# H->bb Weekly Meeting

Ricardo Gonçalo (RHUL)

HSG5 H->bb Weekly Meeting, 15 March 2011

News! News! News!

- First colliding stable beams run in 2011 on Sunday
  - Entire run: 28.34 nb^{-1}
  - Stable beams: 27.38 nb^{-1}
  - Peak lumi: 1.3×10^{30} cm^{-2}s^{-1}
  - Approx. lifetime: 21 h
  - 3 colliding bunches
  - 1.8×10^{11} protons/bunch
- Will progressively increase intensity over next 2 weeks:
  - Should get 32 bunches later today!
- Technical stop early April for 2 weeks of scrubbing – should improve lifetime

Online lumi [ATLAS_PREFERRED] per LB for run 177531

Integrated luminosity: 28.34 nb-1 (all LBs) and 27.38 nb-1 (stable beams)

LBs with stable beams
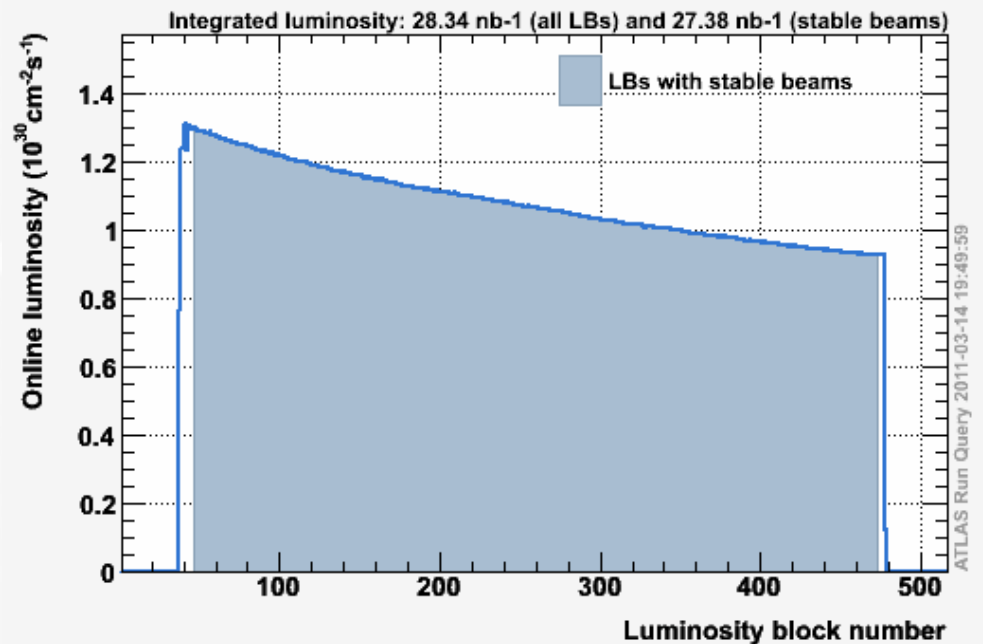
News! News! News!

- New database interface to access data period information: https://atlas-tagservices.cern.ch/RBR/rBR_Period_Report.php

  – Feedback requested to hn-atlas-physicsMetadata@cern.ch.

  – This replaces data prep wiki: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DataPeriods

HSG5 Validation

- Arnaud asks how to update the Trigger menu in the HSG5 validation package

- Triggers currently being monitored:
  - L1 Triggers: L1_MU10, L1_MU15, L1_MU20_XE30, L1_EM10, L1_EM14, L1_EM18_XE30, L1_TAU9I_3J5_2J20, L1_TAU9_XE15, L1_XE40
  - L2 Triggers: L2_mu10, L2_mu15, L2_mu20_xe30, L2_e10_medium, L2_e15_medium, L2_e20_loose, L2_e20_loose_xe30, L2_tau16i_loose_3j30, L2_tau16_loose_xe20, L2_xe40
  - Event Filter Triggers: EF_mu10, EF_mu15, EF_mu20_xe30, EF_e10_medium, EF_e15_medium, EF_e20_loose, EF_e20_loose_xe30, EF_tau16i_loose_3j30, EF_tau16_loose_xe20, EF_xe40

