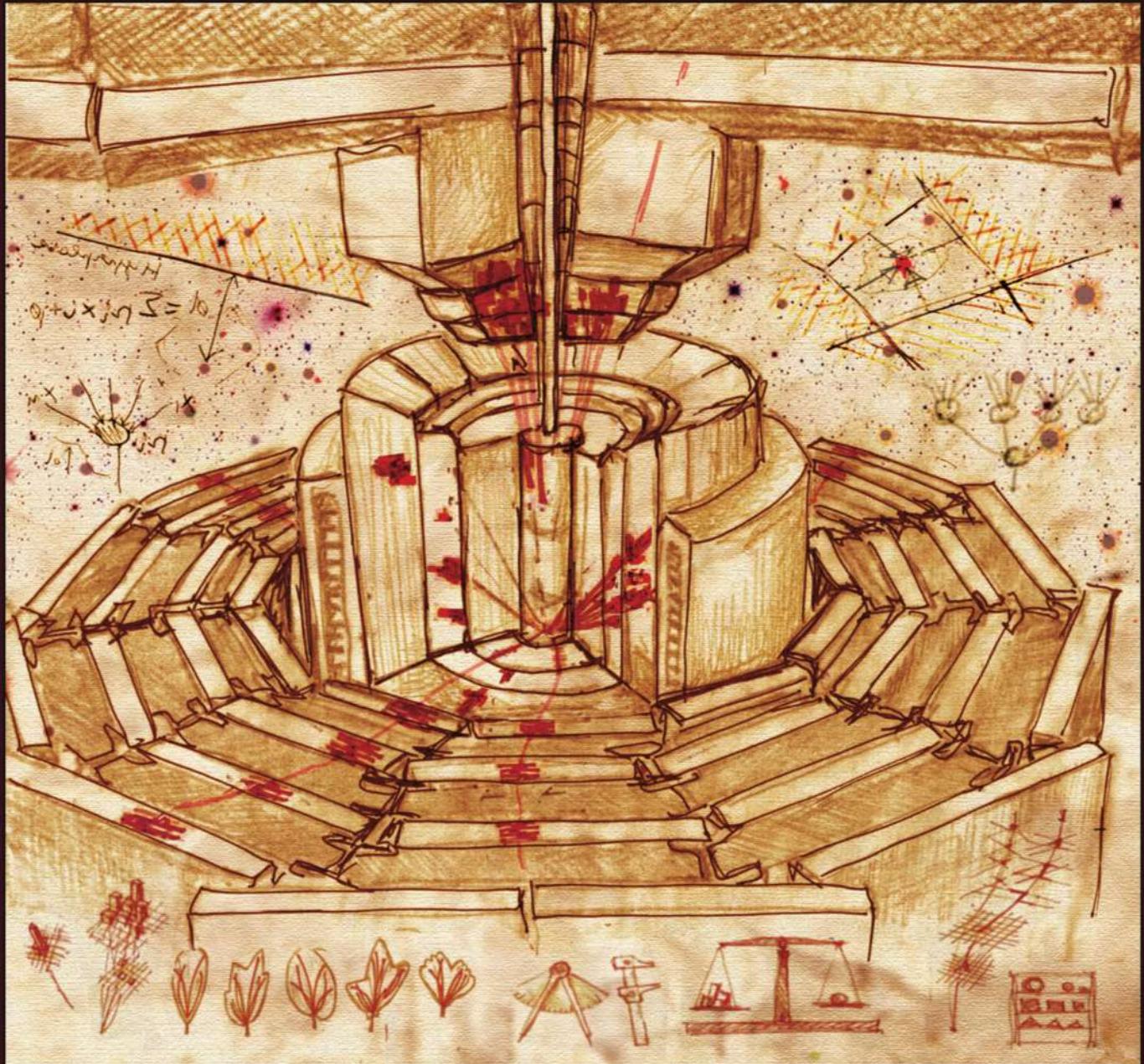
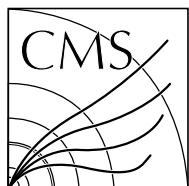


CMS



The Phase-2 Upgrade of the CMS Level-1 Trigger Technical Design Report



CERN-LHCC-2020-004

CMS-TDR-021

10 March 2020

The Phase-2 Upgrade of the CMS Level-1 Trigger

Technical Design Report

CMS Collaboration

Editors

A. Zabi, J. Berryhill, E. Perez and A. Tapper

Chapter Editors

D. Acosta, A. Attikis, M. Bachtis, C. Battilana, J. Berryhill, S. Bhowmik, S. Bologna, C. Botta, J. Brooke, A. Bundock, K. Bunkowski, L. Cadamuro, C. Caillol, M. Cepeda, G. Cerminara, Y. Chen, S. Dasu, J. W. Davies, S. Dildick, J. Duarte, S. Dutta, K. Foudas, B. François, B. Gomber, T. Gorsky, G. Iles, P. Harris, J. Hauser, J. Hegeman, T. Herwig, S. Jindariani, G. Karapostoli, J. Konigsberg, J. M. Langford, T. Laird, T. Lam, J. Li, J. Low, E. MacDonald, A. Madorsky, A. Mallampalli, S. Mallios, N. Marinelli, C. Martín-Pérez, E. Meschi, C. Mironov, I. Ojalvo, S. Paramesvaran, S. E. Park, R. Patel, C. Peña, E. Perez, A. Perloff, G. Petrucciani, F. Ptochos, D. Rabady, D. Rankin, V. Rekovic, A. Rose, T. Ruggles, S. Sarkar, H. Sakulin, J.-B. Sauvan, A. Savin, V. Sharma, A. Shtipliyski, L. Skinnari, D. Sperka, S. Summers, A. Svetek, A. Tapper, M. Tepper, K. Theofilatos, N. V. Tran, K. Ulmer, C. Veelken, C. Vellidis, E. Vourliotis, T. Williams, A. Zabi, Z. Wu.

Cover Design

Sergio Cittolin

Acknowledgements

This document would not exist without the devoted efforts of many colleagues, too many to be named explicitly, who contributed text and figures to this TDR. The very useful input from the CMS Trigger and Data Acquisition project, the CMS internal reviewers (D. Acosta, J. Alcaraz, T. Boccali, G. Boudoul, T. Camporesi, M. Capeans Garrido, S. Donato, S. Gennai, F. Hartmann, J. Kieseler, I. Mikulec, M. Pieri, C. Rovelli, S. Savin, P. Wittich), and the chair of the Phase-2 TDRs editorial board (C. Lourenco) helped to improve the quality of this document and is greatly appreciated. We would like to acknowledge the support of the CMS Management and the Upgrade Project Coordination team. We thank the Offline, Computing, Physics Performance and Datasets, Upgrade Physics Strategy Group, and all the Physics groups for their help in developing and executing the physics and performance studies. We would also like to thank O. Davignon and B. Radburn-Smith for their careful review of all the performance results included in this document, as well as K. Black, M. Konecki, P. Padley, for their help improving the document during the internal review by the project. We would like to thank Sergio Cittolin for providing the cover artwork inspired by this project.

Executive Summary

This Technical Design Report (TDR) describes the ongoing developments and plans towards the upgrade of the CMS Level-1 (L1) trigger for the High-Luminosity Large Hadron Collider (HL-LHC). The HL-LHC era constitutes the Phase-2 of the LHC operation succeeding to the Phase-1 exploitation period currently ongoing and ending in 2024. In its ultimate configuration, the HL-LHC will reach unprecedented performance in terms of instantaneous luminosity ($7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$) potentially leading to a total integrated luminosity of up to 4000 fb^{-1} after ten years of operations, scheduled to start in 2027. This previously unmatched amount of data opens the door to a rich and ambitious physics program including both high-precision measurements and searches for physics beyond the Standard Model.

To achieve the goals of the HL-LHC program, the Phase-2 upgrade of the L1 system utilizes technological advances to enhance the physics selectivity already at the hardware level of the data acquisition system. The intense hadronic environment corresponding to 200 simultaneous collisions per beam crossing, imposes serious challenges to the system requirements in order to maintain performance. To profit from the extended coverage and increased granularity of the upgraded CMS detector, the latency of the system is extended to $12.5 \mu\text{s}$. The use of tracking and high-granularity calorimeter information is possible for the first time at L1. The maximum output bandwidth is 750 kHz to maintain the required performance. Modern processors are used to implement sophisticated algorithms including machine learning-based approaches to target the selection of specific final states.

As a starting point, the trigger algorithms have been studied with the minimum requirement of reaching the performance that CMS needs to perform the studies discussed in Ref. [1], a review that offers a detailed report on the physics program of the HL-LHC. Additionally, the algorithms were further developed to allow studies of several supplementary physics topics. Algorithms were simulated in software, and also prototyped in firmware, to estimate their hardware resource usage and corresponding latency.

Informed by the algorithmic requirements, a program of hardware R&D is ongoing with the aim of identifying hardware platforms and system architectures suitable for implementing the required algorithms. Developments are based on the advanced telecommunications computing architecture (ATCA) standard for electronics, using state-of-the-art field-programmable gate arrays (FPGAs) and serial optical links running at speeds up to 25 Gb/s. Prototypes produced by several groups meet the requirements of the project and in many cases algorithms have been tested directly in prototype hardware. Plans for future development and testing are also summarized.

The physics and environmental considerations that lead to the choice of the conceptual design of the upgraded trigger system along with its technological implementation and its expected performance are provided. The organization of the project is presented, including the institutes participating, the estimated cost, schedule and milestones. The schedule includes the plan for installation, integration and commissioning of the system during the LHC long-shutdown 3 (LS3 starting 2025) in view of delivering triggers for the start of the LHC Run-4 in 2027.

Overall the studies contained within the report show that the physics requirements can be met within the technical constraints of the project with appropriate contingency.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction and overview | 1 |
| 1.1 | Physics motivations and CMS Phase-2 trigger upgrade | 1 |
| 1.2 | HL-LHC upgrade and triggering challenges | 4 |
| 1.3 | Phase-2 Level-1 trigger requirements and design | 6 |
| 1.3.1 | Phase-2 L1 upgrade conceptual design | 7 |
| 1.3.2 | Trigger algorithms for the HL-LHC | 11 |
| 1.4 | Physics reach of the Level-1 Phase-2 Trigger | 18 |
| 1.4.1 | Preserving the CMS discovery potential at the HL-LHC | 18 |
| 1.4.2 | Extending the CMS physics reach with the Phase-2 Level-1 trigger | 24 |
| 1.5 | Outline of the Technical Design Report | 29 |
| 2 | Inputs to the L1 trigger system: the trigger primitives generation | 31 |
| 2.1 | Track finder primitives | 31 |
| 2.2 | Barrel calorimeter trigger primitives | 36 |
| 2.2.1 | Electromagnetic barrel calorimeter | 36 |
| 2.2.2 | Hadron barrel and forward calorimeters | 40 |
| 2.3 | High granularity endcap calorimeter trigger primitives | 42 |
| 2.3.1 | System architecture | 42 |
| 2.3.2 | Reconstruction algorithms | 44 |
| 2.3.3 | Implementation, latency and data volumes | 46 |
| 2.4 | Muon barrel trigger primitives | 47 |
| 2.4.1 | Drift tubes | 47 |
| 2.4.2 | Barrel resistive plate chambers | 49 |
| 2.4.3 | DT+RPC super-primitive combination | 50 |
| 2.4.4 | Muon barrel primitives performance | 52 |
| 2.4.5 | Future improvements | 53 |
| 2.5 | Muon endcap trigger primitives | 55 |
| 2.5.1 | Endcap resistive plate chambers | 55 |
| 2.5.2 | Improved resistive plate chambers | 55 |
| 2.5.3 | Endcap cathode strip chambers | 56 |
| 2.5.4 | Gas electron multipliers | 57 |
| 2.5.5 | Endcap muon trigger primitives performance | 59 |
| 2.5.6 | Future improvements | 59 |
| 2.6 | External Triggers | 61 |
| 2.7 | Summary of trigger input bandwidth and latency | 62 |
| 3 | Trigger algorithms | 65 |
| 3.1 | Introduction: triggering on physics objects at the HL-LHC | 65 |
| 3.2 | Triggering on electrons and photons | 66 |
| 3.2.1 | Introduction | 66 |
| 3.2.2 | Electron and photon reconstruction in the calorimeter barrel | 66 |
| 3.2.3 | Electron and photon reconstruction in the endcap calorimeter | 69 |

