Development of the CMS Phase-2 Outer Tracker Analyzer of Test Outputs software

Matteo Marchisio Caprioglio[1]

Abstract
This report focuses on describing my work in optimizing the Phase-2 Outer Tracker Analyzer of Test Outputs software done during my 2024 summer internship at Fermi National Accelerator Laboratory. The main goal of this optimization consists in the complete re-implementation of the analysis and grading procedures adopted for the silicon detector modules that make up the Phase-2 Outer Tracker Upgrade, by using an Object Oriented Programming (OOP) approach to make the software modular and expandable, and new sections easy to add for current and new developers.

[1] *Università degli Studi di Torino*

Contents

List of Figures

1. The CMS Experiment

1.1 Overview

The Compact Muon Solenoid (CMS) experiment is a multi-purpose detector [2] built at the Large Hadron Collider (LHC) to study head-on collisions of protons (Pb ions) up to a center-of-mass energy of 14 TeV (5.5 TeV per nucleon) to refine our understanding of Standard Model (SM) physics and to investigate Beyond Standard Model (BSM) theories.

The detector is made of several layers (1), the specific functions of each are briefly described as follows:

- **Tracker**: the innermost section of the detector, it is made of several silicon subdetectors (either pixel detectors, used for the Inner Tracker, or microstrip detectors, used for the Outer Tracker) which serve the purpose of reconstructing the paths of charged particles generated in the p-p/Pb-Pb collisions and identifying the position of the interaction vertexes.

- **Electromagnetic Calorimeter (ECAL)**: it is a homogeneous calorimeter made of several $PbWO_4$ scintillating crystals, read by Avalanche PhotoDiodes (APDs) in the barrel region or by Vacuum PhotoTriodes (VPTs) in the endcap region, and it is used to measure the energy of emitted photons or electrons via electromagnetic showering.

- **Hadron Calorimeter (HCAL)**: it is a sampling calorimeter made of alternating layers of brass and plastic scintillator, and it serves the purpose of measuring via hadronic showering the energy of the hadronic component of the particles produced in each interaction.

- **Superconducting Solenoid**: it contains all the layers described above and it is capable of generating a magnetic field of up to 3.8 T, under the action of which charged particles curve: by tracking their curved trajectories one can obtain a momentum measurement.

- **Muon Spectrometer**: it is collocated outside the solenoid (since muons are the most penetrating charged component of all the particles produced in a collision) and it is used to identify muons and measure their momentum; it is made of different types of gaseous detectors, depending on their collocation with respect to the beamline (Drift Tube Chambers in the barrel region, Cathode Strip Chambers in the endcap region, and Resistive Plate Chambers in both regions [4]).

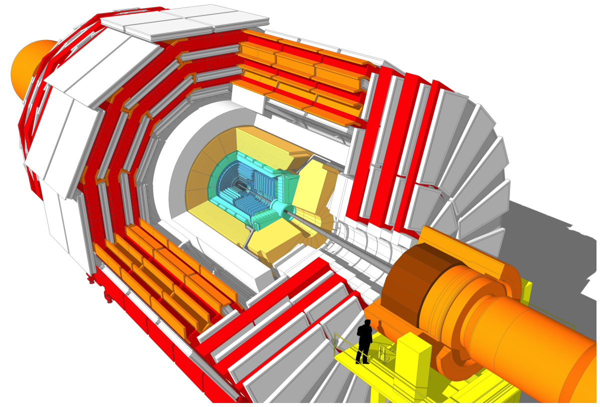

Figure 1. Cutaway view of the CMS detector layers

1.2 Trigger System

The current LHC bunch crossing rate is 40 MHz and this results in a $\sim$ GHz interaction rate during p-p collisions, while the CMS detector can send data to offline storage at an average rate of 400 Hz [3], so a two level trigger system is used to prune the bulk of inelastic collision events from the "interesting" ones:

- **Level 1 Trigger (L1)**: it is a hardware system with a fixed 4 µs latency, within which it has to decide whether to keep or discard an event based on the information from both the calorimeters and the muon detectors.