- Several of these are already obsolete…

- Please check what your triggers will be for 2011 (until Summer) and let us know

Fast Monitoring

- Meeting on Wednesday last week: http://indico.cern.ch/conferenceDisplay.py?confId=130616
- Use Athena ManagedMonToolBase, which is also used in normal monitoring
- Not something really urgent for us, just something to keep an eye on

- Mailing list:

  atlas-phys-fastmonitoring@cern.ch

- Wiki-page:

  https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ FastPhysicsMonitoring

- Indico area:

  http://indico.cern.ch/categoryDisplay.py?categId=3416

Dataflow

From Tier0 → AOD

CAF/TMS — AOD Processing

Relatively clear plan how to do

- Quick-And-Dirty analysis based on AOD in athena to produce ntuples of interesting events and/or histograms
- Ntuples should contain run/event/LB number maybe also the StreamRef
- Integrated in Reco_trf.py

Grid — Intermediate ROOT file

- One root file containing many ntuples (possibly also histograms)
- Should be small enough to open in interactive root for at least one run (better more)

CAF/TMS — Histogram creation + summing

- Root-based SW to process ntuples and histograms
- Combine with data taken so far
- Output: DQM-compliant histograms
- Should be quick to re-run if needed

GRL → Summed histograms → DQM Webdisplay

Largly unclear how to do

The intermediate ROOT file

- Can contain ntuples (preferred) or histograms
    - ntuple gives us more flexibility afterwards. Writing a histogram directly from the AOD should be used only if the ntuple would become too big
- Size estimate for intermediate ROOT file
    - Assume:
        - 10 float/int per analysis
        - 20 different analyses
        - Skim 1% of events (probably much less in reality)
        - 1e9 event/year

    =8GBytes * root-compression

    Small enough to store/process interactively
- Characteristics/Requirements:
    - One ROOT file containing many small, skimmed TTrees and TH*
- What existing infrastructure could be re-used?
    - Considered TAG, D3PD, build something from scratch
    - **AthenaMonitoring** appears to be the most appropriate

Proposed Roadmap for WH Analysis

- After the effort on cut flows, we're ready to start producing results!
  - Concentrating on un-boosted results here only because it's still unclear what would be feasible in boosted analysis until Dubna – commissioning work ongoing
  - BUT: work on boosted VH is starting in parallel – see e.g. Wahid's talk today
- Intended results:
  - Cut-based analysis focusing on WH -> e/μ ν b b
    - I think there should be at least 2 analyses, for cross checking results
    - Ideally using 2 different data formats (AOD vs D3PD)
  - Multivariate analysis in parallel, to improve on cut-based analysis
- Timeline:
  - Analyses should be semi-frozen by Dubna (17 – 19 May)
  - This leaves around 7 weeks
- Results in the form of:
  - Histogram with # events vs m_H
  - Table of # events expected for each value of m_H and background type – including statistical and systematic uncertainties
  - Exclusion plot vs m_H (95% C.L. limit on σ/σ_{SM})
  - …plus control plots etc

Questions to be answered

- Cut-based analysis focusing on WH -> e/µ ν b b
  - Establish analysis selection: why is each cut applied and why at each particular value? Are we convinced this is the right thing to do?
  - Establish set of systematic uncertainties: start from combined performance group recommendations. What are the most important? Are there any hidden pitfalls for us?
  - What b-taggers and why? What calibrations do we expect to be ready in time? What is the corresponding systematic uncertainty?
  - What sort of exclusion limits can we expect for 0.5, 1, 2 fb^{-1} ?

- Multivariate analysis in parallel, to improve on cut-based analysis
  - What event preselection should be used and why?
  - Use for signal-background separation only or target particular backgrounds?
  - What are the possible bias? Where can it go wrong?
  - What improvement can be expected wrt cut-based analysis for 0.5, 1, 2 fb^{-1} ?

