Tau ID in CMS and new developments

Pavel Jež

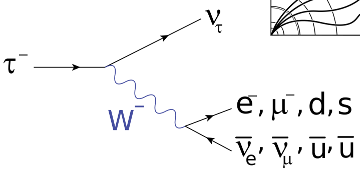
Centre for Cosmology, Particle Physics and Phenomenology - CP3
Université catholique de Louvain

January 24, 2014

S3Be, IIHE Brussels

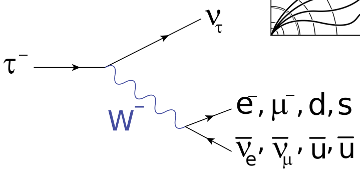
Tau Lepton

- Heaviest charged lepton (1.8 GeV)
- Decays to leptons and hadrons
- Charge conservation⇒ odd number of tracks
  - "1-prong" and "3-prong"



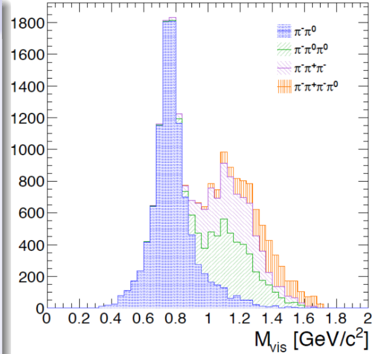
Tau Lepton

- Heaviest charged lepton (1.8 GeV)
- Decays to leptons and hadrons
- Charge conservation⇒ odd number of tracks
  - "1-prong" and "3-prong"



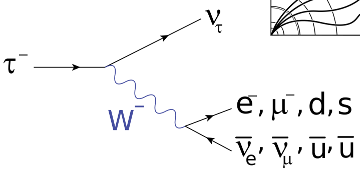
decays of τ^-

Decay Mode	Resonance	BR[%]
$e^-\bar{\nu}_e\nu_\tau$		17.8
$\mu^-\bar{\nu}_\mu\nu_\tau$		17.4
$h^-\nu_\tau$		11.6
$h^-\pi^0\nu_\tau$	$\rho(770)$	25.9
$h^-\pi^0\pi^0\nu_\tau$	$a_1(1260)$	9.5
$h^-h^-h^+\nu_\tau$	$a_1(1260)$	9.8
$h^-h^-h^+\pi^0\nu_\tau$		9.8
other hadronic		3.2



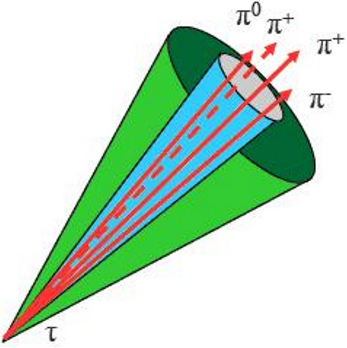
Tau Lepton

- Heaviest charged lepton (1.8 GeV)
- Decays to leptons and hadrons
- Charge conservation⇒ odd number of tracks
  - "1-prong" and "3-prong"



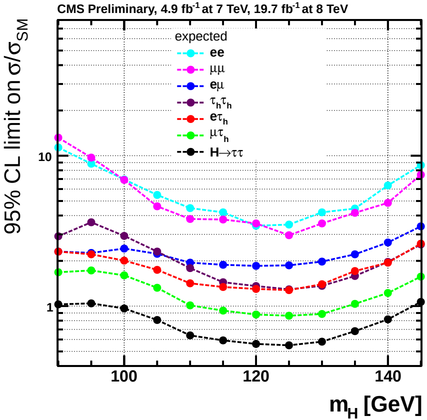
decays of τ^-

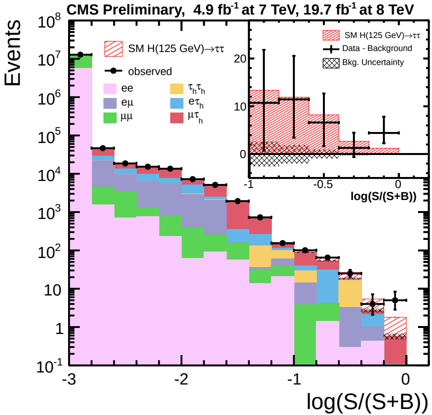
Decay Mode	Resonance	BR[%]
$e^-\bar{\nu}_e\nu_\tau$		17.8
$\mu^-\bar{\nu}_\mu\nu_\tau$		17.4
$h^-\nu_\tau$		11.6
$h^-\pi^0\nu_\tau$	$\rho(770)$	25.9
$h^-\pi^0\pi^0\nu_\tau$	$a_1(1260)$	9.5
$h^-h^-h^+\nu_\tau$	$a_1(1260)$	9.8
$h^-h^-h^+\pi^0\nu_\tau$		9.8
other hadronic		3.2



