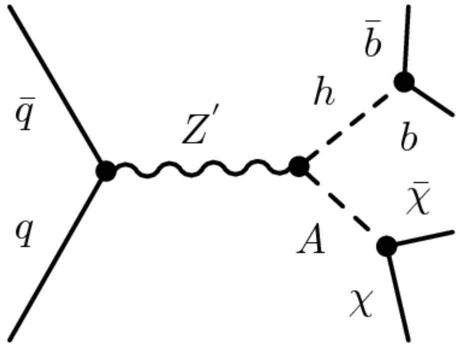
Calculating theoretical uncertainties for MonoH(bb) analysis

Skorda Eleni
Supervisor: Ruth Pottgen

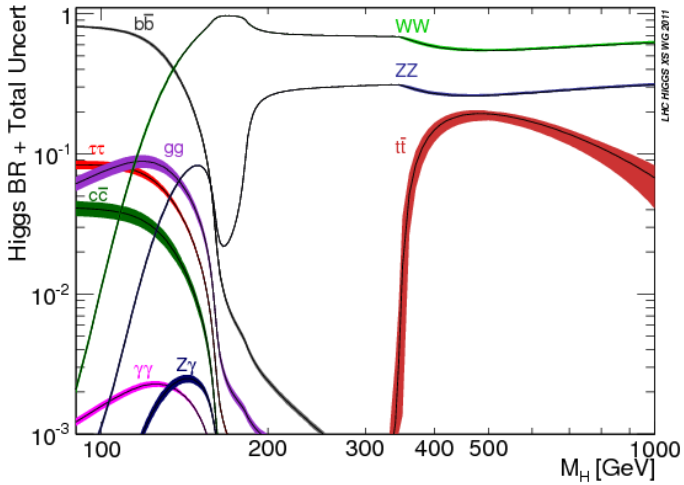
Introduction

- DM particles escape detection producing E_T^{miss}, → additional visible object (jet, γ, W, Z, h)
- The Higgs doesn't originate from initial state radiation → direct probe of the interaction with the dark matter.
- h→bb channel → high BR for m_h of 125 GeV

Final state: E_T^{miss} + h(bb)

