HEP 2022 – 39th Conference on Recent Developments in High Energy Physics and Cosmology

The CMS Trigger System

Milos Dordevic

Vinca Institute of Nuclear Sciences, University of Belgrade

on behalf of CMS collaboration

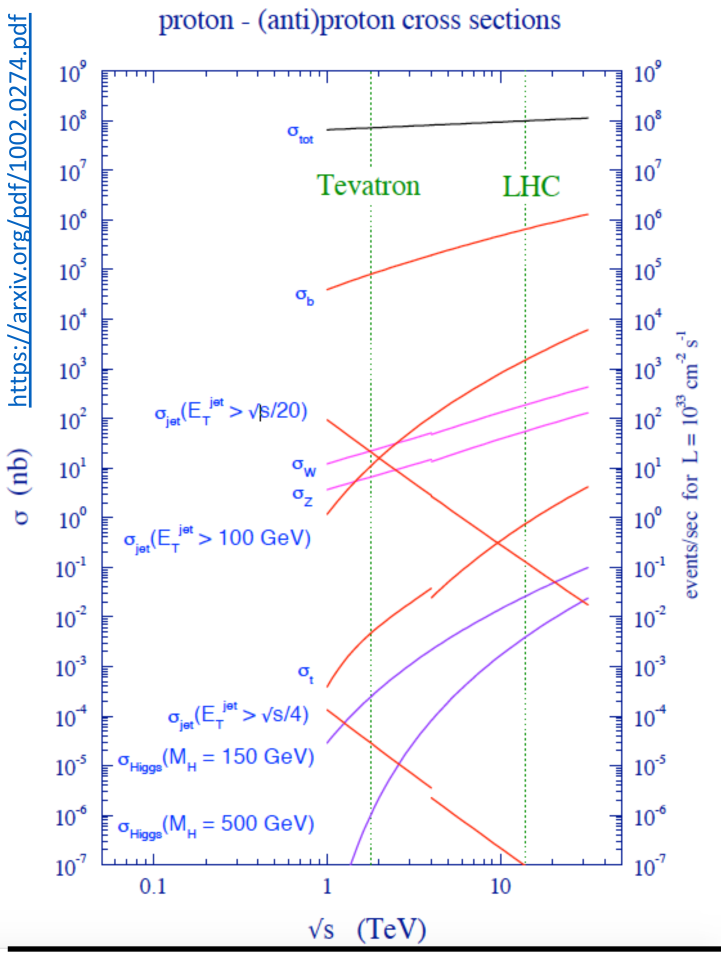
15-19 June 2022, Thessaloniki, Greece

Outline

- Brief overview of the trigger system of CMS experiment
- The CMS Trigger design, architecture & implementation
- The performance of the Level-1 Trigger system at CMS
- The performance of the CMS High Level Trigger system
- The trigger menus at CMS to select the interesting data
- HLT processing time and the GPU - based acceleration
- The another approach: data scouting and data parking
- Summary and Outlook in the view of the Run-3 starting

The CMS Detector and Rates of Physics Processes

- CMS is general purpose detector at the CERN LHC
- Sub - detectors to identify particles & Particle Flow
- Real time decision to store interesting events (Trigger)



Lumi: 2 x 10^{34} cm^2 s^{-1} in the Run2

2556 bunches, 2.5×10^{11} p/bunch

Total collision rate around 2 GHz

b-quark production rate 10 MHz

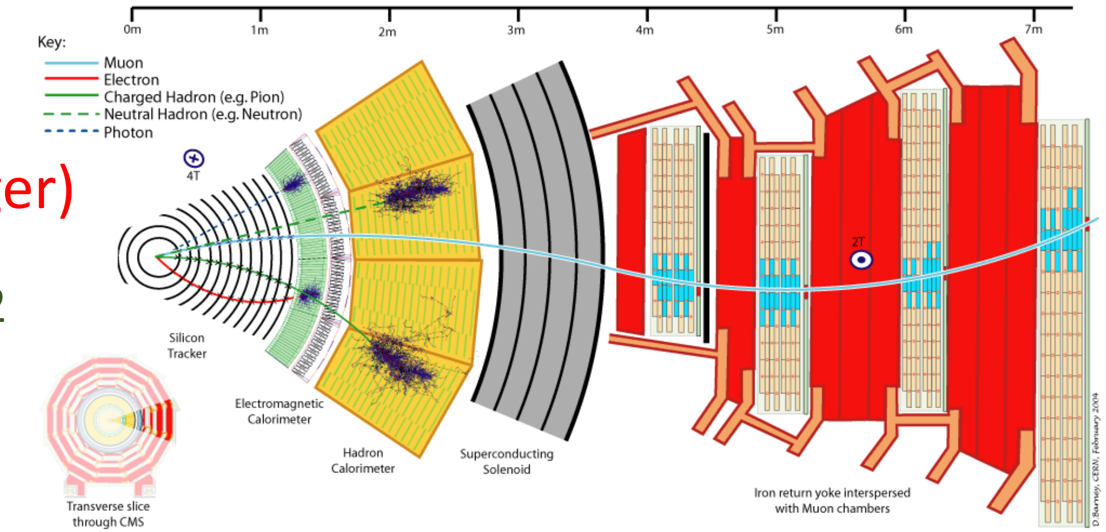
W boson production rate 4 kHz

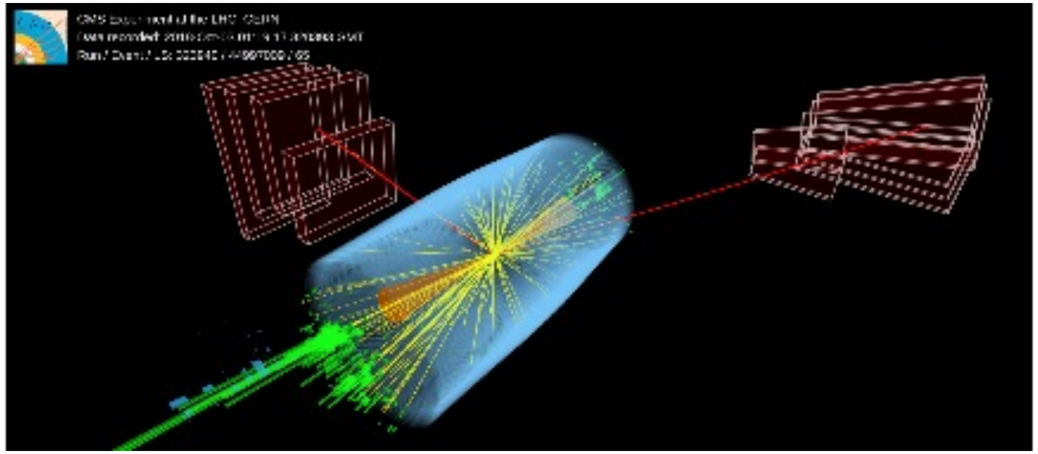
Top quark production rate 20 Hz

Higgs boson prod. rate only 1 Hz

SUSY rate(m@TeV) below 0.1 Hz

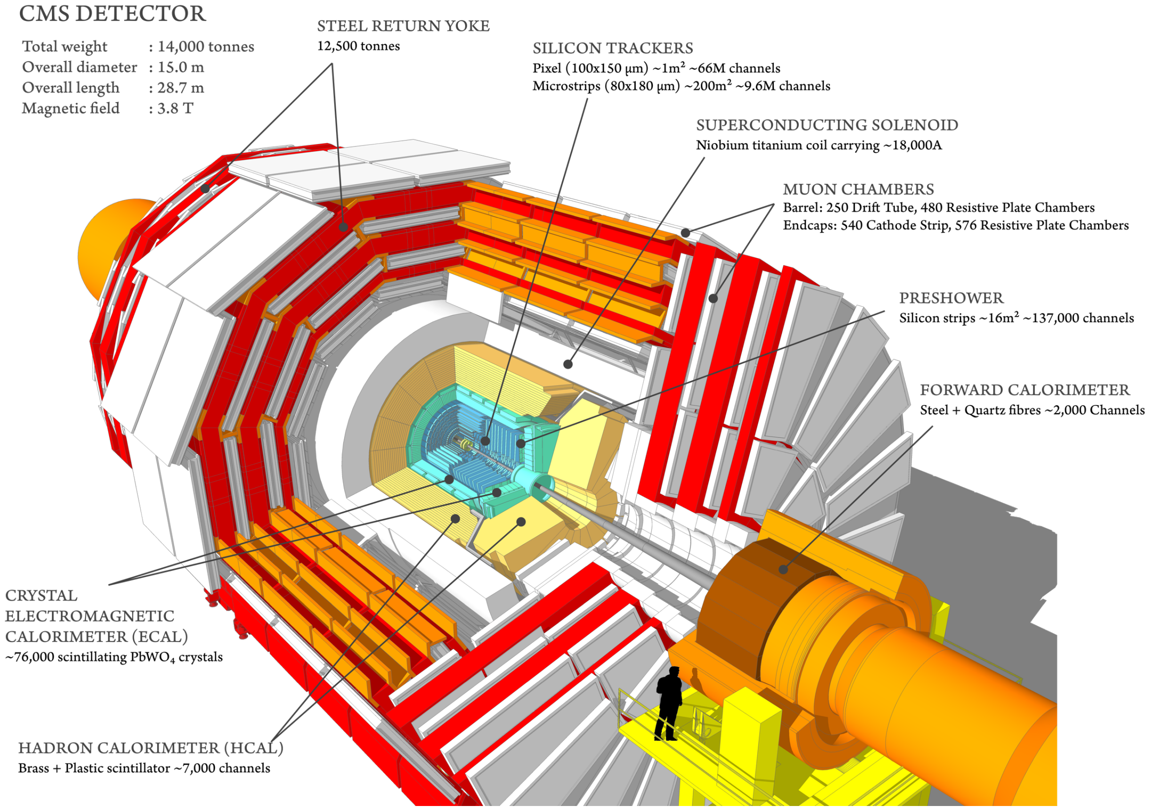
Interesting events at low rates!





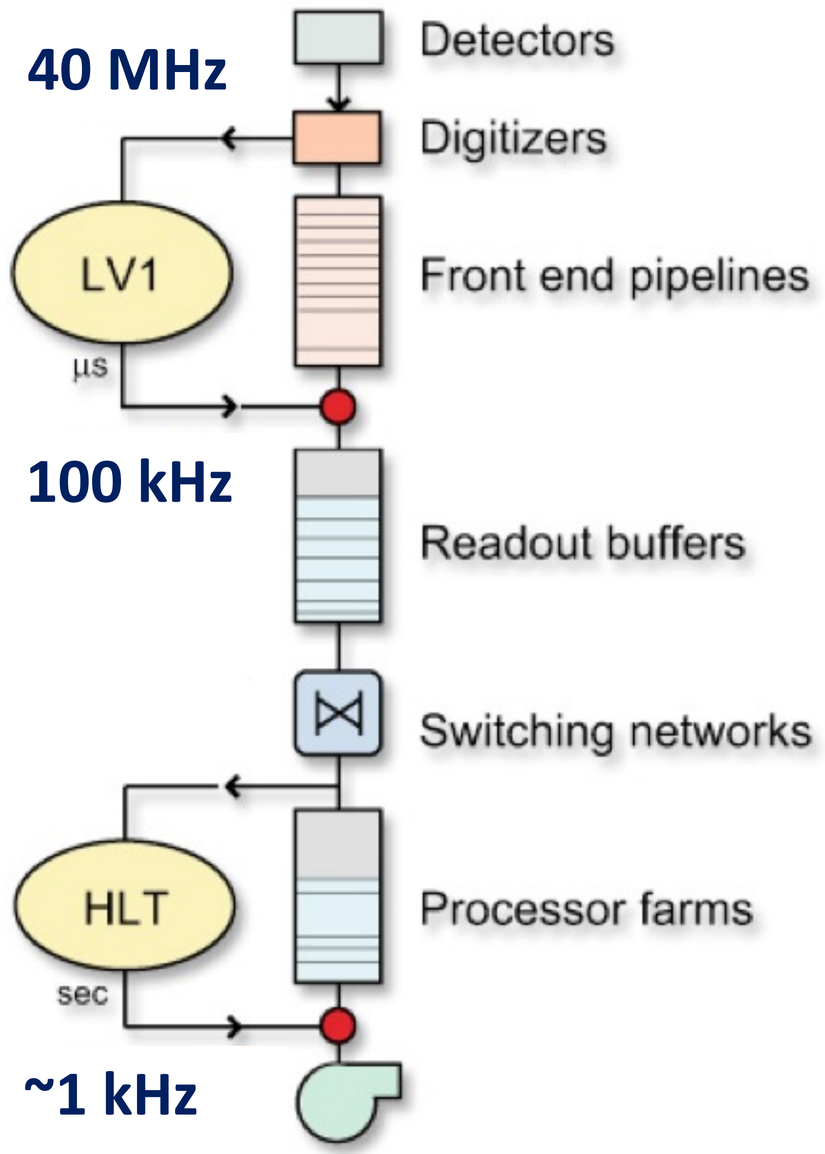
The CMS Trigger System: Overview

- Selective read out of data by experiments in real time is performed by trigger system
- Cannot take all data (storage + processing)
  - bunches collide at 40 MHz rate at LHC
  - may generate 50 terabytes per second



- Each physics analysis starts at the trigger level
- Once event rejected by trigger it is lost forever
  - more than 99.998% LHC data thrown away
  - around 1.5 KHz kept from the 40 MHz rate w/ full event content & prompt reconstruction
- Efficient and clean decision; Trigger universality
- Rate & Time constrains: DAQ bandwidth, buffer

The CMS Trigger System: Design

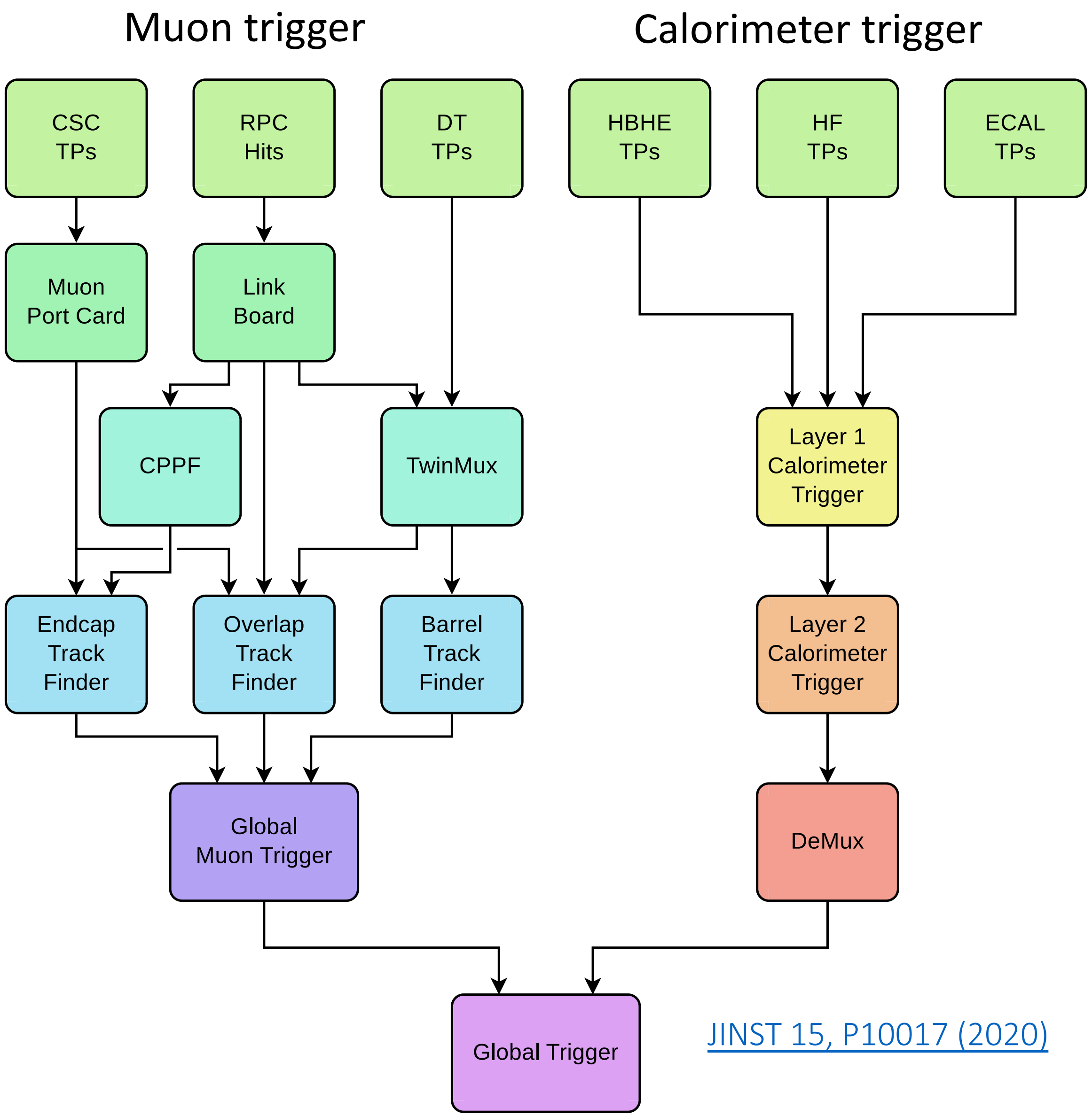


- The CMS Trigger System is organized in two tiers/levels:
  - **Level-1 Trigger** based on custom-made electronics to reduce the data/event rate from the crossing rate of 40 MHz to no more than 100 kHz, with 4µs latency

    ~1 kHz
  - **High Level Trigger (HLT)** filtering events with software running on computing farm based on commercial CPU and now also GPUs, to further reduce the event rate for storage to 1 kHz (in the Run2), now around 1.5 kHz

The Level-1 Trigger: Architecture & Implementation

- Each event analyzed by Muon and Calo trigger
- **Muon trigger** consists of three muon detection systems used early in the processing chain of the trigger, in order to improve the efficiency and resolution, but also to reduce trigger rate
- **Calorimeter trigger** for reconstructing electrons, photons, tau candidates, jets and energy sums
- No tracking readout used (planned for Phase 2)
- **Global trigger** that combines the various objects that are formed by the μGMT and caloL2 triggers
- Set of requirements on trigger objects: L1 menu
  - **around 400 requirements in a logical OR**

