Data analysis of glitches from CQM testing
SPIRE-UOL-REP-002208

Prepared by:

Peter Davis (University of Lethbridge)
peter.davis@uleth.ca

Todd Atkinson (University of Alberta)
todda@ualberta.ca

Trevor Fulton (University of Lethbridge)
trevor.fulton@uleth.ca

Andres Rebolledo (University of Lethbridge)
andres.rebolledo@uleth.ca

Approved by:

David Naylor (University of Lethbridge)
naylor@uleth.ca

Document History:

Issue	Date	Comments
Version 1.0	November 18, 2004	Original document for CQM pre-vibration test campaign
Version 2.0	February 9, 2005	Results from CQM post-vibration test campaign are added New approach to data analysis for data from pre- and post-vibration test campaign.

Table of Contents:

Table of Figures:

1. Scope

Results are reported on flagging and characterizing glitches in the data from the CQM test campaigns in 2004. While a lot of funny things are apparent in the test data, this report focuses on short-lived transients and argues that they are due to instantaneous depositions of energy on bolometers. The goal is to validate the modeled impulse response function due to the thermal properties of the bolometers and the read-out electronics.

After the inconclusive results from the first analysis (version 1.0), a new approach to the data analysis was taken to directly compare the modeled Impulse Response Function to the empirical data. It is detailed below.

2. Reference Documents

#	**Title**	**ID**
RD1	Herschel SPIRE Detector Subsystem Specification Document	SPIRE-JPL-PRJ-000456, issue 3.2
RD2	Herschel SPIRE Detector Control Unit Design Document	SAp-SPIRE- FP-0063-02 , issue 0.3

3. Applicable Documents

#	**Title**	**ID**
AD1	Glitch Simulation	SPIRE-UOL-REP-002207, v0.1

4. Modeling and flagging glitches

PLW's response to a Dirac input was modeled as the transfer function of the thermal response of the bolometers and the electrical filter (see AD1). The specifications for the bolometers and the read-out electronics were used for this model (RD1, RD2). The resulting impulse response function (IRF) was used to synthesize test data and develop an algorithm to identify such glitches reliably (see Figure 1).

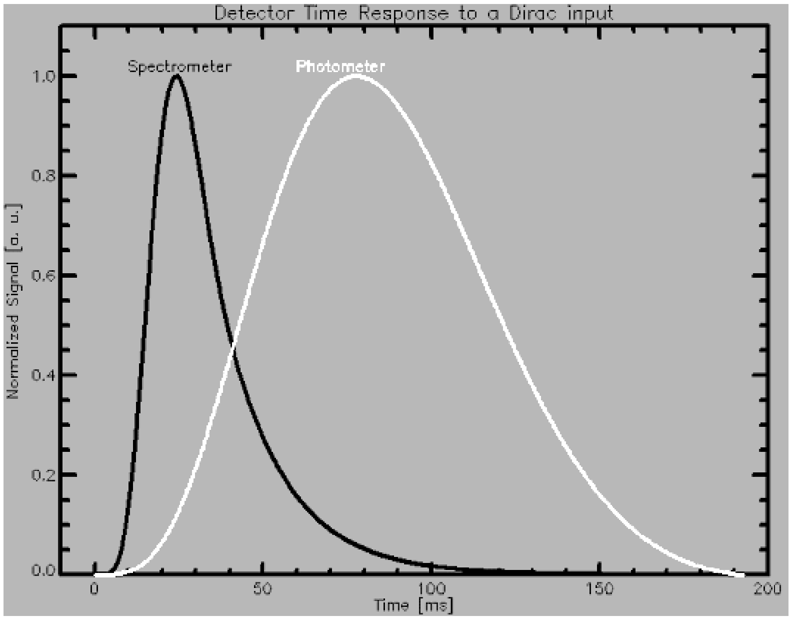

Figure 1: Modeled IRF for the spectrometer and the photometer (PLW).