Tau ID crucial for the evidence of $H \to ff$

CMS-HIG-13-004-005, arXiv:1401.5041, submitted to JHEP





The 3 most sensitive channels contain at least one τ_h

CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

Reconstruction and identification of hadronic taus

CMS: Hadron+strip (HPS) algorithm

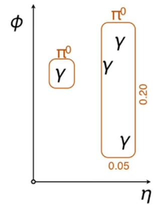
- **input:** Particle flow $\Rightarrow$ list of particles $\Rightarrow$ jets

1. Vertex associated to leading track $\Rightarrow$ jet vertex

CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

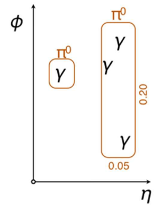
1. Vertex associated to leading track ⇒ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet

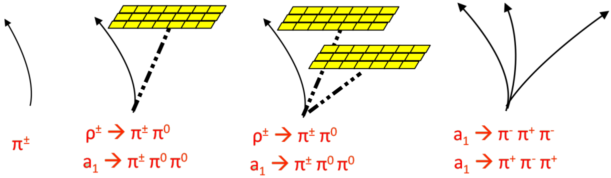


CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

1. Vertex associated to leading track ⇒ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet
3. Combine 1 or 3 charged hadrons with 0-2 π^0 ⇒ τ candidates



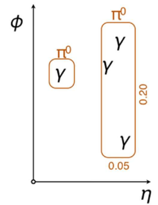


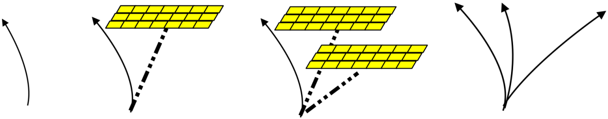
Reconstruction and identification of hadronic taus

CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

1. Vertex associated to leading track ⇒ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet
3. Combine 1 or 3 charged hadrons with 0-2 π^0 ⇒ τ candidates
4. Reject combinations not compatible with τ decay



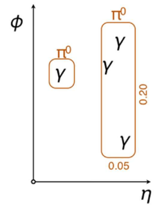


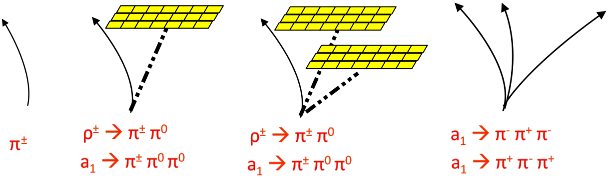
Reconstruction and identification of hadronic taus

CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow $\Rightarrow$ list of particles $\Rightarrow$ jets

1. Vertex associated to leading track $\Rightarrow$ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet
3. Combine 1 or 3 charged hadrons with 0-2 π^0 $\Rightarrow$ τ candidates
4. Reject combinations not compatible with τ decay
5. Require that τ constitutents are very collimated



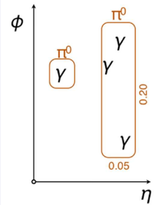


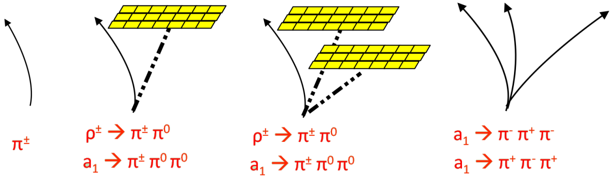
Reconstruction and identification of hadronic taus

CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

1. Vertex associated to leading track ⇒ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet
3. Combine 1 or 3 charged hadrons with 0-2 π^0 ⇒ τ candidates
4. Reject combinations not compatible with τ decay
5. Require that τ constitutents are very collimated
6. Max 1 τ per jet ⇒ keep the most isolated





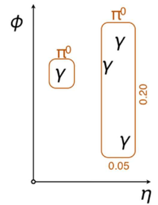
Reconstruction and identification of hadronic taus