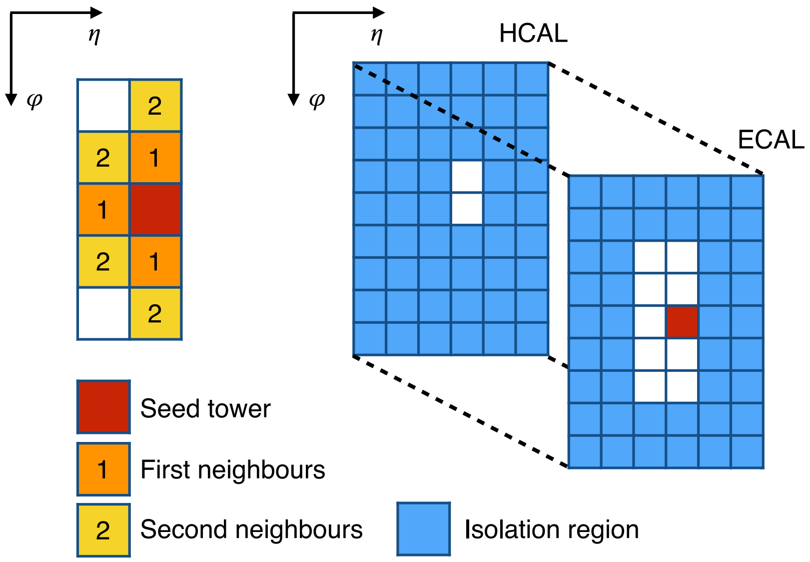


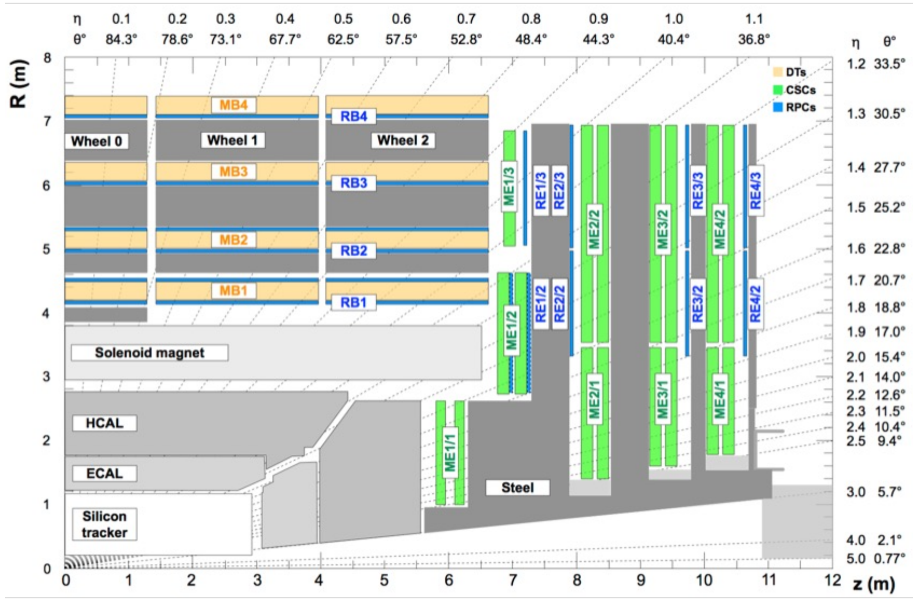
JINST 15, P10017 (2020)

The Level-1 Trigger: Algorithms

JINST 15, P10017 (2020)

- **Electrons and photons** are reconstructed using cluster shape and electromagnetic(EM) fraction to discriminate against jets
- **Jets** reconstructed using sliding window algorithm that looks for trigger tower seeds with an energy over given threshold; 9x9 trigger towers are summed to match offline jets (R = 0.4) after which the jets are also pileup subtracted and calibrated



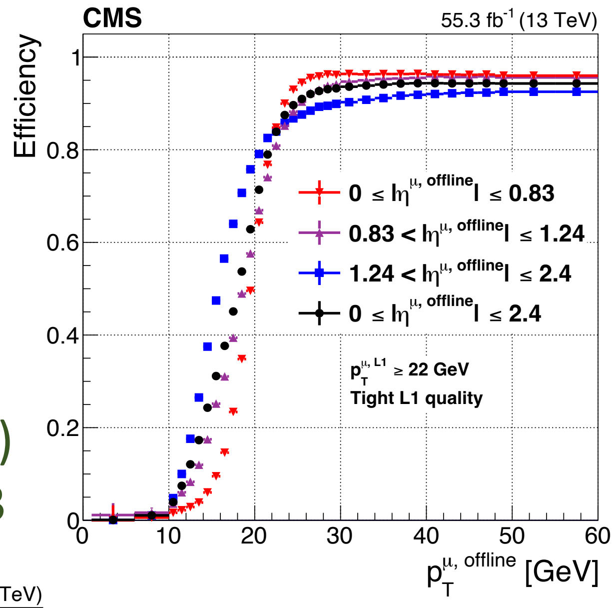


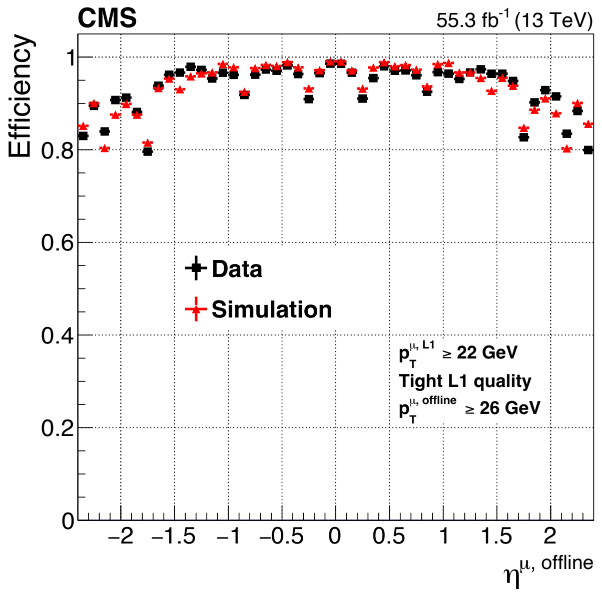
- **Energy sums** are calculated by summing the jet energies with restrictions to jet energy and to pseudorapidity (for the H_T); for MET: all TTs over $E_T(\eta, PU)$ summed (in full η)
- **Muons** reconstruction using an extrapolation based track finding in barrel, pattern based in overlap/endcap region
  - muon p_T assignment based on Δϕ in barrel, patterns in overlap region and BDT regression used in endcap

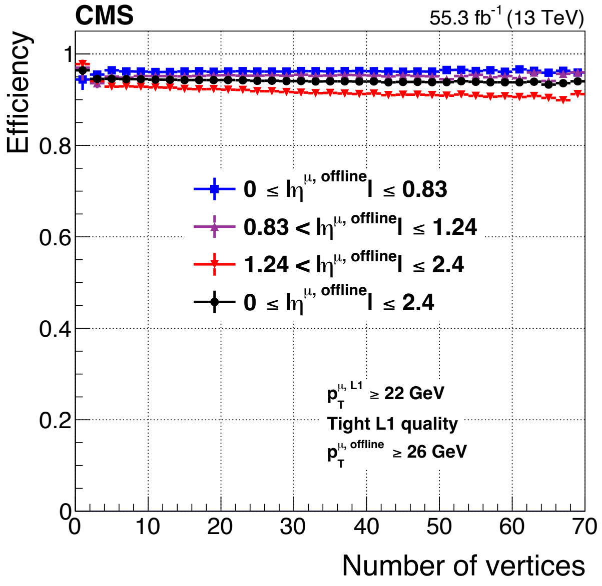
The Level-1 Trigger Performance: muons

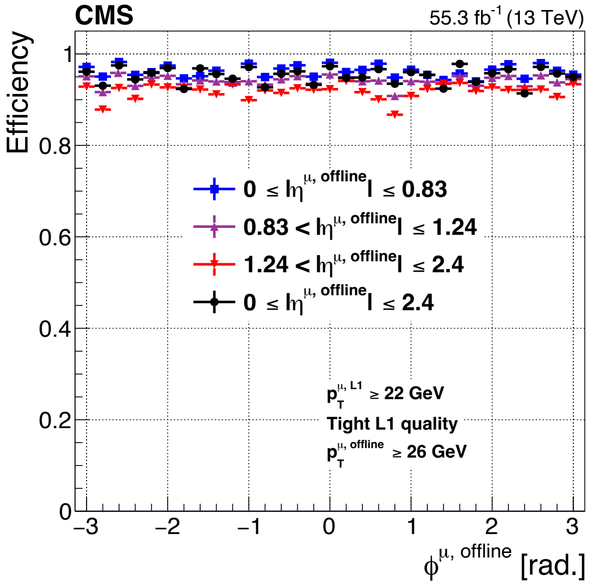
JINST 15, P10017 (2020)

- L1 muon efficiency as function of $p_T^{\mu,\ offline}$
  - sharper in barrel due to better resolution
- Efficiency vs $\eta^{\mu,\ offline}$ falls in forward region
  - due to only CSCs used in EMTF (i.e. no detector redundancy) and p_T assignment more difficult due to reduced lever arm (and more showering)
  - improve with GEM detector added for the Run3







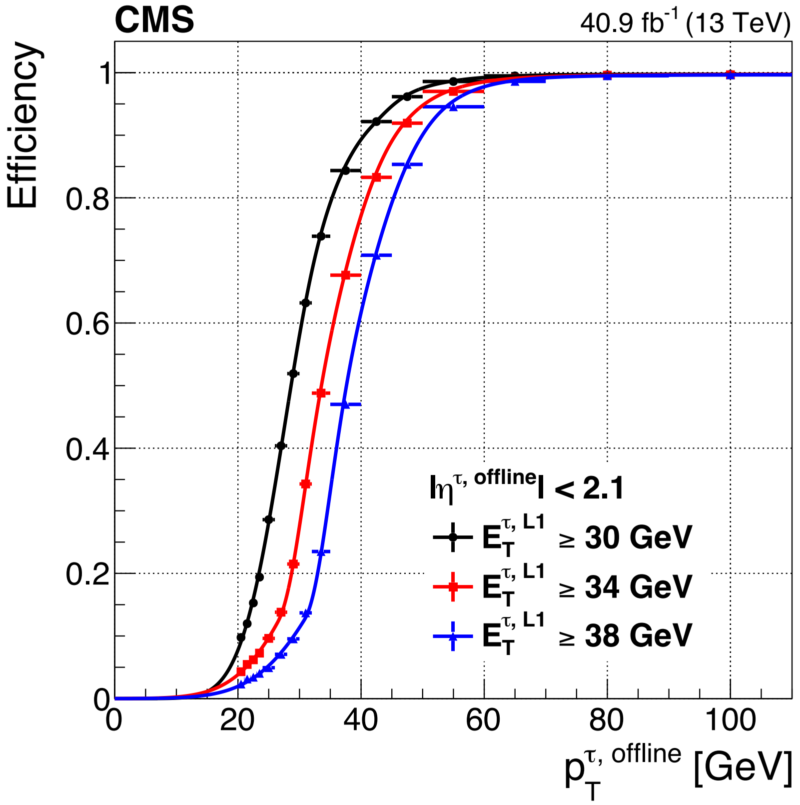


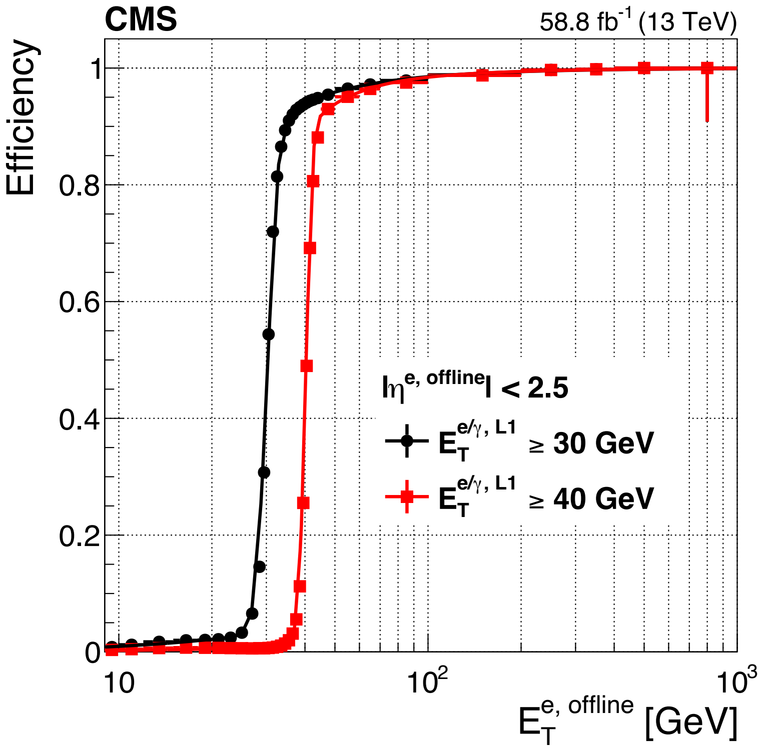
- Level-1 efficiency of the muon track finder flat vs number of offline vertices (left plot) and also flat vs ϕ distribution of the muon
  - higher in detector layers overlap region

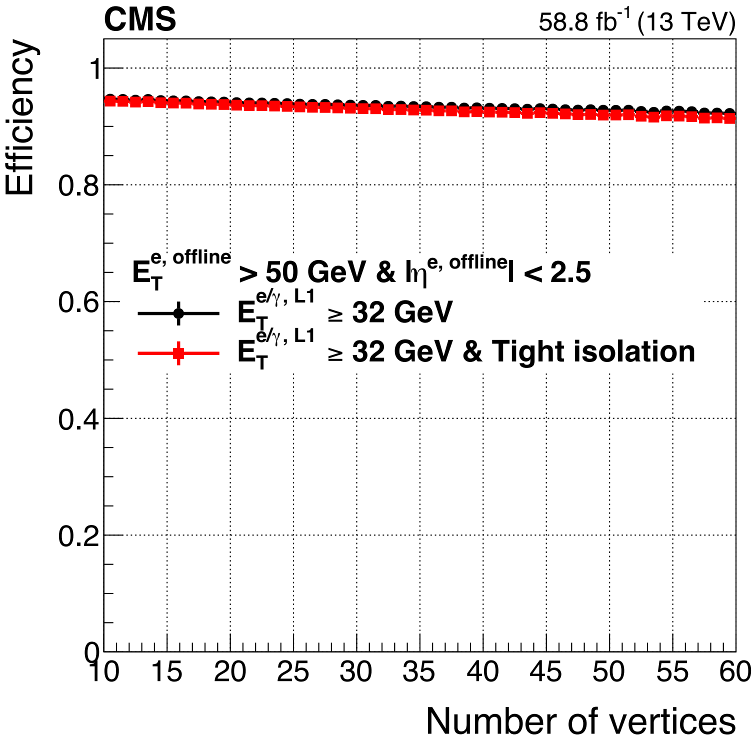
The Level-1 Trigger Performance: ele, pho, tau

JINST 15, P10017 (2020)

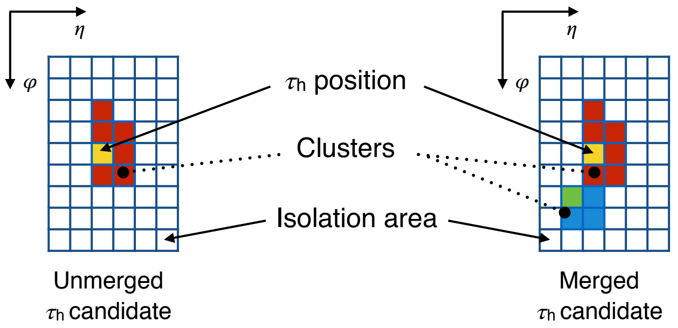
- Level-1 e/γ eff: **sharp turn-on**
- Only small pileup dependence





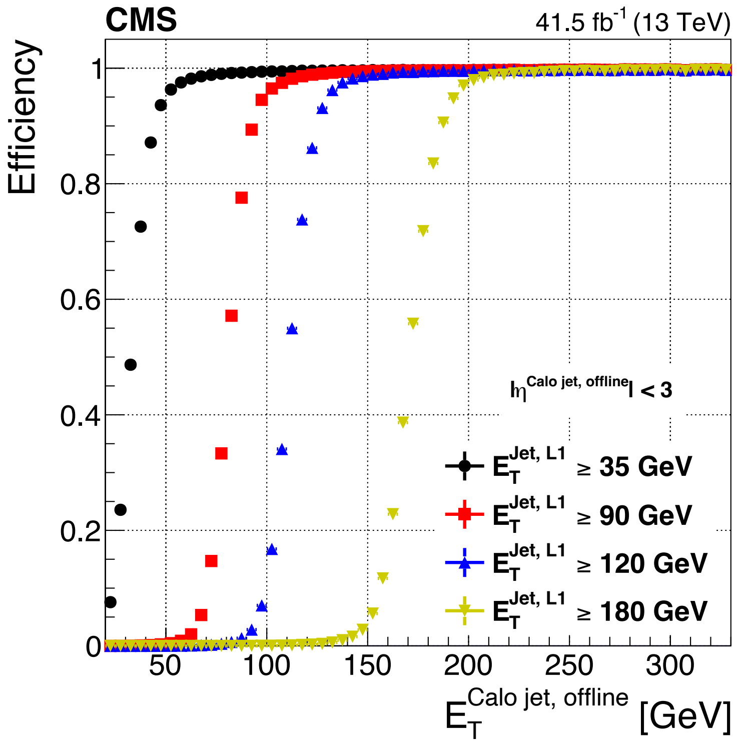


- Level-1 τ clustering & isolation
- Level-1 τ efficiency vs offline p_T



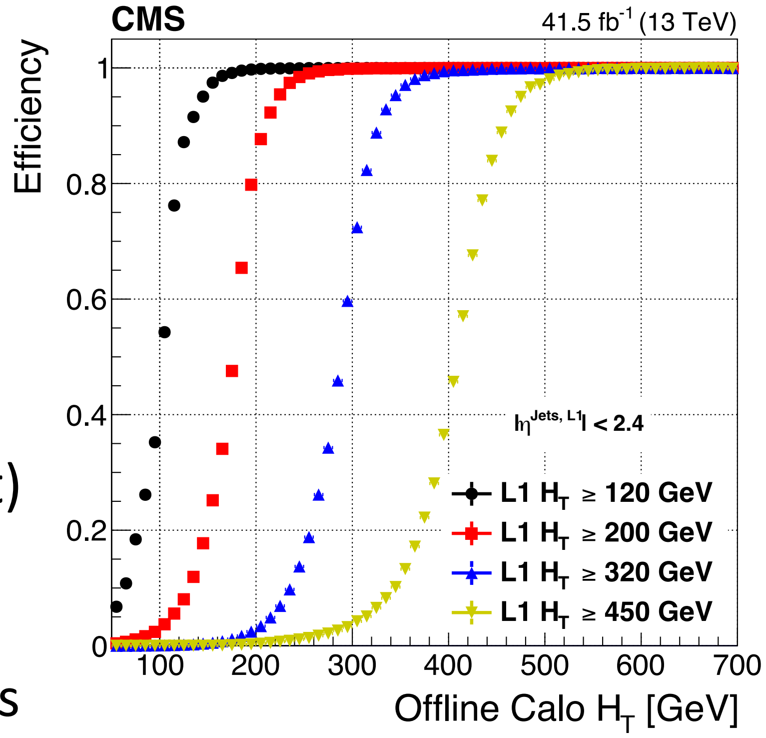
The Level-1 Trigger Performance: jet, HT & MET

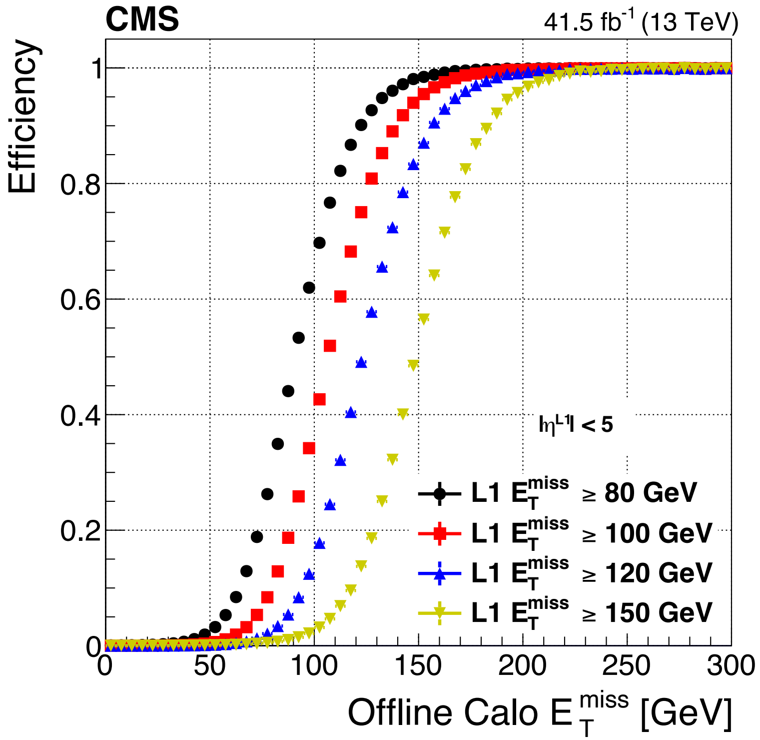
JINST 15, P10017 (2020)



