probability of measuring set of data **d** with *k* samples:

$$p(\mathbf{d}\,|\,\theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^T C^{-1} (\mathbf{d}-\mathbf{h})}$$

$$\Downarrow$$

$$p(\mathbf{d}\,|\,\theta) = \prod_k \frac{2}{\pi T S_{n,k}} e^{-\frac{2|\tilde{d}_k - \tilde{h}_k|^2}{T S_{n,k}}}$$

$$d = n + h$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect

Noise variance is *stationary*

$$\langle \tilde{n}_i \tilde{n}_j \rangle = \sigma_i^2 \delta_{i,j} \equiv \frac{T}{2} S_n(f_i)$$

Probability of measuring set of data **d** with *k* samples:

$$p(\mathbf{d} \,|\, \theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^T C^{-1} (\mathbf{d}-\mathbf{h})}$$

$$\Downarrow$$

$$p(\mathbf{d} \,|\, \theta) \propto e^{-\frac{2}{T} \sum_k \frac{|\tilde{d}_k - \tilde{h}_k|^2}{S_{n,k}}}$$

Noise is zero-mean Gaussian

Noise has known variance

Data are perfectly calibrated

Waveform model is perfect

Noise variance is *stationary*

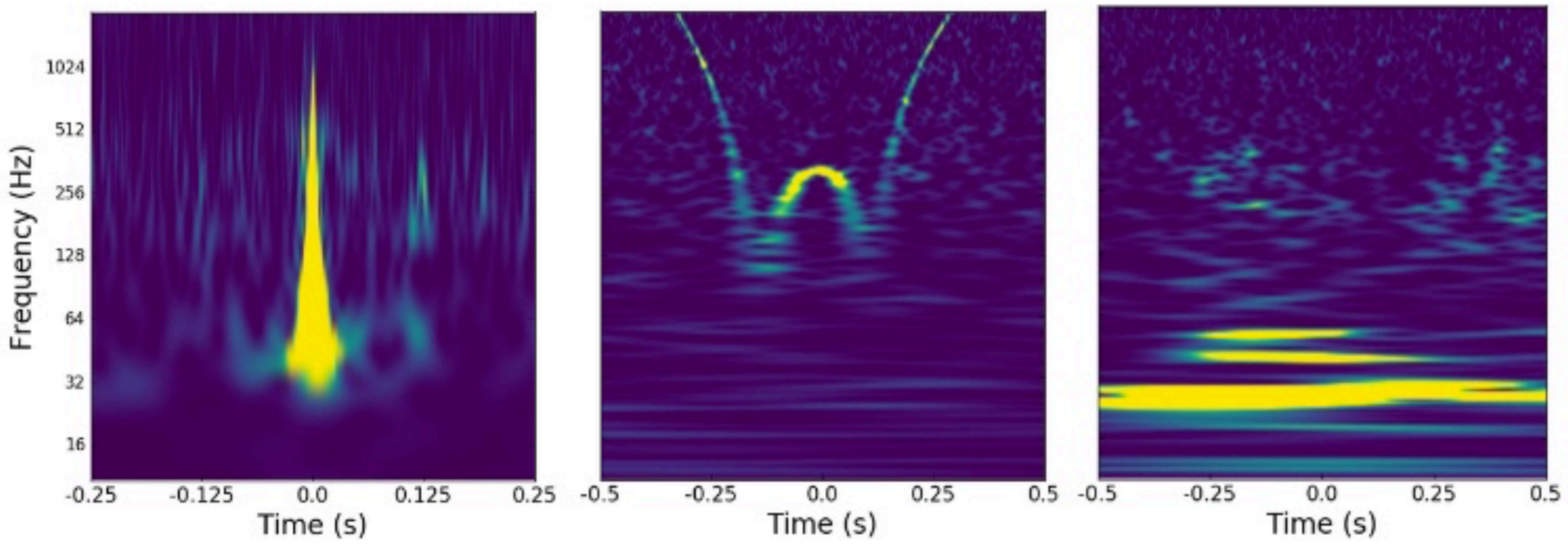
~~Noise is zero-mean Gaussian~~

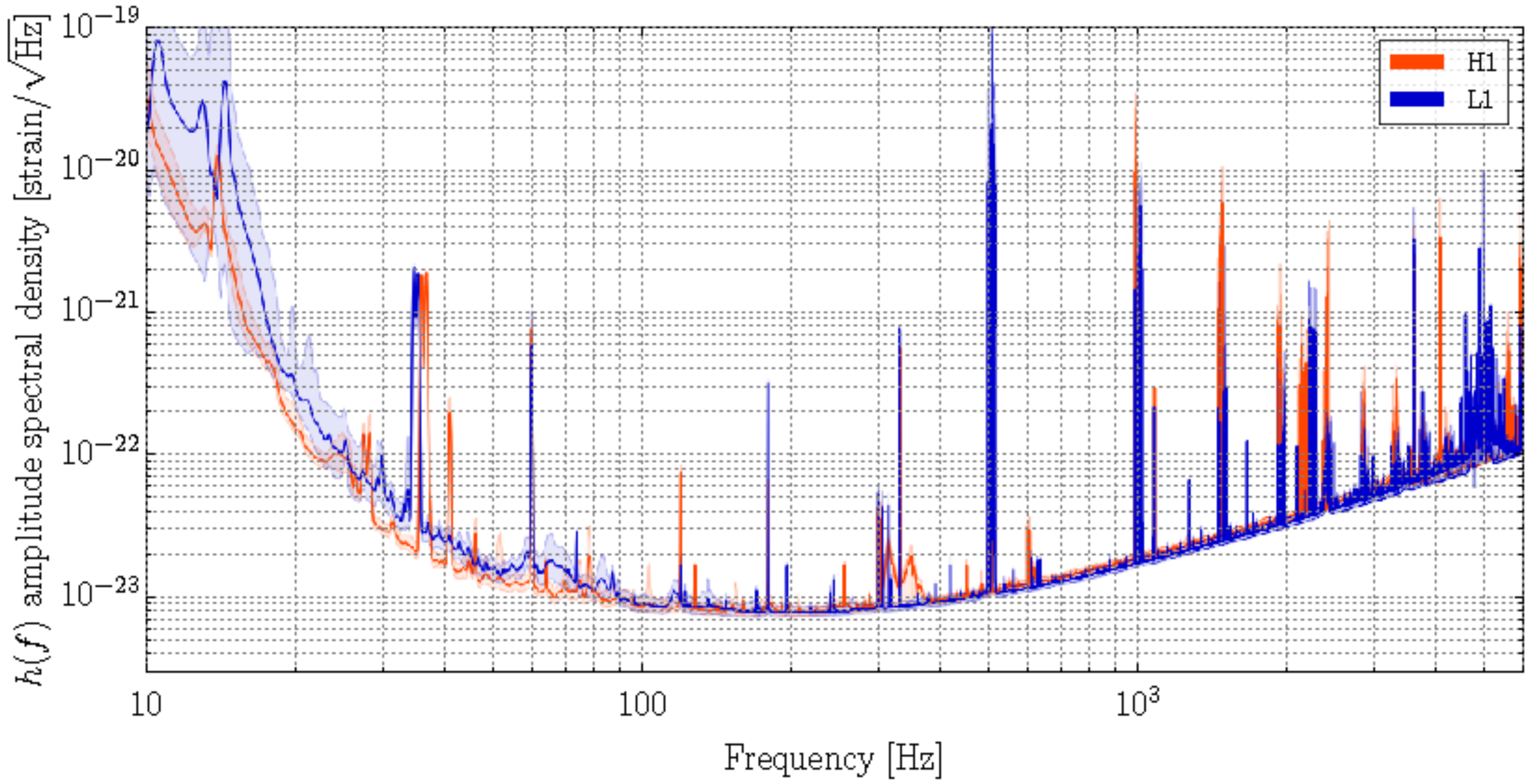
~~Noise has known variance~~

~~Data are perfectly calibrated~~

~~Waveform model is perfect~~

~~Noise variance is *stationary*~~

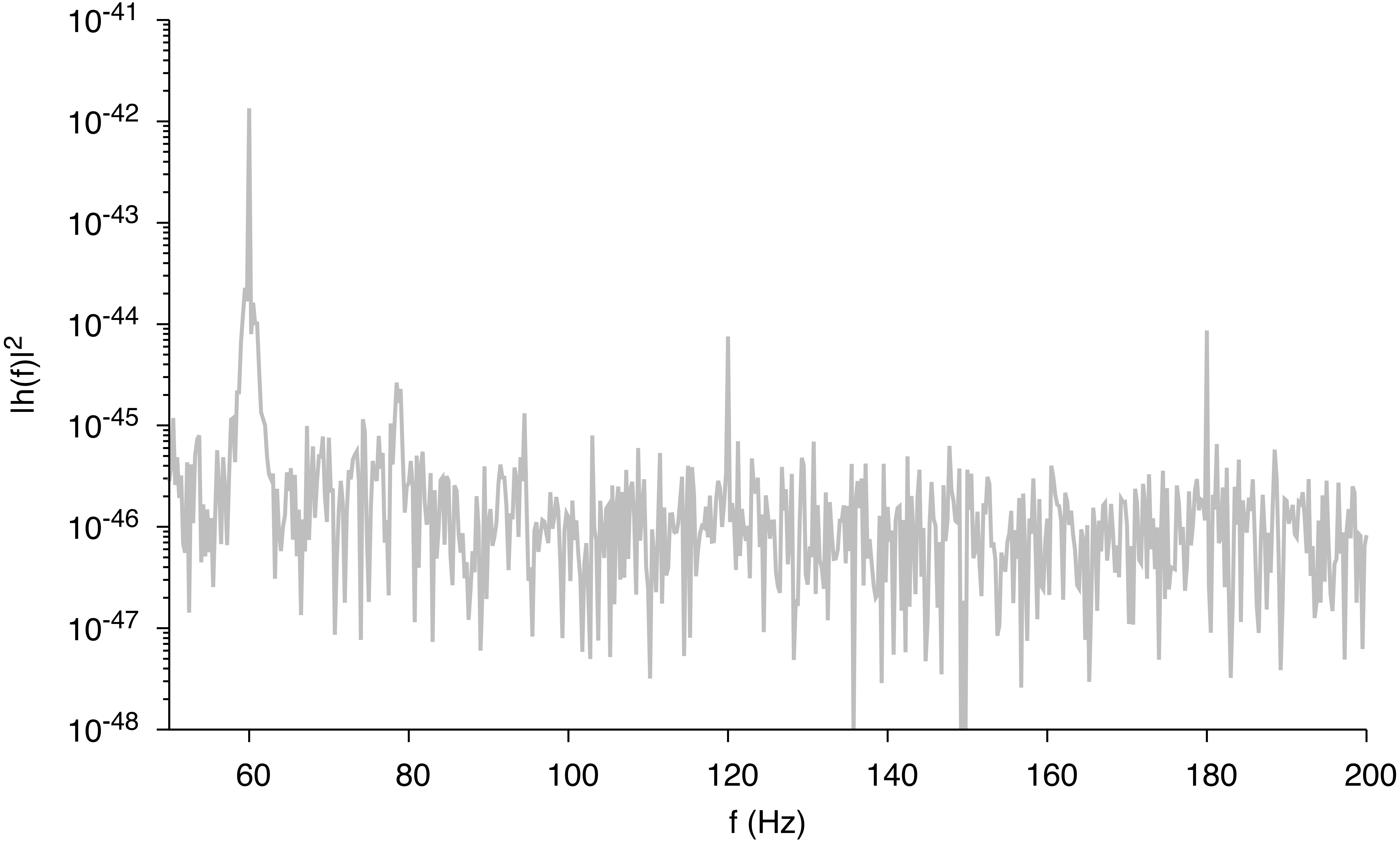




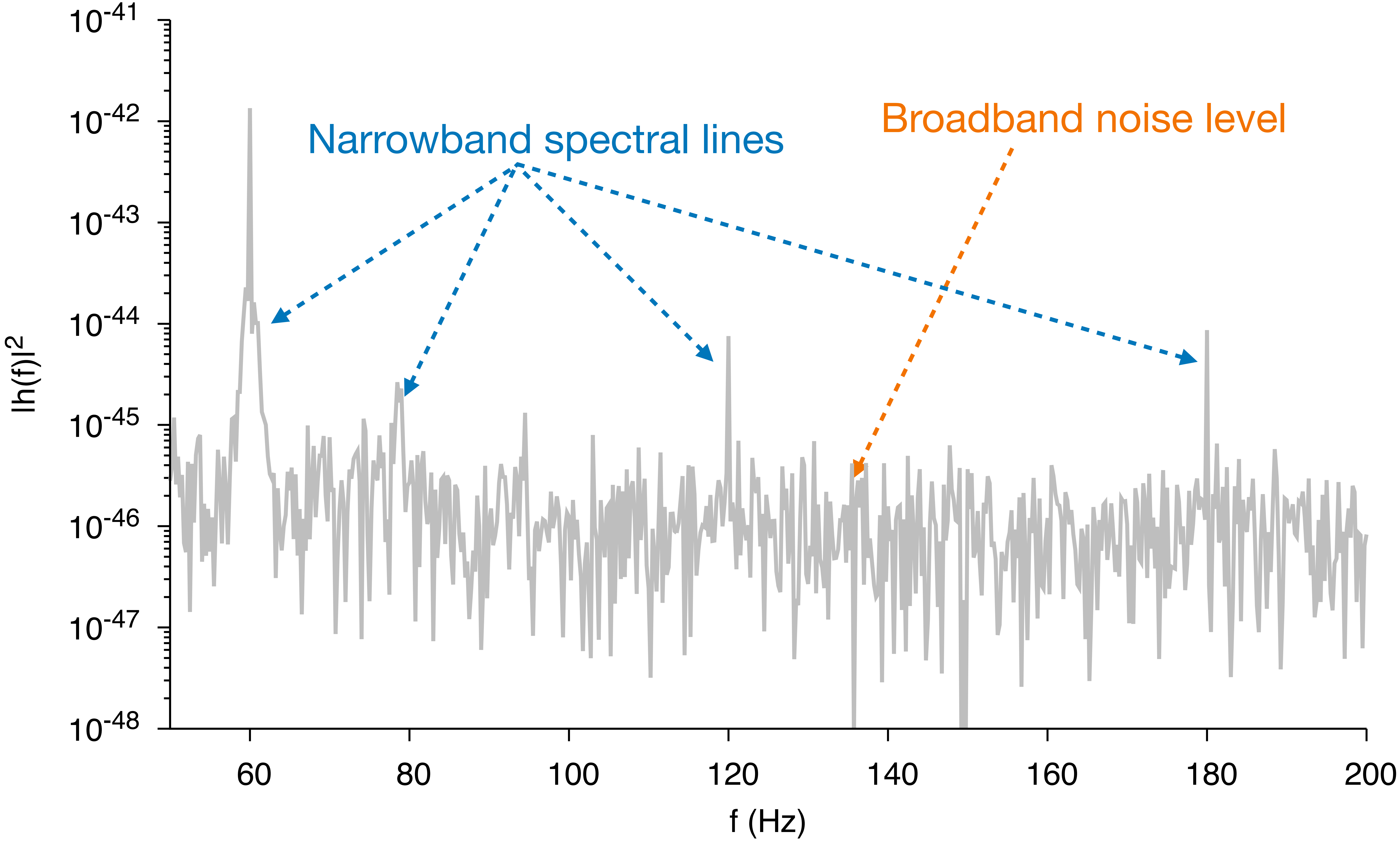
Model everything and let the data sort it out

