The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland
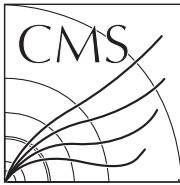
Studies of the CMS Tracker at High Trigger Rate

K. Hahn, S. Nahn, P. Everaerts, K. Sung, S. Tkaczyk, M. Johnson, G. Hall, N. Cripps, J. Fulcher, M. Raymond, and Q. Morrissey

Abstract

During the latter months of 2006 and the first half of 2007, the CMS Tracker was assembled and operated at the Tracker Integration Facility at CERN. One of the important projects undertaken during this period was the assesment of the tracker at trigger rates up to 100 kHz. The test setups, outcome of the tests and subsequent consequences for operation of the Tracker are discussed.

The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland
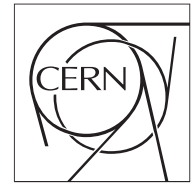
Studies of the CMS Tracker at High Trigger Rate

K. Hahn[1], S. Nahn[1], P. Everaerts[1], K. Sung[1], S. Tkaczyk[2], M. Johnson[2], G. Hall[3], N. Cripps[3], J. Fulcher[3], M. Raymond[3], and Q. Morrissey[4]

[1] Massachusetts Institute of Technology, Cambridge MA, USA
[2] Fermi National Accelerator Laboratory, Batavia IL, USA
[3] Imperial College, London, England
[4] Rutherford Appleton Laboratory, Didcot, Oxfordshire, England

Abstract

During the latter months of 2006 and the first half of 2007, the CMS Tracker was assembled and operated at the Tracker Integration Facility at CERN. One of the important projects undertaken during this period was the assesment of the tracker at trigger rates up to 100 kHz. The test setups, outcome of the tests and subsequent consequences for operation of the Tracker are discussed.

1   Introduction

By any measure (physical dimension, channel count, number of data links, volume of off-detector electronics, etc.) the CMS Tracker [1] is the largest system of its kind ever built. The collaboration paid meticulous attention to detail during the Tracker design and construction and, as a result, detector performance should be more than adequate to provide the quality of data required for success at the LHC. Strict quality control during construction is necessary to guarantee adequate performance in real operation; however, silicon detectors can exhibit unforeseen operational problems after construction [2, 3] that are not amenable to repair. To guard against this possibility, the Tracker collaboration executed a program to test as much of the functionality of the Tracker as possible during operations at the Tracker Integration Facility (TIF) *before* installation at the experimental site. Results from charge collection [4], track reconstruction [5], and alignment [6] studies performed with cosmic rays taken at the TIF indicate that the quality of data collected during Tracker integration is consistent with that established during detector construction.

While this exercise provided important tests of Tracker functionality and operation, running conditions for the Tracker in CMS will be different from those experienced at the TIF. In addition to the challenges posed by the radiation environment, large magnetic field and high multiplicity aspects of Tracker operation, data from the Tracker will be read into the global CMS data acquisition (DAQ) system at high Level-1 trigger (L1) rates of order 50-100 kHz. By contrast, the Tracker performance studies planned for the TIF utilized a VME based DAQ system that is only capable of sustaining L1 rates up to 10 Hz, sufficient for check-out, commissioning and cosmic-ray collection but four orders of magnitude below the expected CMS readout rates.

All relevant components of the final Tracker DAQ had been tested separately at high rate by the summer of 2006. However, an entire readout chain including sensors, front-end readout chips (APVs) and Tracker DAQ electronics had not. In order to bridge this experience gap a DAQ system capable of sustaining high L1 rates with various Tracker systems was assembled. This "DAQ Column" consists of hardware and software components used in the global CMS DAQ. Thus, besides enabling data collection at high rate, the DAQ Column provided operational experience that is directly applicable to the global running of the Tracker in CMS.

The DAQ Column was used to test the high-rate performance of a number of Tracker subsystems, from single TOB modules to $\sim 15\%$ of the Tracker at the TIF. Early tests with a high-rate trigger uncovered a rate-dependent occupancy effect that is common to all Tracker modules. Subsequent investigations on smaller test bench setups revealed that this effect results from sampling particular trigger intervals, corresponding to instantaneous high rate operation. Direct measurement and detailed simulation of the electronic components suggest that the origin of the effect lies within the front-end readout chip.

2   High Rate Data Acquisition Apparatus

This section describes the detector and DAQ systems used in the high-rate studies. The focus is mainly on the S-Link based "DAQ Column" used in the majority of the tests. The DAQ Column persists as a test-stand and development tool for the Pixels and Preshower detectors, as well as a diagnostic laboratory during CMS operation.

2.1   The Tracker and the Slice test detector at the TIF

The CMS Tracking system is composed of an inner pixel detector as well as an outer tracker based on silicon microstrips, the latter of which is the subject of this note. The tracker is further divided up into subdetectors: 4 cylindrical layers comprise the inner tracker barrel (TIB), 3 inner disks on each end the inner disks (TID), 6 outer cylindrical layers the outer barrel (TOB), and 9 disks on each end the end caps (TEC), all together roughly 10 million strips. Each subdetector is composed of silicon sensors capacitively coupled to frontend readout chips, which are called "modules". Modules are further combined into "Strings", "Rods", and "Petals" for the TIB/TID, TOB, and TEC, respectively, which share the same power and readout services. The Tracker Technical Design Report [7, 8] provides more detailed information on the specific subdetectors.

Briefly, the detector measures ionization deposits from charged particles traversing the material in the following manner: the charge information from the strips is sampled every 25 ns and stored in an analog pipeline pending a decision to read out a particular 25 ns sample, which is called a trigger. The readout process entails converting the analog charge into an optical signal, adding header information, and transmitting the optical signal to a FrontEnd Driver (FED) , which digitizes the data, performs pedestal subtraction and zero suppression, and then passes the data further downstream to the CMS DAQ, as described below.

The "Slice Test" at the TIF utilized a subset of the full detector, including pieces of each subdetctor. Approximately 15% of the Tracker was fully instrumented with power supplies and readout electronics, and operated with a scintillator-based cosmic ray trigger as well as programmable triggers used to drive high rate acquisition. More details on the Slice test setup can be found in [4–6].

2.2   Single-rod test bench

As the Slice was not continually available for high rate studies, and the ongoing investigation indicated a systematic effect not related to the scale of the detector under test, a spare TOB rod also at the TIF, was employed in place of the actual detector. This rod is composed of 6 modules with 4 front-end chips per module, connected to the DAQ in exactly the same manner as the Slice test. The use of the rod also allowed access to the front-end electronics for probing during high rate operation, which provided more information regarding operation at high rate. In this manner the high rate studies could be pursued in parallel with other programs during TIF operations.

2.3   Single-module test bench

While the TIF systems allowed data acquisition in a realistic way, as it will happen in the final experiment, a second DAQ system at Imperial College was used to investigate the APV chip and sensor behavior in a more controlled way on the benchtop. This DAQ system employed a programmable digital pattern generator to provide the 40 MHz clock and trigger patterns to a single TEC module. Sending repetitive sequences of triggers relative to a reset trigger pattern, (two triggers separated by a single clock cycle initialize the APV pipeline control and readout

logic) allowed systematic effects to be investigated. The output data streams were digitized with a commercial ADC. It is worth noting that the control and digitizing hardware in the laboratory were entirely different to that used in the TIF system, allowing any effects observed in both systems to be attributed to the front end module alone.

2.4 CMS DAQ

It is instructive to review the design of global CMS DAQ [9] as an introduction to the hardware and architecture employed in the High Rate DAQ at the TIF. CMS subdetectors generally have two interfaces to central DAQ. The FrontEnd Driver/FED ReadoutLink (FED/FRL) interface is an S-Link64 data path through which subdetector event fragments are routed to the first tier of event building PCs (ReadoutUnits, "RU"s) according to resource availability. The FRLs accept input data from one or two FEDs and transmit event fragments on one of two optical links that connect to a Myrinet switch, which routes the fragments to corresponding sets of RUs in the event building system. The switch can be partitioned into "rails" to provide multiple, independent data paths between FRLs and RUs. At the RUs, FED event fragments are concatenated and buffered until requested by a BuilderUnit (BU), which assembles these data into a complete event and passes it to the FilterUnits (FUs), which execute the High-Level Trigger (HLT) algorithms on the complete events. Communication between the RUs/BUs/FUs is managed via Gigabit ethernet switches. Events that satisfy HLT selection criteria are forwarded to the StorageManager for archiving.

The second subdetector/central DAQ interface is the CMS Trigger Throttling System (TTS). FEDs communicate the state of on-board event buffering to FastMergingModules (FMMs) to prevent data overflows. If a FED indicates that its buffers are nearly full, the FMMs throttle the trigger to allow event data to drain. Front-end emulators (the APV emulator (APVE) [10], in the case of the Tracker) prevent similar overflows in the subdetector front-ends.

2.5 DAQ Column at TIF

The DAQ Column is a smaller scale version of the CMS DAQ described above, with several modifications necessary to achieve high rate. Figure 1 is a simple schematic of the DAQ Column system, showing FEDs feeding data through FRLs into a Myrinet switch, which subsequently feeds several PCs. Each PC runs the RU, BU, and FU processes. This single PC approach was chosen to avoid significant overhead in data transfer that distributed RU/BU/FU's would contribute to a small scale system. For the same reason the StorageManager process, which collects, concatenates and writes data to disk, was omitted from the DAQ Column, and alternatively the FU process spooled its output directly to disk.

