Tau tagging in Atlas and CMS

Giuseppe Bagliesi

INFN-Pisa

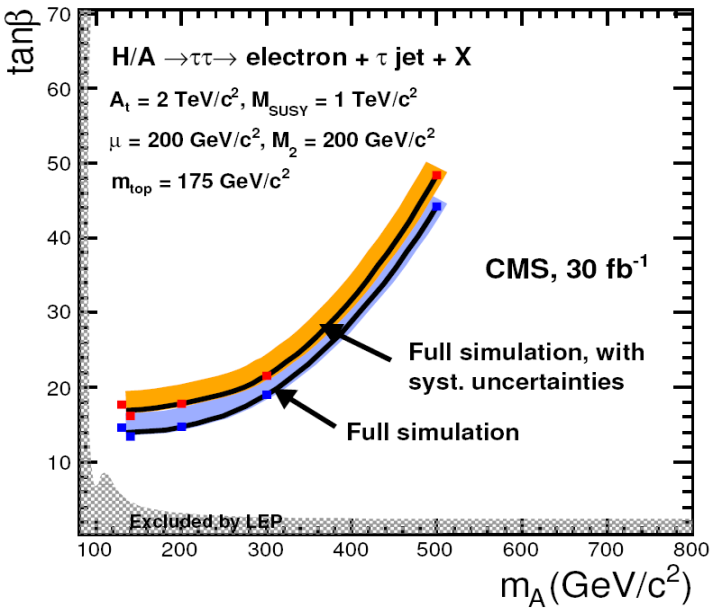
Giuseppe.Bagliesi@pi.infn.it

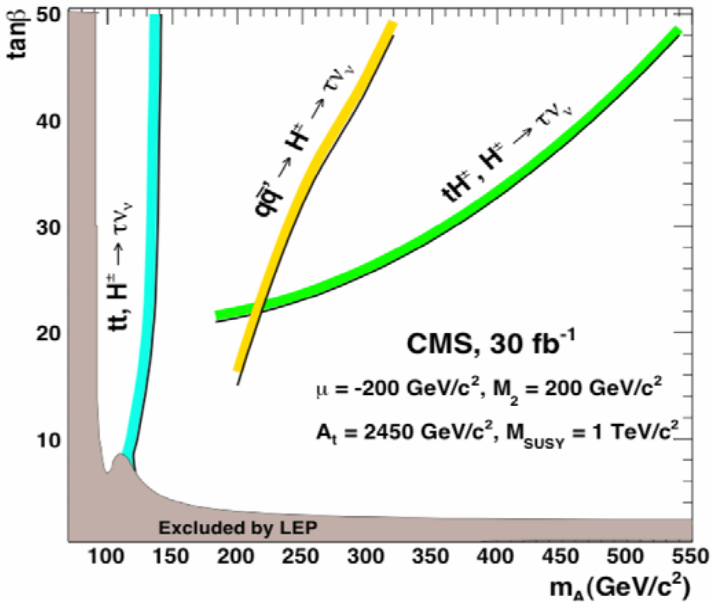
Outlook

- Motivations: tau physics at LHC
- Atlas and CMS
- tau identification
  - algorithms performance
- tau trigger strategy
- conclusions

tau physics at LHC

- Standard model processes
  - Z -> ττ, W -> τν
    - Useful for validation, tuning of the algorithms, calibration
- Searches
  - Many relevant discovery channels involve tau leptons in the final state:
    - MSSM Higgs (A/H, H^{+})
      - A/H -> ττ, H^{+}-> τν
    - SM Higgs
      - qqH->qqττ, ttH->tttτ





tau decay modes

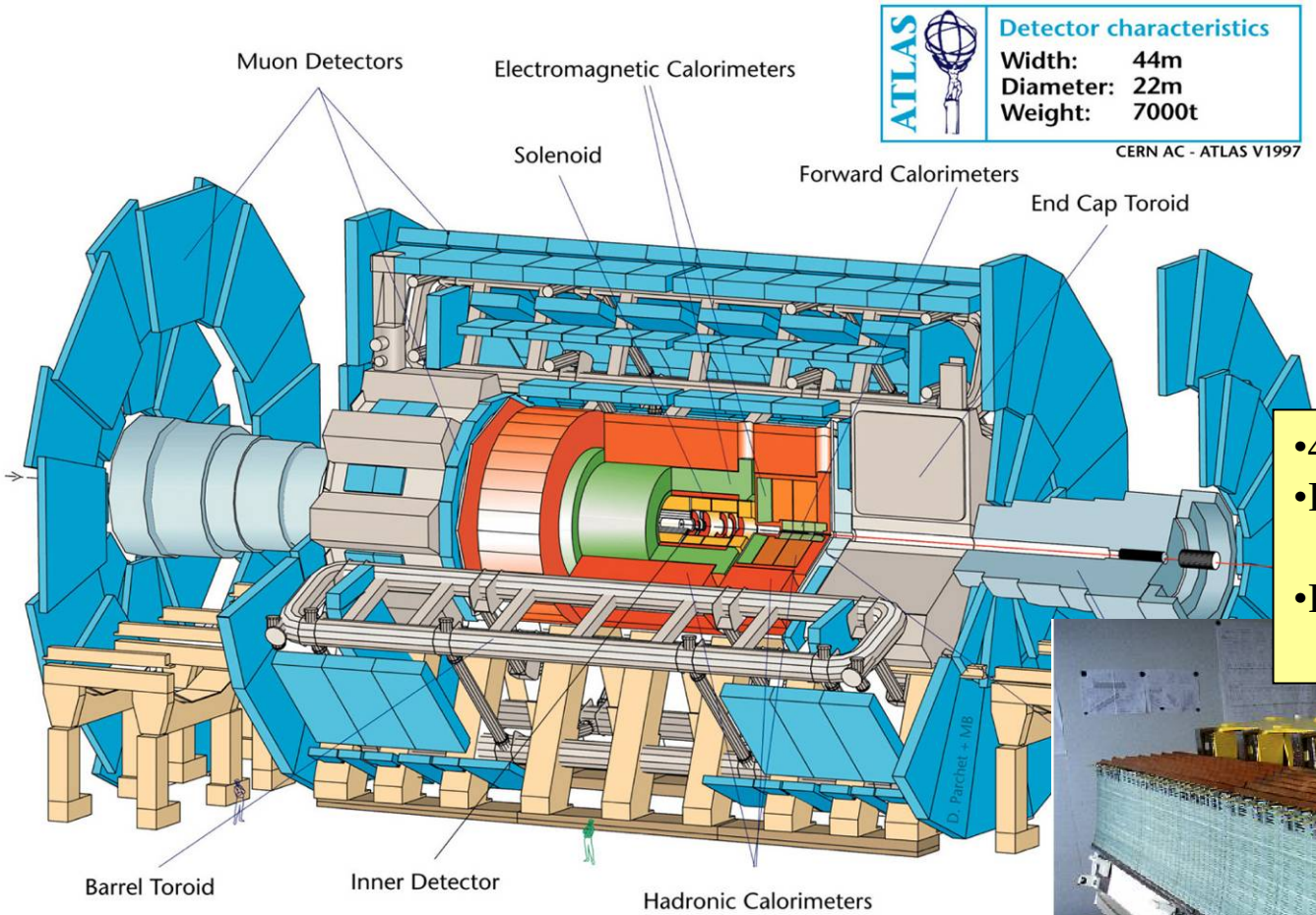
Quite often taus are produced in pairs: 42% of final states contains two "tau-jet"

- $c\tau \sim= 87\mu m$, $m_\tau = 1.78$ Gev/c^2
- Leptonical decays
  - $\tau \rightarrow e(\mu)\ \nu\ \nu$: ~ 35.2 %
    - Identification done through the final lepton
- Hadronical decays
  - 1 prong
    - $\tau \rightarrow \nu_\tau + \pi^{+/-} + n(\pi^o)$: 49.5 %
  - 3 prongs
    - $\tau \rightarrow \nu_\tau + 3\pi^{+/-} + n(\pi^o)$: 15.2 %
  - **"τ–jet" is produced**

$\tau\tau$ decay mode	BR
$\ell\ell\,\nu$	12 %
ℓjet ν	46 %
jet jet ν	42 %

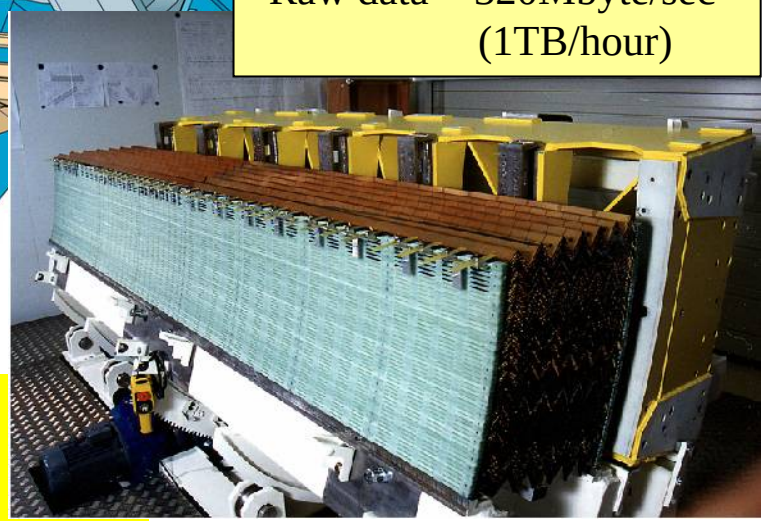
- tau jets at LHC:
  - very collimated
    - **90% of the energy is contained in a 'cone' of radius R=0.2 around the jet direction for ET>50 GeV**
  - Low multiplicity
    - **One, three prongs**
  - Hadronic, EM energy deposition
    - **Charged pions**
    - **Photons from π^o**

ATLAS (A Toroidal LHC ApparatuS)



E,P resolution
(P ~ 100GeV) :
μ ~ 2%
e, γ ~ 1.5%
Jets ~ 8%

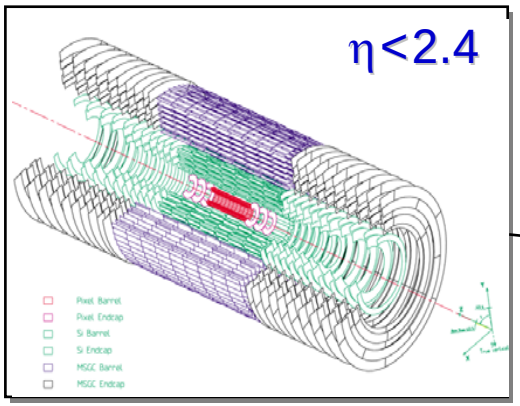
- 40MHz beam crossing
- Readout channel = 160M channels
- Raw data = 320Mbyte/sec (1TB/hour)