Choose a convenient “basis set” to phenomenologically model features in data

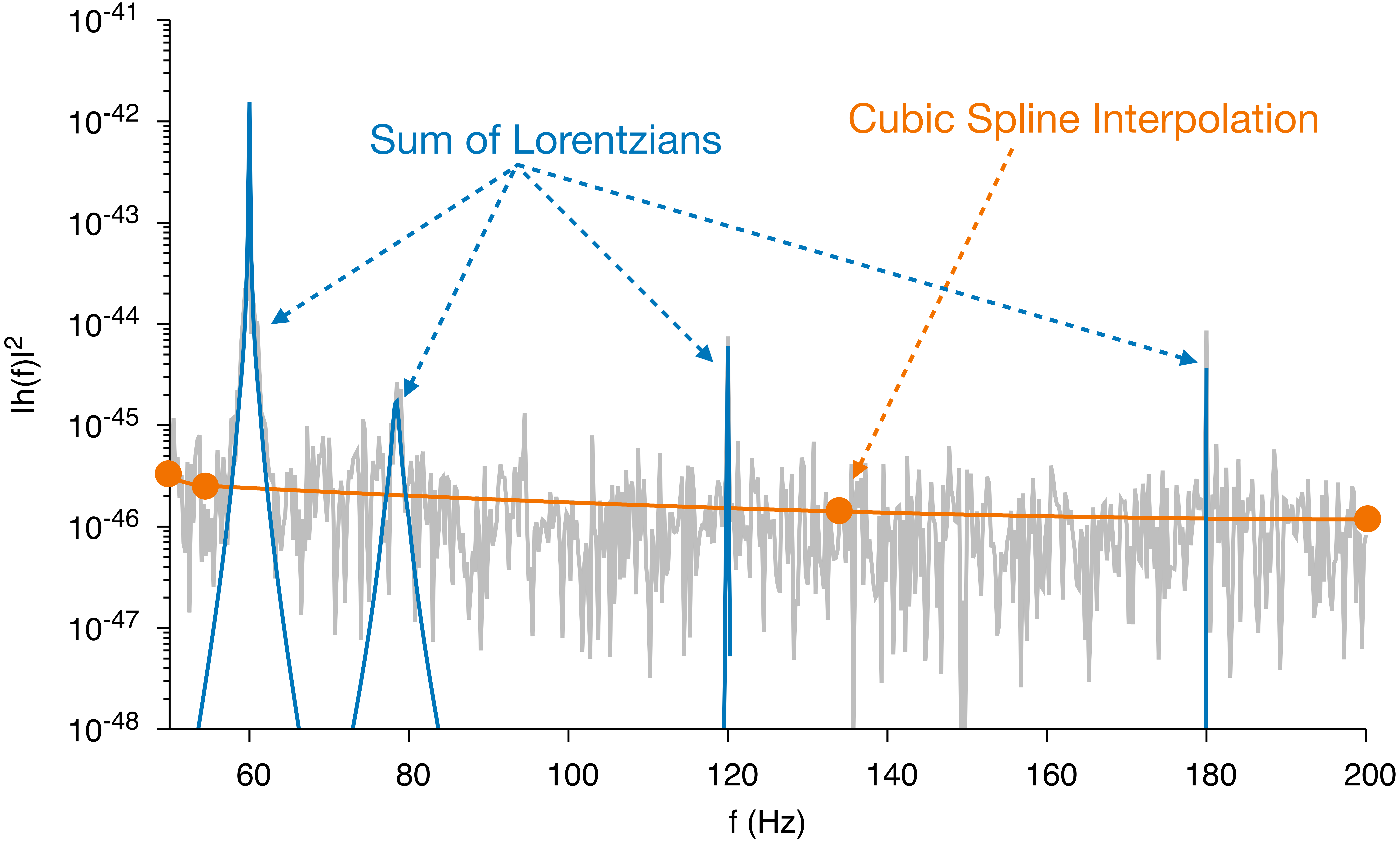
Use *evidence* to determine the number of “basis functions” to use in the model



10^{-41}
10^{-42}
10^{-43}
10^{-44}
10^{-45}
10^{-46}
10^{-47}
10^{-48}
$|h(f)|^2$
60
80
100
120
140
160
180
200
f (Hz)



Narrowband spectral lines
Broadband noise level
10^-41
10^-42
10^-43
10^-44
10^-45
10^-46
10^-47
10^-48
|h(f)|^2
60
80
100
120
140
160
180
200
f (Hz)



Sum of Lorentzians
Cubic Spline Interpolation
10^{-41}
10^{-42}
10^{-43}
10^{-44}
10^{-45}
10^{-46}
10^{-47}
10^{-48}
$|h(f)|^2$
60
80
100
120
140
160
180
200
f (Hz)

