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 \( B \)
\[ p(\theta | d, M) = \frac{p(d | \theta, M)p(\theta | M)}{p(d | M)} \]

Probability (density) of parameters given data \& model
ACME BAYESIAN SAMPLER

\[ p(d | \theta, M) \]

“likelihood”

\[ p(\theta | M) \]

“prior”
ACME
BAYESIAN
SAMPLER

\[ p(d \mid \theta, M) \]  
“likelihood”

\[ p(\theta \mid M) \]  
“prior”

“posterior”

\[ p(\theta \mid d, M) \]

“evidence”

\[ p(d \mid M) \]
\[ p(d | \theta, M) \]

"likelihood"

\[ p(\theta | M) \]

"prior"

\[ p(\theta | d, M) \]

"posterior"

\[ p(d | M) \]

"evidence"
\[ \nu \in \delta V = \frac{\text{# of samples in } \delta V}{\text{# of samples in total}} \]

\[ p(x \mid d, M) = \int p(x, y, z \mid d, M) \, dy \, dz \]

“nuisance parameters”
\[ p(d \mid \theta, M) \]  

“likelihood”

\[ p(\theta \mid M) \]  

“prior”

\[ p(\theta \mid d, M) \]  

“posterior”

\[ p(d \mid M) \]  

“evidence”
\[ p(d \mid M) = \int d\theta \, p(d \mid \theta, M) \, p(\theta \mid M) \]
\[ p(d \mid M) = \int d\theta \, p(d \mid \theta, M) \, p(\theta \mid M) \]

"I've detected Gravitational Waves!"

\[ \hat{O}_{A,B} = \frac{p(M_A)}{p(M_B)} \times \frac{p(d \mid M_A)}{p(d \mid M_B)} \]

"I've measured lots of noise!"

"odds ratio"

"Bayes factor"
\[ p(d \mid M) = \int d\theta \ p(d \mid \theta, M) \ p(\theta \mid M) \]

"I've detected Gravitational Waves!"

\[ \bar{\mathcal{O}}_{A,B} = \frac{p(M_A)}{p(M_B)} \times \frac{p(d \mid M_A)}{p(d \mid M_B)} \]

"odds ratio"

"Bayes factor"

\[ \bar{\mathcal{O}}_{A,B} = X \equiv \text{Model A is preferred over model B with } X : 1 \text{ odds} \]
Bayesian Analyses: Not magic.
Let’s Build a Likelihood Function
\( d \)
d = n
\[ d = n + h \]
Noise is modeled statistically

GWs are modeled coherently (discrete sources) or statistically (backgrounds)
\[ d = n + h \]

Noise is zero-mean Gaussian
Noise has known variance

\[ p(n_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|n_i|^2}{2\sigma^2}} \]
\[ d = n + h \]

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:

\[
p(d | \theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(d-h)^T C^{-1}(d-h)}
\]
\[ d = n + h \]

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:

\[
p(d | \theta) = \frac{1}{\sqrt{(2\pi)^k \det C}} e^{-\frac{1}{2}(d-h)^T C^{-1} (d-h)}
\]
\[ d = n + h \]

- 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:

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

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

- 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:

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

\[
p(d | \theta) \propto e^{-\frac{2}{T} \sum_k \frac{|\tilde{d}_k - \tilde{h}_k|^2}{S_{n,k}}} \]
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.
Broadband noise level

Narrowband spectral lines

<table>
<thead>
<tr>
<th>f (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
</tr>
<tr>
<td>80</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>120</td>
</tr>
<tr>
<td>140</td>
</tr>
<tr>
<td>160</td>
</tr>
<tr>
<td>180</td>
</tr>
<tr>
<td>200</td>
</tr>
</tbody>
</table>

\[ |h(f)|^2 \]
Cubic Spline Interpolation

Sum of Lorentzians
Noise is zero-mean Gaussian
Noise has known variance
Data are perfectly calibrated
Waveform model is perfect
Noise variance is stationary
Short-duration transients

whitened strain

t (s)
Noise is zero-mean Gaussian
Noise has known variance
Data are perfectly calibrated
Waveform model is perfect
Noise variance is stationary

\[ |h(f)| \]