- Inner tracking system : 2T solenoid magnet
- Liq. Ar EM calorimeter
- Muon spectrometer : air-core troidal magnet

CMS (Compact Muon Solenoid)

TRACKER(s)



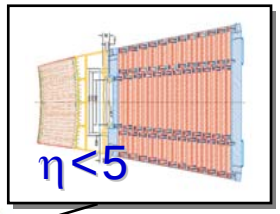
Pixels and Silicon Microstrips

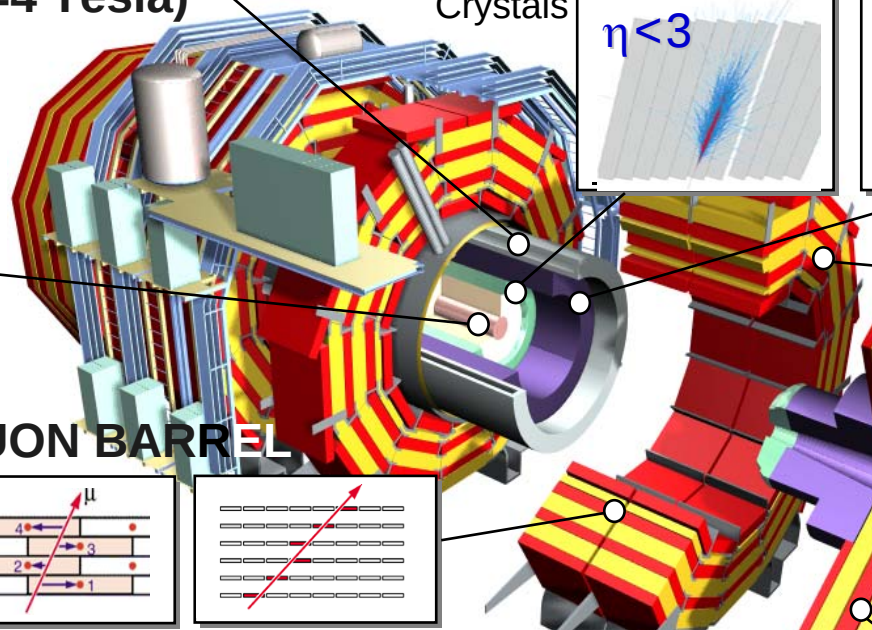
SUPERCONDUCTING COIL (B=4 Tesla)

CALORIMETERS:

ECAL Scintillating $PbWO_4$ Crystals

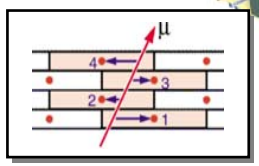
HCAL Plastic scintillator copper sandwich





IRON YOKE

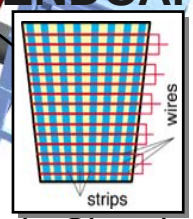
MUON BARREL



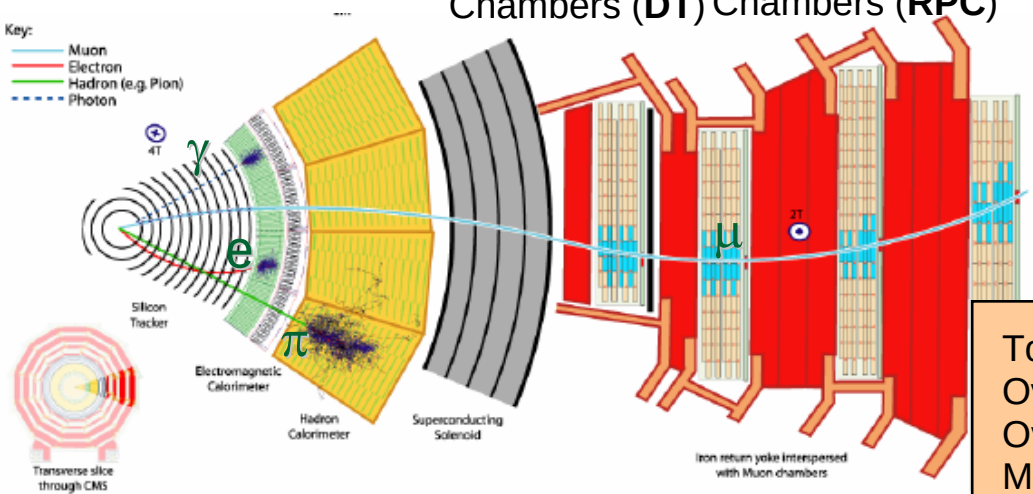
Drift Tube Chambers (**DT**)

Resistive Plate Chambers (**RPC**)

MUON ENDCAPS



Cathode Strip Chambers (**CSC**)
Resistive Plate Chambers (**RPC**)



Total weight : 12,500 t
Overall diameter : 15 m
Overall length : 21.6 m
Magnetic field : 4 Tesla

How hadronic tau decays can be identified

- tau tagging basic ingredients
  - Calorimetric isolation and shape variables
  - Charged tracks isolation
  - Other tau characteristics suitable for tagging:
    - **Impact parameter**
    - **Decay length**
    - **Invariant Mass**

- **Main backgrounds for taus**
  - **QCD jets**
  - Electron that shower late or with strong bremstrahlung
  - Muons interacting in the calorimeter

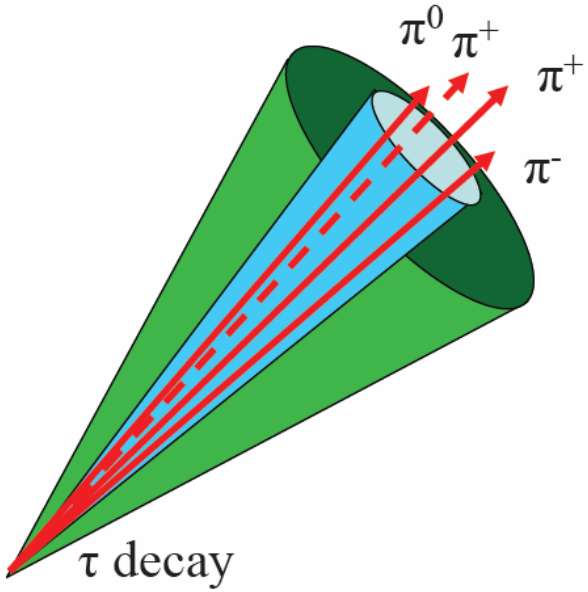
- **Only hadronic tau decays are considered since the leptonic tau decays produce standard electrons/muons**
  - Identification through the final lepton

Atlas: two algorithms have been developed

- TauRec
  - Calorimeter + inner detector are used
  - Energy from calorimeter is calibrated
  - Tracks+clusters are associated
  - Several suitable variables are computed
    - Likelihood or variable cut used for tau-id
- Tau1P3P
  - Intended for studies of low mass Higgs
    - Visible tau hadronic energy 20-70 GeV
  - Identify a leading track
    - Single prong (Tau1P) or three-prong (Tau3P) candidate
  - Use Energy Flow for Energy Calibration
  - Compute set of discriminating variables
    - Apply set of basic cut for tau-id or use Multi-variate classification technique

Atlas: tau-jets reconstruction

- Tau candidate reconstruction
  - Collimated Calorimeter cluster
  - small number of associated charged tracks
  - Selection variables definition
    - R_{em}= EM jet radius
    - ΔE_T^{12}= fraction of E_T in EM/hadronic calo within $0.1<\Delta R<0.2$
    - N_{tracks}, Charge, I.P., strip width in the η-strips