- Exclusion plot vs m$_H$ (95% C.L. limit on σ/σ$_{SM}$)
 - Need someone to implement RooStats workspace

Date	Milestones wish list
17 May	Dubna workshop – analysis frozen After this: add data to un-boosted analysis and prepare for result approval Concentrate more effort on boosted VH with a view to obtaining results quickly
10 May	Review results with 2011 data from cut-based and multivariate analyses
3 May	Margin for dealing with unforeseen problems
26 April	Start looking at 2011 data if enough is available. Any surprises? How does the MC describe the new data? By now we should have a reasonable idea of results from the multivariate analysis
19 April	End of 2 weeks of beam scrubbing. (I'm away for Easter)
12 April	By now we should have a reasonable idea of the exclusion of the cut-based analysis First report on MVA preliminary results – establish plan for getting results by Dubna
5 April	Identify the worst systematics and discuss any possible improvements: •Any changes needed in analysis cuts? •Any study necessary for corrections to some systematic effect? Multivariate analysis: iterate on preselection cuts, methods, questions Assign tasks – divide the work to achieve better results!
29 March	Establish analysis cuts: •If possible as result of optimization •Use 2010 data to develop cuts and show that data is well described by background MC Start evaluating systematics
22 March	Iterate on analysis cuts – why is each cut applied at each particular value? Start iteration on multivariate methods to improve analysis

Backup
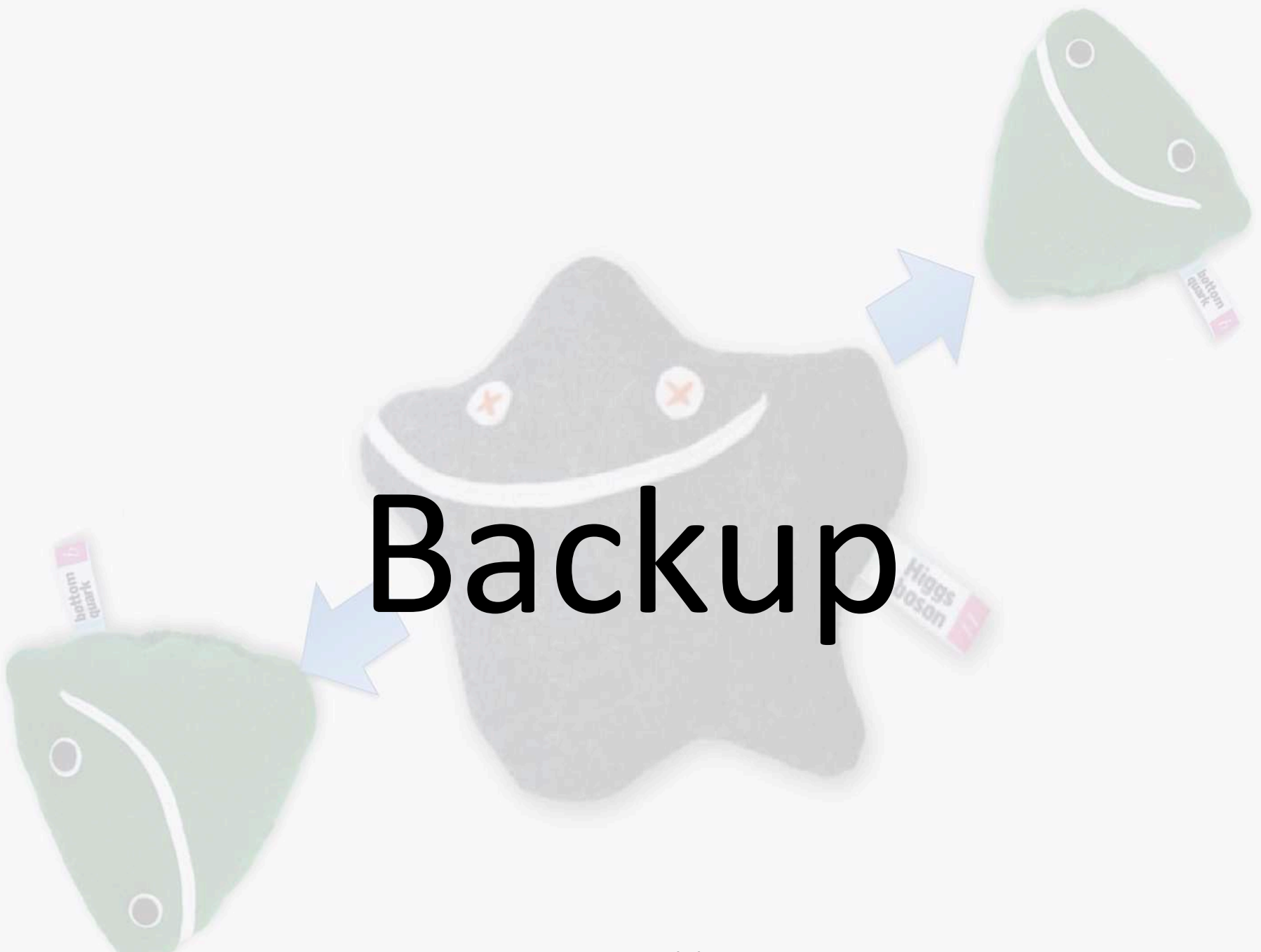
To boost or not to boost...