| | | |
|----------|--|------------|
| 3.2.4 | Exploiting L1 track-finder primitives for e/ γ triggers | 71 |
| 3.2.5 | Firmware implementation and resources | 78 |
| 3.3 | Triggering on muons | 80 |
| 3.3.1 | Introduction | 80 |
| 3.3.2 | Standalone muon reconstruction in the barrel region | 83 |
| 3.3.3 | Standalone muon reconstruction in the overlap region | 90 |
| 3.3.4 | Standalone muon reconstruction in the endcap region | 94 |
| 3.3.5 | Global muon trigger algorithms | 99 |
| 3.4 | Vertex reconstruction | 121 |
| 3.4.1 | Histogram-based vertexing algorithm | 121 |
| 3.4.2 | Track to vertex association | 122 |
| 3.4.3 | Vertex finding performance | 122 |
| 3.4.4 | Firmware implementation of vertexing algorithms | 125 |
| 3.4.5 | End-to-end neural network vertex reconstruction model | 129 |
| 3.5 | Particle-flow reconstruction | 131 |
| 3.5.1 | Calorimeter inputs | 131 |
| 3.5.2 | Tracking and muon inputs | 138 |
| 3.5.3 | Particle-flow and PUPPI algorithms | 139 |
| 3.5.4 | Particle-flow and PUPPI algorithm performance | 143 |
| 3.5.5 | Firmware implementation | 146 |
| 3.6 | Triggering on jets, hadronic tau decays, and energy sums | 155 |
| 3.6.1 | Jet finding algorithms | 155 |
| 3.6.2 | Hadronic τ algorithms | 165 |
| 3.6.3 | Missing transverse energy algorithms | 181 |
| 3.7 | Global trigger algorithms | 188 |
| 3.7.1 | Cut-based trigger algorithms | 188 |
| 3.7.2 | Machine learning based trigger algorithms | 191 |
| 3.7.3 | Overall Global Trigger resource usage | 194 |
| 3.8 | Trigger algorithms for heavy ion collisions | 195 |
| 3.8.1 | Introduction | 195 |
| 3.8.2 | Physics program | 195 |
| 3.8.3 | Trigger considerations | 196 |
| 4 | Level-1 trigger menu | 199 |
| 4.1 | Inputs and design strategy of the Phase-2 Level-1 trigger menu | 200 |
| 4.1.1 | Muon trigger algorithm inputs | 201 |
| 4.1.2 | Electron and photon trigger algorithm inputs | 202 |
| 4.1.3 | Jets and jet-sum trigger algorithm inputs | 203 |
| 4.1.4 | Missing transverse energy trigger algorithm inputs | 204 |
| 4.1.5 | Tau trigger algorithm inputs | 204 |
| 4.2 | Simplified menu with Phase-1 thresholds | 205 |
| 4.3 | New L1 trigger algorithms to extend Phase-1 physics acceptance | 210 |

| | | |
|----------|---|------------|
| 4.3.1 | Physics with the extended pseudorapidity coverage of the L1 lepton triggers | 210 |
| 4.3.2 | Physics with L1 soft and correlated muons | 211 |
| 4.3.3 | Physics of light mesons with L1 tracking | 217 |
| 4.3.4 | Physics with L1 displaced muons | 219 |
| 4.3.5 | Physics with L1 displaced jets | 221 |
| 4.3.6 | Dedicated VBF Higgs boson production L1 trigger algorithms based on machine learning techniques | 226 |
| 4.3.7 | Summary of triggers for extended physics reach | 228 |
| 4.4 | Study of the menu rate evolution with pileup | 228 |
| 4.4.1 | Individual trigger rate evolution with pileup | 229 |
| 4.4.2 | Alternative E_T^{miss} and VBF triggers for extreme pileup | 231 |
| 4.4.3 | Menu rate and bunch-to-bunch pileup variations | 231 |
| 5 | Conceptual design of the Phase-2 L1 Trigger | 235 |
| 5.1 | Introduction | 235 |
| 5.2 | Calorimeter Trigger system | 235 |
| 5.3 | Muon Trigger system | 239 |
| 5.3.1 | Endcap Muon Track Finder | 240 |
| 5.3.2 | Overlap Muon Track Finder | 242 |
| 5.3.3 | Barrel Muon Track Finder | 242 |
| 5.3.4 | Global Muon Trigger | 243 |
| 5.4 | Global Track Trigger system | 244 |
| 5.5 | Correlator Trigger system | 246 |
| 5.5.1 | Correlator Trigger Layer-1: particle-flow candidates | 247 |
| 5.5.2 | Correlator Trigger Layer-2: object reconstruction | 249 |
| 5.6 | Global Trigger system | 251 |
| 5.6.1 | Inputs to the Global Trigger | 251 |
| 5.6.2 | Global Trigger architecture | 253 |
| 5.6.3 | Interfaces to TCDS, DAQ and HLT | 254 |
| 5.6.4 | Using the 40 MHz Scouting System for trigger monitoring | 256 |
| 5.7 | Scouting system | 256 |
| 5.7.1 | General architecture | 256 |
| 5.7.2 | Inputs, memory, and connectivity requirements | 259 |
| 5.8 | Overall trigger system architecture | 261 |
| 6 | Instrumentation of the Phase-2 trigger system | 265 |
| 6.1 | Introduction | 265 |
| 6.2 | System infrastructure and services | 265 |
| 6.3 | Hardware research and development | 269 |
| 6.3.1 | ATCA processor: APx consortium | 269 |
| 6.3.2 | ATCA processor: Serenity consortium | 277 |
| 6.3.3 | ATCA processor: Ocean | 285 |
| 6.3.4 | Specific processing board : Barrel Muon Trigger (BMT) demonstrator | 290 |

| | | |
|-------|---|-----|
| 6.4 | Trigger demonstrators and slice tests | 295 |
| 6.4.1 | Calorimeter trigger demonstrators | 295 |
| 6.4.2 | Muon trigger demonstrators: muon track finder and global muon trigger | 297 |
| 6.4.3 | Global track trigger demonstrator | 308 |
| 6.4.4 | Correlator trigger demonstrators | 315 |
| 6.4.5 | Demonstration of the 40 MHz scouting system | 322 |
| 6.5 | System-wide technical issues | 324 |
| 6.5.1 | Hardware issues | 324 |
| 6.5.2 | Firmware and software deployment | 326 |
| 6.5.3 | Trigger links and system synchronization | 327 |
| 6.5.4 | Software-firmware interface | 328 |
| 6.5.5 | Commissioning and parallel running | 329 |
| 6.5.6 | Monitoring and data validation | 330 |
| 6.5.7 | Latency measurements and checks | 330 |
| 6.5.8 | System health check | 331 |
| 6.5.9 | Demonstrator systems in Run-3 | 331 |
| 7 | Organization, schedule, and costs | 333 |
| 7.1 | Project organization and work breakdown structure | 333 |
| 7.2 | Project schedule and milestones | 334 |
| 7.2.1 | Schedule | 334 |
| 7.2.2 | Milestone list | 336 |
| 7.3 | Institution interests and construction responsibilities | 336 |
| 7.4 | Project costs | 336 |
| | References | 349 |

Chapter 1

Introduction and overview

The HL-LHC [2] presents the opportunity for a very rich and ambitious physics program, exploiting an integrated luminosity of 3000 fb^{-1} . The LHC will undergo major upgrades of its components leading to an increase of the instantaneous luminosity to $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, five times the machine’s original design value. In its ultimate configuration, the HL-LHC will reach a peak instantaneous luminosity of $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, increasing the average number of proton-proton collisions per bunch crossing (pileup) to around 200. The ultimate performance of the HL-LHC would enable the collection of 400 to 450 fb^{-1} of integrated luminosity per year, potentially providing a total of 4000 fb^{-1} to each of the CMS and ATLAS experiments.

The CMS detector requires a trigger and data acquisition system with exceptional performance to collect the required information-rich datasets in these challenging running conditions [3]. The upgrade of the trigger system will enhance the physics selectivity and maintain the performance necessary throughout the 10 year long HL-LHC program, which includes heavy ion operations.

1.1 Physics motivations and CMS Phase-2 trigger upgrade

Physics motivations: The CMS Phase-2 physics program plans to fully exploit the HL-LHC to perform searches for new physics beyond the Standard Model (BSM), achieve unprecedented high-precision measurements of the SM, including a significantly improved characterization of the Higgs Boson sector. A broad spectrum of physics analyses will become possible with the unprecedently large HL-LHC data samples and the new capabilities offered by the detector upgrade, such as its extended coverage. More advanced selection algorithms are required to maintain the effective selection of electroweak-scale processes with 200 pileup events. Sophisticated triggers will be required to select specific topologies such as VBS/VBF production, rare B -meson decays (based on usage of tracks in the L1 trigger for the first time), forward muon trigger for $\tau \rightarrow \mu\mu\mu$ (profiting from extended coverage), etc. Low mass resonances could be identified with a dedicated scouting system. In order to illustrate the discovery potential of the upgraded CMS detector, achieved through the efficient selection of the L1 system, benchmark signals have been selected and are presented in this document.

The CMS trigger system: The CMS experiment currently implements a sophisticated two-level triggering system composed of Level-1 (L1) [4], instrumented by custom hardware processor boards, and software High Level Trigger (HLT). The L1 receives information from calorimeter and muon systems generating an initial selection within a fixed latency of $4 \mu\text{s}$, with a maximum output rate of 100 kHz. Upon reception of a L1 Accept (L1A) signal, the detector is fully read out and the selected event is reconstructed in the HLT. The HLT selection is based on this finer information, reducing the output rate to about 1 kHz on average. A first ma-

major upgrade of the L1 system was conducted during the long-shutdown 1 (LS1 2013–2015). The Phase-1 upgrade consisted in the complete replacement of the system that was deployed and successfully operated during Run-1 of the LHC (2010–2013). A new architecture with improved performance was installed to maintain a high efficiency for collecting the data, under the more challenging conditions experienced during Run-2 (2015–2018) and expected during Run-3 (2021–2024). All the hardware, interconnections, electronics boards, firmware and software were redesigned as described in the technical design proposal (TDR) [5] of the L1 trigger Phase-1 upgrade. Benefiting from the higher input granularity, more sophisticated and innovative object reconstruction algorithms were implemented. High-speed optical links were installed to rapidly collect all the information from sub-detectors contributing to the trigger, hence providing a full field view of the detector, adapted to the precise evaluation of global event quantities, such as pileup or energy sums. In addition, the new global trigger is capable of evaluating complex selection algorithms such as those involving the invariant mass of trigger objects. Using correlations in multi-object triggers has allowed CMS to further enhance the optimization of physics sensitivity while adapting to changing conditions and priorities. The performance of the L1 system during the Run-2 data-taking period is described in Ref. [6]. In spite of the increased luminosity ($2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$) and pileup, changes of proton filling schemes, and the effects of aging of the sub-detectors, the flexibility of the system has allowed the acceptance for physics to be maintained and a rich physics program to be pursued throughout the whole period. Many aspects of the Phase-2 upgrade of the L1 trigger system are inspired by the key technological choices made during the Phase-1 upgrade.