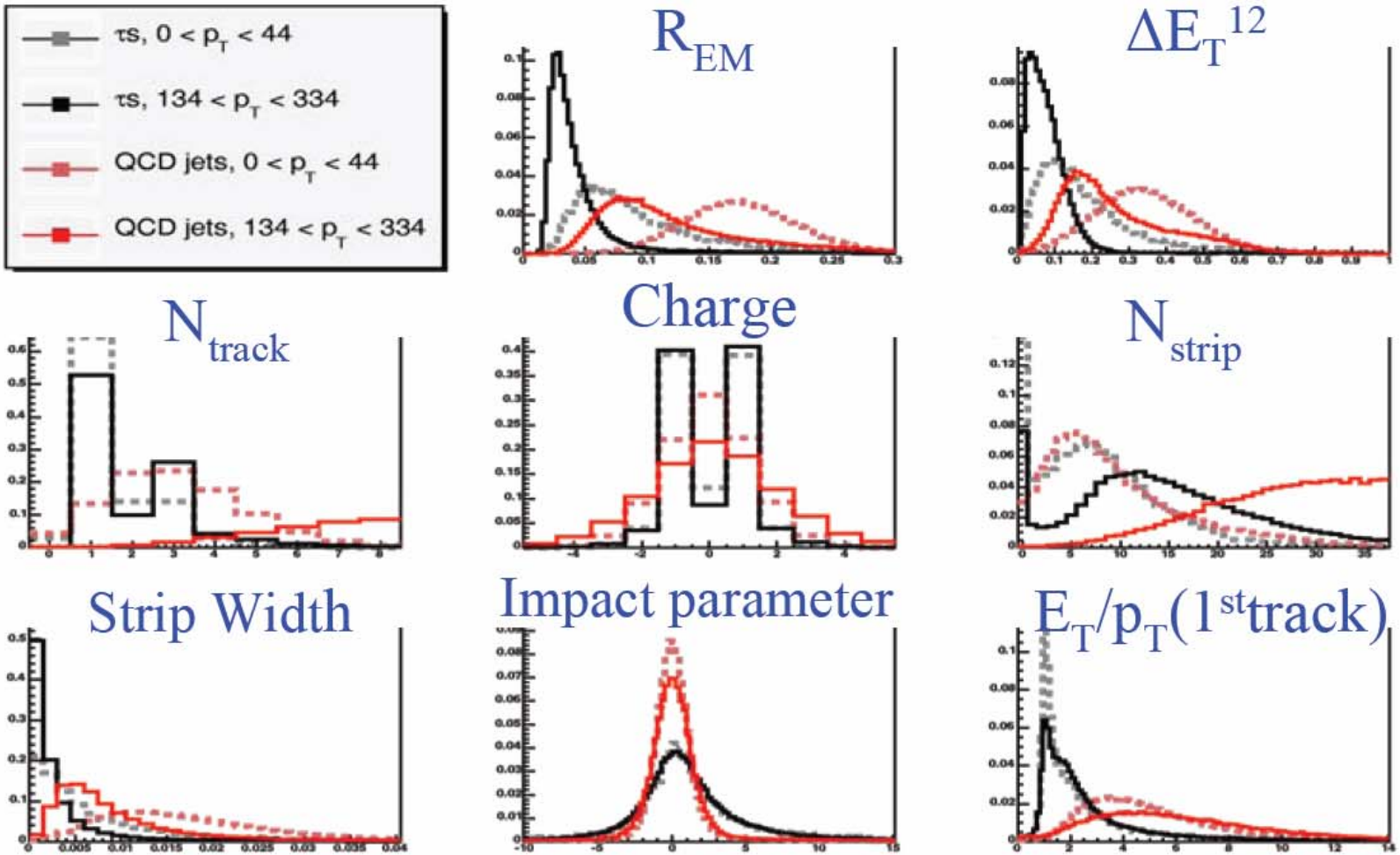


Atlas: TauRec
Default tau-jet reconstruction in Atlas

- Preliminary requirements:
  - abs(η)<2.5
  - Collimated Calorimeter Clusters (E_T>15 GeV)
  - One or three associated "good" charged tracks
    - P_T>2GeV
    - ΔR(track,cluster center)<0.3

- **Variables used:**
  - **R_{EM}:** electromagnetic radius
  - **ΔE_T^{12}**: isolation in calorimeter
  - **N_{track}:** number of assoc. tracks
  - **Charge**
  - **N_{strip}:** #hits in the η-strip ecal layer
  - **Strip width** of η–strip ecal layer
  - **2D signed IP**
  - **E_T/P_T** of leading track

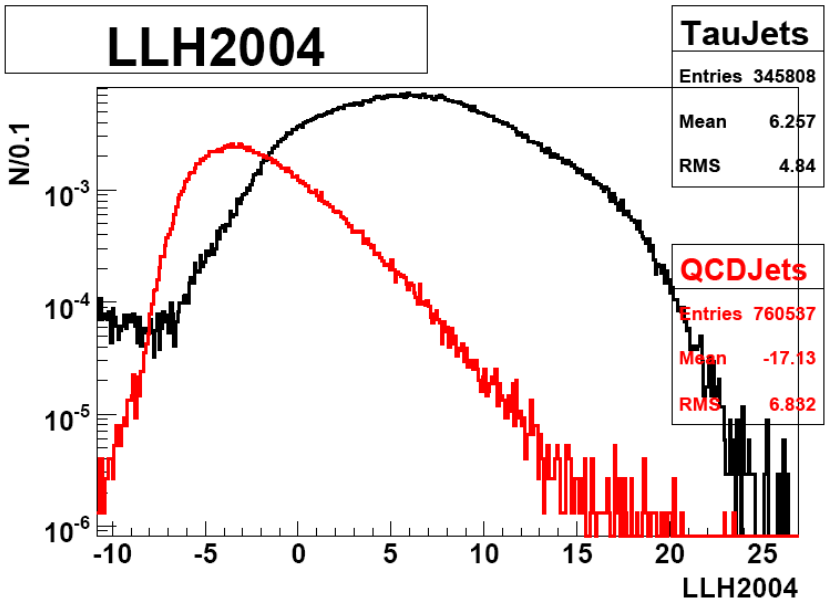
Atlas TauRec: variables distribution

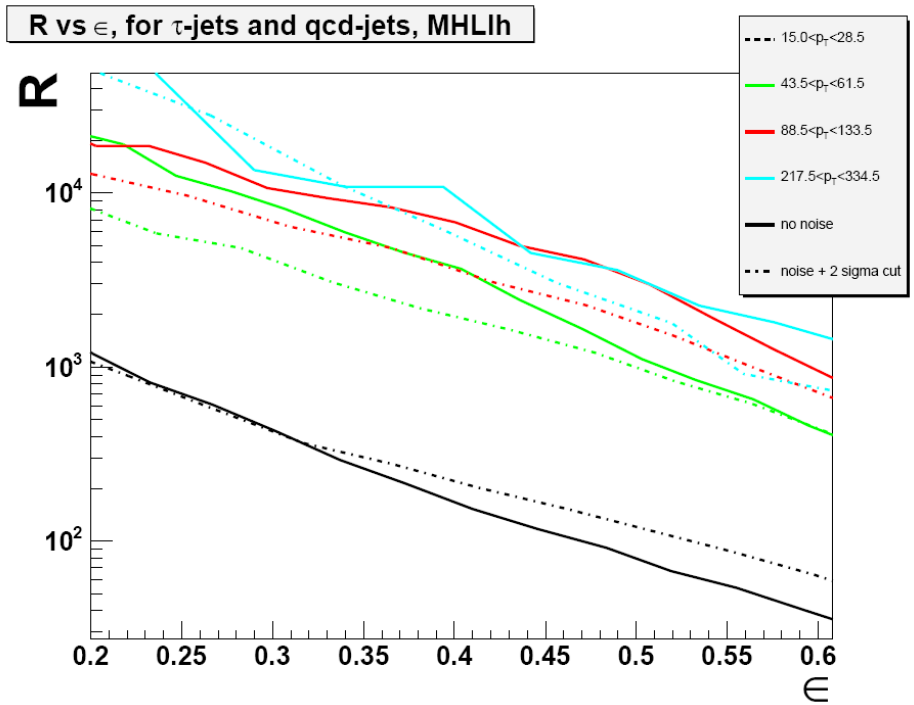


Samples :

- **Signal**
  A/H→ττ
- **Bkg**
  QCD jets

Atlas Taurec: likelihood and performance



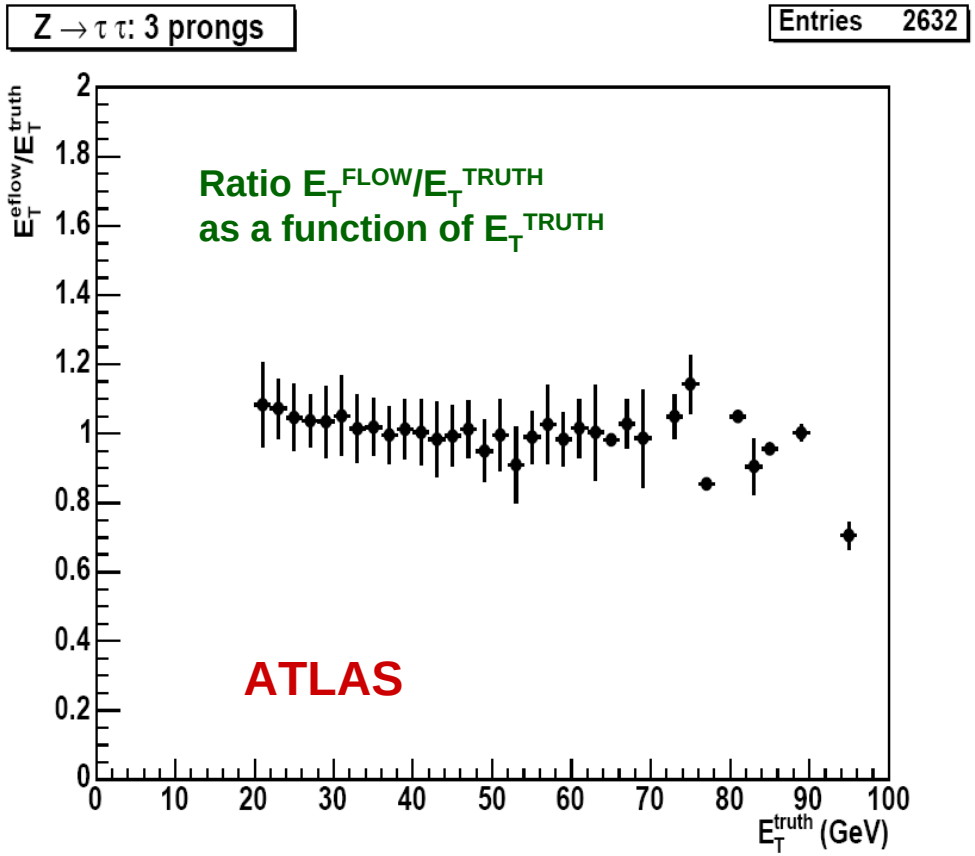


- Likelihood built with the previously described variables
- N.B.:Cut on likelihood depend on E_T

- Rejection against QCD jets vs Efficiency for taus
- various E_T bins are considered

Atlas: Tau1P3P

- Soft taus ($20<E_T^{VIS}<70$ GeV), abs(η)<1.5
- Use of tau-jets features
  - 1(3) track(s) + nπ^o
- Track(s) direction gives a good estimate of the tau direction
- Energy-flow algorithms widely used



Energy flow performance

Atlas: Tau1P3P

- Tau candidates preselection
  - "Good" hadronic leading track P_T>9 GeV
  - 0 or 2 nearby-track: P_T>2GeV, ΔR<0.2
  - Build E_T^{flow} (ΔR<0.2)
  - Require isolation in ring 0.2< ΔR<0.4
  - Calculate several discriminant variables like in TauRec case

Atlas Tau1P3P: performance

Z->ττ–>jet jet sample

- E_{rec}= 82.6%
  - 90.3 (1P) 62% (3P)
- E_{ident}= 59.1%
  - ~same for 1P and 3P
- E_{presel}=39%
  - of all hadronic decays
- Total eff = 18.9 %

