The Road to Discovery

Exploring the Energy Frontier with Hadron Colliders

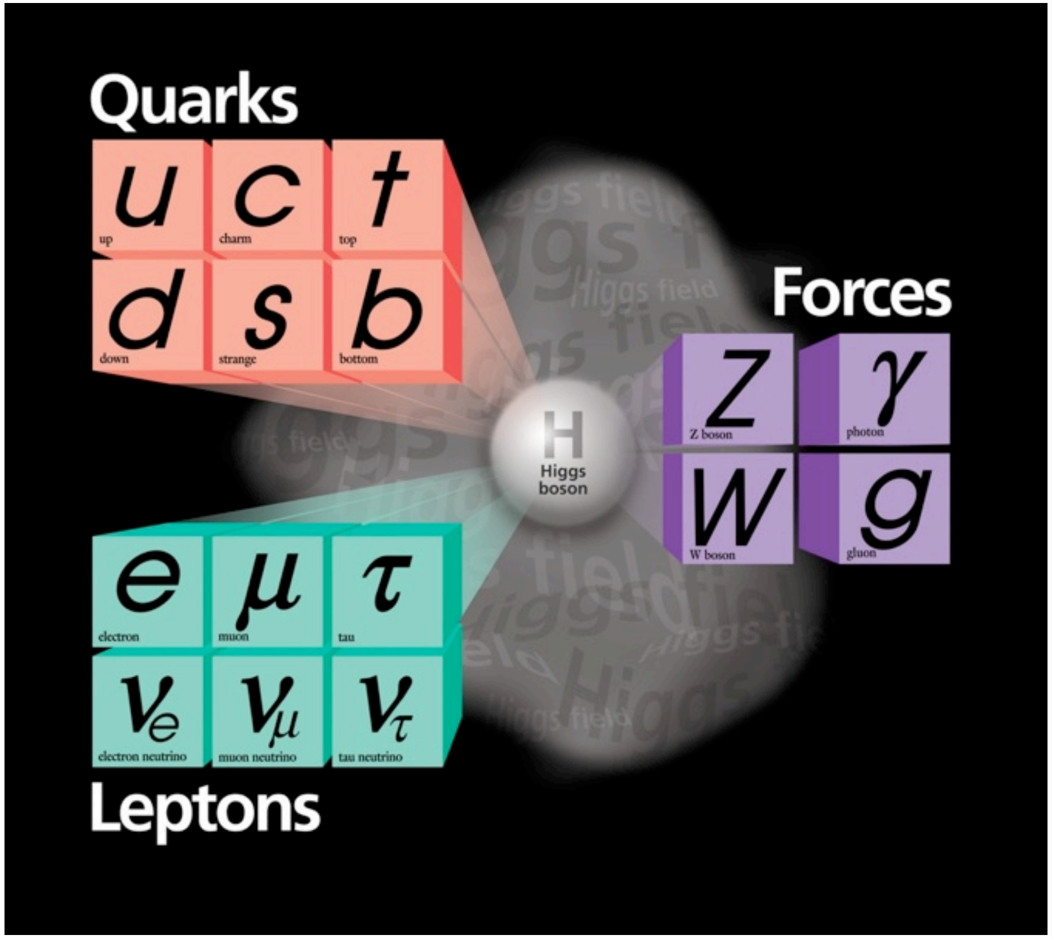
Gustaaf Brooijmans

Columbia University
in the City of New York

Goals

- For theorists (and experimenters):
  - What is it these experimenters *do*?
    - (Except for using fancy equipment to build fancy detectors...)
  - How come it takes them forever to release a result?
  - Why can't they just give us the 4-vectors as they record the data?
  - What can('t) they look for?
- For experimenters:
  - Why does it take so many layers of review to get a result out?   My analysis seems pretty simple at first sight....

HEP in 2013



CKM elements:

Observable	Central ± 1 σ
$\|V_{ud}\|$	0.97430 [+0.00019 -0.00019]
$\|V_{us}\|$	0.22521 [+0.00082 -0.00082]
$\|V_{ub}\|$	0.00350 [+0.00015 -0.00014]
$\|V_{cb}\|$	0.04117 [+0.00038 -0.00115]
$\|V_{ud}\|$ (meas. not in the fit)	0.97444 [+0.00028 -0.00028]
$\|V_{us}\|$ (meas. not in the fit)	0.2257 [+0.0011 -0.0011]
$\|V_{ub}\|$ (meas. not in the fit)	0.00350 [+0.00015 -0.00016]
$\|V_{cb}\|$ (meas. not in the fit)	0.04399 [+0.00069 -0.00397]
$\|V_{cd}\|$	0.22508 [+0.00082 -0.00082]
$\|V_{cs}\|$	0.97347 [+0.00019 -0.00019]
$\|V_{td}\|$	0.00859 [+0.00027 -0.00029]
$\|V_{ts}\|$	0.04041 [+0.00038 -0.00115]
$\|V_{tb}\|$	0.999146 [+0.000047 -0.000016]

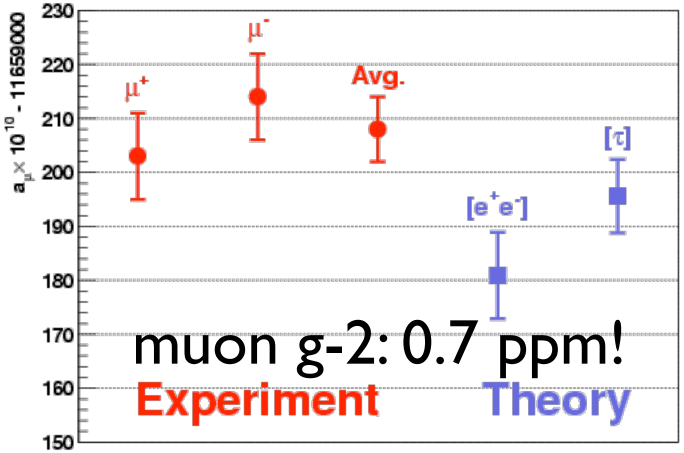
In Words

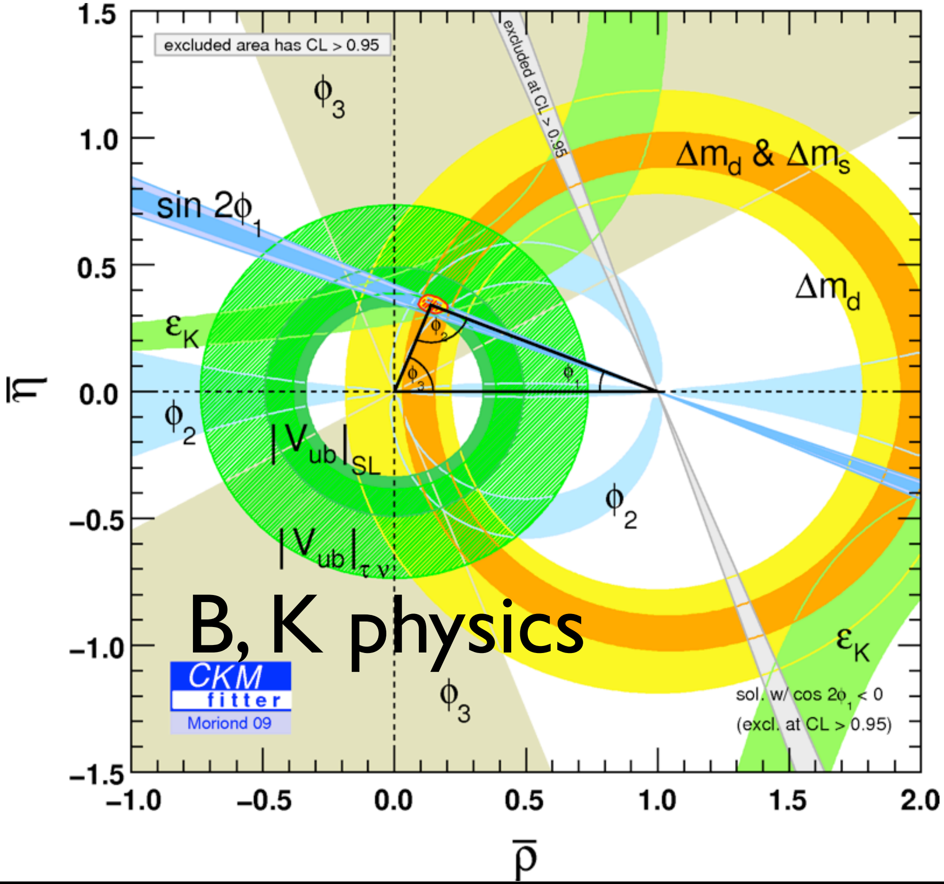
- Matter is built of spin 1/2 particles that interact by exchanging 3 different kinds of spin 1 particles corresponding to 3 different (gauge) interactions
- There appear to be 3 generations of matter particles
- The 4 different matter particles in each generation carry different combinations of quantized charges characterizing their couplings to the interaction bosons
- The matter fermions and the weak bosons have “mass”
- Gravitation is presumably mediated by spin 2 gravitons
- Gravitation is extremely weak for typical particle masses
- There appear to be 3 macroscopic dimensions

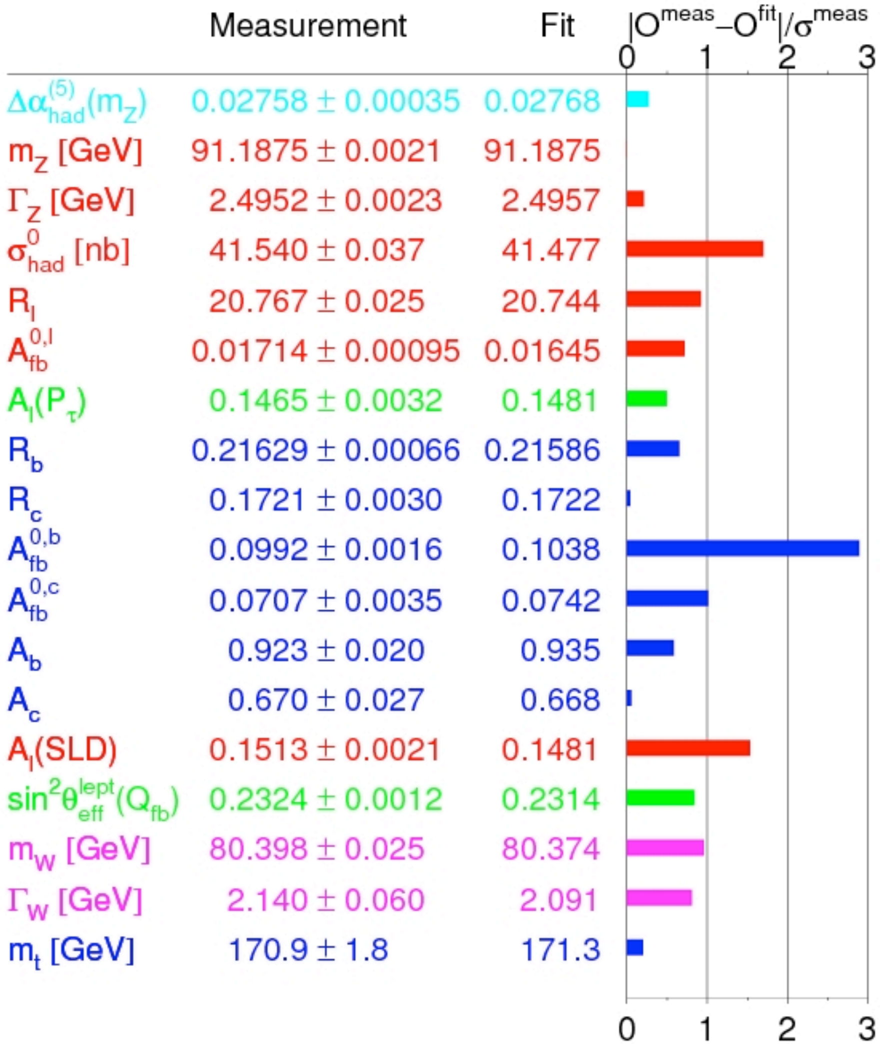
About the Standard Model

- It's a theory of interactions:
  - Properties of fermions are inputs
  - Properties of interaction bosons in terms of couplings, propagations, masses are linked:
    - Measuring a few allows us to predict the rest, then measure and compare with expectation
- It's remarkably successful:
  - Predictions verified to be correct at sometimes incredible levels of precision
  - After ~40 years, still no serious cracks

Precision Results







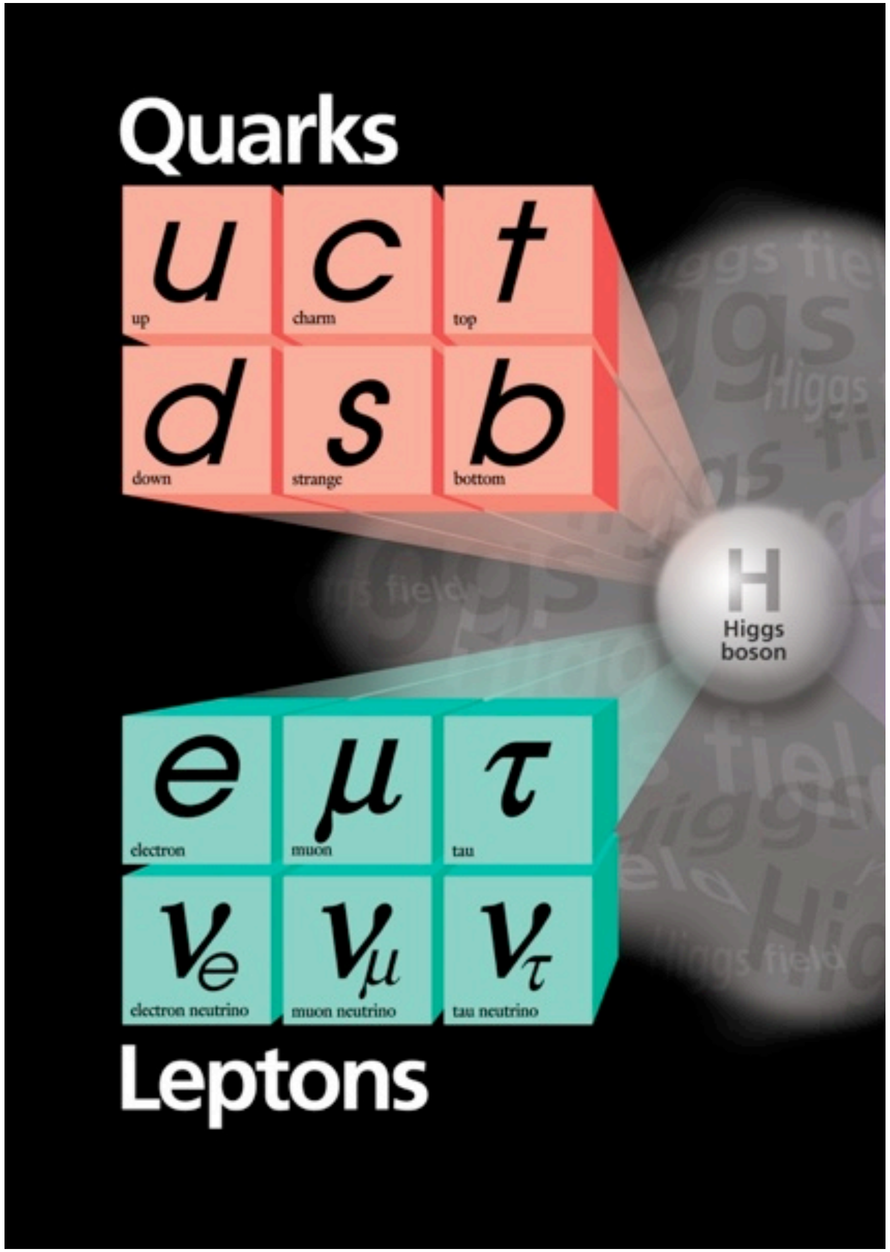
LEP, SLD & Tevatron

Many Fundamental Questions

- What exactly *is* spin? Or color? Or electric charge? Why are they quantized?
- Are there only 3 generations? If so, why?
- Why are there e.g. no neutral, colored fermions?
- What is mass? Why are particles so light?
- Is there a link between particle and nucleon masses?
- How does all of this reconcile with gravitation? How many space-time dimensions are there really?
- …

Lacking in the Standard Model

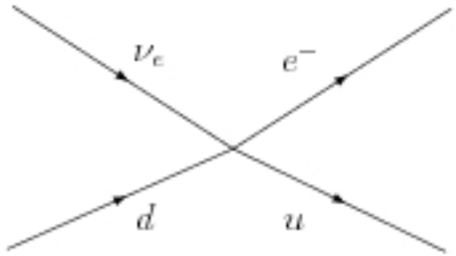
- Clear structure in fermionic sector unexplained
  - Evidence of some selective principle (why are there no neutral colored fermions?)
  - Proton stability, running of couplings suggestive of at least one other scale relevant to SM particles, ~10^{15} GeV
    - Either fine-tuning, or a closer scale



The Old Plot

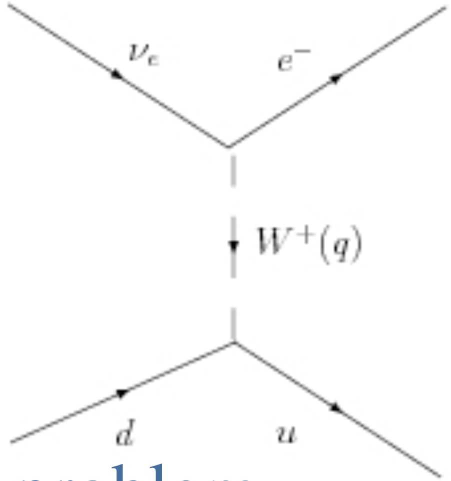
Vector Boson Scattering

- There is in fact one known problem with the Standard Model (+ a second, related, lesser one):
    - If we collide W's or Z's (not so easy...), the scattering cross-section grows with the center of mass energy, and gets out of control (violates unitarity) at about 1.7 TeV: $\sigma(WW \rightarrow WW) \sim s$
- This is similar to "low" energy neutrino scattering:
    - If $q^2 << (M_W)^2$, looks like a "contact interaction", and cross-section grows with center of mass energy: $\sigma \sim s$



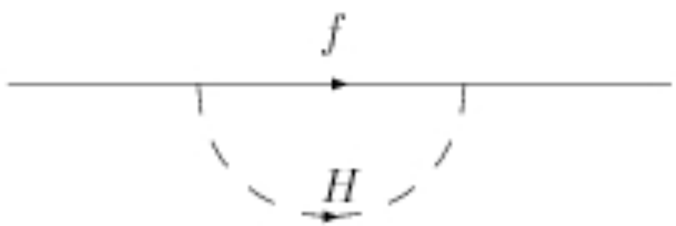
Vector Boson Scattering

- There was in fact one known problem with the Standard Model (+ a second, related, lesser one):
  - If we collide W's or Z's (not so easy...), the scattering cross-section grows with the center of mass energy, and gets out of control (violates unitarity) at about 1.7 TeV: $\sigma(WW \rightarrow WW) \sim s$
- This is similar to "low" energy neutrino scattering:
  - If $q^2 << (M_W)^2$, looks like a "contact interaction", and cross-section grows with center of mass energy: $\sigma \sim s$
  - But when $q^2 \approx (M_W)^2$, W-boson propagation becomes visible, and "cures" this problem



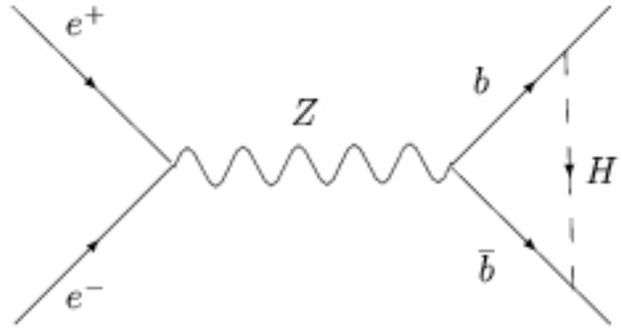
The Higgs Boson

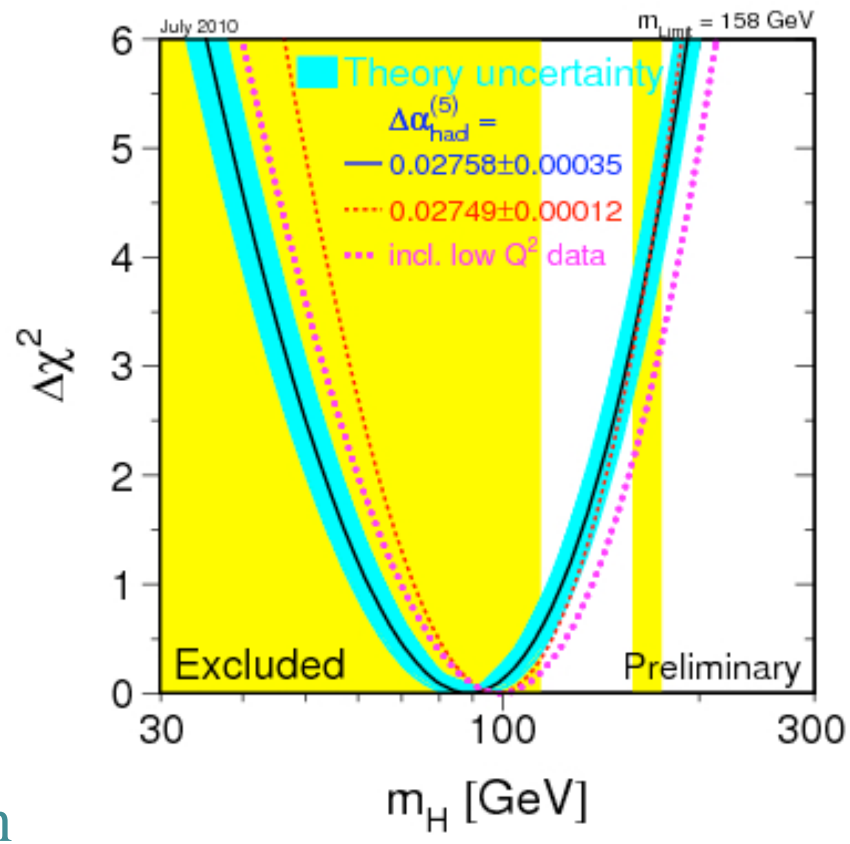
- One way to solve this, is to introduce a massive, spinless particle (of mass < ~1 TeV)
  - Couplings to W and Z are fixed, quantum numbers are known...
  - .... to be those of the vacuum
  - Its mass is unknown, and its couplings to the fermions are unknown.... well, maybe
    - Fermions can acquire mass by coupling to this Higgs boson, so their couplings could be proportional to their masses. This is called the "Standard Model Higgs"



