Searches for Higgs bosons with hadronically decaying τ-leptons

Using Grid and Cloud computing techniques

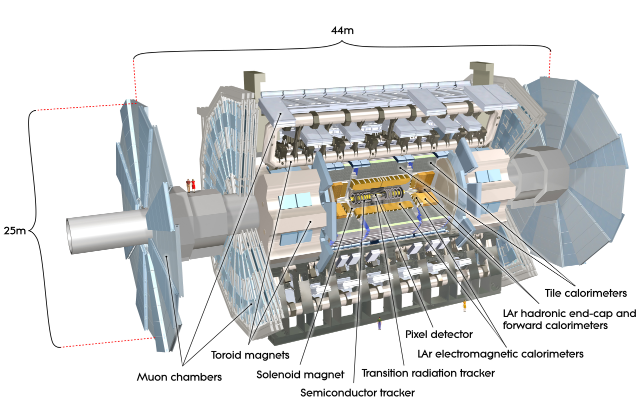
Henrik Öhman (Uppsala University)

Uppsala, 2016-11-18

Introduction

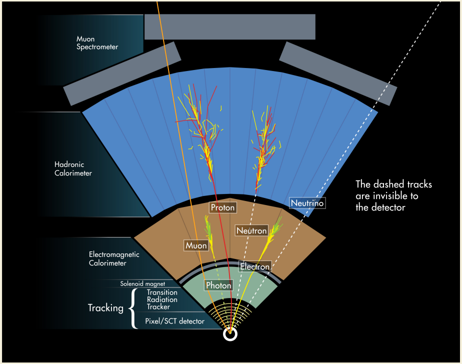
- The ATLAS detector
- Hadronically decaying τ-leptons
- $H \rightarrow \tau\tau$ measurement
- $H^+ \rightarrow \tau\nu$ search

The ATLAS detector



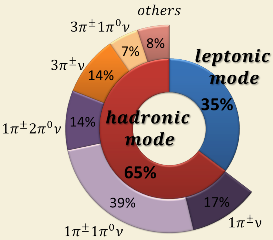
- Inner detector – tracking
- Calorimeters – trigger, particle ID, near full coverage
- Muon spectrometer – tracking, trigger
- Level 1 (L1) trigger – hardware, close to readout
- High-level trigger (HLT) – computer farm, $\sim 50\,000$ cores

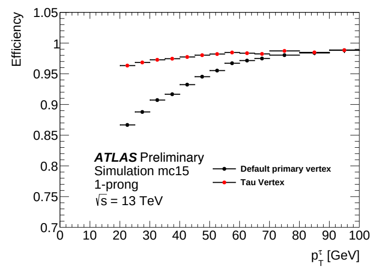
Particle identification

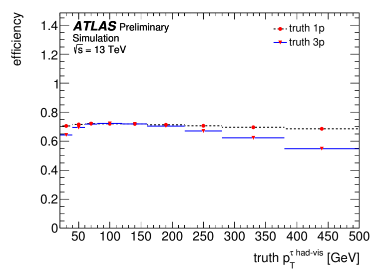


- Electrons – inner detector, electromagnetic calorimeter
- Photons – inner detector, electromagnetic calorimeter
- Muons – silicon detectors, muon spectrometer
- Hadronically decaying τ-leptons – inner detector, calorimeters
- Hadrons (jets) – inner detector, calorimeters
- Missing energy – calorimeters, inner detector

τ-lepton decay and reconstruction



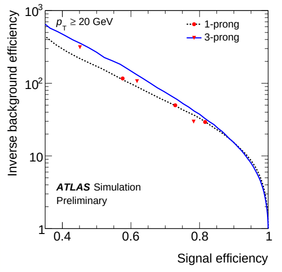


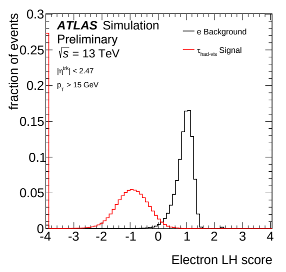


- Short decay time (0.3 ps)
- Leptonic decays: reconstructed as e or μ
- Hadronic decays (τ_{had}): 1 or 3 charged hadrons (prongs), 0 or more neutral hadrons
- Seeded by anti-k_T jet with $R = 0.4$
- Tau vertex identification
- 1 or more associated tracks