The Phase-2 upgraded CMS detector: In order to fully exploit the HL-LHC running period, major consolidations and upgrades of the CMS detector are planned [7]. The collision rate and level of expected pileup imply very high particle multiplicity and an intense radiation environment. The performance required on event object reconstruction to achieve the extraction of physics signatures relies on the implementation of higher granularity detectors along with robust readout electronics. The CMS collaboration plans to replace both the Strip and Pixel tracking detectors, with an Inner Tracker featuring small size pixel sensors and an Outer Tracker equipped with strip and macro pixel sensors, extending their coverage to $|\eta| = 3.8$ [8]. A narrower pitch will provide better transverse and longitudinal impact parameter resolution. The Outer Tracker will implement stacked strip modules, reducing the hit multiplicity and allowing track candidates for the trigger (L1 tracks) to be reconstructed up to $|\eta| = 2.4$, which opens major new possibilities for the L1 trigger. The readout electronics for the barrel calorimeters will be replaced to achieve finer granularity and provide timing information [9]. The endcap calorimeters will be replaced by the high-granularity calorimeter (HGCal) [10], implementing over 6 million readout channels. This sampling calorimeter will provide shower separation and identification adapted to harsher conditions in the forward region of the detector. The muon detection system redundancy achieved through the combination of drift tubes (DTs), resistive plate chambers (RPCs), and cathode strip chambers (CSCs) will remain with consolidated electronics [11]. Additional improved RPC (iRPC) chambers and gas electron multiplier (GEM) chambers will be installed to extend the coverage up to $|\eta| = 2.4$ and 2.8, respectively [11]. A minimum ionizing particle (MIP) timing detector (MTD) [12] placed in front of the barrel and endcap calorimeters will provide precise timing measurement of charged tracks. Along with the sub-detector upgrades, a complete replacement of the trigger (L1 and HLT) and data acquisition (DAQ) system, with increased throughput, is planned. Similarly to Phase-1, the Phase-2 HLT will have access to the full granularity of the detector, with an input rate of 750 kHz in the ultimate HL-LHC scenario; the target for timing is 500 ms, measured on a current (2018) node. The selection algorithms will perform a rate reduction leading to an output bandwidth of 7.5 kHz. The most promising avenue of development is the use of high performance hardware;

for example, porting part of the reconstruction to run on GPUs [7].

The L1 Phase-2 upgrade: The Phase-2 upgrade of the L1 trigger system is designed not only to maintain the efficiency of the signal selection to the level of the Phase-1 performance, but also to significantly enhance, or enable, the selection of any possible new physics manifestations that could lead to unconventional signatures. High-precision measurements of physics processes will benefit from the extension of the available phase space such as enhanced trigger coverage in the forward region of the detector or the ability to exploit fully hadronic final states. Most importantly, state-of-the-art techniques used in offline reconstruction and analyses, such as the global event reconstruction based on particle-flow techniques [13], become possible at the L1 trigger, with the availability of L1 tracks delivered by the upgraded Outer Tracker. Moreover, these algorithms benefit from the increased granularity of the calorimeter information. The trigger processing capabilities allows their hardware implementation to be realized. Based on the Phase-1 experience, this functionality is obtained through the use of Field Programmable Gate Arrays (FPGA) coupled with high-speed optical links to retrieve the detector data and provide the systems interconnections. A flexible and modular architecture, relying on modern technologies, allows the data handling to be optimized to adapt to LHC running conditions and physics requirements. The Phase-2 upgrade of the trigger and DAQ system will keep a two-level strategy while increasing the L1 maximum rate to 750 kHz. The total latency will be increased to $12.5 \mu\text{s}$ to allow, for the first time, the inclusion of the tracker and high-granularity calorimeter information. Moreover, a longer latency will enable higher-level object reconstruction and identification, as well as the evaluation of complex global event quantities and correlation variables to optimize physics selectivity. The implementation of sophisticated algorithms using particle-flow reconstruction techniques or machine-learning based approaches can now be contemplated. In addition to these features, a 40 MHz scouting system harvesting the trigger primitives produced by sub-detectors and the trigger objects produced at various levels of the trigger system is proposed. The concept of trigger scouting has been introduced in CMS at the HLT [14]. It is based on the use of physics objects reconstructed as a by-product of the triggering process to perform data reduction and analysis, only storing high-level information for selected events, thus overcoming the rate to storage limitations of the DAQ. The Level-1 scouting system will use Level-1 trigger reconstructed objects and quantities in a similar way, selecting and analyzing them on the fly at the collision rate. This system has the additional advantage of allowing systematic search of correlations among multiple contiguous bunch crossing, and can be used to scrutinize the collision events and identify potential signatures unreachable through standard trigger selection processes. In order to successfully integrate and commission this complex upgraded L1 trigger, an approach similar to that adopted in the Phase-1 upgrade is chosen, where part of the system will run in parallel with the current system during Run-3 operations. The muon system in place now will remain in Phase-2 and is already used to test new algorithms and gain confidence in their development.

This TDR follows the interim TDR (iTDR) [3], which described the roadmap of the project, as well as the early research and development studies that led to the identification of the key features of this upgrade design. The iTDR reported the evaluation of the technical capabilities in terms of algorithms, firmware, and hardware required to achieve the L1 trigger features. Sophisticated algorithms were implemented in hardware through the extensive use of High-Level Synthesis (HLS) to produce the corresponding firmware. The decisive rate reduction obtained with the inclusion of tracking information was demonstrated along with the first physics menu. A full range of architecture designs were considered, directing the proposal towards the implementation of a central "Correlator Trigger" system where back-end information feeds advanced triggering algorithms. The iTDR also allowed the identification of potential tech-

nologies for FPGAs, optical links as well as the Advanced Telecommunications Computing Architecture (ATCA) as a platform. Advantages of generic processing engines were discussed and presented as an adequate choice for implementation, maintenance and flexibility of the design. Along with the L1 upgrade, major upgrades are planned for the DAQ and HLT.

1.2 HL-LHC upgrade and triggering challenges

In order to extend the LHC discovery potential, consolidations and upgrades of the machine and its injection chain have been planned during long shutdowns as displayed on the roadmap in Fig. 1.1. During LS1, interconnections between the LHC superconducting magnets were consolidated to permit the operation of the machine at 13 TeV of center-of-mass energy. The machine reached a record instantaneous luminosity of $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, with average values of up to 55 proton-proton collisions per crossing during Run-2 operations. During LS2, ongoing at the time of writing, the LHC is optimizing its parameters and luminosity production. The machine is being consolidated to potentially increase the center-of-mass energy to 14 TeV and sustain a maximum instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ for longer periods of time during Run-3 operations. Phase-1 operations extend to 2024 with the plan to deliver between 350 fb^{-1} and 500 fb^{-1} of data to ATLAS and CMS. Major upgrades to the collider and its experiments will take place during LS3, after which Phase-2 operations will start.



Figure 1.1: LHC baseline plan for the next decade and beyond, showing the energy of the collisions (upper red line) and luminosity (lower red lines). The first long shutdown (LS1) in 2013–2014 allowed the design parameters of beam energy and luminosity to be reached. The second long shutdown (LS2), 2019–2020, will consolidate luminosity production and reliability as well as upgrade the LHC injectors. After LS3, 2025–2027, the machine will be in the High Luminosity configuration (HL-LHC).

The HL-LHC project is already half way through its developments targeting a start of operation in the second half of 2027. The baseline configuration of the upgraded collider should allow the collection of 3000 fb^{-1} of integrated luminosity over ten years of operations. The peak instantaneous luminosity will steadily increase during Run-4 to reach a maximum of $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$

as illustrated in Fig. 1.2, with an average of 140 simultaneous collisions per crossing. In this configuration, the instantaneous luminosity may be maintained at a fixed value throughout the duration of a physics fill. As all the equipment is being designed with a 50% margin with respect to instantaneous heat deposition and integrated radiation dose, the machine performance could possibly be pushed to achieve a peak instantaneous luminosity of $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. This would result in increasing the average number of collisions per bunch crossing to around 200. In this ultimate configuration, the integrated luminosity per year would exceed 400 fb^{-1} . Assuming the experiments can handle this level of pileup, 4000 fb^{-1} would be collected at the end of the HL-LHC lifetime. This performance relies on key innovations pushing the limits of accelerator technology by implementing 11–12 T superconducting magnets, beam collimation, and rotation processes described in Ref. [2].

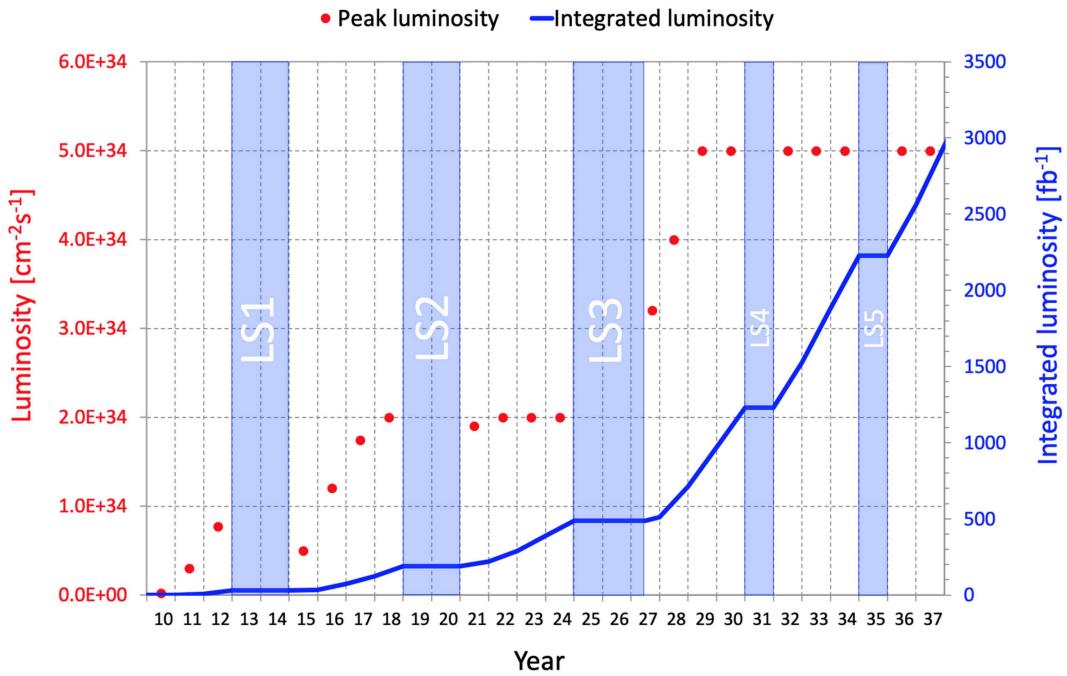


Figure 1.2: Forecast for peak luminosity (red dots) and integrated luminosity (blue line) in the HL-LHC era, according to the nominal HL-LHC parameters. The total integrated luminosity quoted assumes 160 days of physics operation of the LHC complex per year and a physics efficiency of 50%, estimated from the current Phase-1 performance and operation experience. Run-3 will end in 2024, while the LS3 will take place during the years 2025–2027.