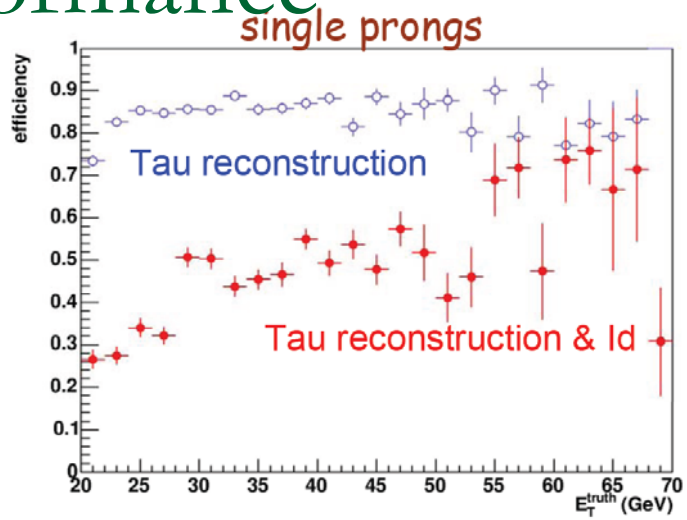
QCD sample

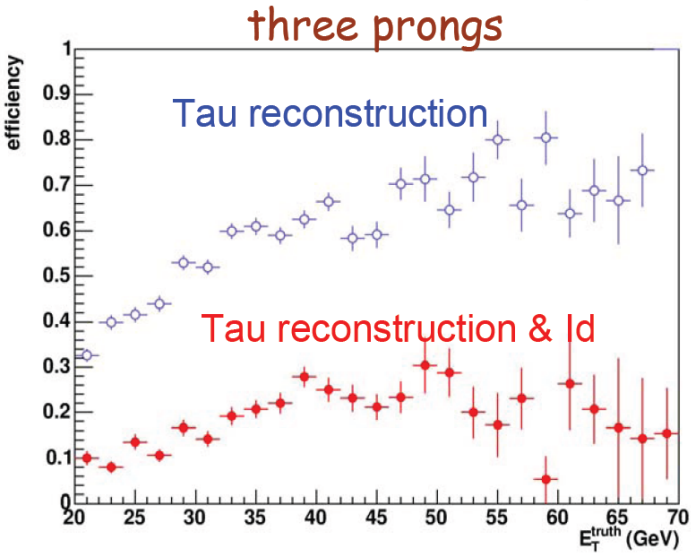
- E_{rec}= 2.0 % (1P) 4.2% (3P)
- E_{ident}= 14.7% (1P) 24.6 (3P)
- Total eff = 0.3 % (1P) 1% (3P)

$eff_{RECONSTRUCTION}$ = Nτ(Reconstruction+τ labeled)/ Nτ(in acceptance)

$eff_{IDENTIFICATION}$ = Nτ(Reconstruction+τ labeled + τIdentification)/ Nτ(Reconstruction+τ labeled)

Rejection = Nτ(Reconstruction+τ nonlabeled)/ Nτ(Reconstruction+τ nonlabeled + τIdentification)





Z–>ττ–>jet jet sample

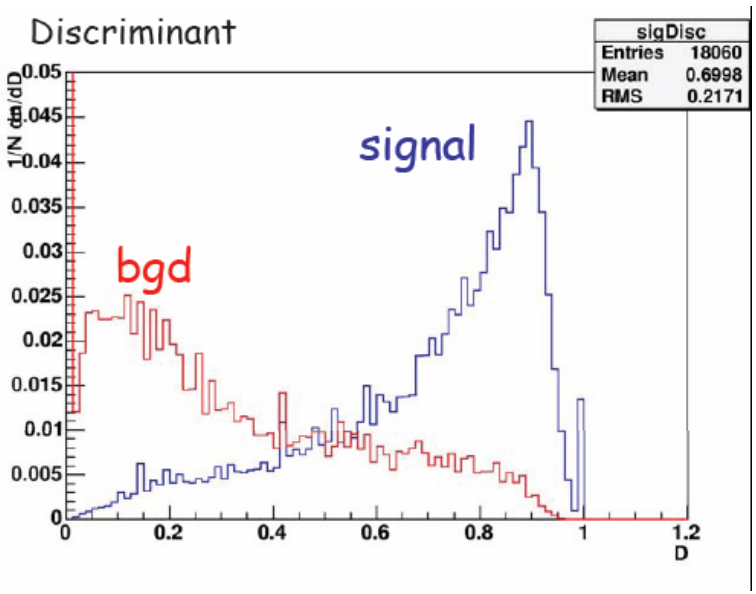
Atlas Tau1P3P: multivariate analysis

- Signal and background samples are used to combine the observables to a single one, called “discriminant”
- It is possible to cut on discriminant to separate signal from background
- Better background rejection is obtained with the same signal efficiency

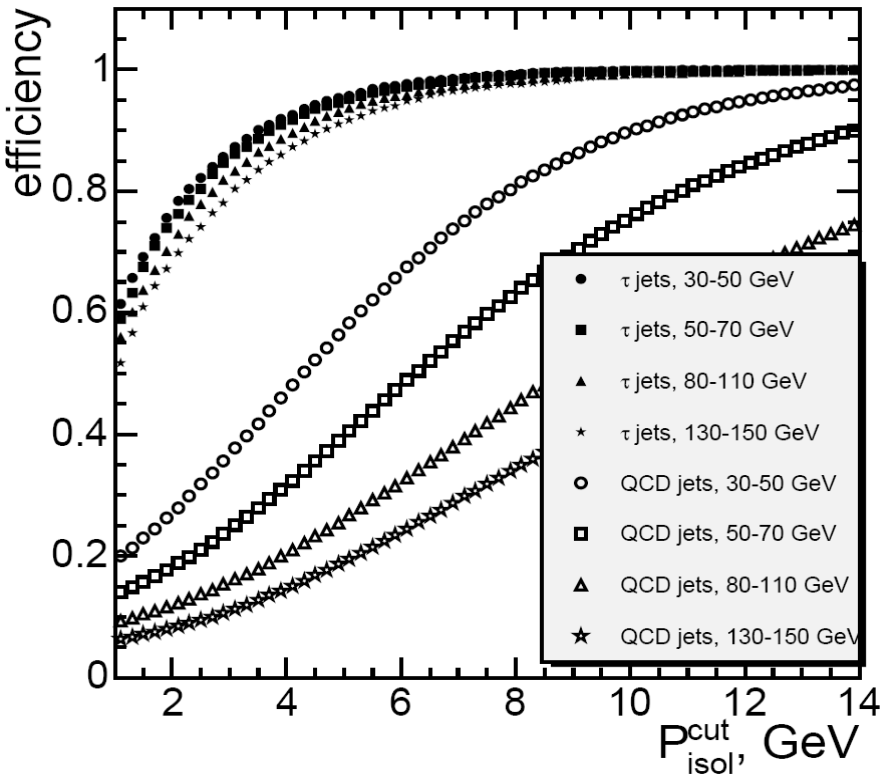
Tau1P results

	Cut analysis		Multvariant Analysis	
	sig.	bg	sig.	bg.
Full sample	58.9%	14.3%	58.9%	9.3%
Only gluon	58.9%	10.0%	58.9%	4.8%
Only quark	58.9%	21.5%	58.9%	11.8%

sign : $Z \rightarrow \tau\tau$
bg : QCD Jets



CMS: Ecal isolation

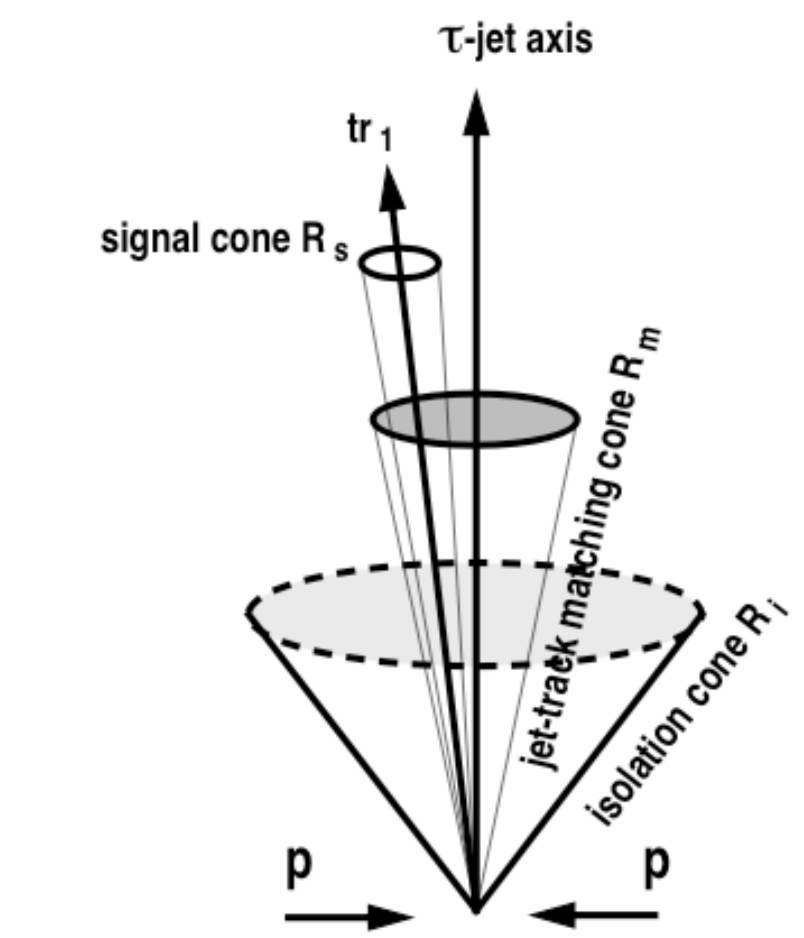


$$P_{isol} = \sum_{\Delta R<0.40} E_T - \sum_{\Delta R<0.13} E_T$$

CMS get a signal efficiency of about 80% with a bkg rejection of 5 for QCD jets with pT>80 GeV/c

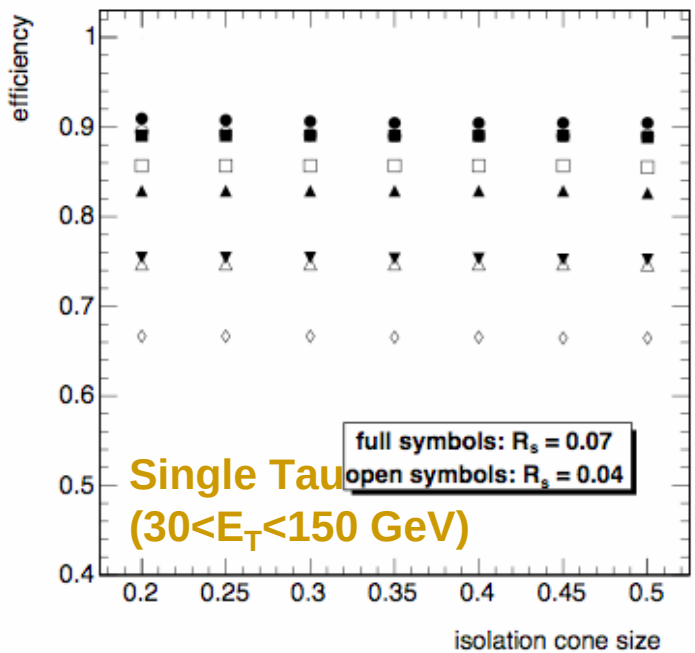
The efficiency of the electromagnetic isolation for τ jets and QCD jets in the different bins of the true transverse energy when the value of P^{cut}_{isol} is varied.