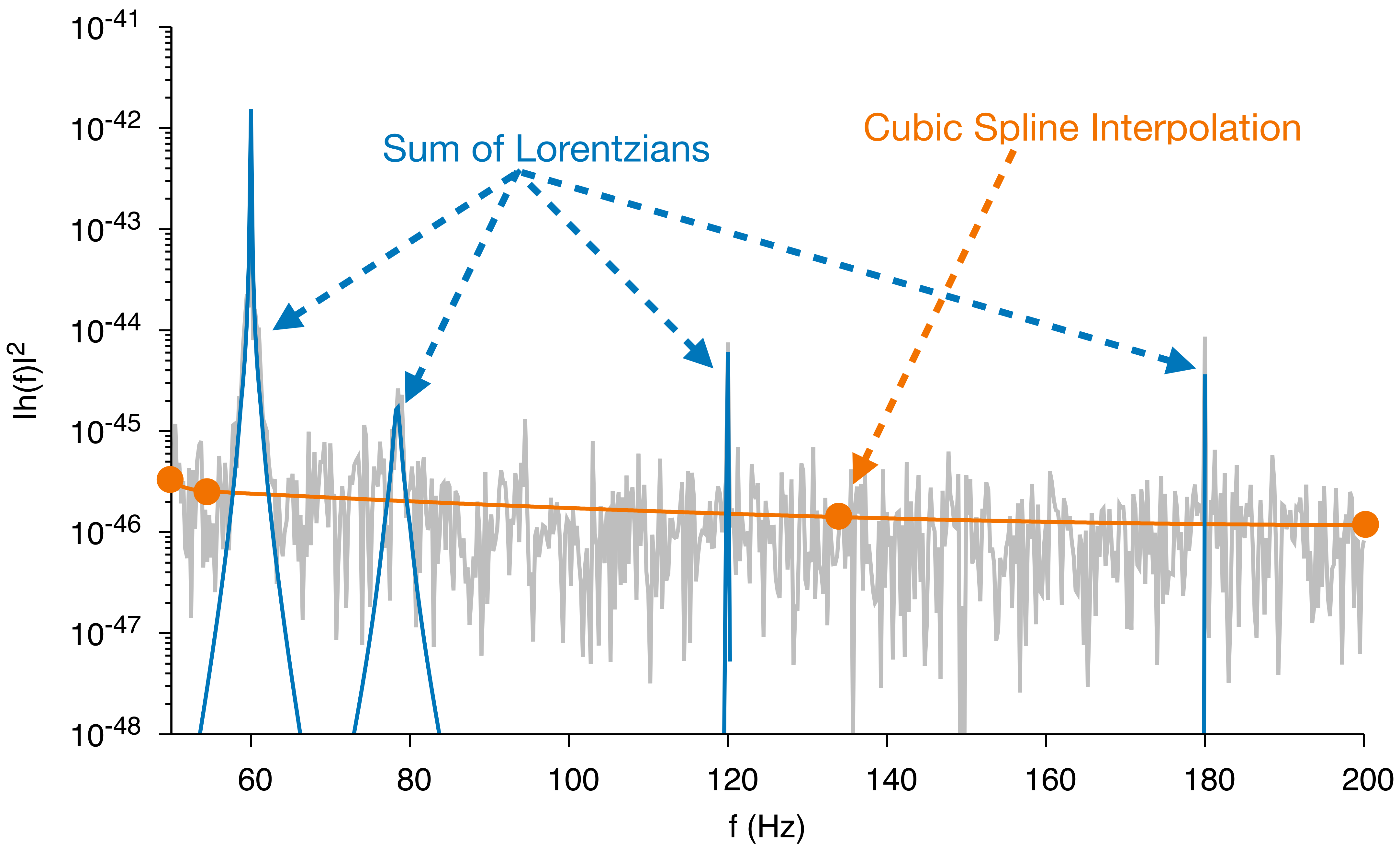
~~Noise is zero-mean Gaussian~~

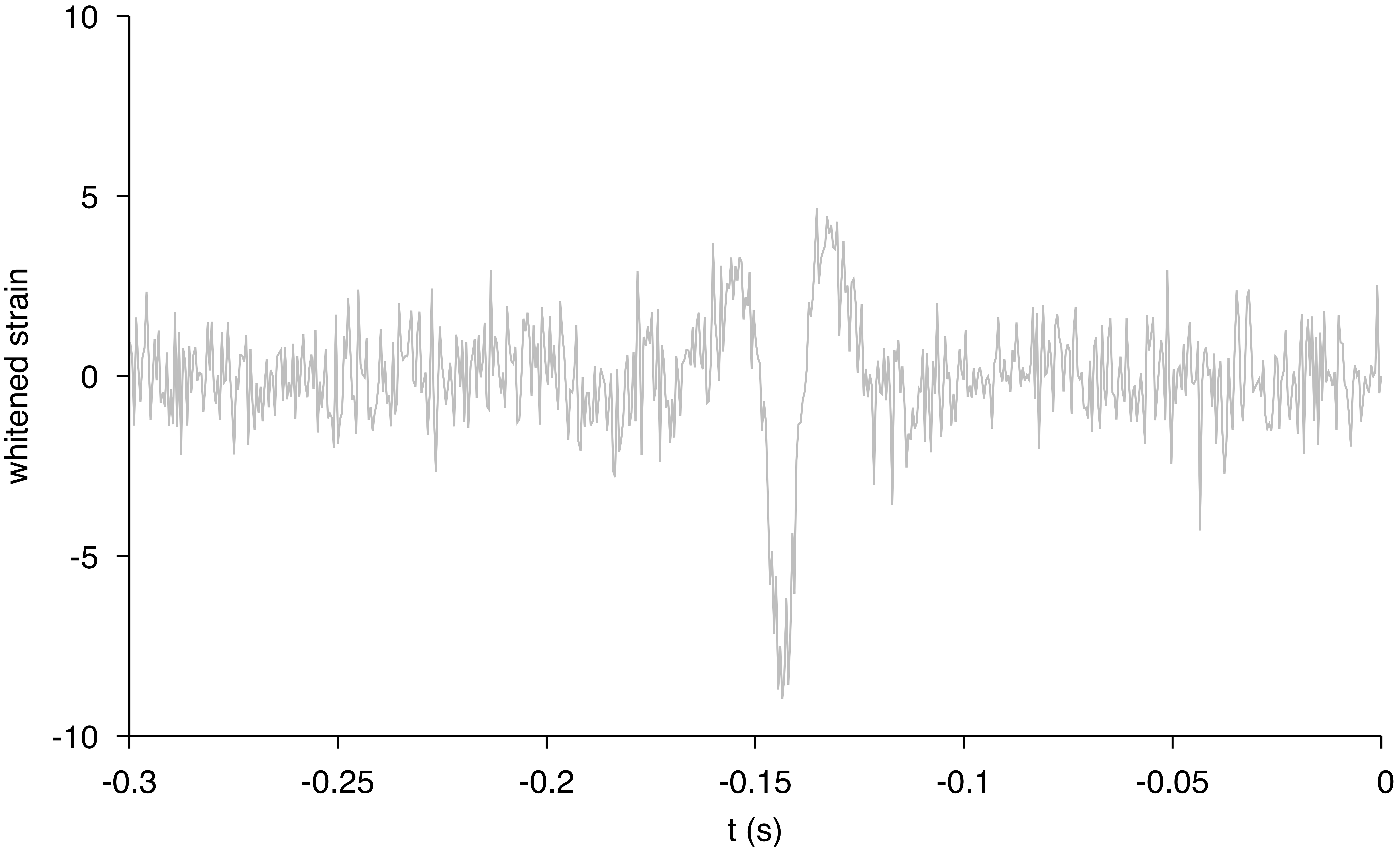
Noise has known variance ✔

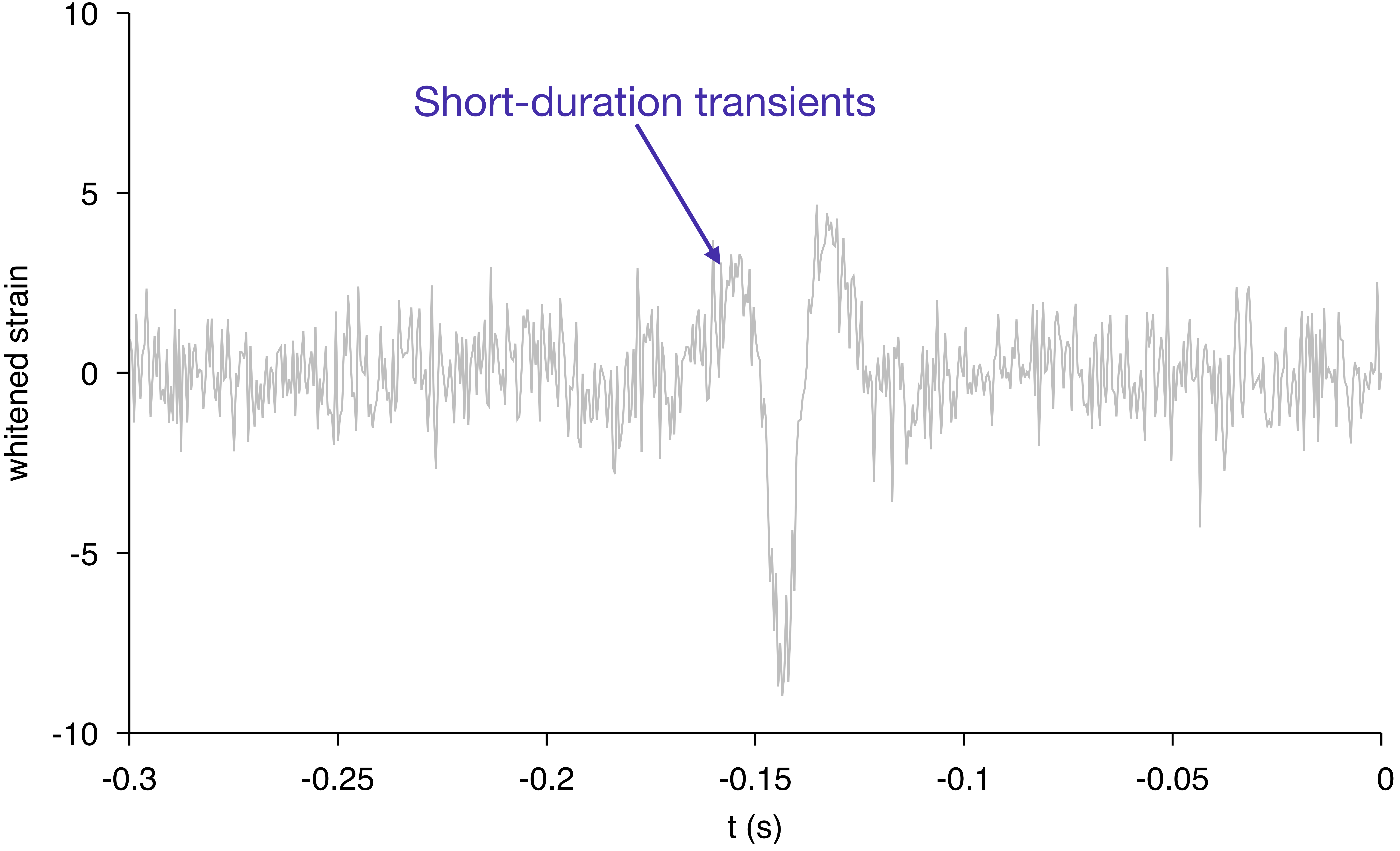
~~Data are perfectly calibrated~~

~~Waveform model is perfect~~

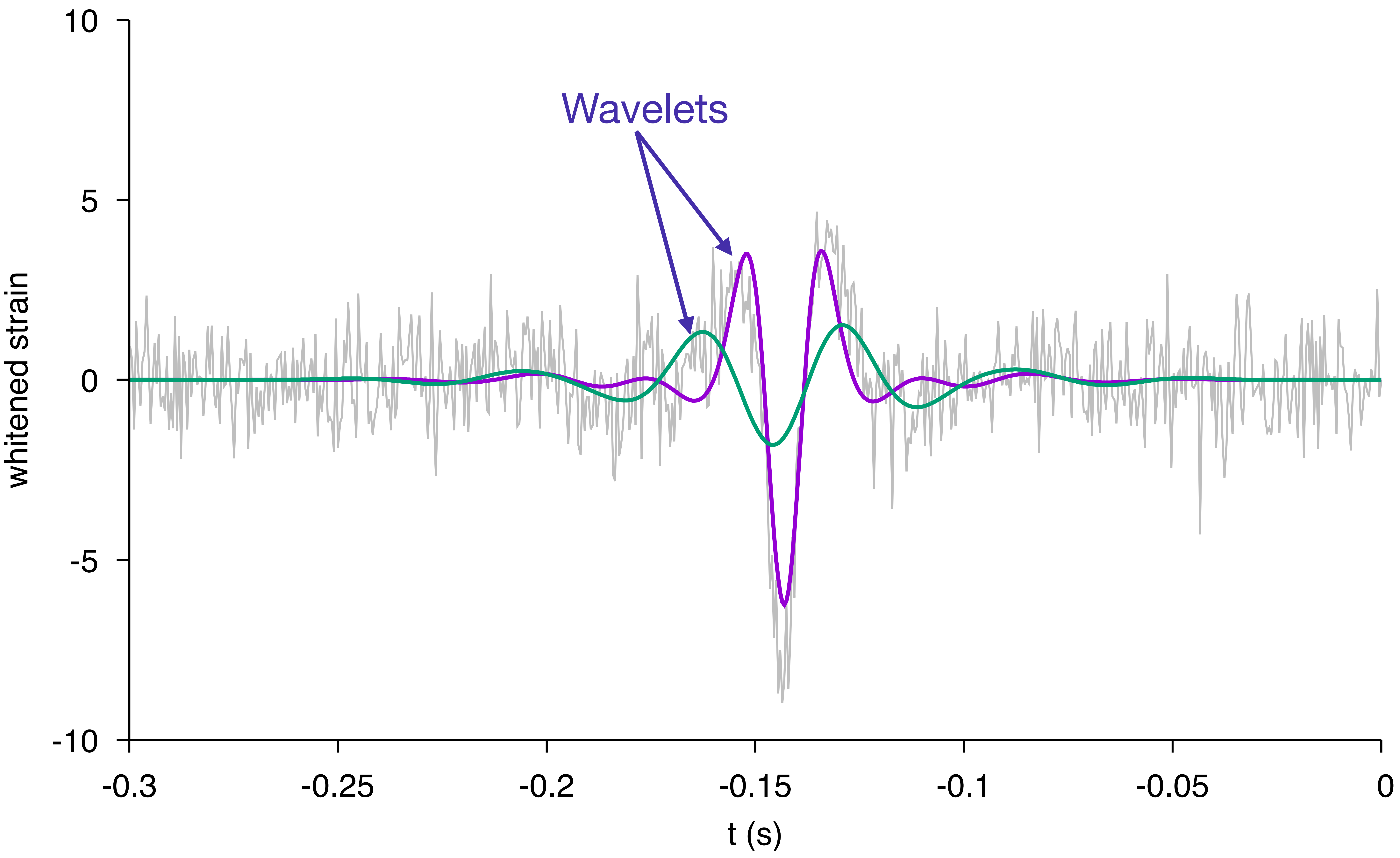
~~Noise variance is *stationary*~~ **OK**



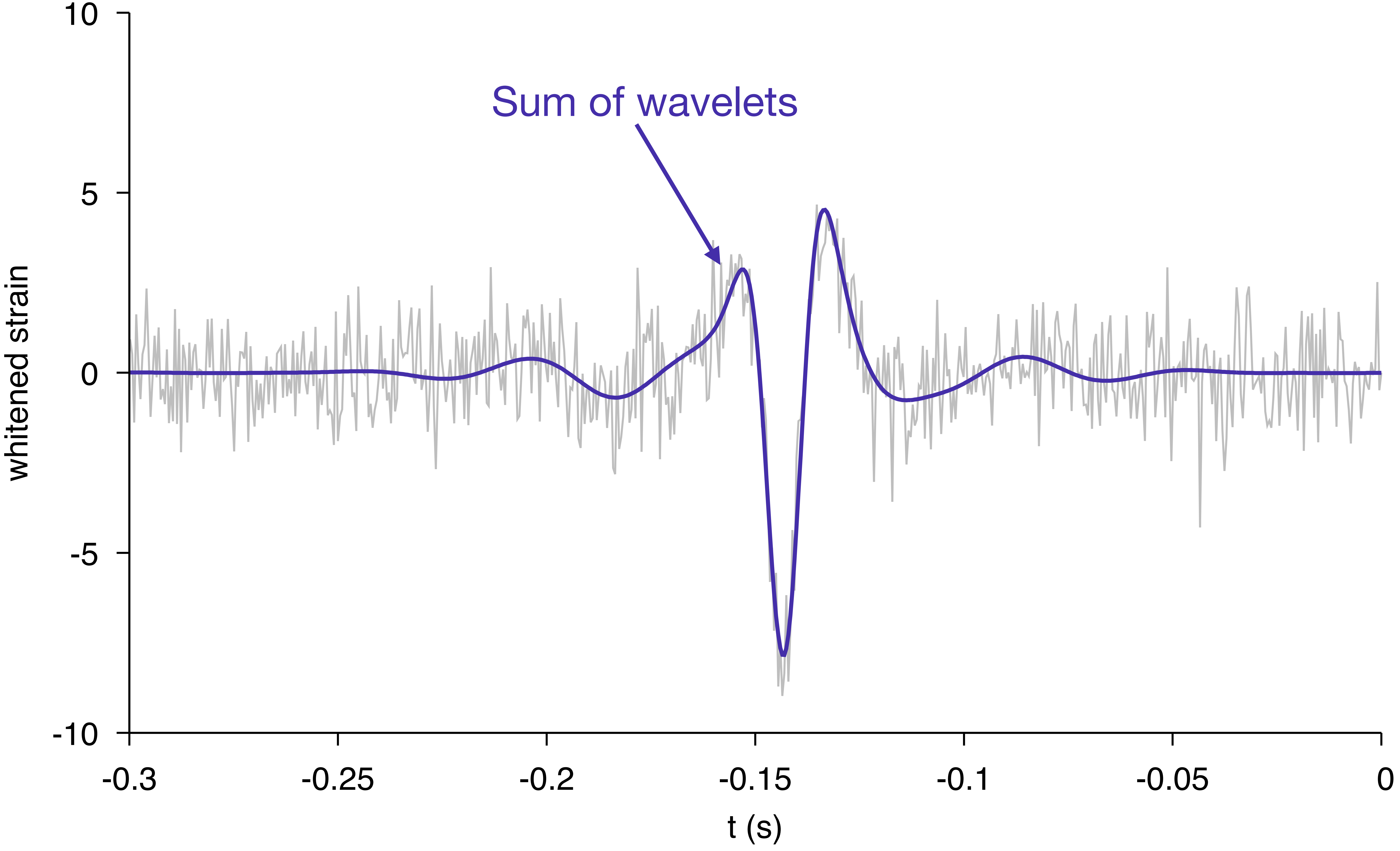

10
5
0
-5
-10
whitened strain
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
t (s)



Short-duration transients
whitened strain
10
5
0
-5
-10
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
t (s)



Wavelets
10
5
0
-5
-10
whitened strain
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
t (s)


Sum of wavelets
whitened strain
10
5
0
-5
-10
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
t (s)