The scenarios considered for the HL-LHC beam parameters have been derived based on the configurations used in Run-2. In 2016, although the number of bunches in the machine was smaller than the design value of 2808, the LHC has attained the nominal instantaneous luminosity of $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ through the reduced emittance achieved by the injectors and a reduced β^* of 40 cm at the interaction points IP1 (ATLAS) and IP5 (CMS). The nominal Run-2 filling scheme had 2556 bunches, grouped in trains of 48 bunches each, with 2544 of these colliding at the CMS interaction point. Due to frequent beam dumps caused by electron clouds in the LHC, a special filling scheme called “8b4e” had to be used during the 2017 running. In this specific scheme, short trains of 8 filled bunches followed by 4 empty bunches (without collisions) are injected in the machine and mitigate the formation of electron clouds. A maximum of 1916 bunches could be inserted in the LHC, forcing the bunch intensity to increase to achieve the same instantaneous luminosities. The instantaneous luminosity was levelled to a value

Table 1.1: The main HL-LHC parameters for proton collision operation are compared with the nominal LHC parameters [2]. The three schemes chosen are "standard 25 ns" (baseline configuration), "BCMS" and "8b4e", already encountered during LHC Run-2 operations (see text). N_b is the bunch population; β^* is the beam beta function (focal length) at the collision point; ϵ_L is the longitudinal acceptance; ϵ_n is the transverse normalized emittance; μ is the average pileup (number of inelastic collisions in the same bunch crossing).

| Parameter | Nominal LHC Design Report | HL-LHC 25 ns standard | HL-LHC 25 ns BCMS | HL-LHC 8b4e |
|--|---------------------------|-----------------------|-------------------|-------------|
| Beam energy in collision [TeV] | 7 | 7 | 7 | 7 |
| $N_b[10^{11}]$ | 1.15 | 2.2 | 2.2 | 2.3 |
| Number of bunches per beam | 2808 | 2748 | 2604 | 1968 |
| Beam current [A] | 0.58 | 1.09 | 1.03 | 0.82 |
| Minimum β^* [m] | 0.55 | 0.2 | 0.2 | 0.2 |
| $\epsilon_n [\mu\text{m}]$ | 3.75 | 2.50 | 2.50 | 2.20 |
| $\epsilon_L [\text{eVs}]$ | 2.50 | 2.50 | 2.50 | 2.50 |
| Peak luminosity with crab cavities $[10^{34}\text{cm}^{-2}\text{s}^{-1}]$ | (1.18) | 12.6 | 11.9 | 11.6 |
| Levelled luminosity for $\mu = 140[10^{34}\text{cm}^{-2}\text{s}^{-1}]$ | - | 5.32 | 5.02 | 5.03 |
| (inelastic) collisions/crossing μ (with levelling and crab cavities) | 27 | 140 | 140 | 140 |
| Maximum line density of pileup events during fill [events/mm] | 0.21 | 1.3 | 1.3 | 1.3 |

of $1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ to avoid more than 60 simultaneous collisions per bunch crossing. The mitigation of beam dumps induced by electron clouds in 2018 allowed the return to the nominal filling scheme. Other configurations, such as "Batch Compression Merging and Splitting (BCMS)", have been used to reduce the transverse emittance, in order to avoid a larger than expected emittance growth in the LHC during injection, ramp and squeeze. Table 1.1 compares the nominal LHC parameters with those of the three potential configurations to be considered for the operation of the HL-LHC. The machine parameters have direct implications on the operation and physics performance of the L1 trigger. Challenges, related to the machine filling scheme and pileup, overcome during Phase-1 are described in Ref. [6]. Studies performed in this TDR address the implications of large pileup conditions on triggering performance and exercise the robustness of the design proposed. Unless specified otherwise, all studies presented in this TDR consider pileup conditions of 200, corresponding to the HL-LHC ultimate configuration (not shown in Table 1.1).

1.3 Phase-2 Level-1 trigger requirements and design

With the intense HL-LHC running conditions in terms of instantaneous luminosity and pileup events, the detectors and their associated readout electronics will grow in complexity. As a result, trigger architectures will also become increasingly complex in order to provide sophisticated selection algorithms to ensure the highest possible acceptance for physics already at the hardware level of the data acquisition system. The key technological features of the CMS L1 upgraded trigger system will be the following:

- The extensive use of state-of-the-art FPGAs and processors to achieve optimized

reconstruction, identification, isolation and energy calibration of trigger candidates using high granularity information from the detector.

- The use of high-speed optical links to facilitate the aggregation of data from across the entire detector to allow complex global processing. The availability of the data on the same processing board provides a complete view of the detector for the purpose of the precise evaluation of global quantities such as missing transverse energy or the implementation of pileup mitigation techniques. Moreover, the use of more sophisticated correlations to select specific topologies such as Vector Boson Fusion, where jets are spread over a large polar angle, are facilitated.
- The implementation of a flexible and modular architecture, which can be reconfigured to adapt to different HL-LHC running conditions and physics needs. Extra resources allow more sophisticated quantities to be computed, to give a richer physics menu and increase selectivity.

This section provides a summary of the features of the baseline architecture for the L1 upgraded system and the main characteristics of the algorithms under development. These topics are further developed in Chapter 3 for the algorithms and Chapter 5 for the system design.

1.3.1 Phase-2 L1 upgrade conceptual design

The CMS L1 Phase-2 trigger system is designed to benefit from the new features provided by the upgraded sub-detectors to sustain a high efficiency of physics event selection in the very high luminosity regime. The functional diagram of the architecture and data flow of the Phase-2 trigger system is presented in Fig. 1.3. With the $12.5\ \mu s$ latency, not only is information from the calorimeters and muon detectors used (as in the Phase-1 system), but the information from the new tracker and high-granularity endcap calorimeter can also be included. The total output bandwidth considered is 750 kHz. Given the complexity and large data volume produced by the detector, a significant fraction of the computing of trigger quantities, such as trigger primitives completed by particle identification variables, takes place in the detector backend electronics.

The key feature of the proposed system is the introduction of a correlator layer, which implements sophisticated algorithms producing higher-level trigger objects resulting from the combination of the information of multiple sub-detectors to achieve enhanced selectivity, approaching that of the HLT. To achieve optimum flexibility of the design with the required robustness, four independent data processing paths are implemented: tracking, calorimetry, muon systems, and particle-flow techniques. This division reflects the need to generate complementary types of trigger objects to achieve the best physics selectivity. Each group of objects targets performance for specific physics requirements that can easily be optimized while providing independent trigger criteria essential to the early commissioning of the CMS detector.

Calorimeter Trigger path: A barrel calorimeter trigger (BCT) and the HGCAL backend are used to process high-granularity information from the calorimeters to produce high-resolution clusters and identification variables to be used for later processing. Outputs from the BCT, HGCAL and the hadron forward calorimeter (HF) are sent to a global calorimeter trigger (GCT), where calorimeter-only objects such as e/γ^1 candidates, hadronically decaying tau leptons (τ_h) candidates, jets and energy sums are built.

Track Trigger path: Tracks from the Outer Tracker are reconstructed in the track finder (TF)

¹Referring here to photons and electrons not distinguishable without tracking information as it is the case in the current L1.

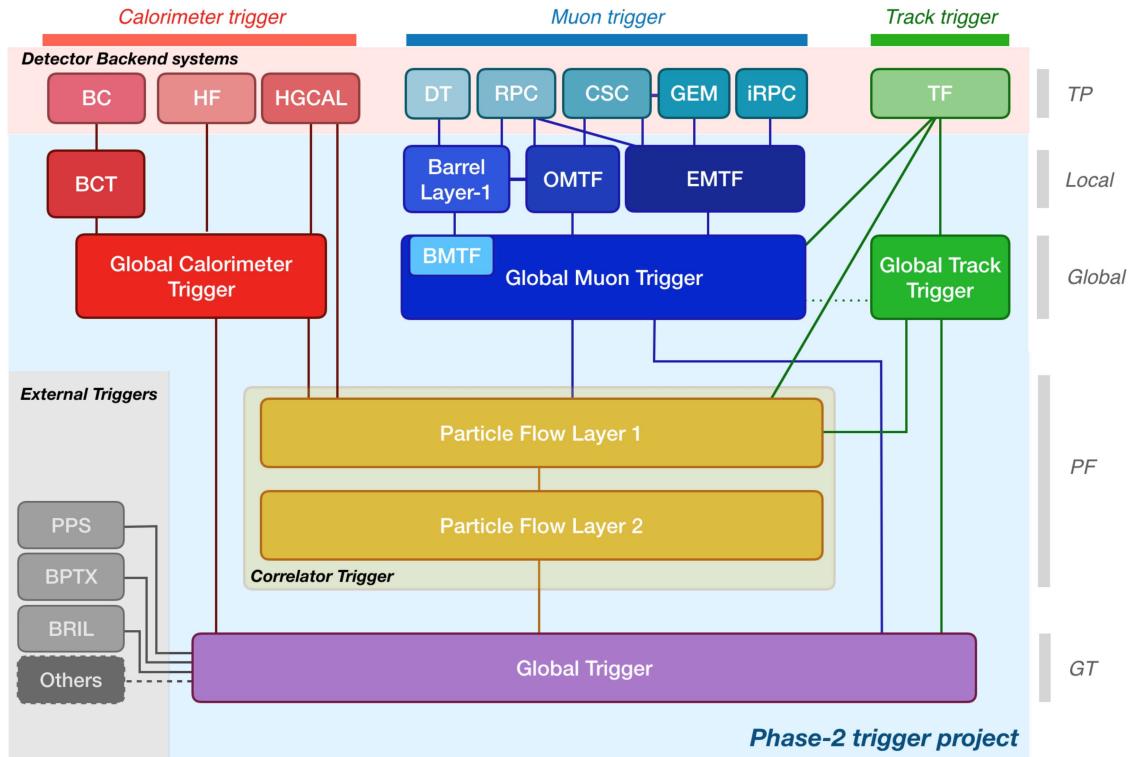


Figure 1.3: Functional diagram of the CMS L1 Phase-2 upgraded trigger design. The Phase-2 L1 trigger receives inputs from the calorimeters, the muon spectrometers and the track finder. The calorimeter trigger inputs include inputs from the barrel calorimeter (BC), the high-granularity calorimeter (HGCAL) and the hadron forward calorimeter (HF). It is composed of a barrel calorimeter trigger (BCT) and a global calorimeter trigger (GCT). The muon trigger receives input from various detectors, including drift tubes (DT), resistive plate chambers (RPC), cathode strip chambers (CSC), and gas electron multipliers (GEM). It is composed of a barrel layer-1 processor and muon track finders processing data from three separate pseudorapidity regions and referred to as BMTF, OMTF and EMTF for barrel, overlap and endcap, respectively. The muon track finders transmit their muon candidates to the global muon trigger (GMT), where combination with tracking information is possible. The track finder (TF) provides tracks to various parts of the design including the global track trigger (GTT). The correlator trigger (CT) in the center (yellow area) is composed of two layers dedicated to particle-flow reconstruction. All objects are sent to the global trigger (GT) issuing the final L1 trigger decision. External triggers feeding into the GT are also shown (more in Section 2.6) including potential upscope (mentioned as "others") such as inputs from the MTD. The dashed lines represent links that could be potentially exploited (more details are provided in the text). The components under development within the Phase-2 L1 trigger project are grouped in the same area (blue area). The various levels of processing are indicated on the right: trigger primitives (TP), local and global trigger reconstruction, particle-flow trigger reconstruction (PF) and global decision.

processors as part of the detector backend. The reconstructed track parameters and track reconstruction quality flags are provided to the trigger system to achieve precise vertex reconstruction and matching with calorimeter and muon objects. This key feature maximizes the trigger efficiency while keeping the trigger rate within the allowed budget. A global track trigger (GTT) will be included, to reconstruct the primary vertices of the event along with tracker-only based objects, such as jets and missing transverse momentum. The GTT can also be used

to propagate extra copies of tracks to any other part of the trigger system if need be. Transiting track information through GTT could allows the number of interconnections to be reduced and therefore simplify the logic on the receiving end. Dedicated studies using hardware demonstrators will help evaluate the implication of such an option in terms of latency.

