Detector Characterization on KISTI/GSDC for KAGRA

Kim, Young-Min (UNIST)
on behalf of KGWG-DetChar
Laser Interferometer

GW channel

Detected Characterization

non-gaussian noise artifacts, “glitch”

Signal Search

Example Inspiral Gravitational Wave
Role of Detector Characterization

Interferometer is so complex!

Detector Characterization

- What GPS time segment to analyze?
- When was the ifo operating?
- How is the data quality?: Is this data segment noisy or not? Glitchy or not? Line noise?
- If noisy or glitchy, what’s the source? How is it coupled to h(t)?
- What channel to see to know the status?

Hardware

- 2/14/19

Data Analysis

- JGW-G1909828

taken from Kokeyama’s slides
Main task of the PEM subsystem

After finishing the installation tasks, we will start the commissioning phase.

One of the important tasks is:

**Possible upgrade of VT**

- **V**: Volume, try to achieve the good sensitivity
  - Search the origin of noise which makes the noise floor dirty and makes the glitch
  - Line noise characterization, time variance of noise floor, glitch noisy period search, veto, ...

- **T**: Time, try to achieve the stable operation
  - Reducing the origin of unlock, quick recovery
  - Safety interferometer control system
The impact of Data Quality Vetoes

The false alarm rate of GW151226 improves by a factor of 567, from 1 in 320 years to 1 in 183000 years, with interferometer data quality information!

If DQ vetoes are not applied, GW151226 is no longer louder than the entire background distribution.

LIGO-Virgo Collaboration, Class. Quantum Grav. 35 (2018) 065010
What we need for the joint Observation

1. On-line Data Quality
   - Online DQ flags
   - online DQ pipeline (e.g. iDQ) - glitch identification and responsible channels

2. Off-line Data Quality
   - deep investigation on glitches, spectral lines, noise sources, etc.
   - improve search background with full DQ

3. Channel Information
   - find out ‘unsafe’ channels
   - ‘safe’ channels to be used for online / offline analysis and vetoes
KGWG DetChar Activities

1. Off-site work
   - glitch investigation with LIGO-Virgo DQ tools
   - Machine learning application to glitch identification
   - (off-line) iDQ operation

2. On-site work
   - short-/mid-term visits to the KAGRA site
   - installation of DQ tools, or development of new tools
   - DQ shifts for the operation

3. Channel safety study
   - find ‘unsafe’ / ‘safe’ channels
Works in KISTI/GSDC

1. Off-site work
   - LIGO/Virgo tools (gwpy, omicron, hveto, bruco), CAGMon, etaGen, etc.

2. Channel Safety Study
   - using Hardware Injections: Photon Calibrator (PCal) Injection
   - Figure out which channels respond to the injections

3. Correlation study
   - CAGMon, …

4. New Event Trigger Generation method: etaGen

5. Machine Learning Application to Glitch Identification
DQ tools in KISTI/GSDC

1. Transient Signal Identification
   - Omicron
   - etaGen

2. veto algorithms
   - hveto
   - iDQ: No low-latent operation for KAGRA due to data transfer issue

3. Correlation and Coherence tools
   - CAGMon
   - BruCo

4. GWPy: a python package for GWDA
Omicron Trigger

   - Burst-type search based on Q-transform (CQG 21, S1809 (2004))

2. **Installation on KISTI/GSDC cluster**

---

![Graph showing signal-to-noise ratio (SNR) over time and central frequency.](image)
etaGen

- EtaGen is an event trigger generator based on Hilbert-Huang Transform (HHT)
- HHT = Empirical Mode Decomposition (EMD) + Hilbert Spectral Analysis (HSA) (for a review, see Huang et al., Rev. Geophys. 2008)
- To reduce computing cost, we are using weighted sliding EMD (wSEMD) instead of Ensemble EMD (EEMD)

Edwin J. Son et al., arXiv.1810.07555
etaGen: how it works

- EMD
  - Generate IMFs
  - Sifting

- HSA
  - Hilbert Transform
  - Complexify to get a(t) & f(t)

- Trigger Generation
  - Unclustered
  - Clustering
etaGen : Performance test
**hVeto**

\[ S = -\log_{10} \sum_{k=n}^{\infty} \left[ \frac{\mu^k e^{-\mu}}{k!} \right] \]

\[ \mu = \frac{N_{\text{main~tot}} N_{\text{aux~tot}} T_{\text{win}}}{T_{\text{tot}}} \]