- The efficiency curves for the Level-1 jet triggers for the barrel plus endcap pseudorapidity range
  - for the thresholds of 35, 90, 120 and 180 GeV

- The efficiency for scalar sum of jet energy with $E_T \geq 30$ GeV (left) and missing transverse energy (right) for the various thresholds

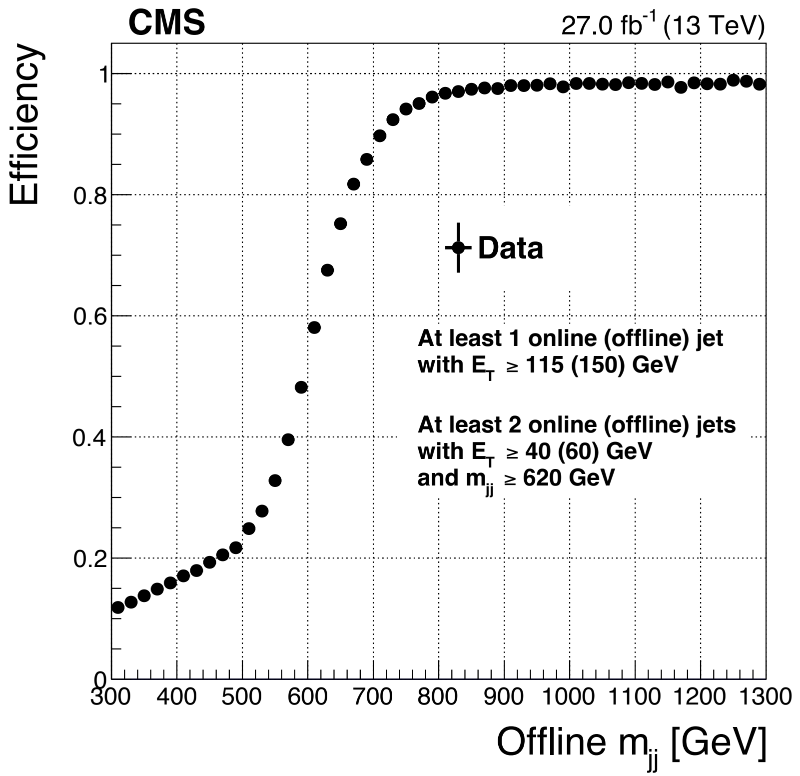




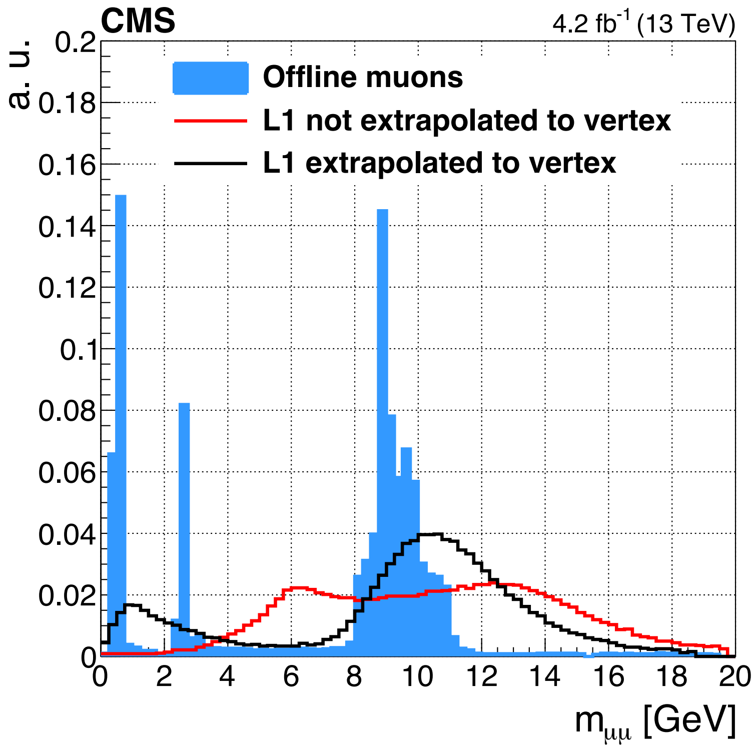
The Level-1 Trigger: Dedicated Analysis Triggers

JINST 15, P10017 (2020)

- Vector Boson Fusion (VBF) dijet trigger with invariant mass cut:
  - required at least two jets with E_T > 115 and 35 GeV and to have at least one pair of jets with each having E_T > 35 GeV and also dijet invariant mass to be greater then 620 GeV



Examples of the possibilities with Level - 1 Global Trigger

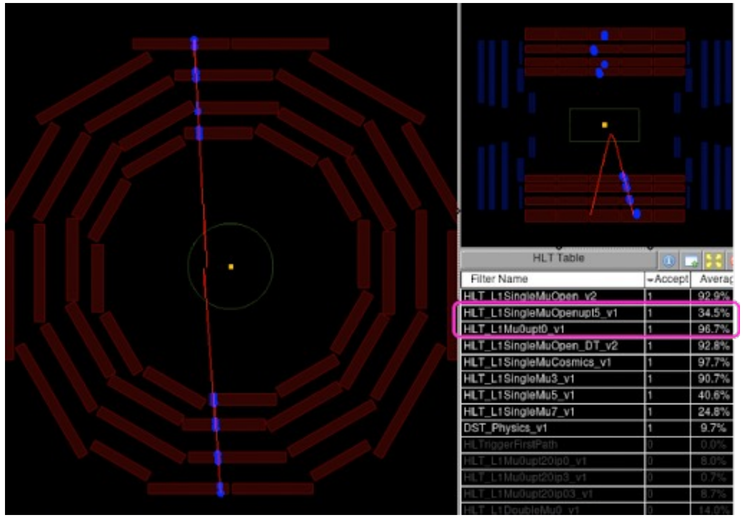


- Low mass di-muon triggers with invariant mass cut (e.g. B trigger):
  - apply lower p_T thresholds with the dimuon invariant mass cut
  - μ p_T extrapolated to vertex for mass calc. from a standalone μ

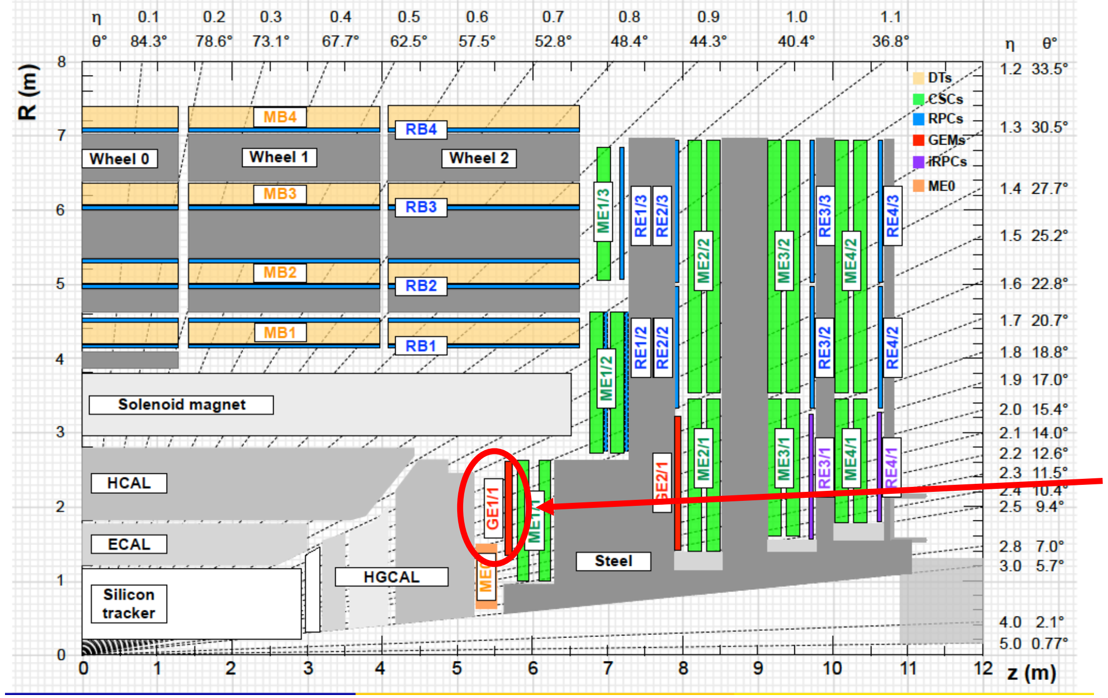
The Level-1 Muon Trigger: Run 3 improvements

PoS Vol. 343 (2019) 139

- Kalman filter for muon tracking incorporated into FPGA logic of **barrel muon trigger**, allowing the possibility to trigger on Long Lived Particles with transverse displacement up to ~1m
  - tested with data obtained from cosmic muon interactions



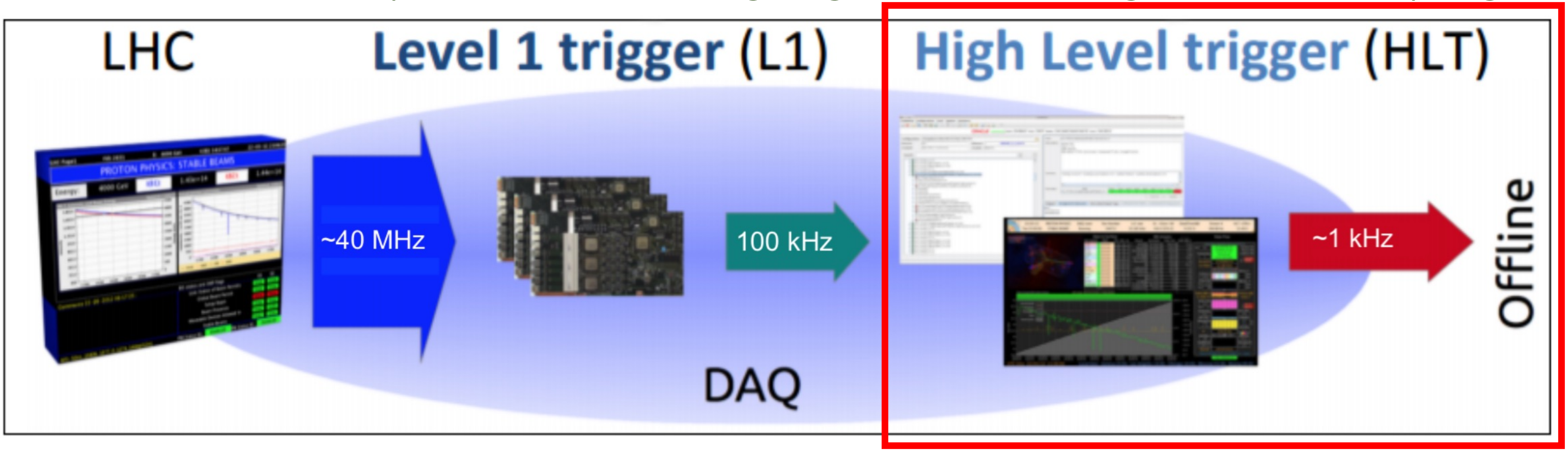
https://indico.cern.ch/event/998052/



- Neural network for measuring displaced muons in endcap ported to the FPGA logic of **endcap trigger**
  - extends possibility to trigger on LLP in endcap
- (Phase-2) GEM detector innermost disk was added
  - to improve efficiency in fwd region 1.6 < η < 2.1

The High Level Trigger: Overview

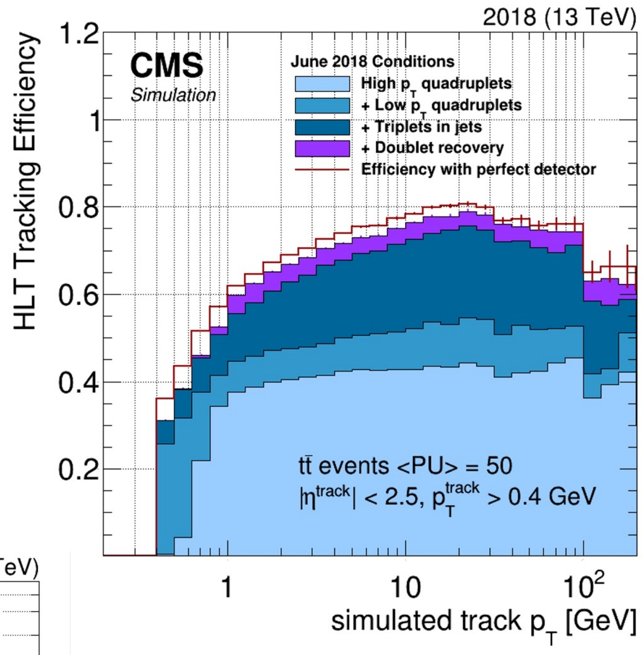
- High Level Trigger is set to reduce the event rate from 100 kHz to ~1(1.5) kHz in LHC Run 2(3)
  - the output rate from offline computing data processing constrains and storage capacities
  - uses offline reconstruction algorithms and code, but optimized so it’s around 100x faster
  - total 30,000 CPU cores used in the High Level Trigger system at the end of the LHC Run 2
  - hundreds of HLT paths constructed, targeting the broadest range of the event topologies

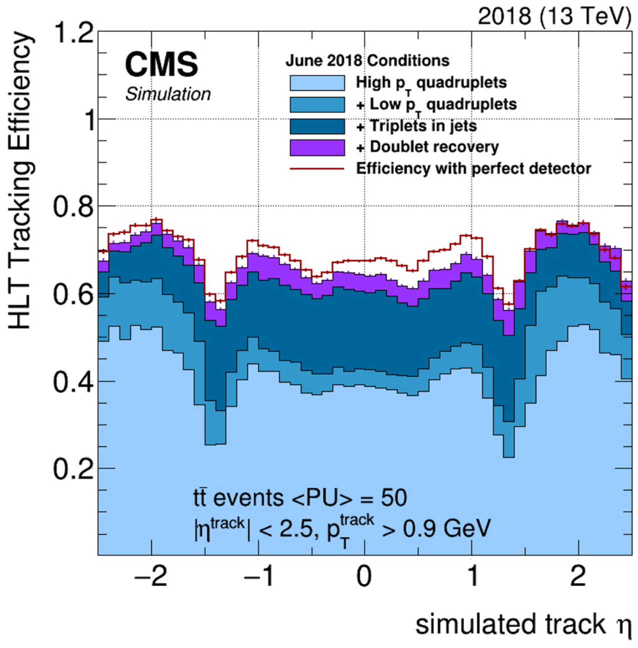


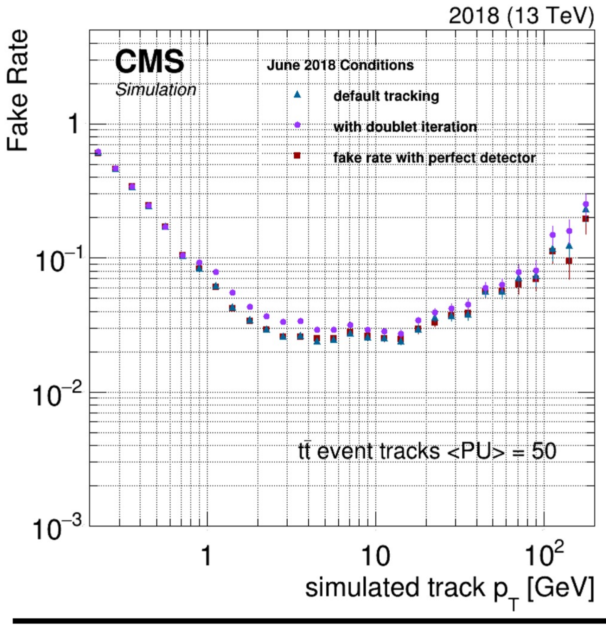
The High Level Trigger: Tracking

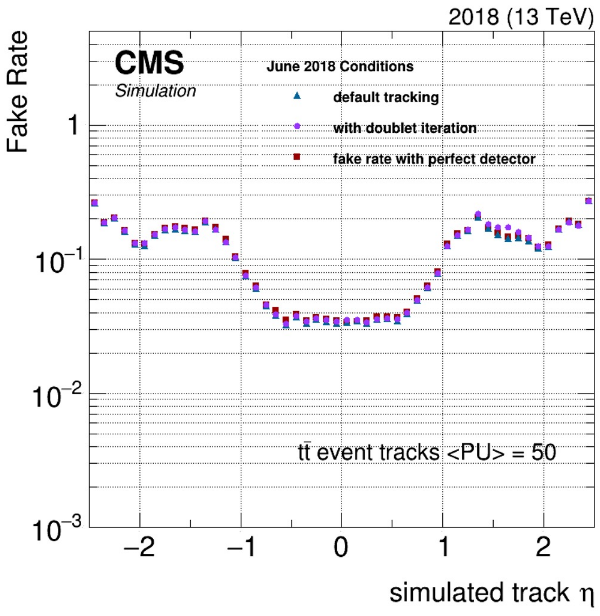
CMS-DP-2018-038

- Tracking at HLT simplified from the offline
  - reduced # iterations (10 in the offline)
  - regional tracking (eg. at high deposits)
  - after all iter. close to perfect detector!
  - efficient for low pT tracks (to 0.4 GeV)







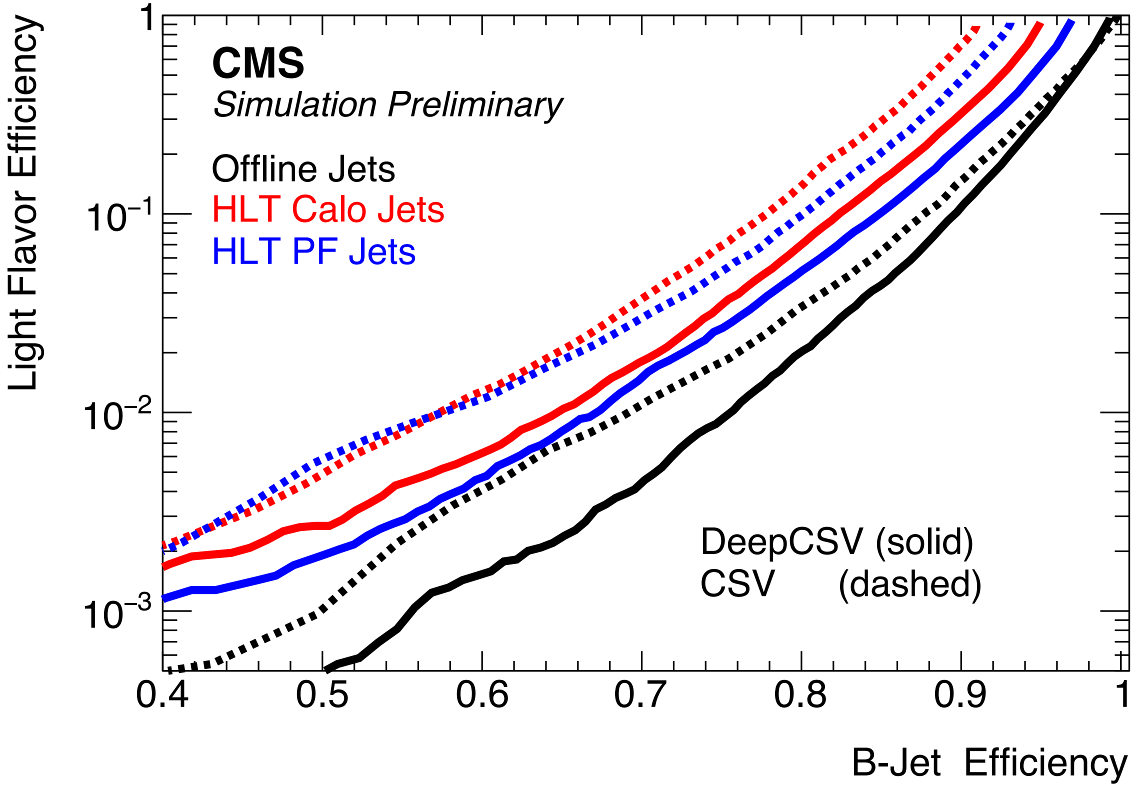


- Tracking fake rate largest in forward region
- Also fake rate grows at the low and high p_T
- Not much increase with doublet recovery iteration