Precision Measurements

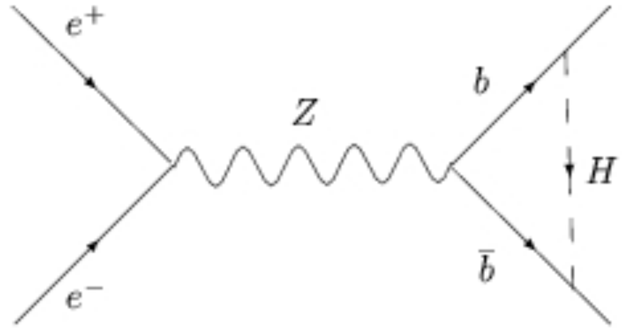
- In fact, we were able to say something about the standard model Higgs mass
  - If the fermions get their masses from the Higgs, we know all couplings and can infer the Higgs mass from precision measurements
  - Result is very sensitive to measured top quark, W boson masses
    - Really wants a “light” Higgs boson

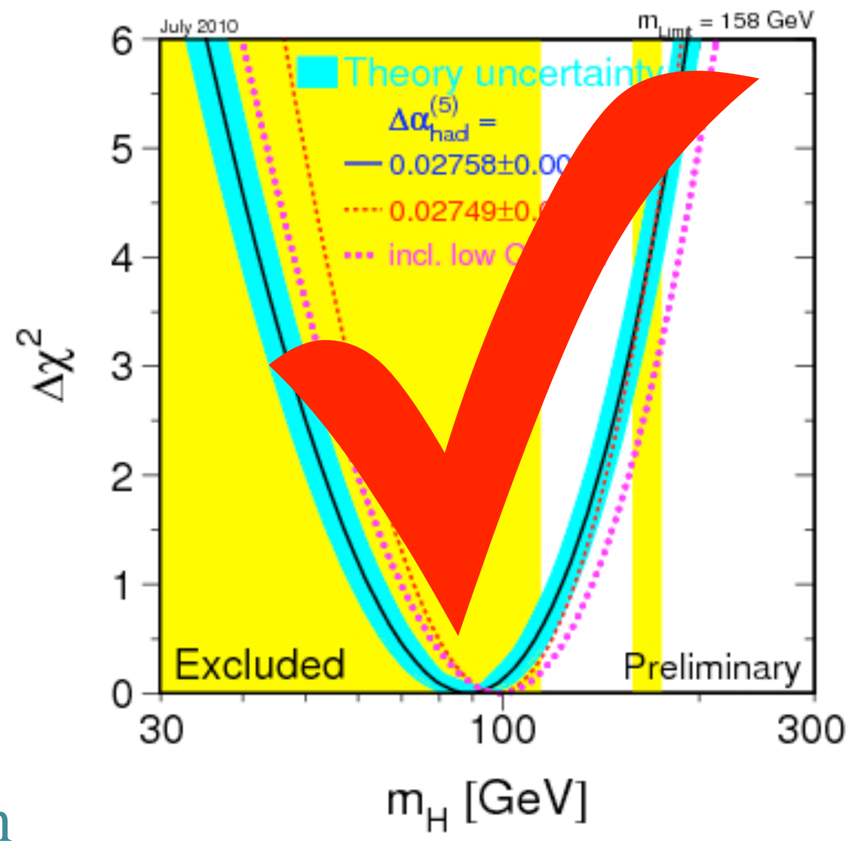




Precision Measurements

- In fact, we were able to say something about the standard model Higgs mass
  - If the fermions get their masses from the Higgs, we know all couplings and can infer the Higgs mass from precision measurements
  - Result is very sensitive to measured top quark, W boson masses
    - Really wants a “light” Higgs boson



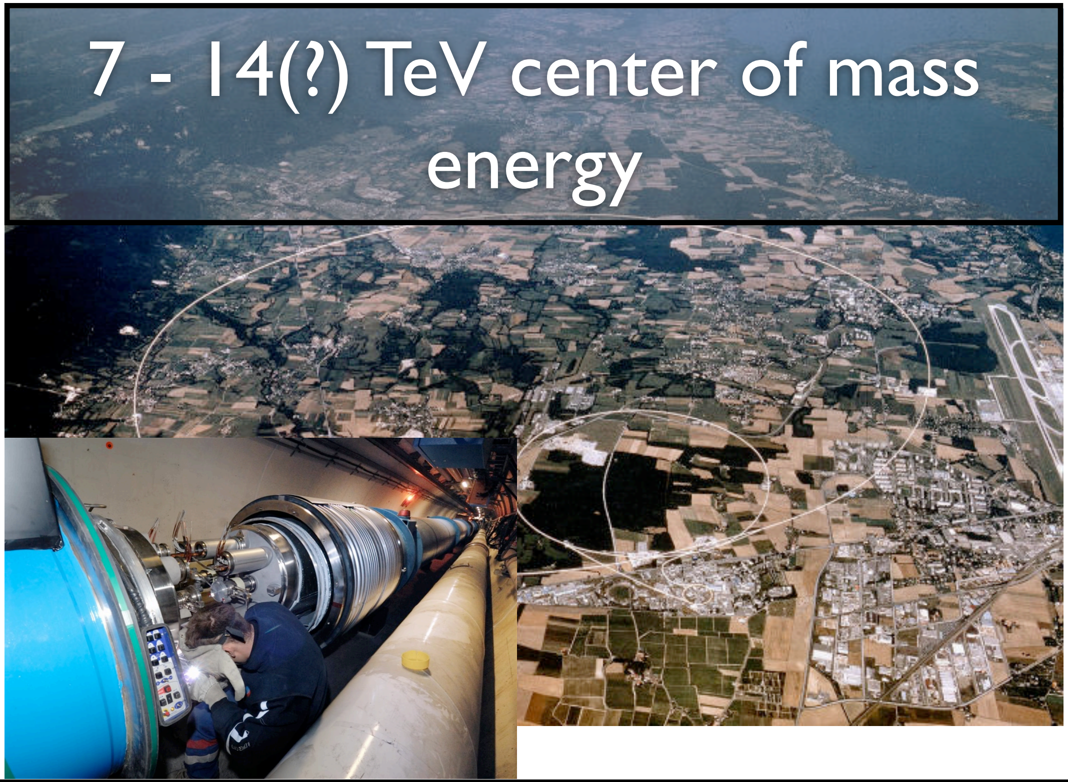


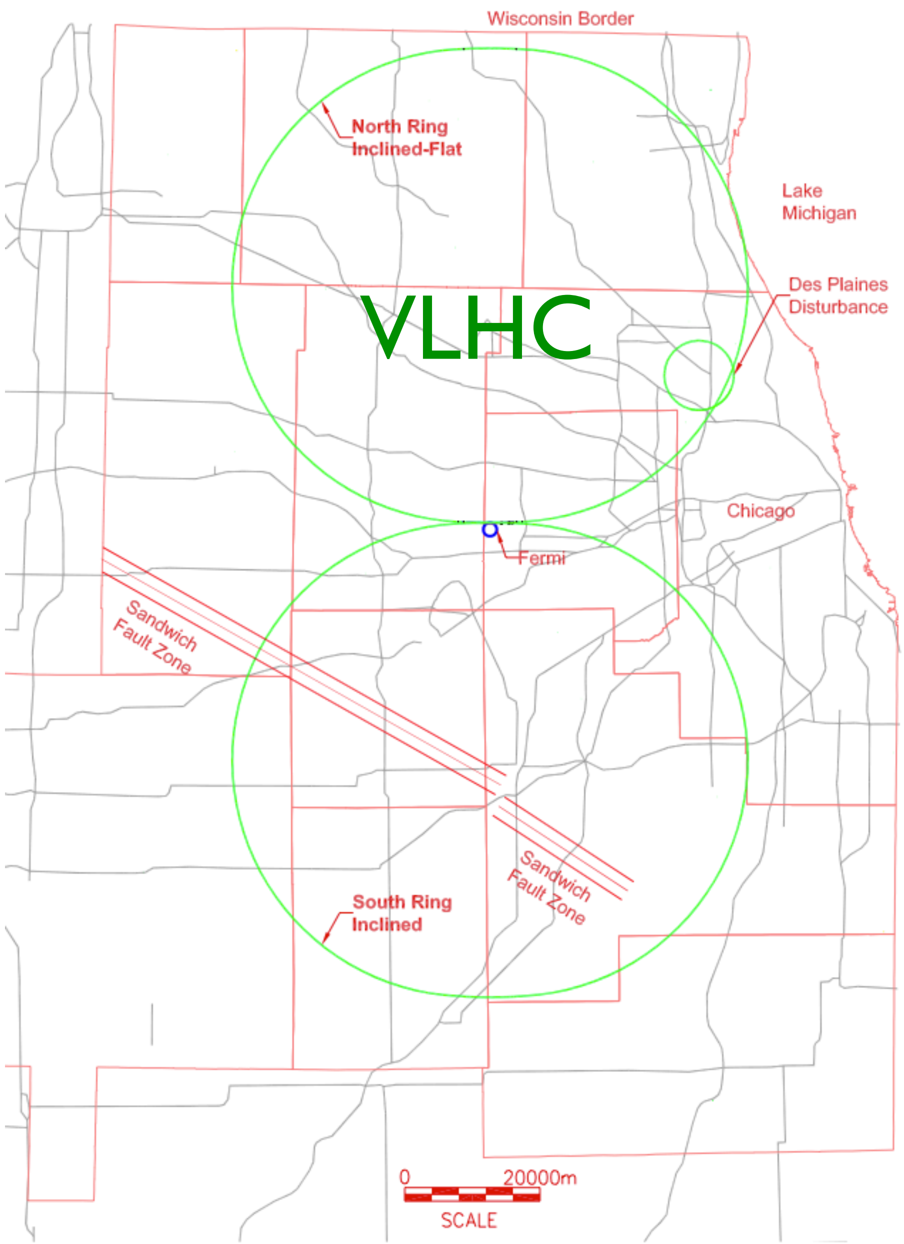
The Tools

Colliders

- Currently, hadron colliders:
  - High energy implies probing of short distances, and production of other, massive particles

LHC

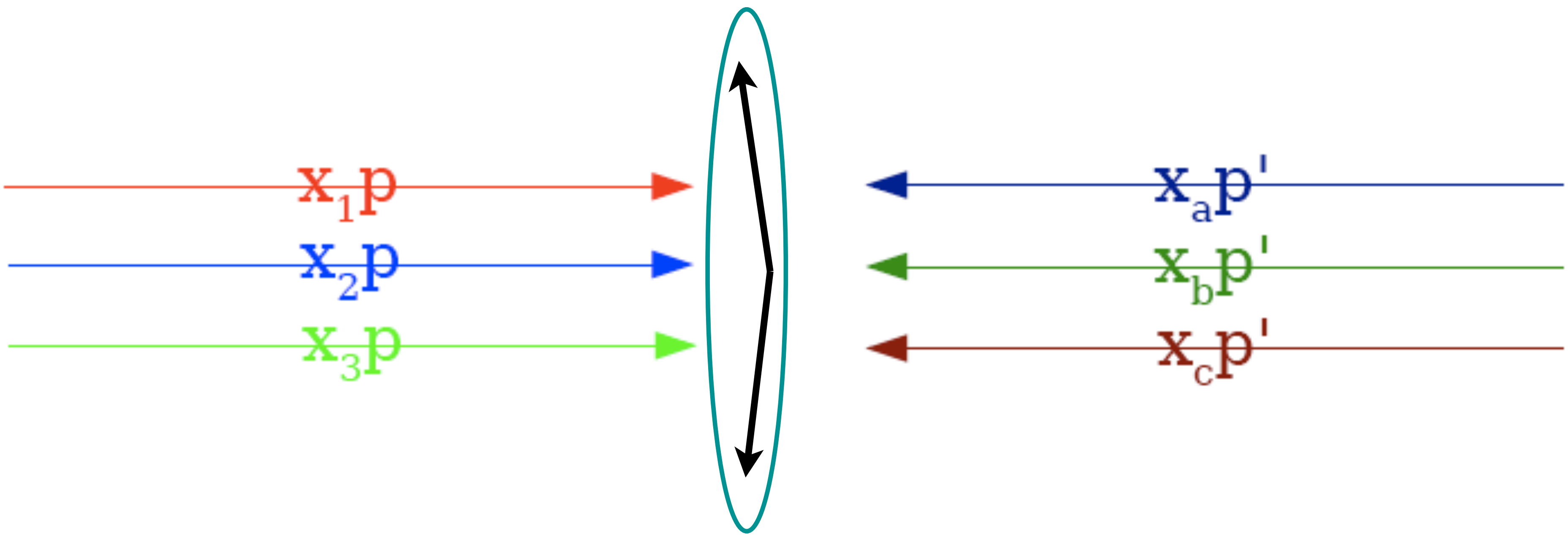




Hadron Colliders

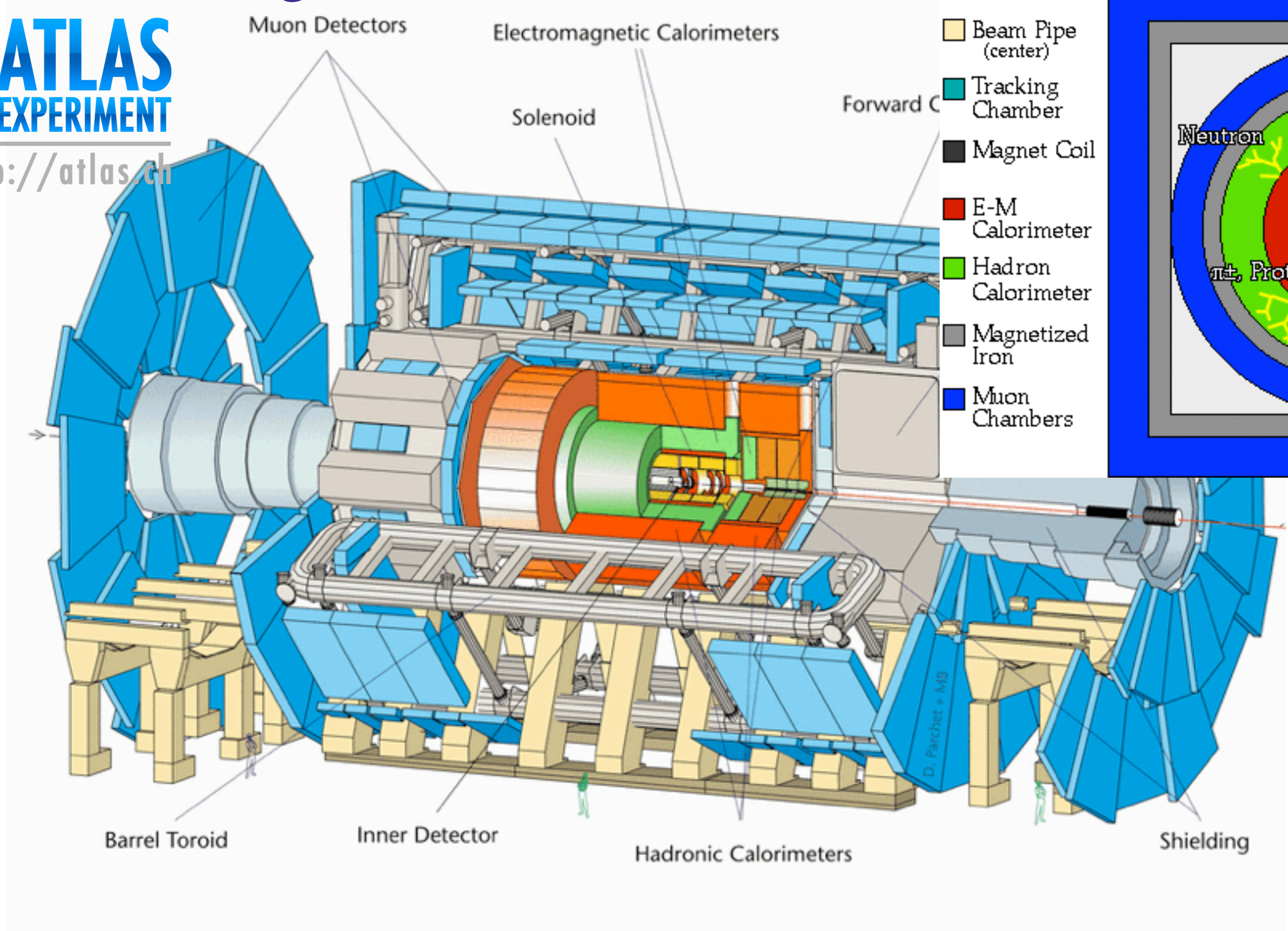
- Incoming longitudinal momentum not known:
  - "Hard interaction" is between one of the quarks and/or gluons from each proton, other quarks/gluons are "spectators"
- Longitudinal boost "flattens" event to a pancake

➡ We usually work in the plane transverse to the beam

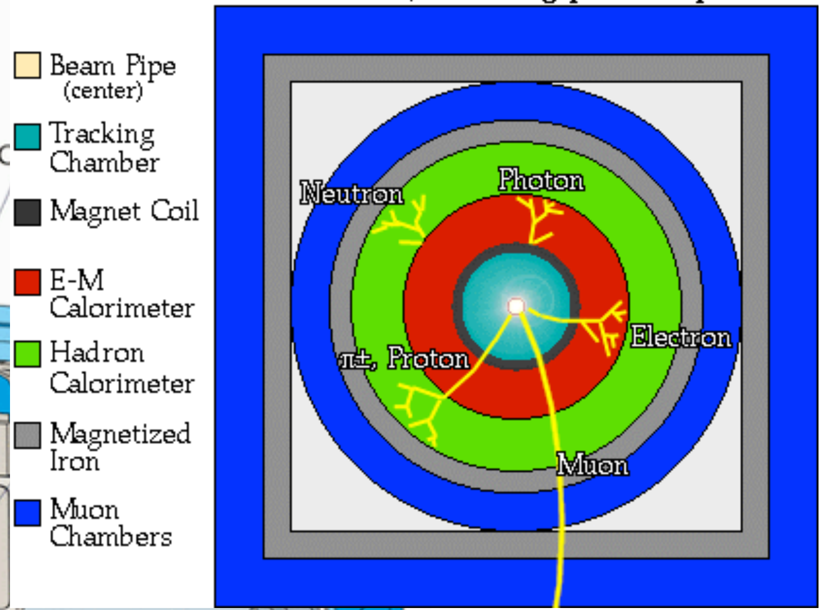


ATLAS

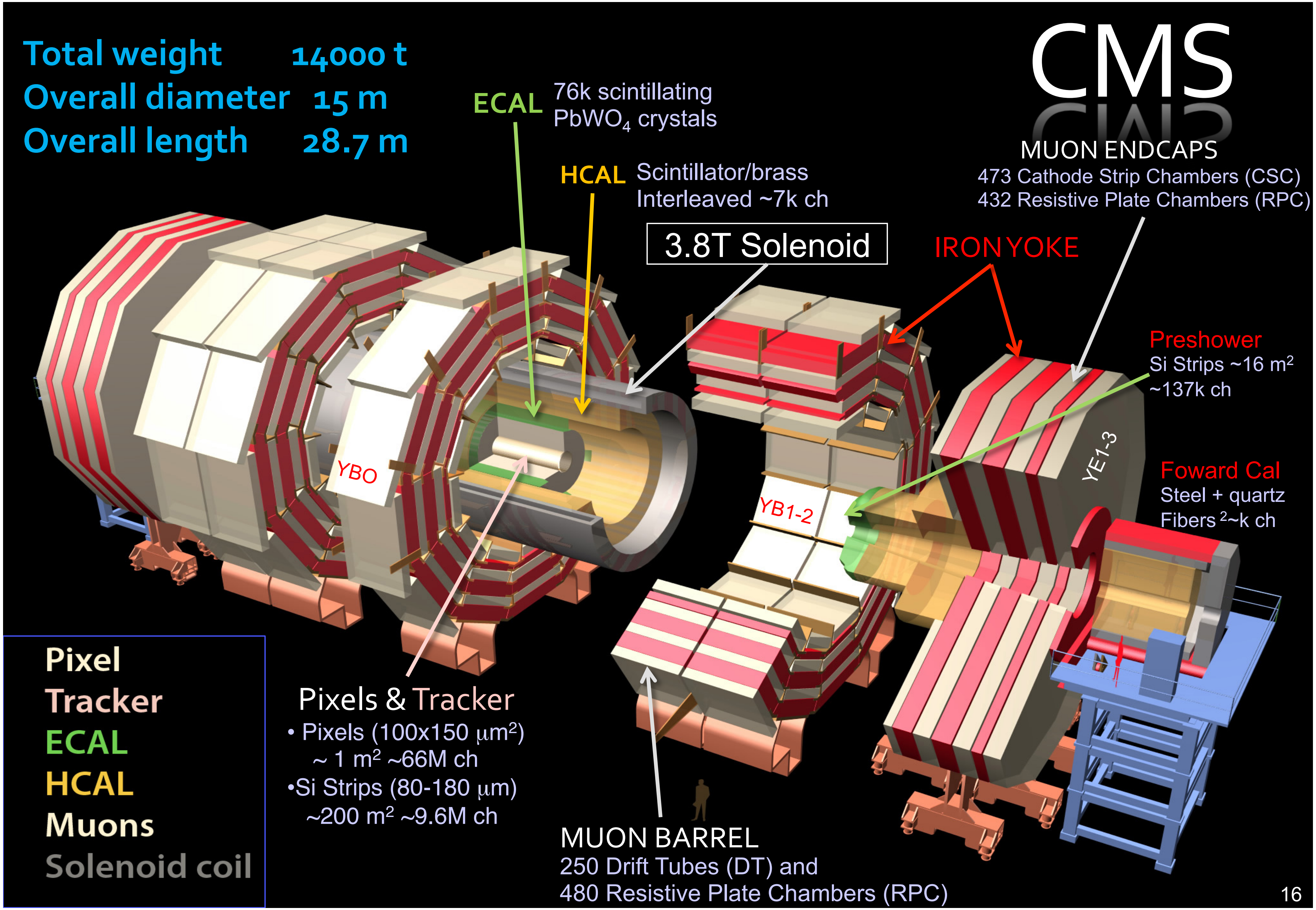
- Make best possible measurement of all particles coming out of collisions