- We expect a sizeable luminosity this year:
  - 0.6 fb^{-1} for PLHC in June; 1 fb^{-1} for EPS in July; 2 fb^{-1} for Lepton-Photon in August
- Need to get boosted VH analyses going again!
- Work is currently ongoing within the SM group – mainly on commissioning the jet substructure reconstruction
- Looking at a H spectrum in WH (by eye!):
  - ≈0.5 of the cross section above p_T^H = 100 GeV
  - ≈7.5% above p_T^H = 200 GeV
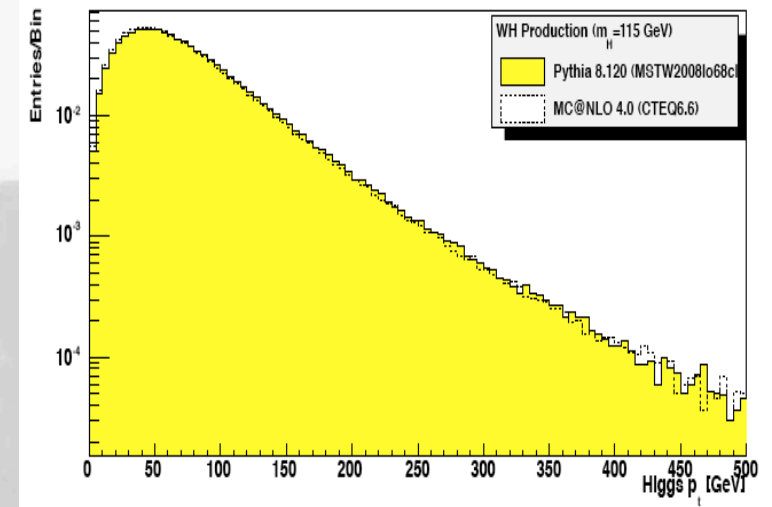
| √s = 7 TeV	p_T^H > 0 GeV	p_T^H > 100 GeV	p_T^H > 200 GeV
Channels	σ x BR(e/μ) (fb) OR # events per fb^{-1}	σ x BR(e/μ) (fb) OR # events per fb^{-1}	σ x BR(e/μ) OR # events per fb^{-1}
WH->lvbb	150	75	11
ZH->llbb	29	14	2.2

To boost or not to boost...

Un-boosted channel sensitivity:

- All we currently have at √s = 7 TeV are Lianliang's numbers from the 15/12/2010 H->bb meeting: http://indico.cern.ch/conferenceDisplay.py?confId=115169

- WH (e and μ channels) gives 56 x SM exclusion with 35 pb^{-1} at m_H = 120 GeV

- Includes 10% systematic uncertainty on signal and background and 11% on luminosity

- Assumes IP3D+SV1 will be usable for this analysis

- For 1 fb^{-1} this translates roughly as a factor 5 improvement, i.e. exclude ≈ 10 x SM

Boosted channel sensitivity:

- Exclude 6.7 x SM at m_H = 120 GeV with 1 fb^{-1} - see ATL_PHYS_PUB-2010-015

- Obtained by scaling 14 TeV numbers – room for optimization at √s = 7 TeV

- Analysis selects p_T^H > 200 GeV

- WH(W->lν) most promising, followed by ZH(Z->νν)
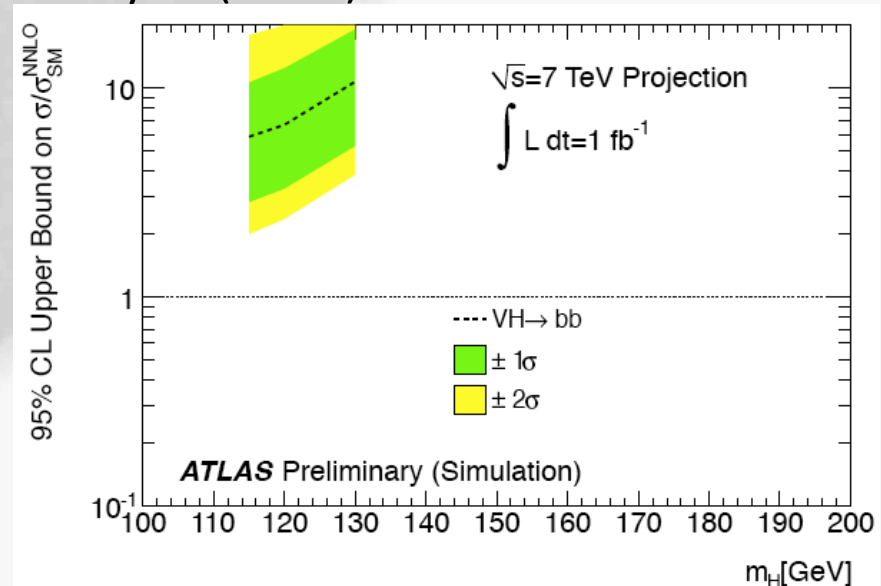
To boost or not to boost…

- Boosted VH Status and Plans:
  - Work ongoing in SM group on commissioning jet substructure technique
    - See Adam's presentation in last week's SM plenary: http://indico.cern.ch/getFile.py/access?contribId=0&resId=0&materialId=slides&confId=127720
    - Jet substructure note should be ready in 1-2 months

  - Work ongoing in parallel to provide b-tagging algorithm tuned for fat jet environment

  - Given infinite time, the plan would be to follow with measurements of g->bb and W/Z+bb with the substructure technique, and then VH(H->bb)

  - But we don't have infinite time! – need to put effort into optimizing the boosted VH analysis for 7 TeV
    - Edinburgh group has started on this – presentation next week – but more effort will be needed

To boost or not to boost…

- My view:
  - We need to have results for Summer conferences:
    - To improve the ATLAS sensitivity at low mass – follow the Tevatron mantra which says that "every channel counts"
    - Develop experience in the group on b tagging, WH reconstruction, setting limits and combining results – pave the way for later
    - Put H->bb on the map of useful channels
  - The un-boosted channels seem to be in a better position for this – remember only 2-3 months left, analysis should be ≈ frozen end of May
  - At the same time boosted VH has better sensitivity and must be our real goal for physics results in associated VH production this year
    - What could be shown from this channel in Summer?
  - How do we organize ourselves to achieve **both** goals?

- My proposal:
  - Split the work so that some people concentrate on un-boosted analysis and some on optimizing boosted analysis and on helping to commission jet substructure method
  - Get results out for Summer and move on fully to boosted analysis

Reconstruction issues

- **Muon** CP group recommendations for release 16:
  - Reconstruction efficiency and isolation efficiency scale factors, momentum smearing functions
  - https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesRel16

- Jet/Etmiss recommendations for **jet cleaning** in release 16:
  - Medium jet cleaning should give similar rejection to rel 15 cleaning but with better efficiency
  - Tight jet cleaning should not be used – still under discussion
  - https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets#Bad_jets_rel16_data

- New!: Final **b-tagging calibrations** for release 16 based on full 2010 data:
  - https://twiki.cern.ch/twiki/bin/view/AtlasProtected/Analysis16

- e/gamma recommendations for **energy scale and resolution** in release 16:
  - https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EnergyScaleResolutionRecommendations
  - And rescaler tool: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EnergyRescaler