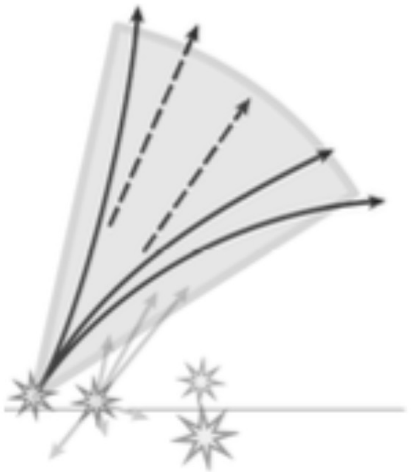
The High Level Trigger: B-Tagging

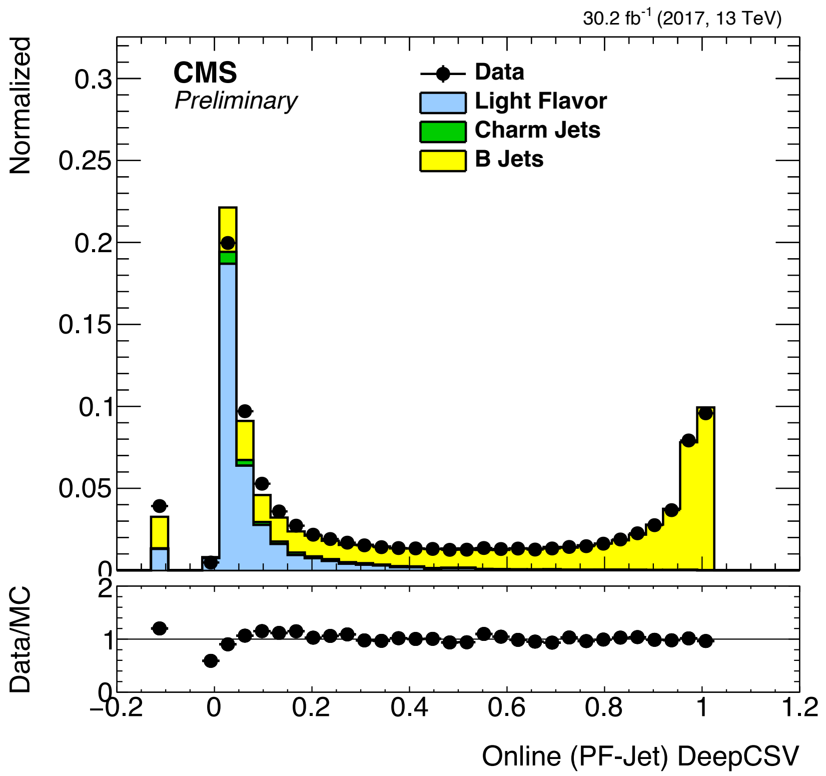
CMS-DP-2019-042

- Neural network based classifier (Deep CSV) used since 2017 to identify the b-tagged AK4 jets



- Online and offline b-jet identification efficiency
- Improved b-tagging efficiency over previous CSV algorithms 5-15% for fixed light flavor efficiency





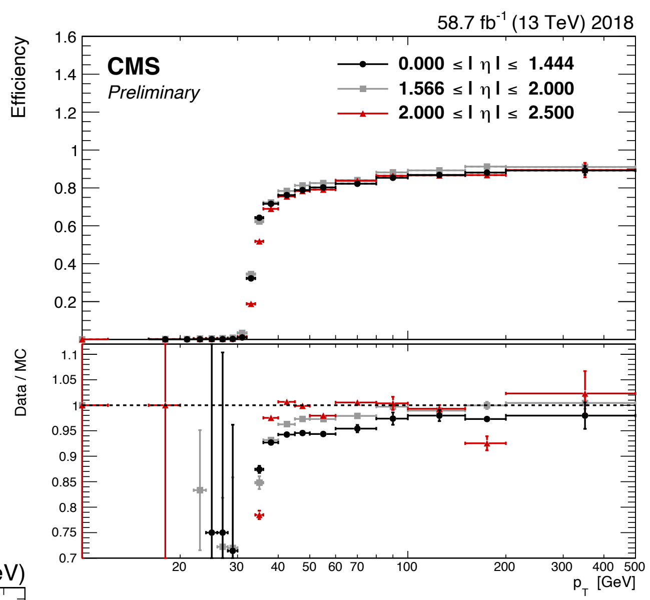
- The DeepCSV discriminator distribution for online (PF-Jets)-> different colors show the contributions in simulations from different jet flavors

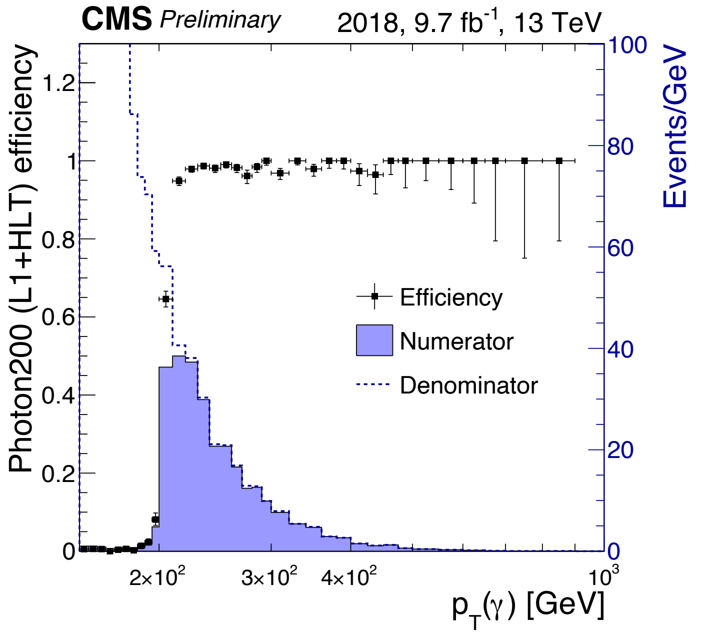
The HLT Performance: Electrons and Photons

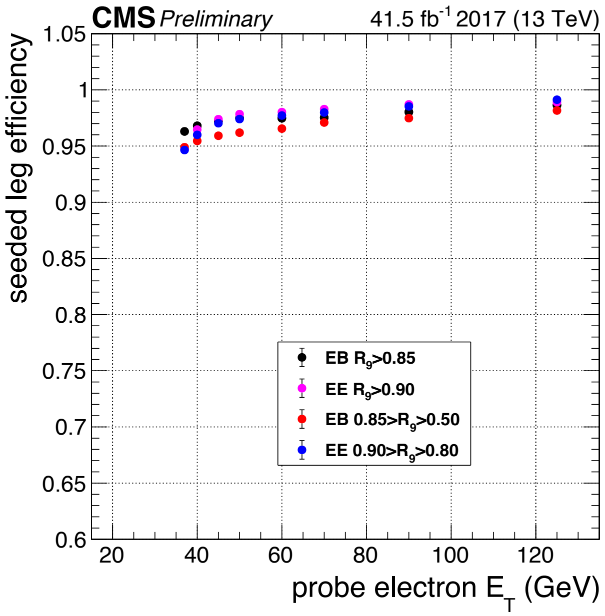
CMS DP-2020/016
CMS DP-2018/049
CMS DP-2018/039

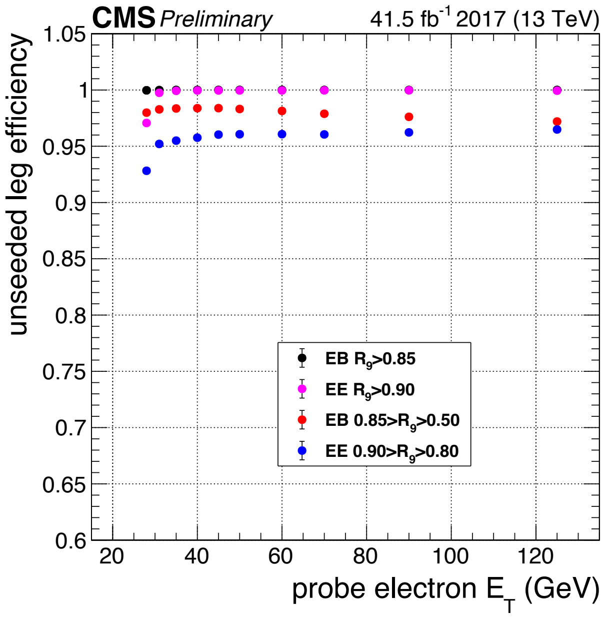
- **Efficiency of HLT_Ele32_WPTight_Gsf** with respect to offline candidates for different η

- **Efficiency of HLT path that requires a photon with p_T > 200 GeV (used in SUSY, also others)**







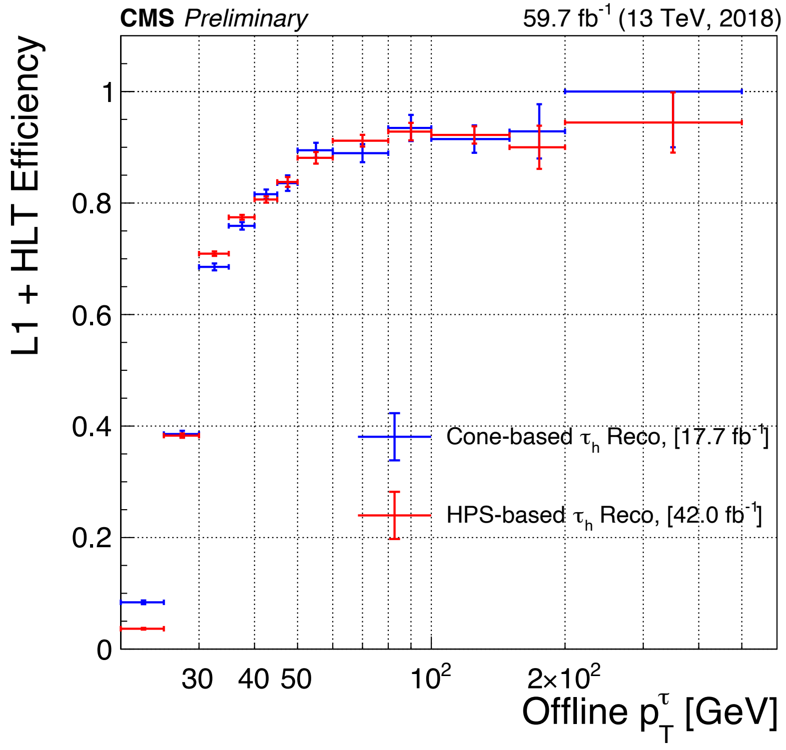


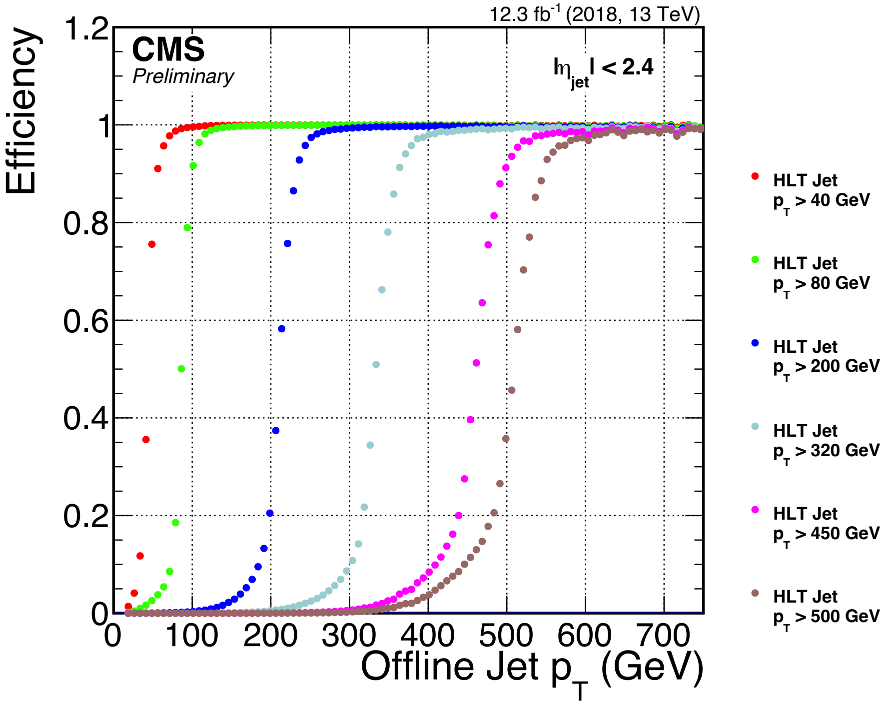
- Efficiency w.r.t probe electron transverse energy of the seeded (left) and unseeded (right) leg of the di-photon trigger for 4 analysis categories, defined w.r.t probe R9 (E_{5x5}/E_{SC}) and η, measured on data for Z->ee events using the tag-and-probe method

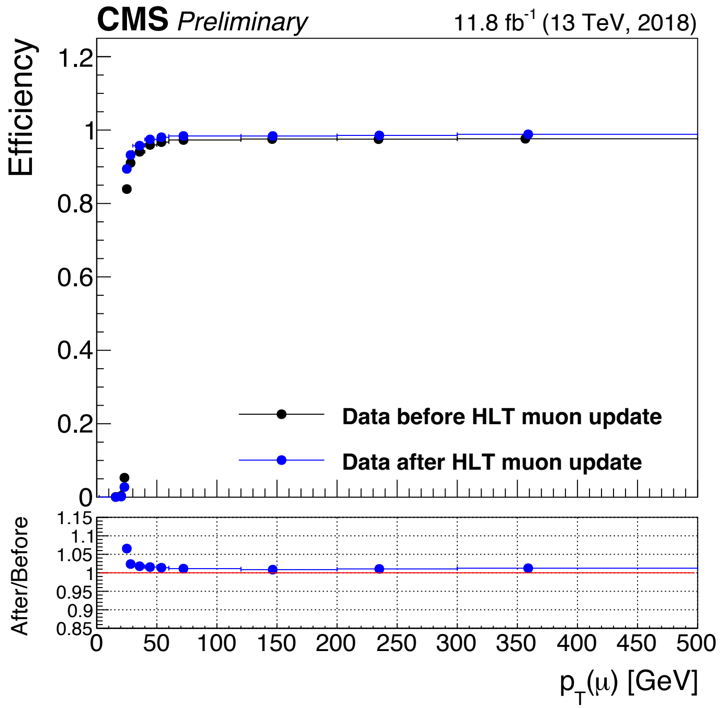
The HLT Performance: Jets, Muons and Taus

CMS DP-2018/034
CMS DP-2019/012
CMS DP-2018/037

- Efficiency of jet triggers w.r.t offline candidates in |η|< 2.4
- Efficiency of trigger requiring iso single μ with p_T > 24 GeV





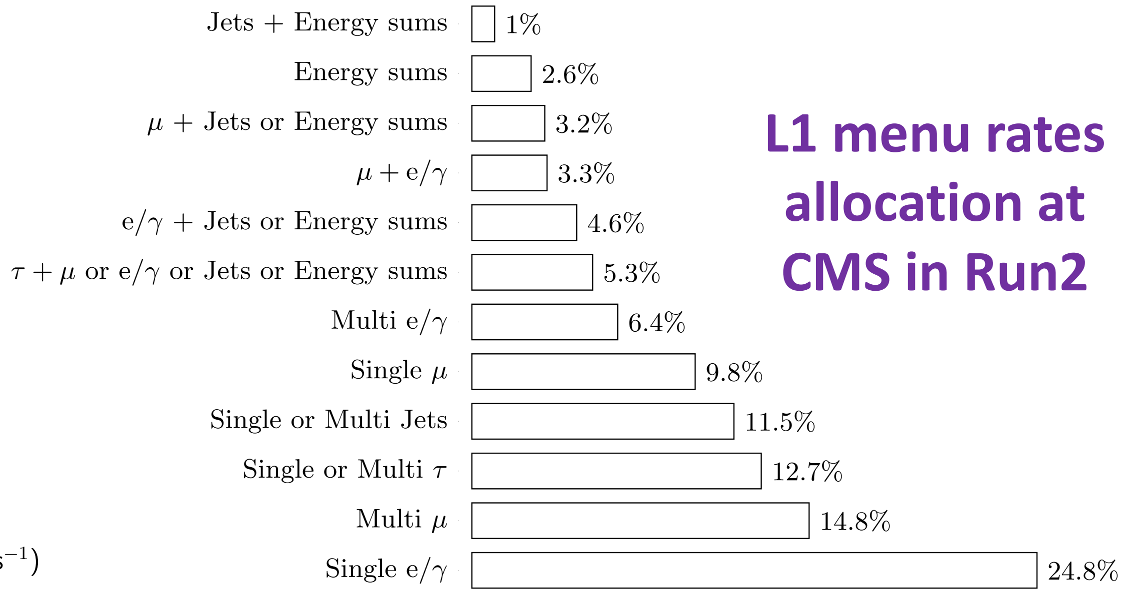


- The efficiency of a hadronic tau leg of the mu-tau trigger shown as a function of offline tau pT for the 17.7 fb^{-1} data taken with the cone-based tau reconstruction and for 42.0 fb^{-1} data collected with the HPS-based algorithm in 2018
- The combined L1 and HLT efficiency of the τ_{had}-leg shown

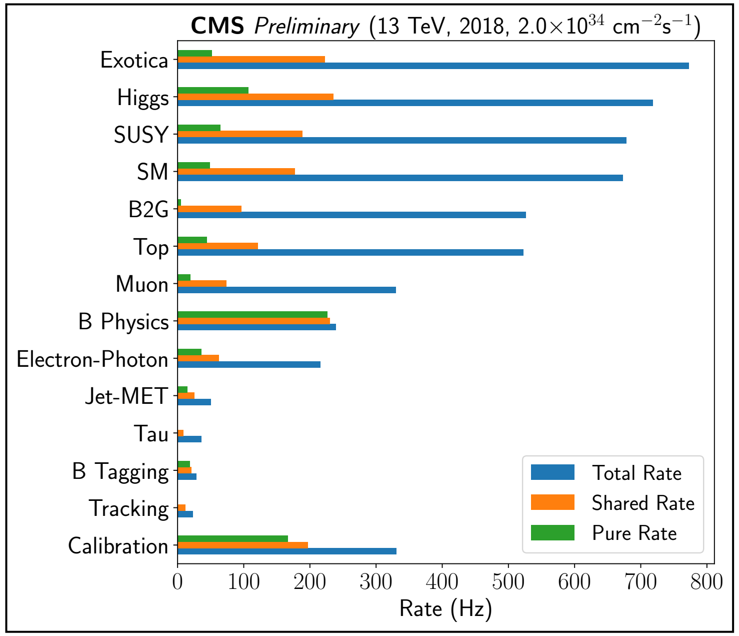
The Trigger Menus: How to Select Your Data

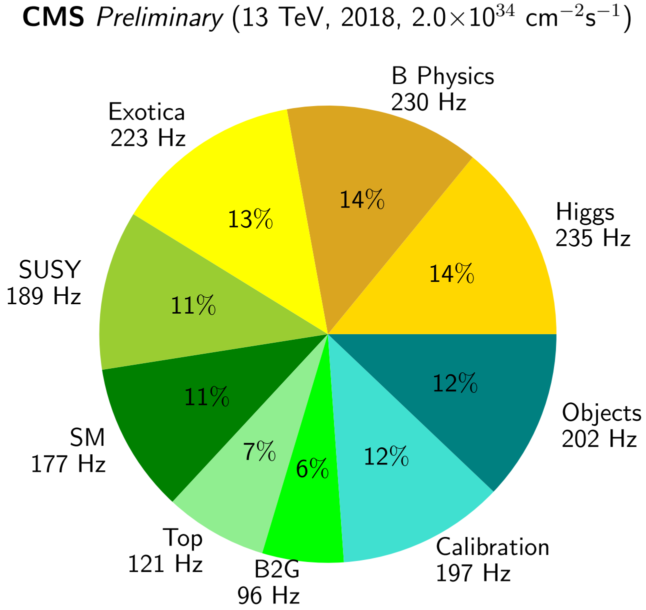
JINST 15, P10017 (2020)
CMS DP-2018/057

- **Trigger menu** represents a large set of selection criteria enabling to fulfill broad physics program
- General triggers, some very specific and backup
- Separate L1 / HLT menus with ~ 300 / 600 items
- Level-1 menu: single/multi/cross path fractions