Empirically, glitches were flagged as outliers in slope. Because a large number of sample points were studied, Gaussian statistics leads to considerable numbers of outliers for a threshold of up to four standard deviations. For low thresholds (2 and 3) one can expect that the majority of flagged outliers is due to white noise. From a threshold of 6 standard deviations on one can expect not to see any chance outliers from white noise at all. A five standard deviations threshold appeared to be a good trade-off between maximizing the sample size of glitches and avoiding false positives (see Figure 2). At the five standard deviations threshold, a negligible number of outliers (seven/nine) are expected for the more than 24/31 million data points from the pre/post-vibration test campaigns.

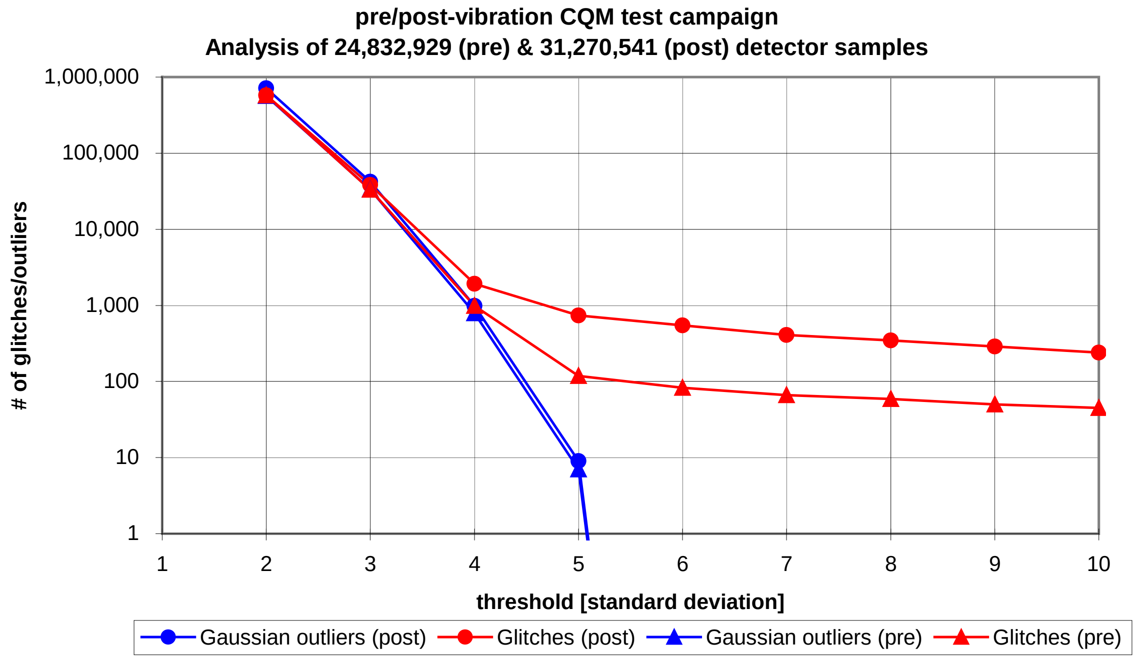

Figure 2: Outliers from a normal distribution (blue) and the actually flagged outliers in the CQM data (red) are shown for data from the CQM pre- and post-vibration test campaign.

The following procedure was used to further study the identified glitches:

1. The baseline for all identified glitches was established as a robust fit of a straight line (IDL's LADFIT) to 31 points with the glitch peak at the center. The number of points used to determine the baseline seems not to have great impact on the results (50 points or 5 points on either side didn't make much difference).
2. The glitch amplitude is determined from fitting a parabola to the glitch peak and its two neighbors. The glitch amplitude is set to the difference between the baseline under the glitch peak and the minimum of this parabola. When IDL detected a small pivot element in the POLYFIT procedure, the glitch was re-considered.
3. For all accepted glitches, the points during the 100ms before and the 150ms after the glitch peak are registered relative to the minimum of the fitted parabola. The baseline is subtracted and the glitch amplitude is normalized to 1.
4. The resulting data points are binned in 0.01 s wide bins. The average values and their standard deviations, as well as the median and the median difference from the median are computed per bin. The result is the empirical IRF for CQM.