τ_{had} identification

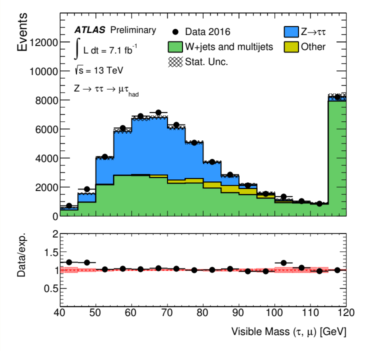
- Very little rejection against background jets from reconstruction
- Boosted Decision Tree (BDT) algorithms trained separately for 1- and 3-prong decays
- Exploit variables related to
  - τ-lepton decay time
  - Charged particle content (1 or 3 collimated tracks, isolation)
  - Neutral particle content (number of π^0)
  - Shower shape variables
- Loose/medium/tight working points
  - 1-prong: 60/55/45 %
  - 3-prong: 50/40/35 %
- Separate electron veto for 1-prong τ_{had}
  - Electron likelihood identification
  - Fixed 95 % efficiency





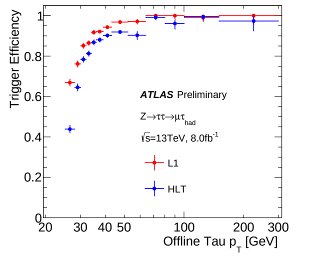
τ_{had} energy scale

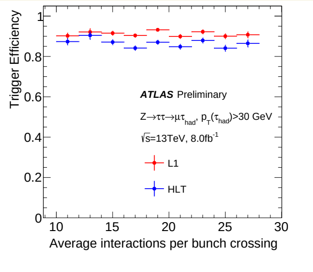
- The reconstructed energy at the local hadronic (LC) scale doesn't account for
  - The narrow cone size of τ_{had}
  - The particle content (typically π^0, $\pi^{\pm}$)
  - Underlying event or pile-up
- Two corrections
  - Pile-up correction (linear in N_{vtx})
  - Detector response (1- or multi-prong, $|\eta|$)
- Performance measured with a *tag-and-probe* analysis in $Z \to \tau\tau$ events
  - Muon used as *tag*
  - τ_{had} (loose) used as probe
  - Kinematic cuts to further reduce background
  - Data-driven multijet and W+jets background



The hadronic τ trigger

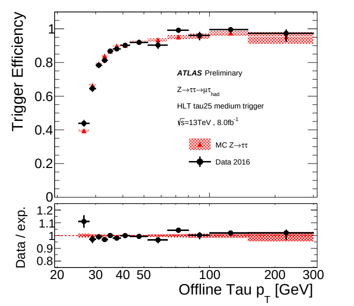
- The hadronic τ trigger is useful for many physics analyses
  - The τ-lepton is the heaviest of the leptons $\rightarrow$ large coupling to Higgs bosons
  - Leptonic decays are cleaner, but hadronic decays have higher branching ratio (BR)
- Trigger with lowest p_{T}-threshold is 25 GeV (prescaled)
- Triggers with the highest yield (only small prescale)
  - 2015: $p_{\mathrm{T}} > 80$ GeV
  - 2016: $p_{\mathrm{T}} > 125$ GeV
- Efficiency of trigger w.r.t. offline reconstructed τ_{had} candidates measured with *tag-and-probe* analyses in $Z \rightarrow \tau\tau$ and $t\bar{t}$ events





Trigger efficiency measurements

- Efficiency is measured for true τ_{had} candidates
  - Background τ_{had} candidates subtracted from the data
- Opposite-sign–same-sign (OS-SS) method used to model background
  - Sign-asymmetric analyses
  - Use shape of SS data
  - Normalize with OS to SS ratio (r_{QCD})
  - Jets misidentified as τ_{had} (mostly) covered by SS data
  - Leptons and rest of jets misidentified as τ_{had} covered by MC
- Signal taken from all true τ_{had} in MC
- Ratio of data - background and MC applied as a scale factor to MC events
  - TOT – events in selection
  - PASS – events that also fire trigger