A detector cross-section, showing particle paths



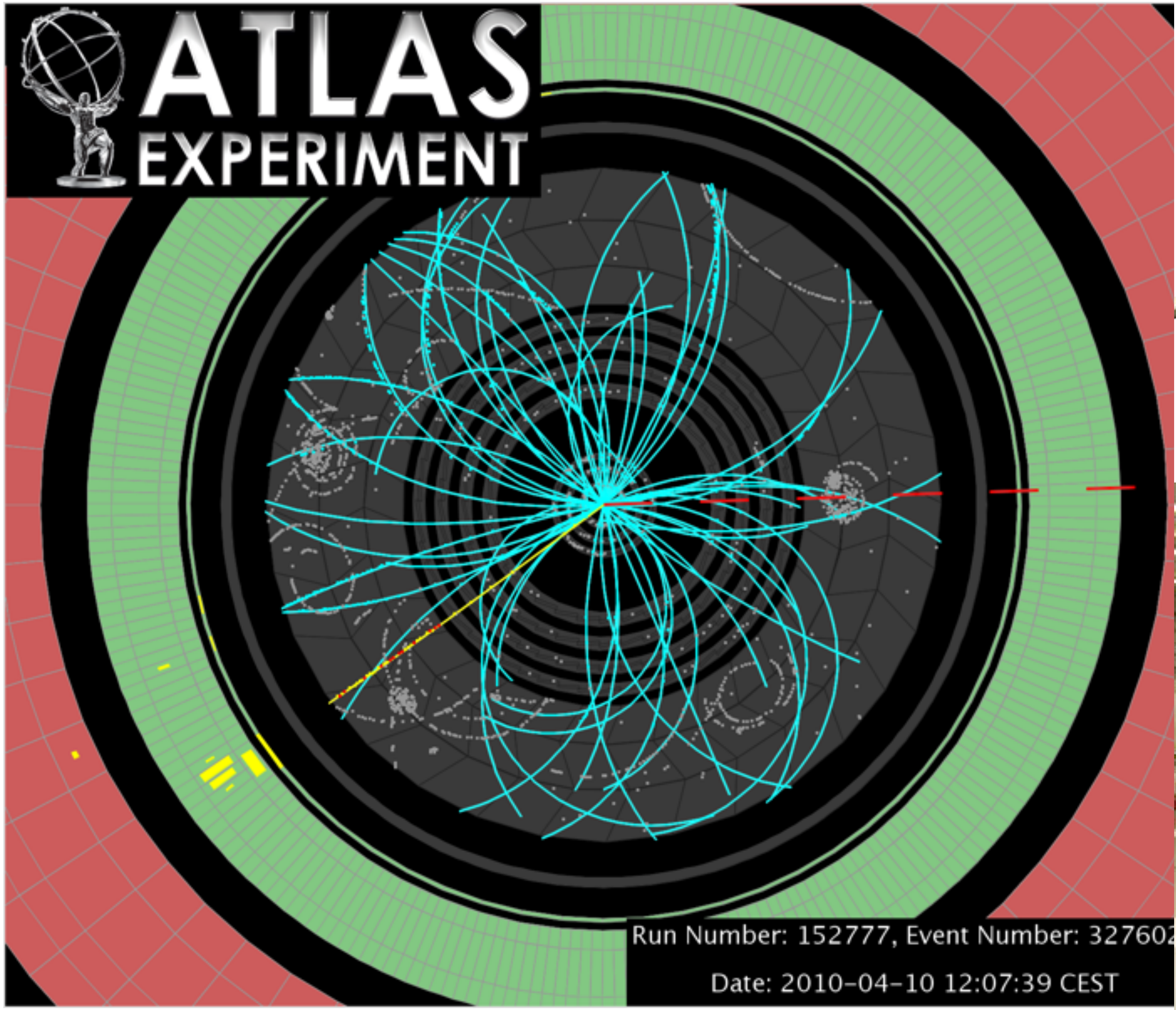
CMS

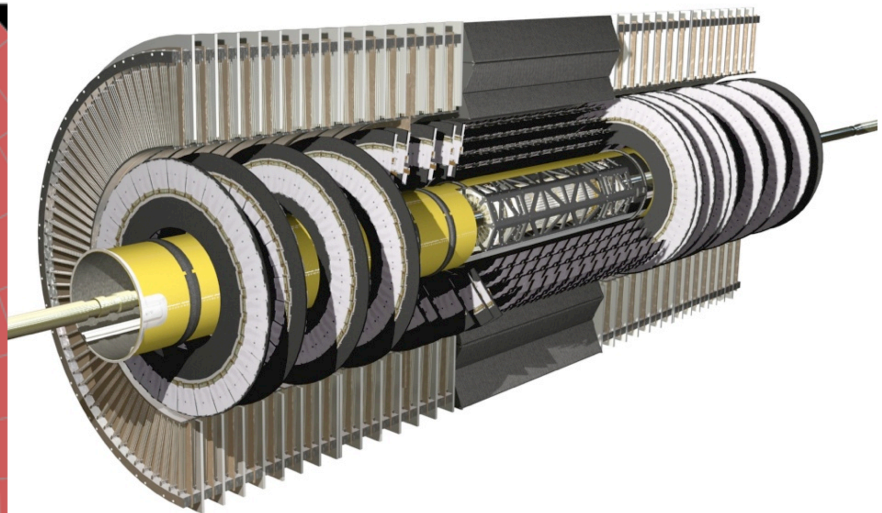


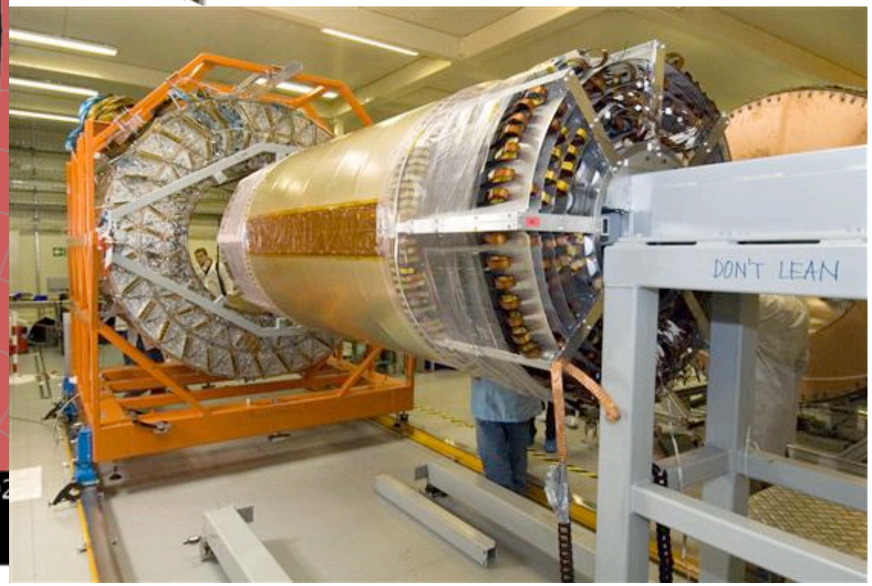
Total weight 14000 t
Overall diameter 15 m
Overall length 28.7 m
CMS
ECAL 76k scintillating
PbWO4 crystals
HCAL Scintillator/brass
Interleaved ~7k ch
3.8T Solenoid
IRON YOKE
MUON ENDCAPS
473 Cathode Strip Chambers (CSC)
432 Resistive Plate Chambers (RPC)
Preshower
Si Strips ~16 m2
~137k ch
Foward Cal
Steel + quartz
Fibers 2~k ch
YBO
YB1-2
YE1-3
Pixel
Tracker
ECAL
HCAL
Muons
Solenoid coil
Pixels & Tracker
• Pixels (100x150 μm2)
~ 1 m2 ~66M ch
•Si Strips (80-180 μm)
~200 m2 ~9.6M ch
MUON BARREL
250 Drift Tubes (DT) and
480 Resistive Plate Chambers (RPC)
16

Tracking

- Combination of pixels, silicon strips ("SCT") and straw tube transition radiation tracker (TRT)

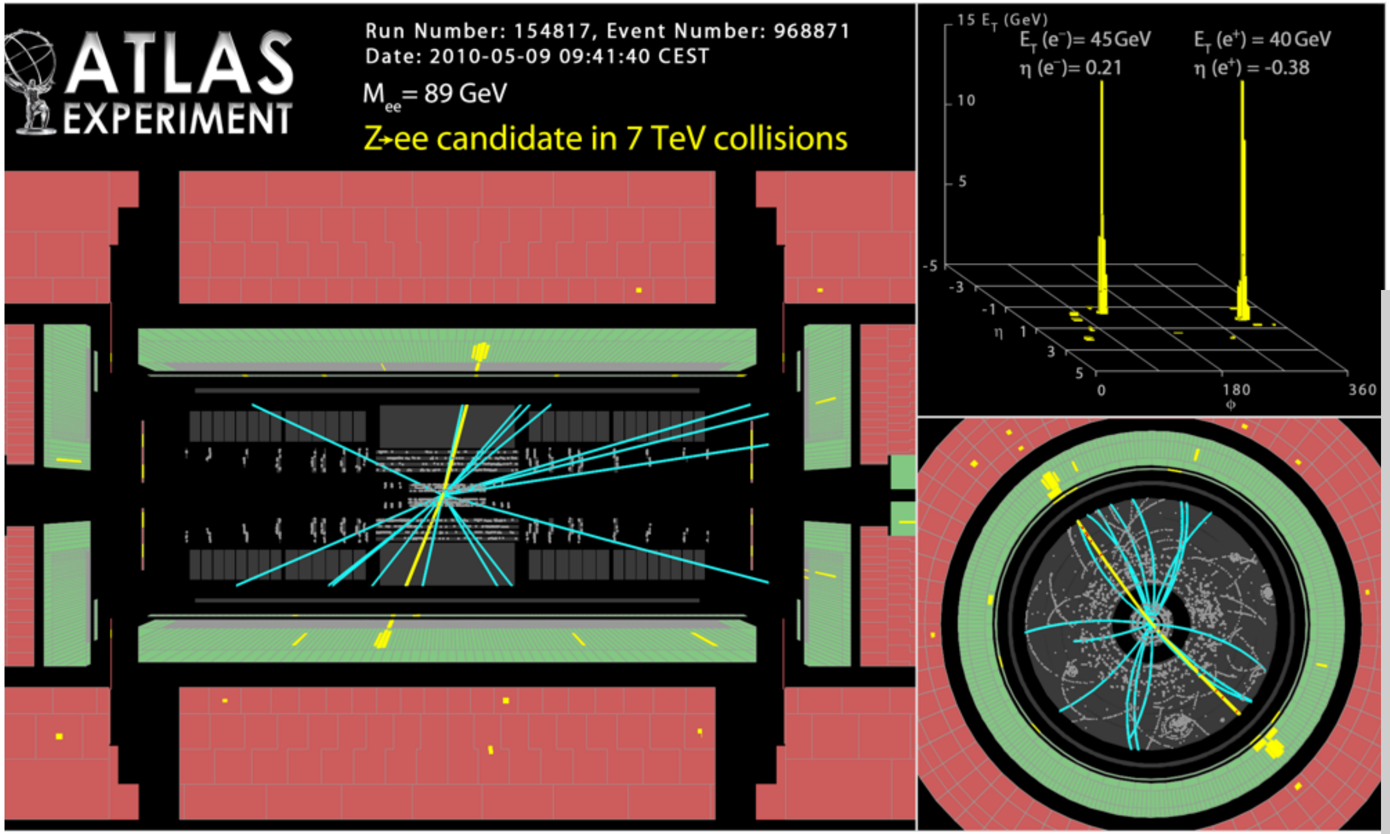


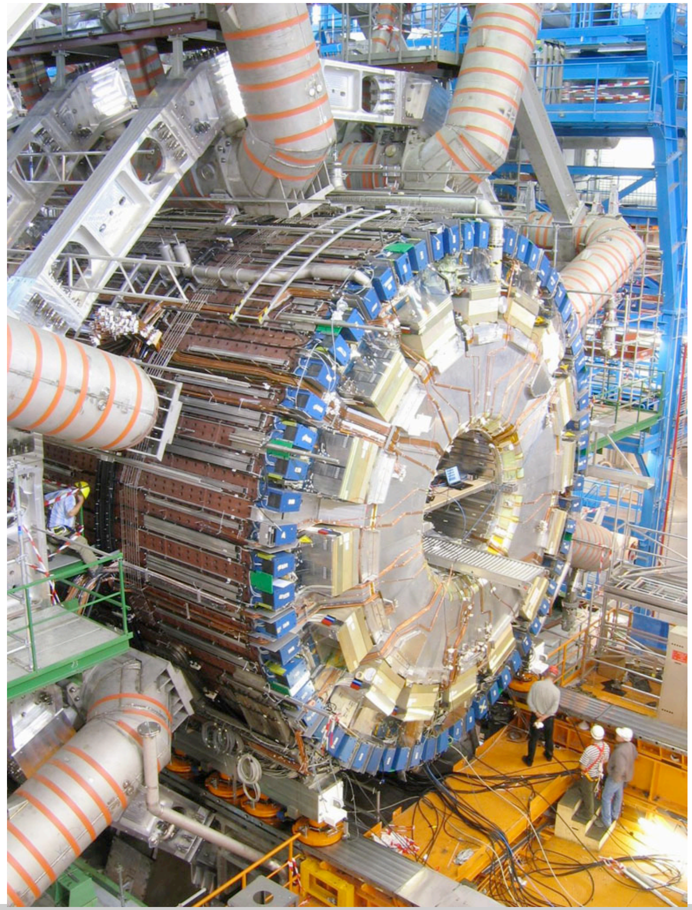


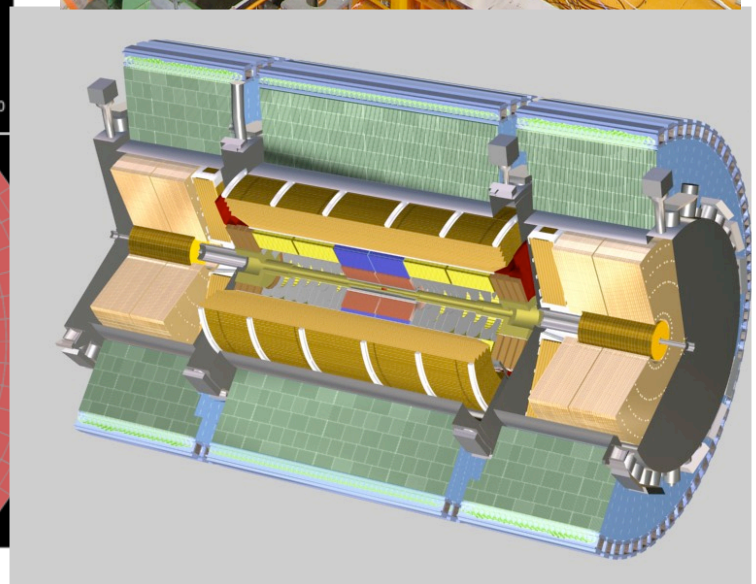


Calorimetry

- Liquid Argon & Pb accordion (EM & forward)
- Scintillator & Pb (hadronic)

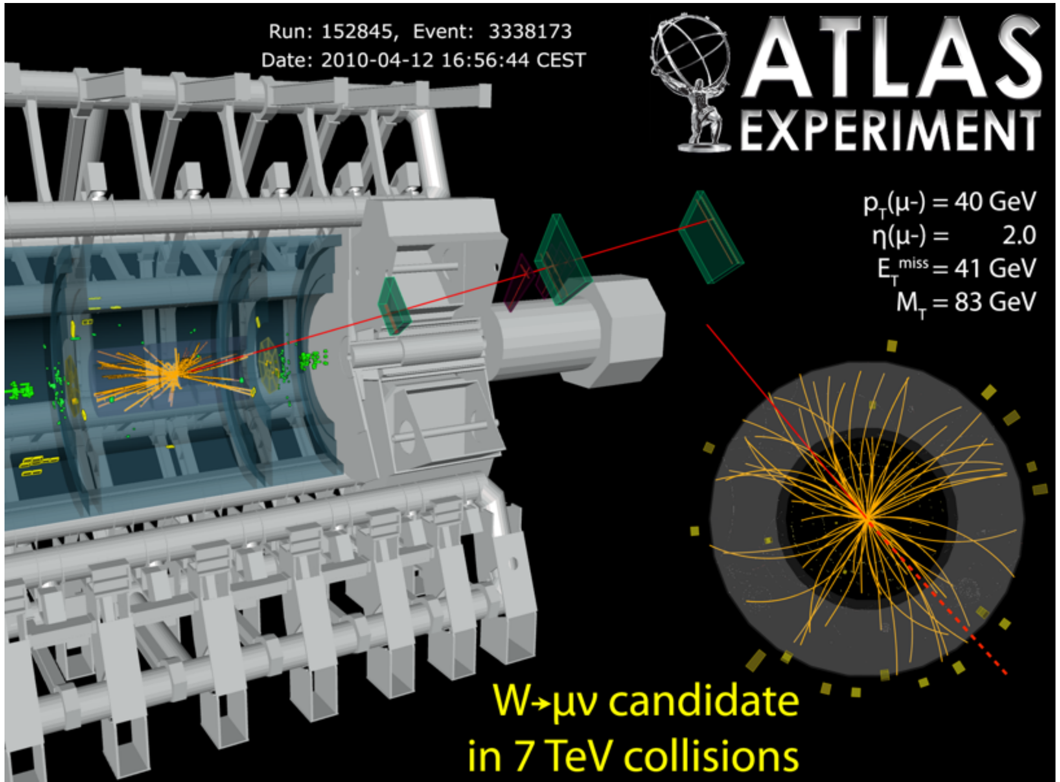


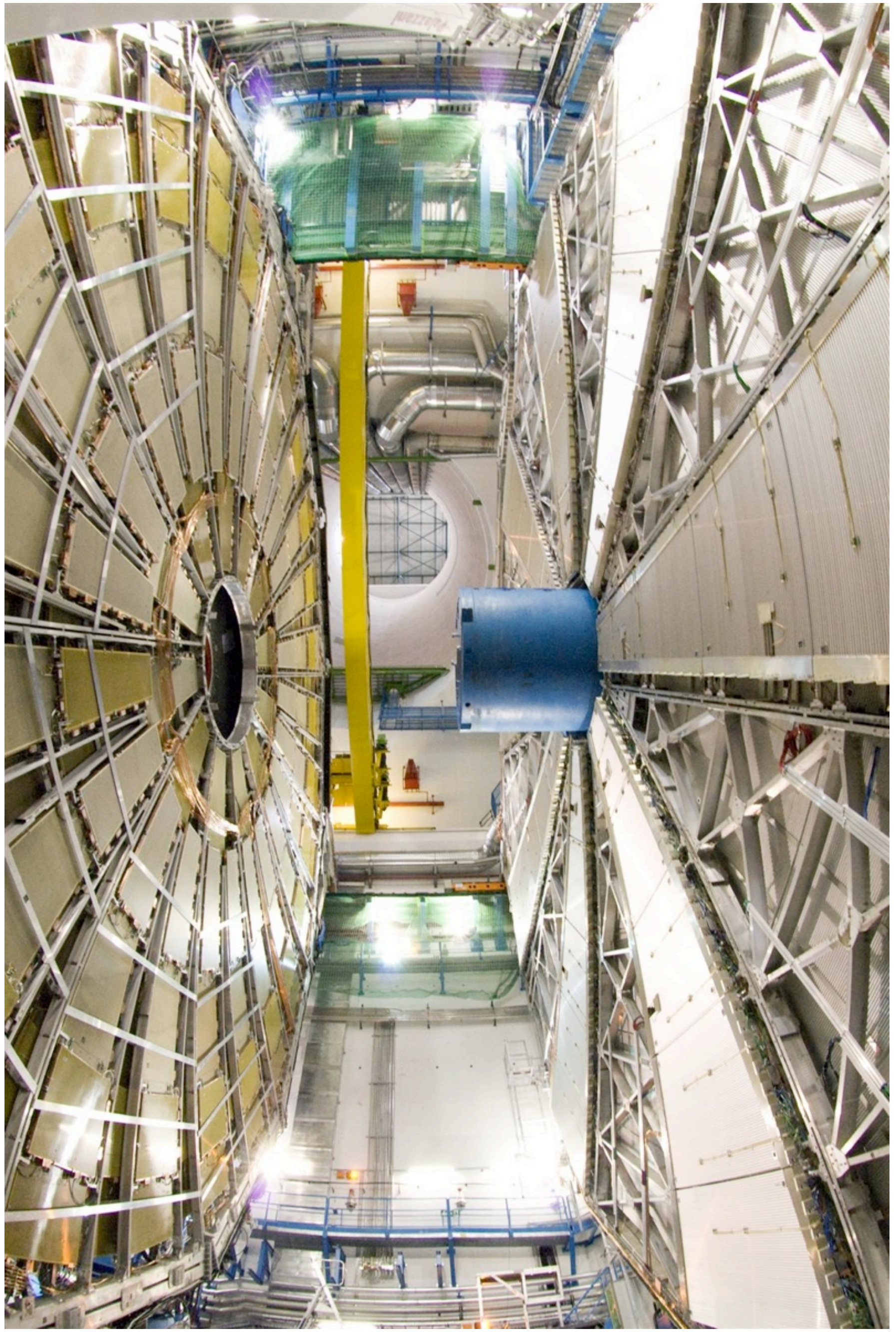




Muons

- Air-core toroids
  - Makes ATLAS big

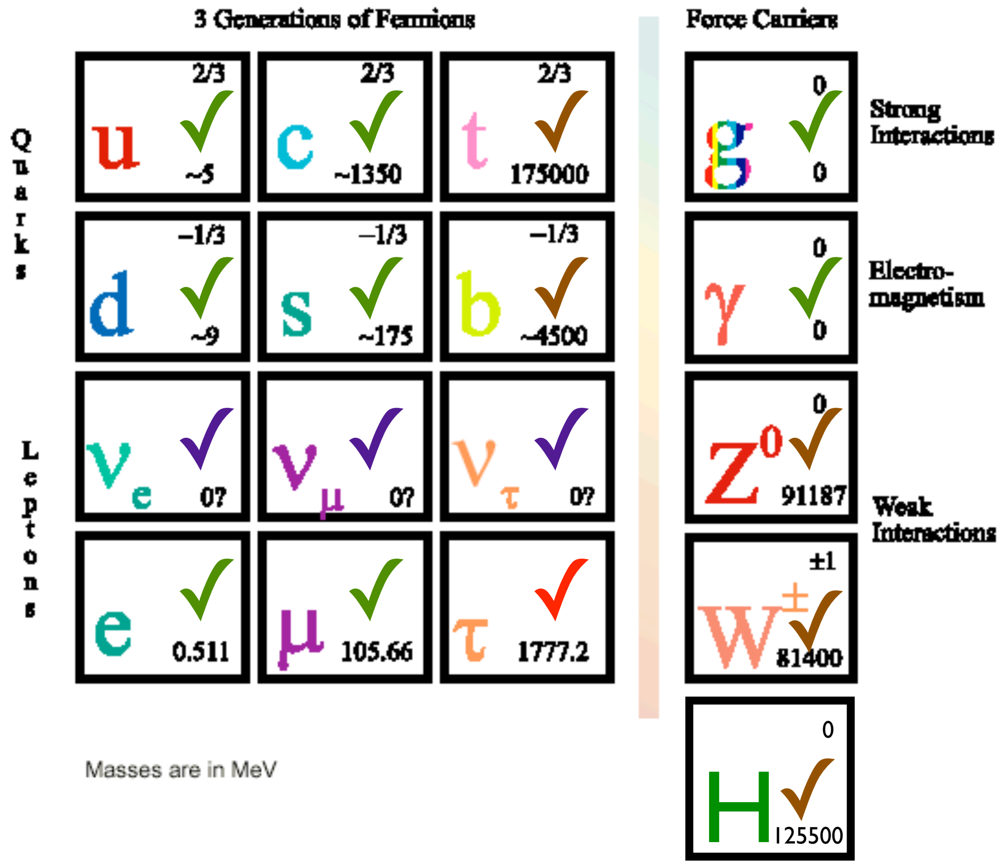




Neutrinos*

*(100% acceptance)