5. Pre-vibration test campaign, February 2004

Data during the performance tests with the Test Facility FTS during the pre-vibration test campaign were taken at a sampling rate of ~43Hz or at time intervals of 23.11 ms. Overall, more than 149 hours of single-pixel data were analyzed. The 5 sigma criterion identified 84 glitches. On average, 0.56 glitches were flagged per hour single pixel data.

IDL's POY_FIT flagged all glitches as dubious fits (small pivot element), but on visual inspection, there was no evidence to suspect problems in the curve fitting. There were some funny 'glitches' where features of a larger scale ('waves') were present in the data. This would occasionally lead to very small amplitudes and therefore huge values after normalization. These excessive normalized signals were discarded.

The histogram of the amplitude distribution for the glitches (see Figure 3) suffers from the relatively small number of glitches. The distribution shows an initial rise in number of glitches due to the 5σ cut-off threshold and a subsequent steep decline in frequency with amplitude.

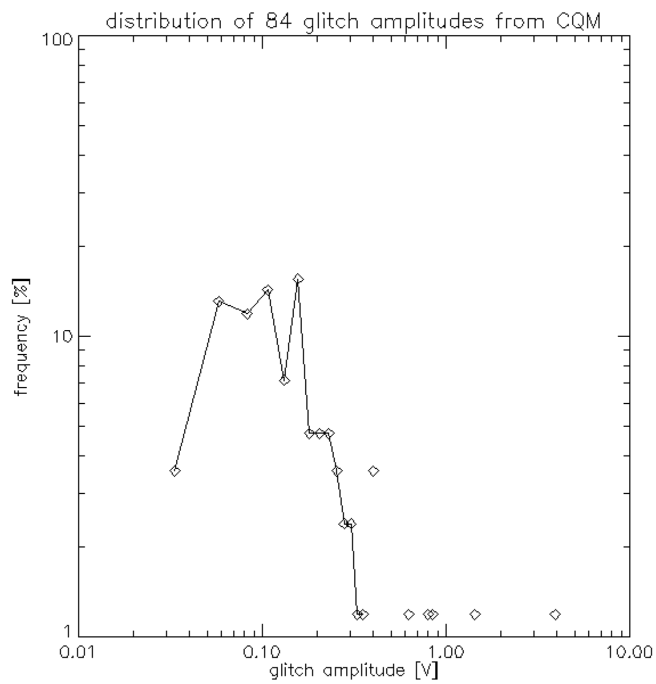

Figure 3: Amplitude distribution of the 84 glitches from the CQM pre-vibration test campaign

The resulting average, normalized response at steep slopes for the pre-vibration test campaign is shown in Figure 4. Note that the glitches are centered at the maximum of the fitted parabola.