- \( n \): the number of coincidences
- \( T_{\text{win}} \): full width of coincidence time window
- \( T_{\text{tot}} \): a given total analysis time

ClassQuantGrav.28.235005(2011)
Counting Experiment

\[ h(t): \quad N_h = 5 \]

Auxiliary Channel: \( N_n = 6 \)  
Coincidences: \( x = 3 \)

Poisson Statistics
Example: ITMX pitch
Example: ITMX pitch
Channel Safety Study with HW injections

1. The safety of a veto is important for veto criteria not to remove accidentally a true gravitational wave signal.

- **unsafe** channels: Auxiliary channels with non-negligible couplings from GW channel. A corresponding response to HW injections is greater than expected by chance.

- **safe** channels: it can be used as a veto or to study glitches in h(t).
How to find unsafe channels?

1. Heuristic Methods
   - Understanding detector itself.

2. Statistical Methods
   - hVeto, UPV, OVL, iDQ, etc…

3. Correlation methods
   - Peason’s correlation, MIC, CAGMon, etc
   - Coherence (BruCo)
Safety check with HW inj.

1. A response of a auxiliary channel to HW injections can be analyzed by using Omicron triggers.
   - Trigger generation rate is larger than KW triggers.
   - Efficiency / dead-time of a safe channel can be estimated.
   - We don’t exactly know auxiliary channels’ responses to $h(t)$ so that we will have to check coincidences of triggers channel by channel.

2. Omicron Scans
   - a time-frequency map
   - It will be used to compare the morphologies between GW channel and an auxiliary channel which is suspicious as unsafe.
Preliminary study of CSS

1. Conducted by Pil-Jong Jung (KGWG —> ICRR)


iDQ - online DQ pipeline

1. Developed by Reed Essick and his colleagues
   - https://docs.ligo.org/reed.essick/iDQ/index.html
2. A (near) real-time statistical data quality pipeline for glitch detection
3. Output: a probability that there is a glitch in h(t) as a function of time
4. 2-class classification
   - 1 for glitch, 0 for clean
   - train : make mapping
   - realtime : evaluate a rank (0~1)
   - calibration : conditioned probability distribution
5. responsible channels to the identified glitch
issue on running iDQ

1. Computing resource
   - Number of Channels: ~1000
   - train: ~1.78 GB/day (37 MB per 1800s) => 454 GB/yr (duty cycle 70%)
   - realtime: ~2.33 GB/day (2.7 GB per 10^5s) => 596 GB/yr

2. Latency of Data transfer
   - off-line iDQ is currently available.
# BruCo (Brute-force Coherence)

| Time | ASC-CHARD P OUT DO | ASC-CHARD P IN1 DO | ASC-CSOFT P OUT DO | ASC-CSOFT P SM DO | ASC-DC1 P OUT DO | ASC-DC1 P SM DO | SUS-RM2 M1 OSEMFIN UR | SUS-ITMY L3 DSCINF P | ASC-OMC B NSUM OUT DO | ASC-OMC A NSUM OUT DO | ASC-SAS C NSUM OUT DO | ASC-POP A RF45 O ERB 239 DO | ASC-MICH CTRL 239 DO | SUS-BS M2 DSCINF L IN1 DO | SUS-PR2 M2 DSCINF L IN1 DO | CAL-CS M2 CTRL DO |
|------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 10.00 | (0.57)             | (0.57)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)             | (0.37)                 | (0.37)                  | (0.37)                  | (0.37)                  | (0.37)                  | (0.37)                  |
| 10.50 | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)             | (0.58)                 | (0.58)                  | (0.58)                  | (0.58)                  | (0.58)                  | (0.58)                  |
| 11.00 | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)                 | (0.59)                  | (0.59)                  | (0.59)                  | (0.59)                  | (0.59)                  |
| 11.50 | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)                 | (0.57)                  | (0.57)                  | (0.57)                  | (0.57)                  | (0.57)                  |
| 12.00 | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)             | (0.64)                 | (0.64)                  | (0.64)                  | (0.64)                  | (0.64)                  | (0.64)                  |
| 12.50 | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)             | (0.57)                 | (0.57)                  | (0.57)                  | (0.57)                  | (0.57)                  | (0.57)                  |
| 13.00 | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)             | (0.59)                 | (0.59)                  | (0.59)                  | (0.59)                  | (0.59)                  | (0.59)                  |
| 13.50 | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)             | (0.27)                 | (0.27)                  | (0.27)                  | (0.27)                  | (0.27)                  | (0.27)                  |
| 14.00 | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)             | (0.43)                 | (0.43)                  | (0.43)                  | (0.43)                  | (0.43)                  | (0.43)                  |
| 14.50 | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)             | (0.15)                 | (0.15)                  | (0.15)                  | (0.15)                  | (0.15)                  | (0.15)                  |

