Phase-2 Upgrade of the CMS Tracker

Stefano Mersi on behalf of the CMS Collaboration

stefano.mersi@cern.ch

Abstract

An upgrade program is planned for the LHC which will smoothly bring the luminosity up to or above $5 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ sometimes after 2020, to possibly reach an integrated luminosity of 3000 fb$^{-1}$ at the end of that decade. In this ultimate scenario, called Phase-2, when LHC will reach the High Luminosity phase (HL-LHC), CMS will need a completely new Tracker detector, in order to fully exploit the highly-demanding operating conditions and the delivered luminosity. The new Tracker should have also trigger capabilities. To achieve such goals, R&D activities are ongoing to explore options and develop solutions that would allow including tracking information at Level-1. The design choices for the CMS pixel and outer tracker upgrades are discussed along with some highlights of the R&D activities and expected detector performance.

Keywords: CMS, Tracker, Upgrade, Front-end, HL-LHC

1. Introduction

The current planning for the LHC and injector chain foresees a series of long shutdowns, designated LS1, LS2, LS3. Between LS1 and LS2, when the total integrated luminosity is assumed to reach about 150 fb$^{-1}$ (2016-17), the pixel detector will be replaced with the “Phase-1” upgrade [1].

Following LS3 (2023-2025) the high luminosity program with the upgraded LHC is referred to as High-Luminosity LHC (HL-LHC) or “Phase-2”. The proposed operating scenario is to level the instantaneous luminosity at $5 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ at the beginning of fills, and to deliver 250 fb$^{-1}$ per year for a further 10 years of operation.

Under these conditions the number of collisions per bunch-crossing, or Pile-Up (PU), is expected to reach or exceed 140 and the total integrated radiation dose is expected to increase by a factor of about ten with respect to the initial LHC design value. In order to maintain or improve the physics performance of the CMS detector, the entire tracking system must be replaced with new detectors featuring higher radiation tolerance and enhanced functionality$^1$. The present Outer Tracker and the Phase-1 Pixel detector would be unable to correctly operate at the HL-LHC: according to simulations [2, 3] the Outer Tracker would suffer thermal runaway of a large fraction of its sensors leading to a significant degradation of track reconstruction performance, as shown in Figure 1, with a large impact on the physics reach of CMS. On top of that the Pixel detector would significantly degrade its resolution after just 500 fb$^{-1}$.

The main requirements for the Tracker Upgrade can be summarized as follows: radiation tolerance up to an integrated luminosity of 3000 fb$^{-1}$, which translates to a maximum particle fluence of $1.5 \times 10^{15} \text{neq cm}^{-2}$ in the Outer Tracker region and $2 \times 10^{16} \text{neq cm}^{-2}$ in the pixel volume; increased granularity to maintain the channel occupancy near or below the 1% level in all tracker regions; improved two-track separation, to overcome current limitations in track finding performance in high-energy jets, and exploit the statistics of HL-LHC; reduced material in the tracking volume, which is cur-

$^1$The present Outer Tracker was designed to operate without any loss of efficiency up to an integrated luminosity of 500 fb$^{-1}$, and PU $\leq 50$ collisions per bunch crossing.
Figure 1: Efficiency for the Phase-1 detector before and after the Outer Tracker has been aged by an equivalent of 1000 fb\(^{-1}\) for charged particles from \(\bar{t}t\) events for which the particles have \(p_T > 0.9\) GeV and are produced within 3.5 cm of the interaction region (in the transverse direction).

Currently a significant limiting factor for the CMS tracking and calorimetry performance; extended coverage of the tracker (and calorimeters) in the forward region, which will benefit the overall CMS physics capabilities; compliance with the Level-1 (L1) trigger upgrade (increase of L1 rate and latency to 750 kHz and 12.5 \(\mu\)s) and adding tracking information to the Level-1 trigger decision; the layout of the upgraded Tracker should enable fast and efficient track finding, which is particularly important for the high level trigger (HLT) at high Pile-Up.

The upgraded Tracker detector will be composed of two sub-detectors: Pixel vertex detector occupying the region\(^2\) \(r < 20\) cm and Outer Tracker in the region \(20\) cm < \(r < 110\) cm as the present one, as shown in Figure 2.

2. Vertex detector: pixels

The design of the detector will preserve the ease of access of the current detector that enables the possibility to replace degraded parts over an Extended Technical Stop, as some of the detector components might not survive the full radiation dose provided by the HL-LHC.