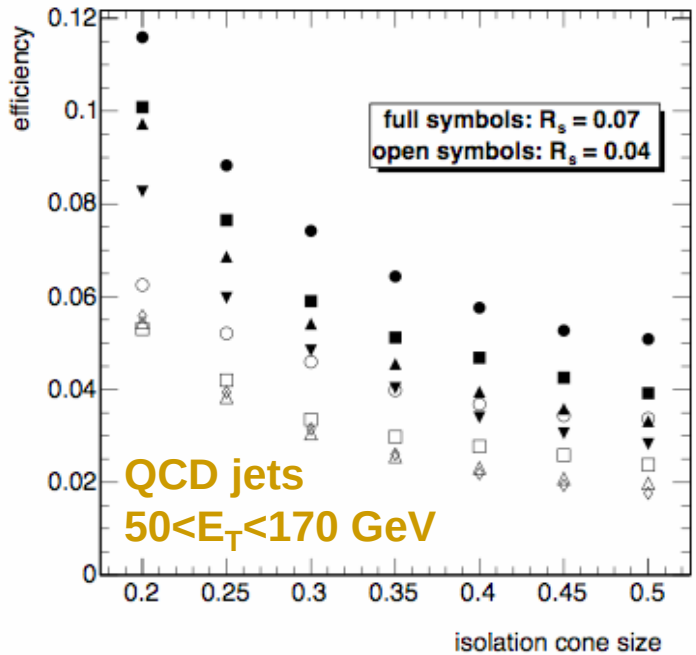
CMS: Tracker isolation



- **Isolation based on the number of tracks inside the isolation cone (R_I) is applied.**
- **Only good tracks are considered:**
  - Associated to the Primary Vertex
  - P_T of the Leading Track (i.e. highest p_T track) must exceed a few GeV/c
  - Leading Track must be found inside the Matching cone
- R_M : calo jet - leading track matching cone
- p_T^{LT} : cut on pT of the leading track in matching cone
- R_S : signal cone around leading track
- R_I : isolation cone (around jet axis or leading track)
- p_T^i : cut on pT of tracks in the isolation cone
- Dz : cut on the distance between z ip of the leading track and z ip of other tracks considered by algorithm (association with pxl primary vertex at HLT; see later)

CMS Tracker Isolation: tau jets and QCD jets efficiency





In the order of decreasing efficiency
symbols correspond to decreasing MC E_T intervals

Cuts used:

8 hits per track, Norm. $Chi^2 < 10$

$P_t^{LT} > 6$ GeV/c, $R_M = 0.1$, $R_I = 0.2\text{-}0.5$, $P_T^I > 1$ GeV, $|Dz| < 2$mm

CMS: other tagging methods

Tracker isolation is a primary requirement for the tau jet identification

All the following tagging methods are applied to jets which preliminarly pass the Tracker Isolation:

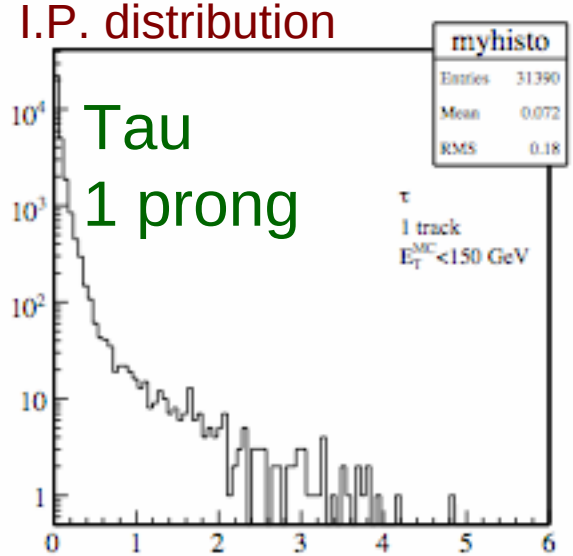
- **Tagging with IP**
- **Tagging with decay length**
- **Mass tag**

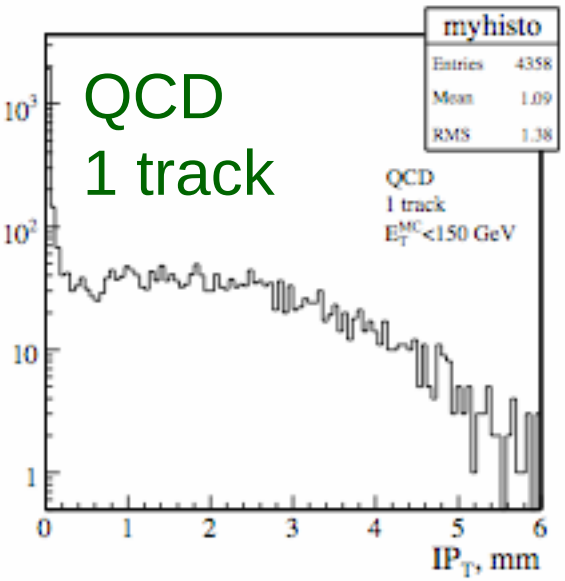
Default isolation parameters

$R_S=0.07$
$R_I=0.4$
$R_M=0.1$
$P_T^{LT}>10$ GeV/c
$P_T^{I}>1$ GeV/c
$|Dz| <2$mm
1 || 3 Tracks inside the signal cone

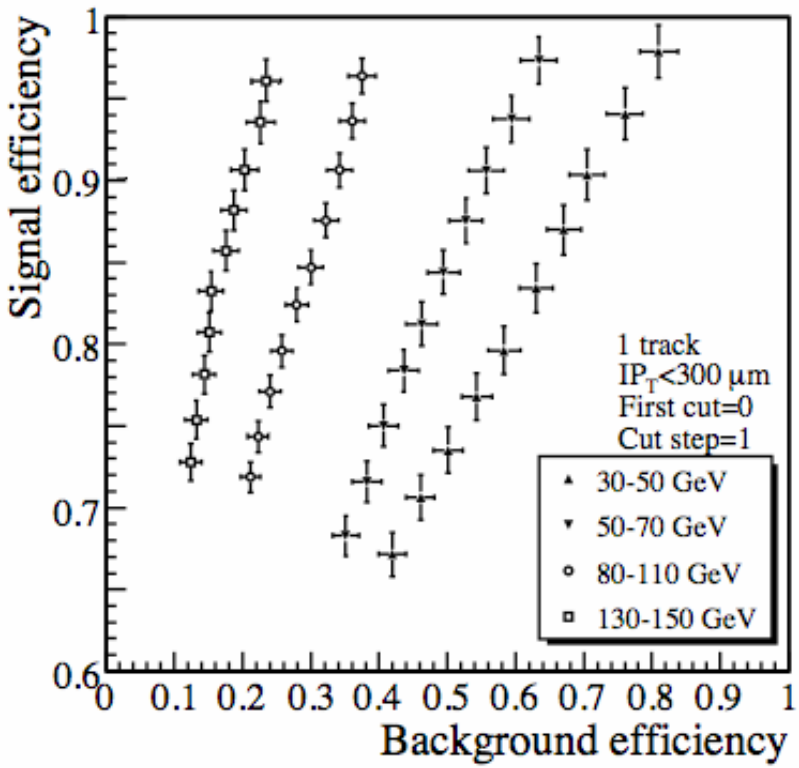
CMS Tagging with IP: distribution and performance

I.P. distribution





Transverse I.P. efficiency

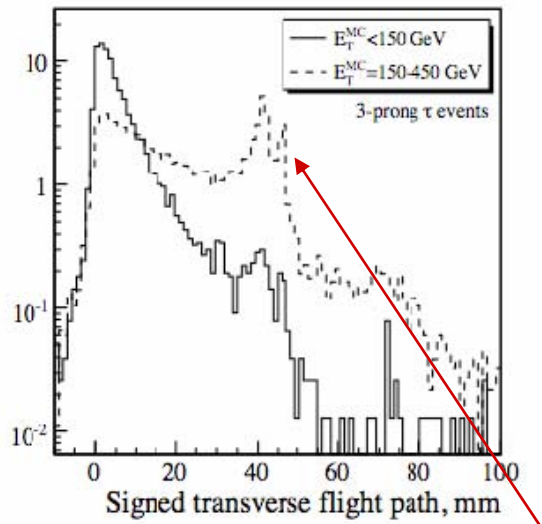


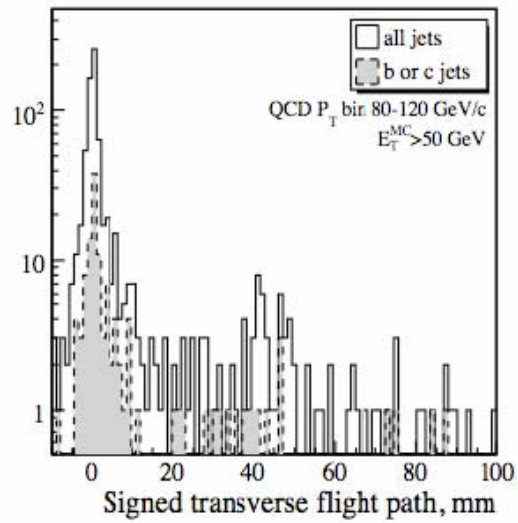
leading track selection:

- **N_{pixels_hits}>= 2**
- **$N_{tracker_hits}$>=8**
- **Chi2<10**
- **P_T>1 GeV**