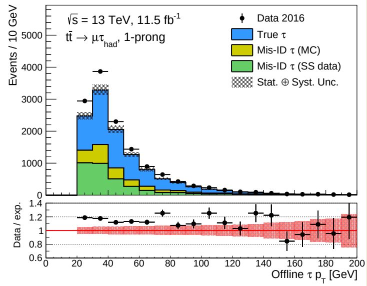
$$\varepsilon_{\text{sig}} = \frac{N^{\text{MC}}_{\text{true }\tau}(\mathrm{PASS})}{N^{\text{MC}}_{\text{true }\tau}(\mathrm{TOT})}$$

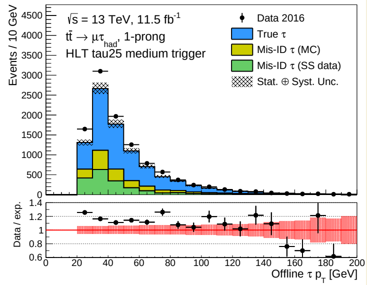
$$\varepsilon_{\text{data}} = \frac{N^{\text{data}}(\mathrm{PASS}) - N^{\text{bkg}}(\mathrm{PASS})}{N^{\text{data}}(\mathrm{TOT}) - N^{\text{bkg}}(\mathrm{TOT})}$$

$$\mathrm{SF} = \frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{sig}}}$$

$t\bar{t}$ trigger efficiency measurement (1)

- τ_{had} p_{T}-spectrum is kinematically limited in $Z \to \tau\tau$ events
- $t\bar{t}$ events offer a higher p_{T}-range, and:
  - Require b-tag $\to$ almost no W+jets or Z+jets background
  - Kinematic cuts and cut on Z mass window unavailable
  - $t\bar{t}$ events with jets misidentified as τ_{had} not entirely covered by SS data
- Event selection 2016
  - Muon trigger
  - $p_{\text{T}}^{\mu} > 26$ GeV, $|\eta| < 2.5$
  - $p_{\text{T}}^{\tau} > 20$ GeV, $0 < |\eta| < 1.37$ or $1.52 < |\eta| < 2.5$, 1 or 3 tracks
  - ≥ 2 jets, $p_{\text{T}} > 20$ GeV, $|\eta| < 4.5$
  - ≥ 1 b-tag, 77 % efficiency, $|\eta| < 2.5$
- Purity ~ 60 % (~ 70 % for $p_{\text{T}}^{\tau} > 60$ GeV)



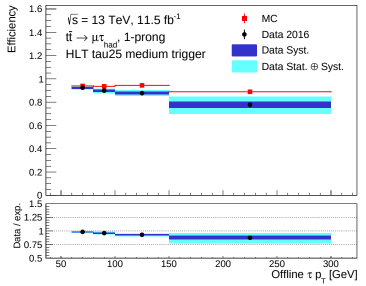


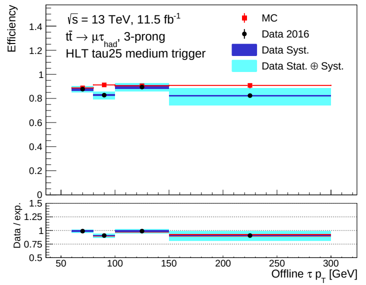
$t\bar{t}$ trigger efficiency measurement (2)