- **High Level Trigger (HLT)**: it is a software system that does a preliminary reconstruction similar to the one done in offline analysis, to determine if an event is sufficiently interesting to be further stored.

As of now **no input is provided to the L1 by the silicon tracker**.

2. The High-Luminosity LHC Upgrade

During the Long Shutdown 3 (planned for the late 2020s), the LHC will be subject to a series of upgrades which will increase its instantaneous luminosity up to $7.5 \times 10^{34}\,cm^{-2}\,s^{-1}$ ($\sim$fivefold with respect to its current value), with the goal of accumulating up to $3000 - 4000\,fb^{-1}$ per year [1].

This upgrade comes with several new operating conditions, which include (but are not limited to) the following:

- **High radiation levels**: fluences of up to $10^{16}\,n_{eq}$ are expected near the interaction region, and this requires higher levels of radiation hardness with respect to the current setup, especially for the innermost parts of each LHC experiment.

- **Pileup of many interactions**: with the increased luminosity, up to 200 p-p collisions per bunch crossing are expected, and this requires a higher granularity for each detector to maintain vertex reconstruction power.

- **High data throughput**: the increased number of interactions results in a higher data rate, and updated trigger systems are required for data reduction.

All the LHC experiments will thus need to be upgraded to comply with these new requirements. The CMS experiment will undergo a series of Phase-2 upgrades, among which the most relevant for this report is the Phase-2 Outer Tracker Upgrade.

3. The CMS Phase-2 Outer Tracker Upgrade

3.1 Overview

This upgrade, which is part of the overall Phase-2 Upgrade of the CMS experiment, is aimed at replacing the current Outer Tracker (OT) with a new system compliant with the new operating conditions of the High-Luminosity LHC. The main characteristics of the new OT are the following [7]:

- **Increased radiation resistance**: the overall radiation hardness of the detector is improved to withstand the harsh fluences of the HL-LHC environment.

- **Augmented granularity**: the usage of silicon pixels improves the overall granularity of the detector, a key requirement in the increased pileup conditions of the HL-LHC.

- **Reduced material budget**

- **L1 trigger input**: the new OT will provide information to the L1 trigger decision, thus mitigating the new increased pileup conditions.

The composition of the new OT and its principle of operation are briefly described in the following section.

3.2 OT modules

The new OT will comprise two types of modules made of two closely spaced silicon detectors connected to the same readout electronics and collocated with respect to the beamline according to their granularity (figure 2):

- **2S (strip-strip) modules**: they are made of two silicon strip detectors with strips parallel to each other, and they are positioned in the outer radial region of the OT (60 to 120 cm from the beamline, figure 3).

- **PS (pixel-strip) modules**: they are made of a strip detector and a macro-pixel array, and due to their better granularity with respect to the 2S modules they are positioned in the inner radial region of the OT (20 to 60 cm from the beamline, figure 4).

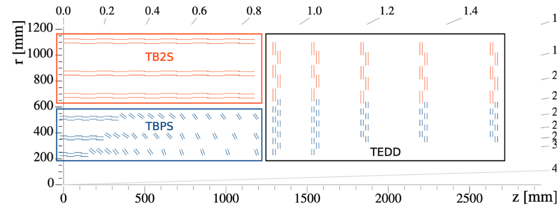

Figure 2. Layout of the new CMS OT [6] in the r-z plane: 2S modules (red) are in the outer radial region, PS modules (blue) are in the inner radial region