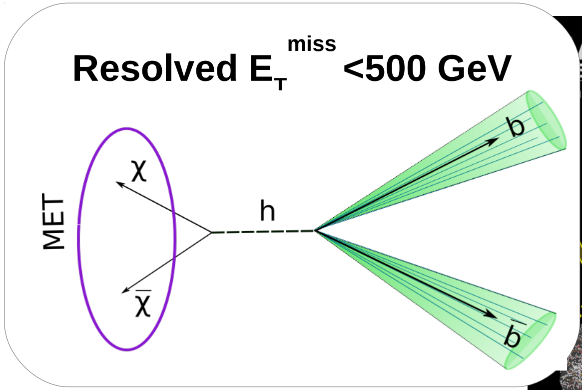


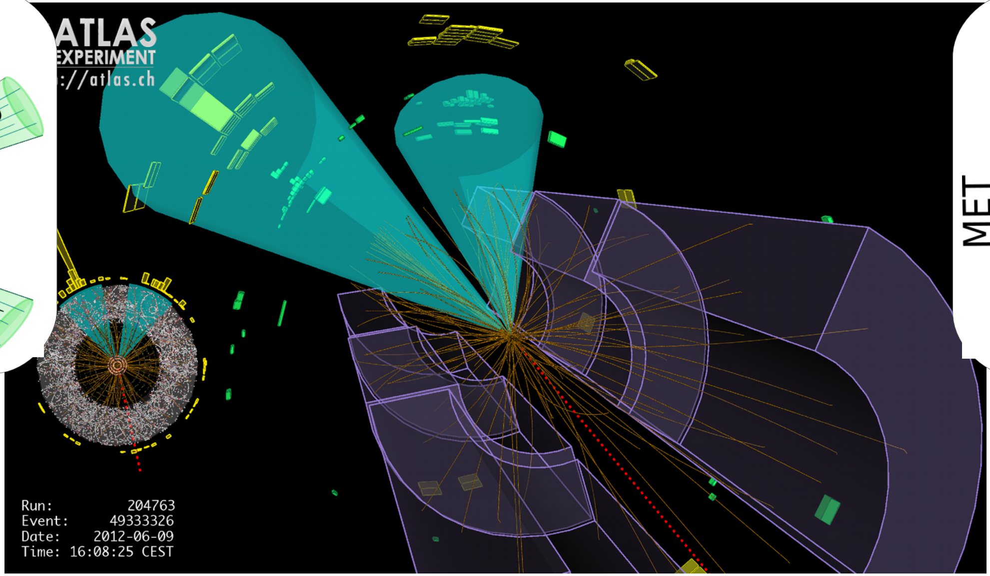
https://cds.cern.ch/record/2632344/files/ATLAS-CONF-2018-039.pdf

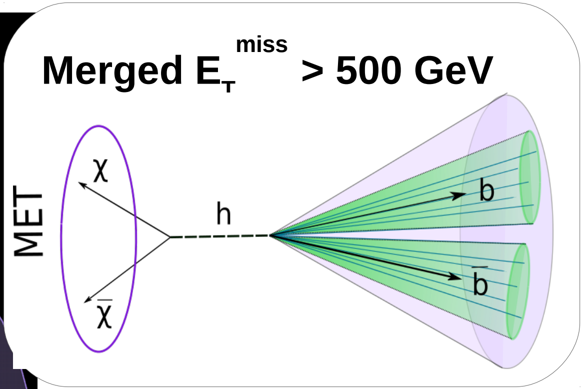


https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsTheoryPlots

Event topologies





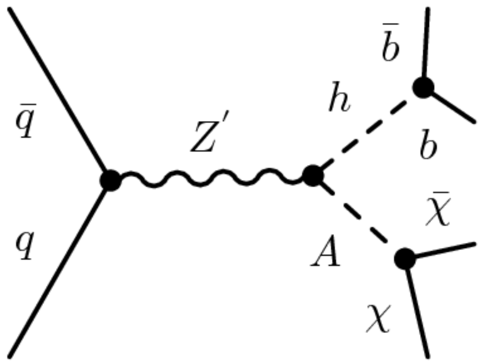


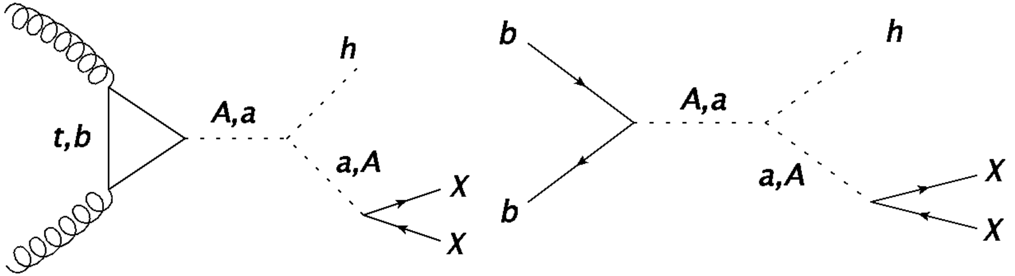
https://atlas.cern/updates/atlas-blog/what-happens-when-energy-goes-missing

2 reconstructed jets with R=0.4 (small-R jets)

Highly boosted Higgs, one jet with R=1 (large-R jet)

Signal Models





Z'2HDM: simplified DM mediator model which involves an extended two-Higgs-doublet sector (2HDM), and an additional vector mediator