- Standard Model **W/Z** group **baseline selection** for release 16 (next 4 slides):
  - See discussion in W/Z group Sharepoint
  - Also, finer points (and perhaps the not so fine) still being discussed

Muon Channel Selection

Muon Object and MET Definition

- STACO Combined Muon
- Author=1 or 6
- MCP Quality Cuts for rel16
- MET Definition: MET_LocHadTopo
  - ex = MET_LocHadTopo_etx + MET_MuonBoy_etx - MET_RefMuonTrack_etx;
  - ey = MET_LocHadTopo_ety + MET_MuonBoy_ety - MET_RefMuonTrack_ety;
  - MET = sqrt(ex2+ey2);

Preselection

1. Vertex Cuts

- $N_{vtx} \geq 1$ with $N_{tracks} \geq 3$

2. Jet Cleaning Cuts

- Reject Events with ≥ 1 loose bad jet
- Reject Events with ≥ 1 ugly jet

3. Muon PreSelection

- ≥ 1 offline Muon with $p_T > 15$ GeV
- $|z_0^{wrtPV}| < 10$ mm for the 0th vertex

4. Trigger

- E4-G1 : EF_mu10_MG (seeded by L1_MU0)
- G2-I1 (up to run 167576) : EF_mu13_MG (seeded by L1_MU0)
- I1 (from run 167607) - I2 : EF_mu13_MG_tight (seeded by L1_MU10)

W → μν Selection

1. Preselection
2. One muon with

- $p_T > 20$GeV
- $|\eta| < 2.4$
- Isolation $\Sigma p_{ID}^{Cone0.2}/p_T < 0.1$ (Tracks must come from selected vertex)
- Veto Events with 2nd cb muon with $p_T > 20$GeV

3. Missing Energy based on MET_LocHadTopo

- $E_T^{Miss} > 25$GeV

4. Transverse Mass

- $M_T > 40$GeV

Z → μμ Selection

1. Preselection
2. Two muons with the largest p_T which fulfill

- $p_T > 20$GeV
- $|\eta| < 2.4$
- Isolation $\Sigma p_{ID}^{Cone0.2}/p_T < 0.1$ (Tracks must come from selected vertex)

3. Opposite Charge
4. Invariant Mass Cut

- $66 < m_{\mu\mu} < 116$GeV

Electron Channel Selection

Electron Object and MET Definition

- use v16 OTX check and only using final map for run 167521
- use simple combination of E of the cluster and η and φ from the track (unless the track has <4 SCT +Pixel hits)
- η-cuts should be performed with cluster coordinates for OTX, |η|<2.47, crack removal
  - More Details under: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronsEnergyDirection
- MET Definition: MET_LocHadTopo
  - To be corrected for energy scaling/smearing

Preselection

1. Vertex Cuts
- $N_{vtx} \geq 1$ with $N_{tracks} \geq 3$
2. Jet Cleaning Cuts
- Reject Events with ≥ 1 loose bad jet
- Reject Events with ≥ 1 ugly jet
3. Electron PreSelection
- ≥ 1 offline electron with p_T>15 GeV with egammaPID::Medium_WithTrackMatch
4. Trigger
- Period ABCDE1-E3 (up to and including run 160879) : L1_EM14
- Period E4-I2: Trigger EF_e15_medium

W → eν Selection

1. Preselection
2. One electron with
- egammaPID::Tight_WithTrackMatch
- Author 1 or 3
- p_T >20GeV
- |η|<2.47 excl. 1.37 < |η| < 1.52
- Veto Events with 2nd medium electron with p_T>20GeV and electron isolation: 4 GeV on corrected E_T^{cone20}
3. Missing Energy based on MET_LocHadTopo
- E_T^{Miss} >25GeV
4. Transverse Mass
- M_T >40GeV

Z → ee Selection

1. Preselection
2. Two electrons with the largest p_T which fulfill
- Author 1 or 3
- egammaPID::Medium_WithTrackMatch
- p_T >20GeV
- |η|<2.47 excl. 1.37 < |η| < 1.52
3. Opposite Charge
4. Invariant Mass Cut
- 66<m_{ee} <116GeV