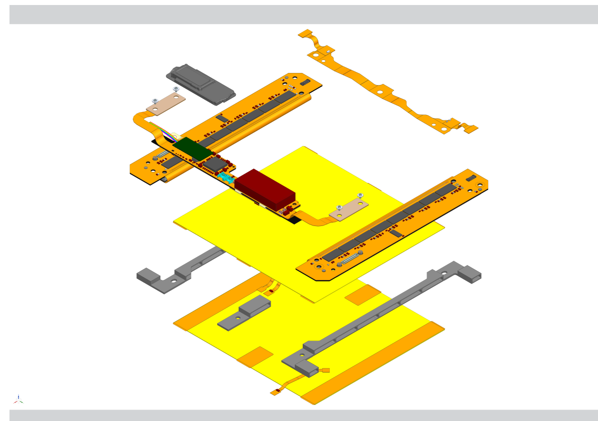

Figure 3. Exploded view of the 2S OT module

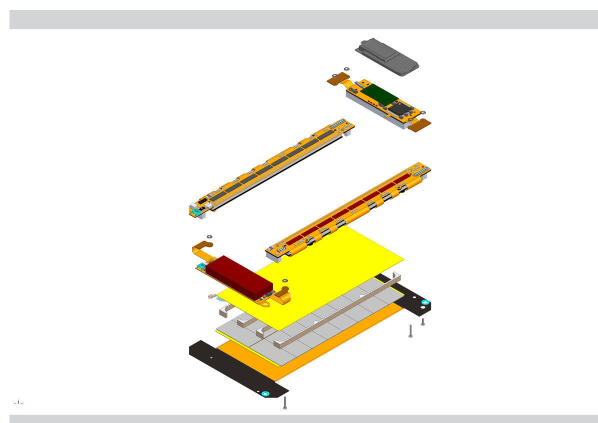

Figure 4. Exploded view of the PS OT module

In the 3.8 T magnetic field in the bore of the superconducting solenoid charged particles bend under the action of Lorentz's force following the well known relation $p = qBr$, where p is the particle's momentum, q its charge, B the magnetic field strength, and r the curvature radius of the particle's trajectory. In a fixed magnetic field high momentum particles curve less than low momentum ones. This is exploited by both the 2S and PS modules which perform a high transverse momentum (p_T) discrimination of incoming charged particles by following the "p_T module" concept: when a particle generates hits on both layers, the module front end electronics correlate them and a tracklet signal (called "stub") is produced if the second hit is within a certain (programmable) spatial window with respect to the first hit (figure 5); the stub is then provided as a L1 trigger input together with informations from the calorimeters and the muon detectors. At the current stage of development [8] a $p_T = 2\,\mathrm{GeV}$ threshold has been identified as sufficient to reduce data volumes by approximately one order of magnitude to acceptable levels with respect to the 40 MHz bunch crossing rate.

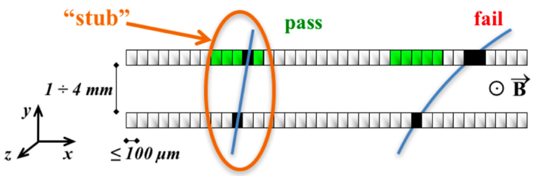

Figure 5. The p_T module concept

While the IT, which is subject to higher fluences, will be built to be disassembled and replaced, the OT is too big and too complex, so the new OT will be built with the goal of surviving the HL-LHC conditions for the entire lifetime of the CMS experiment with little to no servicing once assembled. This requires an extensive campaign of module testing and performance evaluation to make sure each component is up to the highest quality standards.

3.3 Module testing at Fermilab

Many CMS institutions worldwide, including Fermilab, are preparing for production of the 13200 OT modules that will be installed with the Phase-2 OT Upgrade, so a standardized approach to module testing, analysis and performance evaluation is needed across every testing site.

At the Silicon Detector (SiDet) facility OT modules will undergo burnin box tests, and the Phase-2 Acquisition and Control Framework (Ph2-ACF) is used to produce ROOT output files based on those tests. The analysis and evaluation of such outputs is performed by the CMS Phase-2 Outer Tracker Analyzer of Test Outputs (POTATO) software, on which my 2024 summer internship was mainly focused and which is described in further detail in the next section.