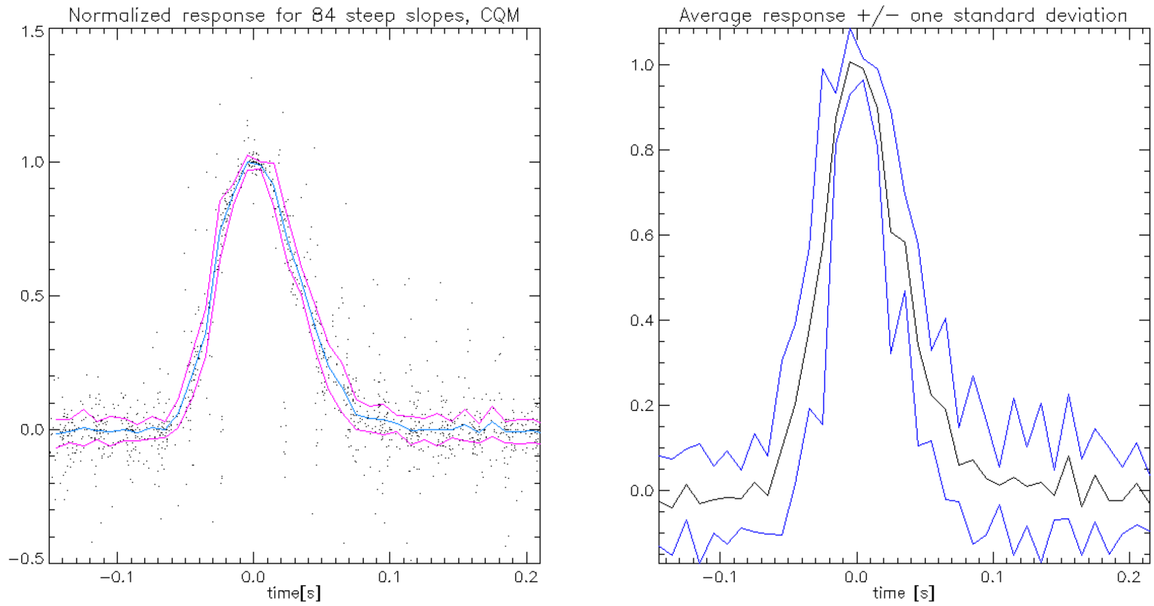

Figure 4: Normalized timeline of measured points at steep slopes. Left: median (blue) ± the median of the differences (magenta). Right: average (black) ± one standard deviation (blue).

Noise has a considerable impact on the averaged response because of the small sample size of only 84 glitches.

6. Post-vibration test campaign, September/October 2004

Data during the performance tests with the Test Facility FTS during the post-vibration test campaign were taken at a sampling rate of ~17.5Hz or at time intervals of 57.24 ms. Overall, more than 497 hours of single-pixel data were analyzed. The 5 sigma glitch identification routine led to 5122 flagged glitches.

85% of all flagged glitches were flagged as a dubious fit (small pivot element) by POLY_FIT. On visual inspection, glitches with dubious ‘glitches’ turned out to result from intermittent system behavior and were discarded. The distribution of the resulting 735 glitches shows a sensible characteristic as the frequency exponentially decreases with amplitude (see Figure 5). At low amplitudes there is an increase in frequency with rising amplitude because of the initial cut-off criterion of 5 sigma to flag outliers. On average, 1.48 glitches were flagged per hour single pixel data – a glitch rate more than twice as high as for the pre-vibration test campaign.

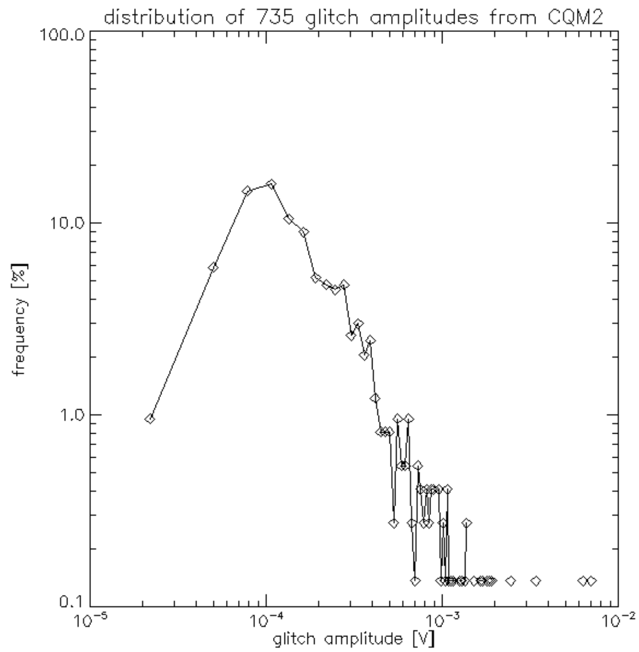

Figure 5: Amplitude distribution of the 735 glitches from the CQM post-vibration test campaign

The resulting average, normalized response at steep slopes for the post-vibration test campaign is shown in Figure 6. The points are not distributed equally on the left pane of Figure 6 because the strong 5 sigma criterion will miss glitches that are not sampled close to the glitch peak.