In order to provide a multi-kHz trigger, a Local Trigger Controller (LTC) was employed to generate either Poisson distributed or fixed-frequency triggers. In contrast to the VME based cosmic data acquisition, high rate operation also required implementation of the CMS TTS system to avoid overwhelming the upstream DAQ processing. Two daisy-chained FMMs to merge and forward FED buffer status as well as an APVE to prevent pipeline overflows in the Tracker front-ends were implemented, with output sent to the LTC to moderate the trigger rate. Note that the triggering scheme does not involve the presence of real signals and the data collected thus represents the noise behavior of the Tracker at high rate.

Although the system evolved as the tests progressed, at its maximum capacity the DAQ Column was capable of running 32 FEDs connected to 16 FRLs with subsequently four Dell 1850s PowerEdge servers acting as the RU/BU/FUs through the Myrinet switch. The bandwidth was partitioned equally among the 4 PCs using both "rails" of the Myrinet to obtain maximum

throughput. This permitted high rate readout of up to one half of the Tracker slice available at the TIF.
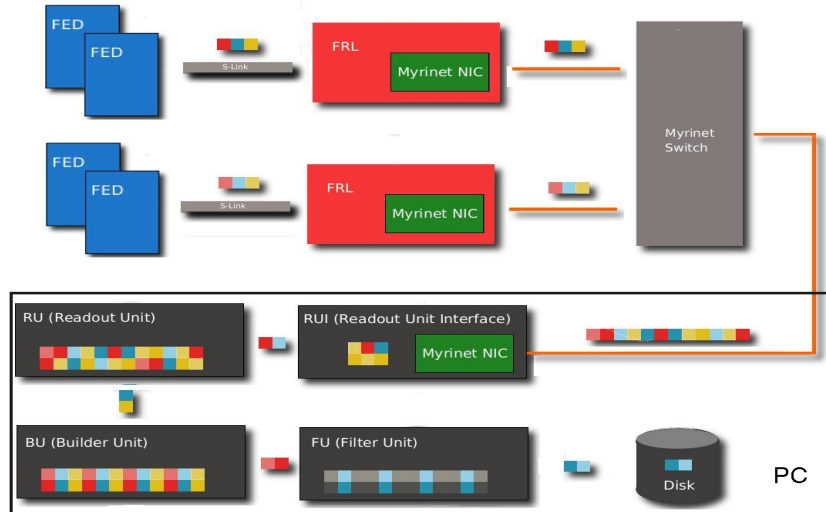

Figure 1: Schematic of the basic DAQ Column, depicting event fragments passing through the system. Different events are represented as different color fragments, and shading indicates different sources. Various modifications made to accommodate sampling at high rate are described in the text.

2.5.1   Increasing the DAQ Column Bandwidth

A single-PC DAQ Column was commissioned using a single FED loaded with emulated, low occupancy data as the first step in the high rate tests. Zero-suppression in the FED allowed the production of simulated events of tunable size around 1 kB. Even for such small events, the system as implemented could only tolerate L1 rates of ~ 10 kHz (see Table 1), motivating additional adaptation of the original DAQ Column concept to be able to reach higher rate.

Occupancy	Max Trigger Rate (kHz)			
(%)	FRL	RU	BU	FU
1	135	89	49	14
2	98	89	49	14
3	73	61	41	13
4	63	59	41	13
5	51	55	41	12
6	45	36	36	11
7	38	37	36	11

Table 1: Maximum trigger rate for various fixed strip occupancies, where the data are discarded at different points along data path, after the FRL, the Myrinet switch (RU), the High Level Trigger (BU), and after a prescale trigger after processing (FU). The dramatic loss of rate inside the HLT motivated applying the prescale further upstream.

The first efforts to relieve I/O bottlenecks focused on the FilterUnit process. The disk I/O operation in the FU process blocks further event processing while writing event data to disk, thus slowing the entire event building chain. This bottleneck was removed with a modification to the FU to store event data in a shared memory ring buffer, from which a standalone

reader application extracted and wrote to disk. This two-stage approach prevented blocking in the FilterUnit by allowing unread data in the ring to be overwritten. The reader process was scheduled with low priority so that disk I/O was performed only when system resources were available. In addition, a prescale[1] in the FU was implemented to limit downstream data flow. The prescale does not jeopardize the goals of the study because the noise behavior of the Tracker is not expected to change instantaneously. These changes to the FU increased the sustained L1 rate to $\sim$ 30 kHz. The next modification was to add a second Myrinet rail on the RU/BU/FU to further increase system bandwidth. The additional NIC card was installed on a separate PCI bus to establish a second, independent path for data input to the PC. At the time, RU software capable of controlling multiple cards had not been released, so a beta-version provided by the central DAQ group was installed in order to operate both cards simultaneously. Once operational, the additional Myrinet input doubled the sustained rate to $\sim$ 60 kHz.

The remainder of the bandwidth limitation was traced to processing inside the RU/BU/FU. To alleviate this, a tunable prescale in the RU software was introduced to reject a large fraction of events before they are injected into the event builder. This required modifications to the RUI software so that "missing" events are not construed as errors. The prescale relieves much of the remaining load from the event building applications and, in combination with the preceding modifications, permits sustained L1 rates of $\sim$ 140 kHz.

The modified DAQ Column was re-commissioned by repeating the fake data throughput tests. The results in Figure 2 show sustained trigger rate as a function of data size for the scenario where data passes through the entire DAQ system. The plot displays two versions of data formatting, zero-suppression ("ZS", filled circles) and zero-suppression with compressed header information ("ZS-lite", empty circles). For this discussion, the most important feature of the plot are the first four entries showing that the modified DAQ Column can sustain O(100 kHz) with low occupancy data. The plot also shows a clear drop at 4 kBytes that stems from a hardcoded maximum packet size in the Event Builder software. This corresponds to a Tracker occupancy at roughly 2%, which is the maximum expected occupancy for normal operation. This effect is noticeable only because the packets were of constant size; it does not arise in real data containing packets that vary in length from event to event.
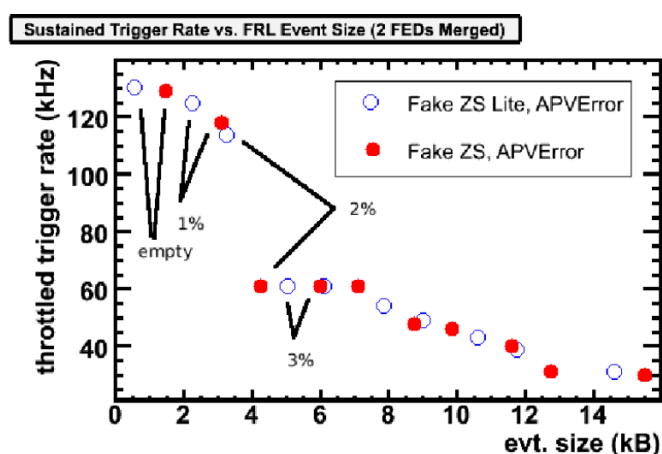

Figure 2: Throttled Event Rate as a function of event size, with events flowing through the FRL, switch, and into the HLT. Event size in terms of strip occupancies of 0-3% is also indicated. The drop in performance is due to a fixed packet size of 4 kB at FRL output.

---

[1]A prescale is a filter which a priori discards a given fraction of the events from the data stream independent of the event content

3   Tests & Results

With the establishment of high rate operations at the TIF, studies of the detector performance at high rate quickly revealed an effect not previously seen. This section describes the investigation which first characterized the dependencies of this new effect through a series of studies, and then turns to the pursuit of the potential causes which match the characterization.

3.1   Tests with the Tracker at High Rate

The TOB was the first Tracker sub-detector available for testing, initially using pedestal and noise measurements to compare S-Link and VME readout, and compare S-Link readout at different trigger rates. These measurements require so-called "Virgin Raw" (VR) data, where the data are not sparsified. The large event sizes that result in this running mode limit the maximum trigger rate to $\sim$ 5 kHz. Figure 3 shows that pedestals and noise are stable over all channels at low and moderate trigger rate. In addition, no differences between VME and S-Link results were detected.
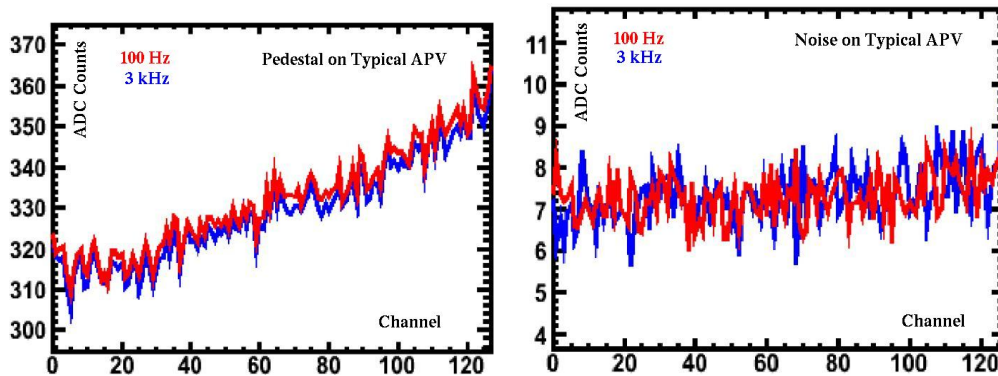

Figure 3: Pedestals and noise versus rate for a typical TOB APV run at 100 Hz and 3 kHz using S-Link readout. The plots show no appreciable differences in results obtained at the two trigger rates.