- Efficiency and scale factors for τ trigger with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$
- Systematic uncertainties arise from background subtraction
- Effects of systematic uncertainty on the weighted-average efficiency:

r_{QCD}	0.8 %
Pile-up	0.1 %
b-jets	$<$ 0.1 %
Muons	$<$ 0.1 %

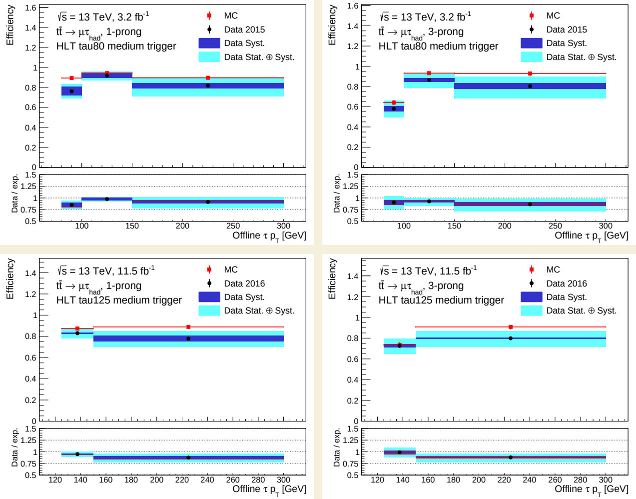
- Scale factors almost equal to 1
- Clear decline in efficiency for 1-prong τ_{had} candidates





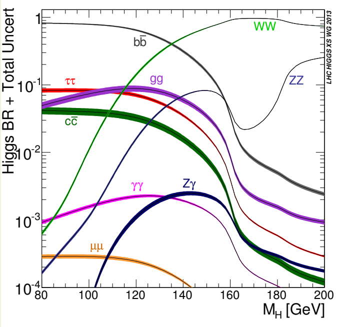
$t\bar{t}$ trigger efficiency measurement (3)

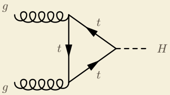
- Efficiency and scale factors "lowest unprescaled" τ trigger
- 2015: $p_{\mathrm{T}} > 80$ GeV
- 2016: $p_{\mathrm{T}} > 125$ GeV
- Scale factors almost equal to 1

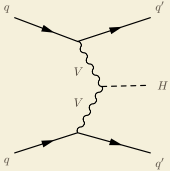


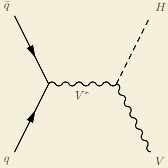
SM Higgs boson production and decay

- Discovery of a neutral scalar in 2012 with mass 125 GeV
  - CMS: Phys. Lett. B 716 (2012) 30
  - ATLAS: Phys. Lett. B 716 (2012) 1-29
- Primarily produced through three processes at the LHC
  - Gluon-gluon fusion (ggF)
  - Vector boson fusion (VBF)
  - Associated production (VH)



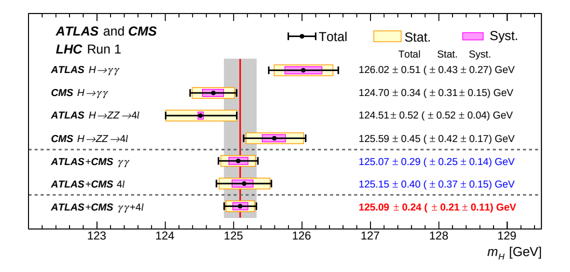






SM Higgs boson measurements

- Higgs mass measured in $H \to \gamma\gamma$ and $H \to ZZ \to 4l$ channels
  - $m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV
  - CMS+ATLAS: Phys. Rev. Lett. 114, 191803
- Coupling to bosons also measured
- Missing: coupling to fermions
  - $H \to b\bar{b}$ – difficult background conditions
  - $H \to \tau\tau$ – lower BR, but semi-leptonic channels help



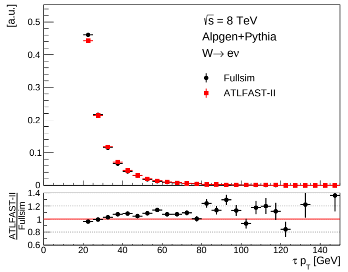
Phys. Rev. Lett. 114, 191803

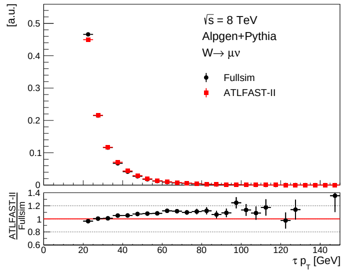
$H \to \tau\tau$ (1)