The geometry of the Phase-1 detector [1] with 4 barrel layers and 3 forward disks is taken as a starting point. The forward extension could be most simply realized by increasing the number of forward disks from 3 to 10, out of which the last 3 consist of the outer part only, in order to be compatible with the conical section of the beam pipe. Such an extended pixel detector will have an active surface of approximately 4 m\(^2\), compared to 2.7 m\(^2\) for the Phase-1 detector.

The target integrated luminosity of 3000 fb\(^{-1}\) corresponds to a hadron fluence of \(2 \times 10^{16}\) n\(_{eq}\) cm\(^{-2}\) and 10 MGy at 3 cm from the interaction region, roughly where the first layer of the Pixel Detector will be located as is the case for the Phase-1 detector. The fluence decreases rapidly with distance and is about \(3 \times 10^{15}\) n\(_{eq}\) cm\(^{-2}\) at a radius of 11 cm.

Preliminary studies show that good results can be obtained by using thin planar silicon sensors, segmented into very small pixels. With such a configuration the detector resolution is much more robust with respect to radiation damage than the present detector, where the precision relies on the ability to reconstruct the tails of the charge deposited in a 300 \(\mu\)m thick sensor. At the same time the required improvement in two-track separation mentioned above is also obtained. Pixel sizes of \(25 \times 100\) \(\mu\)m\(^2\) or \(50 \times 50\) \(\mu\)m\(^2\) are being considered, representing a factor of 6 reduction in surface as compared to the present pixel cells.

The n-in-n planar technology, developed for the current pixel detector and its Phase-1 replacement, is a valid baseline for instrumenting the outer regions, as the higher mobility of electrons compared to holes leads to a higher radiation tolerance.

Adoption of n-in-p sensors could reduce the cost with respect to n-in-n sensors, provided the design can protect from discharge due to the bias voltage being present at the sensor edge.
Thin sensors (150 μm or less) offer advantages in terms of lower bias voltage and lower leakage current. Moreover, the shorter drift distance results in smaller clusters which, when combined with a smaller pixel cell and a reduced signal threshold for the smaller produced charge, can achieve good resolution and improved two-track separation in high-energy jets. The enhanced radiation tolerance should make it possible to extend the use of planar sensors to small radii.

An alternative actively pursued option is the possibility to use 3D silicon sensors, offering intrinsically higher radiation resistance because of the shorter charge collection distance [4]. As the production process is more expensive and so not suitable for large volumes, the use of 3D sensors could be limited to small regions of highest particle fluence.

Compared to the Phase-1 implementation, the Phase-2 ROC will feature six times smaller pixels, will have to cope with about five times higher hit rates, five or ten times higher trigger rates, as well as longer trigger latency. To address the challenges associated with the anticipated hit density and corresponding needs for I/O and buffering capabilities, the cross-experiment collaboration RD53 [5] has been formed, today comprising 19 institutes with a balanced participation between ATLAS and CMS.

The results from initial radiation tests on 65 nm technology test structures are promising. However, some significant radiation effects have been observed for total dose levels above a level of 2 MGy [6], which require additional studies. In particular it is important to understand and quantify the interplay of operating temperature (nominally −20°C) and possible periods of annealing at room temperature.

A common pixel chip architecture has been defined to be fully digital after the basic threshold detection and charge digitization in the analogue pixel cell. Buffering requirements have been analysed with a statistical model [7] and simulations and have been verified to be compatible with the proposed extended CMS trigger latency of 12.5 μs. A buffer depth of 16 pixel clusters for a 4 × 4 pixel region is sufficient to guarantee a hit loss probability below 10−3 for the highest hit rate of 2 GHz/cm².

With an appropriate bump-bonding pattern the same ROC can be used for the two different pixel aspect ratios, and also for larger pixels with compatible dimensions by disabling the unused channels. A second generation of the radiation tolerant GBT link chip [8], the LP-GBT [9], will transfer the signal to high speed optical links at 10 Gb/s. The LP-GBT chip and associated laser and driver are not expected to have sufficient radiation tolerance to be used in the central part of the pixel detector. They will be located on the support cylinders at the outer boundary of the detector volume, in the Phase-1 detector. Electrical links with lengths up to 2 m will connect the modules to the LP-GBTs. For the highest expected hit rates in the first barrel layer a readout bandwidth of up to about 4 Gb/s per ROC may be required.

Special care must be placed in the design and routing of the services in order to avoid large material budget in the pixel region, due to the large needed bandwidth and potentially large power requirements.