4. The CMS Phase-2 Outer Tracker Analyzer of Test Outputs software

4.1 Overview

The Phase-2 Outer Tracker Analyzer of Test Outputs (POTATO) is a software tool built in C++, ROOT, Python and Qt which is currently being developed by CMS teams at Fermilab and at Aachen University. Its main goal is to be used worldwide by CMS module testing sites to assess the quality and performance of Phase-2 OT modules through a unified approach so that production of such modules can be ramped up smoothly. It has a Graphical User Interface (GUI) which makes it easy to use, but can also be run directly from terminal in its lightweight version (POTATO Express). It is able to directly upload testing, analysis and grading results directly to the online CMS module production database, so that relevant data for the Phase-2 OT upgrade is centralized and easily accessible. Users and developers can improve it by implementing plugins in a standardized way to expand its functionalities, the most important of which are described in the next section.

4.2 Functioning

The main workflow of core POTATO usage consists in the following steps:

- **Test output selection**: through the POTATO GUI (figure 6) the user can select the Ph2-ACF output file referring to the module under test. A series of sub-selectors is present to refine the file search by allowing the user to choose the module type, the test date and site and other relevant parameters.

- **Analysis and Grading**: once the test output has been selected, the user can perform both the module analysis and grading by clicking the associated buttons, which call dedicated C++ code to read the ROOT output file and execute the desired actions, thus producing analysis (figure 7) and grading XML documents for the selected ROOT file.

- **CERN login and upload**: after the analysis and grading have been performed, the user can authenticate using their CERN account and proceed to click the Upload button associated with the selected ROOT file, the analysis and/or grading XML document or all the mentioned files together. Upon the release of the button a Python script that compresses the relevant files and loads the CMS database web page for the upload is called. If the chosen files match the database table the upload is successfully performed.

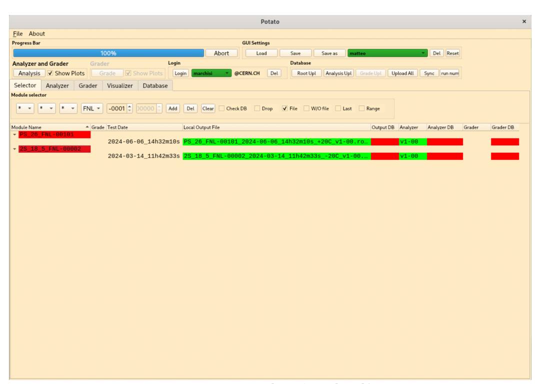

Figure 6. POTATO GUI

```
<ROOT>
 <HEADER>
  <RUN mode="AUTO_INC_NUMBER">
   <RUN_TYPE>mod_final</RUN_TYPE>
   <LOCATION>FNL</LOCATION>
   <INITIATED_BY_USER>marchisi</INITIATED_BY_USER>
   <RUN_BEGIN_TIMESTAMP>2024-03-14 11:42:33</RUN_BEGIN_TIMESTAMP>
   <COMMENT_DESCRIPTION></COMMENT_DESCRIPTION>
  </RUN>
  <TYPE>
   <EXTENSION_TABLE_NAME>MOD_2S_ANL_SMMRY</EXTENSION_TABLE_NAME>
   <NAME>2S Module Analysis Summary</NAME>
  </TYPE>
 </HEADER>
 <DATA_SET>
  <VERSION>v1-00</VERSION>
  <PART>
   <KIND_OF_PART>2S Module</KIND_OF_PART>
   <SERIAL_NUMBER>2S_18_5_FNL-00002</SERIAL_NUMBER>
  </PART>
  <DATA>
   <ROOT_FILE>18889101</ROOT_FILE>
   <ANL_VER>v1-00</ANL_VER>
   <ANL_CUTS_VER>v1-00</ANL_CUTS_VER>
   <LV_CURR_AMP>0.444</LV_CURR_AMP>
   <IV_CURR_MAMP>0.666</IV_CURR_MAMP>
   <IV_RATIO>9999999</IV_RATIO>
   <IV_BREAKDOWN_V>9999999</IV_BREAKDOWN_V>
   <READ_ERR>0</READ_ERR>
   <NOISE_AVG>7.904108</NOISE_AVG>
   <NOISE_RMS>0.779515</NOISE_RMS>
   <NOISE_NOUTL_LOW>0</NOISE_NOUTL_LOW>
   <NOISE_OUTL_LOW></NOISE_OUTL_LOW>
   <NOISE_NOUTL_HIGH>1900</NOISE_NOUTL_HIGH>
   <NOISE_OUTL_HIGH>0,1,732,750,1030,1040,1042,1044,1048,1076,1083,1088,1093,1114,1126,1158,1162
   <NOISE_NOUTL_TOTAL>1900</NOISE_NOUTL_TOTAL>
   <PED_AVG>596.31</PED_AVG>
   <PED_RMS>0.451564</PED_RMS>
   <PED_NOUTL_LOW>5</PED_NOUTL_LOW>
   <PED_OUTL_LOW>171,310,1824,1826,1832</PED_OUTL_LOW>
   <PED_NOUTL_HIGH>12</PED_NOUTL_HIGH>
   <PED_OUTL_HIGH>43,304,382,443,656,784,1161,1256,1266,1278,1345,1497</PED_OUTL_HIGH>
   <PED_NOUTL_TOTAL>17</PED_NOUTL_TOTAL>
   <ALIGNMENT_FEHR_L1>1</ALIGNMENT_FEHR_L1>
```