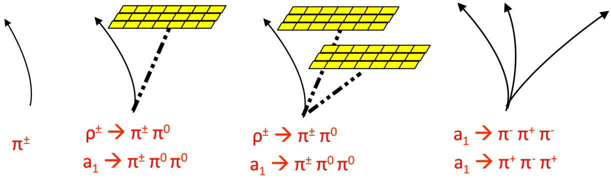
CMS: Hadron+strip (HPS) algorithm

- **input:** Particle flow ⇒ list of particles ⇒ jets

1. Vertex associated to leading track ⇒ jet vertex
2. Create π^0 candidates ("strips") from γ's and e's inside jet
3. Combine 1 or 3 charged hadrons with 0-2 π^0 ⇒ τ candidates
4. Reject combinations not compatible with τ decay
5. Require that τ constitutents are very collimated
6. Max 1 τ per jet ⇒ keep the most isolated

- **Jet rejection:** p_T of particles close to τ (cut on scalar sum or BDT)
- **Electron rejection**: Inverted electron ID + dedicated BDT
- **Muon rejection**: Inverted muon ID

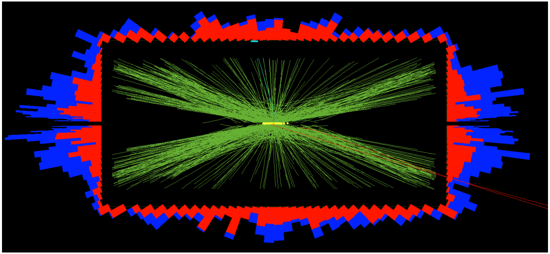




Pileup corrections

Pileup:

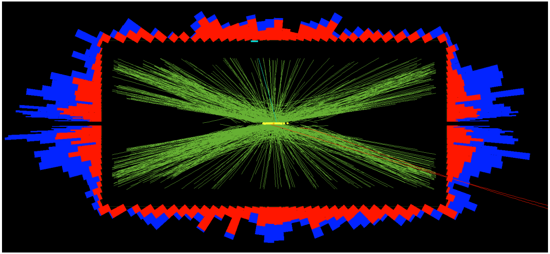
- $\sim$ 30 collisions at the same time $\Rightarrow$ additional particles not connected with primary interaction



Pileup corrections

Pileup:

- $\sim$ 30 collisions at the same time $\Rightarrow$ additional particles not connected with primary interaction



Localized ($\Delta\beta$) pileup corrections

- Using only charged particles compatible with jet/tau vertex $\Rightarrow$ charged pileup robust
- Estimate the neutral pileup contribution (sum of charged pileup around τ $\times$ correction factor $\Delta\beta$) and subtract

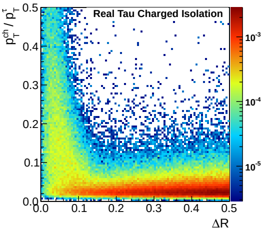
Discrimination by isolation

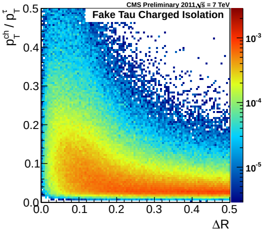
Combined isolation with pile-up correction

- Absolute sum of charged and neutral PF isolation in $\Delta R < 0.5$
- Correction for charged pile-up is "automatic" from construction
- Contributions from neutral pile-up are subtracted using scalar p_T sum of charged pile-up (all PF particles within $\Delta R < 0.8$ of τ **not** coming from τ vertex) and a correction factor

MVA isolation

- Spatial p_T distribution (τ shape) is used to select real τ's
- MultiVariate Analysis (Boosted Decision Trees) variables:
  - scalar p_T sum of charged and neutral PF candidates in 4 annuli and 1 cone around τ
  - p_T weighted mean and variance of $\Delta\eta$ and $\Delta\phi$ between isolation particle and τ





Electron and Muon rejection

Cut based electron rejection

- PF electron ID uses MVA discriminator
- Loose electron rejection inverts the MVA electron ID
- Tighter working points use additional cuts to reject single charged hadrons and high bremstrahnulg electrons

MVA based electron rejection

- Specific MVA dicriminator to distinguish τ's from electrons
- tau candidates are categorized using
  - number of neutral PF among τ constituents
  - pseudorapidity
  - presence of e reconstructed in the tracker and associated to leading τ track
- Dedicated BDT training in each category

Muon rejection

- **Loose:** leading τ track cannot be matched with a track segment in the muon chambers
- **Medium:** The leading τ track not reconstructed as loose muon
- **Tight:** Same as medium + cut on minimum relative p_T deposit for single prong and 0 strips candidates

