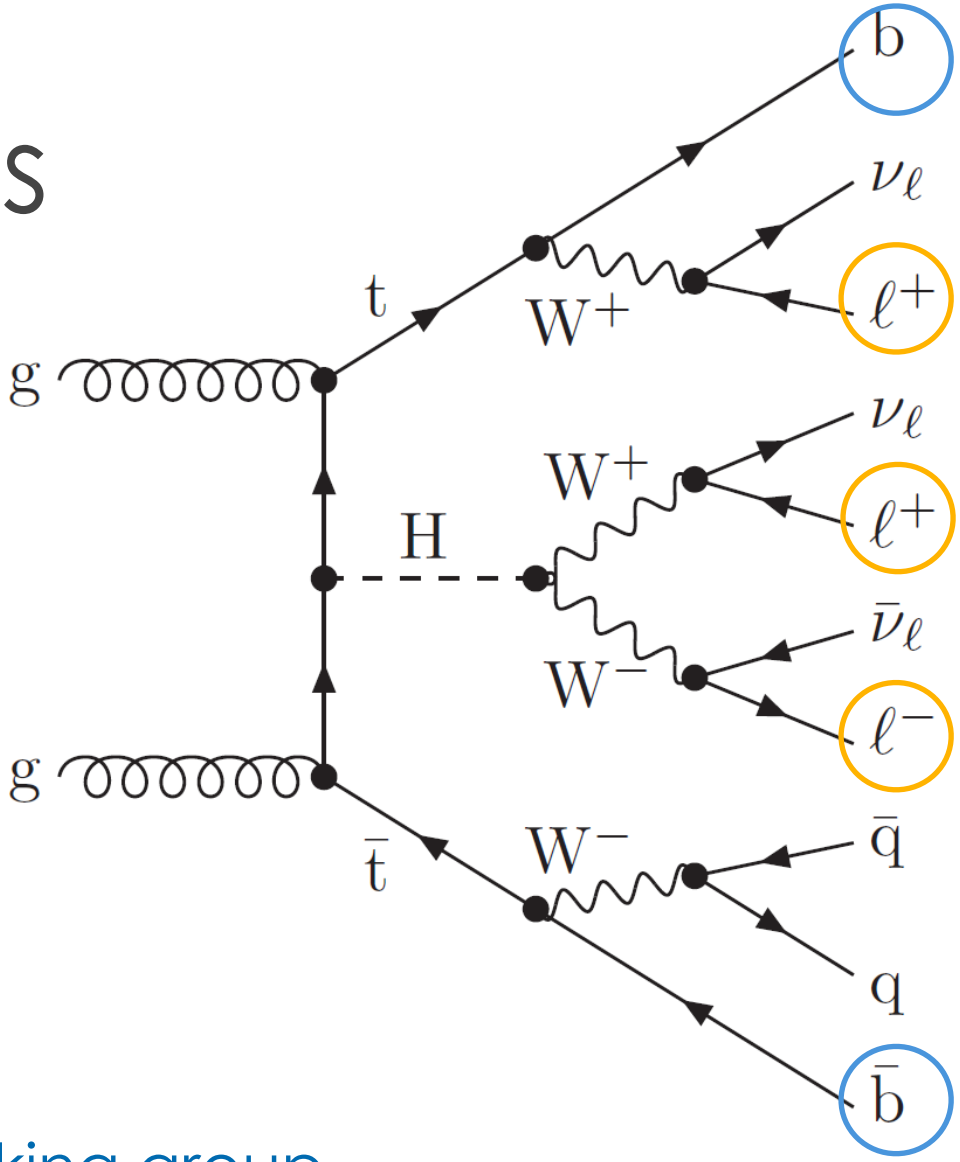
Status and plans towards the November analysis review



Cristina Botta

for the ttH leptonic working group

ttH leptonic working meeting 24.09.2015

Target

- ttH->leptons analysis PAS with 2015 data for Moriond 2016
  - from latest LHC news: we expect between 2.3 and 3.4 fb^{-1}
    - ➡ if similar S/B, σ x 4, L/10 : 1.6 x (8 TeV analysis sensitivity)
    - ➡ expected uncertainty on μ(ttH)~1.9 (we were at 1.2 @ 8 TeV)
      - ➡given that we were measuring μ=4 @ 8 TeV a measurement with a 1.9 uncertainty may be of interest
- First important milestone:
  - Full analysis review at the HWW MiniWorkshop 23/24 November
    - ➡ AtoZ MC strategy
    - ➡ Control of object performances in data and scale factors
    - ➡ Data driven bkg estimates in place