3. Outer Tracker: strips and macro-pixels

The Outer Tracker will provide data both for the L1 reconstruction (for each bunch crossing), and for the global event processing upon reception of a L1 trigger decision. To ease the design, production and operation of the detector, only two module versions are conceived, featuring the same sensor material. Detailed description of the electronics were presented in a dedicated forum [9].

3.1. Sensors

The target integrated luminosity of 3000 fb⁻¹ corresponds to a hadron fluence of 1.5 × 10¹⁵ nₑq cm⁻² at 20 cm from the interaction region.

In order to achieve the required radiation tolerance it is critical to choose appropriate sensor material and processing technology. A comprehensive program has been carried out [10–27] to identify suitable silicon materials and to define the requirements on the operating temperature, using test wafers from a single producer (Hamamatsu Photonics K.K.) with Float-Zone, Magnetic Czochralski and Epitaxial substrate, both p-on-n and n-on-p and with active thickness of 50 μm to 300 μm.

The results collected so far show that sensors with electron read-out are more robust in terms of high field effects after irradiation, and also provide higher Charge Collection (CC) than p-in-n sensors. The p-bulk materials investigated show comparable CC at the highest expected fluence. In addition, 200 μm sensors show a better CC stability for annealing above 20 weeks at room temperature, which is beneficial for operations.

Collecting all the results lead to a clear preference for using 200 μm active thickness n-in-p sensors targeting a sensor operating temperature of −20°C.
3.2. Module design

In order to provide tracking information at Level-1, hits must be shipped off-detector at the collision rate of 40 MHz. Sending all the hit data would require a very large bandwidth that is clearly unaffordable in terms of power and material budget.

The goal is achieved with modules that are themselves capable of rejecting signals from particles below a certain $p_T$ threshold: “$p_T$ modules” [28]. A threshold of around 2 GeV corresponds to a data volume reduction of roughly one order of magnitude, which is sufficient for the purpose of L1 data transmission.

Two versions of the $p_T$ modules are devised (called “2S” and “PS”) both of which will be used in the endcap disks as well as in the barrel region.

The modules are composed of two closely-spaced silicon sensors read out by a common front-end placed at the edges of the sensors, capable of correlating the signals collected in the two sensors. For each hit on the inner sensor an acceptance window is defined on the outer and matching pairs of hits form “stubs”, compatible with particles above the chosen $p_T$ threshold.

The strong magnetic field of CMS provides sufficient sensitivity to measure $p_T$ over the small sensor separation, enabling the use of $p_T$ modules in the entire radial range above $R \approx 20$ cm. Stub data are sent out at every bunch crossing, while all other signals are stored in the front-end pipelines to be read out when a trigger is received.

To achieve efficient rejection of low-$p_T$ particles throughout the Tracker volume the acceptance window must be programmable in the front-end and modules in different regions make use of a few different sensor spacings. For 2S (PS) modules, spacings of 1.8 and 4 mm (1.6, 2.6 and 4 mm) are foreseen.

The design of modules should provide for efficient removal of the heat generated by the electronics and sensors, accurate geometrical positioning, minimal mass, as well as a simple and reproducible assembly procedure. For the thermal performance, the design requirement is to achieve a sensor temperature of $-20^\circ$C or lower with a coolant temperature of $-30^\circ$C for modules irradiated with the full HL-LHC integrated luminosity. To remove heat from electronics and sensors, CO$_2$ two-phase cooling will be used. Such choice of cooling technology helps to reduce the amount of passive material in the tracking volume.

The development of the designs has been guided by detailed Finite Element Analysis of their thermal performance. The concepts chosen enable different sensor spacings to be incorporated with minimal or no impact on the design of the active components.

2S modules are composed by two superimposed strip sensors of approximately $10 \times 10$ cm$^2$, mounted with the strips parallel to each other and segmented in strips of size $90 \mu$m $\times$ 5 cm. They populate the outer regions, above $R \approx 60$ cm (in red in the sketch of Figure 2), which entails approximately 150 m$^2$ of sensing area. Wire bonds at opposite ends of the sensor provide the connectivity of both sensors to the readout hybrid. A single “service hybrid” carries a 5 Gb/s data link, an optical converter, and the DC/DC converter that provides power to the module electronics. The use of one optical link per module provides the bandwidth needed for the trigger functionality, and at the same time offers significant advantages in the overall system design by avoiding additional electrical interconnectivity in the tracking volume.