Current recommendations of the Tau POG

"Legacy" recommendations

- The recipe used for HIG-13-004 paper and shown few moments ago
- Contained in CMSSW_5_3_12 and newer
  - **But:** it is necessary to **re-run PFTau sequence** on 8 TeV AODs
  - **Nevertheless:** contained in AODs from 7 TeV legacy reprocessing
- If you really need to use older recipes, contact Tau POG

"Tau ID 2014" recommendations

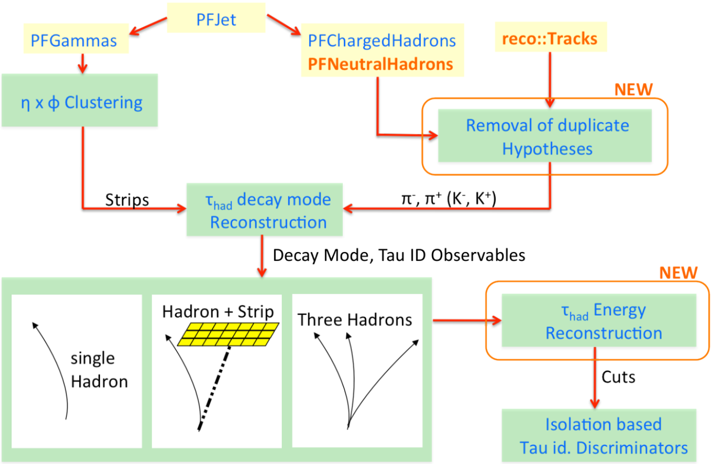
- The code is stabilized and running in 53x and 62x and 70x
- Contains
  - Fix for high p_T behaviour
  - Code for boosted tau reconstruction in subjects
  - Tau lifetime information
  - **Better performing discriminators**
- Dedicated branches in github obtained via `cms-merge-topic`
- Will be included in 7_0_0_pre13

Motivation for new ID

- In December 2012 a problem discovered with performance at very high p_T (1 TeV)
  - significant efficiency decrease
  - underestimation of tau momentum
- A taskforce was set-up and identified causes:
  1. Underestimation of track p_T (small curvature) $\Rightarrow$ fake neutral PF particles $\Rightarrow$ Decay mode finding fails
  2. Kinematical cuts too tight
  3. (Decreased track reconstruction efficiency at high p_T)[3]

Solution to Decay mode finding fails

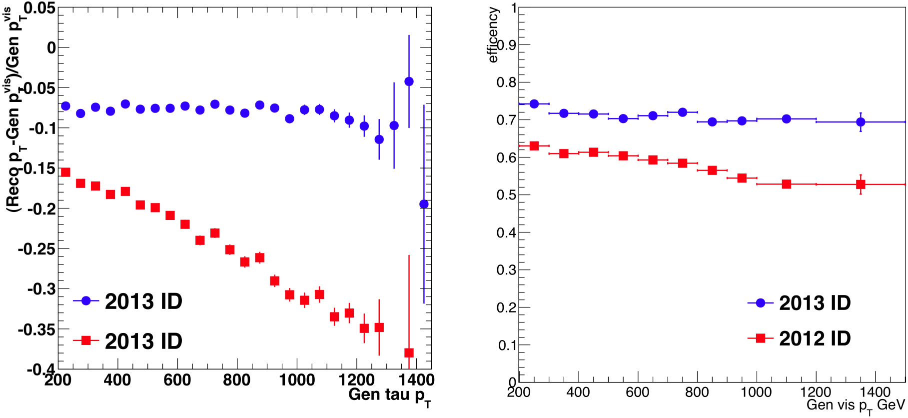
Adding an "afterburner" that reconstructs high p_T π's and K's from PFlow output + tracks



Solution to other issues: change of kinematical cuts

- Mass window cut was changed (Sliding mass window)
- New decay modes were defined
  - Recovering 3-prong decays by looking at **2 tracks + neutral particle**
  - About 10% efficiency gain at very high p_T
  - p_T dependent increase in jet$\rightarrow \tau_h$ fakes
  - Analyst can choose whether to use 2-prong decay modes or not

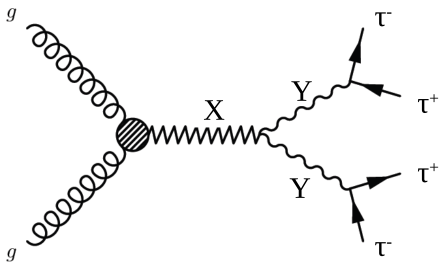
Results



- Stable tau energy response
- Efficiency increased due to new algorithm, new decay modes and relaxed decay mode cuts