- Run 1 $H \to \tau\tau$ result JHEP 04 (2015) 117
  - $4.5\,\mathrm{fb}^{-1}$ 7 TeV
  - $20.3\,\mathrm{fb}^{-1}$ 8 TeV
- $H \to \tau\tau$ measured in three channels ($l = e$ or μ)
  - $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$
  - $H \to \tau_{\mathrm{l}}\tau_{\mathrm{had}}$
  - $H \to \tau_{\mathrm{l}}\tau_{\mathrm{l}}$
- ...and in two categories
  - VBF – VBF/VH, ≥ 2 jets, large $\Delta\eta$
  - Boosted – ggF, ≥ 1 jet, fail VBF
- Several triggers are used
  - Single e or μ trigger: $e\mu$, $\mu\mu$, $e\tau_{\mathrm{had}}$, $\mu\tau_{\mathrm{had}}$
  - Two-e trigger: ee
  - Two-τ trigger: $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$
- Separate BDTs for each channel and category (total: 6)
  - Cross-trained with 8 TeV data (even/odd)
  - Applied to both 8 TeV and 7 TeV data

$H \to \tau\tau$ (2)

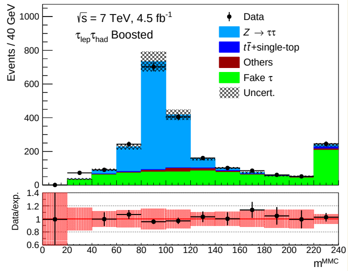
- Semi-leptonic channel exploits the clean signature of the electron or muon
- Irreducible $Z \to \tau\tau$ background
  - Embedding technique
  - $Z \to \mu\mu$ events from data, substitute with simulated τ decays
- Misidentified τ_{had} from jets and e/μ
  - Investigated ATLFAST-II (fast calorimeter simulation) for W+jets background
  - Fake factors derived separately for multi-jet, W+jets, $Z \to ll$+jets, and $t\bar{t}$ backgrounds
- Pre-selection cuts
  - Single e or μ trigger
  - 1 isolated e or μ, no other light lepton
  - 1 τ_{had}, 1 or 3 tracks, opposite sign to e/μ
  - $m_{\text{T}} < 70$ GeV (rejects W+jets)
  - No b-tagged jet with $p_{\text{T}} > 30$ GeV (rejects $t\bar{t}$)

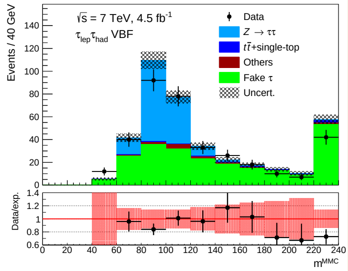




$H \to \tau\tau$ (3)

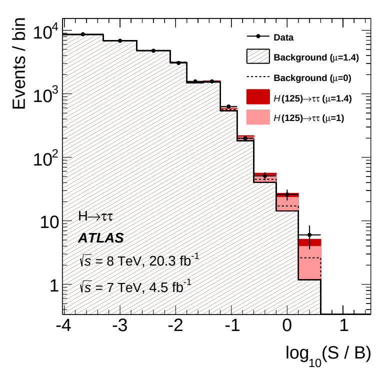
- Reconstruction of $\tau^+\tau^-$ mass not straight-forward due to presence of neutrinos
- 6 to 8 unknowns, but only 4 equations
- Missing mass calculator (MMC) technique scans
  - E_x^{miss} and E_y^{miss}, within resolution
  - Angles between τ-leptons and neutrino systems
  - Masses of neutrino systems (leptonic τ decays)
- Maximizes a likelihood function where PDFs are found from simulated $H \to \tau\tau$ and $Z \to \tau\tau$ events
- Target: good resolution of and separation between Z and Higgs boson mass peaks