Next, zero-suppression (ZS) was enabled in the FEDs to achieve the smaller event sizes needed for reaching higher rates. In ZS-mode, the FEDs subtract pedestals from the raw ADC data in each event and calculate per-APV common-mode levels from the result [11]. After subtracting the common-mode, the algorithm forms clusters by requiring either a single strip with an ADC count above 5σ or two or more adjacent strips with ADC counts above 2σ where σ indicates raw noise, i.e. the RMS of the pedestal for that strip. The FED outputs the corrected ADC data of strips in such clusters, along with information on cluster size and position. This information is used to calculate strip occupancy, defined as the frequency at which a given strip is included in a FED cluster.

Initial tests with the TOB revealed an increase in occupancy at the edge channels of the APV chips with increasing trigger rate, which becomes evident at rates above $\sim$ 30 kHz, both with biased and unbiased sensors. Tests with the TEC and TIB established this high-rate noise effect (HRN) as a universal feature in the Tracker. An example of the growth in occupancy with increased rate from one TEC fiber, which carries data from two APVs, is shown in Figure 4. In addition to increased occupancy, the edge-strip cluster sizes and ADC values also increase with trigger rate. Figure 4 also plots the distribution of cluster charges (summed ADCs) versus rate for two APVs in a fiber. For reference, a minimum ionizing particle (MIP) is expected to deposit clusters of charge around 100 ADC counts.
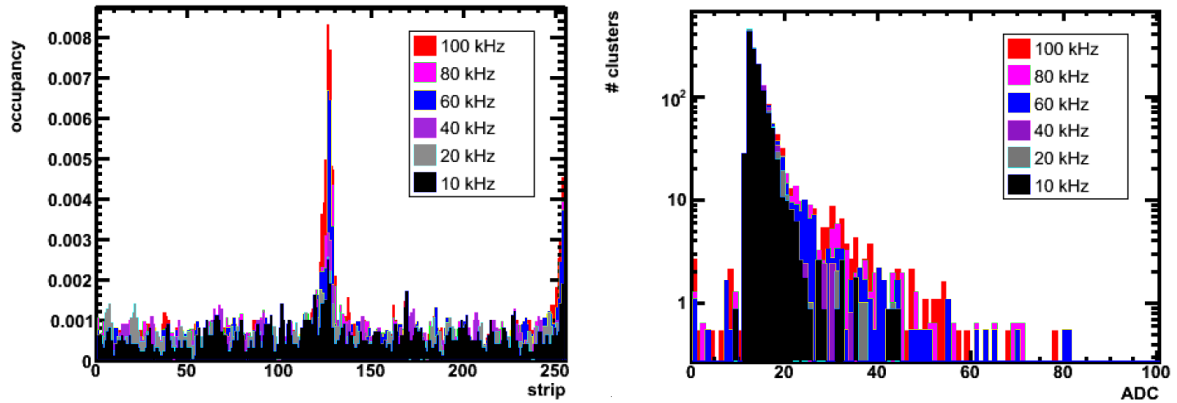
Figure 4: Occupancy (left) and Cluster Charge (right) for a single TEC fiber versus L1A Rate. A clear increase in occupancy with trigger rate is observed again on channels near 127 for both APVs on the fiber, confirming the result found with TOB modules. In addition, the tail of the Cluster Charge distribution also grows with rate.

Once the correlation with rate was established, the studies turned to understanding the source of the effect. First, by selecting events with large ADC counts in strip 127, the time-correlation of events with HRN occupancy peaks was examined. Figure 5 shows that large clusters comprised of strips with high ADC count appear *simultaneously on every APV* when HRN occurs, indicating the effect is independent of particular detector, but is triggered by some global signal or systematic effect from the APV or FED which affects all detectors.
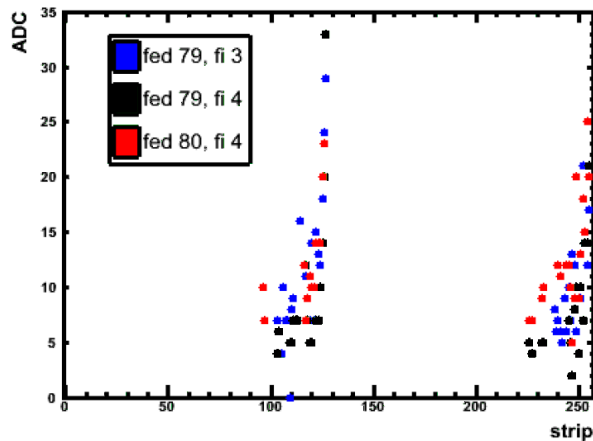

Figure 5: ADC distribution for multiple fibers on different FEDs from the same event, indicating that the HRN occurs simultaneously on every APV independent of the type of module that the APV services.

In order to obtain more information on HRN events, a modification was made to the FED firmware to allow collection of VR data at high rate. Front-end buffers in the FEDs can hold data from at most three VR events, which limits operation in VR-mode to low trigger rates. A prescale in the FED frame-finding logic was implemented to retain data from all strips while triggering at high rate, essentially moving the prescale to the FED input. The modification permitted sampling of VR data at 100 kHz, allowing the examination of ADC values on all strips when HRN occurs. Figure 6 plots ADC values of strips 0 and 127 for events in which the strip 127 ADC exceeds 5σ. The mean of the channel-0 data is well below the chip average,

indicating a strong anti-correlation between the ADC counts of channel-0 and channel-127. In order to confirm that correlation, events where either strip 0 or 127 has a high ADC count are selected. Figure 6 shows the ADC correlation at different trigger rates, which becomes more pronounced as trigger rate increases. Thus the source of the effect must be able to affect all APVs simultaneously and reproduce this characteristic anti-correlation.
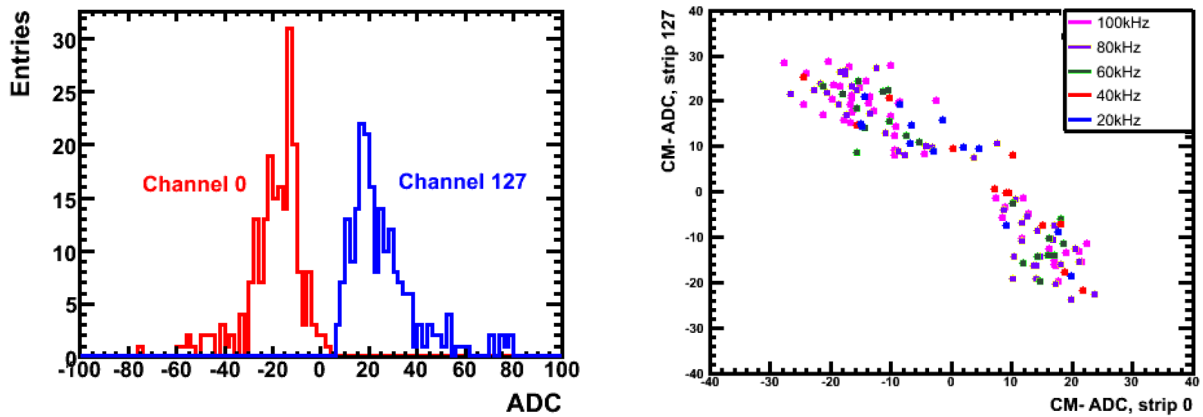

Figure 6: Left: Common Mode subtracted ADC counts for Strip 0 and Strip 127 when Strip 127 has a cluster over threshold. The offset from zero for Strip 0 indicates a strong correlation between the HRN on these two channels. Right: Average ADC count of Strip 0 versus that of Strip 127, when one of them is above threshold, at various different trigger rates, confirming the anti-correlation and its dependence one trigger rate.

In further pursuit of the source of the HRN, the temperature dependence and running mode were investigated. Changing the temperature of the silicon sensor reduces the intrinsic sensor noise, such that a temperature dependent effect could be responsible for the effect. Tests with the TIB at $-10°$ C show increased occupancies at lower temperature, but resulting from a similar amplitude of the HRN effect coupled with lower clusterization thresholds from the lower intrinsic noise. This exonerates the sensor as the source of the noise. In addition, tests using the APV chip in Deconvolution mode [12] rather than Peak mode show little difference between the two running modes with respect to the HRN effect.

3.2 Determining the Origin of High Rate Noise

Armed with the information that all modules behave the same with respect to the HRN, the high rate investigation was pursued in more detail using the single-rod test bench (section 2.2) connected to the DAQ Column. As expected, the HRN effect appeared on the rod modules as well, once they were examined at high rate. The single-rod investigation focused on three possible HRN sources; external noise electrically coupled in through the sensors or cabling, noise generated from pathological trigger conditions, or noise sourced by internal APV operations.