Tight Selection and Jet Definition

1. Additional Tight Selection for W/Z+jets

- Electrons: Require additional relative Isolation (on corrected isolation energy):

$$E_T^{cone20,Corrected}/P_T^{ele} < 0.2$$

  - Note: cut value need to be adjusted
  - this is not yet in the WZ-D3PDs but can be calculated via el_Etcone20
    - https://svnweb.cern.ch/trac/atlasusr/browser/mfiascar/WjetsD3PDskimming/trunk/IsoCorrection.cxx
    - https://svnweb.cern.ch/trac/atlasusr/browser/mfiascar/WjetsD3PDskimming/trunk/IsoCorrection.h
    - float El_Etcone20_pt_corrected = isoTool.correctElectronIsolation (m_el_eta->at(i), m_el_pt->at(i), m_el_Etcone20->at(i));

- $|d_0^{wrtPV}| < 0.1mm$
- W-Channel: Reject events with second lepton
  - Isolated, medium electron with $p_T > 20GeV$
  - Combined muon with $p_T > 20GeV$
- Currently no JVF Cut

2. Overlap Removal

- Lepton/Jet Overlap removal
  - Remove closest jet to a well reconstructed W/Z decay lepton if this jet is closer than $\Delta R < 0.2$
  - Remove event if jet with $p_T \geq 20$ GeV is closer than 0.6 to signal lepton
- Jet-Isolation
  - Apply no Jet Isolation, but apply uncertainties from jet/etmiss group for close-by jets

4. Jet Selection

- Collection: AntiKt jets build from topoclusters
  - Size: results from various jet sizes: 0.4 and 0.6 (add more as calibration get available)
  - Priority: 0.4 Cone Size
- Scale Definition: Use EM+JES jets with offset, and beam spot corrections until GCW and LCW get available
- Remove Jets (which are no signal leptons) with $p_T \geq 20$ GeV and
  - Medium Bad Jets Definition (= tight+ in rel15)
  - Negative Energy
- Jet threshold: Jets are considered if $p_T \geq 30$ (Baseline) and 20 GeV (Optional after careful studies)
  - both cuts should be applied through out the full analysis. Keep in mind that the 20GeV might only be doable for the 2010 data
- Rapidity Rage: $|y| < 4.5$

5. B-Jet Selection

- Collection: AntiKt jets build from topoclusters with cone-size 0.4
- Scale Definition: Use EM+JES jets with offset, and beam spot corrections
- $P_T \geq 30$ GeV (Baseline) and $P_T \geq 20$ GeV
- $|\eta| < 2.1$
- SV0 tagger: Cut at 5.85 in r16
- use BTagging Calibration Tool to reweight MC using SF provided by b-tagging group

Future changes in W/Z baseline selection

To be studied and discussed

To be studied

Electron Selection
- Study various electron isolation criteria
- Study possibility to replace tight with medium +isolation for inclusive measurements

Muon Selection
- Effect of z0 and d0 cuts on the track selection
- Effect of new isolation cut on the inclusive measurements
- Staco / Muid ?

MET Definition
- Performance of MET_RefFinal_em and MET calculation of M.Boonekamp et.al. in VB+jets

Jet/Lepton Overlap removal
- Possibility to resolve jet reconstruction with nearby leptons
- Effect of lepton/jet overlap removal for large jet multiplicities
- Jet Unfolding: Compare unfolding w.r.t. pure truth container or to truth container when applying same lepton/jet/jet removal cuts

Jets:
- Calibration for jets in W/Z+jets events (flavor effects, this is multiplicity dependent)
- Close-by jets in W/Z+jets topology with more than 1 jet.
- Similar effects in B-jets

JVF Cut
- Study impact of new JVF Cut in rel.16
- Study Impact of Mc and data discrepancies
- Is it possible to correct for pile-up during unfolding?

B-Jet Definition
- Study sensible definition of truth b-jet
- Study Calibration constants for b-tagging between 2.1 and 2.5 (get in contact with C. Weiser)

Study S/B ratio of Di-Boson analyses
- Loosen p_T cut on second lepton
- Use medium electron with/without isolation