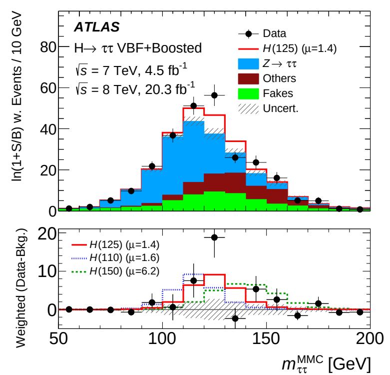




$H \to \tau\tau$ (4)

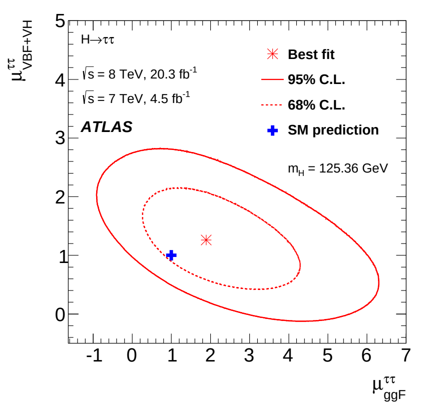
- Simultaneous profile-likelihood fit of the BDT distributions in all 6 channels
- Using $m_H = 125.36\,\text{GeV}$
- Signal strength: $\mu = 1.43^{+0.27}_{-0.26}(\text{stat.})^{+0.32}_{-0.27}(\text{syst.}) \pm 0.09(\text{th.})$
- Deviation from background-only hypothesis of 4.5σ
- Events in MMC plot weighted by $\ln(1 + S/B)$ for visualization





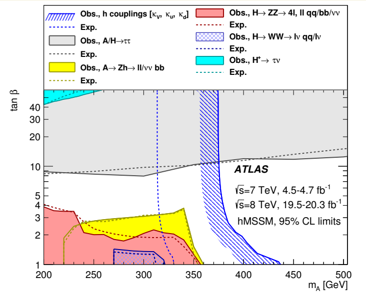
$H \to \tau\tau$ (5)

- Separate data into VBF+VH and ggF production modes
- Simultaneously measure their signal strenghts
- Result within 1σ of the SM prediction



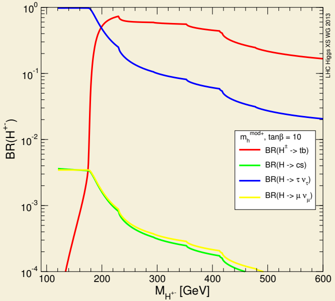
Higgs bosons beyond the SM

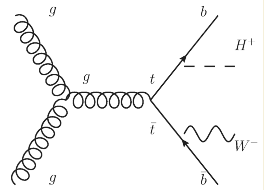
- Two-Higgs doublet models (2HDM)
  - Type I, Type II, lepton specific, flipped
  - Minimal Supersymmetric SM (MSSM) — special case of Type II
- In MSSM benchmark scenarios the Higgs sector parameters are determined by m_A and $\tan\beta$ at tree level
  - Examples: m_h^{max}, $m_h^{\text{mod+}}$, $m_h^{\text{mod-}}$
- hMSSM — a scenario where the mass of the lightest Higgs boson is set to $m_h = 125$ GeV, but where all other parameters are free
  - m_A and $\tan\beta$ decide the Higgs sector parameters to a good approximation (even at NLO and NNLO)
  - Can ignore radiative corrections when computing Higgs masses

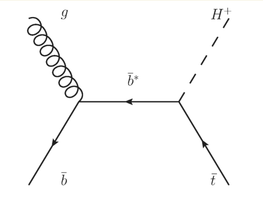


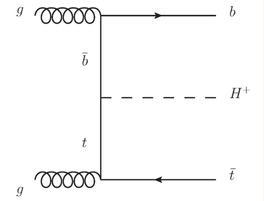
H^+ production and decay

- Charged Higgs bosons are produced
  - In top-quark decays, if $m_{H^+} < m_t$
  - In association with top quarks, if $m_{H^+} > m_t$
- ...and decay primarily to
  - $H^+ \to \tau\nu$, if $m_{H^+} < m_t$
  - $H^+ \to tb$, if $m_{H^+} > m_t$
- Depending on scenario, BR for $H^+ \to \tau\nu$ can still be sizeable, and offers a cleaner signature than $H^+ \to tb$