Detecting Particles



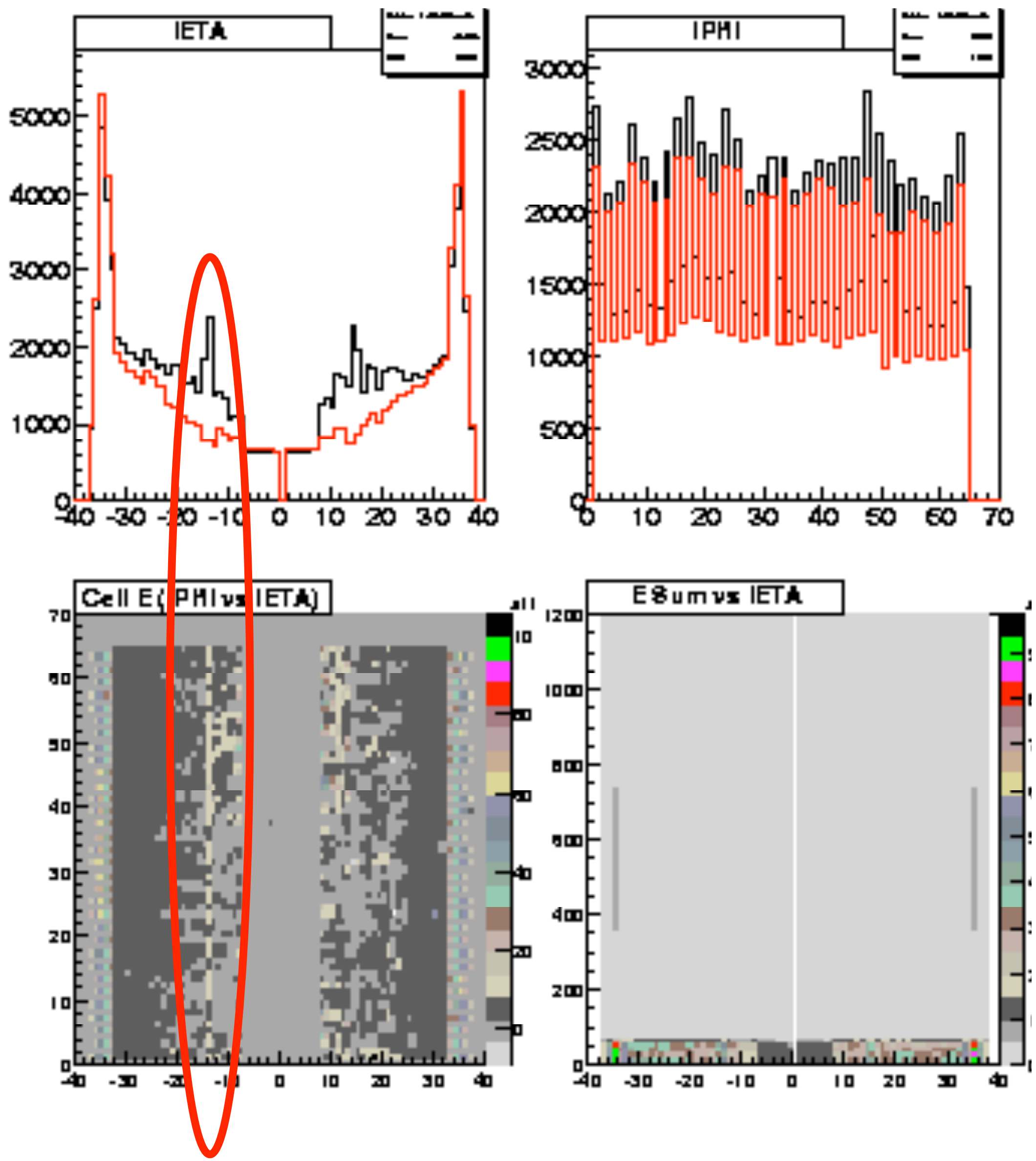
✓: Detect with high efficiency

✓: Detect by missing transverse energy

✓: Detect through decays: $t \rightarrow Wb$, $W/Z \rightarrow$ leptons, ...

Anecdotes From the Field (I)

- DØ's "ring of fire"
  - Noise in a few "eta rings"
  - Occurred on rare occasions
- Originally thought to be a ground fault in HV distribution
- Found to be concurrent with welding in building
  - Finally traced to liquid Argon purity & temperature monitoring



The Work

Steps in a Physics Analysis

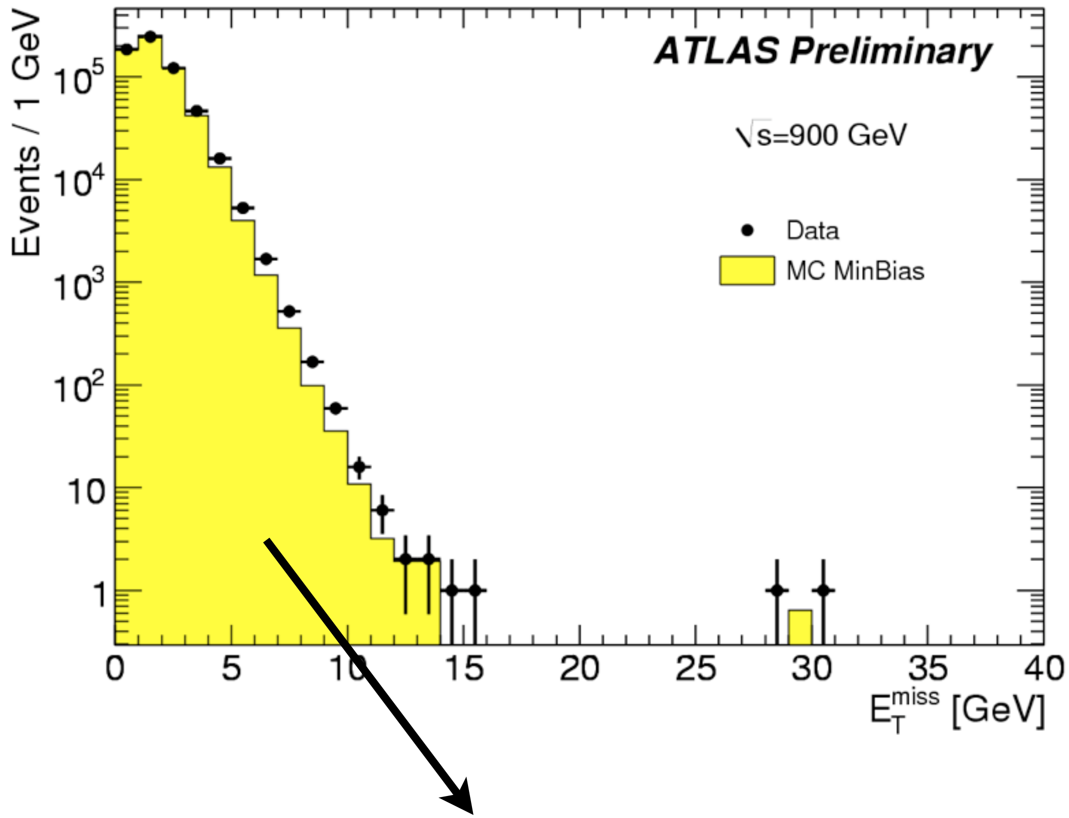
- What is the final state? $\Rightarrow$ "Preselection"
  - For a search, sufficiently loose to be signal-poor
    - Prove you understand the detector response, physics processes contributing
  - But sufficiently tight to have a manageable data volume
    - ATLAS/CMS write 200-400 Hz × 1+ MB/event = 200-400+ MB/s
    - "4-vectors" is not enough, need some amount of detector info
    - In practice, often have preselected sample for frequent analysis, + looser sample for e.g. multijet background with rare passes
- Note that data volume $\propto$ running time, not $\int\mathcal{L}$

Steps (II)

- Determine preselected sample's composition
  - MC and data to understand each contribution
    - QCD multijet background to leptons often extracted from data: rejection factor ~10^{-4}, difficult for simulation to be that accurate
    - MC for most other processes, with corrections from data, since generators are LO or NLO
  - Also need to correct MC for real-life data conditions
    - Different alignment, dead channels etc.
  - As statistics increase, more difficult, since mis-modelings not hidden by large statistical uncertainties anymore

Anecdotes From the Field (II)

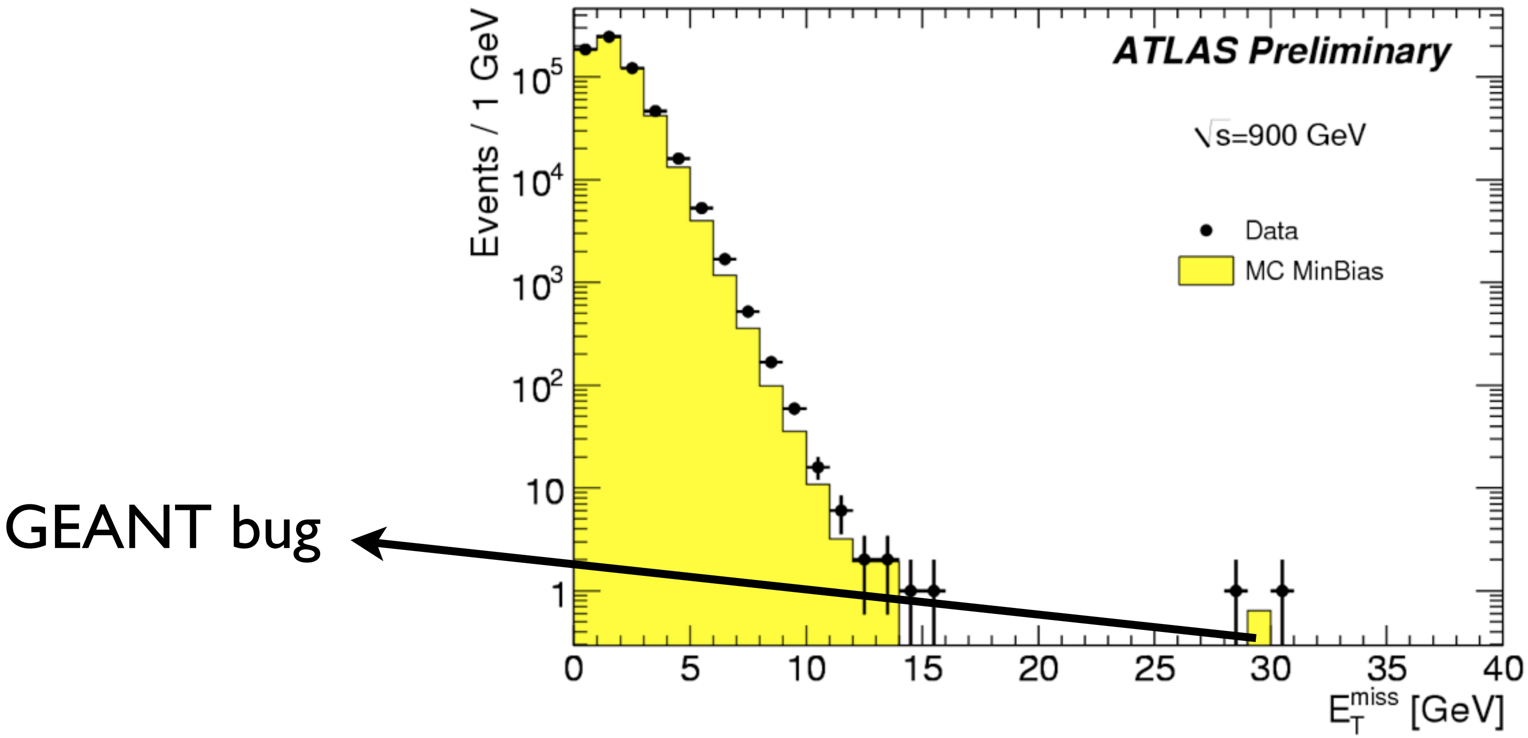
- Everybody wants experimenters to produce results fast
  - Lots of pressure in the early days of LHC...



Only jets, background easy

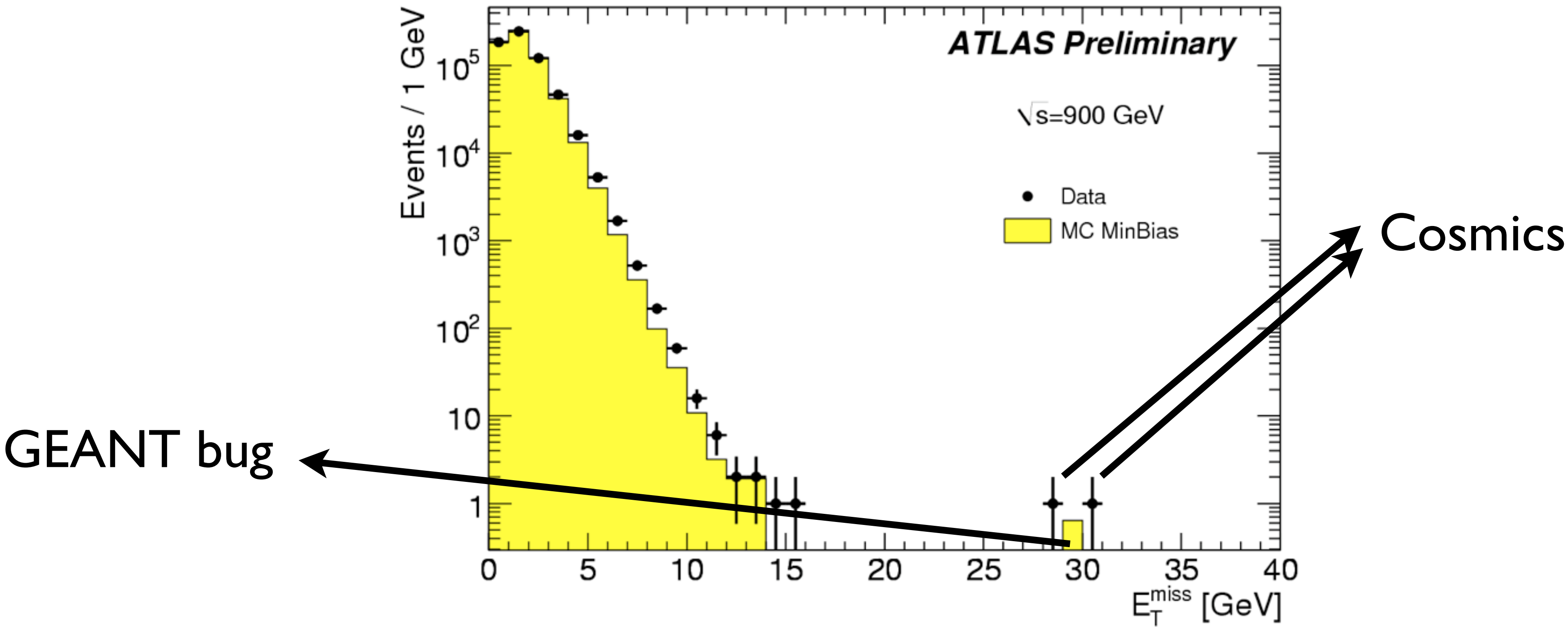
Anecdotes From the Field (II)

- Everybody wants experimenters to produce results fast
  - Lots of pressure in the early days of LHC...



Anecdotes From the Field (II)

- Everybody wants experimenters to produce results fast
  - Lots of pressure in the early days of LHC...

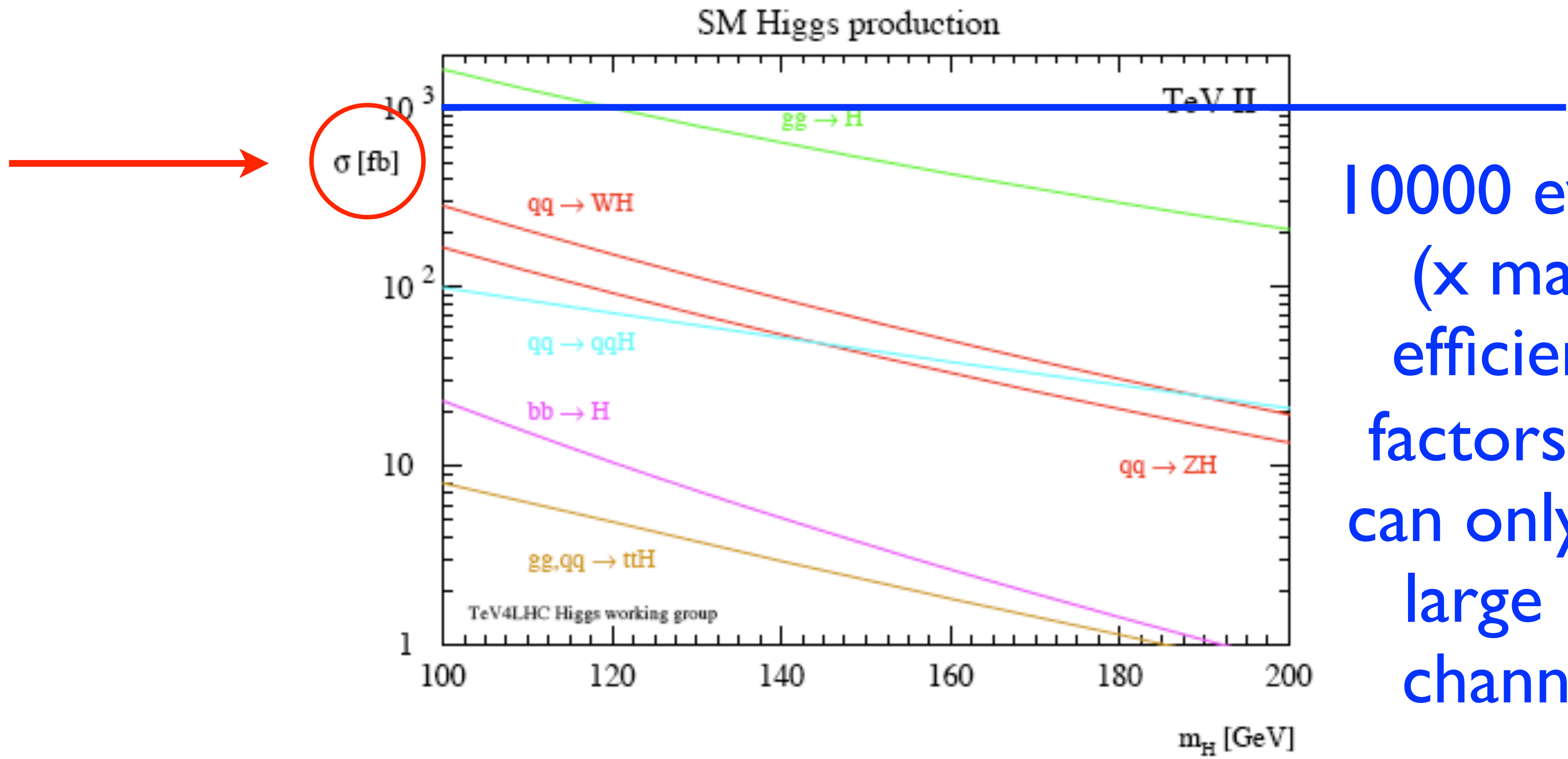


  - Sometimes, it's better to take the appropriate time to investigate

A Semi-Challenging Search: Higgs in WW and γγ Decays

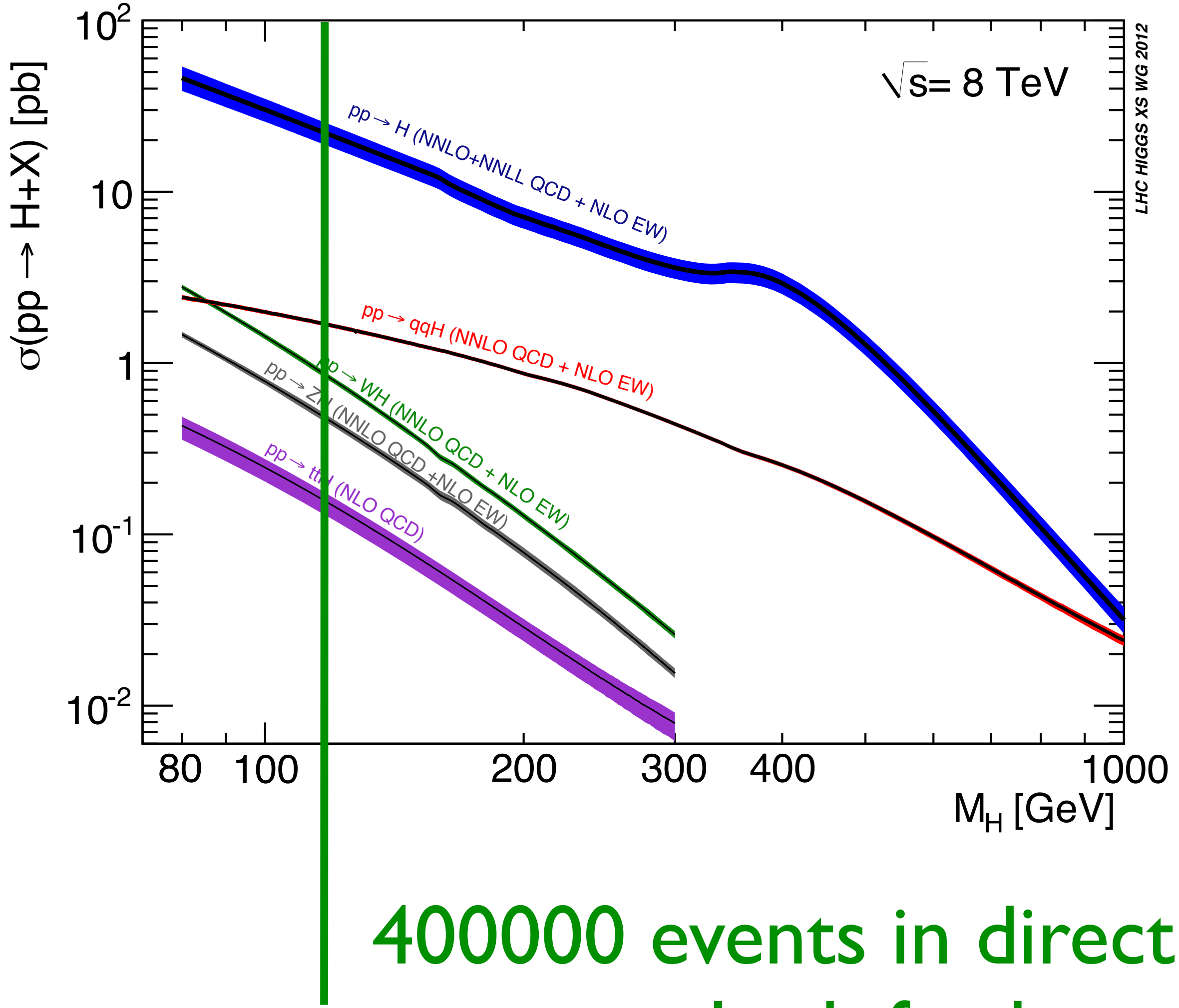
Producing Higgses

- Tevatron experiments collected ~10 fb^{-1} of data on tape
  - (Data taking efficiency was ~90%)