- Larger mZ' and large mZ'-mA mass splittings → More boosted signature

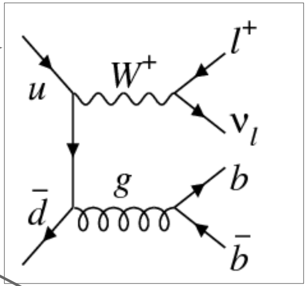
2HDMa : includes an additional pseudo-scalar mediator a

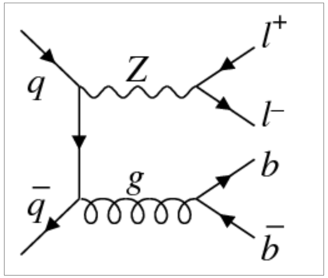
- 2 production modes gg and bb
- Higher b-jet multiplicity in bb induced signal processes
- softer MET / jet p_T spectra

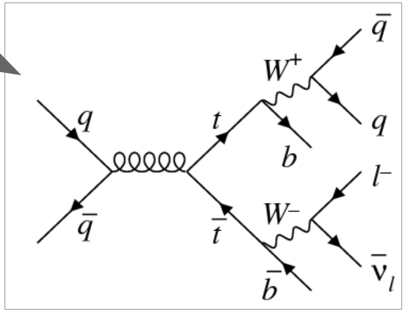
Background

Missing energy production in standard model: neutrinos by Z, W decays SM particles that decay to Z and W

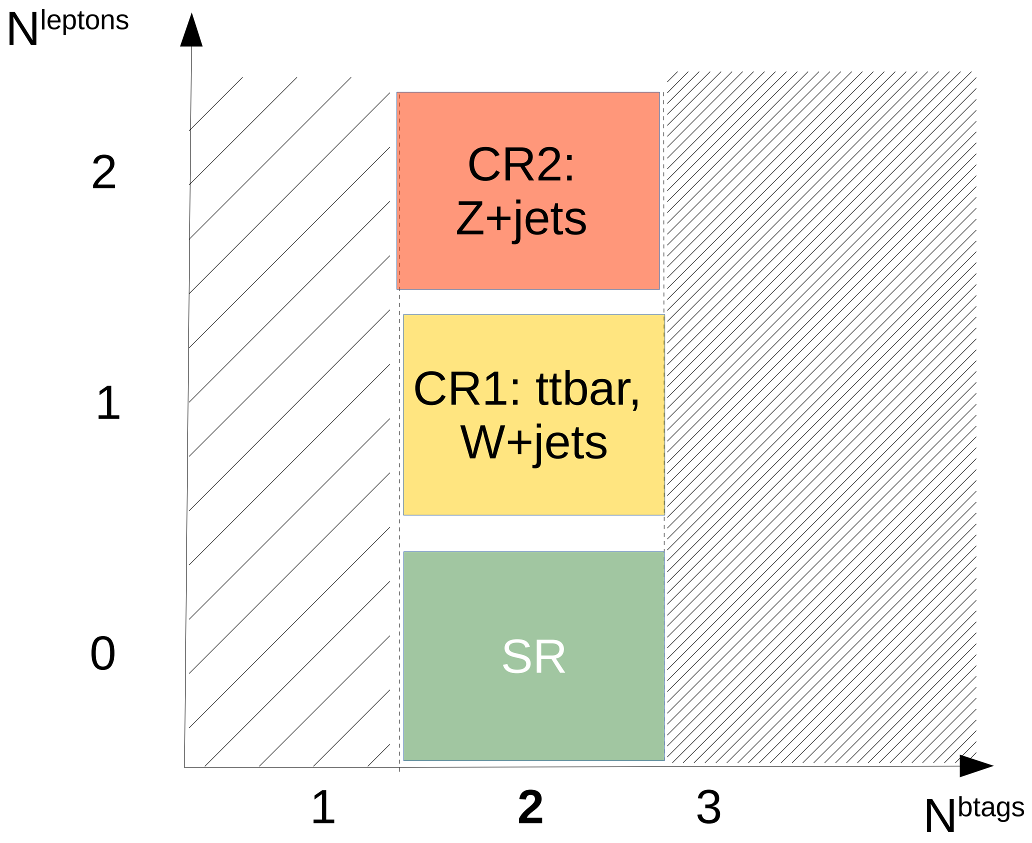
- **V+jets** SHERPA v2.2.1
- **ttbar** (PowhegBox+ Pythia8)
- Single top (PowhegBox+ Pythia8)
- Diboson (SHERPA v2.2.1)
- SM Vh(bb) (PowhegBox+ Pythia8)
- ttbar + h (PowhegBox+ Pythia8)
- ttbar + V (MadGraph5_aMC@NLO v2.3.3)