Samples

- 25ns samples from the 1B Spring15 campaign (waiting for Spring15MiniAODv2)
  - for final signal extraction
    - ttHJetToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8_mWCutfix/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
    - ttHJetToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8_mWCutfix/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9_ext1-v1/MINIAODSIM
    - TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
      - the associated M-1to10 is in preparation
    - TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
    - TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
    - WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
      - or the Madgraph one when it will become available
    - ZZTo4L_13TeV_powheg_pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v1/MINIAODSIM
      - or a new Madgraph one (the current version is bugged)
    - rares: WWsamesign, Tribosons
  - for lepton MVA and final BDT training
    - ttHToNonbb_M125_13TeV_powheg_pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v2/MINIAODSIM
    - TT_TuneCUETP8M1_13TeV-powheg-pythia8/RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9-v2/MINIAODSIM
      - or the semileptonic Madgraph one
    - ttZ, ttW madgraph LO in preparation
  - for bkg data driven estimates
    - all QCD samples MuEnriched and EMEnriched for fake-rate estimate
    - TTJets (maybe the semileptonic Madgraph one), Wjets for fake-rate estimate
    - DYJets for Charge-Flip probability estimate

Triggers

- As for in Frozen 25ns Menu: /frozen/2015/25ns14e33/v4.2/HLT/V1
  - Double Electron Iso: **17/12** instead of 23/12 (available in MC)
  - Single Ele Iso: new Ele23_WPLoose_Gsf instead of 32eta2p1 (but not available in MC)
    - lowest threshold un-prescaled present both in data and MC: Ele**27**_WP85
  - Double Mu Iso: **17/8** (available in MC)
  - Single Mu Iso: new IsoMu18 instead of 24eta2p1(but not available in MC)
    - lowest threshold un-prescaled present both in data and MC: IsoMu**20**

Lepton flavour & p_T categories

8 TeV analysis

- ee (20,20) from DoubleEle triggers
- μμ (20,20) from DoubleMu triggers
- eμ (20,20) from MuEG triggers
- 3l (20,10,5/7) from Double/Triple Lepton triggers
- 4l (20,10,5/7) from Double/Triple Lepton triggers

Lepton flavour & p_T categories

13 TeV analysis

- ee (20,15) from Double and Single Ele triggers + (30,7-15) from Single Ele
  - to be splitted in 0 or >= 1 hadronic taus
- μμ (20,15) from Double and Single Mu triggers + (30,5-15) from Single Mu
  - to be splitted in 0 or >= 1 hadronic taus
- eμ (20,15) from MuEG and SingleMu/Ele triggers + (30,5/7-15) from Single Mu/Ele
  - to be splitted in 0 or >= 1 hadronic taus
- 3l (20,10,5/7) from Double/Triple Lepton triggers
- 4l (20,10,5/7) from Double/Triple Lepton triggers
- **Comments**
  - additional low sub-leading p_T category bring ~10% signal efficiency in each channel but need to check S/B (Geoff, Charlie to finalize the study with new lepton mva)
  - is it worth it to add the tau categories with 2/3 fb^{-1} ? (Jan to answer this question)

Recall: Lepton selection

- **Loose:** for vetoing additional leptons in each categories
  - Common: |dxy| < 500um, |dz| < 1 mm, MiniRelIso (dR=0.3, rho*EA) < 0.4, $SIP_{3D} < 8$
  - Muons: Loose ID
  - Electrons: VLoose WP of Phys14 non-triggering MVA ID, lost hits ≤ 1

- **Tight:** for signal region selections
  - Common: lepMVAttH <0.6-0.8
  - Muons: Medium ID
    - for ss-dilepton only: track p_T/ track p_T error < 0.2
  - Electrons: passConversionsVeto, lost hits = 0
    - for ss-dilepton only: isGsfCtfScPixChargeConsistent

- **Fakable-object:** for fake-rate denominator
  - Same as tight but without lepMVAttH requirement and without Medium Muon ID