3.2.1 Wing Noise Investigation

The strip profile of the high-rate occupancy spikes observed is reminiscent of the TOB-specific "wing-noise" found in earlier studies [13], where the effect arose from noise on the power distribution lines coupled in via the sensors. Those studies demonstrated that the wing effect can be mitigated by the application of copper shielding between modules and the power bus or by the introduction of inductive filtering on the control power cable. These modifications were implemented on the single-rod test setup to quantify the impact on HRN. Figure 7 compares

occupancies obtained with and without the application of shielding. The plot demonstrates that while shielding does diminish low-rate wings (compare the red histograms in the two graphs), it has no impact on the occupancy peaks observed in high-rate data. Similar results in data taken with and without inductive filtering demonstrated that HRN and wing-noise are distinct effects.
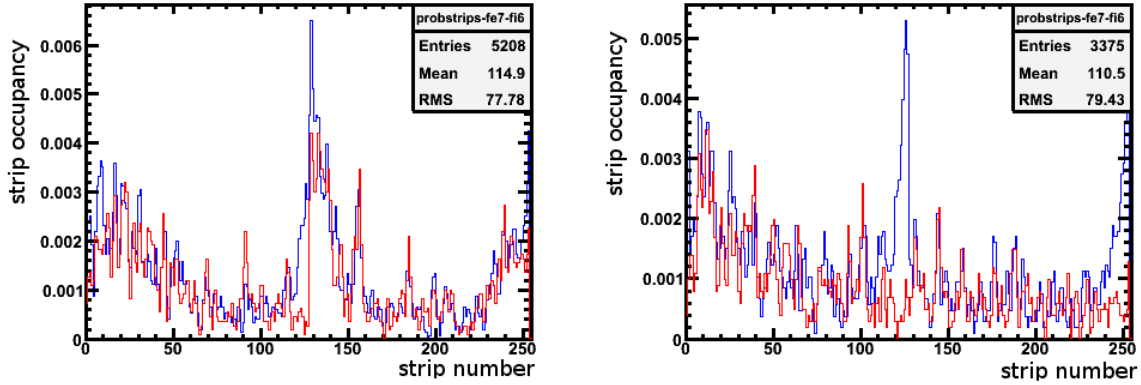

Figure 7: Left: Comparison of strip occupancies taken at low (red) and high (blue) trigger rate from a fiber on the rod. Results from the low-trigger rate data indicate the presence of "wings" (see [13]). Right: The same fiber taken at the same trigger rates as left, but after application of copper shielding. The results indicate that shielding is effective in reducing TOB wings but has no effect on high-rate noise.

3.2.2 Trigger Controller Investigation

The initial studies focused on running detectors with high-rate Poisson triggers because this closely approximates the triggering scenario expected in CMS. Later tests with a fixed-frequency 100 kHz trigger *did not* reproduce the occupancy spikes observed with the Poisson trigger. This result motivated a study of possible differences in APV operation under the two triggering schemes, and in particular the behavior of the LTC in these two modes. The first possibility was that LTC may not respect "trigger rules" (requirements on the minimum number of bunch crossings allowed between multiple triggers) when operating in Poisson-mode. There is a requirement of a minimum spacing of 3 clocks for the operation of the APV [12]. In addition, trigger intervals generated by the LTC in Poisson-mode are not truly random and are instead selected from a look-up table implemented in firmware. It was conceivable that a pathological selection of trigger intervals from the look-up table could cause problems for the APV.

The LTC behavior in Poisson-mode was investigated by enabling its event FIFO. This feature allows the LTC to write the orbit numbers and bunch crossings of triggers to disk at a user-specified frequency, allowing the reconstruction the distribution of triggers that the LTC sends to the APVs. Figure 8 plots the distribution of trigger intervals for the 100 kHz random trigger. The intervals are indeed Poisson distributed. The inset more clearly shows intervals in the 0 - 500 ns range. The absence of intervals below 75 ns confirms that the LTC respects the rules when generating Poisson triggers.

3.2.3 Theory of APV operation

With the LTC absolved, the focus turned to the APV, in particular, the scenario that HRN originates from the reading of particular APV pipeline locations. The APV contains a 192×128 cell
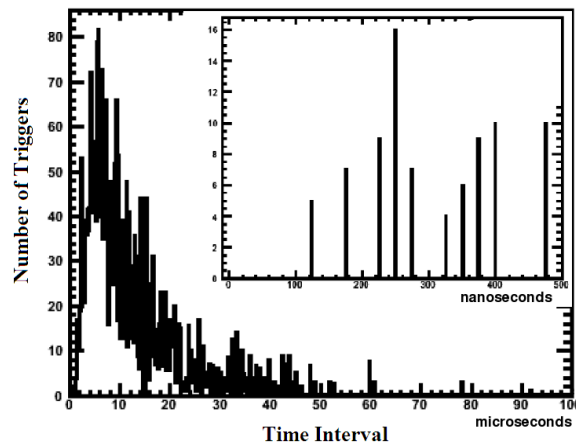
Figure 8: The distribution of trigger intervals coming from the trigger module. The distribution appears Poissonian, and the inset verifies that the trigger rules for which the Tracker was designed are being obeyed, thus the HRN is not the result of a pathological trigger situation.

pipeline that holds charges read from each of 128 silicon microstrips [12]. Following a master reset the internal APV chip logic is initialized and the digital pointers controlling pipeline access are launched and begin to circulate. A write pointer shifts through the pipeline controlling the sampling of the front end amplifier output at 40 MHz. A trigger pointer follows behind with a time delay equal to the programmed latency. When an external level 1 trigger occurs one (three) pipeline cells in peak (deconvolution) mode corresponding to the current trigger pointer location are marked for subsequent readout, and these cells are not overwritten until the readout process has completed. The pipeline readout is governed by a separate cycle with a period of 1.75 μs (70 clocks at 40 MHz), and a phase also determined by the master reset signal. The phase of this pipeline readout cycle is reflected in the output data stream as tick marks, large amplitude signal levels that represent digital data. The tick marks allow external logic to synchronize to the APV output phase so that the start of an output data frame can be detected when, in contrast to a tick mark, the output does not return to the baseline after one clock.

The APV has two main modes of operation, peak and deconvolution. In peak mode one sample per channel is read from the pipeline following a trigger, and then transferred to the output via the multiplexer; the sample should correspond to the maximum amplitude from the CR-RC shaped front end amplifier, which has a time constant of 50 ns. In deconvolution mode three samples are read sequentially and a weighted sum formed. The deconvolution operation [14, 15] results effectively in a re-shaping of the analogue pulse shape to one which peaks at 25 ns and returns to the baseline within one further LHC clock cycle. These operations take 4 pipeline readout cycles, after which transmission of the output data frame can commence. For each APV, a 12-bit digital header precedes the 128 analogue channel samples, creating a data frame of 7 μs total length, the same duration as the 4 pipeline readout cycles. Chip readout is simplified by matching the output data frame and pipeline readout durations, since at high trigger rates triggered data stored in the pipeline can be read out while data from the previous trigger are being multiplexed out.

The Analogue Pulse Shape Processor (APSP) is the part of the APV chip which performs the deconvolution operation. The circuit diagram for the APSP is shown in figure 9. During processing, a series of switches in the feedback network of a high gain amplifier are opened and closed in sequence to apply the appropriate weight to each of the three samples, and then, in a

final cycle, make the weighted sum. The cycle time of the APSP circuit is chosen to be 1.75 μs so the total processing time matches the 7 μs readout time of the APV. As explained earlier, tick marks in the APV output data stream record the transitions of this cycle when there are no output data to be transmitted.
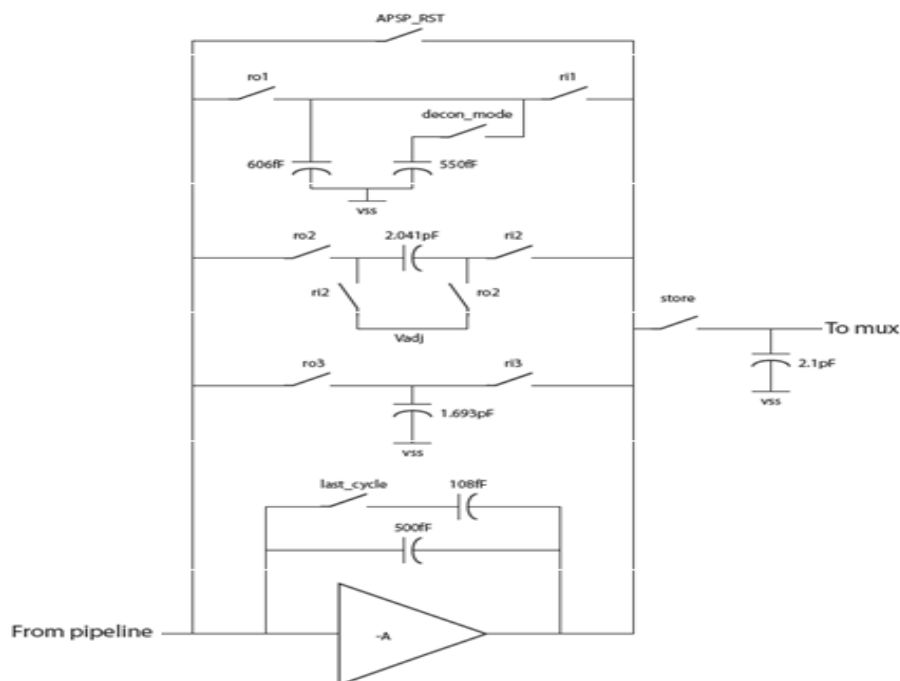

Figure 9: The APSP circuit diagram. The switches labelled ro1,2,3 and ri1,2,3 open and close during the APSP readout cycle for sample processing.