Signal and Control Regions



- Control regions are used to constrain major backgrounds
- 2 topologies for each region:
  - Resolved 150 GeV< E_T^{miss} <500 GeV
  - Merged E_T^{miss} > 500 GeV
- In each region we fit :
  - 0L region → m_{bb}
  - 1L region → muon charge
  - 2L region → yield

Event Selection

Common Selection

- lowest unprescaled E_T^{miss} trigger for SR and CR1 and lowest unprescaled single lepton triggers in CR2
- Lepton veto in SR ,1 μ for CR1 and ee μμ in CR2 with with $|m_{ll} - m_Z| <10$ GeV
- τ-veto
- min ΔΦ(E_T^{mis},Central + Forward jets 1,2,3) > 20°

$$m_T^{b,min/max} = \sqrt{2p_T^{b,min/max} E_T^{\text{miss}}(1 - \cos\Delta\phi(b_{min/max}, E_T^{\text{miss}}))}$$

b_{min} and b_{max} are defined as the b-jet, which is closest to $\vec{E}_T^{\text{miss}}$

Resolved selection

- $E_T^{miss} > 150$ GeV
- $E_T^{miss} < 500$ GeV
- N(central small-R jets) ≥ 2
- E_T^{miss} Significance
- $p_T(jj) > 100$ GeV if $E_T^{miss} < 350$ GeV
- $p_T(jj) > 300$ GeV if $E_T^{miss} \geq 350$ GeV
- $m_T^{b,min} > 170$ GeV
- $m_T^{b,max} > 200$ GeV

Merged Selection

- $E_T^{miss} > 500$ GeV
- N(central large-R jets) ≥ 1

Fitting

- Profile likelihood fit performed simultaneously in signal region (0l) and control regions (1,2l) to extract a possible mono-h(bb) signal
- Benchmark models : Z'2HDM, 2HDMa (ggF, bb)
- Set constraints on the models

Blind: mjj (mJ) 70-140 GeV

Nuisance parameters (NP) θ: dependence of signal and background predictions on the systematic uncertainties

→ Detector related

Systematic for reconstructed objects :