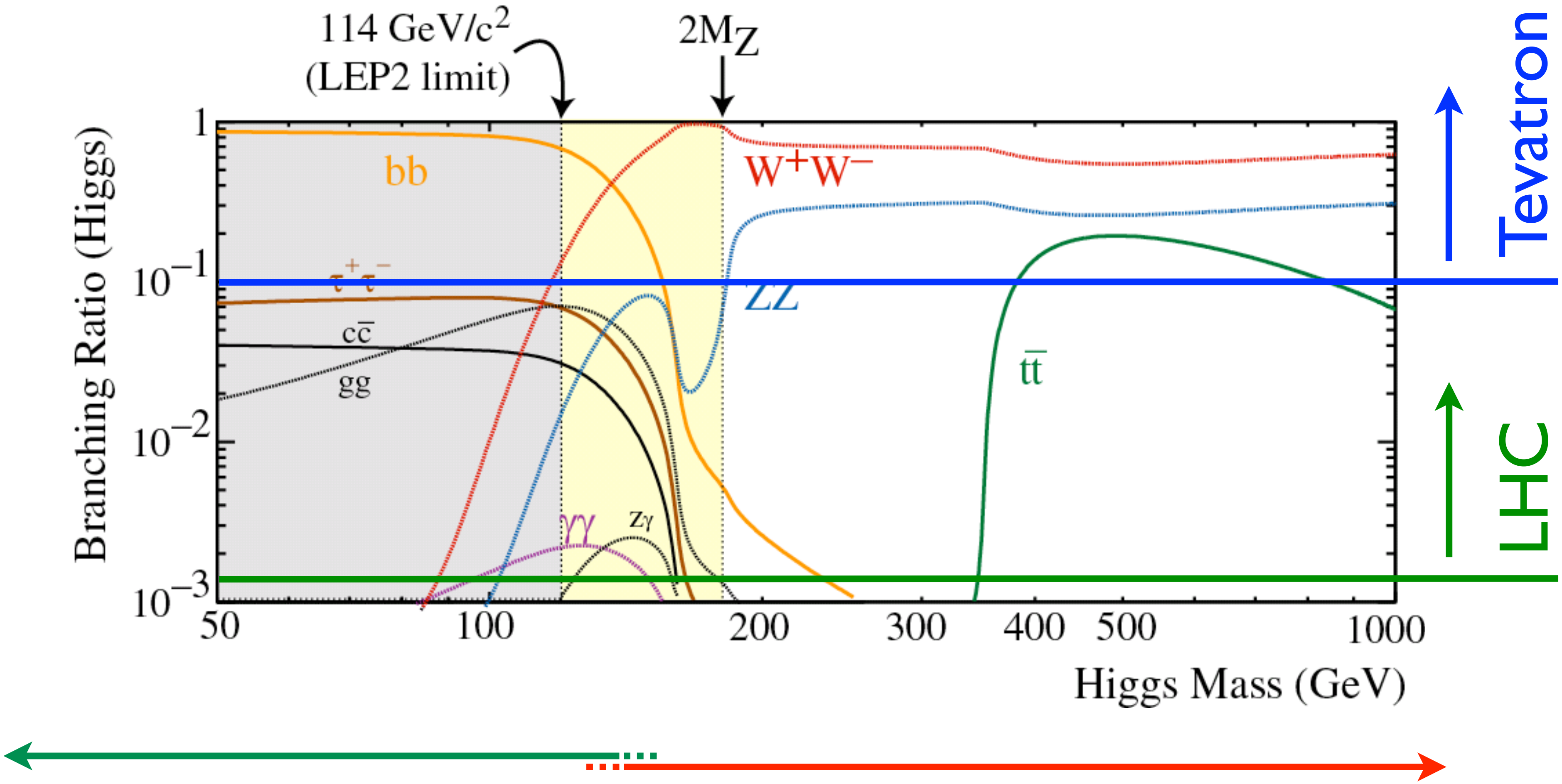
10000 events (x many efficiency factors) → can only use large BR channels

- Picture very different at LHC
  - 20 fb^{-1} collected by end 2012



400000 events in direct production
can look for better S/B!

Higgs Decay



Low Mass

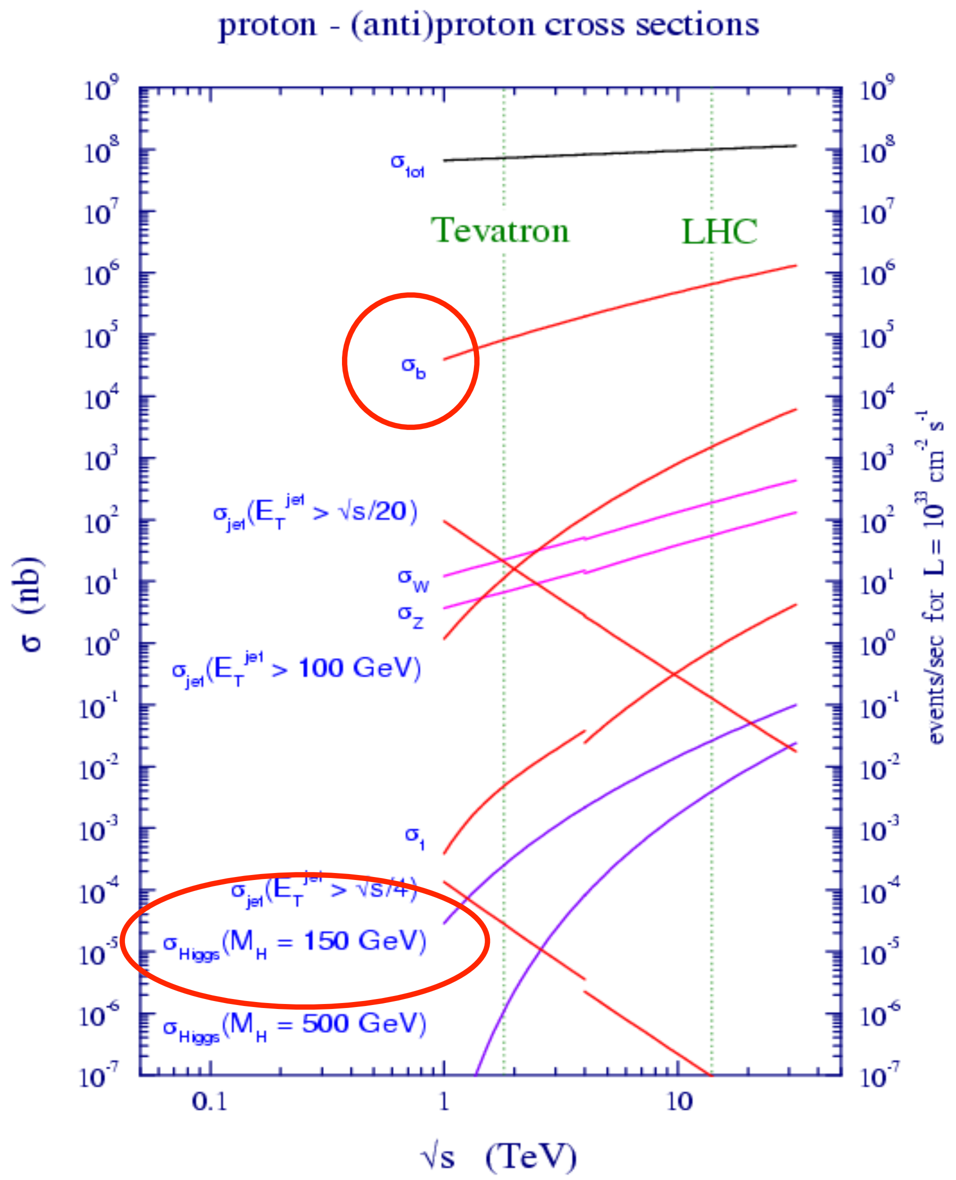
$H \rightarrow bb, \tau\tau, \gamma\gamma$

High Mass

$H \rightarrow WW, ZZ$

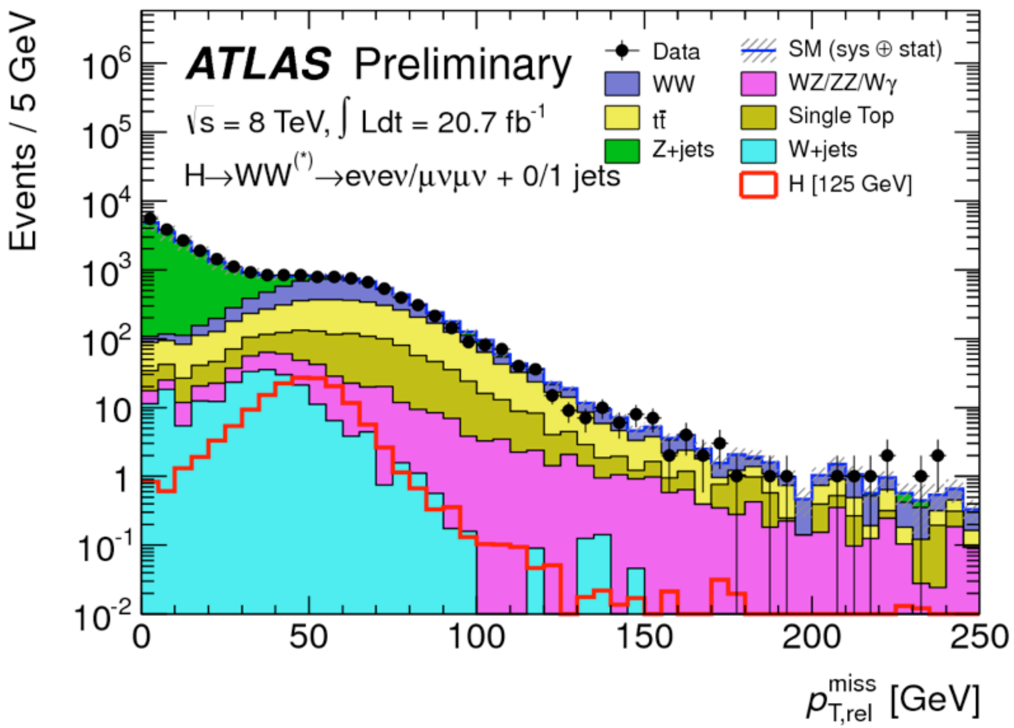
Search Channels

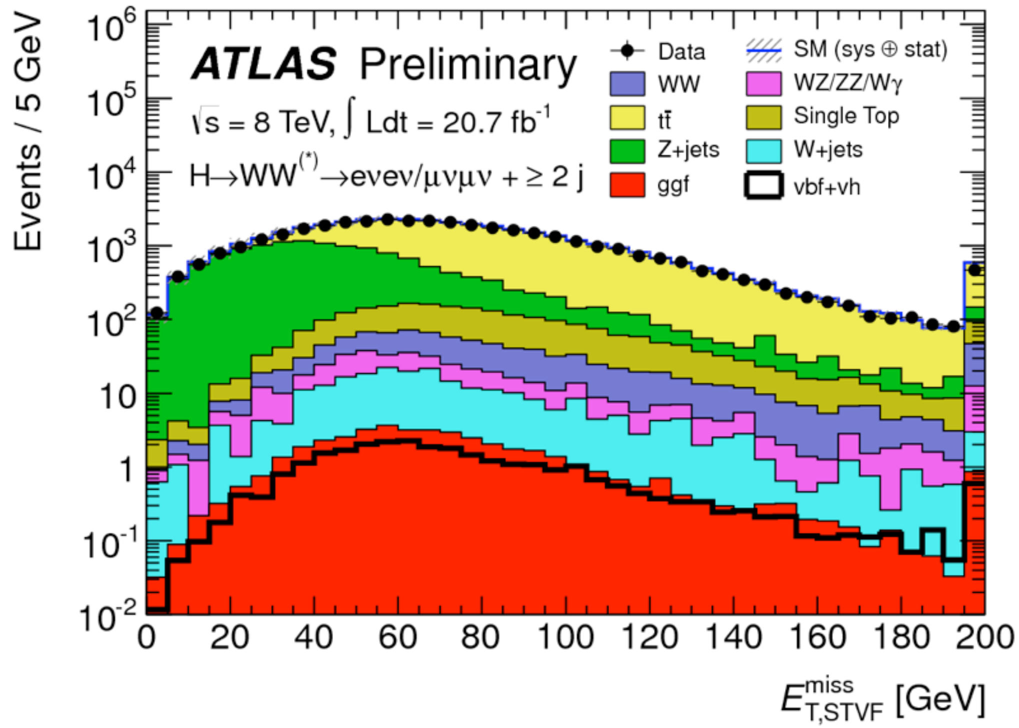
- Hadron colliders
  - $b\bar{b}$ production ~9 orders of magnitude larger than H
  - $gg \to H \to b\bar{b}$ swamped
- ➡ At low mass look for $pp \to WH$ or ZH ($\sigma/10$)$\to W/Z\, b\bar{b}$
  - With leptonic W, Z decay, so # of events /~10!
  - Or... $H \to \gamma\gamma$
- At high mass, $gg \to H \to WW/ZZ$ accessible if at least one W/Z decays leptonically



Dilepton + MET

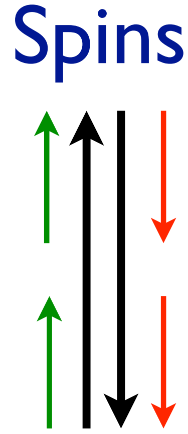
- "Golden" channel:
  - Main background $Z \to \ell\ell$ also a good calibration point
    - Of course, eνμν has lower background and twice the rate....
  - $Z \to \ell\ell$ "easy" to suppress using MET, at least when few jets ...

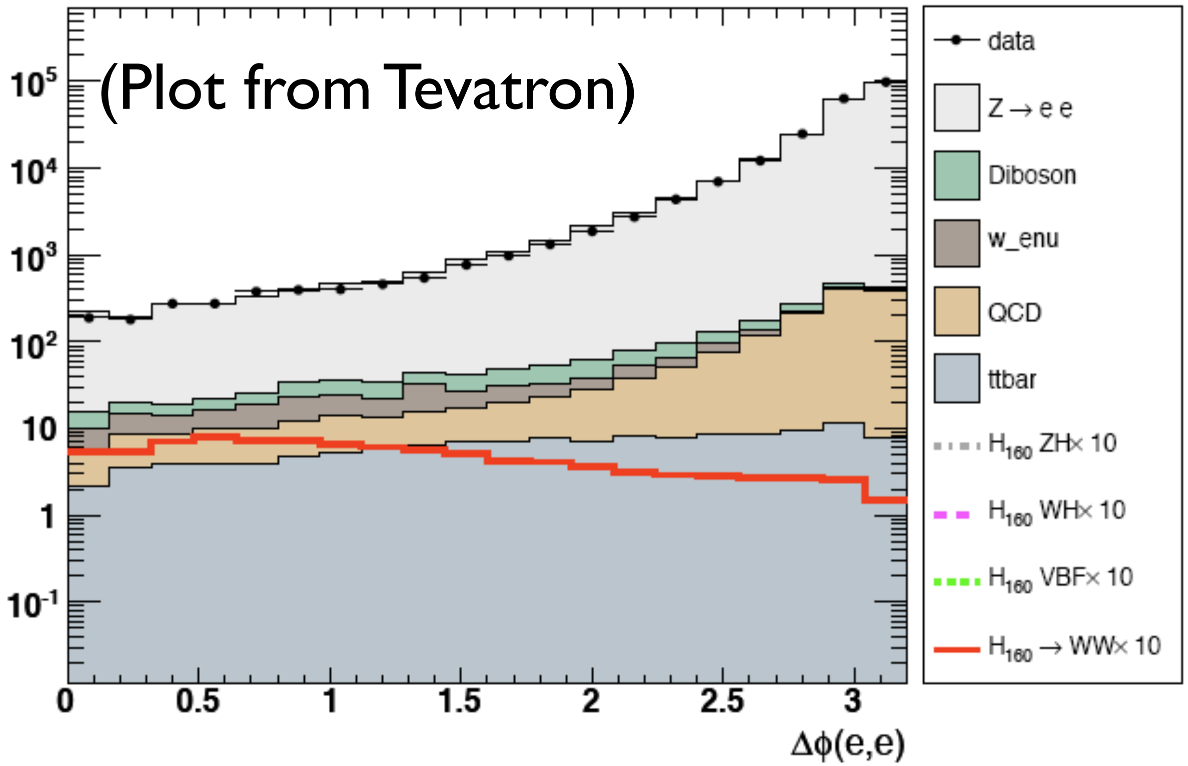


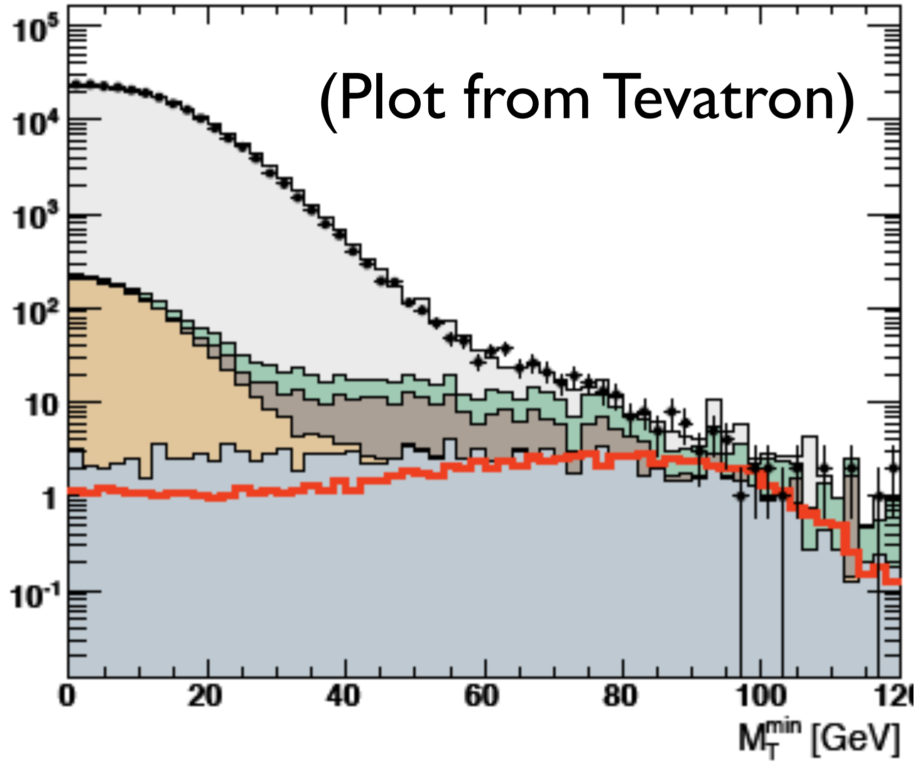


Angles

- In $Z \to \ell\ell$ (and dijets faking leptons), leptons preferentially emitted back-to-back
  - In Higgs decays, W^+W^- spins back-to back, so charged leptons in similar direction! (One LH, other RH)
  - In Z, smallest transverse mass tends to be small