3.2.4   APV behaviour with controlled trigger spacing

With an understanding of the APV logic, one can arrange to read data stored in particular APV pipeline locations by appropriately specifying the intervals between a master reset (RST) command and trigger (T1). A series of runs scanning a series of T1-RST time separations were taken to compare the occupancies measured from each pipeline cell. The left plot of figure 10 shows the maximum occupancy divided by the average occupancy within a fiber, for all fibers in the rod modules, as a function of pipeline position. This occupancy does not differ appreciably, indicating the physical location of the pipeline cells plays no role in generating HRN. Similar analysis as a function of pipeline address rather than position, as well as readout phase relative to the trigger also vindicate these as possible causes of HRN.

In addition, the possibilities that charges stored in given pipeline locations are influenced by the dequeuing of data from adjacent cells or that there was a correlation of the trigger position with the RST command to the APV were investigated, but in both cases no correlation was found. However, in order to investigate the effect of trigger intervals in the Poisson distribution that are not sampled by a fixed-frequency trigger at the same average rate, a second trigger (T2) after T1 was added with a variable delay between the triggers, and the data that T2 samples was analyzed. The right plot of figure 10 shows again the maximum over average fiber occupancy, now obtained as a function of T2-T1 separations. The plot suggests that several trigger intervals (in particular, intervals of 100, 160, and especially 380 clocks) have significantly higher occupancies associated with them, for multiple fibers. This was the first indication that
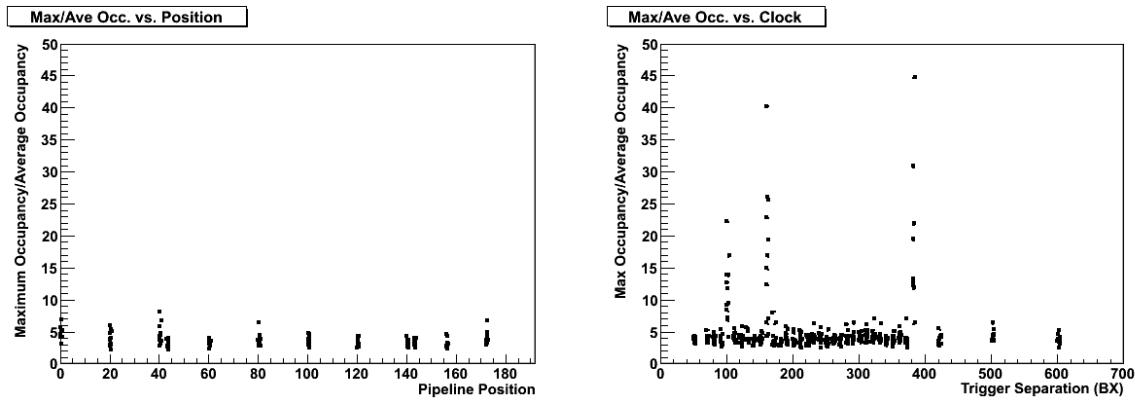
Figure 10: The maximum occupancy divided by the average occupancy within a fiber (2 APVs) for all fibers connected to the rod. On the left, as a function of position of cell in the pipeline, showing no clear indication of the HRN. On the right, as a function of time separation between two consecutive triggers, where certain trigger intervals show the HRN effect for all fibers simultaneously.

the HRN was caused by two triggers with very specific timing between them, but begged the question of what was going on inside the APV causing this interference.

Additional tests with the single-module test bench (section 2.3) were performed to further probe the behavior observed in the TIF system. The digital pattern generator was programmed to cycle repeatedly through a sequence consisting of a reset, a fixed delay to allow the APV pipeline logic to initialize, then two normal triggers where the first trigger time was fixed relative to the reset, but the delay between first and second triggers was varied.

Figure 11 shows the second trigger pedestal data dependence on first and second trigger separation, for an APV edge channel (127 in this case) on a TEC module. The data were taken in the deconvolution mode of operation for different values of programmed latency, and averaged over many triggers, removing the random noise component of approximately 40 rms ADC units, which would otherwise dominate and obscure many of the smaller features. Large pedestal disturbances well above noise levels are evident. Most importantly, the excursions occur at trigger separations which depend on the programmed latency, suggesting that they originate in on-chip activity.

Figure 12 shows a similar picture to figure 11, but for module channels 127 and 128, and for just one value of programmed latency, in this case 130 clocks. This demonstrates an anti-correlation in the pedestal disturbances between edge channels of neighboring APVs, as these are. The pedestals associated with the feature at second trigger positions 157 and 158 have been labeled in the figure to illustrate this anti-correlation more clearly.

The APV output at actual trigger time represents the state of the APV output at the instant the second trigger was applied to the chip. It should be noted that the APV output data frame that can be observed in this representation results from the first trigger. There are no obvious correlations between the pedestal patterns arising from the second trigger and the APV output representation, but if the crosstalk is generated at the chip input, as suspected, it is necessary to take the programmed latency into account. From the APSP circuit schematic (Figure 9), the simultaneous operation of the switches will draw current from the power supply rails, which is liable to generate significant, impulse-like, current fluctuations throughout the system which may couple back into it and potentially generate noise. This APSP transition switching
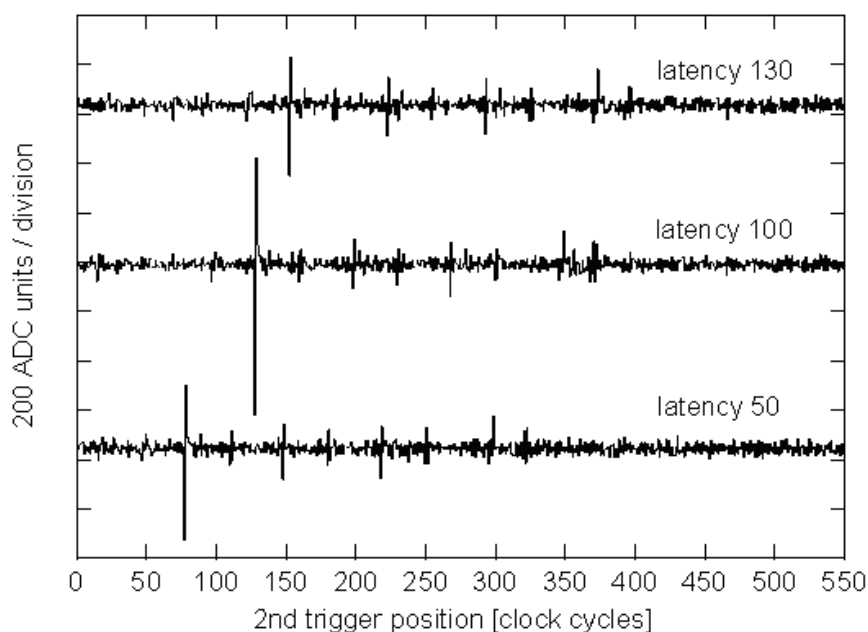
Figure 11: Second trigger pedestal data dependence on first and second trigger separation, for APV edge channel (127 in this case) on a TEC module. Data were taken in the deconvolution mode of operation for 3 different values of programmed latency, and have been averaged over many triggers to remove random noise. An arbitrary vertical offset has been applied to the curves for different latencies for clarity. On this scale, the first trigger was applied at second trigger position -3.

activity causes interference to couple into the APV inputs which will affect data written into the pipeline at that instant, but the subsequent trigger that would access that data would be applied one latency period later.

For example, at second trigger position 380 in figure 12 there is a disturbance visible in the output data retrieved from the pipeline by the second trigger. But these data were actually written into the pipeline 130 clock cycles earlier so should be compared with the APV output state at second trigger position ∼250. At this position the APV output data frame resulting from the first trigger was just beginning. The *APV output offset by trigger latency* represented in figure 12 is just the *APV output at actual trigger time* shifted by the latency (130 cycles here). It therefore represents the state of the APV output when data corresponding to the second trigger were being written into the pipeline. Viewing this trace, the repetitive patterns and spacing between similar features (often 70 clock cycles) imply a connection with the internal APSP cycle of the APV. For example there are small spikes at the times of the tick marks, and the APV output frame header.