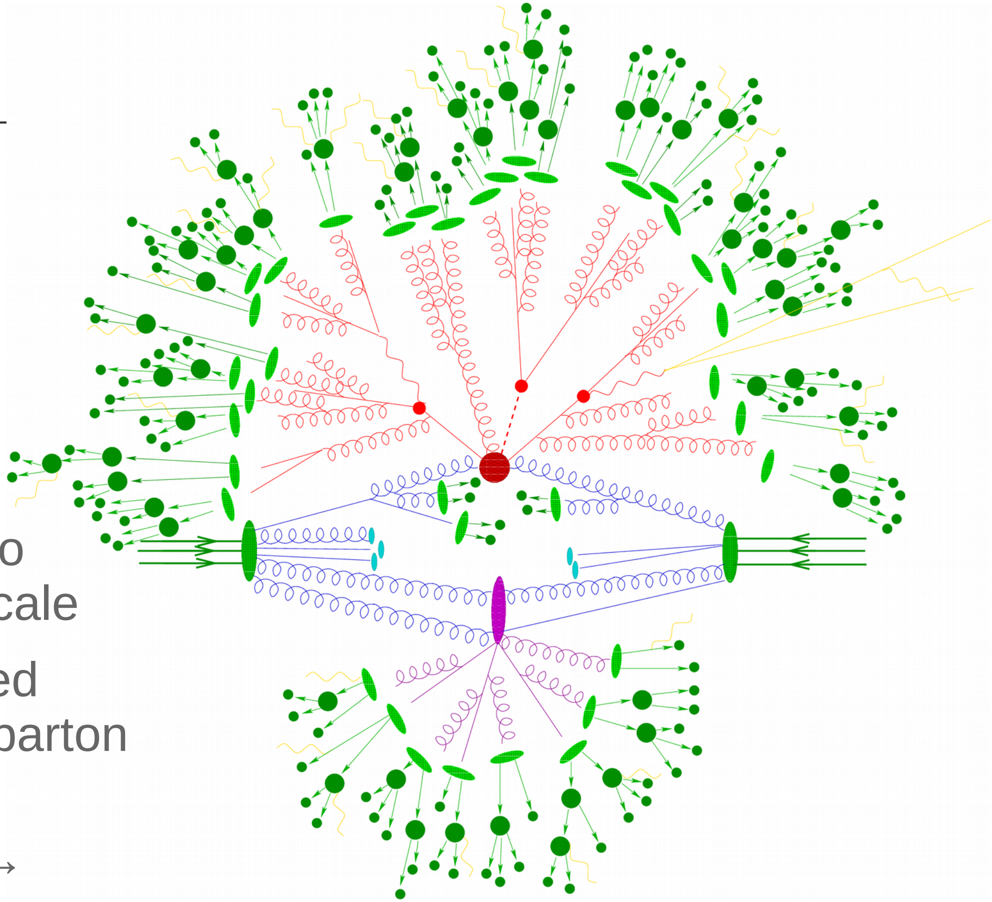
- lepton(e,μ,τ)
- jets→ scale and resolution of jet mass and energy
- b-tagging efficiency

→ Theoretical

Theoretical uncertainties

Sources of uncertainties

- **Missing higher order terms** → variation of renormalization factorization scale
- **Choice of PDFs and a_s**
- **Multijet merging** → for samples generated by merging matrix elements (ME) corresponding to different multiplicities → variation of merging scale
- **Matching uncertainties**: for samples generated using a NLO matrix element and matched to aparton shower → compare different generators
- **Parton shower/hadronization calculations** → compare different generators



http://inspirehep.net/record/1328513/plots

Uncertainties considered in the fit

- Uncertainties on the cross-section used for the normalisation of the MC samples
- Uncertainty on the flavour composition (W,Z+ heavy flavor components – bb, cc, bc, bl)
- Shape uncertainties: compare the shape of the variable between the nominal and alternative MC samples
- Relative acceptance uncertainties: theory uncertainties can alter the shape of observables : eg E_T^{miss} , used to separate regions → change of the acceptances → relative acceptance difference

