Chapter 8

Ranking of Detection Candidates

In this chapter we describe the postprocessing pipeline that has been developed as an extension to the standard pipeline (discussed in chapter 6) to analyze data from GW detectors, such as the LIGO detectors, for GW signals from CBCs. In addition, we will also describe the tuning that was done for the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data, described in this thesis, and referred to as "this search."

At the end of the standard pipeline, we have distilled the data down from one data stream sampled at 16,384 Hz for each GW detector into files containing coincident triggers that have passed our thresholds and vetoes. These files are organized by coincident time segments defined by times when different combinations of detectors' data were being analyzed.

Recalling section 6.8, during these coincident time segments triggers are discarded if they occur during times flagged by the category 1, 2, or 3 DQ flag veto files. As mentioned in section 7.1, these vetoes affect the application of signal-based vetoes, which, in turn, significantly affect the background distribution of triggers. Because of this, we separate the coincident times (and triggers within those times) according to the detectors that were operating without DQ vetoes active. We then recalculate the amount of analyzed time in each type of coincidence time for the in-time data as well as for each time shift, accounting for the dead time incurred from the application of DQ flag vetoes. Next we cluster coincident triggers in order to reduce the final number of triggers to a manageable level, and separate them by categories (mass regions and coincident trigger types). Finally, we calculate the FAR for each of the in-time, software injection, and time-shift coincident

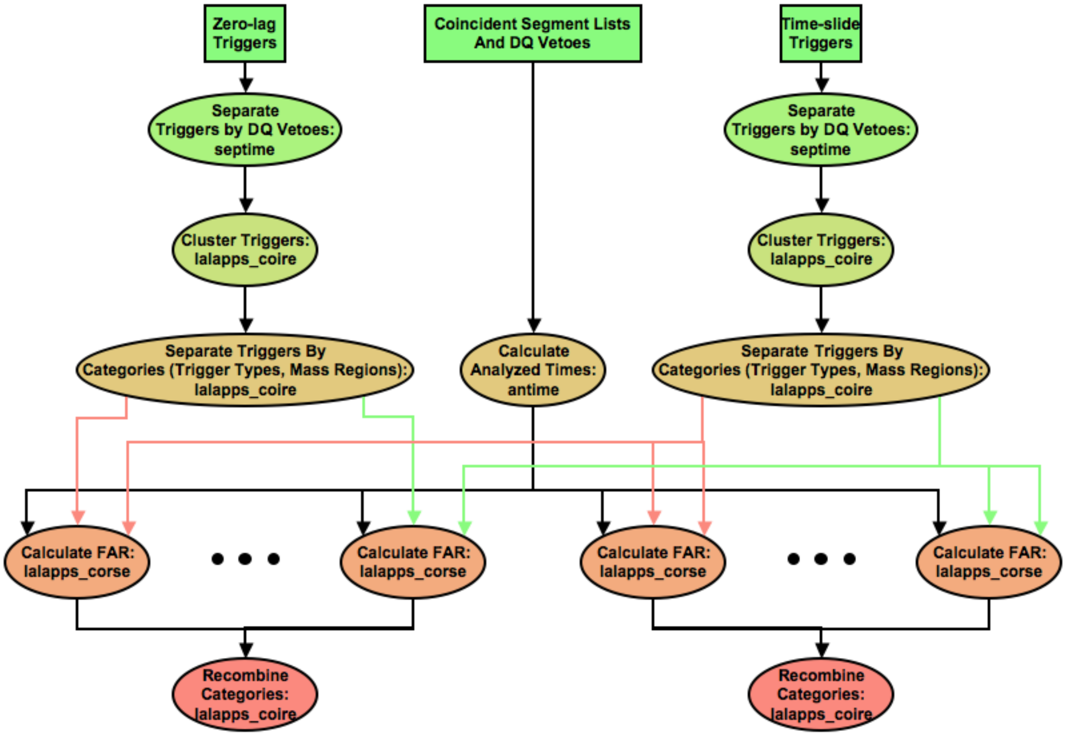

Figure 8.1: Postprocessing Pipeline
The stages of the postprocessing producing the final candidate triggers.

triggers, utilizing the in-time and time shifts' amounts of analyzed time. The postprocessing pipeline is represented in figure 8.1.