Although the tick marks correspond to features in the APV output data, the interference is not produced by a coupling between the output data signals and the chip inputs, but by switching activity within the chip, associated with the pipeline readout phase that is reflected in the tick-marks. Note that the amplitude of the "tick-mark disturbance" is relatively small and would normally be lost in the random noise (which has been averaged out in figures 11 and 12). If this were not the case then high rate noise would also be observed at lower rates (since the tick-marks are present all the time, except during output data frames).
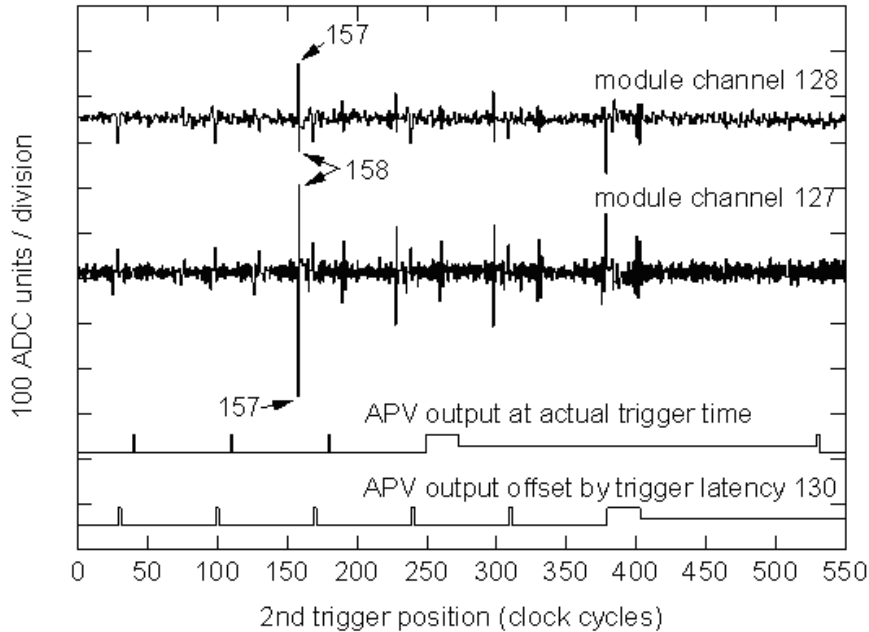
Figure 12: Second trigger pedestal data dependence on first and second trigger separation, for channels 127 and 128 on a TEC module, (corresponding to edge channels on neighboring APVs). Data were taken in deconvolution mode and averaged over many triggers to reduce random noise. The pedestals associated with the feature at second trigger positions 157 and 158 have been labeled to illustrate this anti-correlation more clearly. Two representations of APV output data are shown (see text for explanation). An arbitrary vertical offset has been applied to the curves for different channels for clarity. On this scale, the first trigger was applied at second trigger position -3.

The most prominent feature in the APV output data is the effect in clock cycles 157 and 158. Simulations have shown (section 3.2.6 and figure 17) that this feature corresponds to the closing of switches in the APSP circuit during the period when the first data samples are retrieved from the pipeline. There are smaller effects 70 and 140 clock cycles later, which correspond to the retrieval of the second and third samples. The later features which correspond to the APV header are associated with the APSP readout operation, and the sample/hold stage that precedes the APV output multiplexer. Although the other features in the output stream are evident after signal averaging, they do not represent significant additional noise. However the feature in cycles 157-158 is significant and believed to the main origin of the high rate noise.

3.2.5   Measurement of electrical interferences

The fact that the single-module test bench does not employ the CMS tracker DAQ hardware or software gives confidence that the HRN does not result from that system, but is intrinsic to the front end module. An interesting feature was observed when the module was tested without a sensor (a readout hybrid alone), where the HRN disappeared. This result suggested that the noise correlated with the APV readout cycle is generated within the module, but that the sensor plays a role in coupling the noise to the APV front end.

In order to probe further into the coupling of the HRN, a module was modified with oscilloscope probe connections to the GND, 1.25 and 2.5V supply lines and connected to the readout.

The optical output from the module was input into a photo-diode converter, to monitor the APV data on the same oscilloscope while triggering on the reset signal. Keeping the position of T1 fixed, the position of T2 was set such that its sampling point (T2-latency) aligned with the first clock of T1 readout, concurrent with the transmission of the APV header, with the APVs in peak mode and latency set to 100 clocks. A trigger interval of 362 clocks achieves alignment for this latency. Figure 13 shows the positions of the second trigger and the point at which its data was sampled, 2.5 μs "earlier".
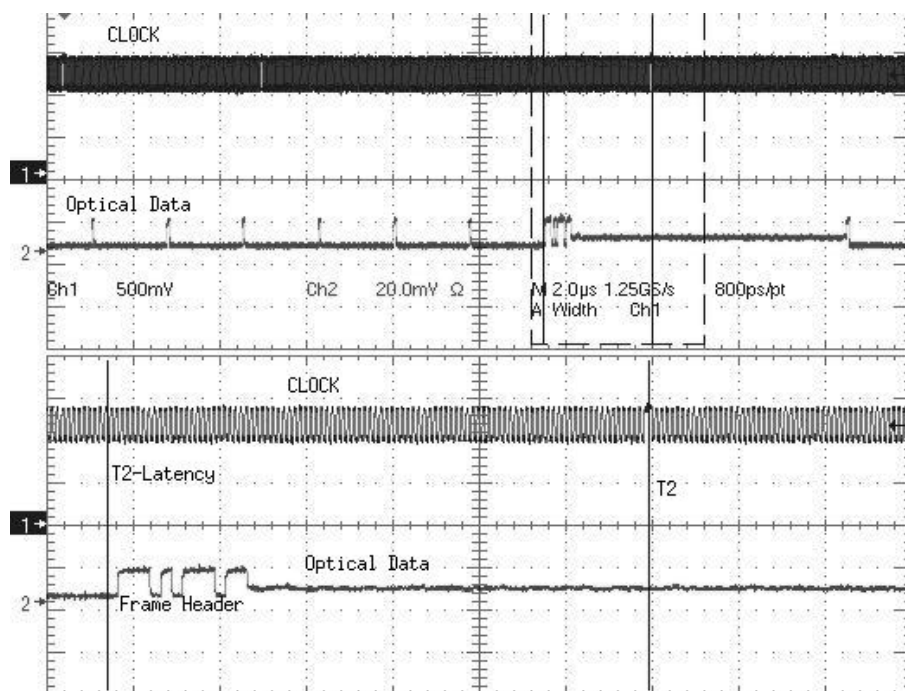

Figure 13: T2 aligned with the T1 Header. The top portion shows the clock trace and the data trace over a period of 100 μs, including several tickmarks spaced by 7 μs and one readout cycle starting roughly 56 μs into the trace The bottom portion is a close-up of the area in the 20 μs wide dashed box in the upper portion corresponding to the start of this readout cycle. The rightmost, solid vertical line indicates the position of T2. The leftmost marks a position that occurs 100 clocks earlier, the time at which data for T2 is sampled. T2 was adjusted such that this time would correspond to the beginning of T1 frame where the header information is sent, shown in the lower trace.

Figure 14 shows the ZS occupancy calculated from a short run taken with these trigger conditions. The figure shows low-numbered channels clearly most affected. The plot on the right shows a clear anti-correlation in the common-mode subtracted ADC counts of channels 0 and 127 from a longer VR run taken with the same trigger separation.

The T2 position was scanned in steps of 1 clock to produce trigger intervals between 362 clocks to 398 clocks. The ZS data acquired in this procedure show that the peaks of Figure 14 diminish and shift toward the middle of the chip. Figure 15 plots data taken at a separation of 370 clocks, corresponding to the readout of the pipeline address section of the APV header.

The effect begins to reappear on the high-numbered channels starting with a separation of 375 clocks. The effect is largest with a 384 clock separation, shown in Figure 16, when the sampling for T2 is close to the end of header readout. As in figure 16 there is a strong anti-correlation in the pedestal and common-mode subtracted ADC counts for channels 0 and 127 for this
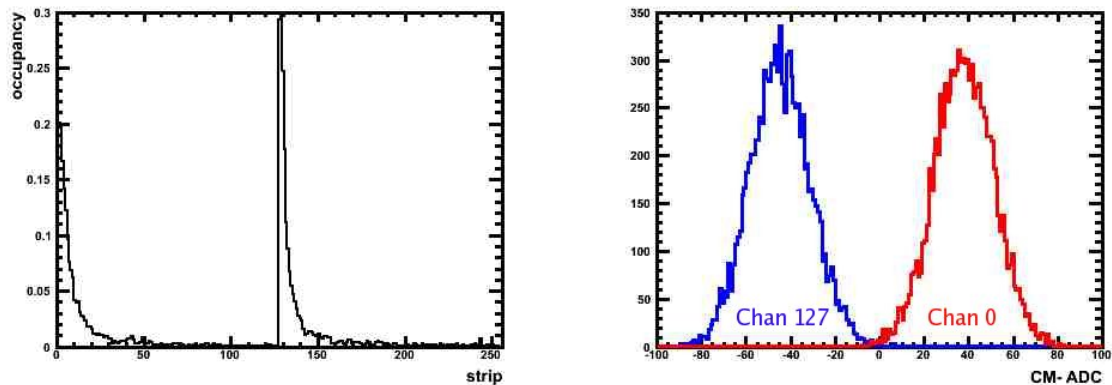
Figure 14: Data for a 362 clock trigger separation. Left: Occupancy spikes occur in the low-numbered channels when setting the trigger interval to 362 clocks, the separation needed to move T2's sampling point near the beginning of T1 readout. Right: Pedestal and common-mode subtracted ADC (channel 0 = red, channel 127= blue) from a run taken with a 362 clock trigger separation. Both channels show large displacements from zero and a per-event anti-correlation.
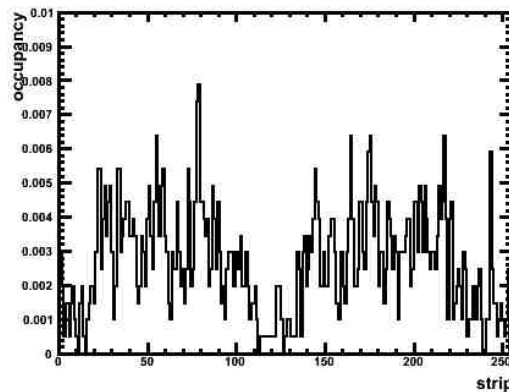

Figure 15: ZS Occupancy for a 370 clocks separation. Occupancy at this interval is greatly reduced from that measured with the 362 clock separation of figure 14). Moreover, the strip position of maximal occupancy has shifted from the low-numbered channels toward the middle of the chip.

separation, but with channel 127 now with positive displacement, as expected.