Recall: lepMVAttH

input variables will be using in RunII

Isolation

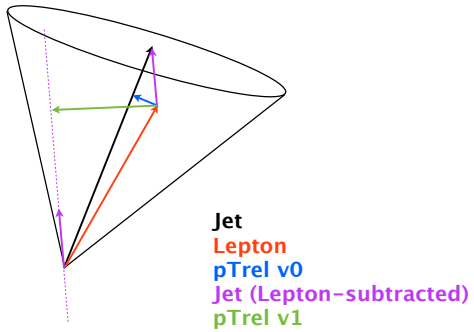
- **PF miniRelIso, charged** had. (R=0.3)
- **PF miniRelIso, neutral** had. & photon (R=0.3, with EA*rho PU corrections, scaled with (0.3/R)^2)

Vertexing

- **3D IP significance** (SIP_{3D})
- **2D IP |dxy|** and **|dz|**

Lep-Jet

- **Lepton's closest jet** (leptons are not removed from jet collection)
  - $p_T(\ell)/p_T(\text{jet})$
  - **Lepton's p_T^{rel} wrt jet**
  - jet CSV b-tag



ID

- **(μ)Segment compatibility/**
- **(e)POG non-triggering MVA Electron Id discriminator**

Lepton ID on going studies

- **Effective Ares for miniIso**
  - for the moment still using those extracted from PHYS14 MC
  - we are deriving those for Spring15 25ns samples, in collaboration with POGs
    - Marco will send final results next week
- **Electron MVA ID**
  - for the moment still using PHYS14 non-triggering training
  - will have to move to EGamma Spring15 25ns non-triggering training
    - available in miniAODv2 (~1/2 weeks)
  - VLoose WP being extracted
- **Commissioning of input variables**
  - a lot of results with 50ns and 25ns data already reported at the HWW meeting
  - new interesting studies on going on "how to apply JEC to the jet associated to the lepton" while computing the lepton-jet associated variables
    - JEC lep-aware proposal being discussed with JET/MET, SUSY PAG etc…
      - will be reported next week in our meeting by Giovanni/Marco

Lepton ID on going studies

- **Lepton MVA new training**
  - with new electron MVA preselection, new electron MVA training as input variable (when available)
  - with new lep-aware JEC for lepton-jet related variables
  - with weighted-lepton p_T as input variable (to partially smooth the difference between the signal and background leptons p_T spectrum)
  - new working point being derived taking into account:
    - that we might want to be looser on low p_T leptons
    - the additional separation power between ttH and tt we can achieve with the final BDT
  - Giovanni and Francesco are working on it

Data driven backgrounds

- Fakes: tight-to-loose method as last year (Marco, Giovanni)
  - with cone-corrected p_T improvement to reduce differences in fake-rate estimate between QCD and ttbar events
    - to be checked if current trigger strategy (~lepton p_T cuts) allows for enough sideband in the fake rate application
  - to be checked with 13 TeV data and MC the tight-to-loose ratio flavour dependence to decide how to approach the **b-tagging categorisation for the signal-extraction**
    - two fake-rate from two different measurement region are needed?
- Charge flip: same method as last year (who?)
  - or possible improvement:
    - extract data/MC SF for charge-flip probabilty in coarse lepton p_T bins
    - then interpolate and extrapolate the p_T dependence using MC
- WZ: same method as last year (who?)

Signal extraction

- **BDT** as in 8 TeV analysis with following improvements
  - might add new input variables
    - ➡ angular variables proposed by Milos (to be studied also with Spring15)
    - ➡ reconstruction of hadronic/leptonic top mass (Giovanni is working on this)
    - ➡ kinematic fits (Benjamin interested in this)
  - new categorisation
    - ➡ b-tight and b-loose
    - ➡ split 3l category in 2-3/4+ jets to better separate ttZ/ttW from ttH
    - ➡ add 3j category in the 2lss final state to control ttW from data
    - ➡ add the 3l category with inverted Z veto to control ttZ from data
  - 2D approach (against tt, against ttW/Z) as tested with Phys14 samples by Predrag and Milos
- **MEM**
  - what is the current status?
  - what is expected for mid November?

Additional topics/proposals to discuss?