The different stages of the postprocessing pipeline used in this search are discussed below:

- separate triggers by veto times (detailed in section 8.1);
- calculate nonvetoed analyzed times (detailed in section 8.2);
- cluster triggers and separate them by categories (detailed in section 8.3);
- calculate the FAR for the triggers (detailed in section 8.4).

8.1 Separating Triggers by Data Quality Flag Vetoes

The application of DQ flag vetoes causes a loss of detector live time that was not accounted for in previous searches (it was instead counted as a loss of efficiency). Accounting for it as a loss of live time requires us to reclassify the coincident category of coincident triggers in the in-time and time-shifted data.

The first thing that is done in the postprocessing pipeline is to separate triggers according to which detectors were producing triggers that did not have DQ flag vetoes active. As mentioned above, because the DQ flag vetoes affect which signal-based vetoes we apply, we want to group double coincident triggers occurring during triple coincident time where the third detector had an active DQ flag veto with double coincident triggers coming from true double coincident time.

We work with single coincident science segments in which we separate triggers in a manner appropriate for both time-shift and in-time triggers. The output is a number of SEPTIME files for each type of time when two detectors were not being vetoed (for H1H2L1 times, the output files can be H1H2 times, H1L1 times, H2L1 times and H1H2L1 times; for H1L1 times, the output is only H1L1 time). In the case of in-time data, it is also dependent on whether or not there was a nonzero amount of time of that type.

Figure 8.2 shows a visual example of extracting the correct time-shift triggers from a single triple coincident segment. The vector used for the time shifts here is $v = \{0, 2, 1\}$ for H1, H2, and L1 respectively. The top black bar at $y = 10$ shows the segment we are working with is $[0 - 10)$, the next bars show the times when each of H1 (red), H2 (blue), and L1 (green) are not vetoed (H1 not vetoed $[1 - 10)$, H2 not vetoed $[1 - 10)$, L1 not vetoed $[1 - 6)$). Just below those lines show the in-time triggers that are shifted when computing coincidence in time-shifted data.

The middle lines and triggers from $y = 5$ to $y = -5$ show the time shifts numbered 5,4, ...,1,-1, ...-5. The colored bars again denote the times the detectors were not vetoed, shifted by the appropriate amount for that time shift and detector. The triggers under the time-shift bars are the in-time triggers that were coincident in that time shift, shifted to their appropriate location. If you pick a coincident trigger in a specific time shift, you can look just above it to tell which detectors were producing nonvetoed triggers.

The final four lines at the bottom denote the four different times triggers could be generated (i.e., H1H2L1, H1H2, H1L1, and H2L1 times from top to bottom). The unshifted trigger times, which can be compared to the input trigger times at the top of the plot, are shown on each of those four lines. The septime program outputs those four lines are four separate trigger files.

At the end of this stage, for each coincident segment, we are left with files whose contents are coincident in-time or time-shifted triggers that occurred during coincident times defined as times when those detectors were analyzed and not vetoed by DQ flag vetoes.