L1 menu rates allocation at CMS in Run2

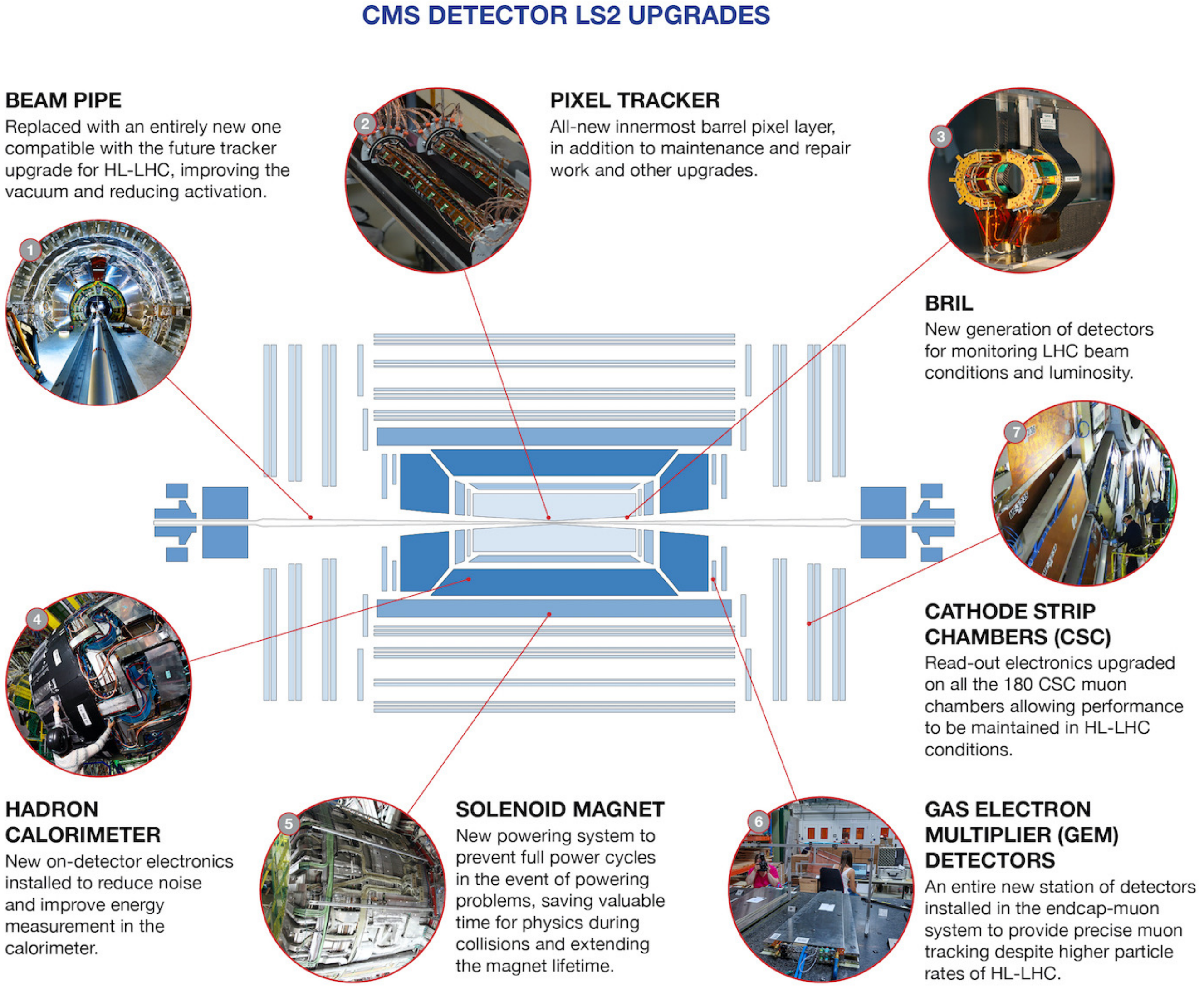




- HLT rates distribution by physics groups
  - measured at 1.2e34, scaled to 2e34
  - total, shared and pure rates shown
- The pie chart also displays rate allocated to each physics group for a lumi of 2e34

The CMS Detector Phase-I Upgrade

- Significant upgrades of the CMS detector during the course of Run 2 and in the LS2
- **Level-1 Trigger upgrades:**
  - Phase-I L1 trigger upgrade CMS-TDR-12 in 2016: finer calorimeter granularity --> improved energy and position resolution while remaining within rate constraints
  - Pileup subtraction in Run-2 at benefit of L1 performance JINST 15 (2020) P10017
- **High Level Trigger upgrades**:
  - Phase-I Pixel Upgrade CMS-TDR-11 during the LHC EYETS in 2016-2017
  - Phase-I HCAL Upgrade CMS-TDR-10
    - ❖ Endcap included in 2018 (not used in trigger); Barrel completed in 2019

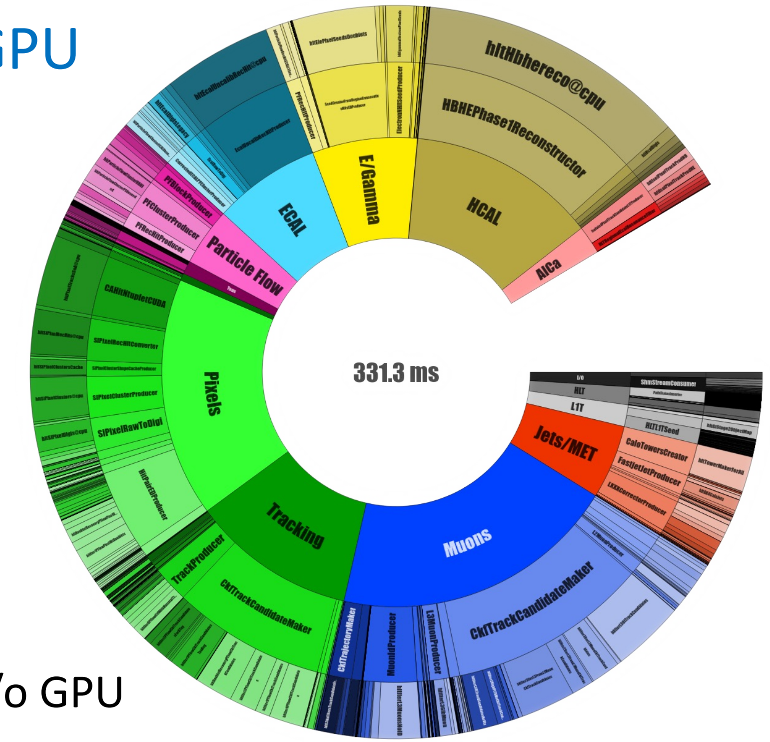


https://home.cern/press/2022/CMS-upgrades-LS2

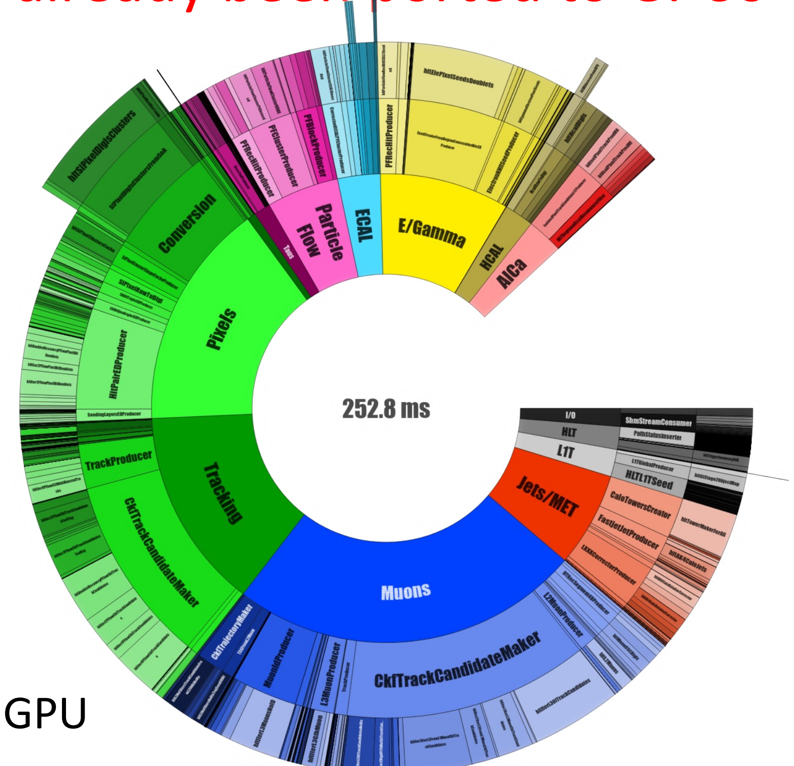
The CMS Trigger System: GPU Acceleration

[CMS DP-2021/013]

- Starting to use a heterogeneous architecture in online reconstruction, based on CPU and GPU
  - pixel and pixel - based tracking, ECAL & HCAL local reco have already been ported to GPUs
  - 25% CPU time offload to GPU
- Planned to **port more reco code (like Particle Flow)** to the GPU in the near future
- **GPU reco much faster** and will allow tracking on more events, improving scouting at Level - 1 for low masses



w/o GPU



w/ GPU

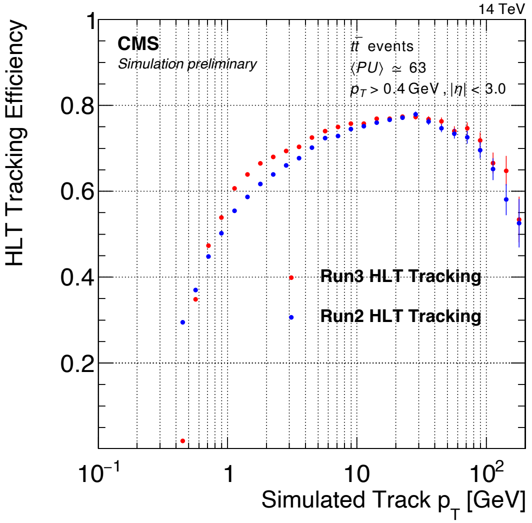
NVIDIA T4

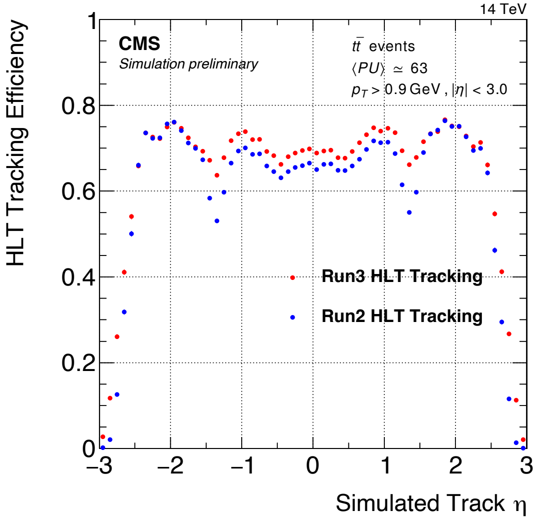
The timing is measured on pileup 50 events from [Run2018D] running 4 jobs in parallel, with 32 threads each, on a full node (2× AMD "Rome" 7502) with SMT enabled.

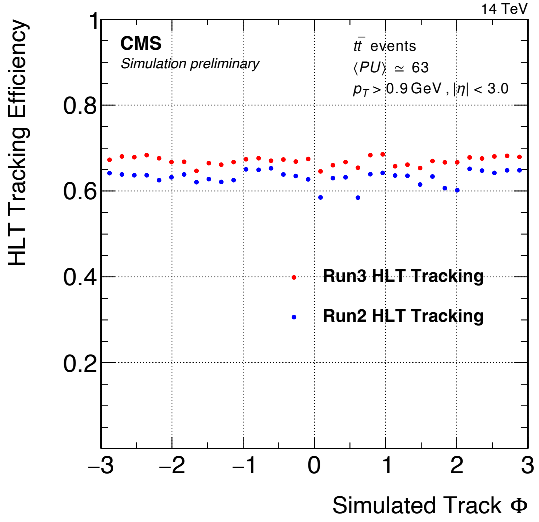
[PhaseIIHLTRecoAndGPUPerformance]

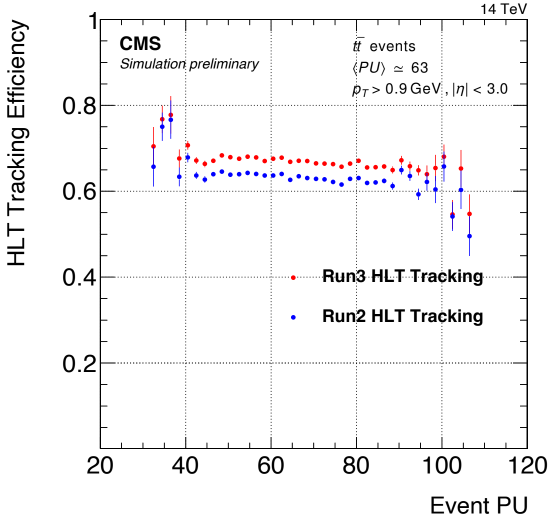
The CMS Trigger System in LHC Run 3: Tracking at HLT