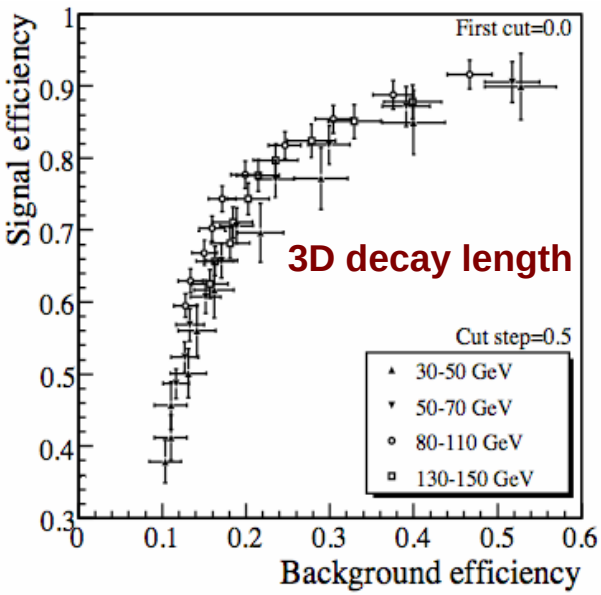
CMS: tagging with decay length

- The lifetime of the tau lepton ($c\tau$ = 87 μm) allows for the reconstruction of the secondary vertex for the 3 (and 5) prongs decay
- Events are required to pass tracker isolation, only tracks inside signal cone are used in the kalman vertex fitter





Fake vertices in the pixel material removed by cutting on the transverse flight distance (< 4 cm)

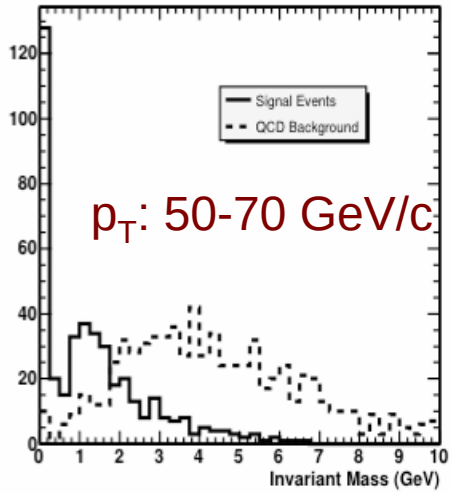


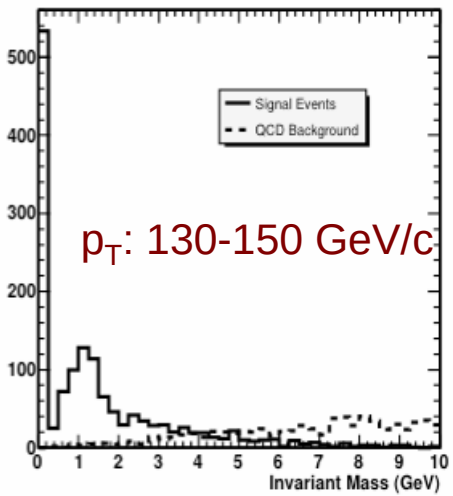
- **Both "Transverse" decay length and 3D decay length considered**

CMS: tagging with invariant mass

- Mass reconstruction is performed using tracks and Ecal clusters

•Only clusters within a cone of 0.4 from the jet axis have been considered.
•Remove clusters which have a track within a cone of 0.08 at the track impact point: avoid double counting



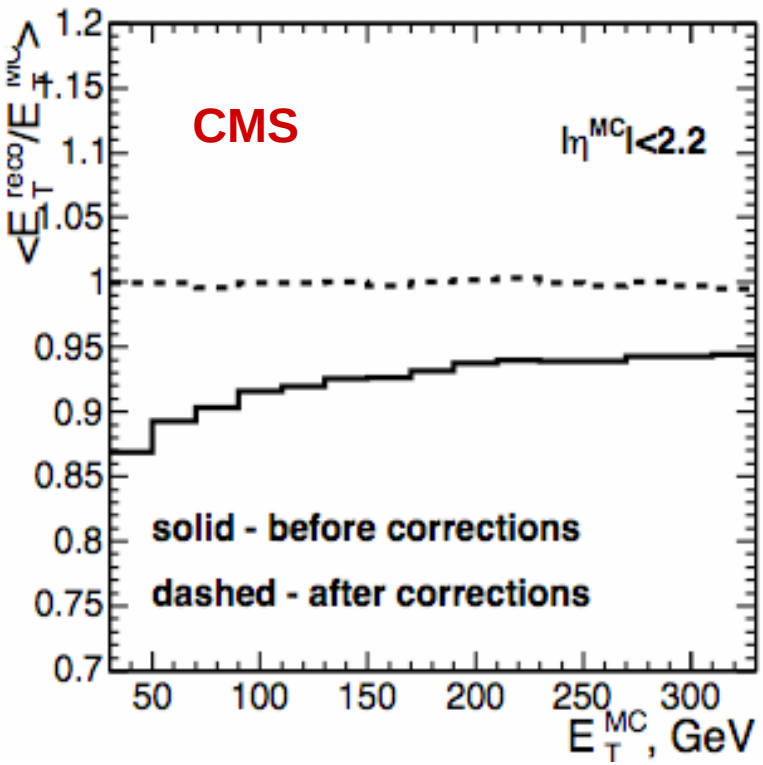


Efficiency requiring M_{TAU}<2.5 GeV/c^2

Pt bins (GeV)	30-50	50-70	80-110	130-150
Signal eff.(%)	86.32	82.27	83.02	80.76
Background eff.(%)	33.67	19.16	6.05	2.47

Tau Jet energy scale

- Both Atlas and CMS apply energy scale corrections
- Tau jets need softer corrections to their energy, wrt QCD jets.
  - At the same transverse energy, pions in Tau jets have harder transverse momentum than pions in QCD jets
  - In Tau jets there is a larger amount of electromagnetic energy (due to the presence of π^0)



The jets corrections optimised for true hadronic taus significantly underestimate energy scale for QCD jets

tau Trigger outline

- Trigger algorithms are developed taking into account some benchmark channels:
  - bbA/H–>ττ–> 2jets
  - gg->tbH$^+$, gb->tH$^+$
    - H$^+$->τν–>jet + MET
- The trigger efficiency is evaluated as a function of the background rejection capability
- The basic ingredients are the same than off-line tau identification
  - <u>For High Level Trigger algorithms speed is a basic requirement</u>

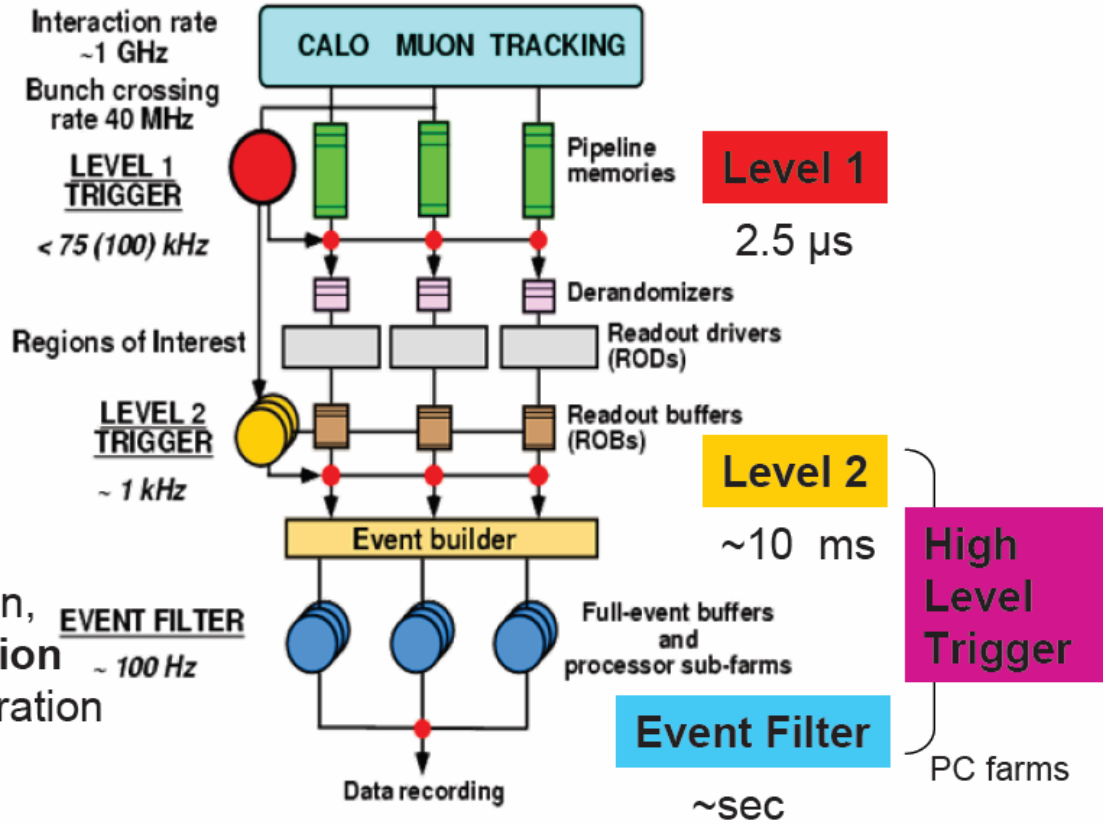
Atlas: general trigger scheme

- Trigger in ATLAS

LVL1 decison made with **Calorimeter+Muon Trigger Chambers,** coarse granularity

LVL2 uses **Region of Interest** **All detectors**, full granularity

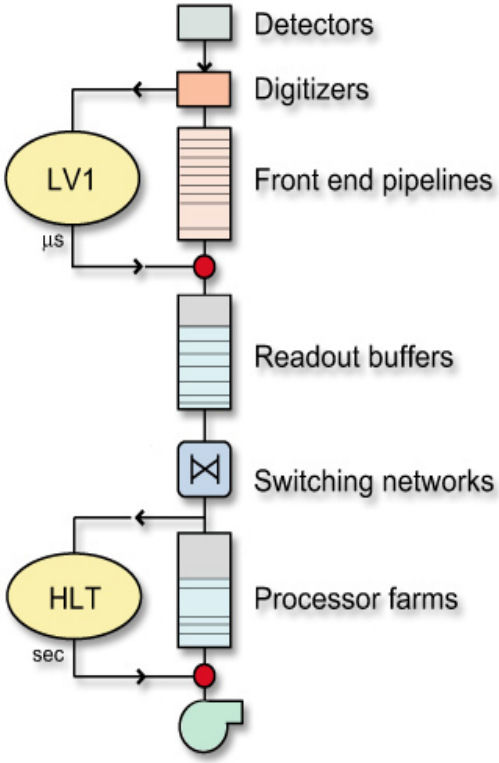
Event Filter refines the selection, can perform **event reconstruction** using latest alignment and calibration data