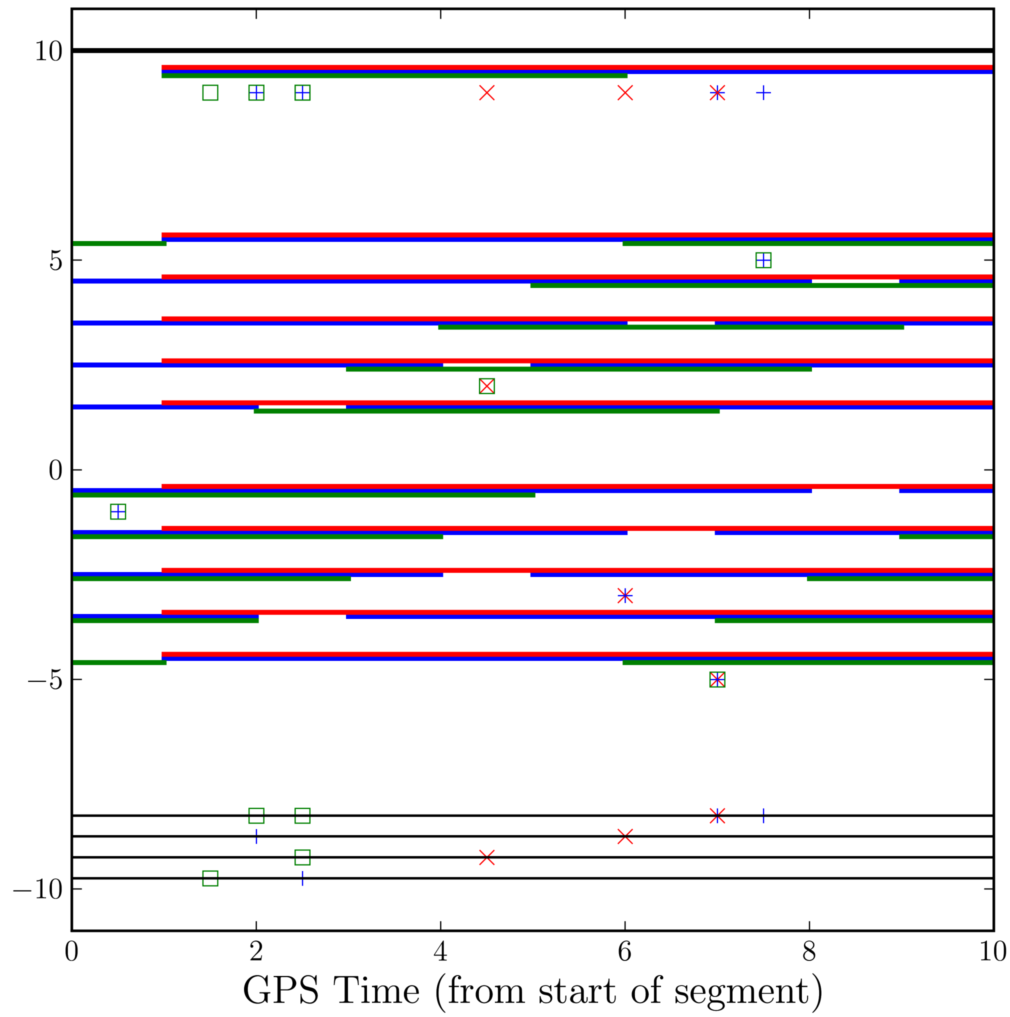

Figure 8.2: SEPTIME Example
An example of septime extracting time-shift triggers from the appropriate coincidence times once DQ flag vetoes are taken into account.

8.2 Calculate Analyzed Times

Now that we have sorted the triggers according to times when detectors were analyzed and not vetoed by DQ vetoes, what we will now refer to as "analyzed times," we also want to calculate the amount of analyzed times per in-time and time-shifted data. We do this using a program called antime that uses the science segment files and appropriate veto files. When detectors are vetoed, that time is removed from that type of time and added to the type of time defined by the detectors not being vetoed (e.g., H1H2L1 time where H2 is vetoed is added to H1L1 time).

Summarizing the in-time analyzed times, before the application of category 2 ans 3 DQ flag vetoes, the in-time analyzed times were 13,237,711 s for H1H2L1 times, 1,179,439 s for H1L1 times, 1,487,038 s for H2L1 times, and 7,337,970 s for H1H2 times. Once the category 2 and 3 vetoes are taken into account we have the following times from triple coincident times: 11,696,445 s for H1H2L1 times (88.4% of the original), 76,446 s for H1L1 times (0.6%), 470,764 s for H2L1 times (3.6%), and 790,885 s for H1H2 times (6.0%). In addition, from the double coincident times we have: 615,994 s for H1L1 times (52.3%), 949,899 s for H2L1 times (63.9%), and 6,951,276 s for H1H2 times (94.7%).

