GEM online data reduction summary for GMn

August 11, 2019

1 GEM online reduction

In the GMn experiment with a 5-layer GEM tracker configuration, there are 40k channels to read using a total of 23 MPDs. The total raw data volume per event is 0.5 MB (= 40k channels * 6 samples * 2 Bytes). At a trigger rate of 5kHz, the raw data rate from the GEMs is 2.5 GB/s. This large data rate is handled by a SSP based system with online data reduction that can reduces the data size by factor of 10 to 15 at the 20% to 25% average GEM occupancy expected in the GMn experiment. The scheme of the system is shown in Figure 1.

1.1 Base version without online data reduction

A preliminary version of the data reduction SSP-MPD system has been established and used to take stable cosmic runs. In this version the SSP simply takes in all raw data from the MPDs and returns it to the readout controller (ROC). The total data rate is limited by the 110 MB/s VME backplane link between the SSP and ROC. In the absence of online SSP-based data reduction, the GMn configuration would be limited by the 110 MB/sec VME backplane transfer rate to roughly 220 Hz in a standard trigger mode.

1.2 Online data reduction algorithm

In the online data reduction version, the SSP uses a firmware based algorithm to remove data that is either from non-fired channels or channels that fire due to random background hits. The removal of non-fired channels is based on zero suppression and the removal of background hits is based on the relative

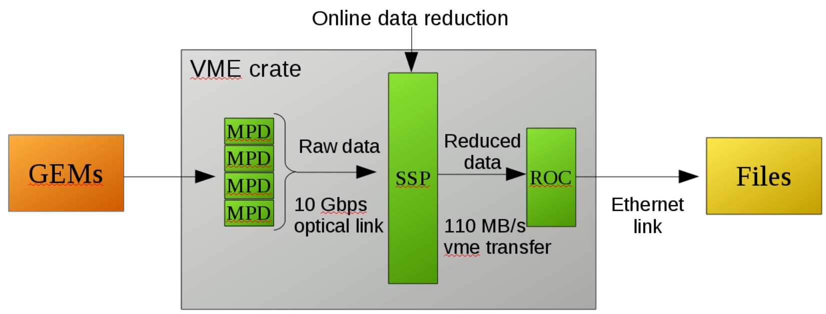

Figure 1: Data flow chart of the current SSP based system. The MPD digitizes signals from APV cards on the GEMs and send it to an SSP. Groups of 4 MPDs share a 10 Gbps optical link which fed into 1 of 8 optical ports on each SSP. Real time data reduction is performed by the SSP which then sends reduced data through VME backplane which has an average transfer rate of 110 MB/s to the crate's readout controller (ROC).

to the trigger. The data reduction algorithm is tested with events that have high local occupancy up to 50% (within a single APV) in the 2016 Hall A GEM test run which had an overall average occupancy of 1.5%.

1.2.1 Zero suppression

Online zero suppression is the first part of the data reduction procedure. It requires computing a common mode baseline of APV samples in real time(Figure 2). The accuracy of the common mode identification is the key to effective zero suppression. This is especially true in high occupancy scenario as every fired channel may shift the common mode by a considerable amount.

In traditional offline GEM analysis, this common mode is found by removing a certain number of channels that have higher rank after sorting on their ADC values aiming to remove fired channels. Since sorting is both computational and memory expensive for the SSP, another simple yet effective algorithm is devised for the online reduction. The method takes the average of all channels that have an ADC value between a wide potential range of common modes and then remove channels that have ADC values 10σ[1] higher than the average. After two iterations, the final average is taken as the common mode. This method uses constant memory and is four times faster than the sorting method. It is implemented and tested using data from GEM test run in Hall A against traditional offline analysis method and shows good accuracy as shown in Figure 3.

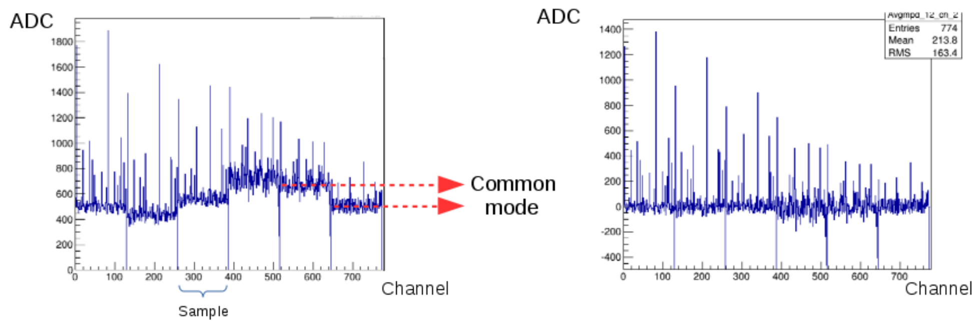

Figure 2: Plot on the left shows how the ADC baseline of channels in one APV fluctuates from sample to sample. These shifted baselines needs to be identified and offset before applying a threshold cut for zero suppression. Channels that legitimately fire within a group need to be identified and excluded from the baseline identification. Plot on the right is after common mode correction.