**Graphs**

- **Coherence**
  - X-axis: Frequency [Hz]
  - Y-axis: Coherence
  - Two traces: Target channel and Noise projection.

- **Spectrum**
  - X-axis: Frequency [Hz]
  - Y-axis: Spectrum
# BruCo Comparisons

1. run "filter4_bruco.py"

## Top 40 coherences at 0.0 <= f <= 4096.0 Hz : coh.=(0.00996898170809,1.0)

<table>
<thead>
<tr>
<th>Frequency [Hz]</th>
<th>Top channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>SUS ETMY M0 DAMP Y IN</td>
</tr>
<tr>
<td>0.25</td>
<td>LSC POPAIR B. RFI3 Q ERR DO (0.15) ref.</td>
</tr>
<tr>
<td>0.50</td>
<td>IMC IM4 TRANS SUM IN</td>
</tr>
<tr>
<td>0.75</td>
<td>ASC Y TR B P</td>
</tr>
<tr>
<td>1.00</td>
<td>ASC CHARD P UV</td>
</tr>
<tr>
<td>1.25</td>
<td>SUS BS M2 OSEMINF UL</td>
</tr>
<tr>
<td>1.50</td>
<td>SUS ETMY L2 WIT Y</td>
</tr>
<tr>
<td>1.75</td>
<td>PEM CS RADIO LVEA NARROWBAND</td>
</tr>
<tr>
<td>2.00</td>
<td>SUS ETMY L2 FASTIMON</td>
</tr>
</tbody>
</table>
Correlations between Auxiliary Channels

- To find a systematic way of the correlation between various auxiliary channels in GW detector
  - Helps us to fix the correlated channels that produces abnormal glitches
  - Important to monitor the dynamical variation of the detector for removing glitches
- Pearson’s correlation coefficient (PCC):

\[ r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \]

- a measure of linear correlation between two random variables
Correlations between Auxiliary Channels

To find a systematic way of computing the correlation between various auxiliary channels in GW detector

- Helps us to fix the correlated channels that produces abnormal glitches
- Important to monitor the dynamical variation of the detector for removing glitches

Mutual Information Coefficient (MIC):

\[ I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]

- a measure of non-linear correlation between two random variables
CAGMon

1. Developed by John J. Oh (NIMS)
   - https://git.ligo.org/john.oh/CAGMon
MLAs for Glitch Identification

Random Forest of Bagged Decision Trees

Support Vector Machine

Artificial Neural Network

LIGO-G1200500

Ch.1  Ch.1  ...  Ch.n-1  Ch.n

t1  x1  x2  ...  xn-1  xn

tm  x1  x2  ...  xn-1  xn

GW channel
MLAs for Glitch Identification

Random Forest of Bagged Decision Trees
Support Vector Machine
Artificial Neural Network

safe Auxiliary channels ~ 1,000
MLA application to DetChar

1. Ordered Veto List (OVL) + 3 Machine Learning Algorithms

   - application to hundreds of channels among 200,000 auxiliary channels
   

\[ => iDQ + MLAs \]
Summary and Future plans

1. LV DQ tools are adopted for KAGRA DQ investigation
   - etaGen, CAGMon were devolved by KGWG

2. Channel Safety Study was partly done and continues

3. For the joint observation with LIGO and Virgo
   - visit to the site
   - online investigation on KAGRA DQ
     • study on recent lock segments of KAGRA commissioning run
   - running off-line iDQ and optimization for KAGRA
   - spectral line tools
   - deep investigation on KAGRA DQ