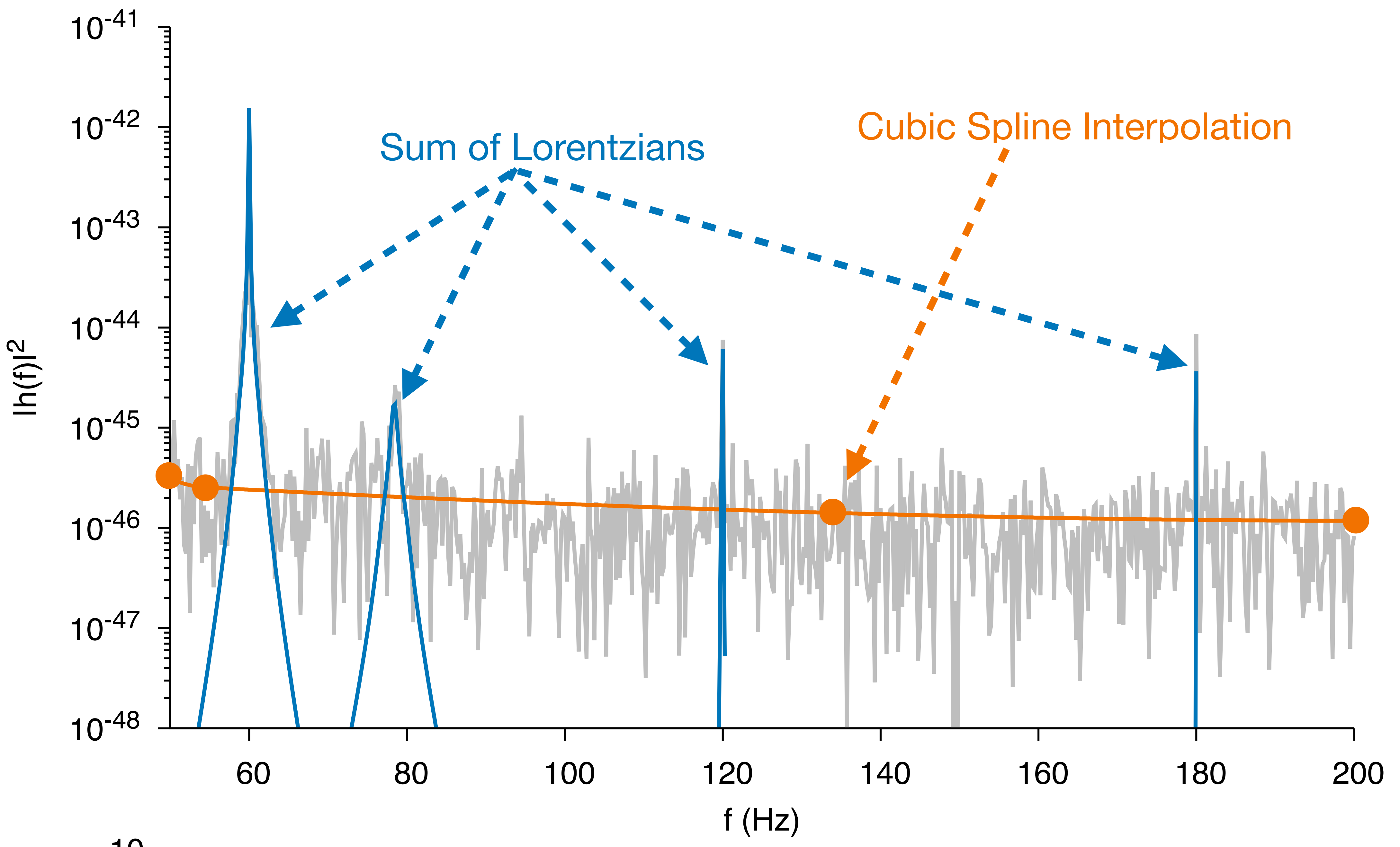
Noise is zero-mean Gaussian ✔

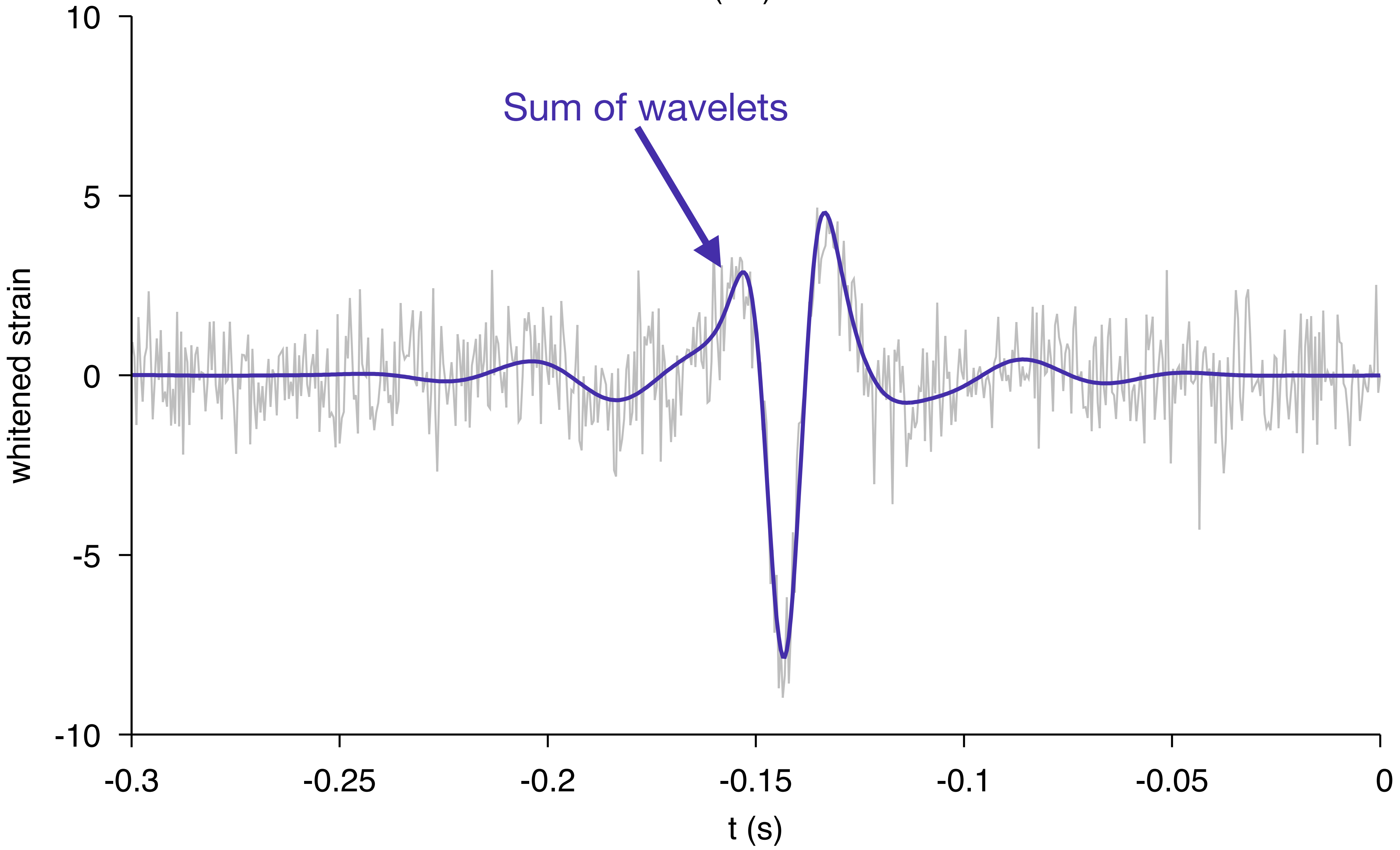
Noise has known variance ✔

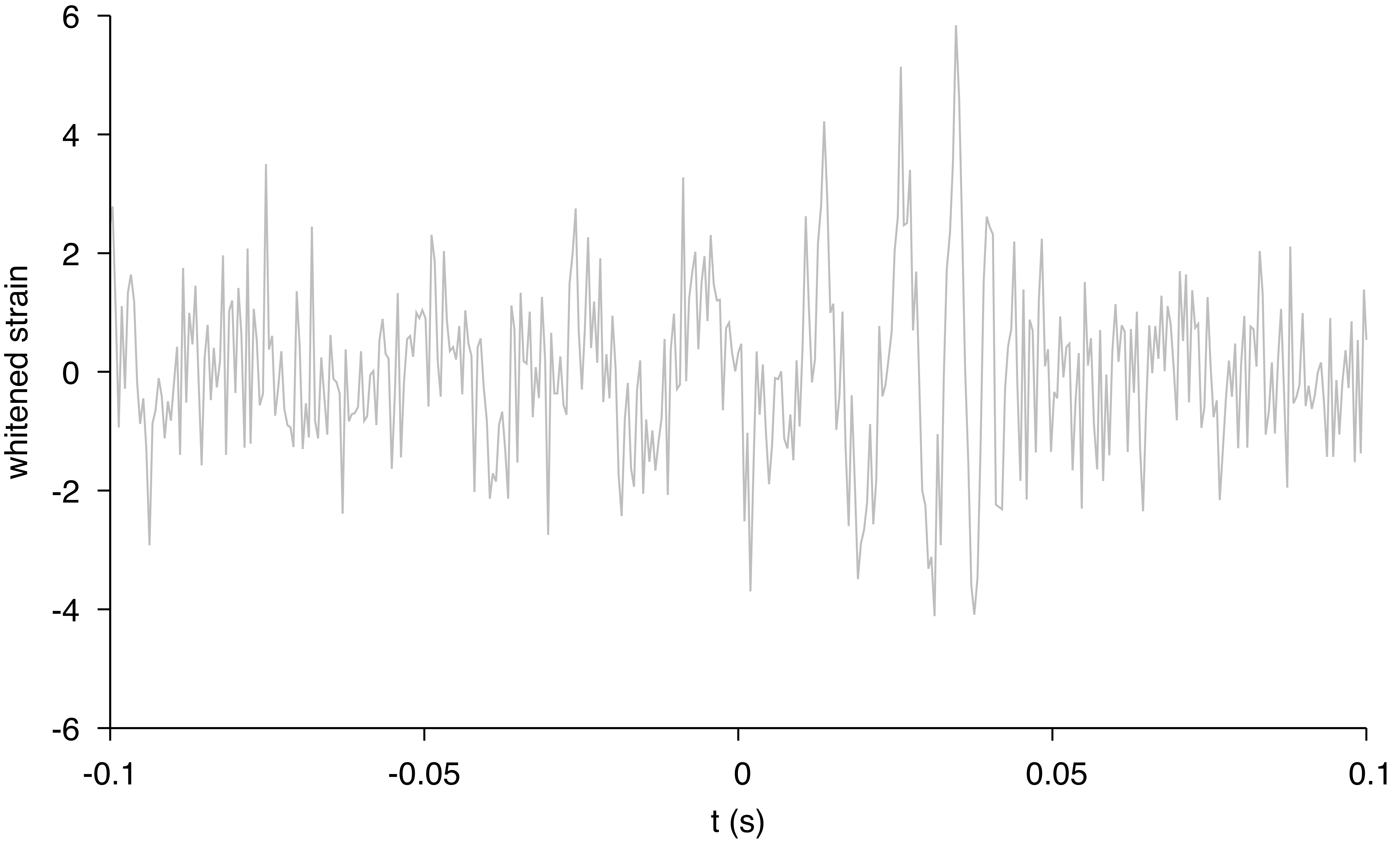
~~Data are perfectly calibrated~~

~~Waveform model is perfect~~

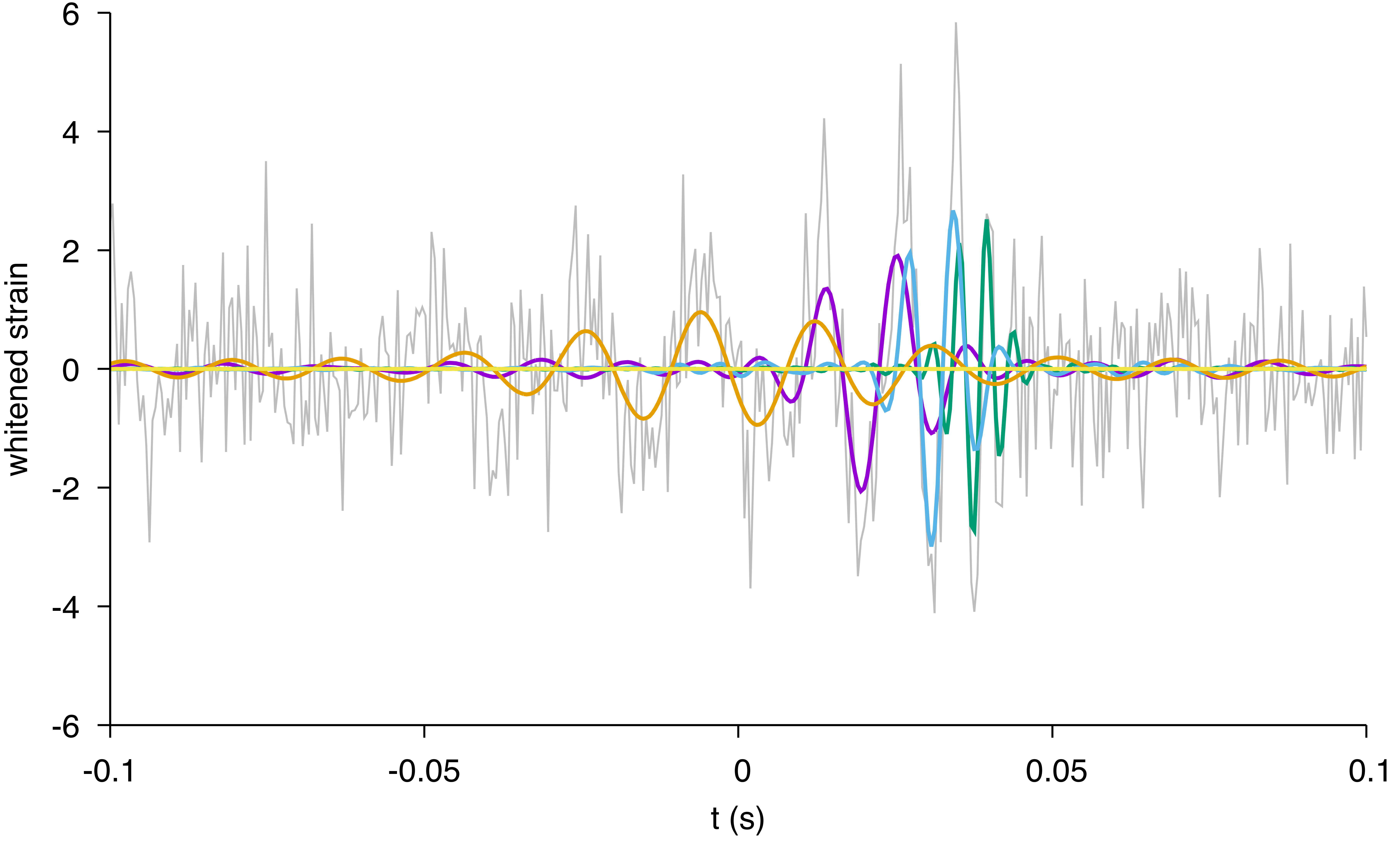
~~Noise variance is *stationary*~~ OK

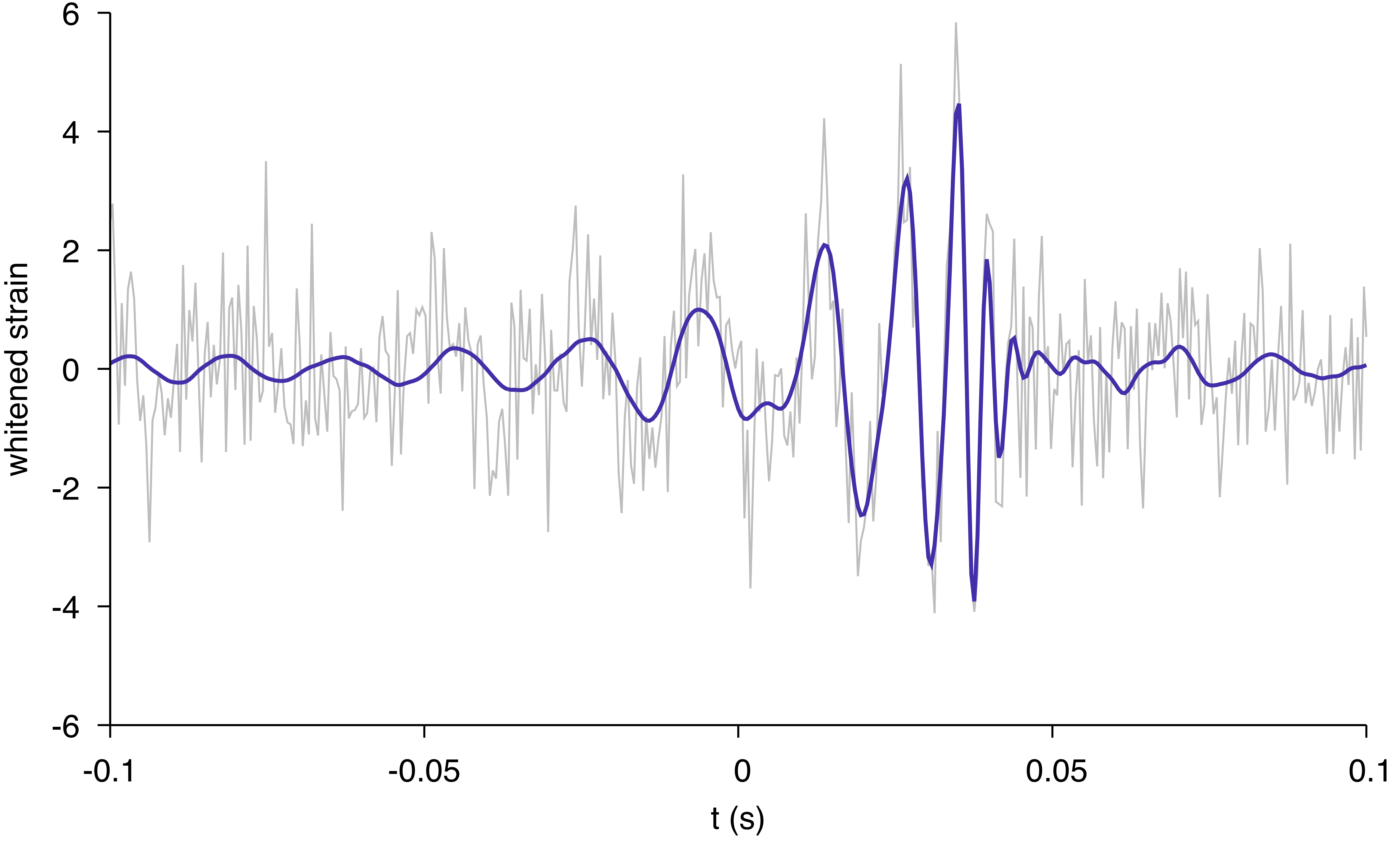




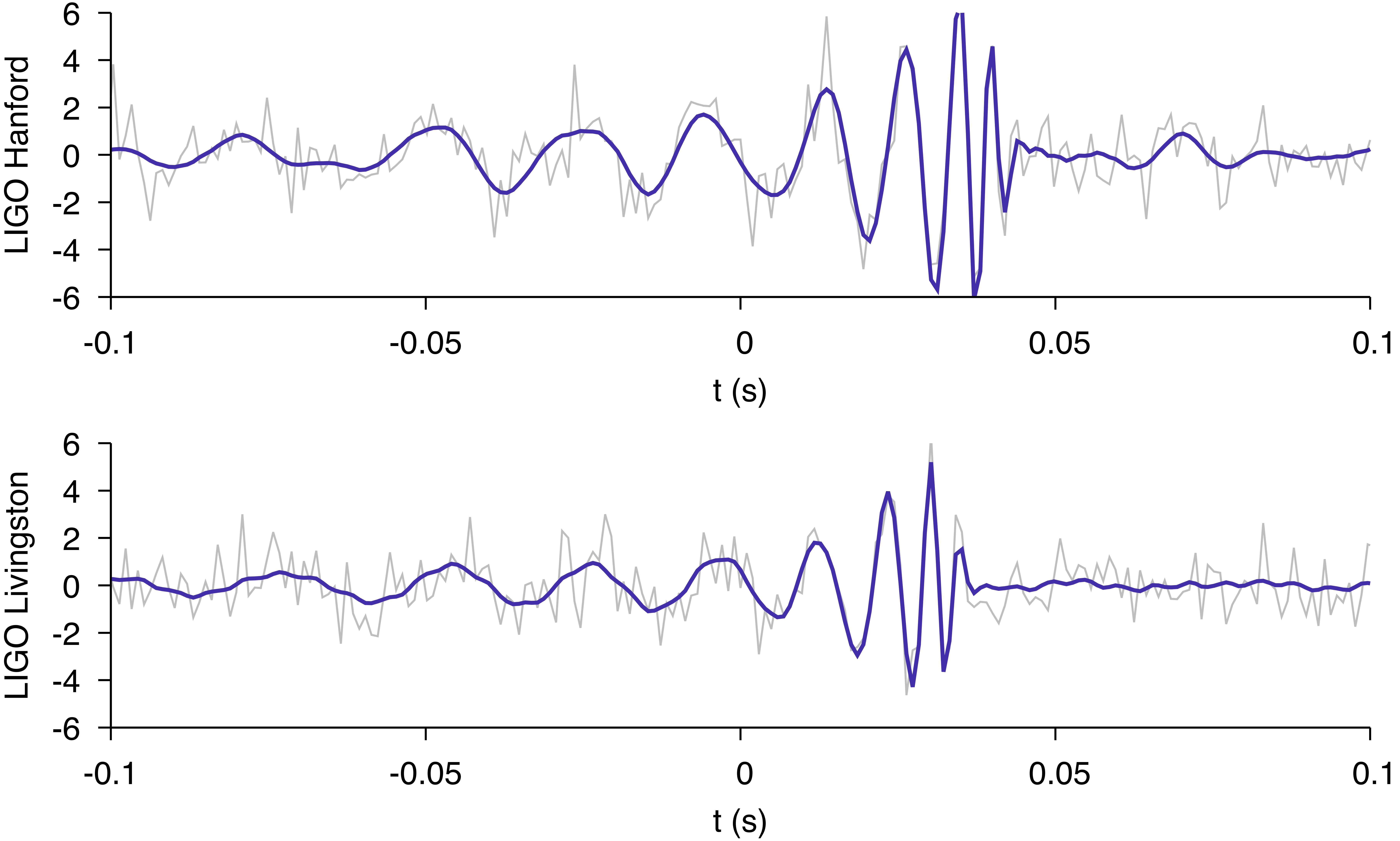


whitened strain
6
4
2
0
-2
-4
-6
-0.1
-0.05
0
0.05
0.1
t (s)


6
4
2
0
-2
-4
-6
whitened strain
-0.1
-0.05
0
0.05
0.1
t (s)



6
4
2
0
-2
-4
-6
whitened strain
-0.1
-0.05
0
0.05
0.1
t (s)



LIGO Hanford
6
4
2
0
-2
-4
-6
-0.1
-0.05
0
0.05
0.1
t (s)