In total, we have 11,696,445 s of H1H2L1 time, 692,440 s of H1L1 time, 1,420,663 s of H2L1 time, and 7,742,161 s of H1H2 time.

Figure 8.3 shows the in-time and time-shifted analyzed times after category 2 and 3 DQ flag vetoes have been removed. The strong dependence on time-shift number seen in H1L1 and H2L1 time is due to correlations between H1 and H2 DQ vetoes. Since these vetoes are positively correlated due to their shared environment, as H1 and H2 times are shifted against each other during H1H2L1 times, time that was initially being vetoed by both H1 and H2 DQ vetoes in the in-time turns into time where only one of the Hanford detectors is being vetoed for the time shifts. Since the H1H2L1 time initially has an order of magnitude more time than H1L1 or H2L1 time, the small effect of sliding correlated H1 and H2 vetoes against each other becomes a more significant amount of time transfered to H1L1 and H2L1 analyzed times in comparison to their initial amount of analyzed times. This effect is more pronounced as one moves away from the nonshifted in-time data because the relative shift between H1 and H2 (10 s) is comparable to the typical duration of DQ flags.

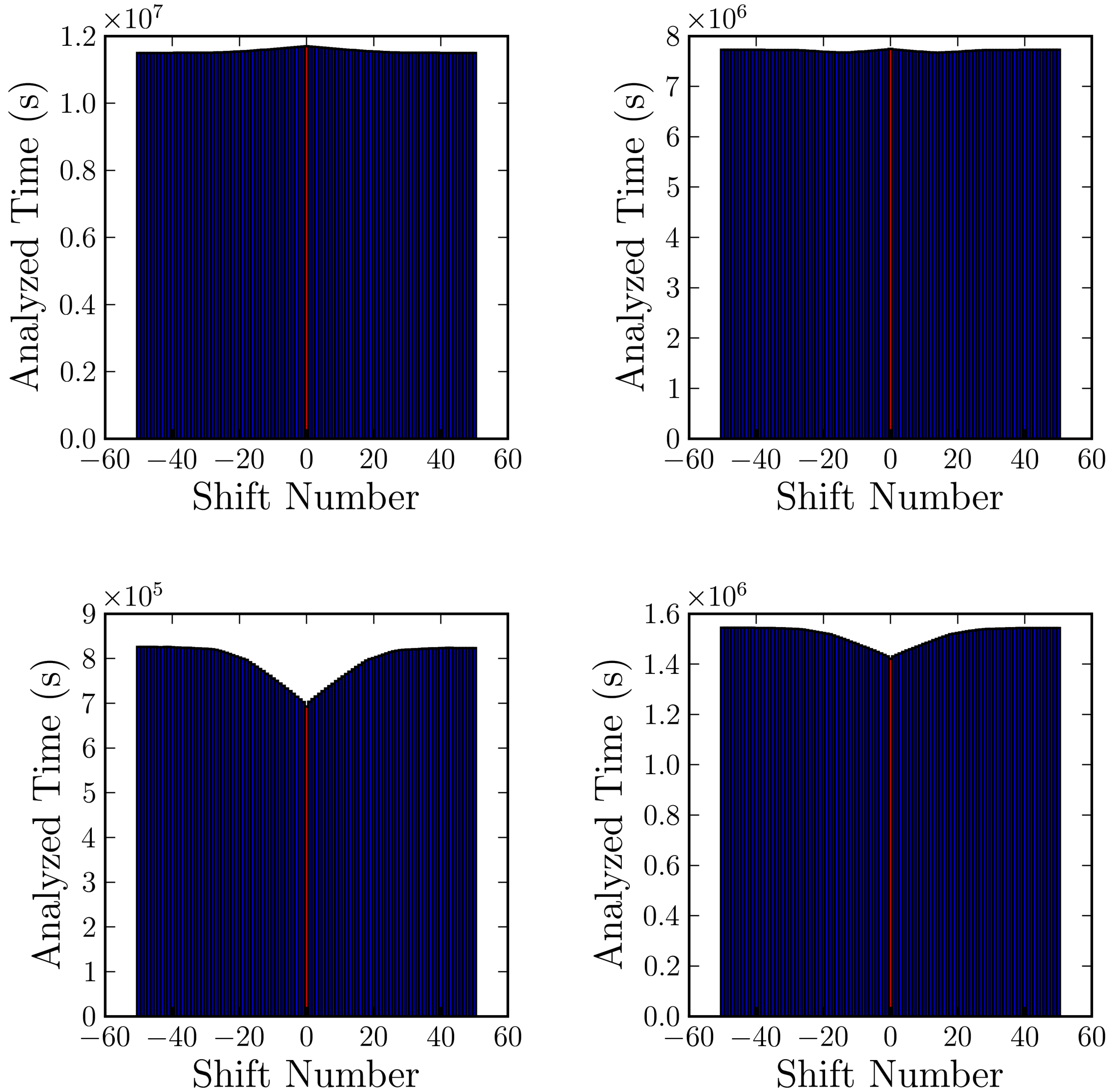

Figure 8.3: Background Analyzed Time
The analyzed times as a function of time-shift number can be seen here after the category 2, and 3 DQ flag vetoes are taken into account for H1H2L1 (top left), H1H2 (top right), H1L1 (bottom left), and H2L1 (bottom right) times. The red bar in each plot denotes the in-time analyzed time.

8.3 Trigger Clustering and Separation by Category

In the next stage of the postprocessing pipeline, we prepare the triggers for the FAR calculation. In order to reduce the total number of triggers from the search to a reasonable size, we cluster the SEPTIME files using a sliding cluster window of 10 s, using the combined effective SNR as the intermediate detection statistic. The output of this thep is one COIRE file per SEPTIME file. This incurs an insignificant loss of signal recovery efficiency since, generally, when there is a glitch in the data, the whole template bank