PS modules are composed of two sensors of approximately $5 \times 10$ cm$^2$, one segmented in strips of size $100 \mu$m $\times$ 2.3 cm and the other segmented in “macro-pixels” of size $100 \mu$m $\times$ 1.4 mm. The chosen pixel size permits the use of the “C4” bump-bonding technology, an industrial process that is expected to be affordable for a large-scale production. As for the 2S module, wire bonds provide the connections from the strip sensor and from the macro-pixel readout chip to the front-end hybrid, and, in turn, to the auxiliary electronics for powering and readout, all of which is integrated in the module assembly. PS modules are deployed in the radial range between $R \approx 20$ cm and $R \approx 60$ cm (blue in the sketch of Figure 2), resulting in a sensor surface area of about 60 m$^2$ (30 m$^2$ short strip sensors and 30 m$^2$ macro-pixel sensors). The macro-pixel sensors provide sufficiently precise measurements of the $z$ coordinate for tracking to enable primary vertex discrimination at L1. At the same time, three additional layers of unambiguous 3D coordinates, each with an associated estimate of the particle $p_T$, are of particular use for track finding.

The development of electronics for the 2S module has already completed several prototype cycles for the front-end chip (CBC) [29, 30] and hybrid circuits [31], to the point that working modules could be assembled and subject to a beam test [32]. This test demonstrated the $p_T$ discrimination front-end capability (as shown in Figure 4, bottom), the assembly technology and the full readout chain based on the GLIB multi-purpose AMC board [33].

4. Layout and expected performance

The layout shown in Figure 2 has been adopted as baseline design. The Outer Tracker layout has been the subject of extensive studies and detailed modeling,
both using full Monte-Carlo simulations and a dedicated software package [34], exploring several variants, including geometries with barrels only, and geometries with different numbers of barrel layers, and/or different numbers and size of end-cap disks. The pixel layout is currently under development, but it is assumed to have an extended forward region, to achieve an extension of the tracking acceptance up to \(|\eta| \approx 4\). Pixel modules are dedicated to the accurate measurement of impact parameter and coverage of tracking in the forward region, both of which will be available only in the full event reconstruction, at High-Level trigger and beyond.

The Phase-2 Tracker offers a significantly extended tracking acceptance compared to the Phase-1 Tracker, with slightly fewer hits per track over the common \(\eta\) range and a coverage up to \(\eta \approx 4\) (Figure 3, top). The inactive material inside the tracking volume is substantially reduced, as shown in the bottom plot of Figure 3, where the distribution of radiation length versus \(\eta\) for the Phase-1 Tracker is compared to that expected for the current model of the Phase-2 Tracker. A summary comparison between the current and Phase-2 tracker is given in Table 1.

Enabling trigger functionality at L1 requires an additional step in the back-end processing beyond what is needed in the present system. At each bunch crossing the “stub data” are processed to form “L1 tracks”. The L1 tracks are tracking primitives that are combined with information from the other sub-detectors to form L1 triggers. The cabling of the detector and the overall architecture of the back-end system must be optimized for efficient track finding with an affordable amount of data traffic.

The performance of the stub finding capability of \(p_T\) modules is characterized in terms of the efficiency achieved for particles above the chosen threshold, and the rate reduction obtained. In Figure 4 (top) the data reduction rate is shown for the barrel region (a similar performance is achieved in the end-cap region).

The reduction rate drops below ten at the largest radii because the rate of hits from low-\(p_T\) tracks decreases faster than the total rate (due to the CMS magnetic field), which poses no problem in terms of bandwidth. The reduction rate also drops for the inner layers because the selection purity is affected by combinatorial background and secondary particles producing a larger fraction of “fake” stubs.

The resulting stub rate fits well within the specified bandwidth for all layers, except for the one closest to the interaction point, where data volumes are near to the limit. Reducing the \(p_T\) resolution or increasing the \(p_T\) threshold are available methods to reduce the data volume.

Simulations of the resulting stub finding efficiency for muons demonstrated that the efficiency plateau is

<table>
<thead>
<tr>
<th>Current</th>
<th>Phase-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silicon [m²]</td>
<td>200</td>
</tr>
<tr>
<td>Strips</td>
<td>(9 \times 10^6)</td>
</tr>
<tr>
<td>MacroPixel</td>
<td>—</td>
</tr>
<tr>
<td>Modules</td>
<td>15 148</td>
</tr>
<tr>
<td>Readout rate [kHz]</td>
<td>100</td>
</tr>
<tr>
<td>Stub readout rate [kHz]</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1: Size of the current and Phase-2 upgrade tracker according to current design. The final size of the pixel detector is not fully defined yet.
reached at 2 GeV in all the tracker detecting layers, with the appropriate tuning of the front-end parameters.