\[ f \text{ (Hz)} \]

\[ \ln(x^2) \]

\[ f_\text{cubic} \]

\[ f_\text{Lorentz} \]

\[ f_\text{wavelet} \]
whitened strain

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

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

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

\( N_S \times \{ f_0, t_0, A, Q, \phi_0 \} \cup \{ \alpha, \delta, \psi, \epsilon \} : \) Generic Signal Model

and/or

\( \{ m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0 \} : \) CBC Model

\( N_{\text{cal}} \times \{ \delta A_I, \delta \phi_{IFO} \} : \) Calibration Model
Model everything… likelihood = $p(d \mid \text{signal, noise, glitch})$

Marginalize the stuff you don’t care about…

$$p(\text{signal} \mid d) = \int_{\text{glitch,noise}} p(d \mid \text{signal, glitch, noise})$$
Transdimensional (Reversible Jump) MCMC
Transdimensional (Reversible Jump) MCMC
Transdimensional (Reversible Jump) MCMC

$$\frac{\text{counts in model } A}{\text{counts in model } B} \equiv \mathcal{O}_{A,B}$$
Transdimensional (Reversible Jump) MCMC …are notoriously tricky to get mixing…

propose adding parameter set \( x : \{ x_0 \} \to \{ x_0, x_1 \} \)

\[
\alpha_{M_0 \to M_1} = \min \left[ 1, \frac{p(d | x_0, x_1)}{p(d | x_0)} \frac{p(x_0)p(x_1)}{p(x_0)} \frac{1}{q(x_1)} \right]
\]
Transdimensional (Reversible Jump) MCMC

...are notoriously tricky to get mixing...

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

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

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

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

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

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

and/or

$\{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\} : \text{CBC Model}$

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

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

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

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

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

and/or

$\{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\}$: CBC Model

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

Point estimate of PSD for LIGO-Virgo CBC PE
\( N_{\text{spline}} \times \{ f_0, S_n \} : \text{Spline Model} \)

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

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

\( N_S \times \{ f_0, t_0, A, Q, \phi_0 \} \cup \{ \alpha, \delta, \psi, \epsilon \} : \text{Generic Signal Model} \)

and/or

\( \{ m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0 \} : \text{CBC Model} \)

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

Point estimate of PSD for LIGO-Virgo CBC PE

Point estimate of Glitch Model for some CBC PE
\begin{align*}
N_{\text{spline}} \times \{f_0, S_n\} & : \text{Spline Model} \\
N_{\text{lines}} \times \{f_0, A, Q\} & : \text{Line Model} \\
N_{G,I} \times \{f_0, t_0, A, Q, \phi_0\} & : \text{Glitch Model} \\
N_S \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\} & : \text{Generic Signal Model}
\end{align*}

and/or

\{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\} : \text{CBC Model}

\begin{align*}
N_{\text{cal}} \times \{\delta A_I, \delta \phi_{IFO}\} & : \text{Calibration Model}
\end{align*}

Point estimate of PSD for LIGO-Virgo CBC PE

Point estimate of Glitch Model for some CBC PE

Template-free CBC waveform reconstructions
$N_{\text{spline}} \times \{f_0, S_n\} : \text{Spline Model}$

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

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

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

and/or

$\{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\} : \text{CBC Model}$

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

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

Point estimate of PSD for LIGO-Virgo CBC PE

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

Point estimate of Glitch Model for some CBC PE

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

Template-free CBC waveform reconstructions

\( N_S \times \{f_0, t_0, A, Q, \phi_0\} \cup \{\alpha, \delta, \psi, \epsilon\} : \text{Generic Signal Model} \)

Burst search/characterization

and/or

\( \{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\} : \text{CBC Model} \)

New use-cases in development for O4

\( N_{\text{cal}} \times \{\delta A_I, \delta \phi_{IFO}\} : \text{Calibration Model} \)
\[ N_{\text{spline}} \times \{f_0, S_n\} : \text{Spline Model} \]

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

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

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

and/or

\[ \{m_1, m_2, S_1, S_2, L, \alpha, \delta, D_L, t_0\} : \text{CBC Model} \]

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

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

Similar algorithms under development for LISA