Boosted tau reconstruction

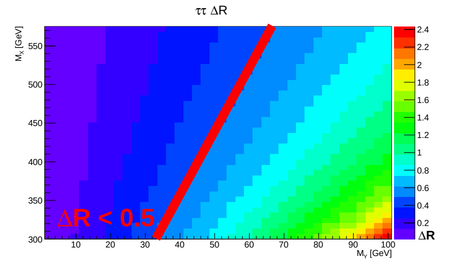
Motivation



- Taus difficult to reconstruct when $\Delta R < 0.5$
- $\Delta R = \frac{2M_Y}{p_T^Y}$, $p_T^Y \approx \frac{M_X}{2}$
- If $p_T^Y > 5M_Y$ or $M_X > 10M_Y$ $\Rightarrow$ difficult kinematic regimes with standard tau reco

Boosted tau reconstruction

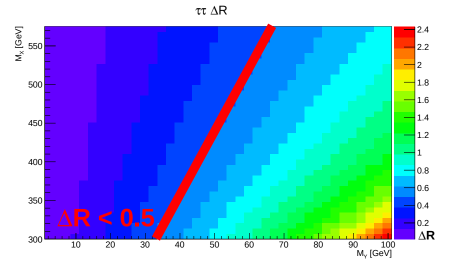
Motivation



- Taus difficult to reconstruct when $\Delta R < 0.5$
- $\Delta R = \frac{2M_Y}{p_T^Y}$, $p_T^Y \approx \frac{M_X}{2}$
- If $p_T^Y > 5M_Y$ or $M_X > 10M_Y$ $\Rightarrow$ difficult kinematic regimes with standard tau reco

Boosted tau reconstruction

Motivation



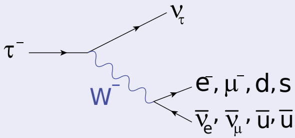
- Taus difficult to reconstruct when $\Delta R < 0.5$
- $\Delta R = \frac{2M_Y}{p_T^Y}$, $p_T^Y \approx \frac{M_X}{2}$
- If $p_T^Y > 5M_Y$ or $M_X > 10M_Y$ $\Rightarrow$ difficult kinematic regimes with standard tau reco

Solution

- Subjet techniques are used to get input for tau reconstruction
- Algorithm functional and running
- Everybody is welcome to use it

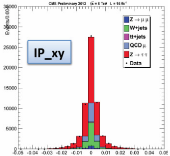
Tau lifetime information

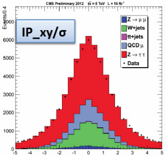
Motivation



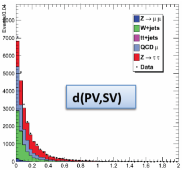
- Tau decays via weak interaction⇒ measurable length-of-flight
- 1 mm (20 GeV)/ 4.9 cm (1 TeV)
- additional handle to identify τ_h

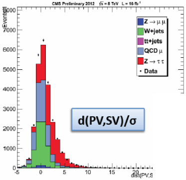
1-prong: impact parameter





3-prong: |PV-SV|

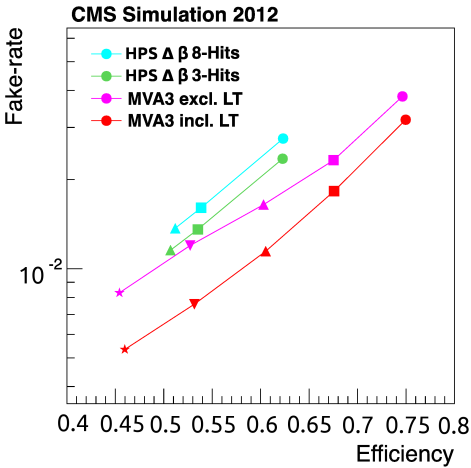




- Good Data/MC aggreement
- Improves tau ID performance

Evolution of the tau ID performance

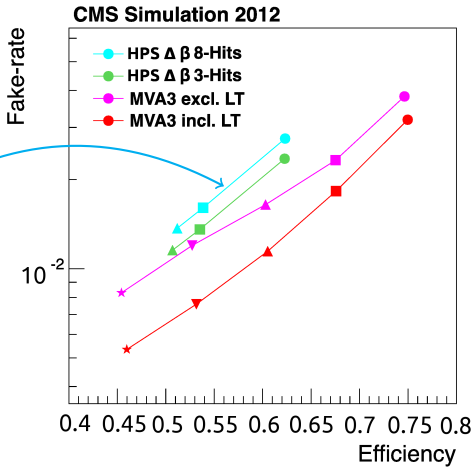
Performance is improving every year!