As mentioned above, the working principle of $p_T$-modules has been also demonstrated in a beam test at DESY, using 2S module prototypes. Different values of the local curvature of tracks due to different $p_T$ have been emulated by changing the incident angle between the 2 GeV positron beam and the prototype $p_T$-module under test. The plot shown in Figure 4 (bottom) has been obtained by stepping through several incident angles and translating the value of the angle to the equivalent $p_T$ at a radius of 75 cm. The resulting efficiency curve corresponds to a measured effective $p_T$ threshold of $2.2 \pm 0.1$ GeV, which compares well with the nominal threshold of 2.14 GeV, as calculated from the module parameters (sensor spacing, acceptance window and strip pitch).

Each bunch crossing produces on the order of 10,000 stubs. Only about 5 to 10% of these stubs actually belong to primary tracks with $p_T > 2$ GeV (approximately 125 per bunch crossing). The remainder are random combinations of hits, or are due to lower-$p_T$ particles and secondary particles produced by interactions of primary particles with the material in the detector. The goal of the L1 Track Finding system is to perform pattern recognition to reconstruct the tracks of primary particles with $p_T > 2$ GeV, and discard as many as possible of all the other stubs.

Several approaches are currently under consideration to process stub data in the back-end electronics and construct tracks for the Level-1 trigger, with the pattern recognition problem being addressed either through conventional track finding methods (and taking advantage of the close spacing of the measurement points to reduce the combinatorics) or by making use of Associative Memories.

The performance of these methods on single particles is evaluated from samples of single muons, pions, and electrons, uniformly distributed in $\phi$, $\eta$ and $p_T$, overlaid with an average of 140 PU events. The track finding efficiency is shown in Figure 5 (top) for one of the algorithms (called “tracklet”), where the efficiency is defined with respect to truth-level tracks and corresponds...
to loose track quality criteria for muons/pions and electrons. Muons have an overall efficiency of about 99%, while pions and electrons have lower efficiency due to interaction with detector material. It is expected that those figures (notably for electrons) can be further improved.

The pattern recognition performance has also been verified in even more dense environments, namely events with four high energy top jets overlaid with 140 pile-up events, and single particles overlaid with 200 pile-up events. In both cases the efficiency was found to remain approximately the same, demonstrating the robustness of the tracking.

The L1 track $z_0$ and $p_T$ resolutions are shown for single muons in Figure 5 (middle and bottom, respectively) as function of $|\eta|$ for three ranges of $p_T$. The momentum resolution decreases as expected beyond $|\eta| > 1.5$ due to the lower lever arm of the Tracker. A $z_0$ resolution of about 1 mm is achieved in the central region, similar to the average separation of pile-up vertices.

Associative Memory-based algorithms were also tested with the same method, leading to similar performance. The actual implementation to be chosen will thus depend mainly on technological relative advantages (like cost, level of processing parallelism, etc.).

5. Conclusions

The HL-LHC upgrade imposes extreme conditions on the CMS experiment in terms of radiation tolerance, pile-up and occupancy, to the point that the entire tracking system will have to be replaced during Long Shutdown 3 (Phase-2 Tracker). A comparison with the current tracker is shown in Table 1.

It was shown that the experiment would greatly benefit from having tracking information available for the Level-1 selection to maintain or improve its physics reach while keeping the trigger rates under control. This poses an additional challenge for the Phase-2 Outer Tracker: sending data for each bunch crossing to the back-end and performing on-line track reconstruction with a resolution good enough to perform a track-based event selection and telling apart most of the tracks coming from different pile-up vertices, which was proven to be achievable with the current design of $p_T$ modules.

The main challenge for the pixel detector design under way is the combination of high integrated radiation dose (up to 10 M Gy) with hit rates up to 2 GHz/cm², which makes the readout scheme and distribution of power difficult to achieve without compromising the material budget.

References

[22] M. Draganovic, The New Silicon Strip Detectors for


G. Hall, M. Raymond and A. Rose, 2-D PT module concept for the SLHC CMS tracker, JINST 5 C07012, 2010.


[33] P. Vichoudis et al., First results with the Gigabit Link Interface Board (GLIB), JINST 6, C12060, 2011.

[34] G. Bianchi et al., tlLayout: a design tool for innovative silicon tracking detectors, JINST 9, C03054, 2014.