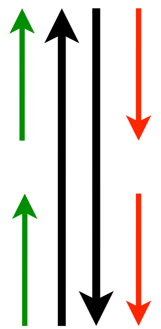




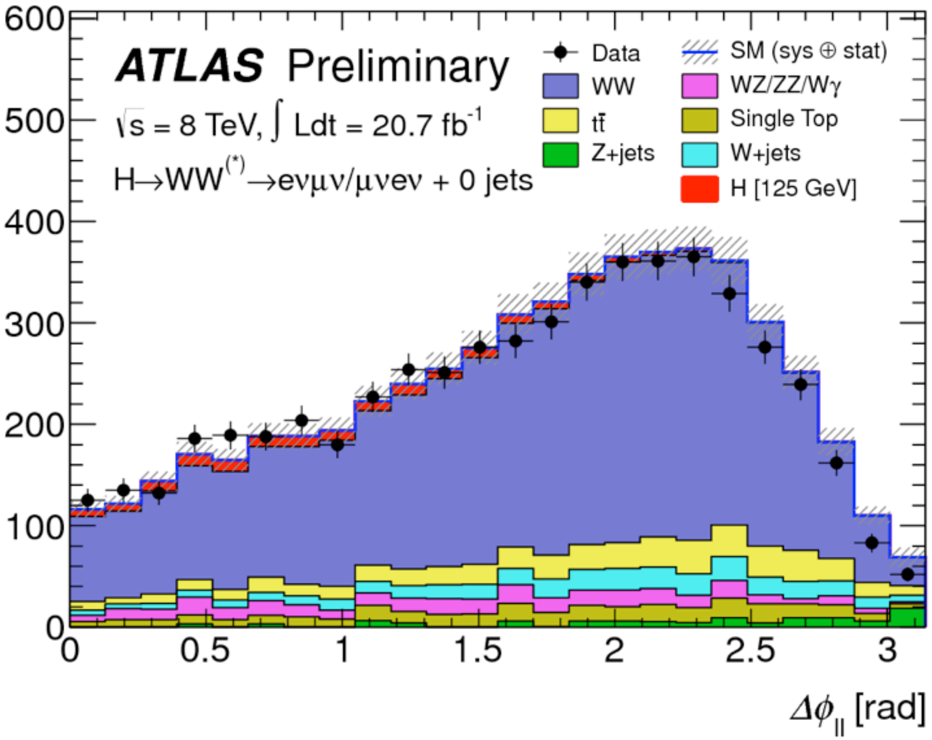


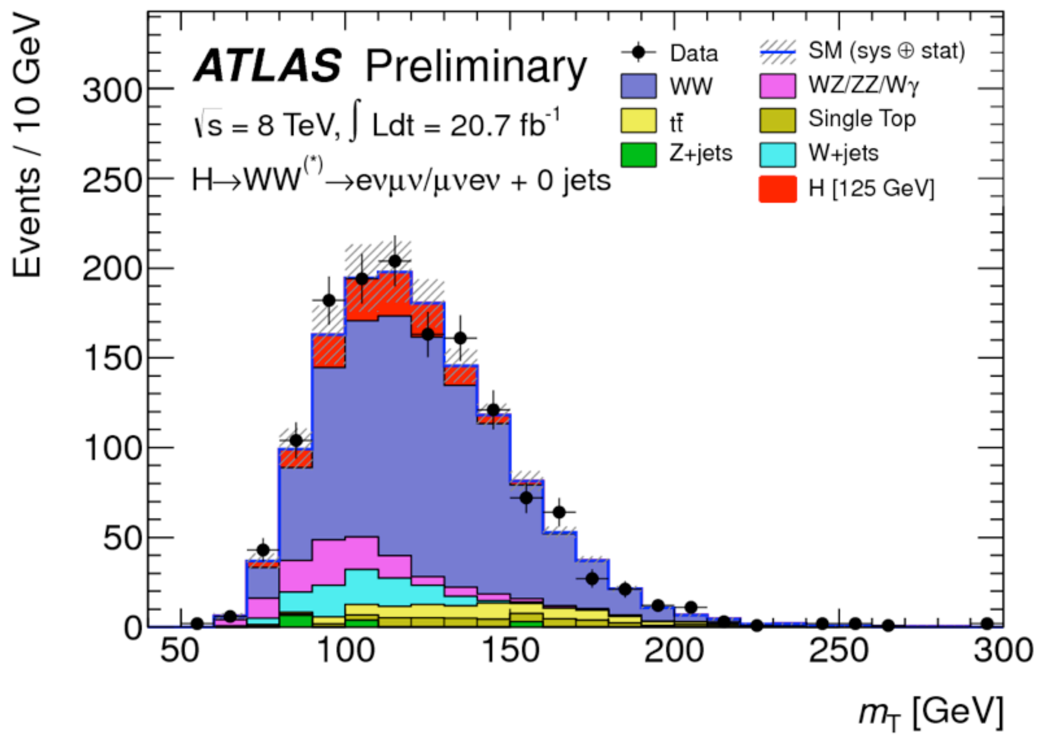
Angles

Spins



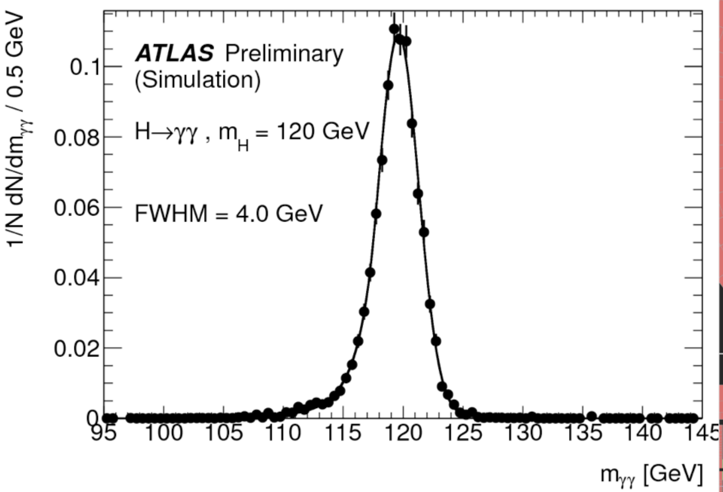
- In $Z \to \ell\ell$ (and dijets faking leptons), leptons preferentially emitted back-to-back
  - In Higgs decays, W^+W^- spins back-to back, so charged leptons in similar direction! (One LH, other RH)
  - In Z, smallest transverse mass tends to be small

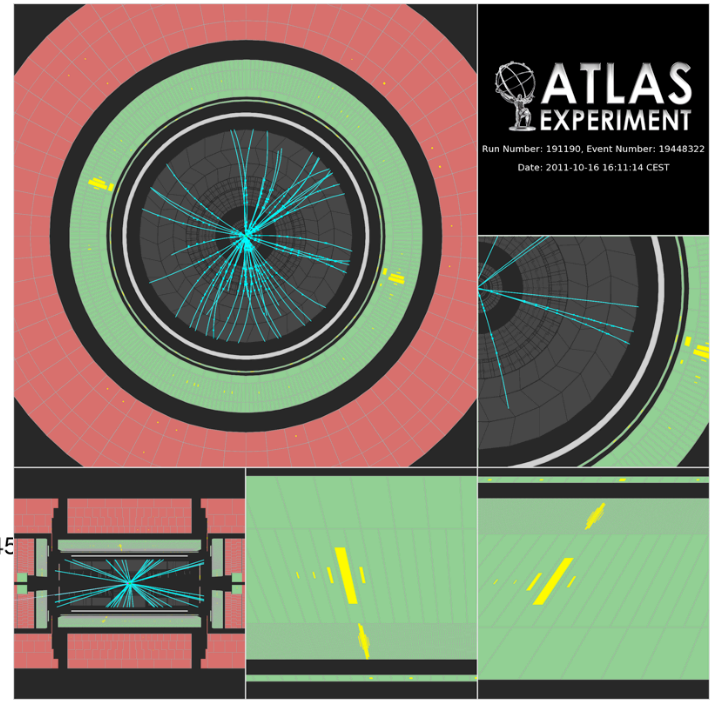


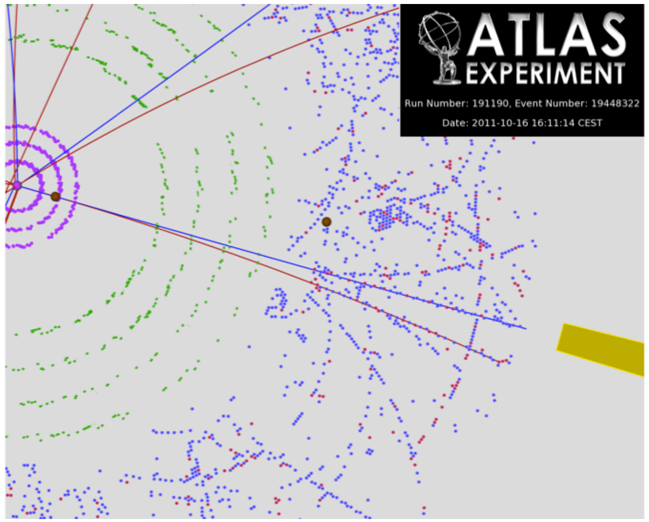


$H \rightarrow \gamma\gamma$

- Low branching ratio ("rare decay"), but:
  - Simple final state with excellent mass resolution
  - Large, but not huge SM "continuum" background
  - Excellent "calibration" reference from $Z \rightarrow e^+e^-$ at 91.2 GeV
  - In our detectors, many photons convert before calorimeters

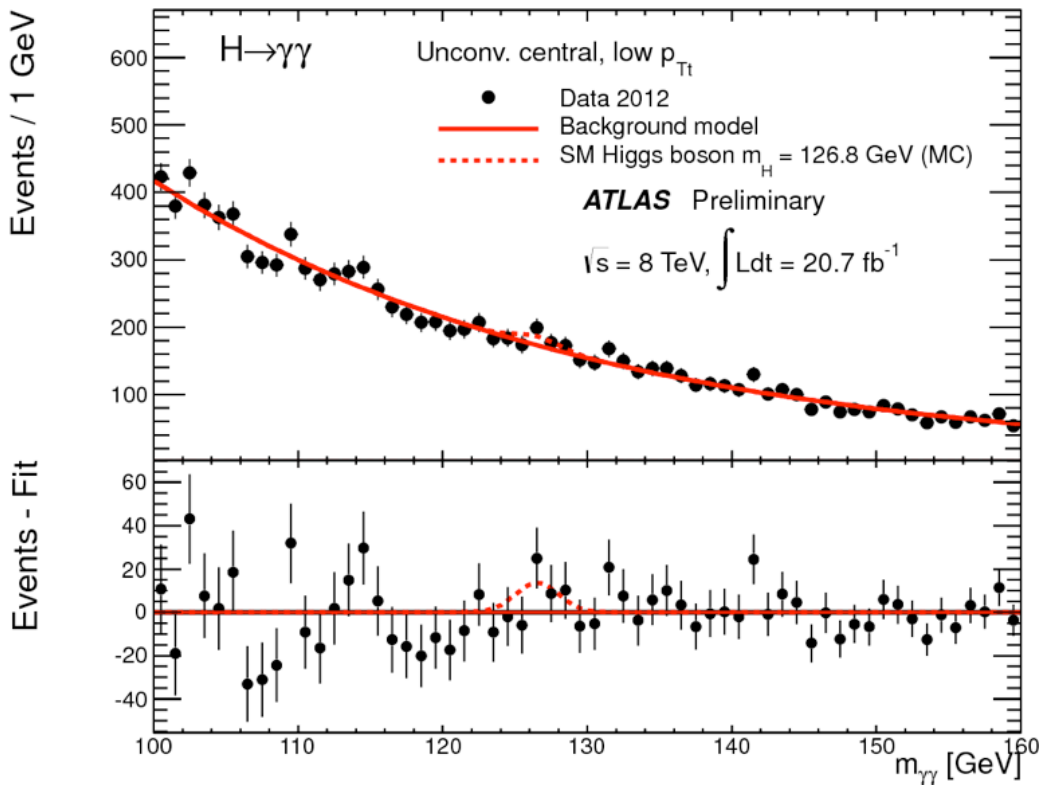


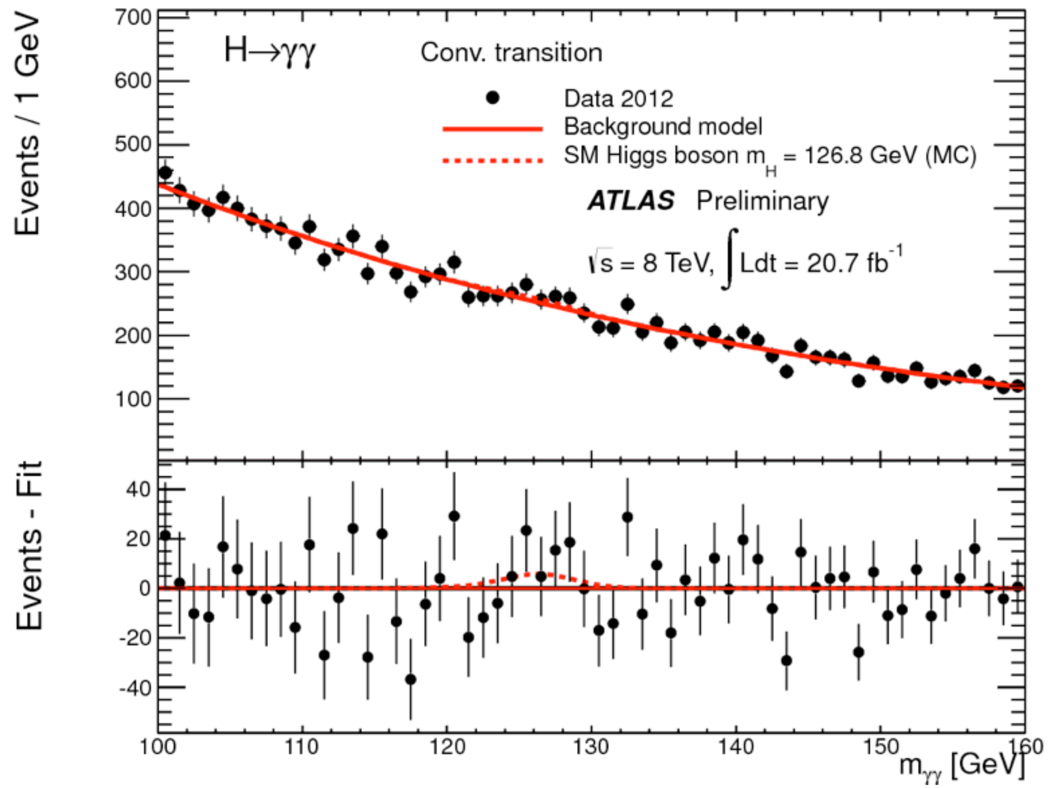




Categorize!

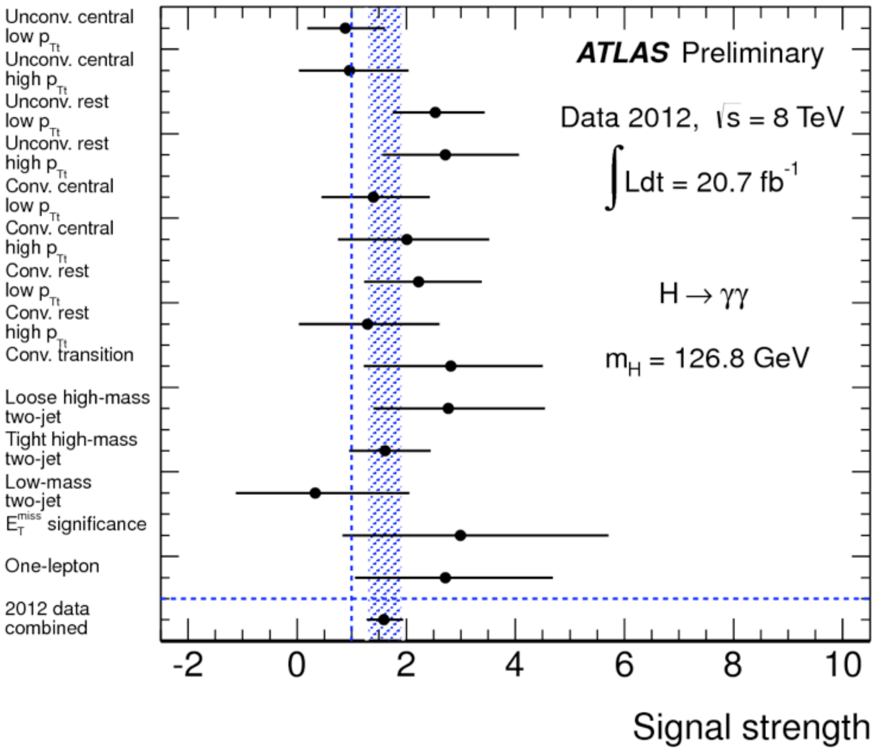
- γγ event subsamples have different S/B
  - Valuable to split to assign corresponding weights (typically ln(1+S/B)), will increase sensitivity





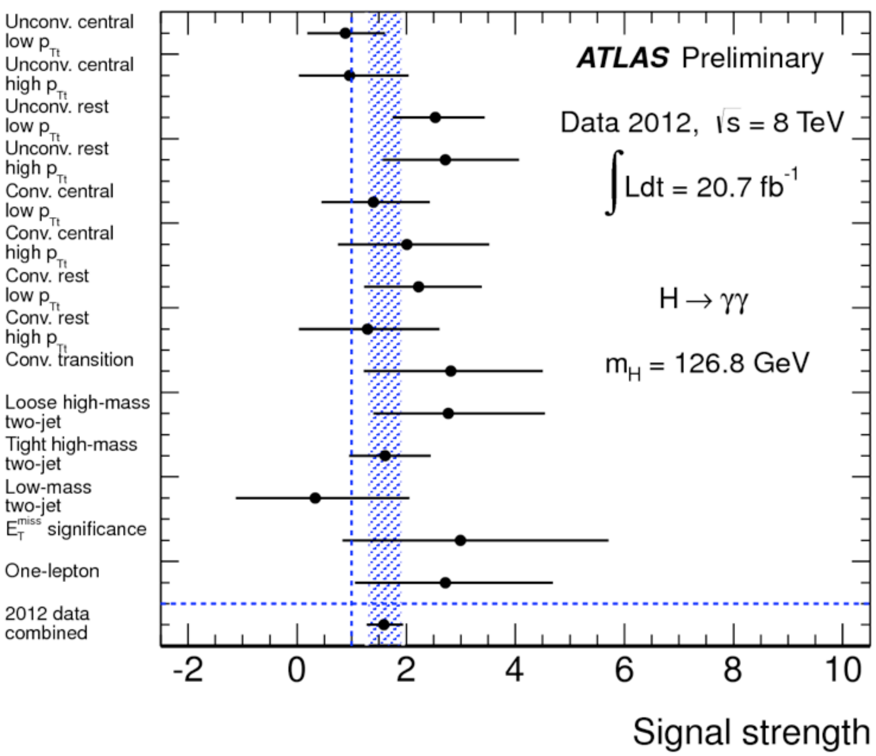
Categorize!

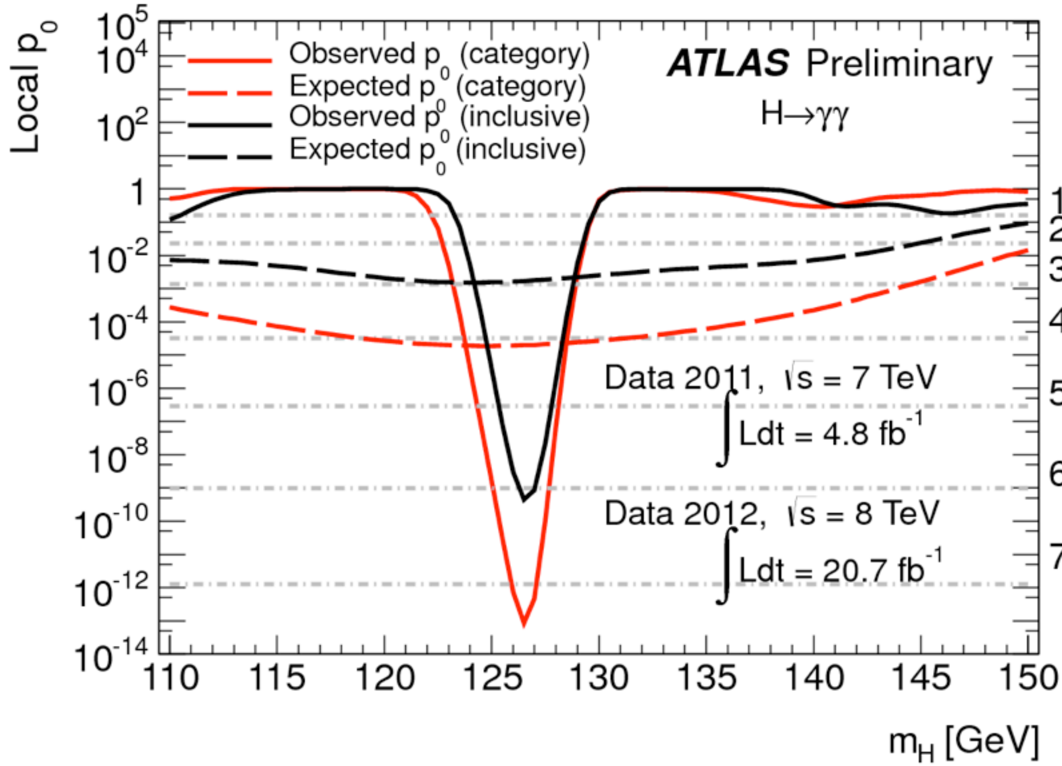
- $\gamma\gamma$ event subsamples have different S/B
  - Valuable to split to assign corresponding weights (typically ln(1+S/B)), will increase sensitivity



Categorize!