produces triggers, masking possible signals during that time. The clustering we do focuses the glitch into the loudest trigger during that time, which generally has the largest combined effective SNR in the high mass region of the template bank. This effect could be the cause of the louder background combined effective SNR distribution for the high mass region, described in chapter 7.1.

Next we separate triggers into different categories due to the differences between the background trigger combined effective SNR distributions (chapter 7.1). In that chapter we found that the background combined effective SNR distributions can be affected by multiple factors including the coincidence level (i.e., double versus triple coincidence), the signal-based vetoes applied, and the mass of the templates producing the triggers. The first two of these three factors leads us to separate triggers by the different observation times and the different coincidence types (i.e., H1H2L1, H1H2, H1L1, and H2L1 triggers in triple coincident time and H1H2, H1L1, and H2L1 triggers in double coincident times). In addition, we separate these triggers at the chirp mass divisions of $\mathcal{M} = 3.48$ $M_{\odot}$and $\mathcal{M} = 7.40$ $M_{\odot}$.

These trigger categories are implemented by taking the above clustered COIRE files, grouping them together by observation time, separating the triggers into the three mass regions, and subdividing the triggers in triple times further into coincidence types (H1H2L1 triggers, H1H2 triggers, H1L1 triggers, and H2L1 triggers).

8.4 FAR Calculation

The final stage of the postprocessing pipeline is the FAR calculation. The FAR is calculated using the time shifts and analyzed times as described in section 7.2. This is done separately for each trigger category using LALAPPS_CORSE (the COincident Rate Statistic Estimator). Once the FAR is calculated, the categories for a particular observation time are recombined as separate trials and their FARc is calculated (described in section 7.3). In the end, the IFARc is used as our detection statistic where triggers with larger IFARc values are louder.