Muon Trigger path: While the muon system will maintain its redundant infrastructure of detection chambers, additional stations will be installed to extend the coverage up to $|\eta| = 2.8$. The processing of trigger primitives (TP) by muon track finder algorithms is organized as in the Phase-1 system covering three separate regions: barrel (barrel muon track finder, BMTF), overlap (overlap muon track finder, OMTF) and an endcap (endcap muon track finder, EMTF). Standalone muons and stubs as well as L1 tracks are sent to a global muon trigger (GMT). A muon stub contains reconstructed local information extracted from the detector hits in each of the muon stations. It includes position, bend angle, and timing, depending on the station. Beyond the removal of muon duplicates and misreconstructed muons, the main feature of the GMT is the generation of track-matched muons and L1 tracks matched to muon stubs, the so-called tracks plus muon stubs. This is achieved through the propagation of tracker tracks through the layers of the muon system. Tracks can either be received directly from the TF or through the GTT. Interconnections established between GTT and GMT offer the possibility to provide the vertex information to the GMT algorithms if required. Motivated by robustness and latency considerations, the baseline design foresees the tracks to be sent to the GMT directly from the TF.

Particle-Flow Trigger path: The correlator trigger (CT) occupies a central role in the design. The CT implements sophisticated algorithms to produce higher-level trigger objects, applies particle identification, and provides a sorted list of objects to the global trigger. Among the algorithms considered to achieve global event reconstruction, the particle-flow reconstruction algorithm, widely used by CMS in other areas, has been chosen and its performance and implementation studied in depth. The structure of the CT is organized in two layers with a first layer, referred to as "Layer-1" producing the particle-flow candidates, which are constructed from the matching of calorimeter clusters and tracks, and a second layer, called "Layer-2", building and sorting final trigger objects and applying additional identification and isolation criteria. Layer-1 could also host a simplified version of the Pileup Per Particle Identification [15] (PUPPI) algorithm, which can be used to mitigate the degradation of the energy resolution caused by pileup. PUPPI candidates are also transmitted to Layer-2 to be used in trigger object reconstruction (see Section 1.3.2). The GTT transmits the event primary vertex candidates to the CT Layer-1. In addition, the Layer-2 can receive tracker-only objects from the GTT, track-matched muons (and/or muon track plus stubs) from the GMT, as well as calorimeter-only objects from the GCT, and apply further isolation criteria, for example. Isolation of muons may also be calculated in the GMT, depending on available resources.

Global Trigger: Outputs from the GCT, GMT, GTT, and CT are combined in the global trigger (GT), which calculates a trigger decision based on a menu of algorithms. The GT has resources to evaluate sophisticated correlation variables among various types of objects to increase the selectivity. The Level-1 Accept signal is transmitted to the Trigger Control and Distribution System (TCDS), which distributes it to the detector backend systems, initiating the readout to the DAQ. The GT provides the interface to external triggers (more in Section 2.6) such as triggers for the precision proton spectrometer (PPS), beam position and timing monitors (BPTX), and luminosity and beam monitoring (BRIL) detectors. A potential upscope of the trigger would include other external inputs, discussed later in this paragraph.

Scouting System: The 40 MHz Level-1 Scouting System captures part or all of the intermediate

trigger data streams from the different trigger layers using spare optical outputs of the different processing boards, and enables trigger-less analysis of Level-1 objects at the bunch-crossing rate. The scouting input boards, hosted on standard computers, perform zero suppression and local preprocessing of trigger data, and feed it to host memory. Subsequent processing is carried out asynchronously in a distributed fashion across a high-performance interconnect, and produces feature streams that can be stored and used for analysis or monitoring. These feature streams only contain high-level information derived from the trigger objects, and thus retain the limitations of the Level-1 reconstruction in terms of resolution and purity. Thanks to the limited size of the information stored per bunch crossing, they can however be stored at high-rate and also allow analysis of multi-bunch-crossing correlations. In particular, capturing the inputs and outputs of the Global Trigger at 40 MHz will enable detailed diagnostics of the trigger system at large. It will also enable the detection of anomalies in quasi-real-time in most of the lower level systems, ease the testing of novel GT algorithms, and provide alternative luminosity measurements based on physics processes. Besides its diagnostic and monitoring capabilities, the Scouting System will be used to study physics channels lacking a well defined signature for effective Level-1 rate reduction, in cases where the full detector acceptance and/or full detector resolution is either already approached with L1 primitives and L1 algorithm outputs or not strictly necessary for a competitive measurement.

Architecture overview: This architecture proposal is the result of studies conducted to optimize the number of processing boards and board interconnections, as well as latency, while complying with the need for flexibility and robustness. A division of labor achieved through the implementation of global triggers (GCT, GMT and GTT), allows the reduction of the FPGA resources required to implement the particle-flow algorithm in the CT. Enough headroom is available to further optimize the algorithms, as was the case during the Phase-1 L1 trigger operation. The physical implementation of the architecture is described in Chapter 5. Regional and time-multiplexed² architecture options are both evaluated. The design exploits the availability of new hardware technologies to ensure the computing power along with high-speed data transfer (up to 28 Gb/s) to provide a global detector view. The system will handle a huge amount of input data bandwidth (63 Tb/s compared to \sim 2 Tb/s today). Inspired by the technological choices made during the Phase-1 upgrade of the L1 trigger, the data processing units are mainly designed as generic stream-processing engines instead of custom-designed processors. Generic processors can be used to perform any task within the system, allowing to consider original trigger architectural options adapted to the requirements inherent of a flexible Correlator design. The electronics boards under development feature large FPGAs such as Xilinx VU9P equipped with 28 Gb/s transceivers. A small level of customisation remains to address specific requirements such as lighter processing tasks or fast memory access, etc. More on this topic is described in Chapter 6. Results of hardware demonstrators are included as well. The paragraph 1.3.2.1 describes the inputs to the L1 Phase-2 trigger system and their use by the sophisticated trigger reconstruction algorithms under development.

Potential future upscopes: The design of the Phase-2 L1 trigger system provides provision for more input bandwidth and latency for later potential upscopes, for example from the MTD [12]. Timing information from MTD could help selecting displaced signals and complement the timing information provided by the calorimeters and muon systems.

² In a time-multiplexing (TM) approach, N_{TM} processors run identical algorithms, on different events. The same data may also be sent to multiple boards, which run different algorithms as described in Ref. [16].

1.3.2 Trigger algorithms for the HL-LHC

The envisaged L1 system will more closely replicate the full offline object reconstruction, instead of making use of simple subsystem variables, to make a better optimized selection. The Phase-2 trigger algorithms foreseen can be used to reconstruct a large variety of objects: standalone, which are reconstructed from single detector information (including tracker-only objects), standalone matched to L1 tracks, and particle-flow [13]. The trigger decision can rely on the complementarity of these objects to achieve the best possible efficiency while keeping the trigger rate under control. The trigger system described in this document aims to achieve the most optimized selection of collision events through a global event reconstruction that has proven to be efficient to reach the highest sensitivity in offline data analyses. Given the physics drivers listed in Section 1.1, the trigger object requirements are not only driven by the need to maintain physics selection thresholds to match those of Phase-1, but also by having to provide the selection of exotic signatures, including displaced objects. The algorithms described here are developed keeping in mind potential features that could help expand the physics reach while providing robustness in view of effects introduced through detector ageing, for example. The algorithm implementation in firmware (see Chapter 3) greatly benefits from the introduction of High-Level-Synthesis software that could be used to design advanced machine learning trained variables or even iterative processes in the core of the trigger system. This section provides an overview of the baseline algorithms that have been developed with the minimum requirement of meeting the challenges of the HL-LHC. New developments and further improvements are certainly foreseen beyond the scope covered by this TDR. A short description of the trigger primitives inputs is reported as well. In view of the potential ultimate configuration of the HL-LHC machine, all studies were carried out with simulated data samples containing 200 pileup events, unless specified otherwise.

1.3.2.1 Inputs to the L1 trigger

The Phase-2 upgraded calorimeter trigger will benefit from the enhanced granularity provided by the upgraded barrel [9] and endcap calorimeters. The ECAL barrel (EB), Very Front-End (VFE), and Front-End (FE) electronics will be fully replaced with boards equipped with high-speed optical links (5 Gb/s, compared to 800 Mb/s in the Phase-1 system) in order to transmit the single crystal energy information for triggering purposes, achieving a granularity increased 25 times with respect to the Phase-1 system. The HCAL barrel (HB) on-board electronics readout will be replaced to retrieve the depth information, providing 4 times more granular information with respect to the current system. Trigger primitive information from both ECAL and HCAL will carry timing information with an expected resolution of 30 ps and 500 ps for objects with $p_T > 50$ GeV, respectively. Similar trigger primitives information as in Phase-1 is expected from the hadron forward (HF) calorimeter. The endcap calorimeters will be replaced by the HGCAL instrumented with over 6 million readout channels [10]. It is a 3D sampling calorimeter with alternating layers of silicon sensors/scintillator and copper/lead absorbers. The goal is to achieve unprecedented spatial resolution and shower separation to optimise the matching with tracks. About half of the 50 layers (28 electromagnetic and 22 hadronic), representing a total of roughly 990 000 trigger channels, can be exploited for triggering purposes, hence increasing the granularity by a factor of more than 500 compared to the Phase-1 system. The HGCAL trigger primitives consist of the summed energy of 4 adjacent channels. Channels correspond to a readout cells of roughly 1×1 cm 2 paving each calorimeter layer. The trigger cell information are grouped and sent to the backend electronics responsible for producing the trigger primitives. This system is organised in two layers where the first layer is used to calibrate and reorganise the data in time-multiplexed fashion and the second layer reconstructs

3D clusters. The 3D clustering algorithm is performed on a full depth view of the detector, hence exploiting the longitudinal and transverse shower profiles specific to each physics object. Clusters are transmitted to the Phase-2 L1 trigger along with variables to discriminate electromagnetic, hadronic and pileup-induced energy deposits.

One of the key features of the Phase-2 trigger upgrade system is the addition of charged particle track reconstruction. Standalone reconstructed trigger objects will suffer from high trigger rates and, therefore, reconstruction algorithms will greatly benefit from matching with tracks. The higher p_T resolution of the tracks reduces contributions to the rate coming from p_T below the trigger threshold. Tracking information can also be used to compute photon and lepton isolation, improve jet and energy sum trigger reconstruction, provide vertexing, estimate the level of pileup (not only calorimeter-based estimators as in Phase-1), etc. CMS plans to replace both the silicon Pixel and Strip tracking detectors. The Outer Tracker on-detector electronics will generate "stubs" that the track finder system will associate into track candidates to be transmitted to the L1 trigger system. Stubs are computed from closely-spaced silicon-sensor modules composing each tracker layers. The bending of tracks between each side of the modules is used to discard hits from low p_T tracks and reduce, on detector, the hit rate to a manageable level. The track finder coverage in pseudorapidity extends up to $|\eta| = 2.4$. The overall latency is estimated to 5 μs (including 1 μs for data transmission). A total of around 1000 tracks with $p_T > 2 \text{ GeV}$ can be reconstructed per event at 200 pileup in the Phase-2 LHC environment. The track finder trigger primitives are reconstructed with a hybrid algorithm performing a Kalman filter fit on associated stubs to a given seed or "tracklet" (also formed from adjacent layers stubs). These L1 tracks are transmitted to the Phase-2 L1 trigger and contain all the fitted parameters (p_T , η , ϕ , d_0 , and z_0), along with the track quality. Another approach, referred to as "Extended L1 tracking" in the text, could be used to reconstruct displaced trajectories originating from potential BSM signatures. The principle lies in the combination of three layers and/or disks, called triplet seeds, that can be formed without beamspot constraint in addition to the tracklet seeds. While not in the current baseline, studies are being conducted to evaluate the possibility of such an algorithm even for standard objects.