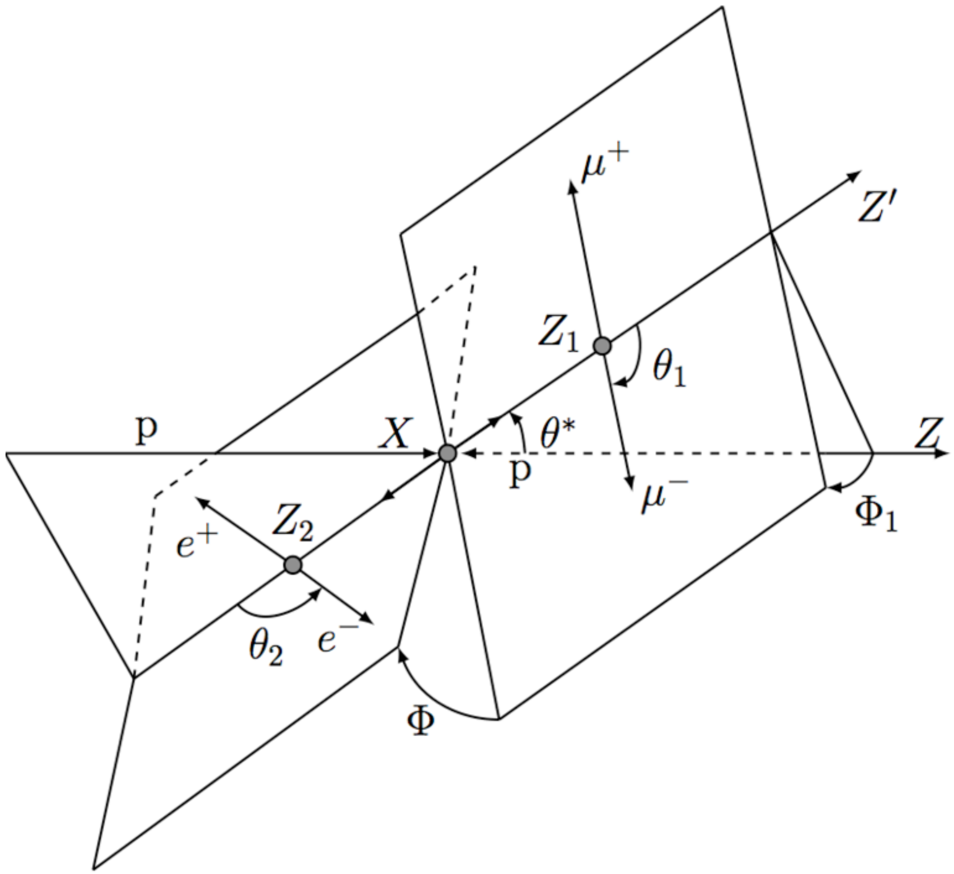
- Event subsamples have different S/B
  - Valuable to split to assign corresponding weights (typically ln(1+S/B)), will increase sensitivity

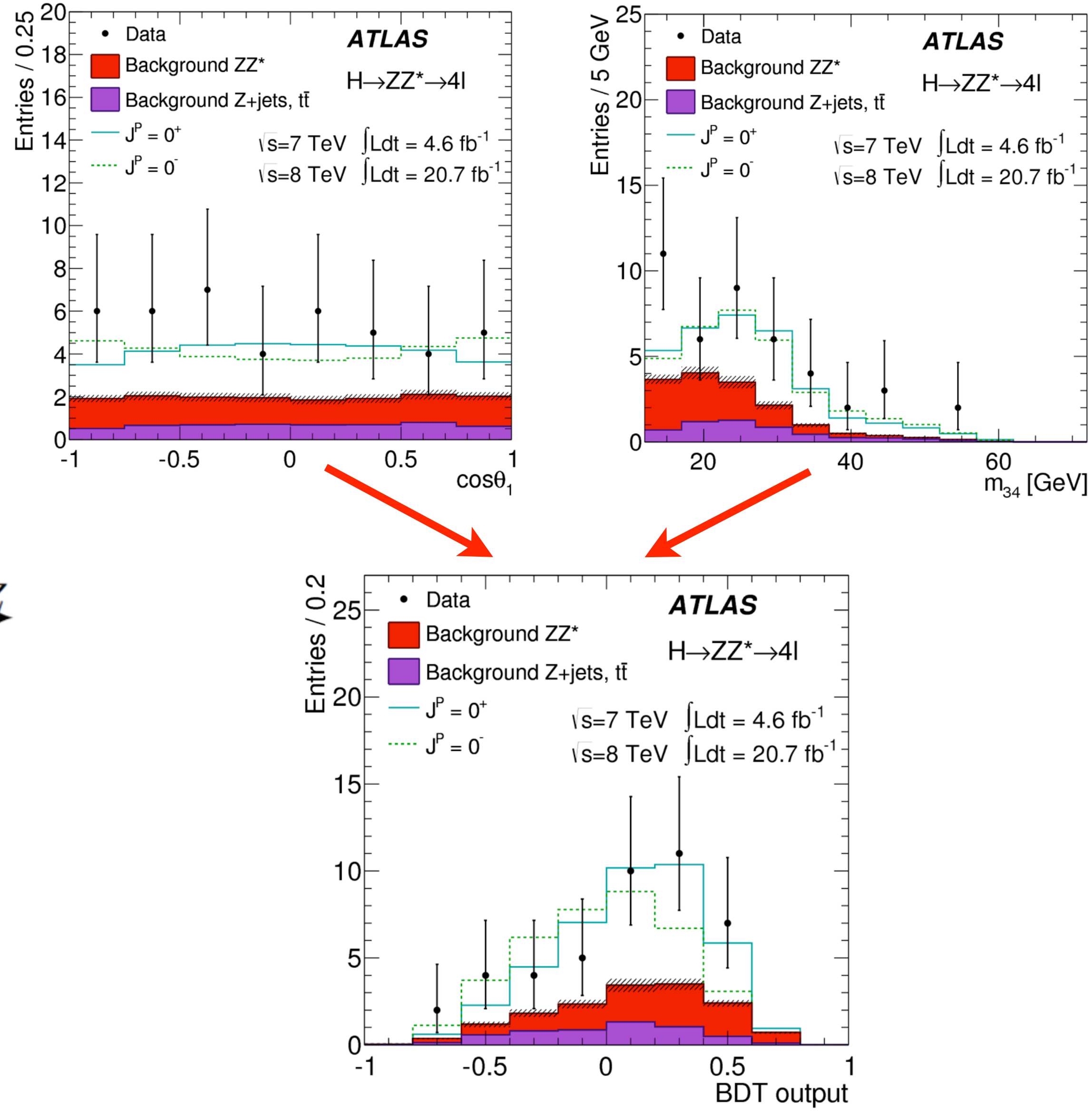




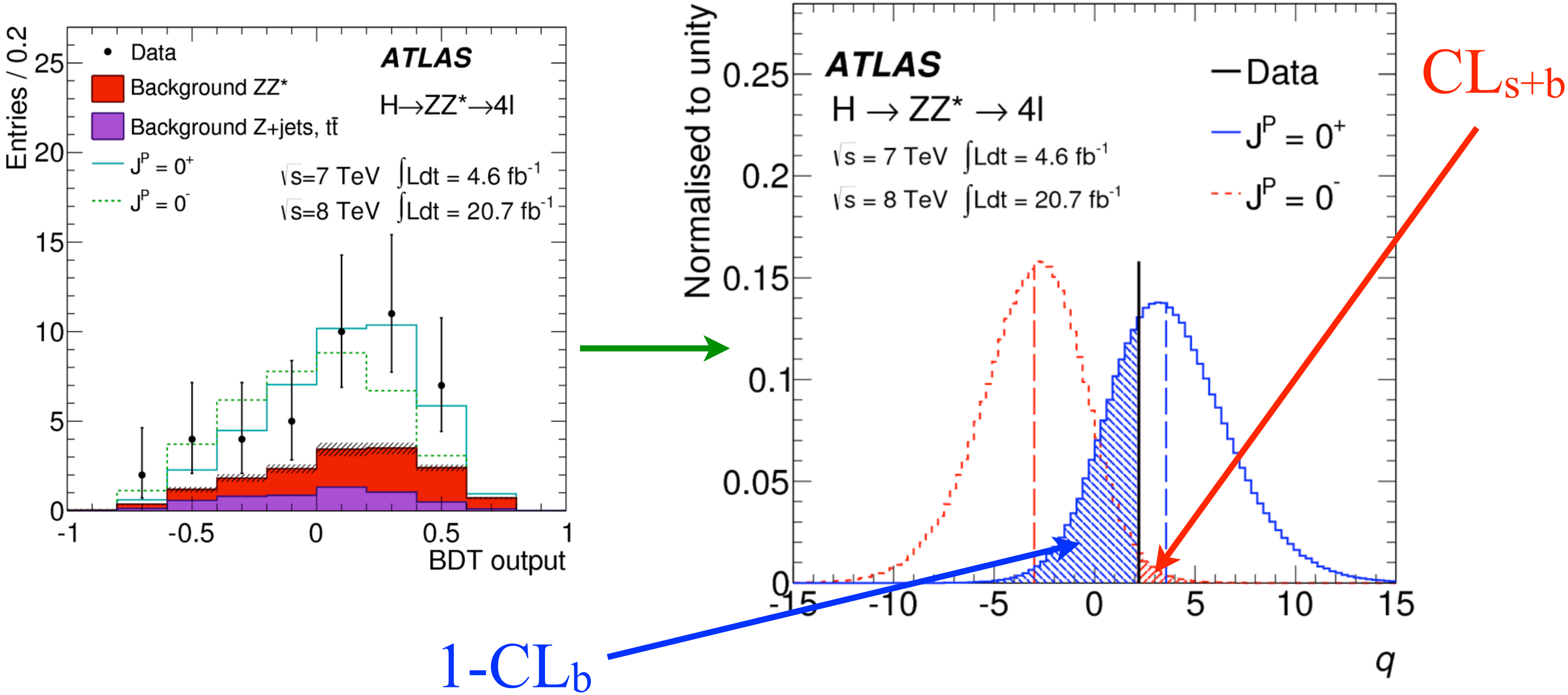
(Pseudo)-Scalar?

- Pseudo-scalar = smoking gun of compositeness





CLs

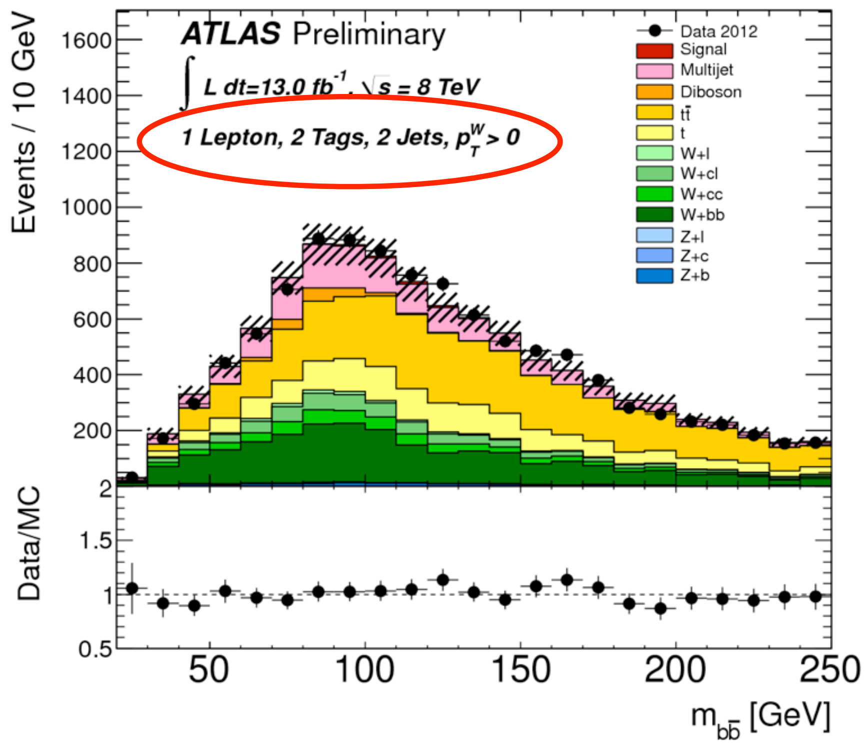


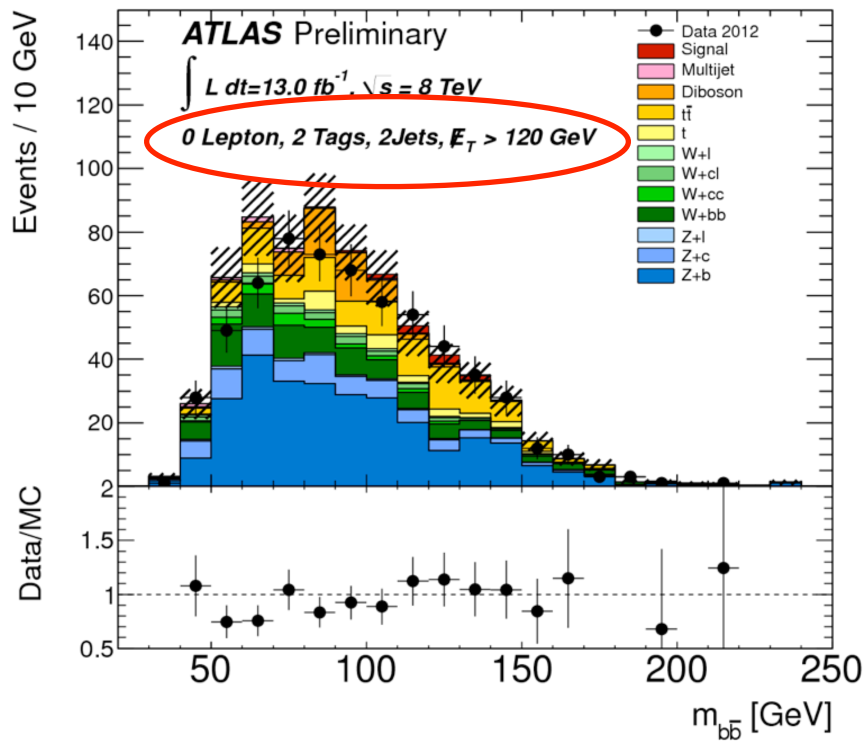
In principle, CL_{s+b} is all you need...
Unless your background description is wrong → $$CL_s = \frac{CL_{s+b}}{CL_b}$$

More Challenging: Add Jets

Jetty Backgrounds and the Higgs

- Final states with jets can't be ignored: H-b coupling!
  - But tons of background jets at the LHC, so require leptons
    - (Including neutrinos....)
- And in addition dijet mass resolution is so-so:





Sample Composition

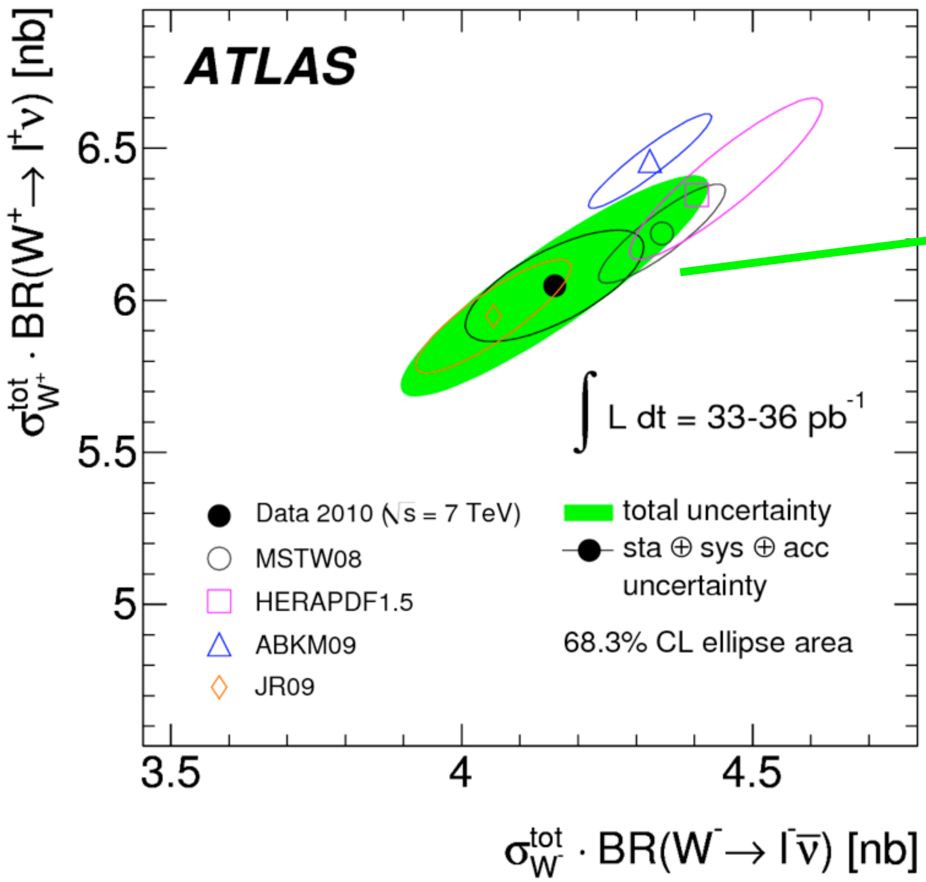
- After preselection, low S/B allows to verify shapes of dominant backgrounds
  - E.g. for WH, first before *b*-tagging (W+light), then with 1 tag (W+b), then 2 tag but more jets (top)
- Determining the sample's composition
  - I.e. which processes contribute, and how

Increasing difficulty ↓

- Diboson from MC simulation (usually small, + "trust" MC)
- Z+jets from data & MC ("easy" to get a clean sample, correct MC)
- QCD multijet from data (no choice)
- Top from MC + data
- W + jets from MC + data, but ....

Sometimes Physics Helps

- At the LHC, produce more W^+ than W^-
  - Can exploit that to normalize W+jets



$$(N_{W^+} + N_{W^-})^{\text{exp}} = \left(\frac{r_{\text{MC}} + 1}{r_{\text{MC}} - 1}\right)(N_{W^+} - N_{W^-})^{\text{data}}$$

But what about shape??

Generators Used

- We use four kinds of Monte Carlo generators
  - "Calculators" (often NNLO) do not actually generate events, they just calculate some (limited) distributions, like W p_T
  - Traditional $2 \rightarrow 2$ generators: LO, e.g. $q\bar{q} \rightarrow WZ$
    - Include parton shower, i.e. QCD radiation, and hadronization to jets
  - "Matrix Element" $2 \rightarrow n$ ($n < 9$): LO, e.g. $q\bar{q} \rightarrow e\nu jjjj$
    - Necessary to generate events with multiple hard jets
    - Require matching to parton shower to avoid double counting
  - NLOwPS $2 \rightarrow 2$ generators: include NLO corrections
    - I.e. in a sense they are $2 \rightarrow 2$ & 3 with virtual corrections

Correction Factors

- Of course, the ME's are LO, so "K-factors" needed
  - Different ones for heavy flavor etc..... (DØ) convention to avoid confusion....

- K-factor is purely theoretical, and denotes a (N)NLO/LO ratio of cross sections;
- K'-factor is also theoretical, and denotes a (N)NLO/LL ratio of cross sections. According to Steve, ALPGEN cross sections are Leading Log;
- S-factor is empirical, and comes on top of K or K' to bring MC in agreement with data. MC should be initially normalized to luminosity, and all correction (a.k.a. scale) factors should be applied (trigger, ID...);
- HF-factor is, in principle, theoretical, but in practice only theory inspired. It tells you by how much heavy flavor production should be increased, on top of K or K', and possibly S;
- S_HF-factor is empirical, and comes on top of K or K', S, and HF, to bring MC in agreement with data, after b-tagging.

Correction Factors

- Of course, the ME's are LO, so "K-factors" needed
  - Different ones for heavy flavor etc..... (DØ) convention to avoid confusion....

• K-factor is purely theoretical, and denotes a (N)NLO/LO ratio of cross sections;

• K'-factor is also theoretical, and denotes a (N)NLO/LL ratio of cross sections. According to Steve, ALPGEN cross sections are Leading Log;

• S-factor is empirical, and comes on top of K or K' to bring MC in agreement with data. MC should be initially normalized to luminosity, and all correction (a.k.a. scale) factors should be applied (trigger, ID...);

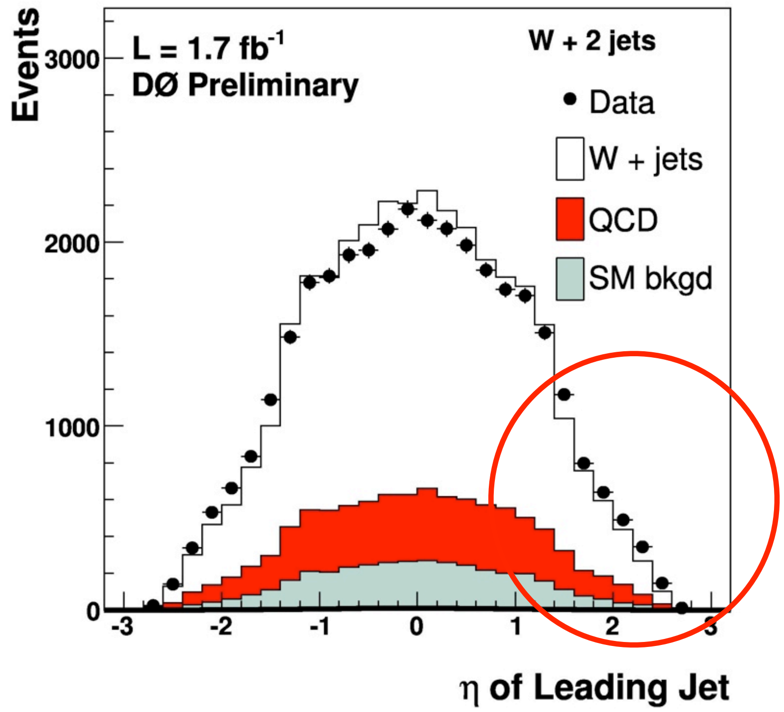
• HF-factor is, in principle, theoretical, but in practice only theory inspired. It tells you by how much heavy flavor production should be increased, on top of K or K', and possibly S;

• S_HF-factor is empirical, and comes on top of K or K', S, and HF, to bring MC in agreement with data, after b-tagging.

In addition to W/Z PT reweighting

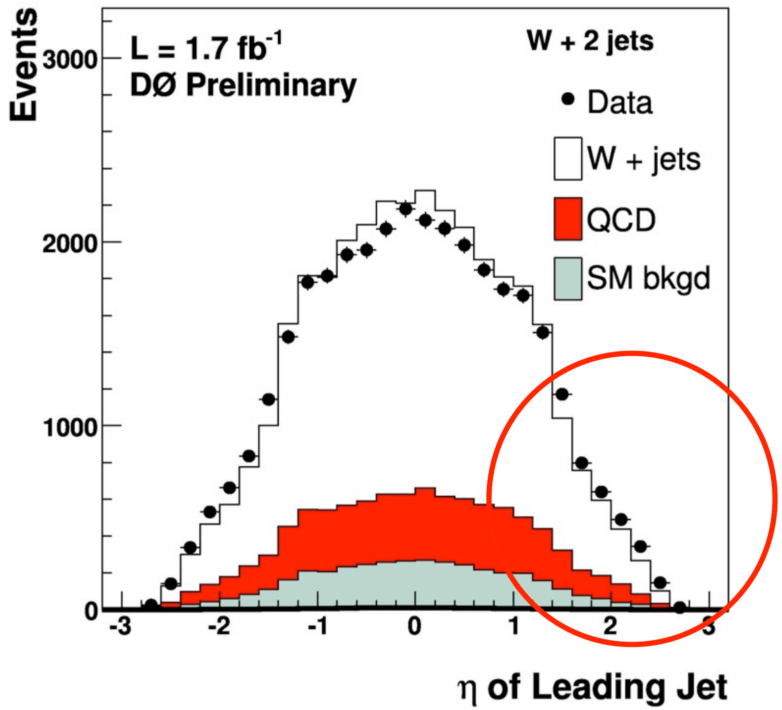
Anecdotes From the Field (III)

- Pile-up events ("minimum bias") do produce jets



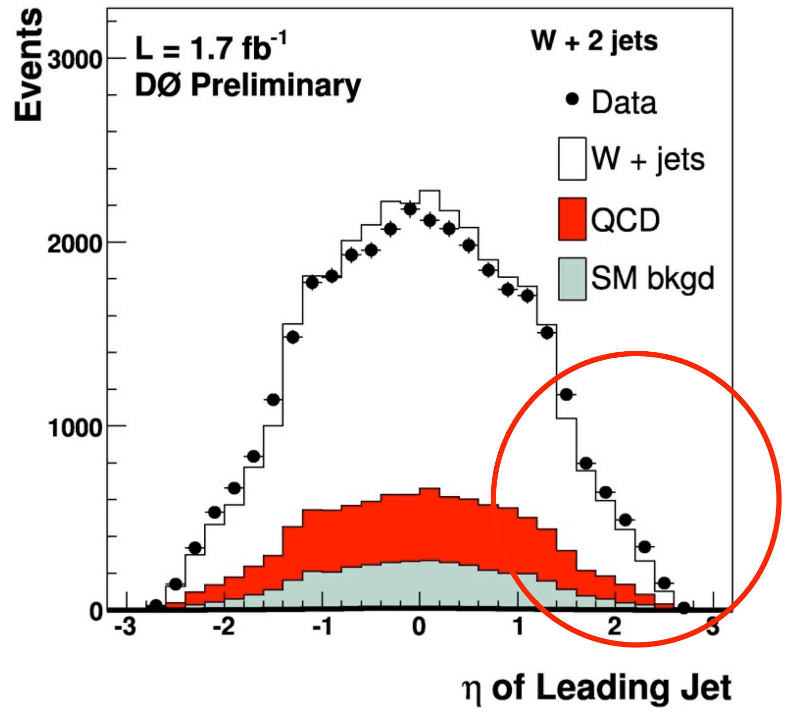
Anecdotes From the Field (III)

- Pile-up events ("minimum bias") do produce jets
  - At high $\mathscr{L}$, require that tracks pointing to jets originate from same vertex as lepton
  - High η excess disappeared!