Evolution of the tau ID performance

Performance is improving every year!

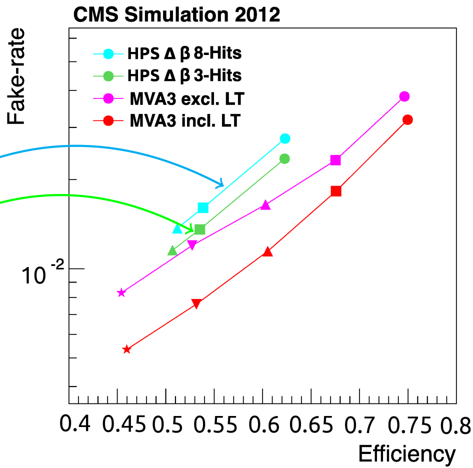
- 2011



Evolution of the tau ID performance

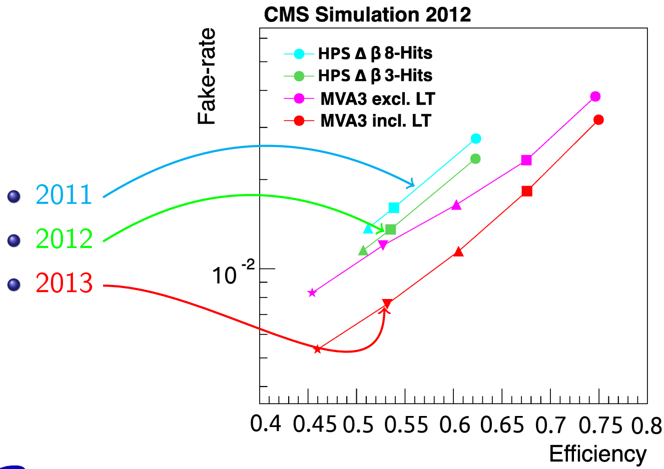
Performance is improving every year!

- 2011
- 2012



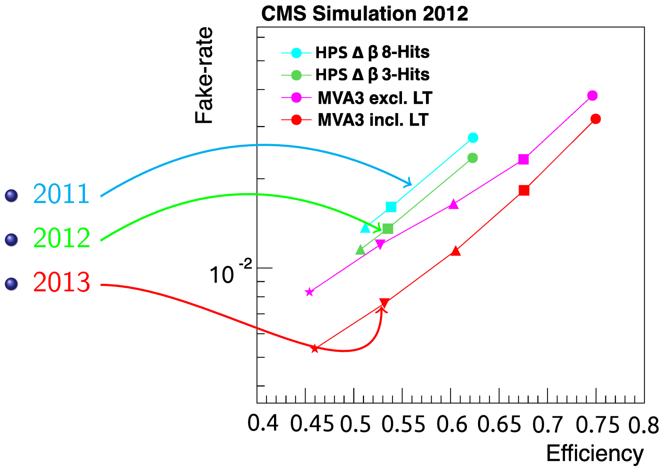
Evolution of the tau ID performance

Performance is improving every year!



Evolution of the tau ID performance

Performance is improving every year!



Addition of the lifetime information seems to lead to a large performance improvement

Plan for 2014

1. Move to the new tau ID
   - Everything seems to indicate that the new tau ID is vastly superior to the current one in all categories
   - The algorithm will be made part of official CMSSW releases and extensively validated in various analyses
2. Prepare for Run 2 (continuation of present studies)
   - Different running conditions $\Rightarrow$ re-tune $\Delta\beta$ corrections
   - Retuning of cuts and retraining MVA's
   - Trying to maintain our Run 1 performance or better