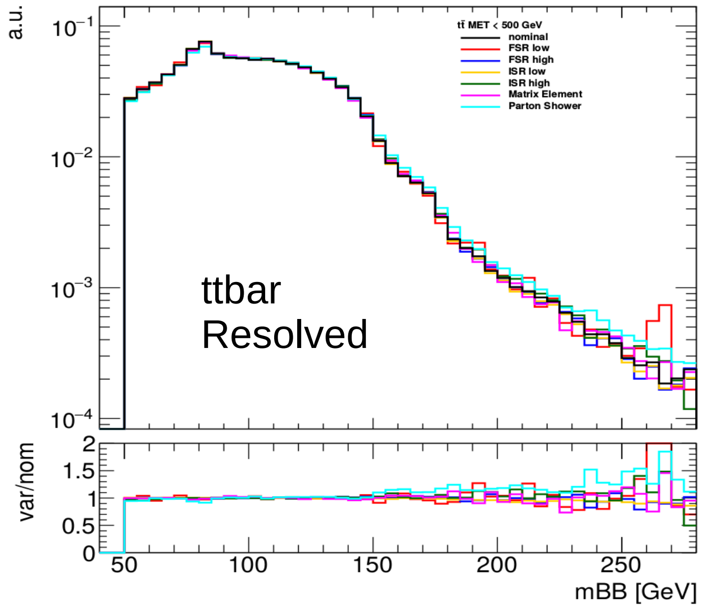
How are they estimated

The full monoHbb selection is applied at the truth level

- For each theory parameter variation (M total number of variations) and for two regions A,B the uncertainty related to relative acceptance difference is

$$\sigma_{\text{acceptance}} = \sqrt{\sum_{i}^{M} \left(1 - \frac{n_A^{\text{var }i} / n_B^{\text{var }i}}{n_A^{\text{nom}} / n_B^{\text{nom}}}\right)^2}$$

- For the shape uncertainties the ratio of the histograms between each variation over the nominal is fit



How is it done ?

- Produce truth ntuples with some loose p reselection for all regions
- Apply the same selection as we have for the reconstructed objects
- Compare the results between Reco and truth
- Calculate uncertainties and figure out how the can be implemented in the fit

Courses

List of courses

Past Year:

- Learning and teaching in higher education AKA “LATHE” (4.5 credits)
  Note: long weighting list and compulsory
- Scientific Computing with Fortran and python AKA “Scifopy” (7.5 credits).
  Note: very organised and useful, organised by COMPUTE

Coming up :

- Quantum Field Theory (7.5 credits) (available for master and PhD students)
- Jupiter notebooks (I always wanted to learn about that but never look it up)
- CERN school of computing
  23 Aug - 5 Sep 2020, Cracow

Useful links :

- CSC home page: registration for main school not open yet
- COMPUTE : courses coming up all the time :)
- LATHE: Registration for the period 29 january – 28 February 2020 closes on 16th of December

Back-up Slides

Object definitions

- **electron**: loose ID and isolation , $|\eta| < 2.47$, $p_T > 7$GeV
- **muons** :loose selection and isolation for baseline ($|\eta| < 2.7$) medium, tight for signal ($|\eta| < 2.5$), $p_T > 7$GeV
- **taus**: BDT wp: loose, ($|\eta| < 2.5$), $p_T > 20$ GeV
- **small-R jets**: anti-kt , EMTopo, R=0.4
- **large-R jets**: anti-kt , LCTopo, R=1, $p_T > 200$ GeV
- **b- tagging** : AntiKt4EMTopo/AntiKtVR30Rmax4Rmin02 , MV2c10, Eff = 77
- **E_T^{miss}** : Calo-based, Soft term:track-based, tight operating point

- **VR track Jets**

$$\mathrm{R}(p_\mathrm{T}) = \frac{30\,\mathrm{GeV}}{p_\mathrm{T}}$$

R in (0.02, 0.4)

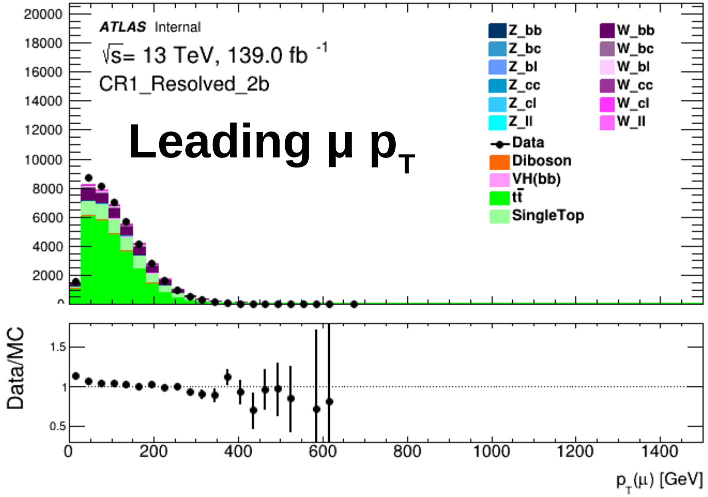
- **E_T^{miss} significance**

$$\mathcal{S} = \frac{\left|\overrightarrow{E_\mathrm{T}^\mathrm{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}.$$