Although most of the muon detectors will remain in place, their associated readout electronics will be replaced, providing finer information and improved timing from 25 ns to 1.5 ns. This information can then be used to develop more sophisticated trigger primitives algorithms with enhanced performance, close to present offline reconstruction. Improved DT trigger primitive generation is based on hit information in each layer and correlations among super-layers allowing enhanced bunch crossing assignment. Each RPC hit will carry time information with a granularity of one sixteenth of a 25 ns BX period. Trigger primitives from DT and RPC are combined into super-primitives to exploit the redundancy and complementarity of these two sub-detectors, thus achieving better performance and being resilient to ageing-induced effects. The CSC on-chamber electronics, located in the inner rings will also be replaced to handle higher trigger rates and potentially improve robustness of the trigger primitive generation. A larger improvement in the bending direction resolution could be achieved by combining CSC and GEM trigger hits. With improved RPC chambers, the full redundancy of the muon spectrometers is performed in the forward region of the detector. Trigger primitives are computed from clustered strip hits providing excellent timing resolution. The GEM chambers are essential to achieve a precise measurement of the muon bending angle in the forward region $1.6 < |\eta| < 2.8$. GEM trigger primitives are obtained from the clustered information generated by dedicated trigger pads mounted on the detectors. These TPs can then be combined with CSC or directly transmitted to the Phase-2 endcap muon track finder (EMTF). Both the GE1/1 chambers installation and the CSC on-chamber electronics replacement are being performed

during LS2.

1.3.2.2 Triggering on electrons and photons

Many standalone electron and photon trigger reconstruction techniques are being investigated to optimise both the response and the position resolution for the purpose of achieving the highest possible track matching efficiency. The proposed algorithms utilize both identification criteria and isolation variables based on calorimeter, as well as tracking information to reduce the background level. Given the intense running conditions foreseen, the algorithms are designed to be pileup resilient.

The electron finder in the barrel region uses the crystal information from the ECAL. A 5×3 crystal matrix ($\Delta\phi \times \Delta\eta = 0.087 \times 0.052$) is used to define the maximum size of the electron footprint in the ECAL. As in the Phase-1 algorithm approach, the extension in ϕ is motivated by the necessity to recover energy lost through bremsstrahlung. An improved position resolution is achieved using a weighted-energy sum around the seed crystal (the seeding threshold used is $E_T > 1$ GeV). Extra shower shape features are used as identification criteria and the matching of the clusters with tracks is performed using an extrapolation to the ECAL surface.

The starting point of the electron reconstruction algorithm in the endcap region is the cluster reconstructed in the backend electronics of the HGCAL. Further identification of the electromagnetic object is performed through a multivariate approach optimized to exploit the input variables transmitted from the HGCAL. Dedicated boosted decision trees (BDTs) are trained on signal and background to achieve an optimal signal efficiency while rejecting pileup-induced clusters. Bremsstrahlung recovery is performed as well as an energy calibration of the final e/γ candidate. The availability of the tracking information facilitates the reconstruction of isolated photon candidates. The reconstruction of electron tracks with the TF is inefficient because of the radiation induced by bremsstrahlung throughout the tracker material. The extended tracking, originally designed to reconstruct displaced trajectories, could help recover track reconstruction efficiency.

1.3.2.3 Triggering on jets, taus and energy sums

Triggering on hadronic signals has always represented a challenge for detectors operating in an intense hadronic environment. Algorithms developed for the Phase-1 Level-1 trigger system are optimized to provide thresholds adequate for physics using calorimeter-only information [6]. The rate of such triggers measured on Run-2 collision data and, in particular, trigger conditions based on missing transverse energy, display a strong dependence on the level of pileup events and, to a certain extent, the filling scheme used by the machine. Dedicated pileup mitigation techniques have been developed to maintain the trigger rate to acceptable levels, preventing a dramatic rise of thresholds. The level of pileup expected at the HL-LHC will seriously impact the performance of hadronic triggers. This motivates the detailed studies conducted here to provide pileup mitigation algorithms exploiting the full capabilities of the Phase-2 detectors. As calorimeter-only algorithms are expected to have high thresholds, complementary approaches are proposed with tracker-only information, track-matched calorimeter objects and higher-level objects reconstructed with particle-flow and PUPPI inputs. Calorimeter-only jet finding algorithms use barrel ECAL and HCAL information, endcap HGCAL and forward HF information. Although various configurations were considered, a simple square geometry of 7×7 trigger towers (see Section 3.6 for the trigger tower definition) around a local maximum gives acceptable performance while keeping the pileup contribution to a minimum. More details are provided in Section 3.6. The jet window definition corresponds approximately to the cone size of 0.4 used by the offline anti- k_T algorithm [17]. Similarly to the

Phase-1 algorithm, a tower-by-tower pileup correction depending on the level of pileup and η is applied prior to jet clustering.

Tracker-only jet finding is performed on a set of tracks from the track finder [8] passing purity requirements to keep the trigger object resilient to pileup. Track clustering makes no use of the primary vertex information. In order to optimize the latency, primary vertex computation and jet clustering are performed in parallel. Note that more robustness against pileup is obtained by considering a smaller z -range of tracks from the main interaction point. The clustering of tracks in the $\eta - \phi$ plane is performed using a nearest-neighbor approach in two one-dimensional steps. The maximal jet size corresponds to $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ while the jet p_T is computed as the sum of each track p_T associated to it. The extended tracking mentioned previously could provide displaced tracks that could be clustered to reconstruct displaced jet objects. The performance of such objects is studied in this document as well. The particle-flow based jet finding consists of building jets from particle-flow candidates grouped into pseudo trigger towers (equivalent to 0.083×0.087 in $\eta - \phi$), which are then clustered into a 7×7 tower window around a local maximum. The jet momentum is computed as the sum of the objects' momentum within 7×7 window and the jet position is that of its seeds (η, ϕ) coordinates. Various pileup mitigation techniques were investigated based on jet energy correction scale factors or pileup energy subtraction. Each jet finding algorithm has a dedicated energy calibration, detailed in Chapter 3. The jet p_T and direction can then be used to compute H_T (scalar sum of jet p_T) and H_T^{miss} (vector sum of jet p_T).

The benefit of developing a dedicated hadronically decaying tau (τ_h) reconstruction and identification algorithm has been fully demonstrated during Run-2 of the LHC [18]. As for the jet finding algorithm, calorimeter-based τ_h are built from trigger towers. Given that τ_h are narrow jets and that several decay products may be producing more than one cluster spatially separated along the ϕ direction due to the magnetic field, a 3×5 tower window ($\Delta\eta \times \Delta\phi = 0.261 \times 0.435$) is chosen to optimise p_T resolution. Similarly to the Phase-1 algorithm, a window of 7×7 trigger towers around the core objects defines the isolation region that is used to identify further the τ_h candidate while maintaining the rate under control. Although this approach performs well within the entire calorimeter acceptance, the enhanced granularity of the HG-CAL detector allows the implementation of advanced identification techniques exploiting the τ shower characteristics. The shower profile being different to that of pileup-induced particles, stringent discrimination can be achieved³. A dynamic clustering of 3D-clusters gives optimum response, while trained BDTs can provide dedicated energy calibrations for each of the τ decay modes. Algorithms matching the current offline reconstruction of τ_h , called "hadron plus strips" [19] have also been developed. These algorithms have been studied considering both PUPPI or particle-flow particle candidates. In the case of tau leptons, the isolation variable is essential to further reduce the background. The track isolation is calculated from tracks that are within $|\Delta z| < 1.0$ cm of the τ_h , where z is the distance of the tracks along the beam line from the interaction point. When PUPPI is utilized, this can further be enhanced by including the neutral candidates as well. Identification can also rely on neural network approaches that, despite the complication of firmware implementation, give excellent performance when combined with the PUPPI-based approach.

Triggering on missing transverse energy (E_T^{miss}) is a particularly challenging task for detectors operating in hadronic environment, especially when the average expected pileup is 200. This quantity is a key input for many signatures, including beyond Standard Model processes, in the L1 trigger. The use of L1 tracks is essential to achieve manageable rates for moderate

³To a certain extent, the HCAL longitudinal segmentation could be used to perform similar discrimination of pileup-induced jets (more in this discussed in Section 2.2.2)

thresholds. The algorithms pursued are either tracker-based or PF-based. The tracker-based approach considers tracks originating from the primary vertex and applying dedicated selection to reject misreconstructed tracks. The rate is considerably reduced using this approach. With particle-flow, the information of all sub-systems is used and further mitigation of pileup contributions is obtained with PUPPI. Thresholds applied to particle-flow and PUPPI inputs in various η regions can be adjusted to optimize performance.

1.3.2.4 Triggering on muons

The overall structure of the muon system for Phase-2 remains similar to the current one. It is composed of three partially-overlapping sub-detectors (CSC, DT, and RPC), whose signals are combined to reconstruct muons and measure their transverse momenta. Additional muon stations, such as iRPC, GEM and ME0, are installed in the forward region to extend the acceptance to $|\eta| = 2.4$ and $|\eta| = 2.8$, respectively. Following the approach of the Phase-1 trigger upgrade, the reconstruction of standalone muons uses information from all available sub-detectors simultaneously to build tracks in three distinct pseudorapidity regions, improving the muon reconstruction and increasing signal efficiency while reducing background rates. Given the improved sub-detector electronics readout for Phase-2, the muon chambers will provide finer information along with precise timing (~ 1.5 ns) that can be exploited by the muon track finding algorithms. Each track finder uses an optimised track reconstruction algorithm and p_T assignment logic, and assigns a track quality corresponding to the estimated p_T resolution. Similarly to the existing system, the BMTF uses DT and RPC trigger primitives to reconstruct segments merged to obtain a muon candidate. The RPC fired strips are clustered before being used by the muon track finders. A track finding approach based on a Kalman Filtering technique called Kalman barrel muon track finder (KBMTF) has been developed and tested on Run-2 collision data. As tracks can be reconstructed by KBMTF with and without a constraint forcing them to originate from the primary vertex, displaced muons can be reconstructed with acceptable p_T resolution resulting in higher acceptance. In the overlap region, the OMTF receives data from DT, RPC and CSC stations and reconstructs tracks by associating hits, using generated patterns from simulated events. This naive Bayes-classifier approach identifies the most likely muon p_T . The muon endcap track finding algorithms exploit the information from up to 12 muon stations. In addition to CSC and RPC trigger primitives, the Phase-2 EMTF++ system proposed for this upgrade receives information for GEM (including ME0) and iRPC detectors. The standalone reconstruction algorithm looks for correlated CSC trigger primitives through multiple stations compatible with a muon track corresponding to predefined patterns. Consistent RPC primitives are associated to this track candidate and a trained deep neural network (DNN) for p_T assignment, with and without beam constraint, is implemented. GEM detectors cover $|\eta| > 1.6$ (GE1/1) and $|\eta| > 1.8$ (GE2/1). For $|\eta| > 1.6$, EMTF++ can use combined GEM-CSC bend angle. The muon reconstruction in the forward region beyond 2.1 in pseudorapidity uses the position and bend angle from ME0 (up to $|\eta| < 2.8$) and CSC (up to $|\eta| < 2.4$).