1.2.2 Random background hit removal using time information of hit

Removal of background hits in the online reduction procedure is designed to very conservative to avoid discarding good hits. It is based on the fact that the hit of interest is correlated in time with the trigger and its peak can be placed in the middle of the readout window, as shown in Figure 4. Since it is extremely rare for a valid signal to have its pulse overlap with the first and last time sample, such out-of-time events can be recognized as hits from random background and be removed. Careful consideration needs to be taken here to avoid accidentally removing good hits that pile up with a background hit in a high-occupancy situation. However, usually when a background hit

[1] σ = pedestal noise RMS

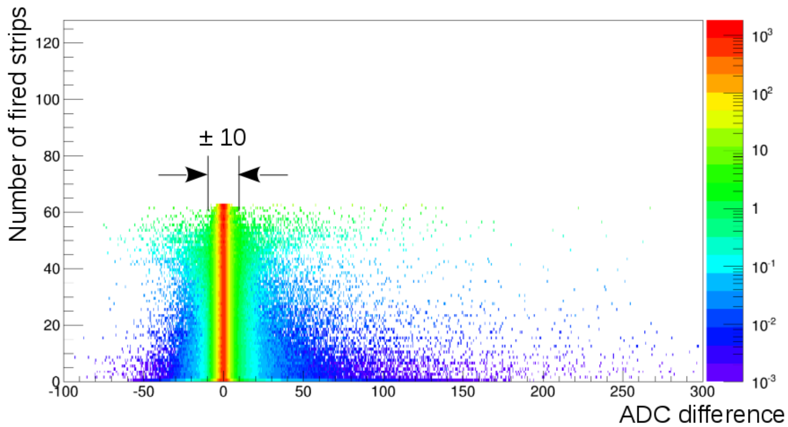

Figure 3: The difference between common mode results from an offline analysis method and the online data reduction system at different occupancy levels. The comparison is done with up to 50% occupancy, where 64 channels fire out of all 128 channels in a single APV. These data are from a GEM test in Hall A where the average occupancy over a full GEM detector was 1.5%, but local occupancy within one APV can often be much larger. APVs were categorized by the total number of fired channels to look at the difference of the two methods at different occupancy levels. As shown in the figure, the difference is well under 20 ADC bins for occupancy up to 50% when the ADC value of a typical MIP hit in the GEM is more than 400.

is large enough to place a channel's peak value into the first time sample, the information of any good hit on that channel is not retrievable, and the channel information is useless anyway. If this data reduction feature is enabled, it can reduce the total data size by about 60%, with some modest reduction in efficiency due to some good hits being rejected in pile-up events.

1.3 Hardware implementation status

As stated above, the correctness of the data processing algorithm has been tested using Hall A GEM test data by running the algorithm in parallel with traditional offline analysis and making direct comparisons. As a next step, the full implementation of the algorithm into the SSP needs to be done. A limited scale version of the SSP system has been built and tested.

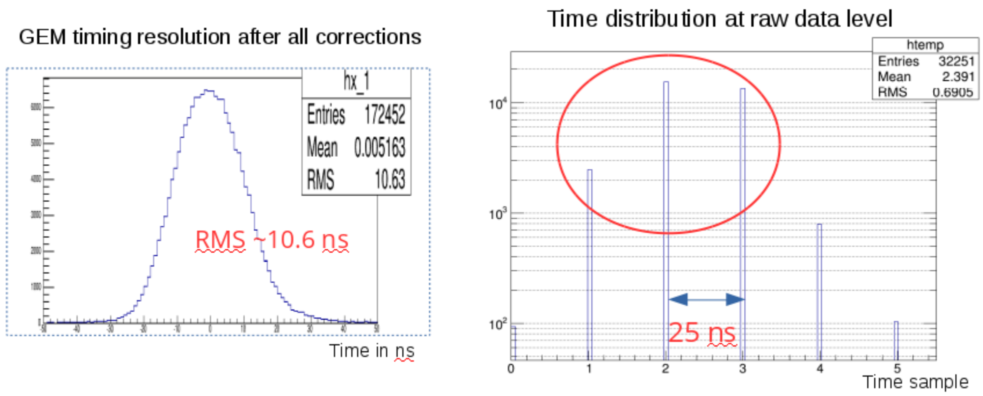

Figure 4: The final timing correlation of good hits with trigger has RMS around 10 ns after corrections (left plot). In the raw channel data before any correction, the timing of a good hit is localized to within 3–4 25 ns time samples. Channels with timing out of this range can be identified as background and suppressed.

In the current system the SSP supports up to 4 MPDs. This system has been used to take stable cosmic runs and the result shows little difference from running an analysis using offline data processing. The final version will support support a full complement of 32 MPDs per SSP.