Figure 7. Example of 2S analysis XML document: each field corresponds to a relevant variable extracted from the Ph2-ACF output file

4.3 My work

During my 2024 summer internship I worked on optimizing existing and new POTATO code to make it modular and easily expandable by other developers: my main focus was on almost completely re-implementing the analysis and grading processes through an Object Oriented Programming (OOP) approach fully compatible with C++, ROOT and Qt built-in implementations. I also helped in adding a new ROOT histogram to both the analysis and grading procedures and took part in the debugging of the CERN login process. In the following sections I describe in detail my code contributions [5].

4.3.1 Generic and specific 2S/PS C++ classes

Some parts of the analysis (or grading) process are common to both the 2S and PS modules, while others are different because of the construction parameters of the two types of modules, so I chose an hybrid approach: I implemented a generic Analyzer (Grader) class from which specific Analyzer2S/PS (Grader2S/PS) publicly inherit. In this way analysis (grading) methods to write XML fields can be added accordingly:

- If the process is exactly the same for both types of modules, then a base class method is implemented, and both inherited classes call it in the same way.
- If the process differs in some way (from just different statistical variables to a completely customized approach for each type of module), then an inherited 2S/PS class method is implemented.

This approach made it possible to completely segment both the analysis and grading processes by implementing a single method to write a single XML field (or group of strictly related fields). In this way the existing code was made extremely more human-readable and expandable.

By following this improved paradigm I helped in adding a new Ph2-ACF histogram that analyzes the efficiency of a module in reconstructing a known injected pattern in a specific set of pixels/strips (figure 8).

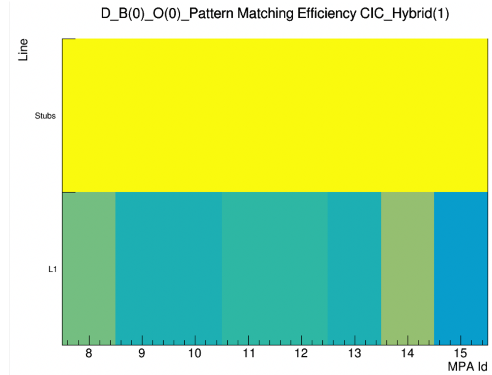

Figure 8. Pattern matching efficiency histogram for a PS module

4.3.2 Versioning system

As mentioned before, POTATO performs the upload of selected files if they match exactly the database structure. In the first stages of the OOP re-implementation of the analysis and grading procedures a mismatch between local XML files and online database tables emerged and the upload could not be completed, so I expanded upon an already present skeletal versioning system for the Analyzer and Grader classes to account for this mismatch. I did so by differentiating the Analyzer/Grader versions based on the chosen development stage:

- **v1-00**: XML documents created with this version are up to date with the online tables and can be successfully uploaded to the CMS database, while new fields are simply not added nor written into the XML documents.
- **v1-0x**: XML documents created with these versions are used to test the implementation of new fields based on updated Ph2-ACF outputs and cannot be uploaded until the database developers add the new fields to the online tables.