Atlas: tau trigger scheme

- Lepton trigger
  - Trigger with the electron or the muon
- Hadronic tau trigger
  - LVL1 Tau Trigger (Calo)
    - R.O.I. defined with EM and hadronic towers
    - Isolation in ECAL and HCAL
  - LVL2 Tau Trigger (Calo + Tracking)
    - Evaluate offline variables:
      - Cluster EM radius, width of energy deposition, isolation fraction, tracks
  - Event Filter
    - Based on TauRec code
- Good progresses in last year
  - not yet fully finalized

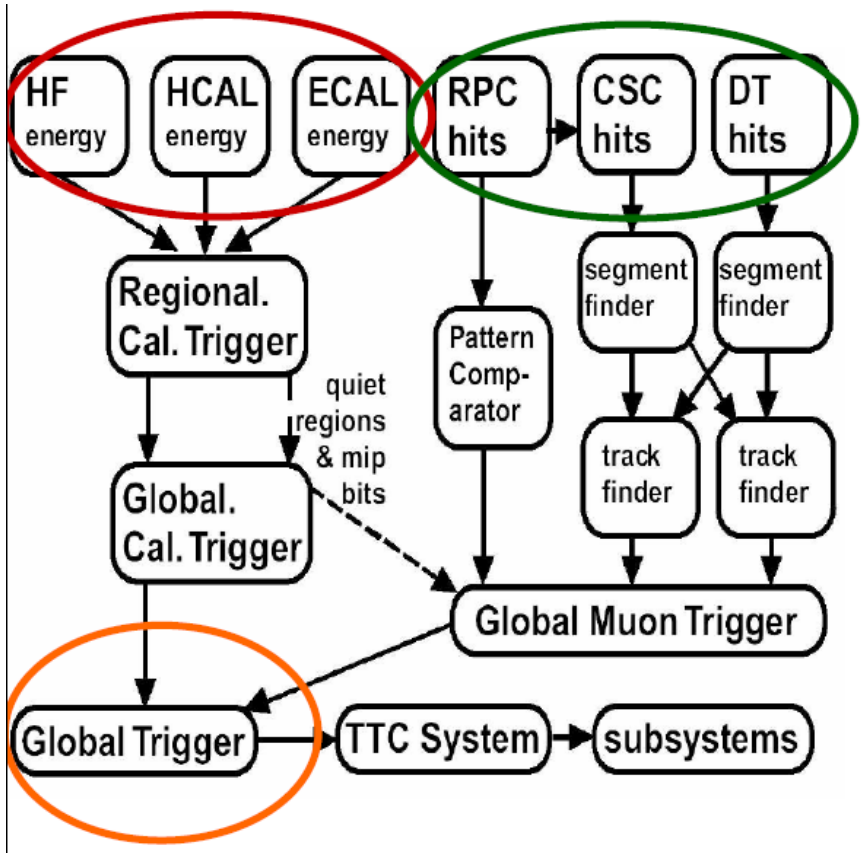
CMS: general trigger scheme



40 MHz
Clock driven
Custom processors

100 kHz
Event driven
PC network
Totally software

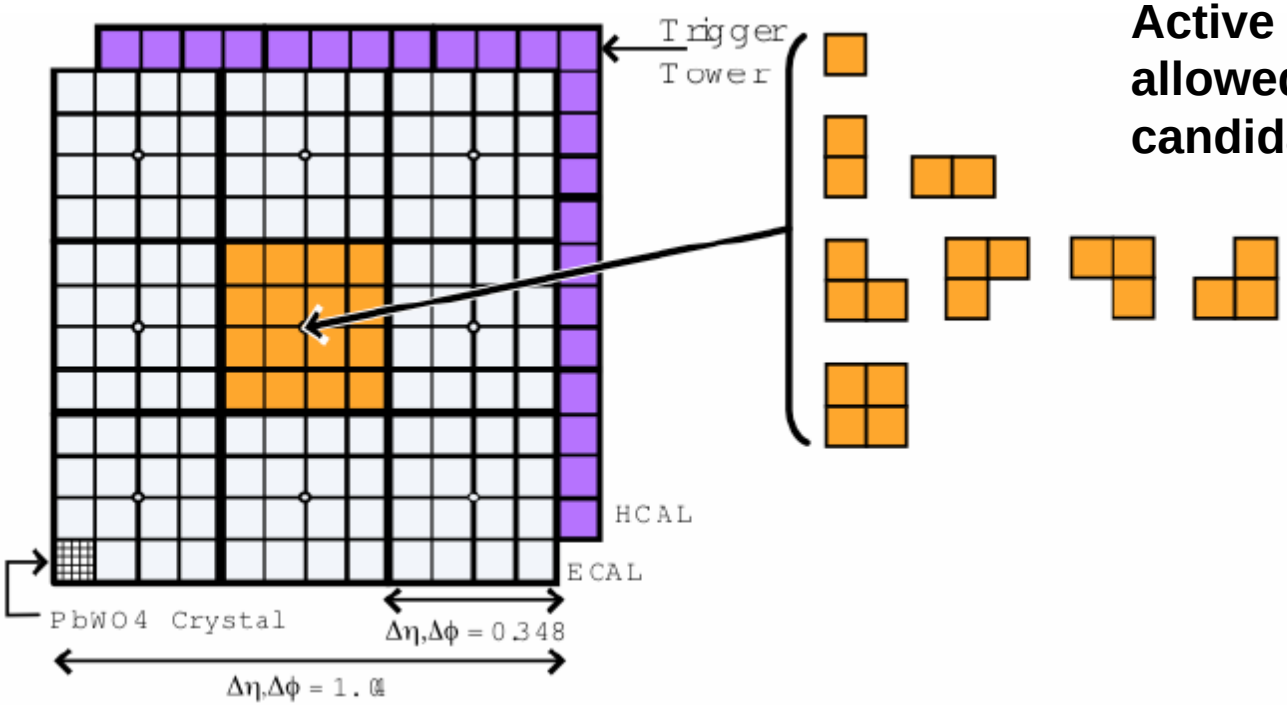
100 Hz
To mass storage



two trigger levels

Level-1 (~µs) 40 MHz
High-Level (ms-sec) 100 kHz
Event Size ~ 10^6 Bytes

CMS: L1 trigger for tau jets



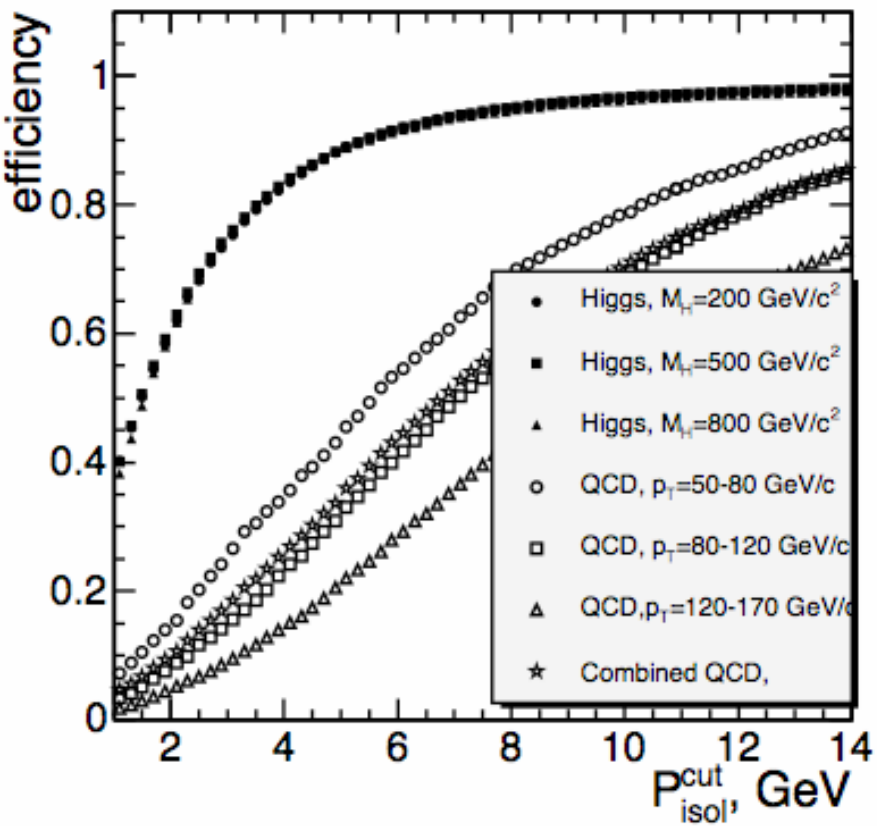
Active towers patterns allowed for tau jets candidates

- QCD E_T samples in the range 50-170 GeV have been used for the HLT studies. They represent more than the 90% of the total L1 Rate
- A factor ~10^3 of QCD background rejection is required at HLT
  - Reduce rate from ~kHz -> ~Hz

CMS: Ecal isolation at HLT

Calorimeter isolation is applied to the 1st (i.e. most energetic) HLT Calo jet.

$$P_{isol} = \sum_{\Delta R<0.40} E_T - \sum_{\Delta R<0.13} E_T$$



Efficiency has been evaluated for bbH->bb+tautau sample.

The main rate is represented by QCD di-jet events (p_T^{hat}:50-170 GeV/c), which is our bkg reference sample.

Rejection factor 3 is provided with P_{isol} < 5.0 GeV

It is used in association with the Pixel isolation (see following slides)

CMS HLT: Calo+Pxl or Trk Trigger?