The power lines of the APV were probed with the same scope monitoring the optical data, looking for potential electrical pickup correlated with the HRN. There are small glitches associated with the reception of trigger T1 and the reset earlier, but these are brief and distant with respect to the time at which the T2 data are sampled, where there are no obvious problems on any of the power lines.
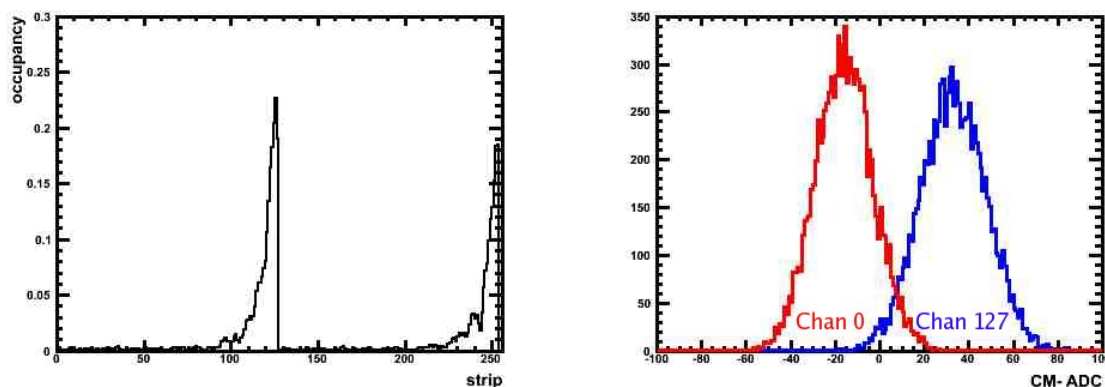
Figure 16: Data for a 384 clock separation. Left: After passing through a minimum, the occupancy spikes begin to reappear on the high-numbered channels as the trigger interval is increased. For a 384 clock separation, the effect is maximal. Right: Pedestal and common-mode subtracted ADC (channel 0 = red, channel 127= blue) from a run taken with a 384 clock trigger separation. Again, both channels show an anti-correlation and large displacements from zero, but this time in direction opposite to those of Figure 14.

3.2.6   APV simulation and Coupling Mechanism

The scope measurements of the power lines are inconclusive but APV simulations find a noise effect that is correlated with readout. Figure 17 graphs the simulated response of the *GND*, 1.25 and 2.5 V supply and analog output lines during the APV's readout sequence. The intermediate current spikes appearing on the supply lines are correlated with the falling edges of the three control signals for the APSP input phase and the reset signal to the APSP amplifiers, *ri1,ri2,ri3*, and *APSP_RST* in figure 9. The two largest spikes in each cycle come during the output stage and are correlated with the three control signals (*ro1/ro2/ro3*) of the APSP output phase. The positive-going spike on the +2.5V supply is caused by connecting the signal storage capacitor to the APSP output. The negative-going spike occurs when the internal sampling capacitors are disconnected from the APSP. The magnitudes of the largest spikes are below the level of sensitivity achievable with the scope probes used in the measurements above.

The supply-line current spikes that appear in the APV simulations are the likely origin of the HRN. Several studies investigated how this noise couples to the APV front end. First, a hybrid with no sensor, pitch adapter or wire-bonds was tested on the rod, creating a system similar to that which was used for production hybrid testing. The APVs in other modules in the rod remained connected to their sensors and acquired data using a 100 kHz Poisson trigger. Figure 18 shows the pedestal-subtracted, common-mode-subtracted ADC counts for a particular event from the two APVs on the bare hybrid, in which there is no HRN, consistent with the single-module measurement. For reference, Figure 18 also shows data from a fully-connected module on the rod for the same event, where it is clear the occupancy effect is present.

To further isolate how the HRN effect is introduced into the APV, several runs were taken where the connection between the hybrid and module was interrupted at different locations along the path for the APV channels 126/127 near the edge, where the HRN effect is maximal. Figure 19 shows the same quantity as figure 18, but now with the wire bonds between the pitch adapter and sensor detached at the sensor side, the bond between the pitch adapter and APV detached at the pitch adapter side, and the same bond completely removed, going left to
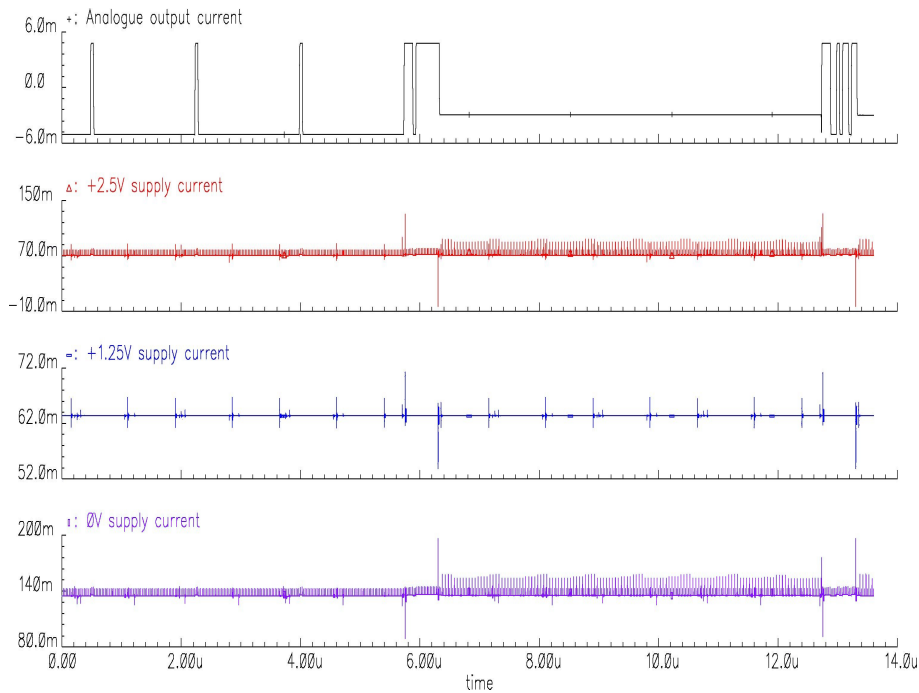
Figure 17: Simulation of APV chip behavior, showing short current spikes correlated with APV readout activity.

right.

The sharp signal drop when the bond between the pitch adapter and sensor is removed is likely due to a change in capacitance. The magnitude of the change is too large however for the effect seen with the connected chip to be entirely capacitive in nature. The sharp drop in ADC counts between channel 127 on one chip and channel 0 on the next suggests that it is unlikely that noise is coupling through the silicon sensor. If this drop is due to a smaller capacitance on the input, then the small change between the pitch adapter and the chip with a wire bond alone indicates that most of the signal is picked up by the wire bond. This in turn indicates that the likely coupling is between the power and ground wire bonds and the adjacent channels.

The leading hypothesis which emerges from these tests is that the HRN is a consequence of inductive coupling to the current spikes produced during the APV readout cycle. The physical coupling is between the power bonds and nearby signal bonds, thus affecting only a small subset of channels. The susceptibility of the channel to this coupling is greatly enhanced by the increased capacitance and therefore decreased impedance when the sensor is attached to the APV, while the pitch adapter and wire bond itself plays a less important role. The anticorrelation between the two channels observed is due to the direction of the current flow, which adds to the signal for one channel and subtracts from it for the other.
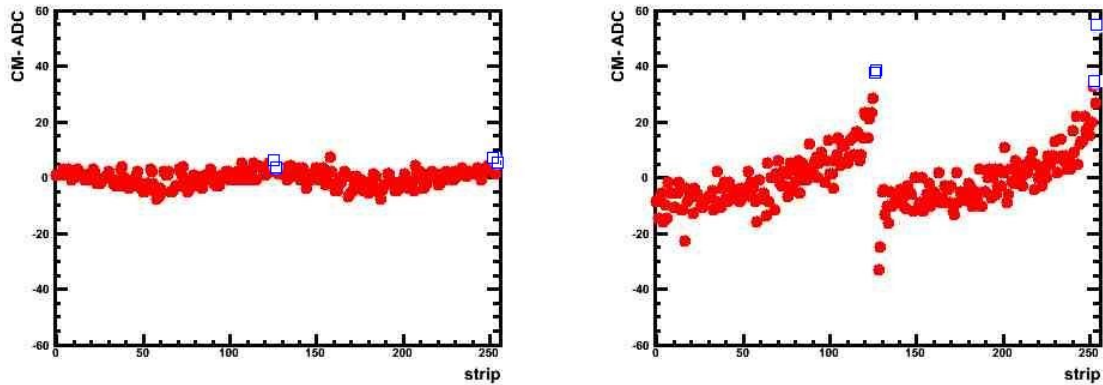
Figure 18: Bare Module and Reference. These show the pedestal and common-mode subtracted ADC counts for the two APVs (channels 0-127 and channels 128-255) of different modules but the same event. Left is the bare hybrid, right is another (connected) module on the rod. In both cases channels 126/127 on the two APVs are drawn in open blue squares. The HRN effect does not appear on a bare hybrid, but needs the module connected.
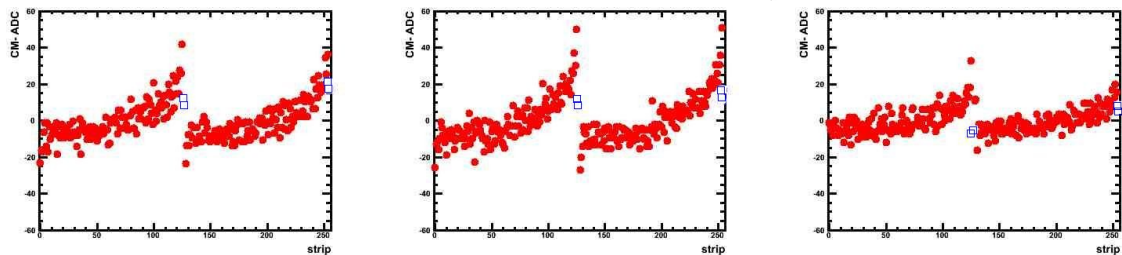