LIGO Livingston
6
4
2
0
-2
-4
-6
-0.1
-0.05
0
0.05
0.1
t (s)

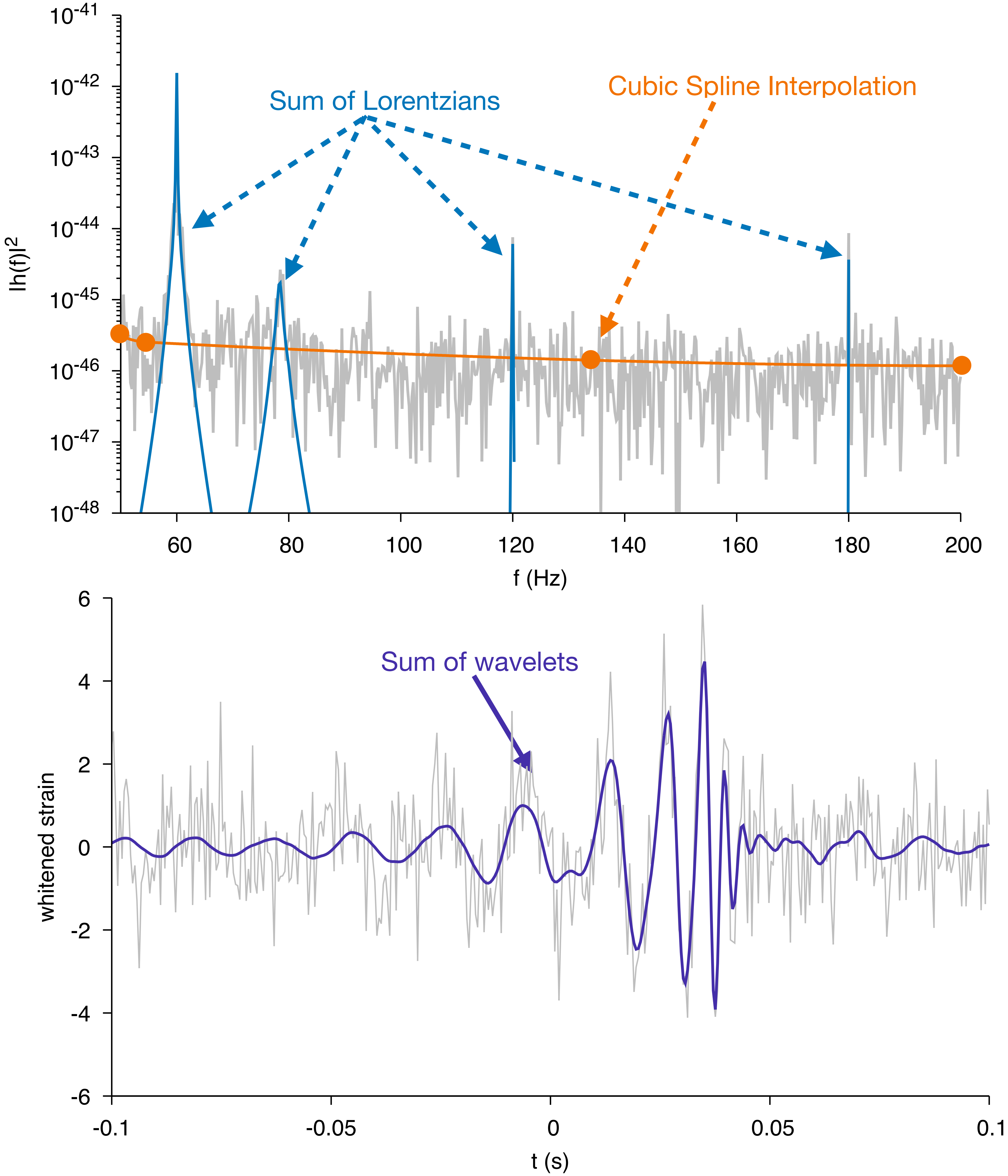
Noise is zero-mean Gaussian ✓

Noise has known variance ✓

~~Data are perfectly calibrated~~

~~Waveform model is perfect~~ **OK**

~~Noise variance is *stationary*~~ **OK**



$N_{\mathrm{spline}} \times \{f_0, S_n\}$: Spline Model

$N_{\mathrm{lines}} \times \{f_0, A, Q\}$: Line Model

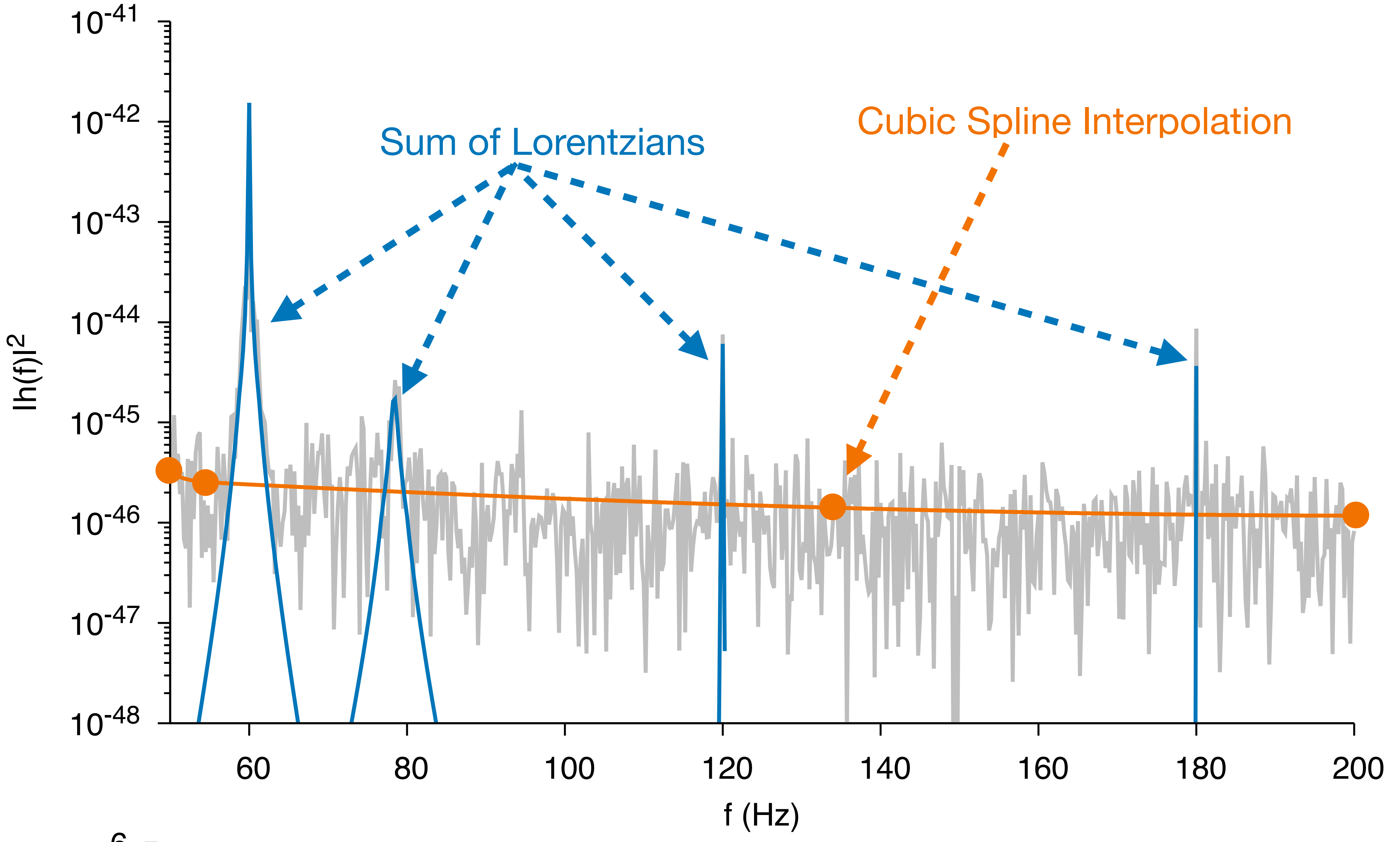
$N_{\mathrm{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

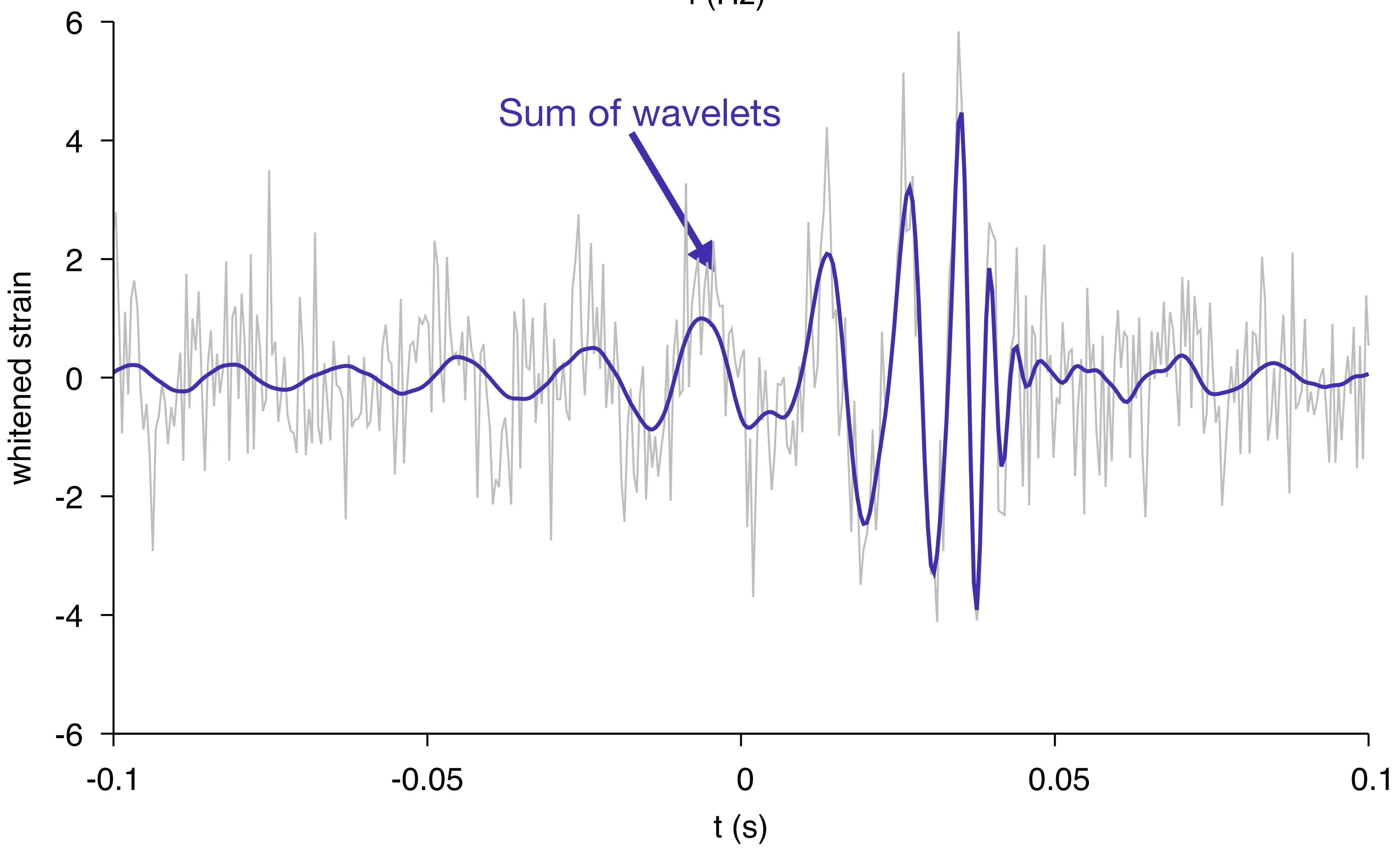
$N_{\mathrm{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\mathrm{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