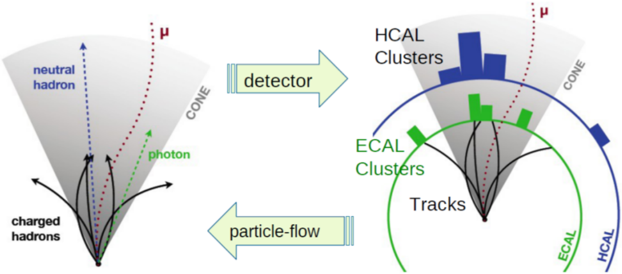
Summary

- Tau POG offers 2 flavours of tau ID
  1. "Legacy" recommendation
     - ★ standard, well tested, well performing
     - ★ in CMSSW > 5_3_12
  2. "2014 Tau ID" recommendation
     - ★ state-of-the-art, not that much tested yet, everybody is invited to try
     - ★ `git cms-merge-topic -u cms-tau-pog:CMSSW_5_3_X_boostedTaus`
- The recipes for installation are kept at Twiki
  - https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID
- The recommendation for analyses are also in Twiki
  - https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation
- Questions? Comments? Suggestions?
  - Write to hypernews hn-cms-tauid@cern.ch
  - Come to meet us: Mondays @ 16:00

Additional Material

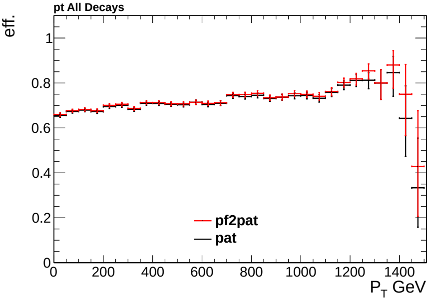
Particle Flow

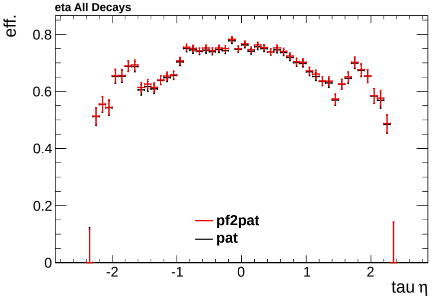
- Algorithm to reconstruct all stable particles in the event
  - e, μ, γ, charged and neutral hadrons
- allows parton-level-like analysis
- individual particles are used e.g. to
  - reconstruct jets and τ's
  - determine E_T^{miss}
  - quantify particle isolation
  - tag b-jets
- possible due to high precision tracker and EM calorimeter (allows to separate γ's and charged particles deposits in the calorimeter)



Status of PFTau→GED integration

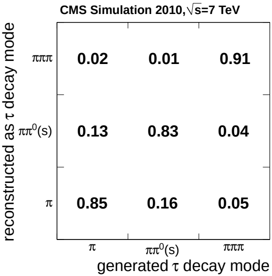
- The plan of the RECO/AT group is to eventually move to PFBRECO and GED
- In every official release, tau sequence is working in PFBRECO
- The PFBRECO is also working on the *_highPt/boostedTaus branches of cms-tau-pog
- PFBRECO was **successfully validated** against simple PFTau sequence in 53x
- Validation and clean-up in 62x/70x is ongoing

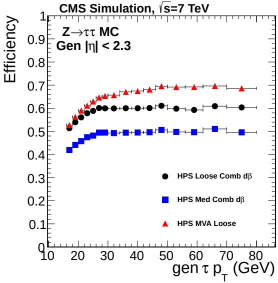


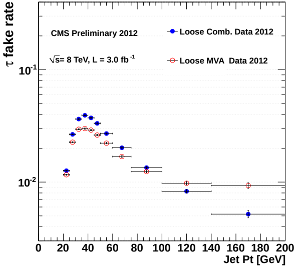


Performance of CMS Tau ID

Decay mode finding and isolation







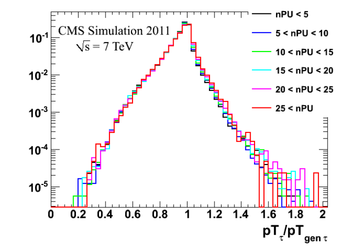
Efficiency and Fake rate definitions

- Tau selection: $p_T > 20$ GeV, $|\eta| < 2.3$
- In efficiency and FR plots, τ's are passing Decay mode selection, loose e and μ rejection
- For efficiency - reconstructed τ matched to generated τ ($\Delta R < 0.5$)
- Fake rate measured in jet data