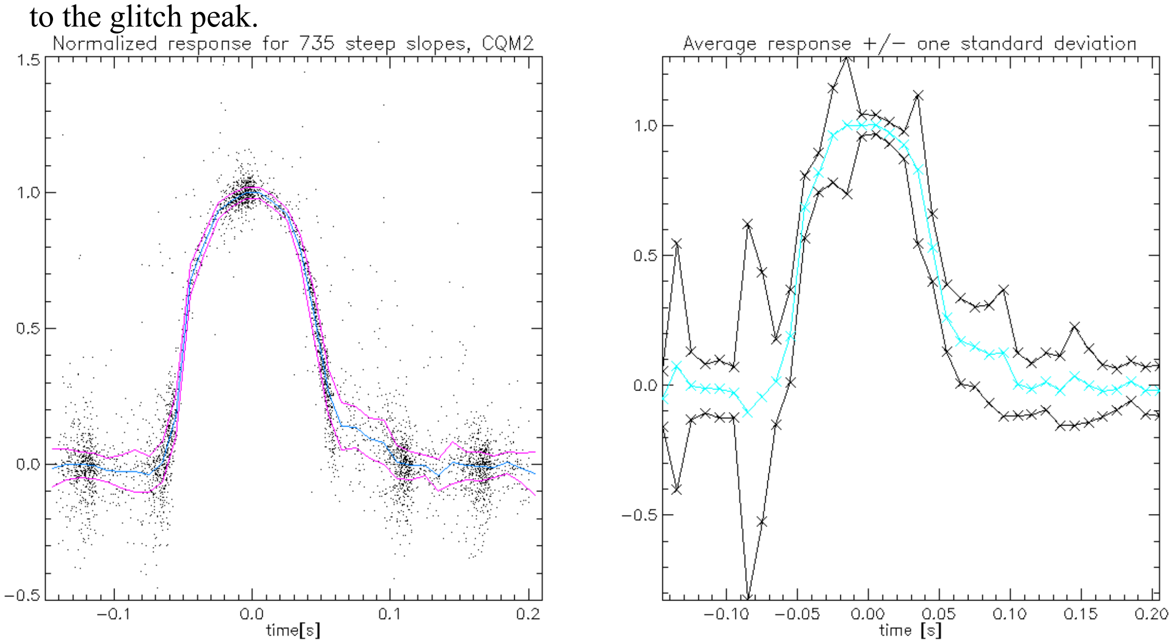

Figure 6: Normalized timeline of measured points at steep slopes. Left: median (blue) ± the median of the differences (magenta). Right: average (cyan) ± one standard deviation (black).

The noise level is rather high and no distinct feature of the empirical response function can be seen, other than a general, parabolic shape. The sparse sampling makes it impossible to trace the shape of the impulse response function in more detail.

7. Comparison between model and measurements

Unfortunately, during the post-vibration test campaign, which resulted in a much larger number of identified glitches, data were sampled at a low frequency. The sampling interval of 57ms is too large to resolve detailed features of the IRF which is of the order of 130ms. During the pre-vibration test campaign, where the sampling interval was smaller by almost a factor of 3, much fewer glitches were identified. These data, however, are sufficient to allow a comparison to the expected impulse response function.