The OOP approach described above thus has the side effect of facilitating the version differentiation: a single method that writes a single field (or group of strictly related fields) is called according to the chosen version. At the current stage of development the Analyzer and Grader versions should always match, although a “missing field” system could be implemented to account for local mismatches due to different types of tests (e.g. new vs. old or full vs. quick) or updated grading decisions (A, B, C etc. vs. A+, B–, N/A etc.).

4.3.3 CERN account login

From October 1st every CERN personal account is mandatorily required to switch from Single Sign On (SSO) to Two Factor Authentication (2FA) access, and this change broke the

CERN login process in the Python script called by POTATO upon clicking the login button.

Together with my supervisor Lorenzo Uplegger we simplified the login procedure to accept only SSO accounts and we tested it with a CERN services account. Once functional we updated POTATO and suggested that every CMS OT module testing site creates a CERN services account for the purpose of logging into POTATO and uploading to the database.

5. Conclusions

During my 2024 summer internship I had the great opportunity of working in close contact with CMS experts who greatly helped me in understanding the current stage of development of the CMS detector and in improving the POTATO software. I learned how collaborate on a common project with other people and how to manage my time efficiently to meet deadlines, and I improved my coding skills by implementing an OOP approach to the POTATO software. I also deepened my understanding on how to debug code and and learned how to interact with the online CMS database. These are the main takeaways of this experience:

- With its unified approach to OT module analysis and grading, POTATO is a very powerful tool which will certainly help in the overall production and installation of the Phase-2 Outer Tracker upgrade.
- Some work still needs to be done on the Analyzer classes, since their current implementation is not fully OOP compliant and new graphs are constantly being added to Ph2-ACF ROOT outputs.
- The grading decisions still need to be fully expanded and agreed upon by OT module testing sites: specifically, the overall module grade assignment process is yet to be defined, and this is of crucial importance to choose which modules will eventually be installed permanently.

Acknowledgments

I would like to thank my supervisors, Lorenzo Uplegger and Fabio Ravera, and SiDet colleagues (including Irene Zoi and Nicolò Salimbeni), without whose constant support, understanding and and friendliness I would not have been able to fully exploit the immense potential of this internship.

I would like to thank Fermilab and the University of Pisa for accepting my application and giving me the opportunity to work on such an important project.

I would also like to thank Stefano Argirò of the University of Turin for informing me about the Italian Summer Student Program back in January.

Finally, I would like to thank my family and friends (both those at home and those I made while in the United States) for their support and encouragement throughout the entire duration of this internship.

References

[1] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont, and L. Rossi. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2015.

[2] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[3] V. K. et al. The cms trigger system. *Journal of Instrumentation*, 12(01):P01020, jan 2017.

[4] L. Guiducci. Cms muon system towards lhc run 2 and beyond. *Nuclear and Particle Physics Proceedings*, 273-275:1014–1022, 2016. 37th International Conference on High Energy Physics (ICHEP).

[5] M. Marchisio Caprioglio. Potato gitlab repository branch. `https://gitlab.cern.ch/cms_tk_ph2/potato/-/tree/marchisi?ref_type=heads`, 2024.

[6] K. Nash. The phase-2 upgrade of the cms outer tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1058:168788, 2024.

[7] S. Roy Chowdhury. The Phase-2 Upgrade of the CMS Outer Tracker. Technical report, CERN, Geneva, 2020.

[8] I. Zoi. Design and construction of the CMS Outer Tracker for the Phase-2 Upgrade. *PoS*, VERTEX2023:021, 2024.