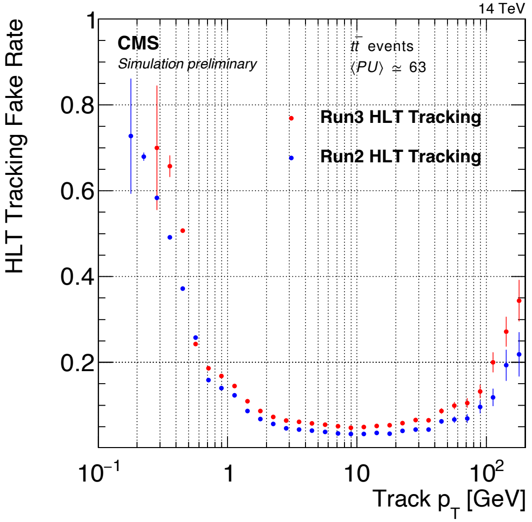
CMS DP-2022/014

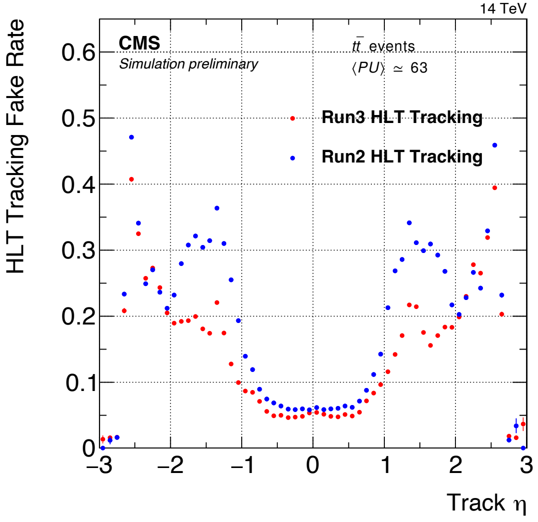


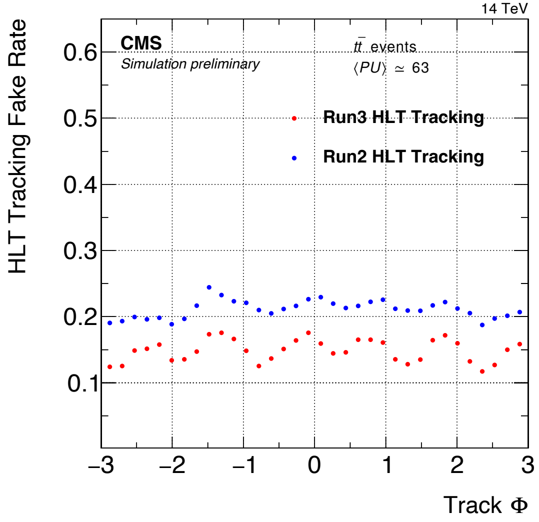


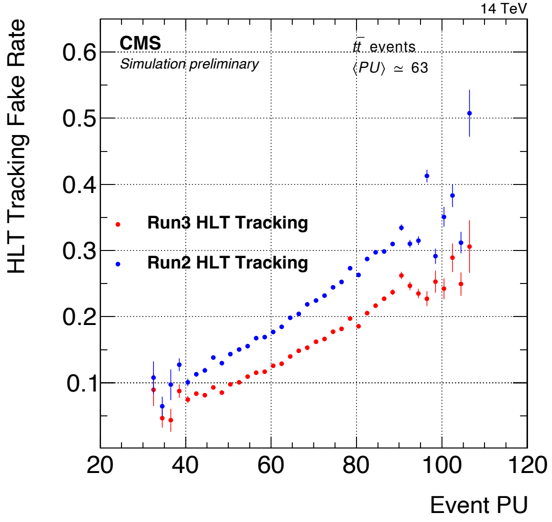








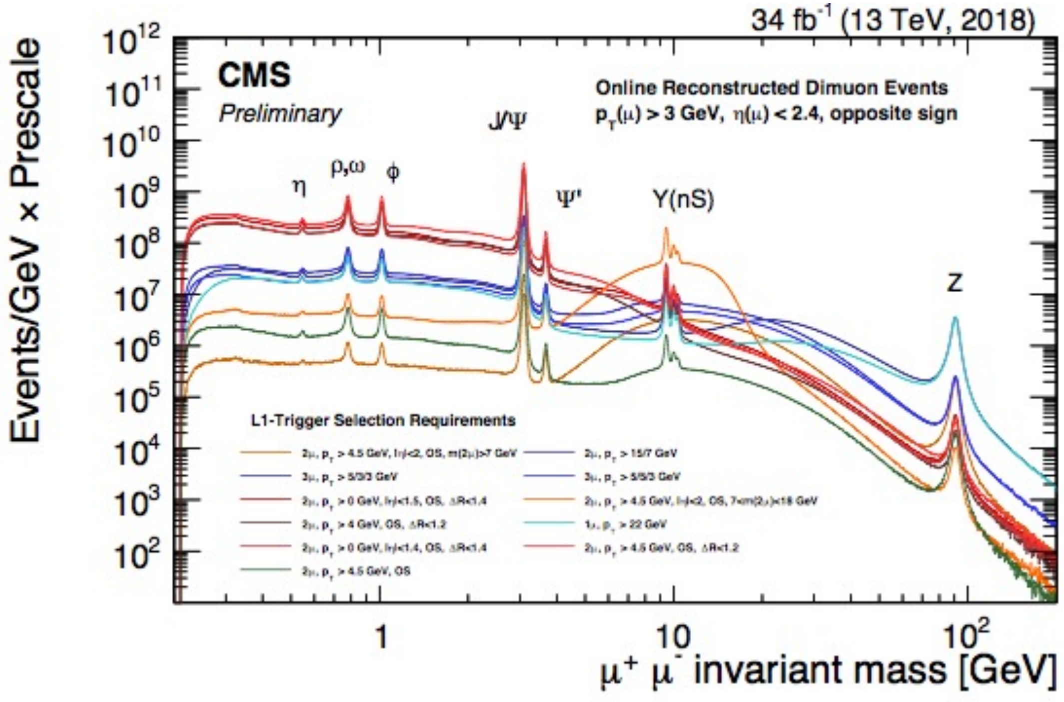


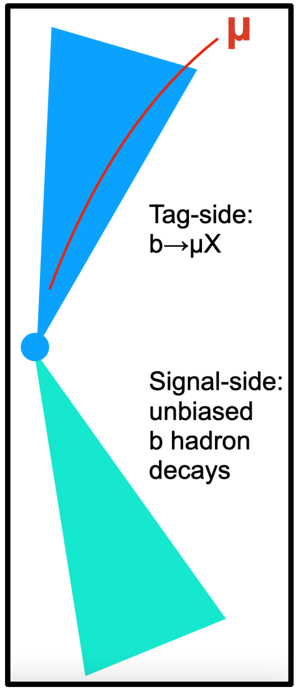


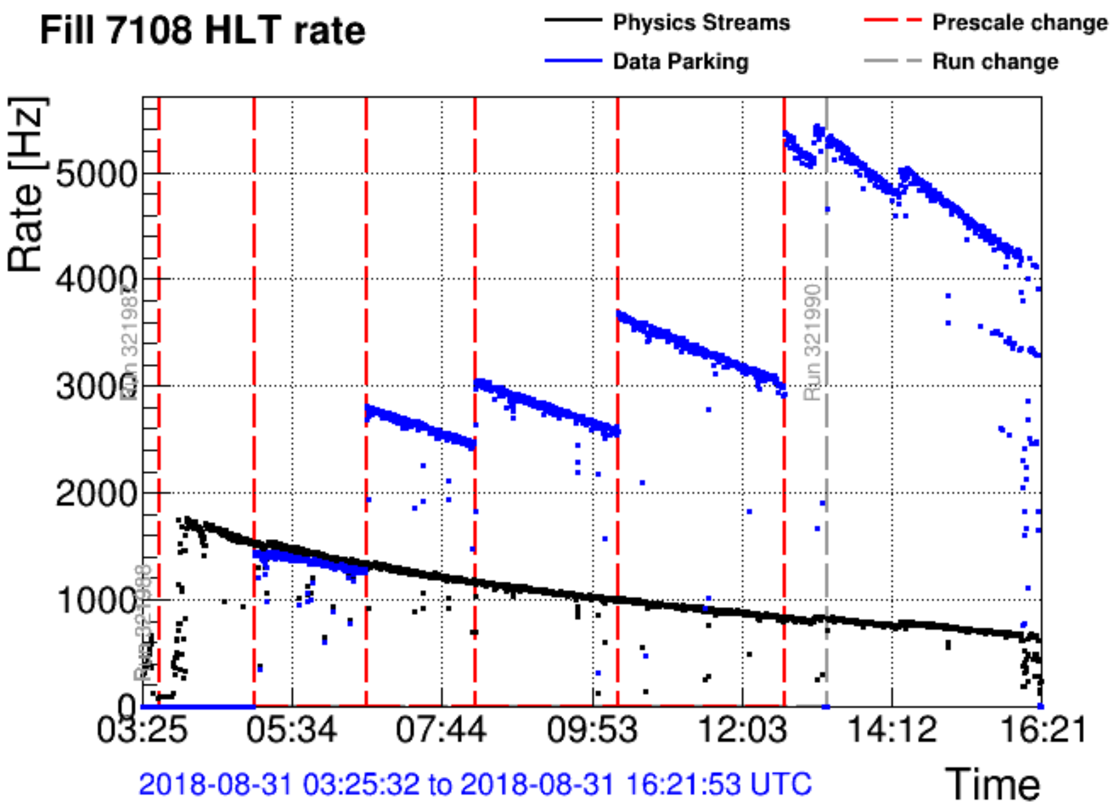
The CMS Trigger System in LHC Run 3: Parking & Scouting

CMS-DP-2019-043
CMS-DP-2018-055

- Computing resources limited at HLT:
  - up to 1 kHz in Run2, 1.5kHz Run3
  - limited bandwidth to ~5Gb/s and also storage (disk & tape) limited
- Reduce event size w/ Scouting: save online obj. used directly in analyses







- Reduce computing resources w/ Parking, i.e. "park" data on tape, skip prompt reco & reconstruct it later
- In Run 3, expanding reach to high rates and more exotic phase spaces; dedicated LLP triggers at L1

Summary and Outlook

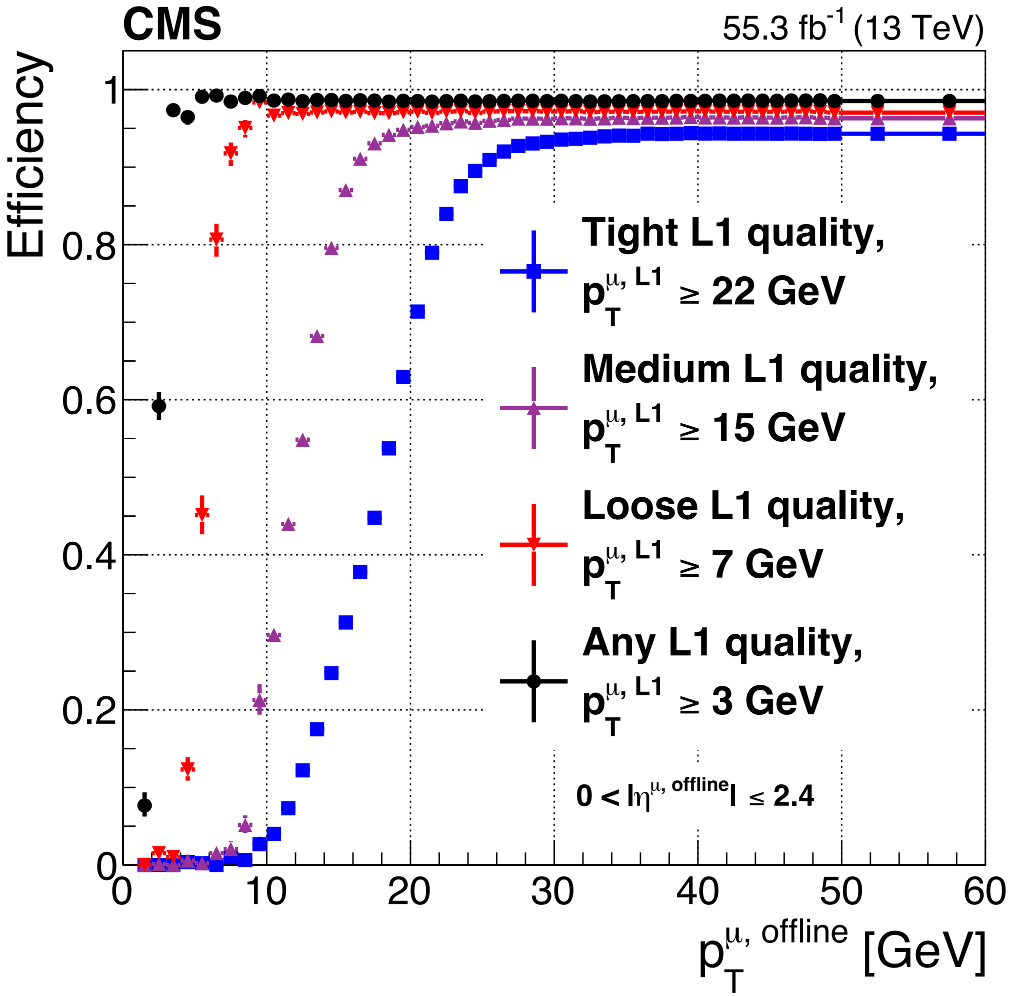
- The CMS Trigger System is robust, flexible and proven in Run 1 and Run 2
  - able to deal with large number of events to fulfill the CMS physics goals
- Excellent performance in Run2: sharp turn-ons, small pileup dependence
- Integrated new technologies (LS2), improved/inovated trigger algorithms
- Additional improvements for the Run3, eg heterogeneous reconstruction comprising on CPU and GPU; further non-conventional triggers (like LLPs)

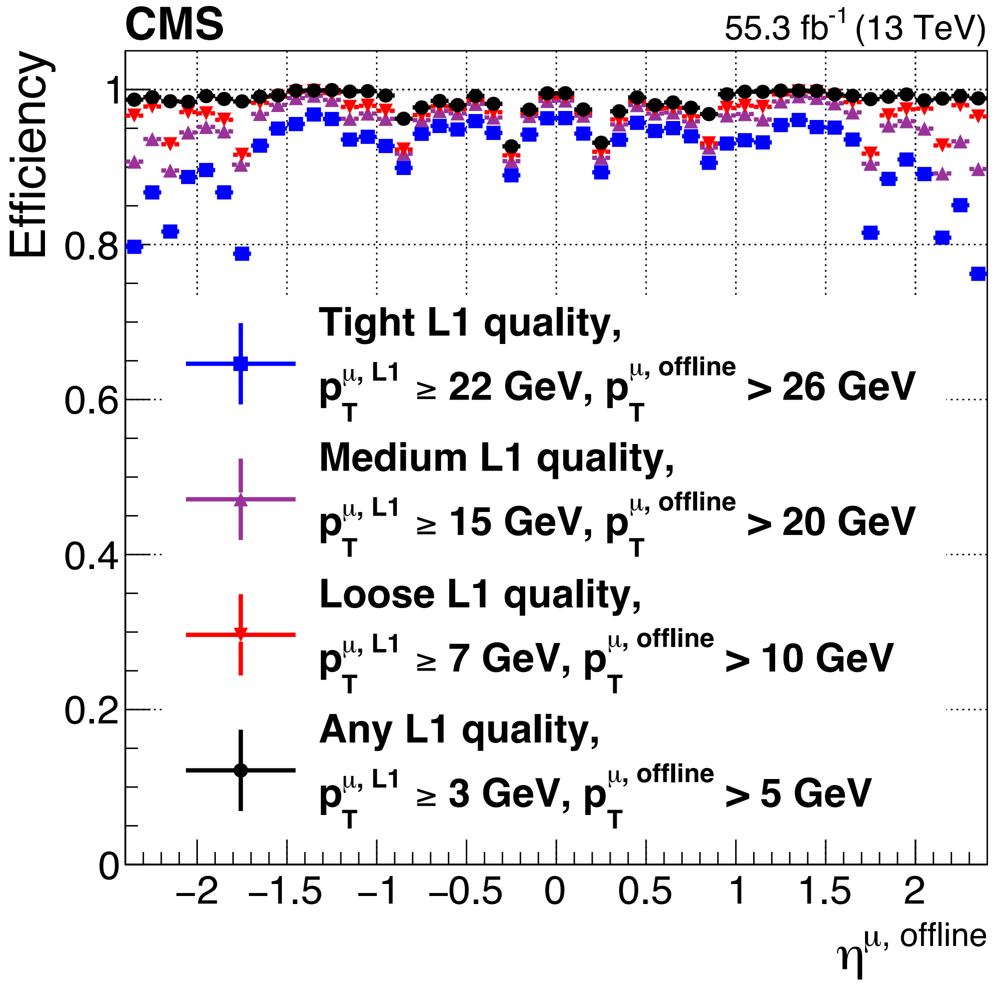
- Run 3 is starting right now -> new data (& exciting physics) around the corner!

BACKUP SLIDES

The Level-1 Trigger Performance: muons

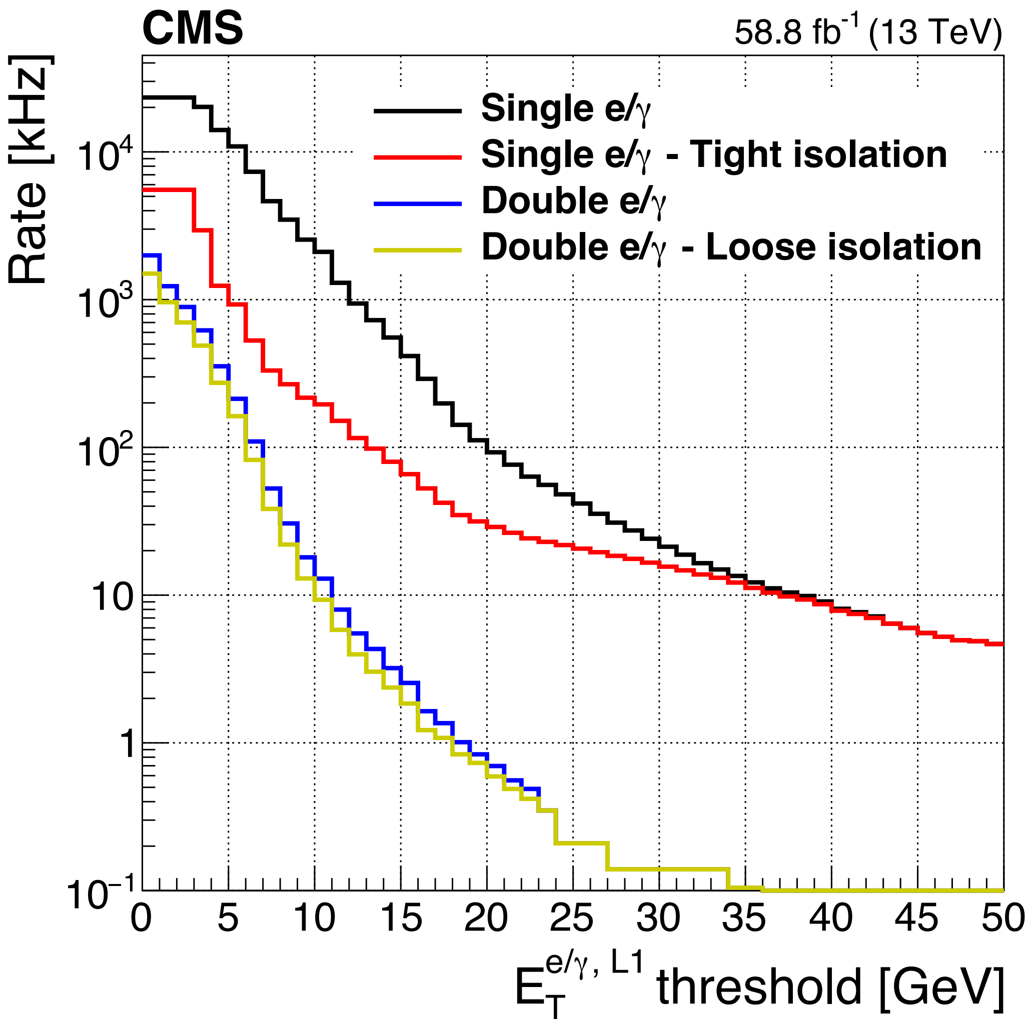
JINST 15, P10017 (2020)

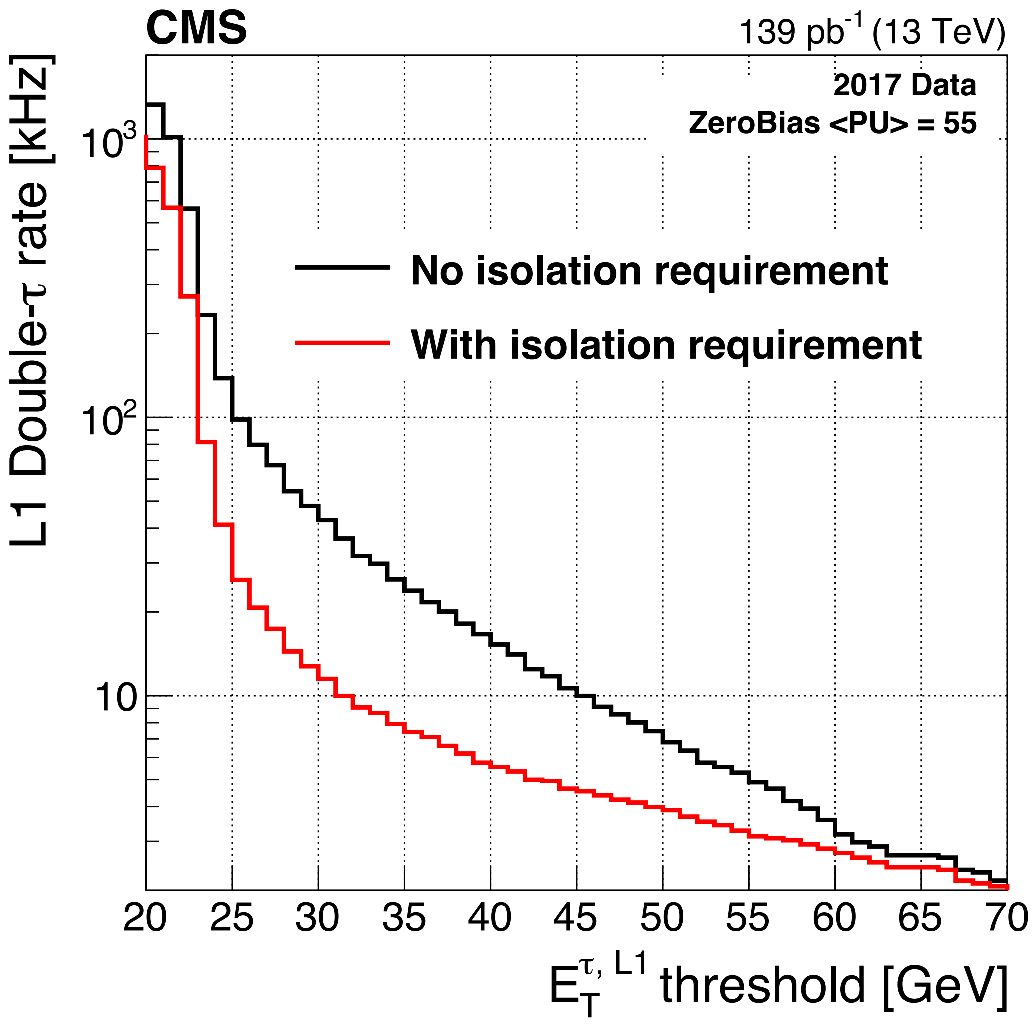




The Level-1 Trigger Performance: ele, pho, tau

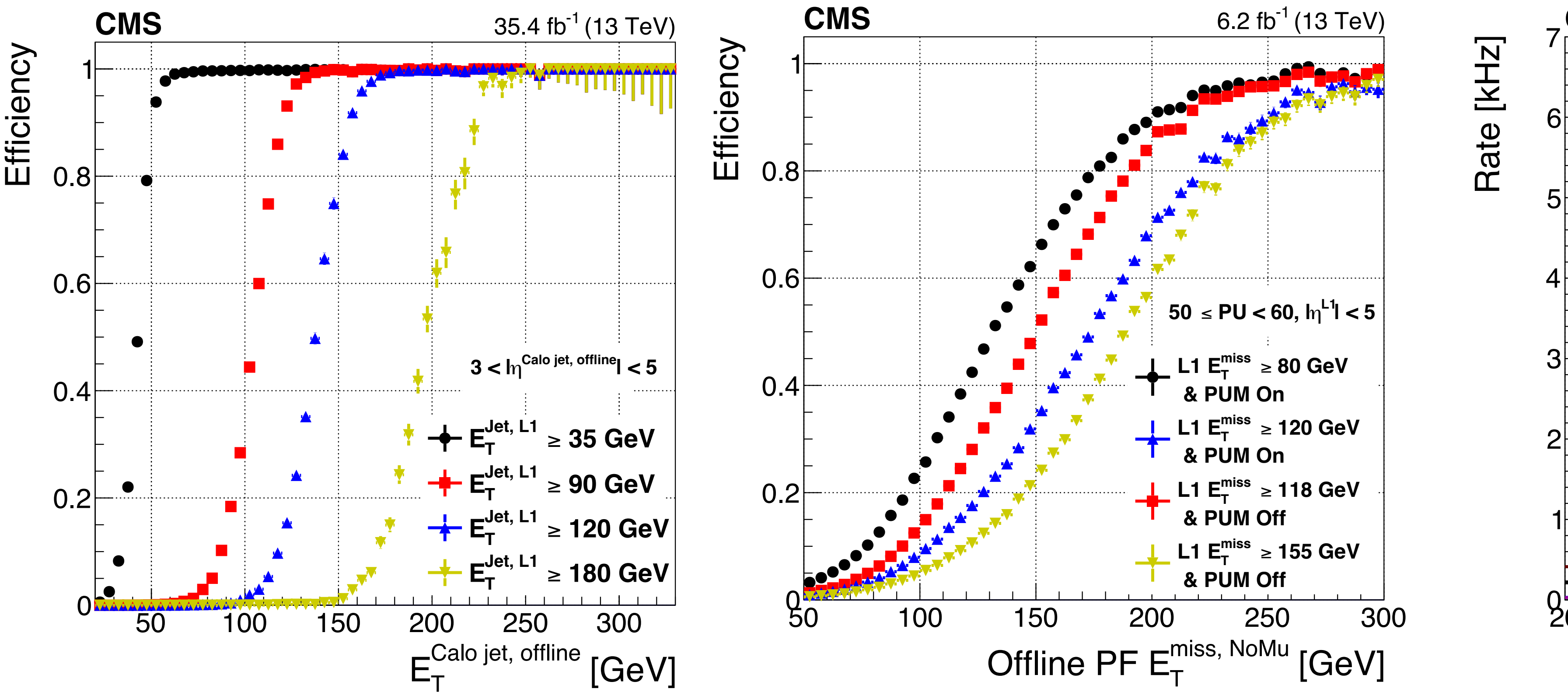
JINST 15, P10017 (2020)