Total variances in the longitudinal and transverse directions to E_T^{miss}

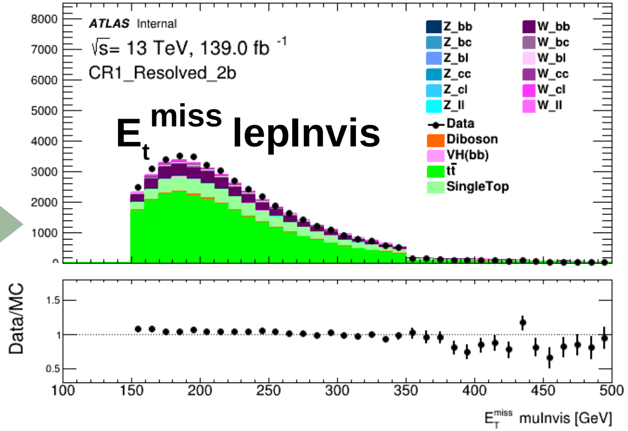
is the correlation factor of the longitudinal L and transverse T measurements

E_T^{miss} definition for CR1,CR2



Leading μ p_T

Included in the E_T^{miss} as invisible



E_t^{miss} lepInvis

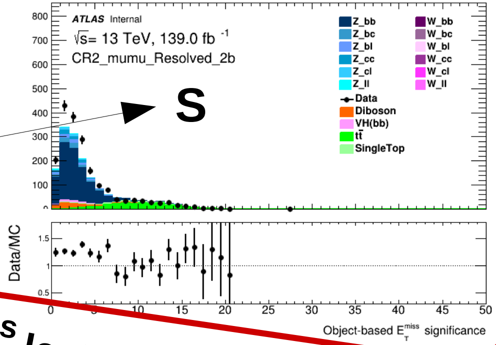
Used to seperate merged and resolved

Same for CR 2 : instead using $p_T(ee/\mu\mu)$

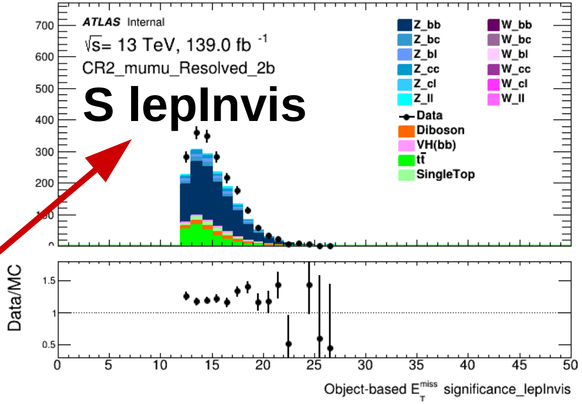
- E_T^{miss} significance

$$S = \frac{\left|\overrightarrow{E_T^{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}.$$

"E_T^{miss}" → **S**



E_T^{miss} lepInvis



S lepInvis

Extended list of courses

- Statistical tools in Astrophysics – 7.5 credits
- Scientific Writing – 1.5 credits:
  - 3 Days Long
  - All the work is done during these days
  - Very useful
- Geant 4 tutorial – 3 credits:
  - 1 week of lectures and Hand On
  - More than 1 week project
  - Very useful but time consuming
- Detector school in Copenhagen/Helsinki – 10 credits
- Phenomenology – 7.5 credits
  - Full semester
  - Lots of homework and studying
  - Very useful → in understanding concepts around MC processes