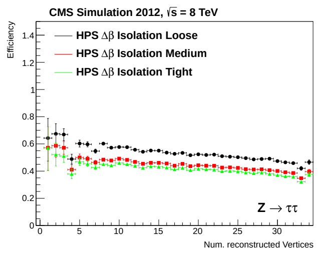
Performance of Tau ID

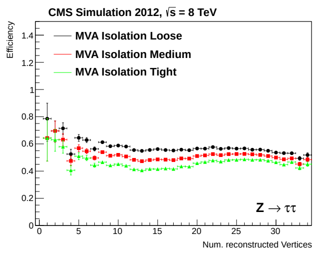
Stability with pile-up

p_T resulution for 1 prong + π^0 events



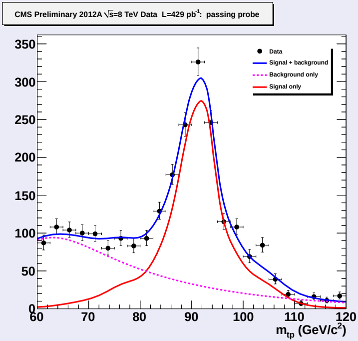
Isolation efficiency as a function of reconstructed vertices





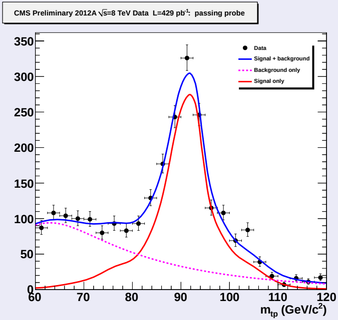
Light lepton rejection

Electron fake rate



- Electron fake rate measured in sample of isolated electrons from $Z \to ee$ using tag and probe
- Measured fake rate in barrel 0.5%/13% (MVA/loose) in barrel and 0.9%/29% in endcap
- Efficiency is 85%/100% (65%/100%) in barrel (endcap)

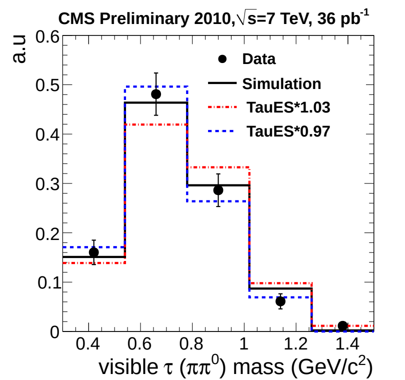
Muon fake rate



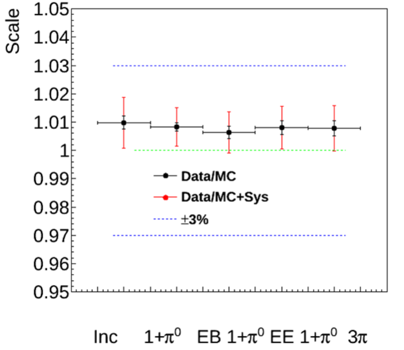
- Muon fake rate measured in sample of isolated electrons from $Z \to \mu\mu$ using tag and probe
- Upper limit 0.24% on $\mu \to \tau$ fake rate

Tau energy scale

2010



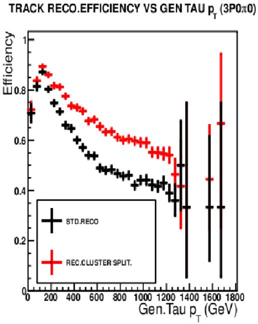
2011

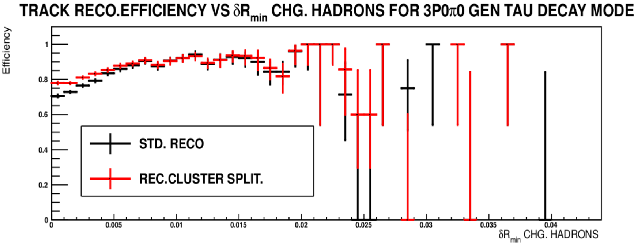


- Hadronic taus composed of charged hadrons and γ's that are reconstructed with high precision in PF
- Tau energy scaled is well described in data and our understanding is steadily improving

Study of high p_T tracking

- Tracks lying close to each other can have their cluster merged
- Affects tracking efficiency quality
- Tried to apply cluster splitting algorithm and check its effect on the tau performance
- 15% efficiency gain for 3-prong taus ¿ 800 GeV, 5% gain overall for $p_T > 800$ GeV





Study of the post-LS1 performance

- Already started to look on the tau performance after the LS1
- No proper MC samples available, so only very preliminary studies:
  - Study 1
  - Study 2
- Not much conclusions but it will be probably necessary to re-tune $\Delta\beta$ corrections
  - Present value is slightly overcorrecting $\Rightarrow$ safe for 25 ns
  - The tau performance is not affected much by changing the $\Delta\beta$