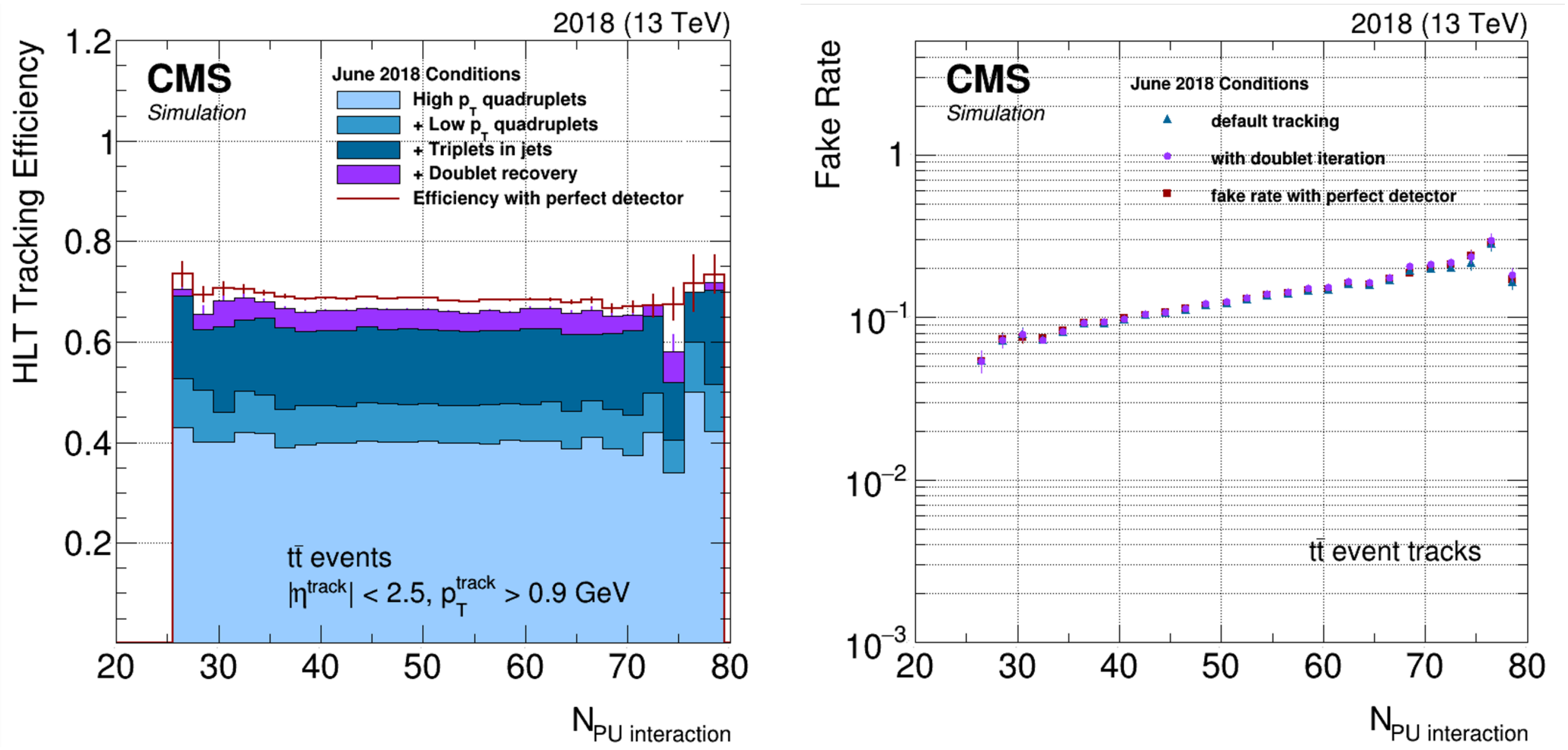
The Level-1 Trigger Performance: jet, HT & MET

JINST 15, P10017 (2020)



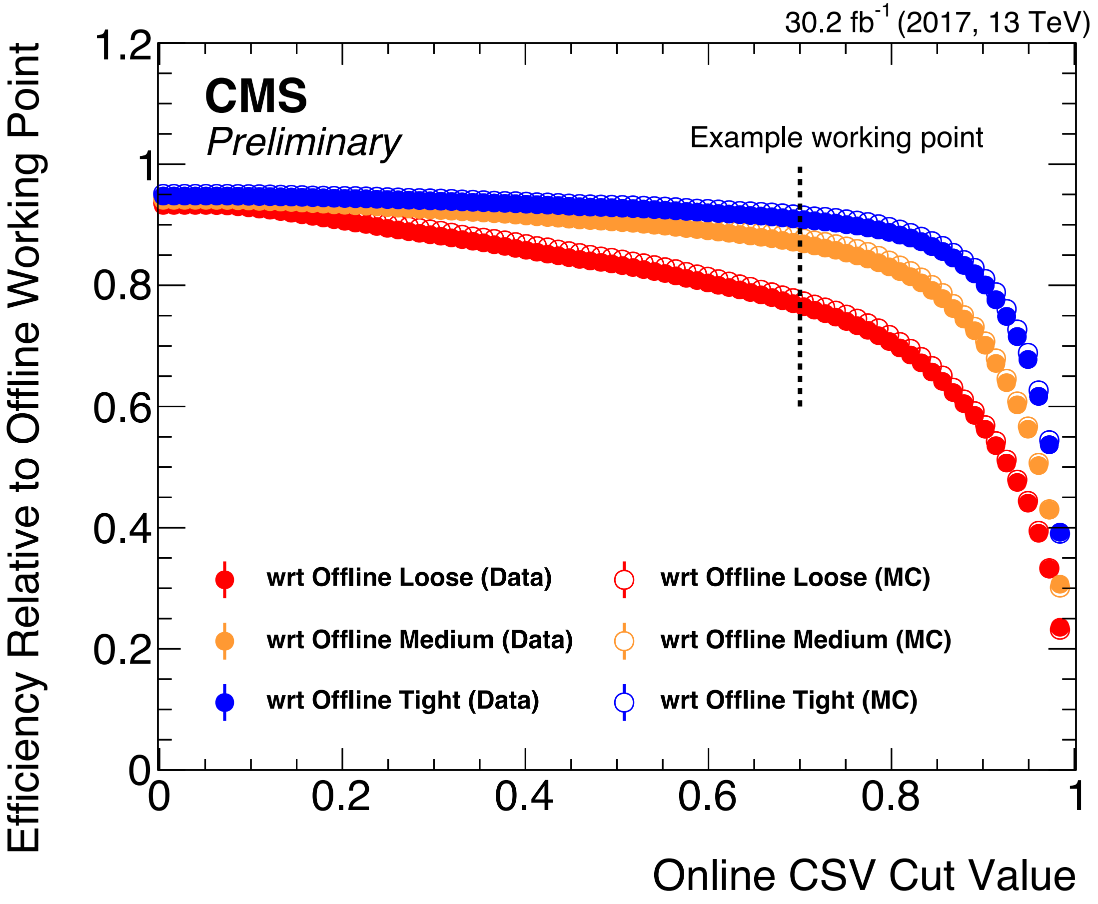
The High Level Trigger: Tracking

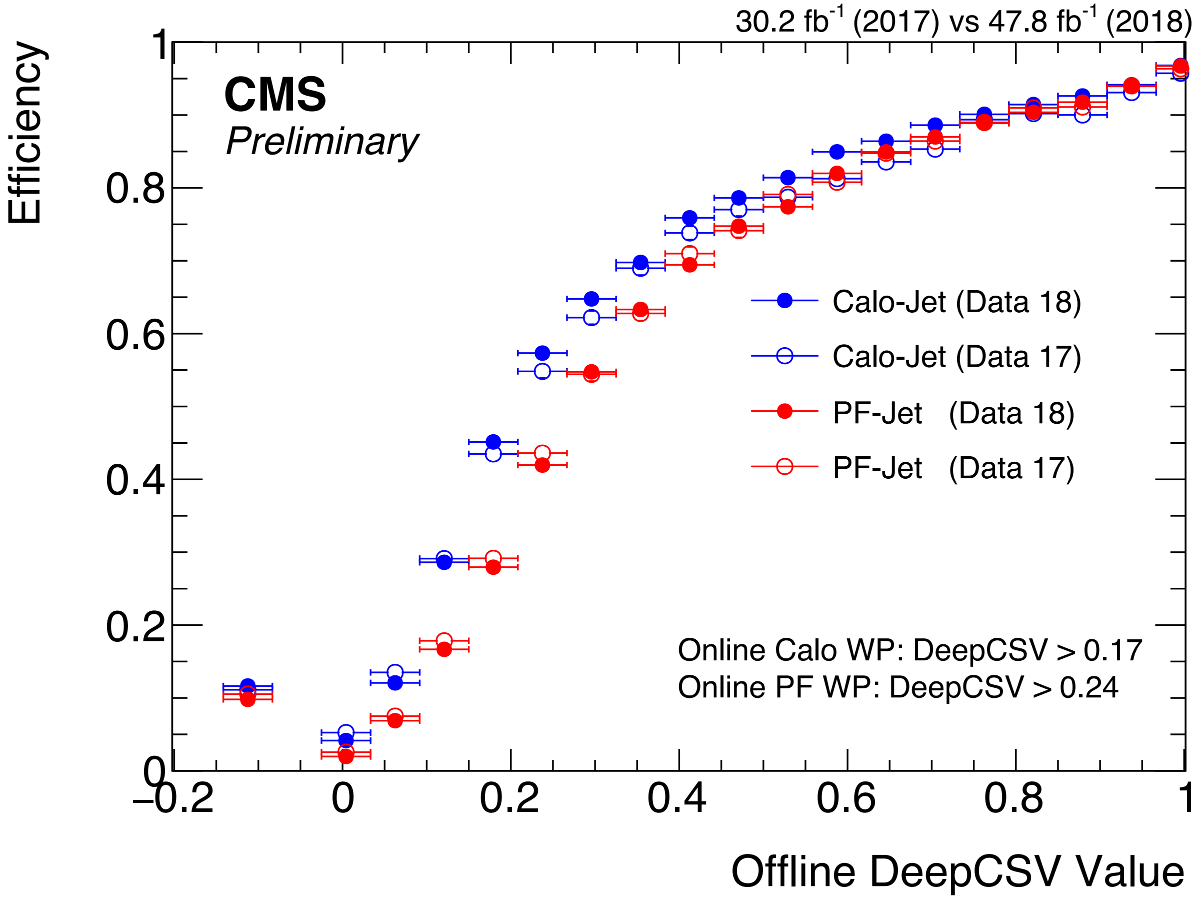
CMS-DP-2018-038



The High Level Trigger: B-Tagging

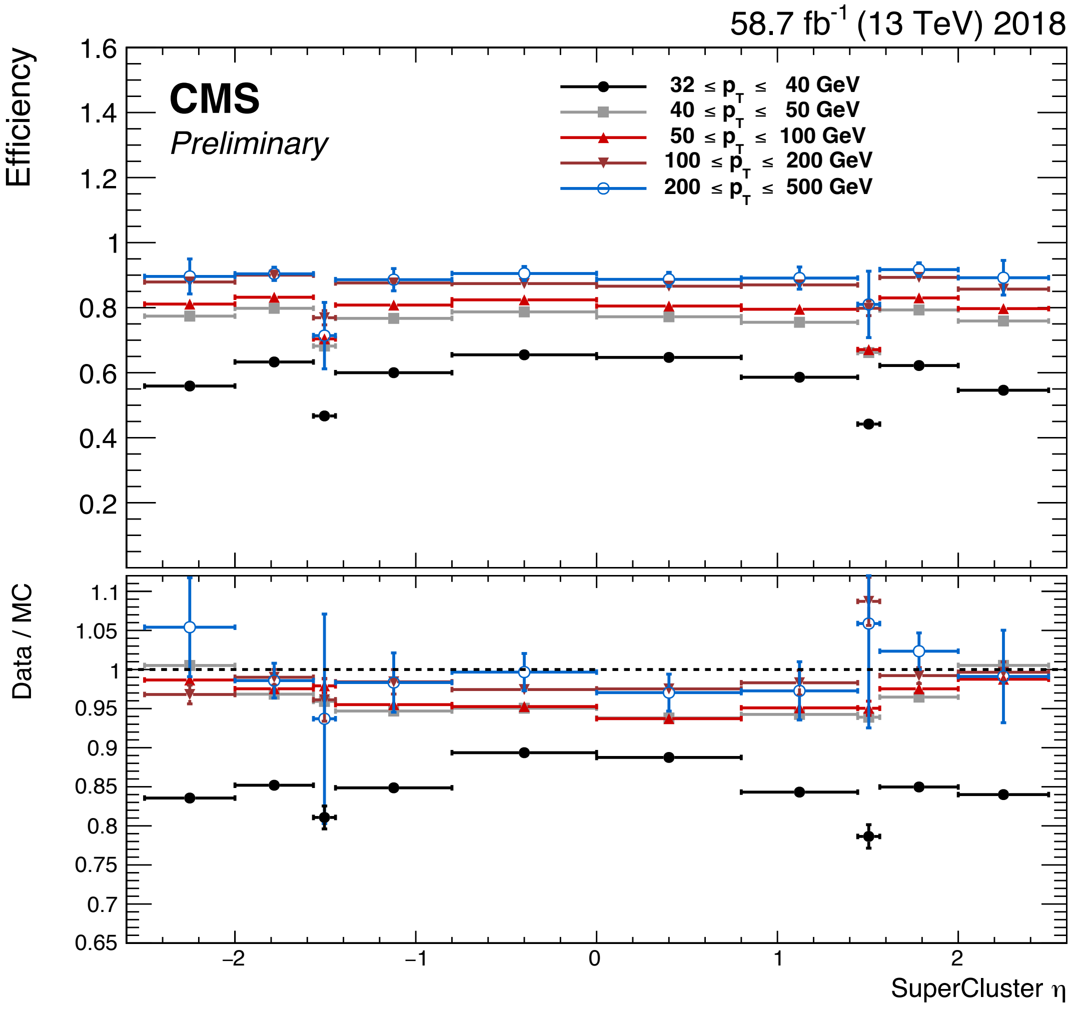
CMS-DP-2019-042

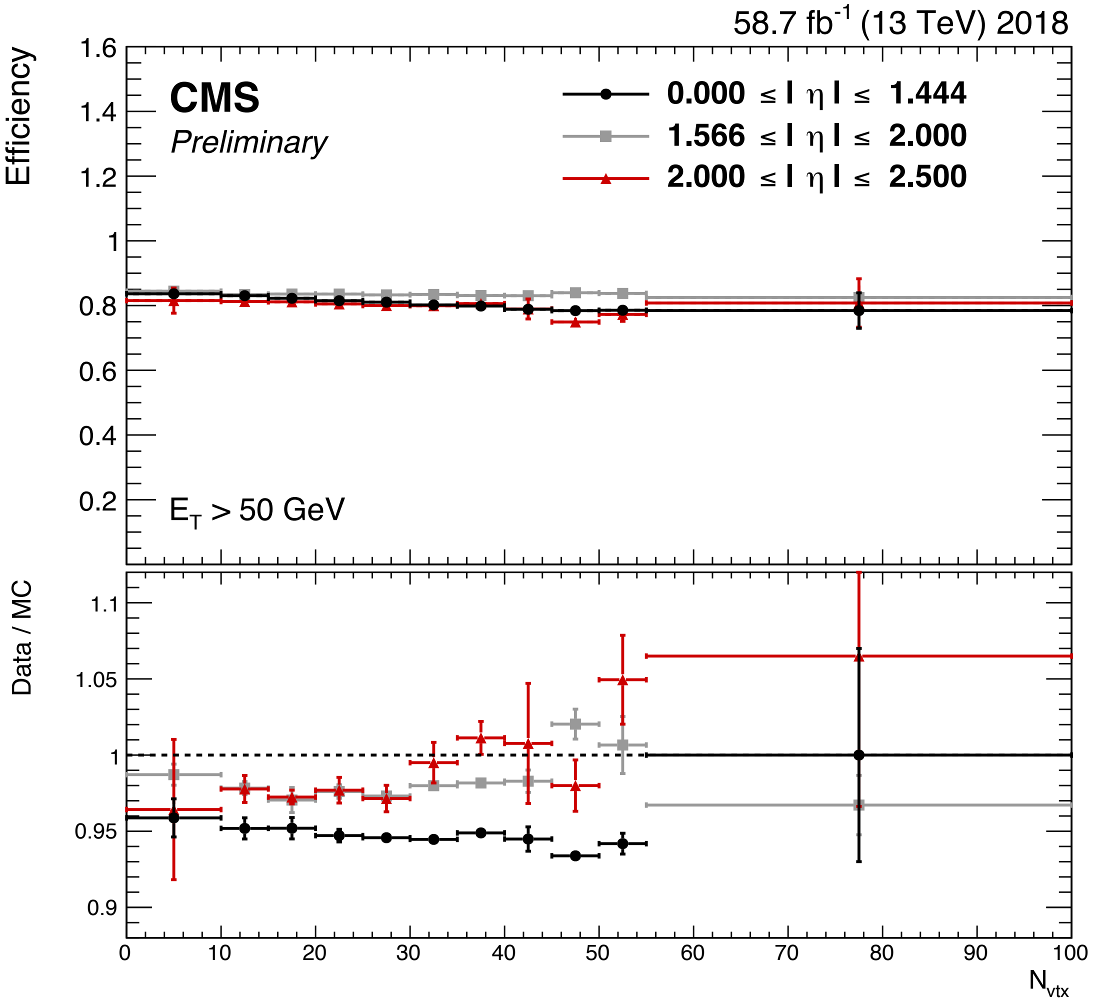




The HLT Performance: Electrons and Photons

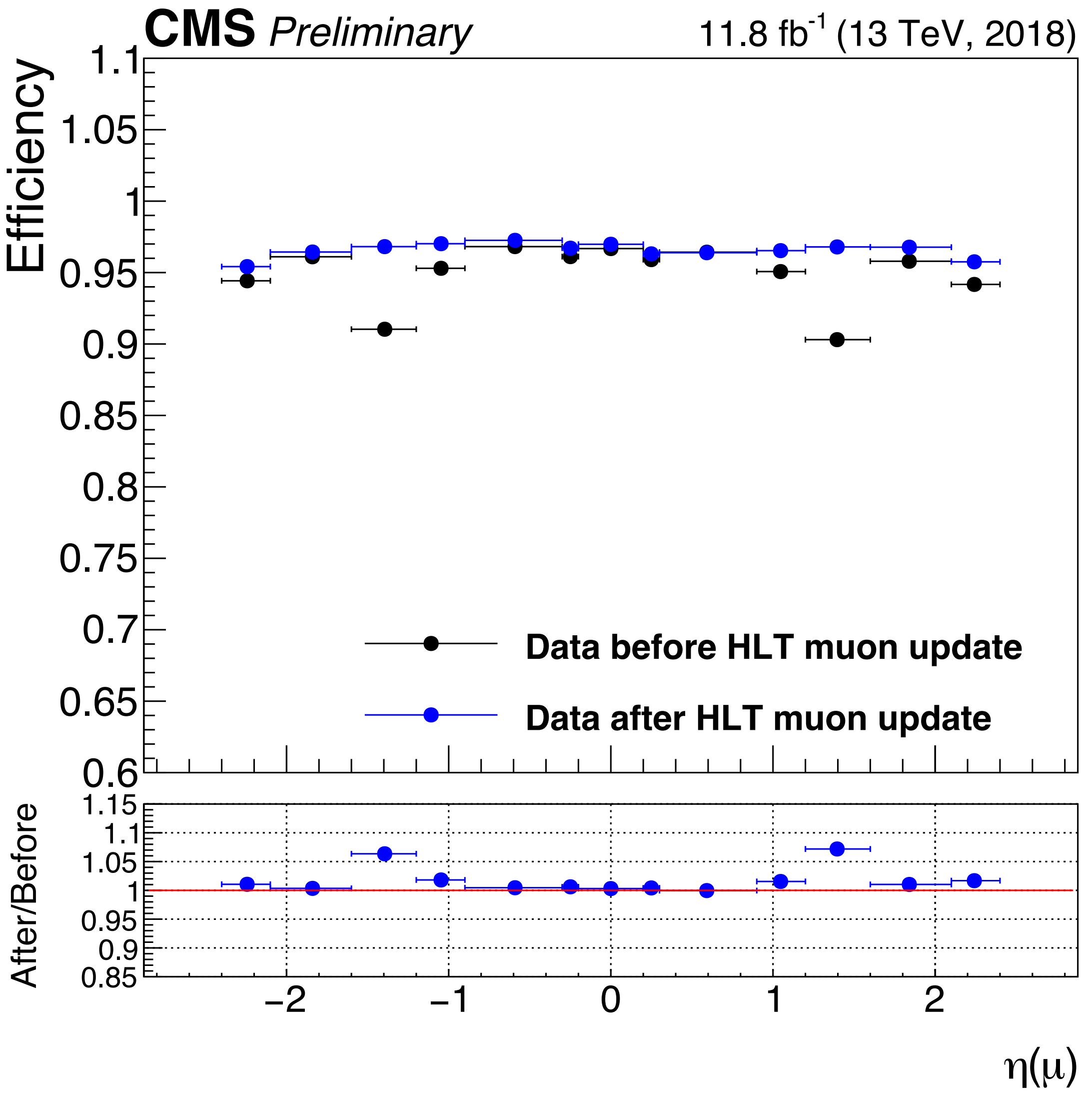
CMS DP-2020/016
CMS DP-2018/049
CMS DP-2018/039