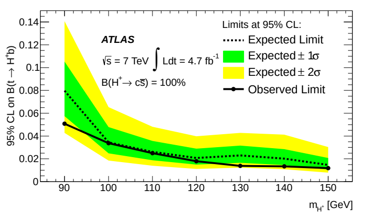


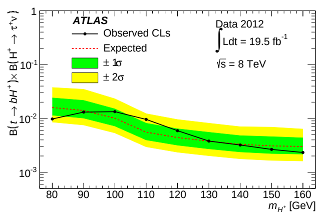


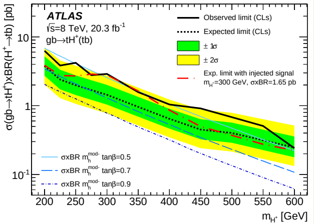


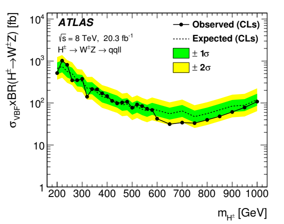
Run 1 H^+ searches

- Charged Higgs bosons searched for in multiple channels in Run 1
  - $H^+ \to cs$, 7 TeV
    Eur. Phys. J. C, 73 6 (2013) 2465
  - $H^+ \to \tau\nu$, 8 TeV JHEP03 (2015) 088
  - $H^+ \to tb$, 8 TeV JHEP03 (2016) 127
  - $H^+ \to W^+Z$, 8 TeV
    Phys. Rev. Lett. 114, 231801 (2015)









$H^+ \to \tau\nu$ (1)

- Search for $H^+ \to \tau\nu$ with $14.7\,\text{fb}^{-1}$ 13 TeV data from 2015 and 2016 ATLAS-CONF-2016-088
- Cross section increases 4–10x (m_{H^+} and model dependent)
- τ_{had} +jets channel (hadronic W decay)
  - $E_{\text{T}}^{\text{miss}}$ trigger
  - 1 τ_{had}, $p_{\text{T}} > 25\,\text{GeV}$, 1 or 3 tracks, $0 < |\eta| < 1.37$ or $1.52 < |\eta| < 2.5$
  - ≥ 3 jets, ≥ 1 b-tag
  - $E_{\text{T}}^{\text{miss}} > 150\,\text{GeV}$
  - No light leptons
- Use transverse mass as discriminating variable

$$m_{\text{T}} = \sqrt{2p_{\text{T}}^{\tau} E_{\text{T}}^{\text{miss}} (1 - \Delta\phi_{\tau,\text{miss}})}$$

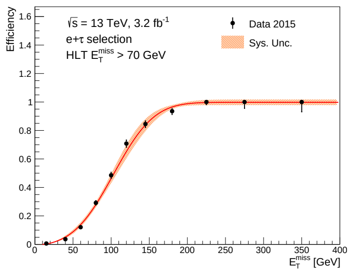
- Extra cut on $m_{\text{T}} > 50\,\text{GeV}$ to avoid events with mismeasured $E_{\text{T}}^{\text{miss}}$
- Dominant background is jets misidentified as τ_{had} from
  - $t\bar{t}$ events, in *low*-m_T region
  - multi-jet events, in *high*-m_T region

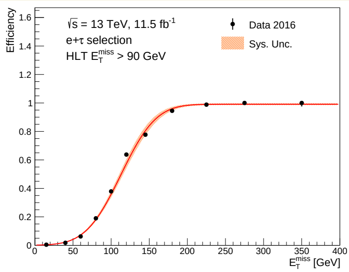
$H^+ \to \tau\nu$ (2)