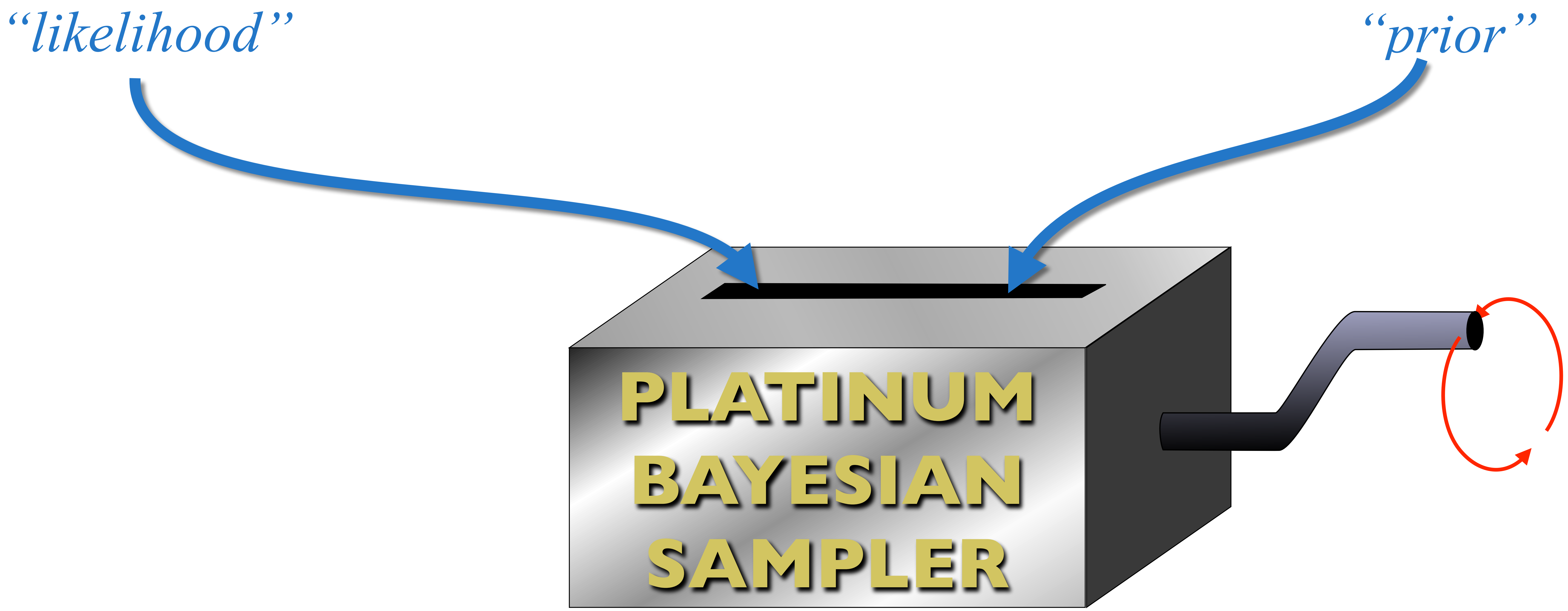




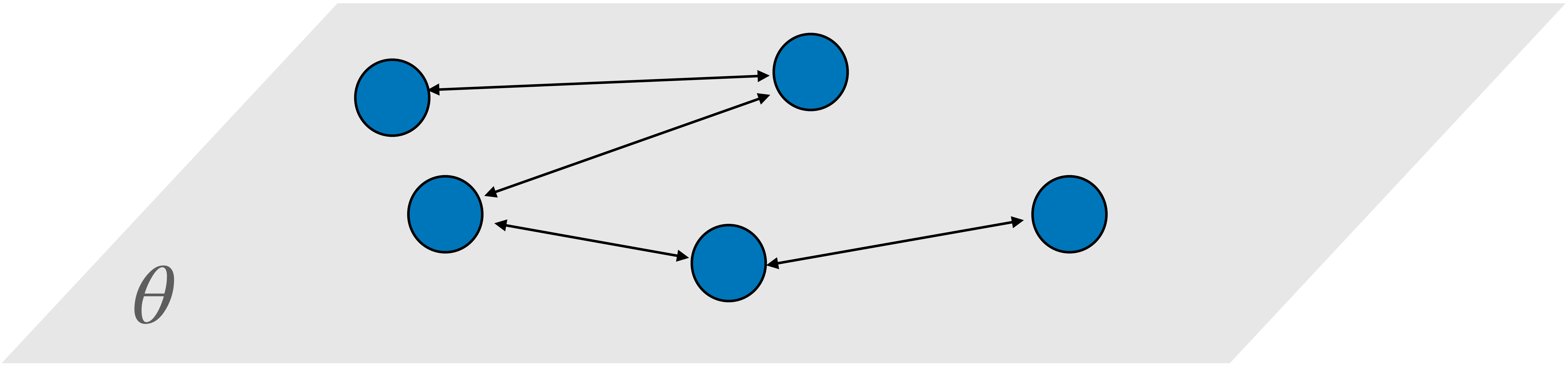
Model everything… likelihood $= p(\mathbf{d} \,|\, \text{signal, noise, glitch})$

Marginalize the stuff you don't care about…

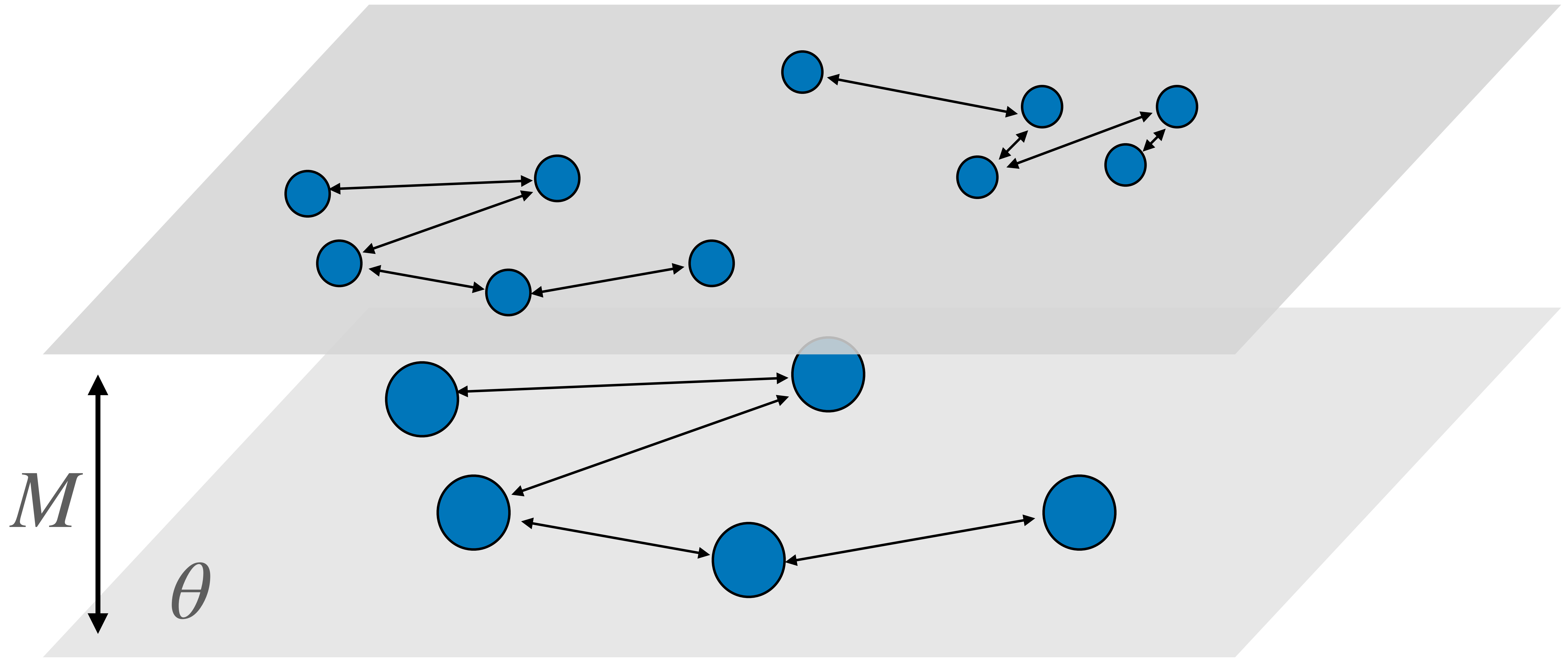
$$p(\text{signal} \,|\, \mathbf{d}) = \int_{\text{glitch,noise}} p(\mathbf{d} \,|\, \text{signal, glitch, noise})$$