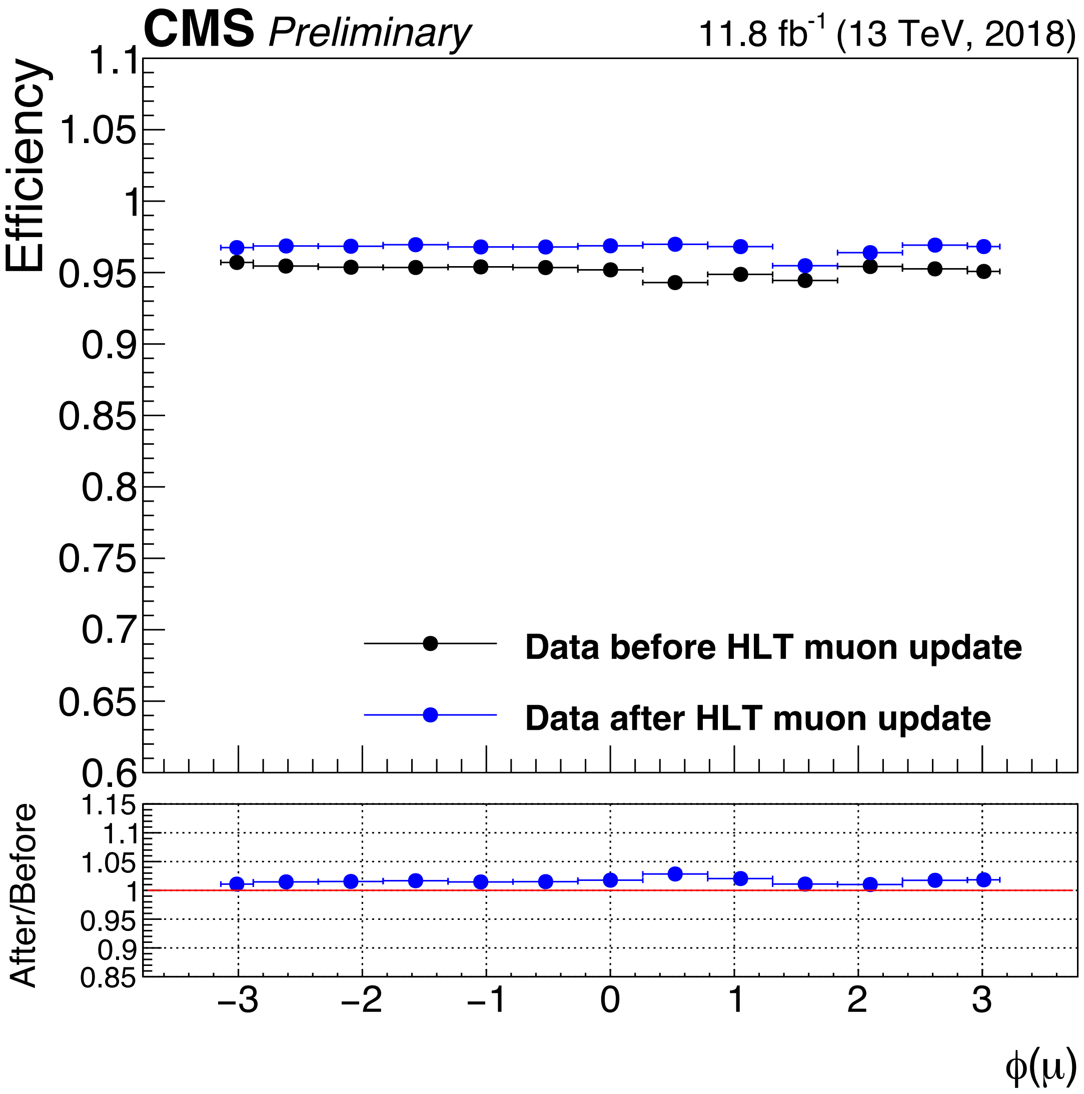




The HLT Performance: Jets, Muons and Taus

CMS DP-2018/034
CMS DP-2019/012
CMS DP-2018/037





The Trigger Menus: How to Select Your Data

JINST 15, P10017 (2020)
CMS DP-2018/057

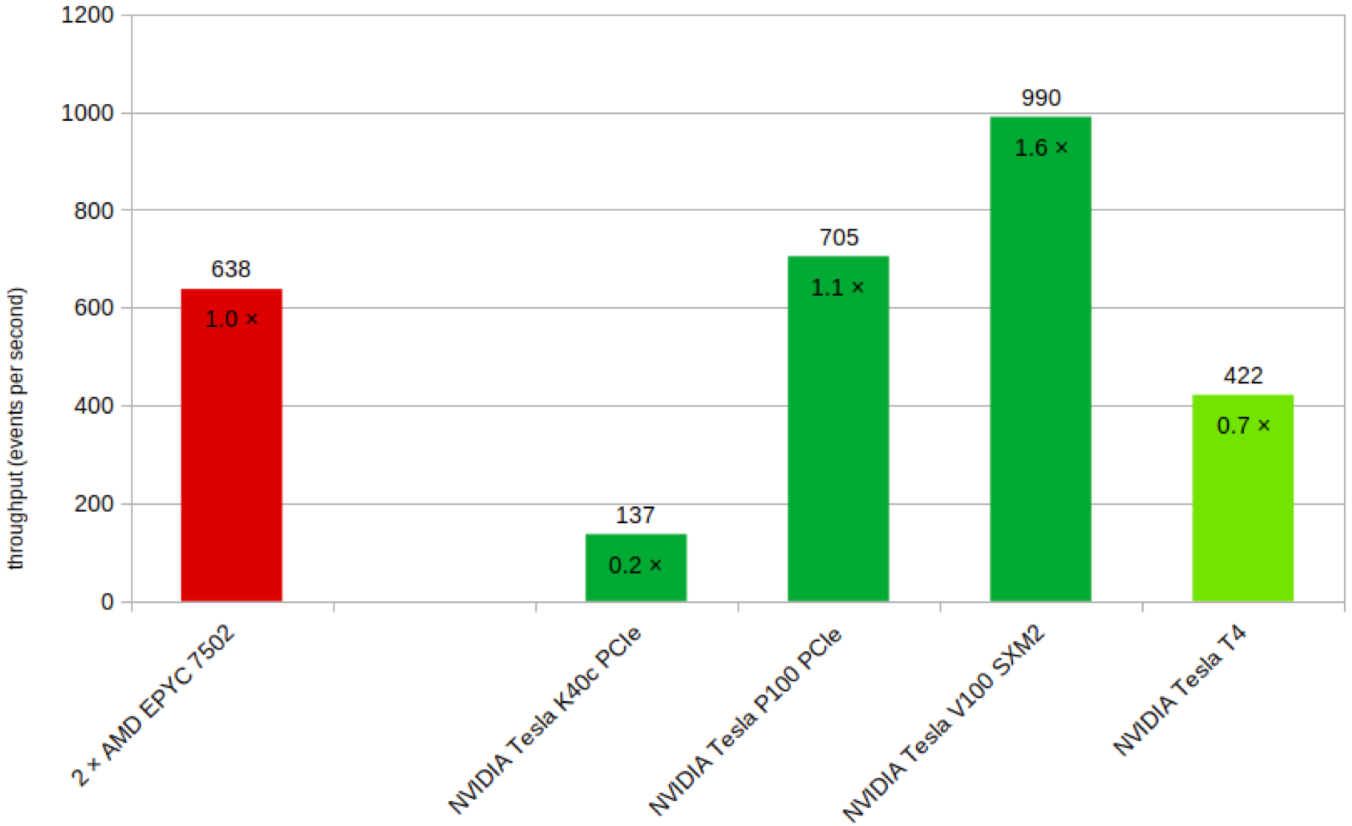
Algorithm	Requirements (p_T, E_T, $m_{\mu\mu}$, and m_{jj} in GeV)
Muons	
Single μ	$p_T > 22$ & Tight quality
Double μ	$p_T > 15, 7$ & Medium quality
Double μ	$p_T > 15, 5$ & Tight quality
Double μ	$p_T > 8, 8$ & Tight quality
Double μ + mass	$p_T > 4.5$ & $\lvert\eta\rvert < 2.0$ & Tight quality & OS & $m_{\mu\mu} > 7$
Double μ + ΔR	$p_T > 4$ & Tight quality & OS & $\Delta R < 1.2$
Double μ + ΔR	$p_T > 0$ & $\lvert\eta\rvert < 1.5$ & Tight quality & OS & $\Delta R < 1.4$
Double μ + BX	$p_T > 0$ & $\lvert\eta\rvert < 1.4$ & Medium quality & Non-colliding BX
Triple μ	$p_T > 5, 3, 3$ & Medium quality
Triple μ	$p_T > 3, 3, 3$ & Tight quality
Triple μ + mass	$p_T > 5, 3.5, 2.5$ & Med. qual.; two μ OS & $p_T > 5, 2.5$ & $5 < m_{\mu\mu} < 17$
Triple μ + mass	Three μ any qual.; two μ & $p_T > 5, 3$ & Tight qual. & OS & $m_{\mu\mu} < 9$
Electrons / photons	(e/γ)
Single e/γ	$p_T > 60$
Single e/γ	$p_T > 36$ & $\lvert\eta\rvert < 2.5$
Single e/γ	$p_T > 28$ & $\lvert\eta\rvert < 2.5$ & Loose isolation
Double e/γ	$p_T > 25, 12$ & $\lvert\eta\rvert < 2.5$
Double e/γ	$p_T > 22, 12$ & $\lvert\eta\rvert < 2.5$ & Loose isolation
Triple e/γ	$p_T > 18, 17, 8$ & $\lvert\eta\rvert < 2.5$
Triple e/γ	$p_T > 16, 16, 16$ & $\lvert\eta\rvert < 2.5$
Tau leptons (τ)	
Single τ	$p_T > 120$ & $\lvert\eta\rvert < 2.1$
Double τ	$p_T > 32$ & $\lvert\eta\rvert < 2.1$ & Isolation
Jets	
Single jet	$p_T > 180$
Single jet + BX	$p_T > 43$ & $\lvert\eta\rvert < 2.5$ & Non-colliding BX
Double jet	$p_T > 150$ & $\lvert\eta\rvert < 2.5$
Double jet + $\Delta\eta$	$p_T > 112$ & $\lvert\eta\rvert < 2.3$ & $\Delta\eta < 1.6$
Double jet + mass	$p_T > 115, 35$; two jets $p_T > 35$ & $m_{jj} > 620$
Double jet + mass	$p_T > 30$ & $\lvert\eta\rvert < 2.5$ & $\Delta\eta < 1.5$ & $m_{jj} > 300$
Triple jet	$p_T > 95, 75, 65$; two jets $p_T > 75, 65$ & $\lvert\eta\rvert < 2.5$
Energy sums	
E_T^{miss}	$E_T^{miss} > 100$ (Vector sum of p_T of calorimeter deposits with $\lvert\eta\rvert < 5.0$)
H_T	$H_T > 360$ (Scalar sum of p_T of all jets with $p_T > 30$ and $\lvert\eta\rvert < 2.5$)
E_T	$E_T > 2000$ (Scalar sum of p_T of calorimeter deposits with $\lvert\eta\rvert < 5.0$)

Terms used
Tight quality: muons with hits in at least 3 different muon stations.
Medium quality: muons with hits in at least 2 different muon stations.
The "non-colliding BX" requirement selects beam-empty events.
$\Delta R \equiv ((\Delta\phi)^2 + (\Delta\eta)^2)^{1/2}$, and phi is the azimuthal angle in radians.
OS: Opposite Sign (of electric charge).
E_T: Scalar sum of p_T of calorimeter deposits.
H_T: Scalar sum of p_T of jets.
Isolation and loose isolation: The isolation requires an upper limit on the transverse calorimeter energy surrounding the candidate. The limit depends on the pileup, the Level-1 candidate E_T and $\lvert\eta\rvert$. Details are given in Sections ?? and ??.

Algorithm	Requirements (p_T, E_T, $m_{\mu\mu}$, and m_{jj} in GeV)
Two objects	
Single μ + Single e/γ	$p_T(\mu) > 20$ & Tight quality(μ) & $p_T(e/\gamma) > 10$ & $\lvert\eta(e/\gamma)\rvert < 2.5$
Single μ + Single e/γ	$p_T(\mu) > 7$ & Tight quality(μ) & $p_T(e/\gamma) > 20$ & $\lvert\eta(e/\gamma)\rvert < 2.5$
Single μ + Single τ	$p_T(\mu) > 18$ & $\lvert\eta(\mu)\rvert < 2.1$ & Tight quality(μ) & $p_T(\tau) > 24$ & $\lvert\eta(\tau)\rvert < 2.1$
Single μ + H_T	$p_T(\mu) > 6$ & Tight quality(μ) & $H_T > 240$
Single e/γ + Single τ	$p_T(e/\gamma) > 22$ & $\lvert\eta(e/\gamma)\rvert < 2.1$ & Loose isolated(e/γ) & $p_T(\tau) > 26$ & $\lvert\eta(\tau)\rvert < 2.1$ & Isolated(τ) & $\Delta R > 0.3$
Single e/γ + Single jet	$p_T(e/\gamma) > 28$ & $\lvert\eta(e/\gamma)\rvert < 2.1$ & Loose isolated(e/γ) & $p_T(jet) > 34$ & $\lvert\eta(jet)\rvert < 2.5$ & $\Delta R > 0.3$
Single e/γ + H_T	$p_T(e/\gamma) > 26$ & $\lvert\eta(e/\gamma)\rvert < 2.1$ & Loose isolated(e/γ) & $H_T > 100$
Single τ + E_T^{miss}	$p_T(\tau) > 40$ & $\lvert\eta(\tau)\rvert < 2.1$ & $E_T^{miss} > 90$
Single jet + E_T^{miss}	$p_T(jet) > 140$ & $\lvert\eta(jet)\rvert < 2.5$ & $E_T^{miss} > 80$
Three objects	
Single μ Double jet + ΔR	$p_T(\mu) > 12$ & $\lvert\eta(\mu)\rvert < 2.3$ & Tight quality(μ) & $p_T(jet) > 40$ & $\Delta\eta(jet,jet) < 1.6$ & $\lvert\eta(jet)\rvert < 2.3$ & $\Delta R(\mu, jet) < 0.4$
Single μ + Single jet + E_T^{miss}	$p_T(\mu) > 3$ & $\lvert\eta(\mu)\rvert < 1.5$ & Tight quality (μ) & $p_T(jet) > 100$ & $\lvert\eta(jet)\rvert < 2.5$ & $E_T^{miss} > 40$
Double μ + H_T	$p_T(\mu) > 3$ & Tight quality(μ) & $H_T > 220$
Double μ + Single jet + ΔR	$p_T(\mu) > 0$ & Medium quality(μ) & $\Delta R(\mu, \mu) < 1.6$ & $p_T(jet) > 90$ & $\lvert\eta(jet)\rvert < 2.5$ & $\Delta R(\mu, jet) < 0.8$
Double μ + Single e/γ	$p_T(\mu) > 5$ & Tight quality(μ) & $p_T(e/\gamma) > 9$ & $\lvert\eta(e/\gamma)\rvert < 2.5$
Double e/γ + Single μ	$p_T(e/\gamma) > 12$ & $\lvert\eta(e/\gamma)\rvert < 2.5$ & $p_T(\mu) > 6$ & Tight quality(μ)
Double e/γ + H_T	$p_T(e/\gamma) > 8$ & $\lvert\eta(e/\gamma)\rvert < 2.5$ & $H_T > 300$
Four objects	
Double μ + Double e/γ	$p_T(\mu) > 3$ & Medium quality(μ) & OS(μ) & $p_T(e/\gamma) > 7.5$
Double μ + Double e/γ	$p_T(\mu) > 5$ & Medium quality(μ) & OS(μ) & $p_T(e/\gamma) > 3$
Five objects	
Double μ + E_T^{miss} + Single jet OR Double jet	$p_T(\mu) > 3$ & Tight quality(μ) & $E_T^{miss} > 50$ & ($p_T(jet) > 60$ & $\lvert\eta(jet)\rvert < 2.5$) OR ($p_T(jet) > 40$ & $\lvert\eta(jet)\rvert < 2.5$)
H_T + Quad jet	$H_T > 320$ & $p_T(jet) > 70, 55, 40, 40$ & $\lvert\eta(jet)\rvert < 2.4$

The CMS Trigger System in LHC Run 3: GPU Acceleration

CMS DP-2021/013



GPU vs CPU throughput

The histogram shows the absolute and relative throughput of the part of the HLT reconstruction that can be offloaded to GPUs, running on different hardware:

- the reference, in red, is a dual processor machine with 2× AMD “Rome” EPYC 7502 CPUs (from 2019);
- three generations of high power (250 W), dual-slot NVIDIA datacenter GPUs are shown in dark green: a Tesla K40 (from 2013), a Tesla P100 (from 2017), and a Tesla V100 (from 2018);
- the performance of a low power (70 W), single slot NVIDIA datacenter Tesla T4 (from 2019) is shown in light green.