Figure 19: Broken Pitch Adapter – Sensor Connections. Theses plots are similar to those of figure 18 however here the wire bonds between the pitch-adapter and sensor have been detached or removed. Left: The wirebond is detached from the sensor for Channels 126/127, indicating that the sensor plays a large role in the noise coupling. Center: The wirebond is detached from the pitch adapter. Little change is seen relative to the left plot. Right: The wirebonds are completely removed for channels 126 and 127, while for channels 254 and 255 they are connected to the chip but not to the pitch-adapter. These channels still register some noise while the disconnected chip appears similar to the bare hybrid.

4   Consequences and Potential Mitigation

4.1   Consequences of the High Rate Noise Effect

Although the HRN effect has the potential to generate significant disparity in subdetector occupancies relative to those for which the CMS global DAQ was designed, the net increase in occupancy is negligible when integrating over all channels on a chip. The FED and event building systems were designed with specific maximum subdetector occupancies in mind ($\sim 3\%$ for the Tracker). The strip occupancy generated by high-rate noise is of the same order, $\sim 1\%$, and could consume a large portion of the bandwidth allocated for the Tracker if the effect extended across all APV channels. However, the magnetic field that couples supply line noise to the APV front-end falls off quickly and only a few channels are affected on average.

The impact of HRN on track reconstruction is more serious. A simple simulation in which the ADC values of edge-strips were fluctuated to emulate the addition of HRN into TIF cosmic data showed that the reconstruction of a single, fluctuated event requires 15 minutes to complete when using default offline clusterization thresholds. The algorithm identifies $\sim 30,000$ phantom track segments. Reasonable event processing times were only achievable by increasing thresholds to unrealistic levels.

4.2   Mitigation of the High Rate Noise Effect

While at this point there is no way to eliminate the high rate noise effect, there are several possible solutions to mitigate the consequences. One possibility is to exploit the universal character of HRN events to distinguish them from normal data offline. Once identified, HRN events can be individually processed before track reconstruction. Most likely "processing" will entail the rejection of such events given that stripping HRN clusters from data would be difficult and its impact on track reconstruction efficiency hard to quantify. HRN event rejection would require a dedicated software filter to run during reprocessing or as part of any job that performs track reconstruction.

In a variation on this proposal, one could separate the identification of HRN events from their rejection by implementing a recognition algorithm in the FED firmware. The algorithm could access raw data input to the FED and could therefore utilize the degree of anti-correlation between edge channels in its determination. This might provide for more accurate HRN identification than can be achieved by methods relying on the presence of time-correlated, edge clusters alone. The FED could communicate its findings to the offline environment by setting a header bit in its output data, for example. The offline filter would then have the simple task of checking this bit to determine whether an event should be rejected.

Both solutions are safe in that HRN events are still written, unaltered, to raw datasets *offline*. Other possibilities include the rejection or correction of HRN data *on-line*, before they are written to disk. For example, a specialized FED pedestal/common-mode subtraction algorithm could be implemented to apply when HRN events are identified. This approach would allow for the removal of HRN clusters on-line so that HRN events can be input to normal reconstruction processes. HRN cluster removal would inevitably impact signal too however and information on the amount of charge subtracted from signal clusters would not be available offline. Furthermore, it would be difficult to fully understand the impact of such a procedure on reconstruction efficiency.

The most straightforward way of addressing HRN involves preemptive rejection of these events on-line. The FED identification and error bit designation scheme described above could be used to inform the event-building system that such events should be discarded. Rejection would oc-

cur in the BU because fragments from all FEDs would have to be discarded in a given HRN event. On the other hand, event building applications are efficient in part because they do not access subdetector payloads. The impact of modifications to the BU to allow access to FED error bits would need to be small for this approach to be feasible.

Another on-line approach would prevent pathological trigger intervals using the APVE. The APVE receives both trigger and reset signals and can use these to determine the bunch crossings in which data readout will occur. By programming the APVE to reject triggers that occur at specific times during the APSP cycle, but only when a previous trigger has been applied, it should be possible to efficiently prevent triggers which would give rise to a high-rate noise event. This approach requires non-trivial extensions to the APVE firmware however and will result in Tracker-imposed dead-time in the trigger. The cost of this dead-time to the experiment can be equivalent to the offline rejection of HRN events. The firmware modifications for this approach are currently being implemented and tested.

To summarize, the tests with the Column-DAQ reveal a noise effect that appears on edge strips of all Tracker APVs when running with a high-rate, Poisson random trigger. This triggering scheme is very similar to what can be expected during normal operation in CMS. As the single-module studies clearly show, the effect results from particular trigger intervals in the Poisson distribution that align a trigger with the readout of the frame of a previous trigger. Simulations indicate that current spikes correlated with the readout of data from the first trigger magnetically couple to APV inputs. This leads to elevated ADC counts on edge channels in data queued by the second trigger. The effect occurs simultaneously on every APV because the chips are synchronized during data taking. This effect poses serious problems for reconstruction, and several possible mitigation schemes are currently under investigation.

5   Acknowledgments

References

[1] **CMS** Collaboration, R. Adolphi et al., "The CMS experiment at the CERN LHC," *JINST* **0803** (2008) S08004. `doi:10.1088/1748-0221/3/08/S08004`.

[2] **CDF** Collaboration, S. Nahn, "Status of the CDF Run II silicon detector," *Nucl. Instrum. Meth.* **A511** (2003) 20–23.

[3] **D0** Collaboration, M. S. Weber, "The D0 silicon micro-strip tracker," *Nucl. Instrum. Meth.* **A560** (2006) 14–17.

[4] **CMS Tracker** Collaboration, W. Adam et al., "Performance studies of the CMS Strip Tracker before installation," *JINST* **4** (2009) P06009, `arXiv:0901.4316`. `doi:10.1088/1748-0221/4/06/P06009`.

[5] **CMS Tracker** Collaboration, C. Noeding and W. Adam, "Tracker Reconstruction with Cosmic Ray Data at the Tracker Integration Facility," *CMS Note* **2009/003** (2009).

[6] **CMS Tracker** Collaboration, A. Gritsan and G. Flucke, "CMS Tracker Alignment at the Tracker Integration Facility," *CMS Note* **2009/002** (2009).

[7] **CMS** Collaboration, "CMS, tracker technical design report,". CERN-LHCC-98-06.

[8] **CMS** Collaboration, "Addendum to the CMS tracker TDR,". CERN-LHCC-2000-016.

[9] **CMS** Collaboration, e. . Sphicas, P., "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger,". CERN-LHCC-2002-026.

[10] G. Iles, W. Cameron, C. Foudas, G. Hall, and N. Marinelli, "The APVE emulator to prevent front-end buffer overflows within the CMS Silicon Strip Tracker,". Prepared for 8th Workshop on Electronics for LHC Experiments, Colmar, France, 9-13 Sep 2002.

[11] **CMS Tracker** Collaboration, K. Bell et al., "User Requirements for the Final FED of the CMS Silicon Strip Tracker," *CMS Note* **2001/043** (2001).

[12] M. J. French et al., "Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker," *Nucl. Instrum. Meth.* **A466** (2001) 359–365. `doi:10.1016/S0168-9002(01)00589-7`.

[13] J. Lamb, "CMS silicon strip tracker integration and slice test," *Nucl. Phys. Proc. Suppl.* **177-178** (2008) 300–301. `doi:10.1016/j.nuclphysbps.2007.11.135`.

[14] N. Bingefors et al., "A Novel technique for fast pulse shaping using a slow amplifier at LHC," *Nucl. Instrum. Meth.* **A326** (1993) 112–119. `doi:10.1016/0168-9002(93)90340-N`.

[15] S. Gadomski et al., "The Deconvolution method of fast pulse shaping at hadron colliders," *Nucl. Instrum. Meth.* **A320** (1992) 217–227. `doi:10.1016/0168-9002(92)90779-4`.