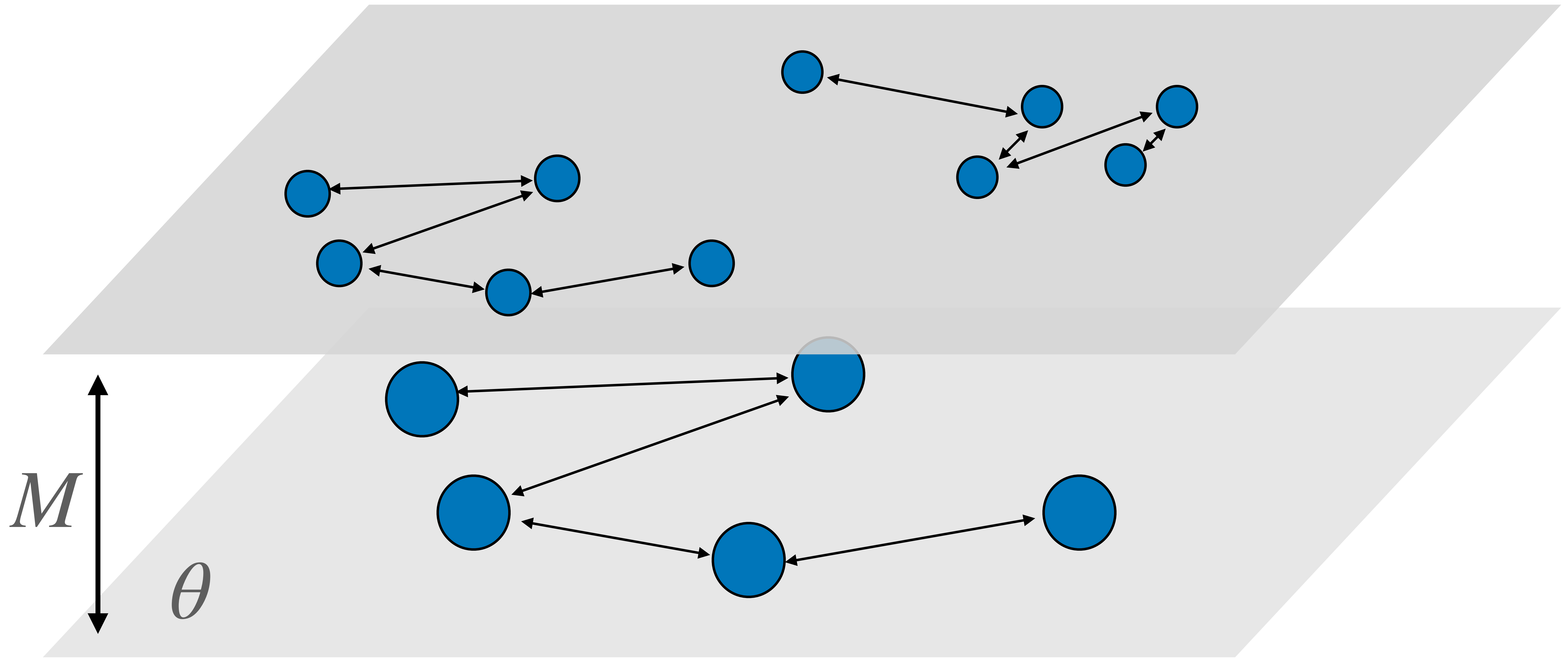
Transdimensional (Reversible Jump) MCMC



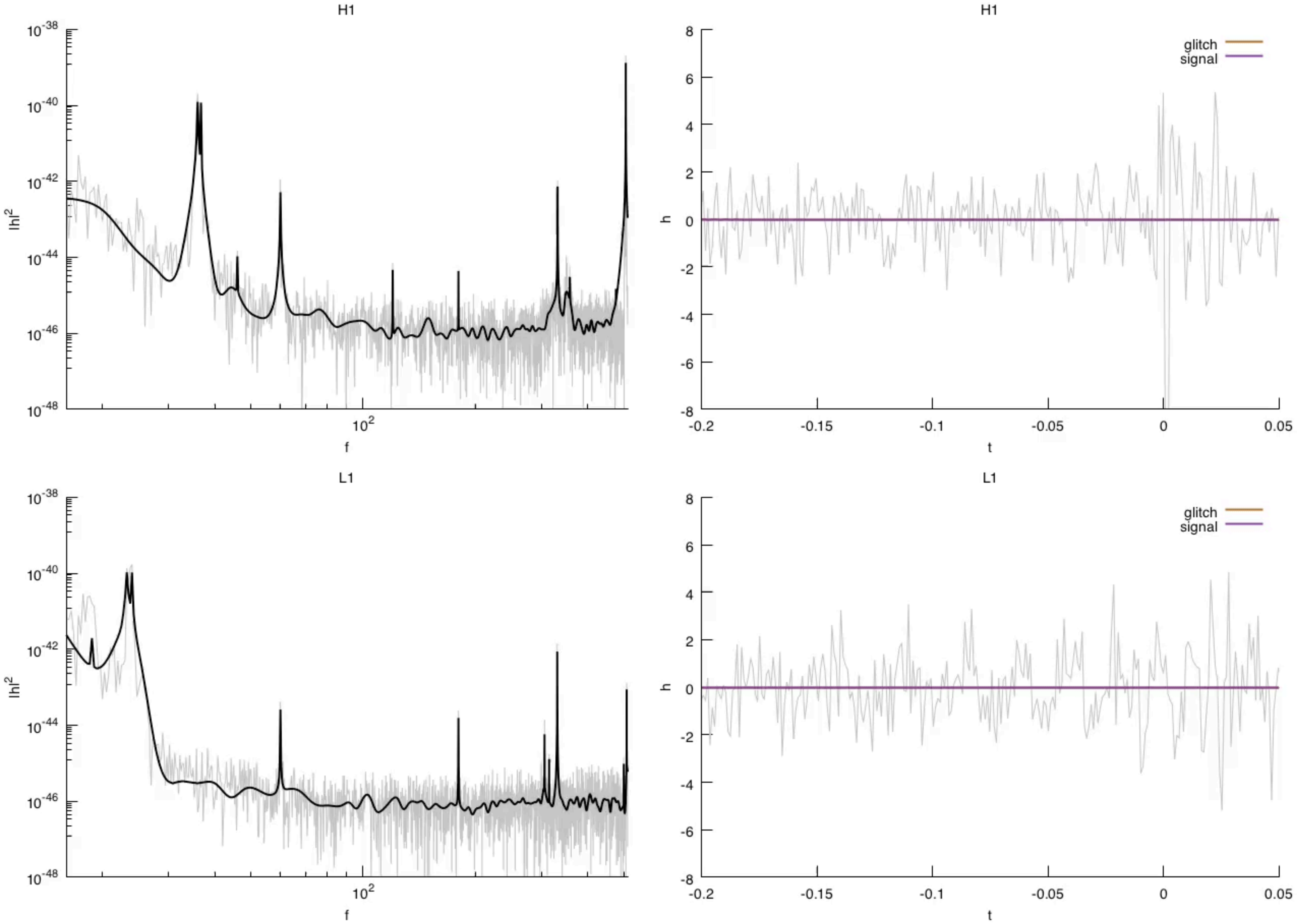
Transdimensional (Reversible Jump) MCMC

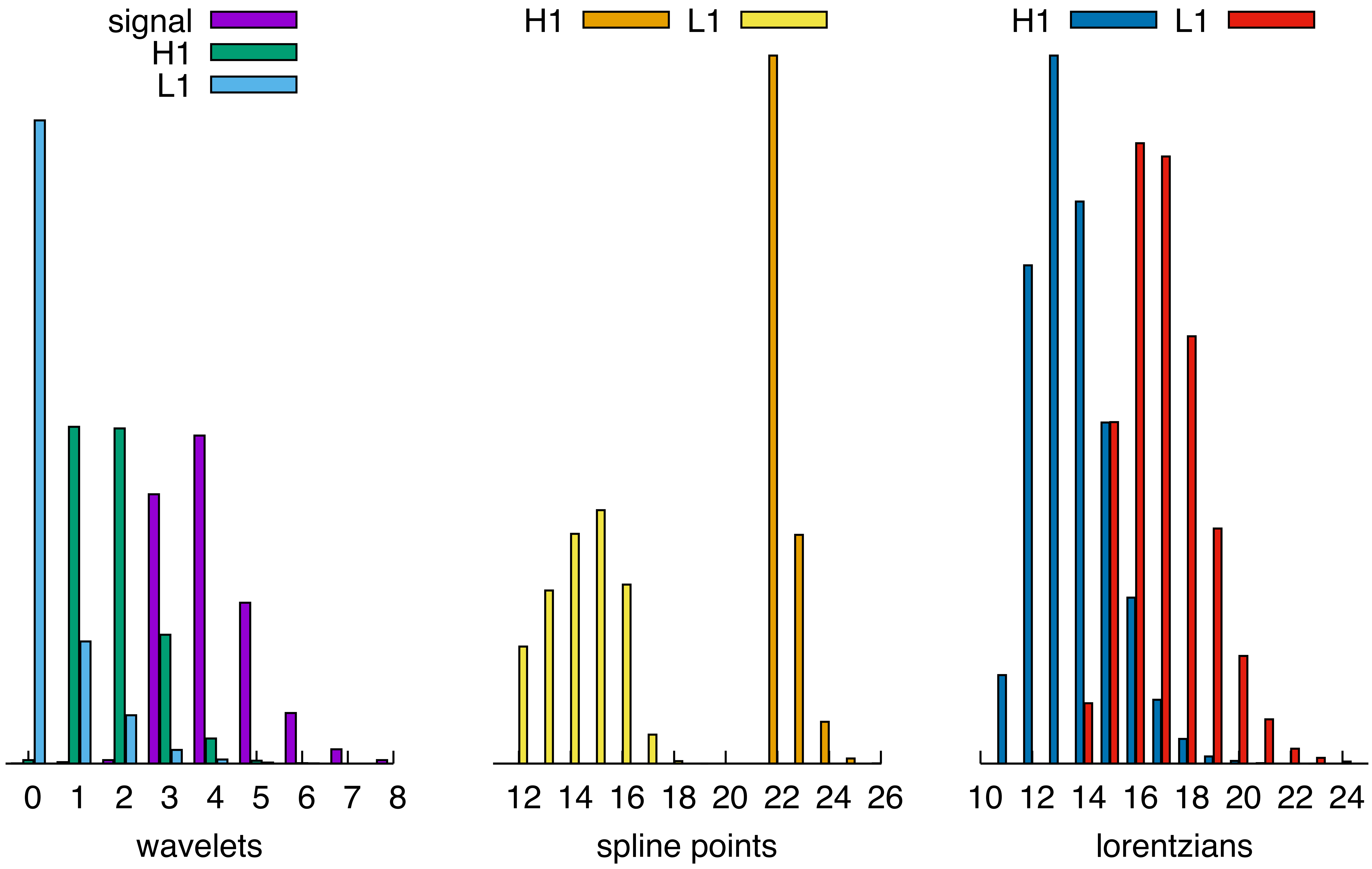


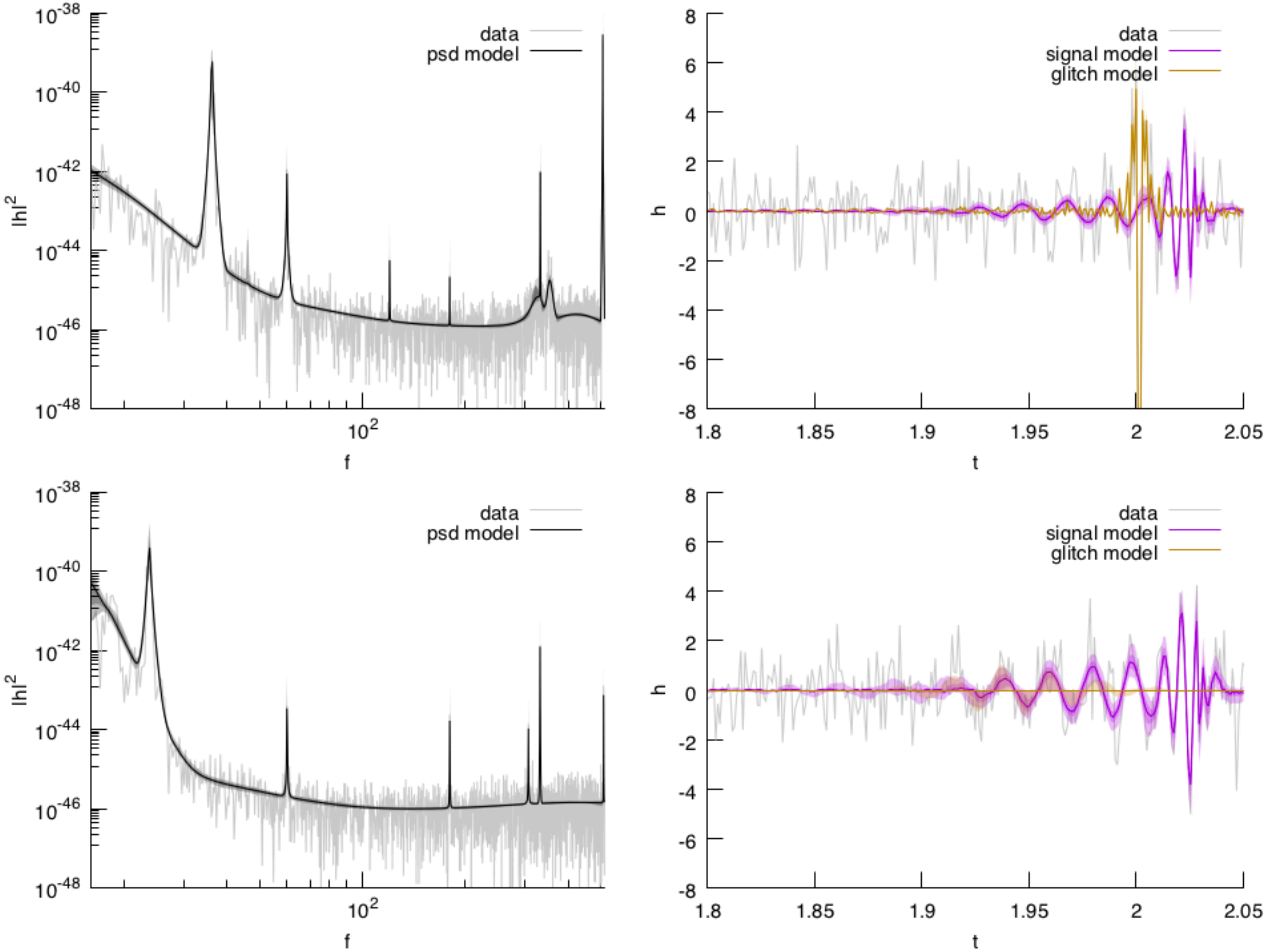
Transdimensional (Reversible Jump) MCMC



$$\frac{\text{counts in model A}}{\text{counts in model B}} \equiv \mathcal{O}_{A,B}$$


H1
|h|^2
f
10^2
H1
glitch
signal
h
t
L1
|h|^2
f
10^2
L1
glitch
signal
h
t


signal
H1
L1
0 1 2 3 4 5 6 7 8
wavelets
H1
L1
12 14 16 18 20 22 24 26
spline points
H1
L1
10 12 14 16 18 20 22 24
lorentzians

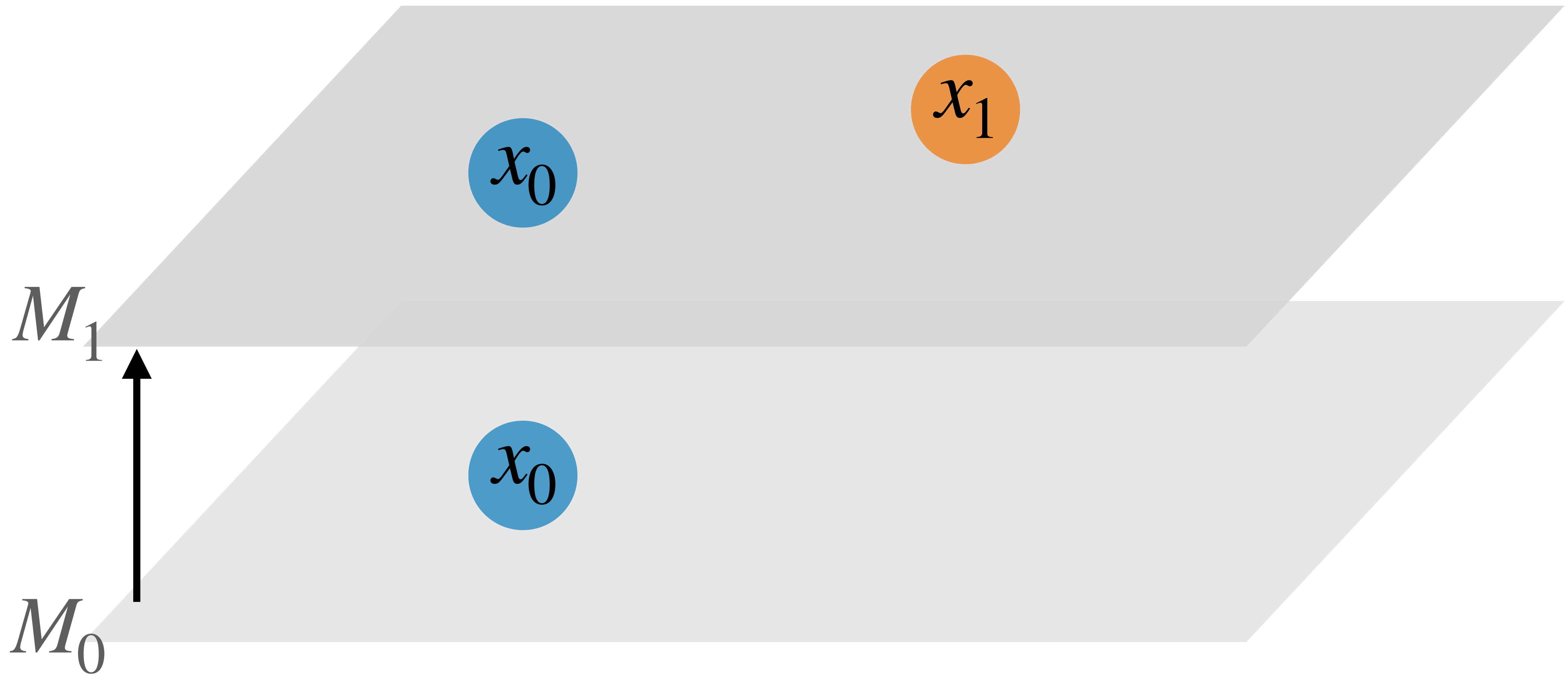


data
psd model
|h|²
f
data
signal model
glitch model
h
t

Transdimensional (Reversible Jump) MCMC
…are notoriously tricky to get mixing…

propose adding parameter set $x : \{x_0\} \rightarrow \{x_0, x_1\}$

$$\alpha_{M_0 \rightarrow M_1} = \min \left[1, \frac{p(\mathbf{d} \mid x_0, x_1)}{p(\mathbf{d} \mid x_0)} \frac{p(x_0)p(x_1)}{p(x_0)} \frac{1}{q(x_1)} \right]$$