The availability of tracks from the Outer Tracker allows another category of muons, with increased acceptance at low p_T or originating from regions with limited detector coverage, to be considered. The matching of standalone muons and tracks is performed optimally in each pseudorapidity region, so that, as in the offline or Phase-1 HLT cases, misreconstructed muons are reduced, and the p_T measurement accuracy is improved. Another complementary approach consists of propagating the tracks from the Outer Tracker into the muon detectors and associate stubs from at least two layers of the muon stations. This algorithm shows optimum performance for a large variety of physics signals while maintaining efficiency and providing robustness against detector aging. The possibility to correlate tracking and muon stubs information is used to produce trigger objects adequate to identify heavy stable charged particles

(HSCPs). Given the particularity of this signal, the candidate L1 track is matched to muon stubs from the same event or subsequent ones. Other combinations of L1 tracks and muon stubs can form topological objects appropriate for the selection of $\tau \rightarrow \mu\mu\mu$ events, for example.

1.3.2.5 Reconstruction of the primary vertex and particle-flow objects

The reconstruction of the primary collision vertex is an essential ingredient of the physics object reconstruction. The provision of tracks from the track finder allows the main interaction point and its associated tracks to be identified. Various algorithms were developed and tested on different types of final states. The studies considered busier events such as $t\bar{t}$ producing jets with high track multiplicity and processes with less tracks such as $Z \rightarrow \mu\mu$. The algorithms range from a simple histogram-based approach to more sophisticated ones based on machine learning techniques or even iterative implementations. Firmware considerations help select the optimum one. While performance can reach 85% of vertex identification efficiency for $t\bar{t}$ events, its lower score for $Z \rightarrow \mu\mu$ can be recovered if multiple vertex candidates (at least 3) are considered. The consequence in terms of firmware resources needs careful studies.

The availability of tracking information and the enhanced calorimeter granularity at the Level-1 trigger allow us to contemplate the exciting possibility to propose algorithms matching the performance of those implemented at the HLT. The benefit of introducing a full event description (such as the one obtained through the particle-flow reconstruction algorithm [13]) have been demonstrated with the Phase-1 HLT system. The Phase-2 trigger Correlator processor (see Fig. 1.3) will be used to combine all detector information to produce a list of candidates from which higher-level trigger objects are constructed, such as identified prompt leptons, photons and jets, as well as global quantities such as missing transverse energy or hadronic transverse energy sum (H_T) at L1. Dedicated energy calibration factors are derived for each sub-detector input leading to an improved response for single particles. In addition to the particle-flow algorithm, a simplified version of the offline PUPPI algorithm can be implemented. PUPPI relies on the knowledge of the vertex position but can use other mechanisms to discriminate against pileup outside the tracker volume. In addition to the particle-flow candidates, PUPPI particle candidates can also be used to reconstruct complex objects or isolation variables. The firmware design is simple and able to process all input objects in parallel for pipelined execution on FPGA, in contrast to the sequential approach used offline, adapted to CPU. The firmware structure chosen has direct consequences on the correlator architecture organization and, in particular, the regionalization of the detector to perform optimal processing with a short and fixed latency.

1.3.2.6 Global Trigger Algorithms

Global trigger algorithms refer here to algorithms based on correlations among physics objects or using more sophisticated variables such as invariant masses. The ability to trigger on such variables has significantly enhanced the selectivity of the Phase-1 trigger system and therefore this feature is foreseen for its Phase-2 upgrade. Naturally, these algorithms are implemented in the GT, where tailored triggers for specific physics analyses are provided. In addition, the GMT, GCT and GTT systems are also used to generate quantities based on information only available upstream of the correlator trigger and the GT. Specific objects or variables can be propagated through direct connection and combined with other physics objects. For example, the GTT can compute invariant masses of combinations of tracks passing certain quality cuts to trigger on particular light resonances. Provision is made to eventually receive timing information from MTD to flag out-of-time physics objects predicted in beyond Standard Model processes. Other global quantities such as centrality, commonly used for heavy ion triggering purposes,

can be implemented. Machine learning approaches are also pursued as alternatives to simple cut-based trigger conditions on object combinations. Specific software such as hls4ml [20] has rendered this type of techniques implementable in FPGAs. Preliminary results on selecting signatures such as the production of the Higgs boson through VBF, with the $H \rightarrow b\bar{b}$ and invisible decay ($H \rightarrow \text{inv}$) channels, show significant improvements with respect to standard triggers.

1.3.2.7 Examples of the Phase-2 L1 trigger algorithms performance

In this Section, examples of the Phase-2 Level-1 trigger algorithm performance are presented. The performance of all algorithms proposed are presented in Chapter 3. Figure 1.4 (left) displays the H_T trigger efficiency for simulated $t\bar{t}$ events with 200 average pileup events as obtained for a fixed rate⁴ of 10.5 kHz. The performance of H_T computed with jets reconstructed with the PUPPI algorithm is compared to calorimeter-only, tracker-only quantities. The PUPPI-based H_T approach outperforms the other algorithms by optimally exploiting both the tracking information and enhanced detector granularity. The muon trigger performance presented in Fig. 1.4 (right) demonstrates the performance of the barrel standalone muon trigger reconstruction algorithm. This algorithm can provide a sustainable rate for muons originating from long-lived particles that appear as displaced trajectories in the detector. The impact parameter in the transverse plan is referred to as L1 D_{xy} .

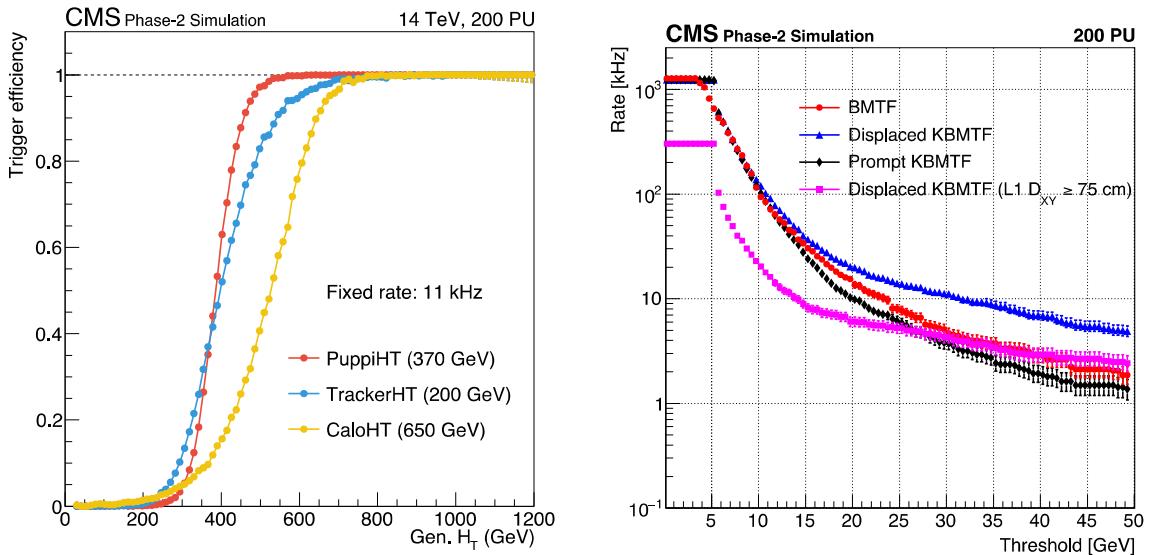


Figure 1.4: Left: The signal efficiency is shown for $t\bar{t}$ events selected by H_T triggers utilizing different sets of inputs at a fixed selection rate of 10.5 kHz. The performance of the H_T trigger quantity computed with PUPPI jets, tracker-only jets and calorimeter jets are compared. Right: Single muon rate as a function of the L1 muon p_T for the Phase-1 BMTF (red), the Phase-2 prompt (black) and the displaced KBMTF (blue) reconstruction algorithms. The displaced KBMTF muon algorithm takes the maximum of the vertex constrained and unconstrained p_T . Displaced muon rate for $L1 D_{xy} \geq 75$ cm is also shown (magenta) as a function of the L1 muon p_T threshold. D_{xy} refers here to L1 trigger reconstructed quantity.

⁴The method used to compute the rates throughout this document is explained in Section 3.1.

1.4 Physics reach of the Level-1 Phase-2 Trigger

The HL-LHC will provide the experiments with the largest dataset ever collected. Prospect studies of the direct and indirect discovery potential offered by HL-LHC have been documented in a "Yellow Report" [1]. For some of the studies reported in this document, it is assumed that the selection criteria used in Run-2 analyses can be maintained, while for others, larger physics acceptance compared to the current one was considered. Consequently, the analyses reported in Ref. [1] manifest how crucially the HL-LHC physics program relies on the capabilities of the trigger system to maintain comparable physics acceptance to that of Phase-1 and to collect further enriched datasets covering an extended phase-space, not reachable previously. These requirements are discussed in Section 1.4.1. The extended physics reach achieved through the new features of the Phase-2 L1 trigger system is presented in Section 1.4.2. The gain in acceptance with respect to reference Phase-1 physics channels is described, followed by a presentation of the improvement expected on several SM measurements and BSM searches. Throughout this section, benchmark physics channels were chosen for these studies. Targeted trigger thresholds delivered by the upgraded Phase-2 Level-1 trigger system are specified.

1.4.1 Preserving the CMS discovery potential at the HL-LHC

In order to substantiate the capabilities of the Phase-2 L1 Trigger to maintain the current physics acceptance in the extreme scenario of 200 average pileup events per crossing, a simplified L1 trigger menu, covering most of the CMS Run-2 physics selection strategy, was developed (see Chapter 4). The development of the trigger menu represents an important step in demonstrating the successful implementation of the Phase-2 trigger strategy. A typical L1 trigger menu is composed of a set of criteria applied to single-, multi-, and cross-objects targeting a variety of final states. Typical criteria include lower thresholds on the object's transverse momentum and conditions on its pseudorapidity. Trigger algorithms include: single lepton triggers, double lepton and photon triggers, hadronic triggers (based on jet p_T , scalar sum of jet p_T (H_T) or missing transverse energy (E_T^{miss})), cross lepton triggers, cross hadronic-lepton triggers, specific soft leptons triggers (targeting B-physics signals) and specific hadronic triggers (targeting VBF topologies). The menu presented in Chapter 4 contains an ensemble of such algorithms, with p_T thresholds set to the ones that were used at the end of Run-2. A similar exercise was performed previously in Ref. [21], with equivalent Phase-2 LHC running conditions but using the Phase-1 trigger design and algorithms, hence not including tracking capabilities. The total Level-1 trigger rate attained was approximately 1500 kHz for 140 pileup events and almost 4000 kHz for 200 pileup events, well beyond the 750 kHz originally foreseen. This study alone motivated the inclusion of tracking information into the Phase-2 Level-1 trigger. What follows is not intended to summarise the menu developments but to illustrate the capability of the upgraded trigger to provide low enough thresholds for a few selected CMS Phase-2 physics benchmarks. To emphasise further the need for this upgrade, the L1 trigger thresholds used in Run-2, and allowed by the proposed system, are compared to the reference thresholds the current trigger would need to apply to reach the same target rate under HL-LHC running conditions.