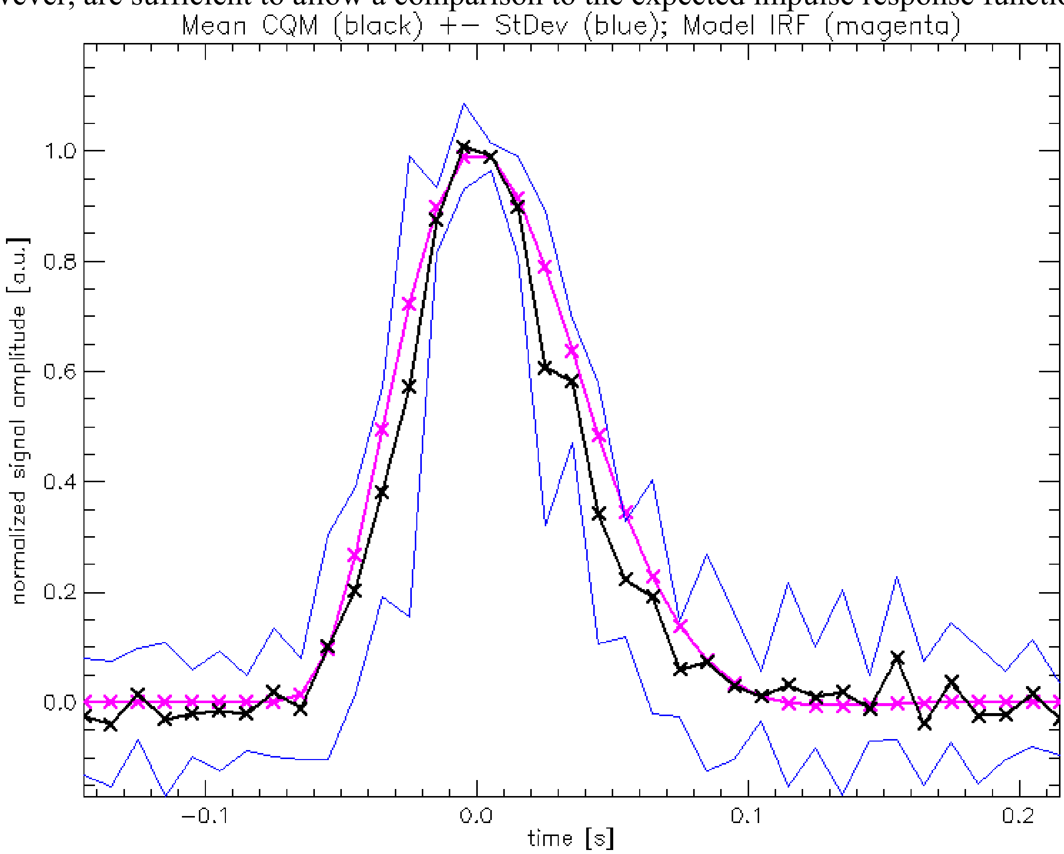

Figure 7: The average measured IRF from the CQM pre-vibration test campaign (black) ± one standard deviation (blue) and the model IRF (magenta).

8. Conclusions

Some conclusions can be drawn from the above analysis:

- The empirical evidence from CQM testing does not invalidate the model IRF which is based on design specifications. The deviation of the actual system behavior from design specifications is not significant.
- The glitch rate increased by a factor of ~ 2.6 after the vibration tests. The measurement error on this factor is probably of the order of + 0.1 and – 0.3.

- Bolometer saturation, i.e. consecutive readings of the same, extreme signal after a steep slope, was not seen in the data.
- Commonly, steep slopes occurred in various pixels around the same time.
- The assumption that 'glitches' in the data can be modeled as the dissipation of an instantaneous deposition of energy has been validated.
- In order to repeat this analysis on other instrument models, it is necessary to collect data at a high enough sampling rate. 'High enough' means that the sampling interval must be smaller than the dissipation time constant by a factor of at least three. For the spectrometer this suggests a sampling interval of less than 10ms and less than 20ms for the photometer.