Transdimensional (Reversible Jump) MCMC
…are notoriously tricky to get mixing…

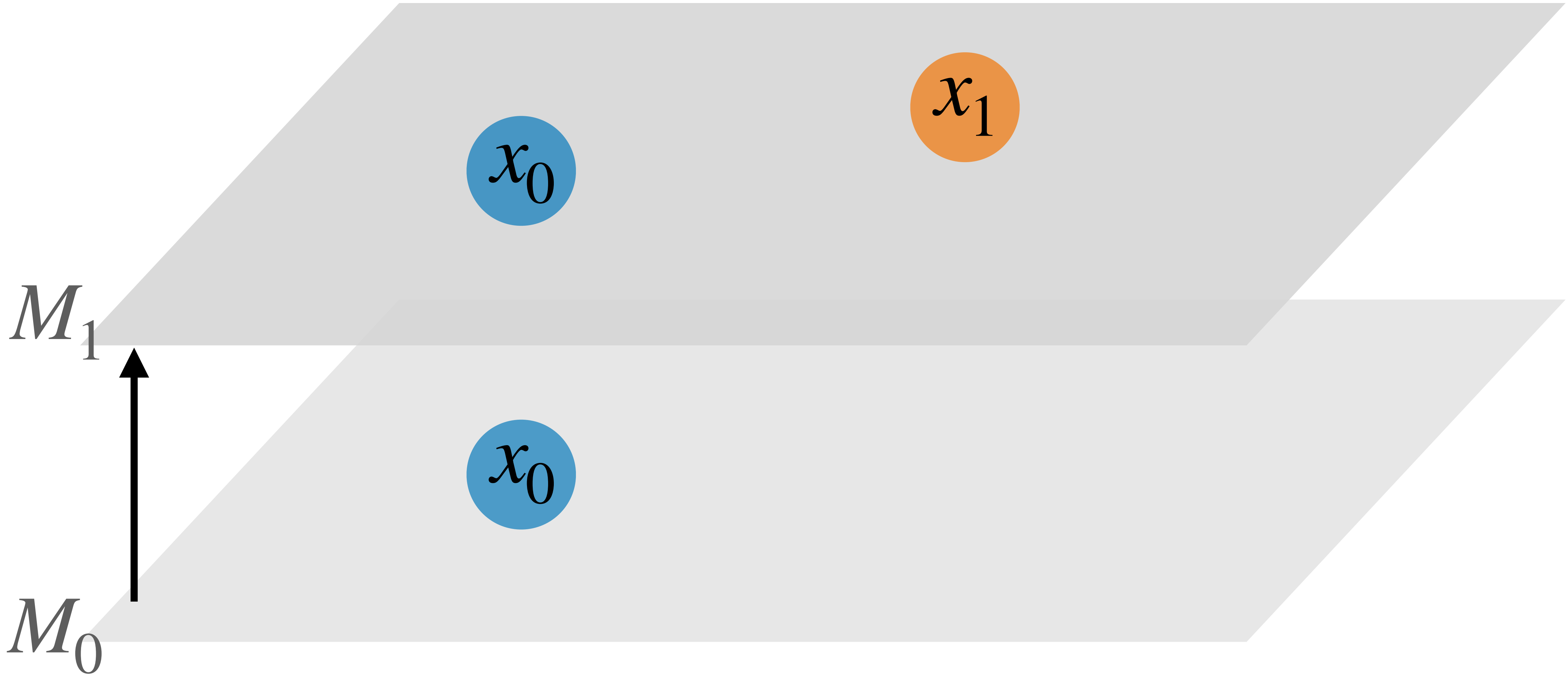
propose adding parameter set $x : \{x_0\} \rightarrow \{x_0, x_1\}$

$$\alpha_{M_0 \rightarrow M_1} = \min \left[1, \frac{p(\mathbf{d} \,|\, x_0, x_1)}{p(\mathbf{d} \,|\, x_0)} \frac{p(x_0)\, p(x_1)}{p(x_0)} \frac{1}{q(x_1)} \right]$$

<<1 unless close to bulk of posterior

penalty for adding prior volume

penalizes "heavy handed" proposal



Transdimensional (Reversible Jump) MCMC …are notoriously tricky to get mixing…

(i) Mixing benefits from domain knowledge / data-driven proposals

Transdimensional (Reversible Jump) MCMC …are notoriously tricky to get mixing…

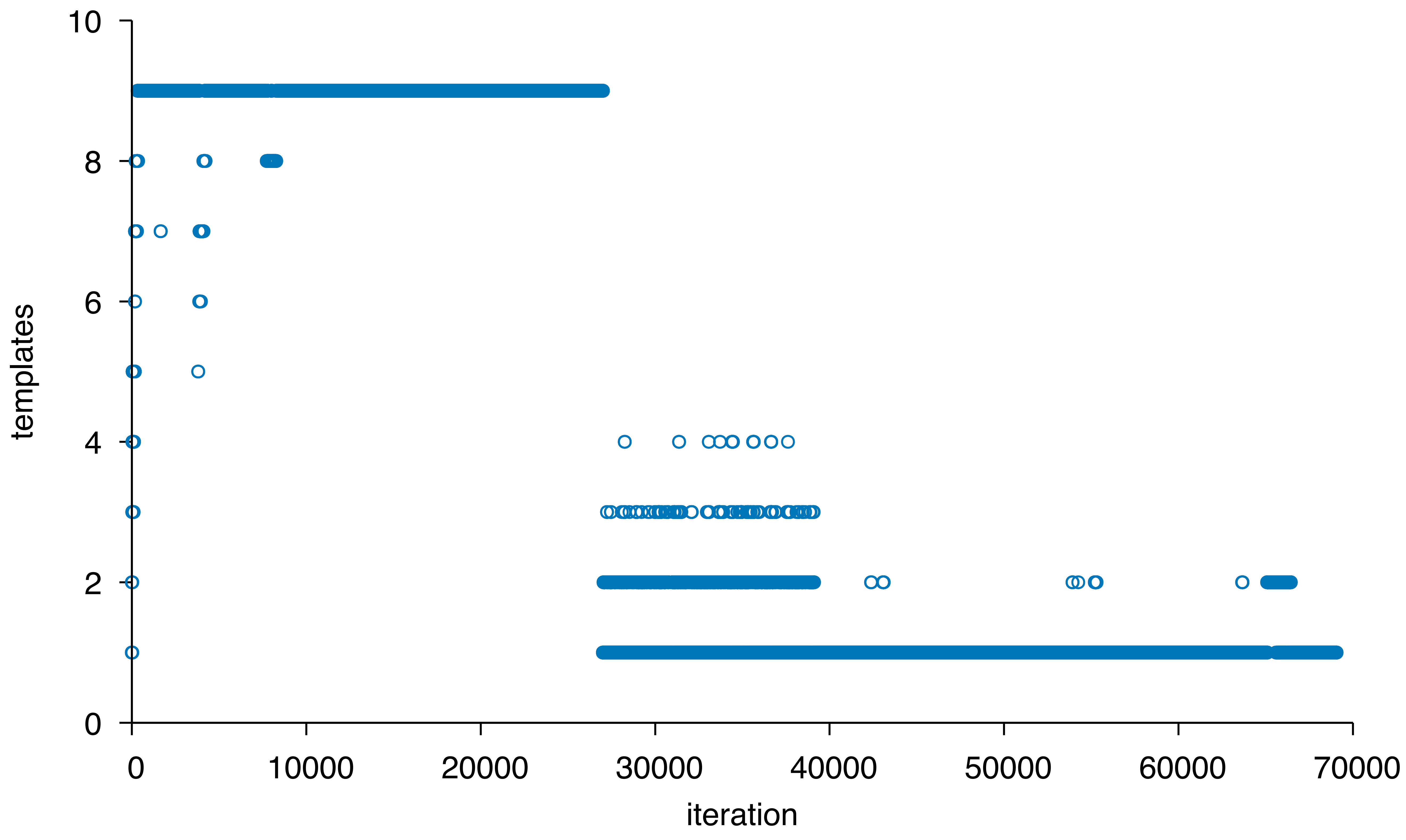
(i) Mixing benefits from domain knowledge / data-driven proposals

(ii) Convergence benefits from a helping-hand during burn-in

Transdimensional (Reversible Jump) MCMC …are notoriously tricky to get mixing…

(i) Mixing benefits from domain knowledge / data-driven proposals

(ii) Convergence benefits from a helping-hand during burn-in

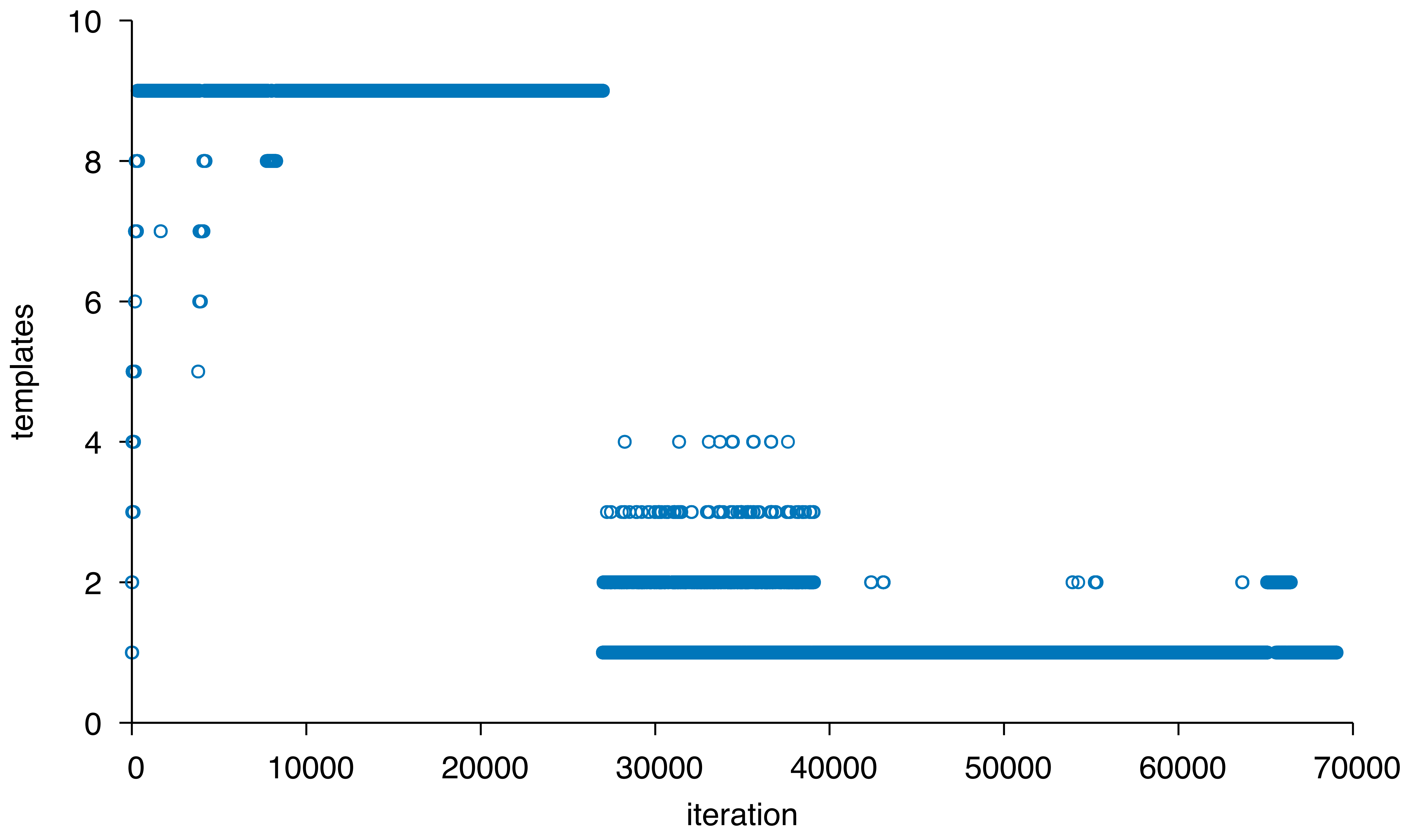


Transdimensional (Reversible Jump) MCMC …are notoriously tricky to get mixing…

(i) Mixing benefits from domain knowledge / data-driven proposals

(ii) Convergence benefits from a helping-hand during burn-in

(iii) Having (ii) should make you nervous about the robustness of the sampler



$N_{\text{spline}} \times \{f_0, S_n\}$: Spline Model

$N_{\text{lines}} \times \{f_0, A, Q\}$: Line Model

$N_{\text{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

$N_{\text{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\text{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

$N_{\mathrm{spline}} \times \{f_0, S_n\}$: Spline Model

$N_{\mathrm{lines}} \times \{f_0, A, Q\}$: Line Model

$N_{\mathrm{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

Point estimate of PSD for LIGO-Virgo CBC PE

$N_{\mathrm{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\mathrm{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

$N_{\rm spline} \times \{f_0, S_n\}$: Spline Model

Point estimate of PSD for LIGO-Virgo CBC PE

$N_{\rm lines} \times \{f_0, A, Q\}$: Line Model

Point estimate of Glitch Model for some CBC PE

$N_{\rm G,I} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

$N_{\rm S} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\rm cal} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

$N_{\mathrm{spline}} \times \{f_0, S_n\}$: Spline Model

Point estimate of PSD for LIGO-Virgo CBC PE

$N_{\mathrm{lines}} \times \{f_0, A, Q\}$: Line Model

Point estimate of Glitch Model for some CBC PE

$N_{\mathrm{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

Template-free CBC waveform reconstructions

$N_{\mathrm{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\mathrm{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

$N_{\rm spline} \times \{f_0, S_n\}$: Spline Model

Point estimate of PSD for LIGO-Virgo CBC PE

$N_{\rm lines} \times \{f_0, A, Q\}$: Line Model

Point estimate of Glitch Model for some CBC PE

$N_{\rm G,I} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

Template-free CBC waveform reconstructions

$N_{\rm S} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

Burst search/characterization

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\rm cal} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

$N_{\text{spline}} \times \{f_0, S_n\}$: Spline Model

$N_{\text{lines}} \times \{f_0, A, Q\}$: Line Model

$N_{\text{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

$N_{\text{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

and/or

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

$N_{\text{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model

Point estimate of PSD for LIGO-Virgo CBC PE

Point estimate of Glitch Model for some CBC PE

Template-free CBC waveform reconstructions

Burst search/characterization

New use-cases in development for O4

$N_{\mathrm{spline}} \times \{f_0, S_n\}$: Spline Model

Point estimate of PSD for LIGO-Virgo CBC PE

$N_{\mathrm{lines}} \times \{f_0, A, Q\}$: Line Model

Point estimate of Glitch Model for some CBC PE

$N_{\mathrm{G,I}} \times \{f_0, t_0, A, Q, \phi_0\}$: Glitch Model

Template-free CBC waveform reconstructions

$N_{\mathrm{S}} \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\}$: Generic Signal Model

Burst search/characterization

and/or

New use-cases in development for O4

$\{m_1, m_2, \mathbf{S}_1, \mathbf{S}_2, \mathbf{L}, \alpha, \delta, D_L, t_0\}$: CBC Model

Similar algorithms under development for LISA

$N_{\mathrm{cal}} \times \{\delta A_I, \delta\phi_{IFO}\}$: Calibration Model