CMS has developed two similar HLT trigger algorithms:

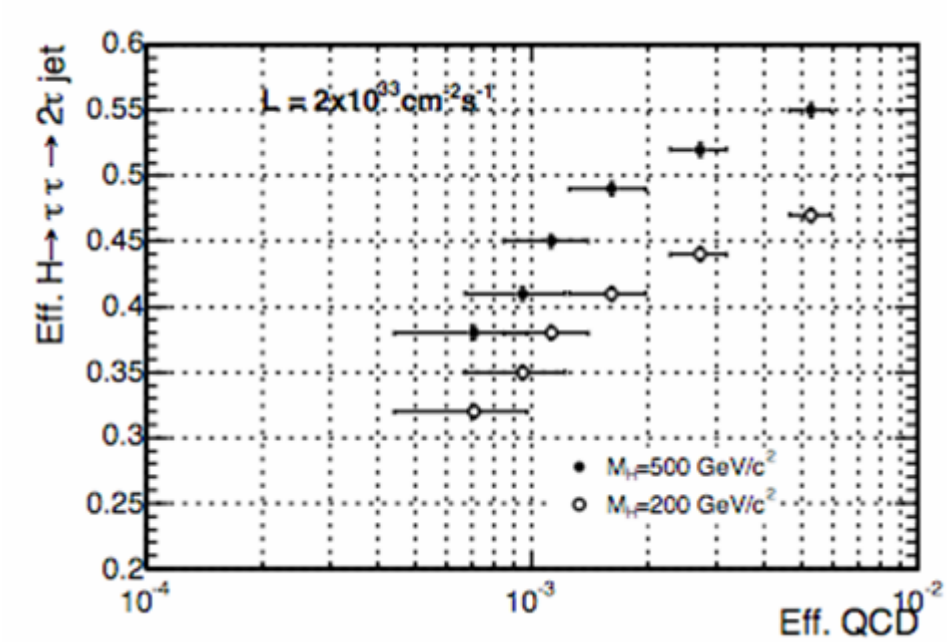
"Calo+Pxl trigger " and "Trk trigger"

- "Calo+Pxl trigger": only the pixels hits and the Level2 calo isolation used
  - **fast**
  - **good performance for isolation**
  - **preferred for decays with two taus in the final state (like A/H->tautau)**

- "Trk trigger": (some) hits of the microstrip inner tracker used, no Level2 calo isolation
  - **slower than Calo+Pxl**
  - **much better resolution for track momenta**
  - **useful in channels like charged Higgs boson decay**
    - **tight cut on the pT of the leading track**

CMS: HLT with Tracker isolation

- HLT tracker isolation is ~same than offline isolation
- Primary Vertex is made using only pxl
- Pixels + layers of silicon tracker is used to reconstruct tracks
  - Regional seeding with 2 out of 3 pxl hits
- Isolation cone is around Leading Track
- Tracks are associated to the 1st PV in the list
  - PV are sorted by the Σp_T^2 of the associated tracks.
- P_T Lead Tk > 6 GeV/c



Selection efficiency vs QCD efficiency
Benchmark channel:

- **H–>ττ–>2τ–jets**
- **HLT selection is applied to both taus**

CMS: Charged Higgs HLT trigger

Channel considered

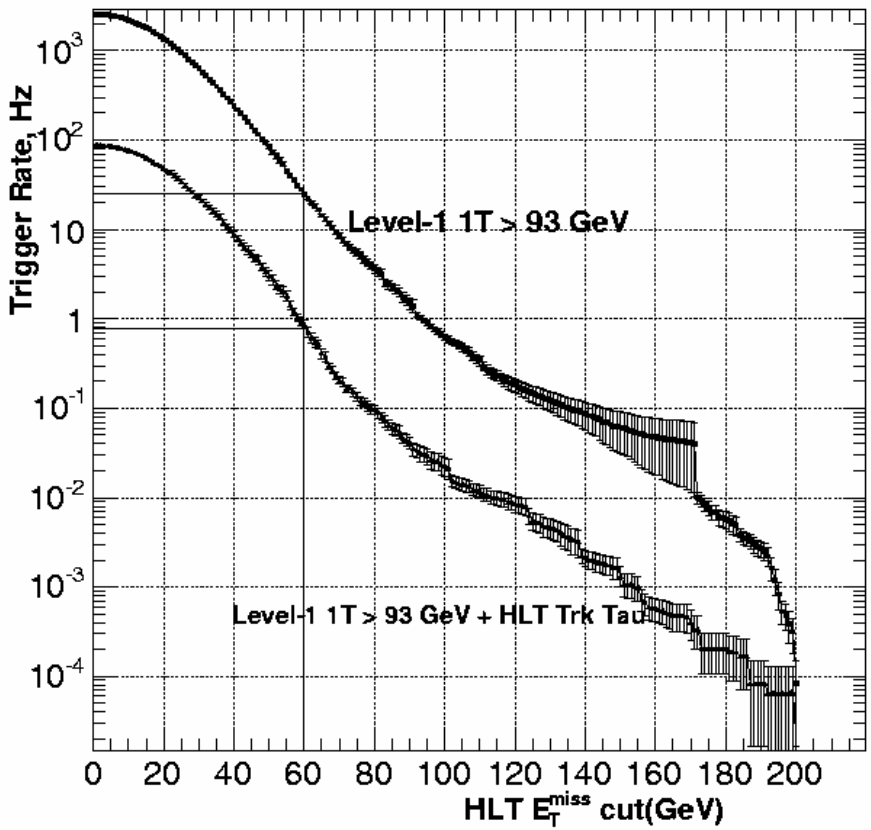
- **gg->tbH+, gb->tH$^+$**
  - **H$^+$ ->$\tau\nu$ (tau hadronic decay)**

- **L1 output rate: ~ 3kHz**

HLT selection

- **E_T^{miss}>65 GeV: output rate ~30 Hz**

- **After applying Tracker isolation + momentum cut (P_T^{LT}>20GeV):**
  - **output rate: ~ 1Hz**



CMS: tau tag efficiency from data

- The tag efficiency can be estimated from the ratio: $\frac{Z \rightarrow \tau\tau \rightarrow \mu + \tau_{\rm jet} + X}{Z \rightarrow \mu\mu}$

$$N_{Z\rightarrow\tau\tau\rightarrow\mu+\tau_{jet}} = \sigma_{DY} \times L \times BR(Z \rightarrow \tau\tau \rightarrow \mu + \tau_{jet}) \times \epsilon_{HLT_{1\mu}} \times \epsilon_{\tau\ tag} \times \epsilon_{other} = N^{meas}_{\mu+\tau_{jet}} - N^{bkg.}_{\mu+\tau_{jet}}$$

$$N_{Z\rightarrow\mu\mu} = \sigma_{DY} \times L \times BR(Z \rightarrow \mu\mu) \times \epsilon_{HLT_{1\mu}} \times \epsilon_{mass\ reco} = N^{meas}_{\mu+\mu} - N^{bkg.}_{\mu+\mu}$$

$\epsilon_{\tau\ tag}$ - *τ isolation in cone, requirement for 1 or 3 tracks in the signal cone*

ϵ_{other} - *selections rejecting background to the $Z \rightarrow \tau\tau \rightarrow \mu + \tau_{jet}$: $t\bar{t}$ and $W + jet$ and mass window selection*

$\epsilon_{mass\ reco}$ - *efficiency to find the second muon and mass window selection efficiency*

$$\epsilon_{\tau\ tag} = \frac{N^{meas}_{\mu+\tau_{jet}} - N^{bkg.}_{\mu+\tau_{jet}}}{N^{meas}_{\mu+\mu} - N^{bkg.}_{\mu+\mu}} \times \frac{BR(Z\rightarrow\mu\mu)}{BR(Z\rightarrow\tau\tau\rightarrow\mu+\tau_{jet})} \times \frac{\epsilon_{HLT^{\mu\mu}_{1\mu}}}{\epsilon_{HLT^{\mu\tau_{jet}}_{1\mu}}} \times \frac{\epsilon^{\mu\mu}_{mass\ reco}}{\epsilon^{\mu\tau_{jet}}_{other}}$$

Things in black - known from experiment (introduces statistical error)

Things in red - to be obtained from Monte Carlo and data (introduces systematical error)

CMS: tau tag efficiency from data

Selection table for 30 fb^{-1}

Signal

Background

Process:	$Z \to \mu\mu$	$Z \to \mu + \tau_{jet}$	$t\bar{t}$	$W + jet$	bb
$\sigma \times BR$ [pb]	2127	468	830	41457	22904478
Events for 30 fb^{-1}	63810000	14038200	24900000	1243720312	687134340000
HLT (single μ, $p_T > 19\ GeV/c$)	12113522	532024	1714285	12825148	1658017
only one, μ $p_T > 20\ GeV/c$	not applied	449087	1093926	11256357	1044637
τ_{jet} $E_T > 45\ GeV$	not applied	82211	534301	2174934	51228
leading τ_{jet} $tk.$ $p_T > 40\ GeV/c$	not applied	17030	204978	466396	4609
τ tag	**not applied**	**12379**	**137657**	**190844**	**2499**
$m_T(\mu, MET) < 30\ GeV/c^2$	not applied	8588	21407	32485	1464
$\Delta\varphi(\mu, \tau_{jet}) < 175°$	not applied	6820	20107	25663	1266
$E_{\nu_1,\nu_2} > 0$	not applied	3665	8531	5633	694
Electron veto	not applied	3032	2664	3839	504
Jet veto (1 central jet)	not applied	1666	682	1460	116
Mass window	**7191968 ± 28658**	**988 ± 99**	**63±22**	**42 ± 30**	**61 ± 34**
Mass window (wide)	**7191968 ± 28658**	**1291 ± 113**	**206±40**	**271 ± 75**	**75 ± 37**

- Error contribution on ε_{tag} for 30 fb^{-1}

Source:	Number of $\mu\tau_{jet}$ events	Calorimtery scale uncert.	background uncertainty	**Total**
Contribution [%]:	3.4	8	1	**8.7**

Atlas vs CMS

ATLAS

- TauRec is the standard tool: flexible and powerful
  - Possibility to use different seed
    - Clusters, tracks…
  - Likelihood approach
- Tau1P3P gives good performance for low momentum taus
  - Multivariate analysis helps to improve the performance
  - Good energy flow algorithm
    - Very important for low energy tau tagging
- Triggering path is established…
  - But not yet finalized

CMS

- Triggering path well understood and established
  - HLT algorithms based on isolation work very well
- Off-line tau tagging based on isolation + additional methods
  - Good results
  - Likelihood/multivariate approach would improve overall performance
- A first method to evaluate tag efficiency from data has been studied
- General energy flow is being developed in CMS…
  - But not yet available
  - Would improve things for low energy taus and mass tag

Conclusions

- Tau jet identification at LHC is crucial for many discovery channel
- Off-line tau tagging methods are well advanced in Atlas and CMS
- Trigger to tau-jet starting from Level 1 is possible
  - needed for several discovery channels
- LHC data taking is (finally!) approaching
  - we are ready to switch from simulation to real data!