1.4.1.1 Final states requiring lepton trigger algorithms

- **Higgs boson associated production with a W/Z boson (HW/Z).** Muons and electrons from the decay of the W/Z bosons are particularly important to trigger the selection of events in which the Higgs boson is produced in association with a W or a Z boson and decays into hadronic objects with soft p_T . Triggers targeting this final state generally do not include conditions on the hadronic objects produced through the Higgs decay, given the large associated background rate. As a result of their

clear signature, lepton triggers are of utmost importance in a hadronic environment. Hence, triggers based solely on leptons allow for an inclusive selection that can target a broad range of final states; they have been successfully used for several Phase-1 analyses. This is particularly relevant for searches for physics beyond the SM, such as searches for Higgs boson decays into light pseudoscalars that subsequently decay into hadrons. The analyses targeting such rare processes see their acceptance significantly limited by any increase of the lepton p_T thresholds of the single- or double-lepton triggers used to select these events.

- **Single top-quark or top-quark pair production ($t\bar{t}$)**. Precision measurements of the top-quark properties are an important test of the Standard Model. In particular, flavor changing neutral current (FCNC) couplings of the top-quark, which are highly suppressed in the SM and for which the current experimental constraints are far from approaching the SM predictions, offer interesting opportunities to probe BSM models, for example via studies of single top-quark production. Other processes involving the top-quark such as the associated production of the Higgs boson with a pair of top-quarks (tth) or the production of four top-quarks, would equally rely on lepton triggers.
- **Higgs boson pair production $HH \rightarrow \tau\tau bb \rightarrow \ell\nu_\ell\nu_\tau\tau_h\nu_\tau bb$** . The study of Higgs boson pair (HH) production is a crucial test of the Standard Model as it grants direct experimental access to the Higgs boson self-coupling and thus the characteristics of the Higgs potential itself. The most recent prospect for the study of HH production at HL-LHC with the CMS detector results in an expected significance for the Standard Model signal of 2.6σ [1], assuming the capability of the Phase-2 L1 trigger to achieve comparable thresholds as during LHC Run-2 operations. The measurement of the Higgs boson self-coupling requires the detection of Higgs bosons pairs with an invariant mass close to the kinematic production threshold of $2 \times m_H$, where a deviation of the self-coupling from its SM value gives the largest variations. This results in soft objects in the final state where leptons can provide a clean signature for triggering. In the $HH \rightarrow \tau\tau bb$ state, one of the leading channels for HH production observation at the HL-LHC, final states where the $\tau\tau$ system decays to $\ell\tau_h$ and neutrinos can give access to the low HH invariant mass region for the self-coupling determination and can be recorded with single-lepton triggers.

Figure 1.5 (left) shows the inclusive p_T distributions of electrons and muons coming from the $HH \rightarrow bb\tau\tau \rightarrow bb\tau_\ell\tau_h$ ($\ell = e, \mu$) process, single-lepton $t\bar{t}$ and single-top production at a center-of-mass energy of 14 TeV; τ_ℓ refers to leptonically decaying τ leptons. The single electron/muon trigger thresholds allowed by the upgraded Phase-2 L1 trigger are compared to the ones that would be expected without the use of tracking information when reconstructing those objects. By definition, the meaning of these trigger thresholds is that the efficiency of the single-lepton trigger, as a function of the p_T of offline reconstructed leptons, reaches 95% of its plateau efficiency at the quoted value. The target thresholds for single leptons of 15 GeV (muon) and 28 GeV (electron) can be sustained with the Phase-2 algorithms, which associate L1 tracks with standalone calorimeter or muon objects, while with standalone objects only, these thresholds would have to be increased to 24 GeV and 62 GeV respectively, for an equivalent rate of 42 kHz (muon) and 24 kHz (electron). The significant gain in acceptance for the aforementioned analyses is clearly visible in Fig. 1.5 (left). This gain is particularly essential to the search for Higgs boson pair production $HH \rightarrow bb\tau\tau \rightarrow bb\tau_\ell\tau_h$ where the lepton from the τ decay displays a soft p_T spectrum. Additionally, Fig. 1.5 (right) illustrates how the expected exclusion limit at 95% confidence level (CL) on the FCNC $t \rightarrow ug$ branching fraction would degrade with an

increased lepton p_T threshold used in the analysis [1]. See Section 4.2 for more details.

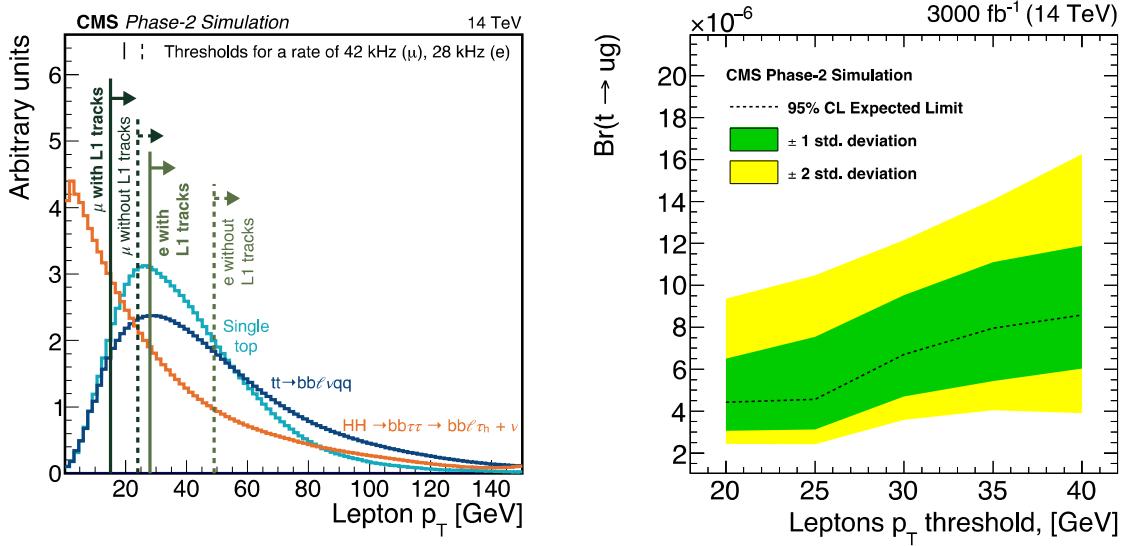


Figure 1.5: Left: Simulated distributions of transverse momentum of the electron and muon produced in HH , single top-quark, and the semileptonic decays of $t\bar{t}$. The vertical lines correspond to the offline p_T thresholds at which the single object trigger efficiency reaches 95% of the efficiency plateau. The solid vertical lines correspond to the trigger thresholds provided by the Phase-2 L1 trigger system (at 200 pileup) matching the thresholds currently deployed by the L1 menu for Run-2. The dashed vertical lines correspond to the trigger thresholds required to achieve the same rate using trigger algorithms that do not make use of L1 tracks. Right: The expected exclusion limits at 95% CL on the FCNC $t \rightarrow ug$ branching fraction as a function of the offline leptons p_T threshold.

1.4.1.2 Final states requiring double-photon trigger algorithms

- **Higgs boson pair production ($\text{HH} \rightarrow \gamma\gamma b\bar{b}$)**. As shown in Ref. [1], one of the most sensitive decay channels to access di-Higgs production is $\text{HH} \rightarrow \gamma\gamma b\bar{b}$, where the event selection relies on a double-photon trigger with thresholds as low as those used in Phase-1. Harvesting these rare events would contribute to obtaining evidence of the HH process, which constitutes one of the main goals of the CMS Phase-2 physics program.
- **Higgs boson decay into photons ($H \rightarrow \gamma\gamma$)**. This final state benefits from the complete reconstruction of the Higgs boson kinematics and from the clean signature of the diphoton invariant mass. Hence, this channel is particularly suited to perform the measurement of the Higgs boson differential cross sections and in particular of the Higgs boson p_T spectrum, including the low p_T regime. During Phase-1, this major discovery channel relied on the double-photon trigger. The baseline strategy to pursue this analysis remains similar to Phase-1 and therefore the trigger thresholds applied should sustain a selection as inclusive as possible.

Figure 1.6 displays the inclusive p_T spectrum of the sub-leading photon in single and double Higgs boson final states. In the case of Higgs boson pair production, one of the Higgs bosons decays into photons. The trigger threshold on the sub-leading leg of the double-photon trigger allowed by the upgraded Phase-2 L1 trigger is compared to the one expected without any tracking information used. The Phase-2 photon objects reconstruction exploits both

the full granularity of the Phase-2 calorimeters and the L1 tracking information. The latter allows track-isolation requirements to be used to select isolated photons, which gives improved performance compared to isolation criteria based on calorimeter information alone as implemented by the Phase-1 L1 trigger. Exploiting the new features of the Phase-2 L1 trigger, a threshold of 12 GeV can be implemented while a threshold of 20 GeV would be needed without the tracker information, for an equivalent rate of 50 kHz (see Section 4.2 for more details).

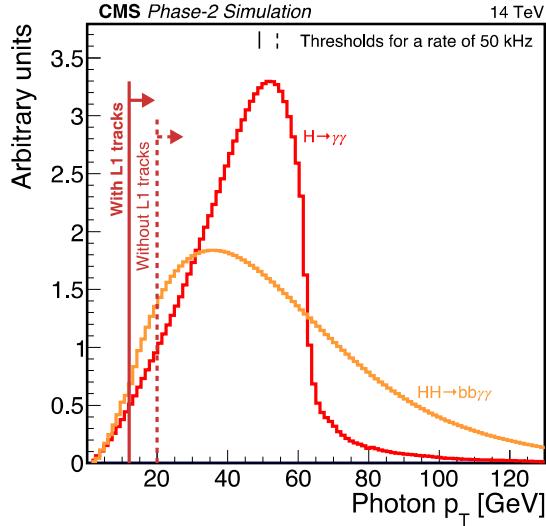


Figure 1.6: Simulated distribution of transverse momentum of the lowest p_T photon from $\text{HH} \rightarrow \gamma\gamma b\bar{b}$ and $H \rightarrow \gamma\gamma$ processes. The vertical lines correspond to the offline p_T thresholds at which the single object trigger efficiency reaches 95% of the efficiency plateau. The solid vertical lines correspond to the trigger thresholds provided by the Phase-2 L1 trigger system (at 200 pileup) matching the thresholds currently deployed by the L1 menu for Run-2. The dashed vertical lines correspond to the trigger thresholds required to achieve the same rate using trigger algorithms that do not make use of L1 tracks.

1.4.1.3 Final states requiring double- τ lepton trigger algorithms

- **Higgs boson pair production ($\text{HH} \rightarrow \tau\tau b\bar{b}$)**. As mentioned, one of the most sensitive channels for the search of double Higgs boson production at HL-LHC is $\tau\tau b\bar{b}$. Most of the analysis sensitivity is provided by the final states where both τ leptons decay into hadrons and a neutrino. The HH signal acceptance in this final state is limited by the L1 trigger thresholds. During the Phase-1 running, this channel relied on a double- τ trigger with low enough thresholds to provide an event selection as inclusive as possible. While more features of this specific final state could be exploited at Level-1, using in particular the additional jets from the $H \rightarrow b\bar{b}$ decay, a similar strategy is used here for the sake of evaluating the impact of Phase-2 τ reconstruction on the signal acceptance.

Figure 1.7 (left) shows the inclusive p_T distribution of the sub-leading τ in events where two Higgs bosons are produced, one decaying into a pair of τ leptons and the other into a pair of b -quarks ($\text{HH} \rightarrow \tau\tau b\bar{b}$). The threshold on each leg of the double- τ trigger, allowed by the Phase-2 L1 trigger, is compared to the expected threshold without any tracking or particle-flow information used by the reconstruction algorithm. These thresholds are defined as the p_T of the offline reconstructed tau at which 50% of the single object trigger efficiency plateau is reached. The τ trigger algorithm deployed by the Phase-2 L1 trigger is based on a particle-flow