- Trigger efficiency measured in data
  - $e + \tau_{\text{had}}$ + jets selection
  - Fitted to the error function

$$F(x) = p_0 \cdot \left[1 + \left(\frac{x - p_1}{p_2}\right)\right] + p_3$$

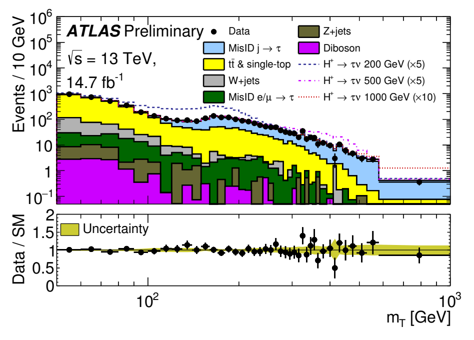
- Systematic uncertainties derived by varying
  - e identification
  - τ_{had} identification
  - Event selection
  - Statistical precision
- Fitted efficiency aplied to MC events
- Triggers used
  - 2015: Topo-cluster, $E_{\text{T}}^{\text{miss}} > 70$ GeV
  - 2016: Jet-based, $E_{\text{T}}^{\text{miss}} > 90$ GeV



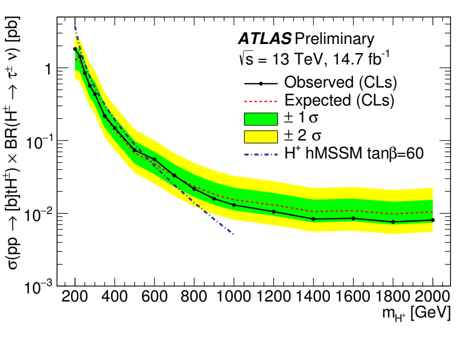


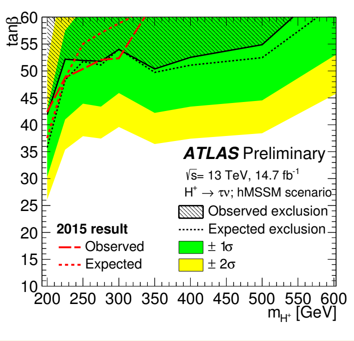
$H^+ \to \tau\nu$ (3)

- Background from jets misidentified as τ_{had} treated by the fake factor method
- Fake factors parametrized by p_{T}, separately for 1- and 3-prong τ_{had}
- m_{T} plot after fit to background-only hypothesis
- Signal hypotheses added on top
  - Cross section as predicted by the hMSSM scenario at tan$\beta = 60$
  - 200 GeV (5x), 500 GeV (5x), 1000 GeV (10x)



- Profile-likelihood fit of m_T distribution in signal region
- Systematic uncertainties included as nuisance parameters
- No deviation from SM background
  - 95 % CL exclusion limits on $\sigma(pp \to [b]tH^+) \times \mathcal{B}(H^+ \to \tau\nu)$: 1.8 pb–14 fb for $m_{H^+} = 200$ GeV to 2000 GeV
  - Exclusion in m_{H^+}–$\tan\beta$ plane as interpreted in hMSSM





Conclusion

- Hadronically decaying τ leptons are important to many analyses
  - Reconstruction, identification, and energy calibration
- Hadronic τ trigger
  - Efficiency measurement in $Z \to \tau\tau$ events
  - New method to measure efficency in $t\bar{t}$ events
- SM Higgs boson measurement in $H \to \tau\tau$
  - Full 7 TeV and 8 TeV datasets
  - Best fit signal strength $\mu \approx 1.4$
  - Corresponds to 4.5σ
- Charged Higgs boson searches
  - Many searches performed in Run 1
  - Search performed in $H^+ \to \tau\nu$ with 14.7 fb^{-1} 13 TeV data from 2015 and 2016
  - Limits on $\sigma(pp \to [b]tH^+) \times \mathcal{B}(H^+ \to \tau\nu)$:
    1.8 pb–14 fb for $m_{H^+} =$ 200 GeV to 2000 GeV
  - Exclusion in m_{H^+}–tanβ plane as interpreted in hMSSM

Bonus

Cloud computing R&D

- ATLAS analysis cluster on Google Compute Engine (GCE) IaaS
- Using Puppet to configure the cluster
- Using cvmfs, Condor, XRootD, AutoPyFactory to run ATLAS jobs