Impact of calibration uncertainties on cosmological measurements from gravitational wave sources

Yiwen (Eva) Huang, Carl-Johan Haster, Hsin-Yu Chen, Salvatore Vitale, Jeffrey Kissel, Lilli Sun, Ethan Payne
Goal: understand the impact of calibration errors and uncertainties on astrophysical parameter estimation (PE) and cosmology

Calibration: produce response functions that convert the photodetector output in the interferometers to the strain data from which we can extract gravitational wave signals

Method: add artificial calibration errors, motivated by detector behavior of Hanford and Livingston in the third observing run (O3) of aLIGO-VIRGO

**PE results** for individual events

Plans to combine events for Hubble constant \( H_0 \) measurements
Background

- We need **luminosity distances** $D_L$ for Hubble constant $H_0$ measurement.
- $v_H = H_0 D_L$ at small redshifts, $v_H$ is local Hubble flow velocity.
- Biases in luminosity distance can lead to bias in Hubble constant, more significant when we combine multiple events.
- We assume there are electromagnetic (EM) counterparts for binary neutron stars (BNSs) and neutron star-black holes (NSBHs) in our study.
Calibration uncertainties in PE

- **Spline interpolation**
  - Fits the response function using a cubic spline polynomial
  - Determines errors at each node of the polynomial in frequency
- **physiCal**[2009.10192]
  - Uses a distribution of response function curves
  - Each curve is a possible posterior sample
Simulation Set-up

- From calibration team: model response functions $R_{\text{model}}$ used in parameter estimation and the response functions $R_{\text{miscal}}$ used for miscalibration

![Graphs showing magnitude and phase versus frequency for $R_{\text{model}}$ and $R_{\text{miscal}}$]
Simulation Set-up

• From calibration team: model response functions $R_{\text{model}}$ used in parameter estimation and the response functions $R_{\text{miscal}}$ used for miscalibration

• Run parameter estimation with $R_{\text{model}}$

• Experiment runs: add artificial calibration errors using one curve from $R_{\text{miscal}}$ to mimic “bad scenarios” where we do not manage to capture all the features when modeling the response function

• Control runs: no calibration error, to disentangle other causes for bias
A curve for the misbehaving detector, picked by minimizing:

\[
\int_{f_{\text{min}}}^{f_{\text{max}}} \left( A_{i,\text{miscal}}(f) - A_{\text{model}}(f) \right) S(f) \, df
\]

Weighted by PSD(sensitivity)

- A curve within the 68% of the model response function distribution used for PE (physiCal), for the other detector
A curve for the misbehaving detector, picked by minimizing:

\[
\int_{f_{\text{min}}}^{f_{\text{max}}} \frac{A_{i,\text{miscal}}(f) - A_{\text{model}}(f)}{S(f)} \, df
\]

Weighted by PSD(sensitivity)

A curve within the 68% of the model response function distribution used for PE (physiCal), for the other detector.
Simulation Set-up

• From calibration team: model response functions $R_{\text{model}}$ used in parameter estimation and the response functions $R_{\text{miscal}}$ used for miscalibration

• Run parameter estimation with $R_{\text{model}}$

  • Experiment runs: add artificial calibration errors using one curve from $R_{\text{miscal}}$ to mimic “bad scenarios” where we do not manage to capture all the features when modeling the response function

  • Control runs: no calibration error, to disentangle other causes for bias

• 4 typical compact binary coalescence signals $h(t, \theta)$

  • Assume we know the sky localization (ra, dec) of potentially EM bright coalescences that include a neutron star

• Add miscalibration $S_{\text{miscal}} = (\text{noise}(t) + h(t, \theta)) \times R_{\text{miscal}, i}$

• Worst-case scenario: the same calibration error is not accounted for but present for all events
- BNS
  - \(m_1 = 2M_\odot\), non spinning
  - \(m_2 = 1.5M_\odot\), non-spinning
- Sky localization known
- SNR 50, physiCal*, Large calibration error ("mis") vs No calibration error ("control")
  - Lines are quartiles (25%, 50% and 75%)
  - *Spline results are very similar to physiCal, thus not shown here
LIGO PE Results - BNS

- BNS
  - $m_1 = 2M_\odot$, non spinning
  - $m_2 = 1.5M_\odot$, non-spinning
  - Sky localization known
- SNR 50, PhysiCal with uniform $\eta_{\text{NIST}}$ prior, Large calibration error ("mis") vs No calibration error ("control")
  - Relative biases on distance ($\Delta D_{L,\text{med}} / D_{L,\text{true}}$) for mis (control)
    - 4.7% (0.9%) for SNR50
    - 4.1% (0.2%) for SNR35

Normalized by the true value
• NSBH
  • $m_1 = 5M_\odot$, $a_1 = 0.8$, $t_1 = 40^\circ$
  • $m_2 = 1.4M_\odot$, non-spinning
  • Sky localization known
• SNR 50, PhysiCal* with unif $\eta_{\text{NIST}}$ prior, Large calibration error ("mis") vs No calibration error ("control")
  • Lines are quartiles (25%, 50% and 75%)
  • *Spline results are very similar to physiCal, thus not shown here
• NSBH
  • $m_1 = 5M_\odot$, $a_1 = 0.8$, $t_1 = 40^\circ$
  • $m_2 = 1.4M_\odot$, non-spinning
  • Sky localization known

• SNR 50, PhysiCal* with unif $\eta_{\text{NIST}}$ prior, 
  **Large calibration error ("mis") vs No calibration error ("control")**
  • Relative biases on distance 
    ($\Delta D_{L,\text{med}}/D_{L,\text{true}}$) for mis (control)
  • $4.1\%(0.8\%)$ for SNR50
  • $4.6\%(1.0\%)$ for SNR35
Summary

- Single event level
- Systematic bias \(~4-5\%\) in luminosity distances, smaller than statistical uncertainties for all individual events here
- If the same effect is present in multiple events, the bias on combined PE will become more significant
Next Steps

- 100 BNS coalescences
  - Random luminosity distances uniform-in-volume \(~ D_L^2\)
  - Random inclinations and sky localizations (ra, dec) uniformly distributed
  - Assume we know the sky localization (ra, dec)
  - Use time #1 for “worst” calibration error realization
  - Most significant biases in distance, etc
  - Worst-case scenario: same calibration error is not accounted for but present for all events
Relative uncertainties on luminosity distances (sorted by SNR) show a big spread among events.

Systematic biases become more comparable to statistical uncertainties at higher SNRs.

At SNRs of potential O4 events.