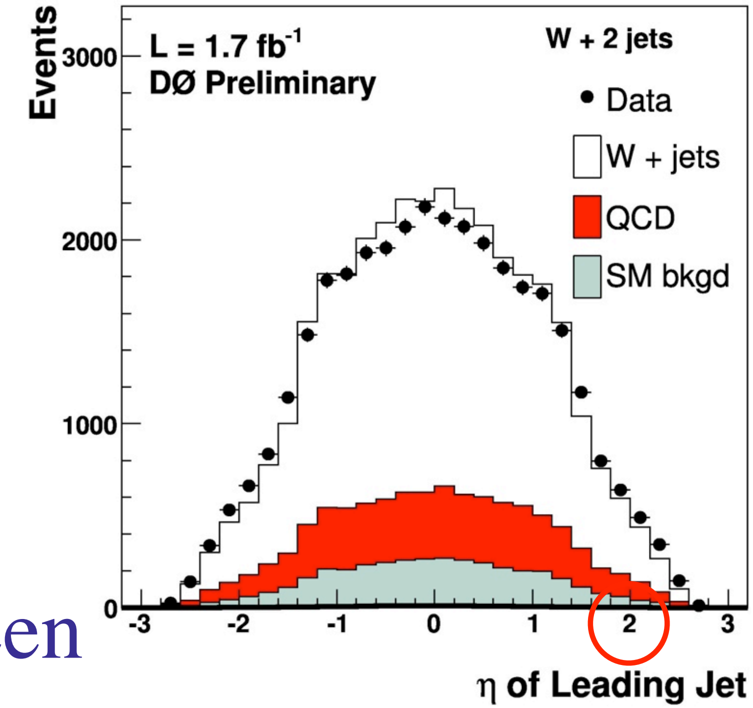
Anecdotes From the Field (III)

- Pile-up events ("minimum bias") do produce jets
  - At high $\mathscr{L}$, require that tracks pointing to jets originate from same vertex as lepton
  - High η excess disappeared!
  - Eta-dependence of jet-vertex match turns out to have shape very very similar to excess
  - After correcting for this, excess is back....

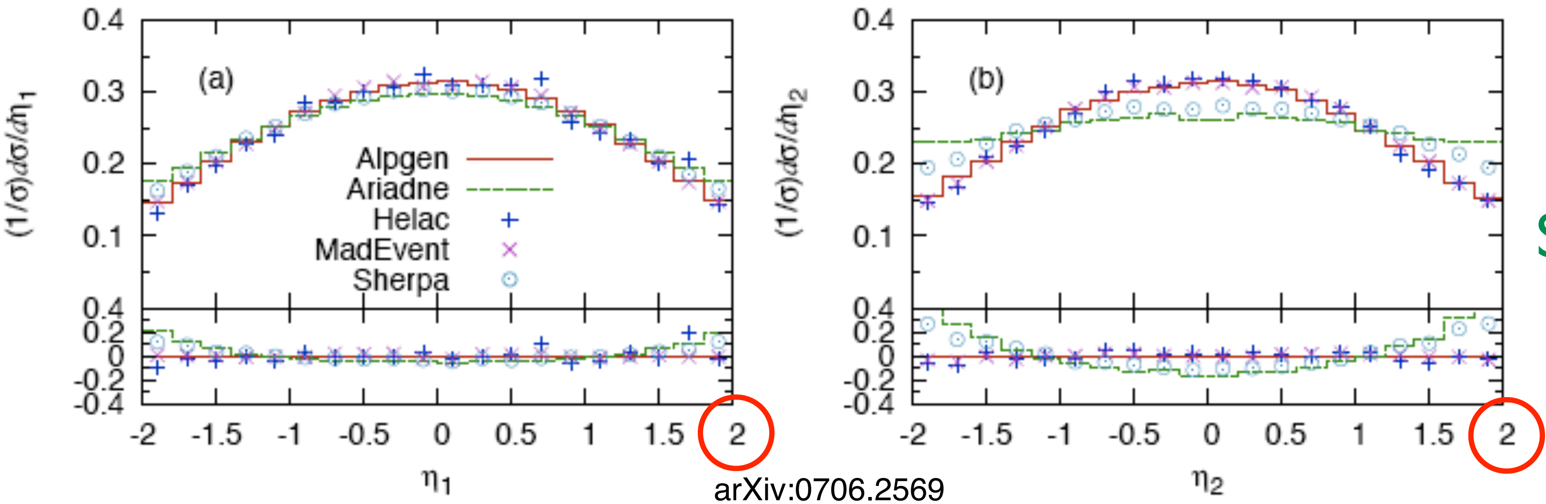


So...

- After all K/K'/S/HF-factors and boson p_T reweighing:



- Similar angular differences between generators: reweigh alpgen to sherpa



arXiv:0706.2569

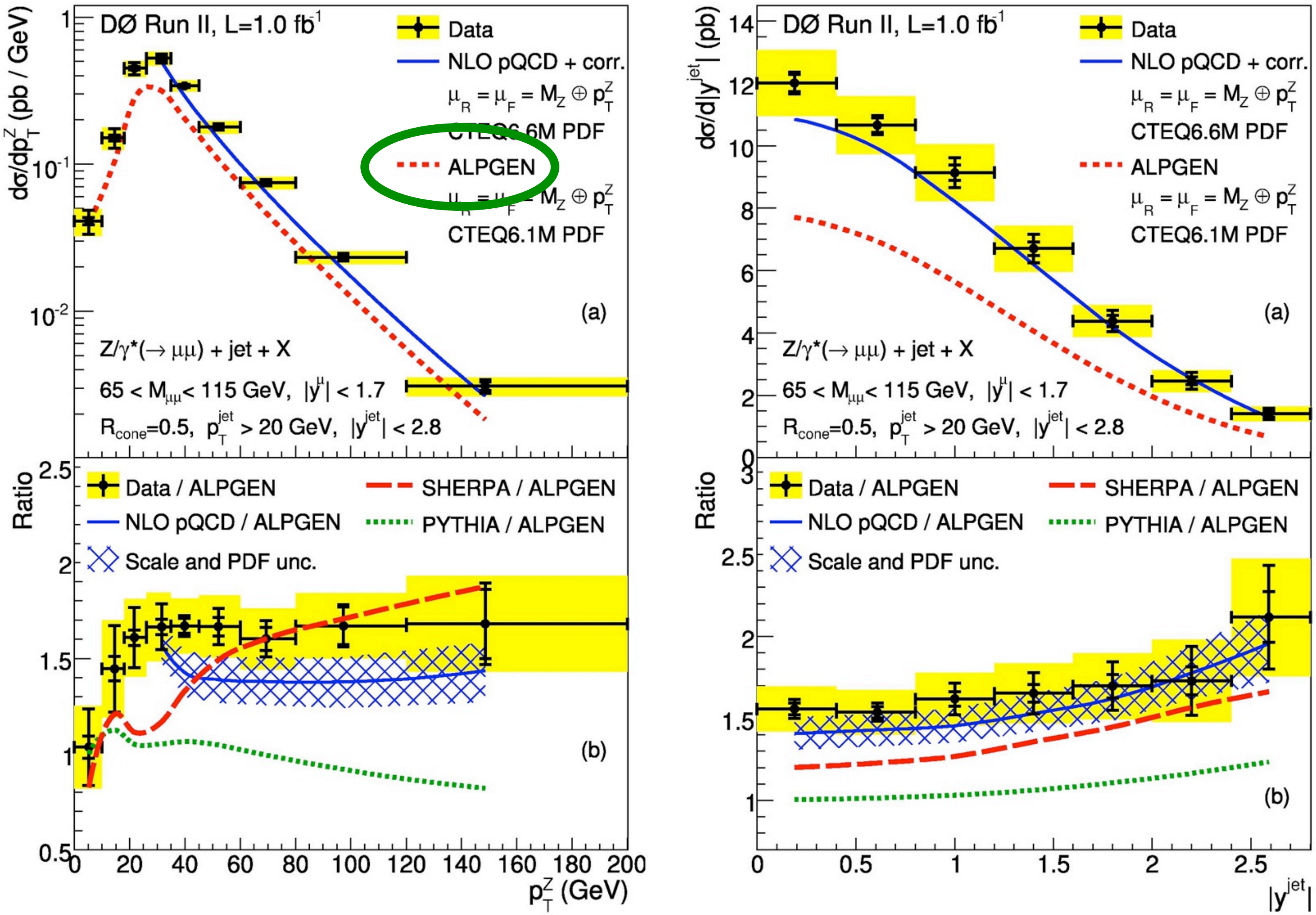
Alpgen, MadEvent, Helac with MLM, Sherpa and Ariadne with CKKW matching

Z (→ℓℓ) + jets

- Can get a clean sample, check if our simulation reproduces the data

⇒ yes, with ~expected deviations

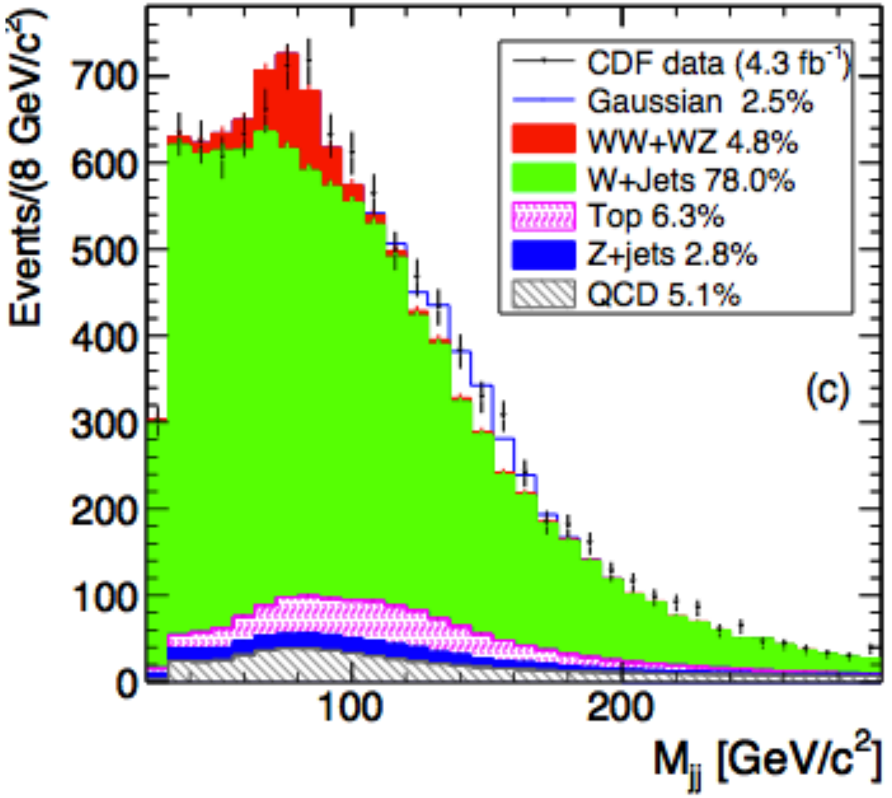
Need reweighing of MC

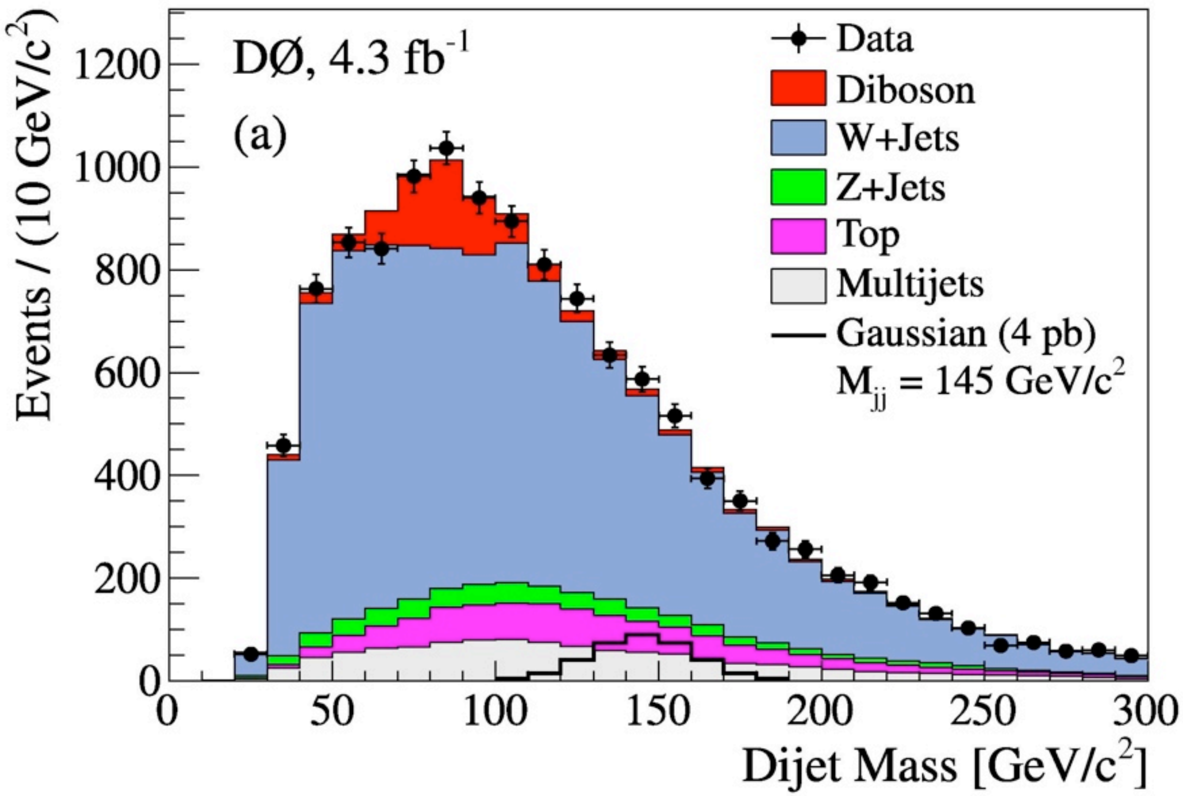


Anecdotes From the Field (IV)

- Searched for WW/WZ in $\ell\nu jj$

Phys.Rev.Lett.106:171801

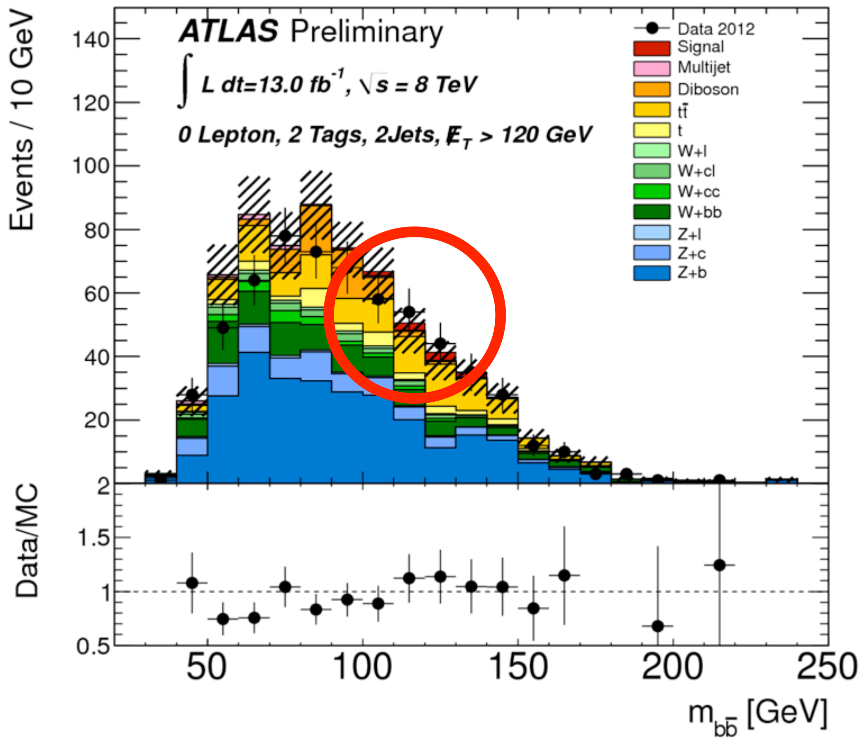




- The background here is not SM, it is uncorrected alpgen!!
  - But this is not the issue.....

Jetty Backgrounds and the Higgs

- Final states with jets can't be ignored: H-b coupling!
  - Many backgrounds
  - All have uncertainties!
- And in addition we have experimental uncertainties!

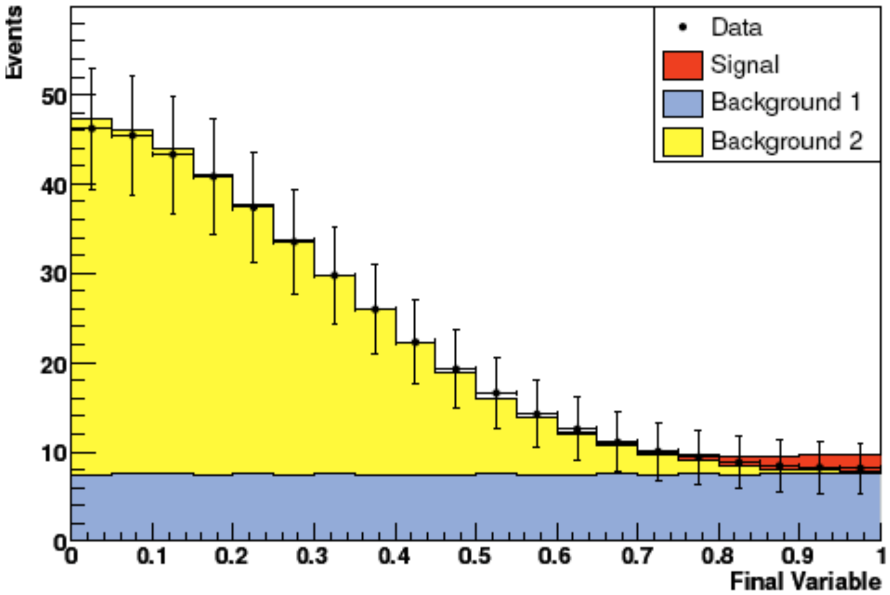


Systematics Profiling

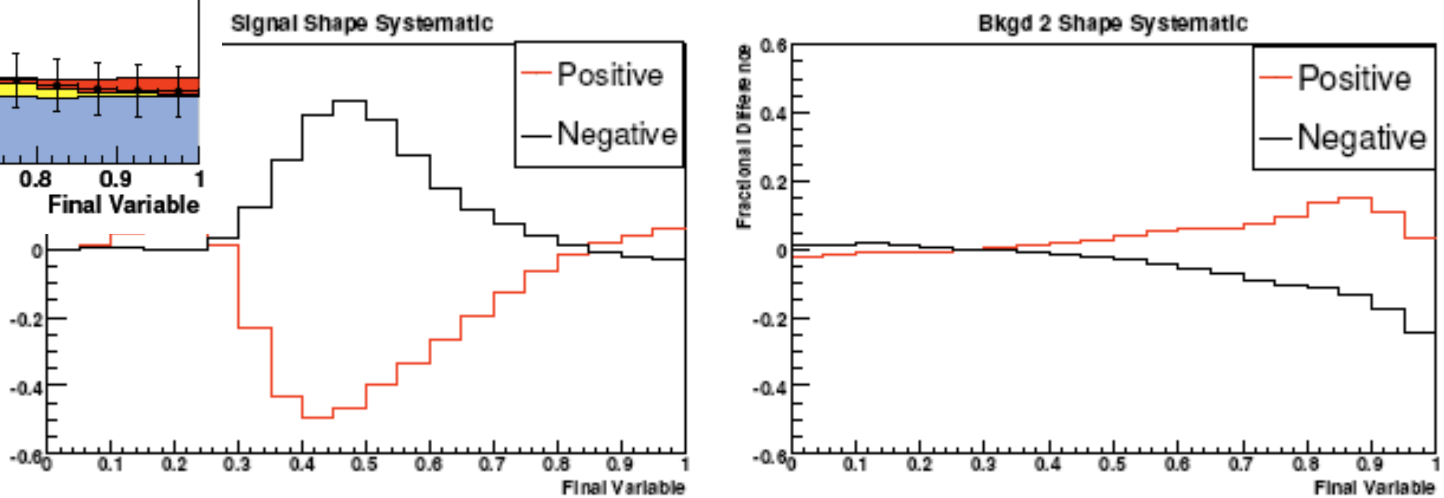
- Systematic uncertainties are propagated through the full analysis chain to the discriminating distribution
  - E.g. we repeat the analysis with jet energy scale shifted up & down by 1σ
  - Some systematic uncertainties affect shape (jet/lepton/photon reconstruction efficiency, energy scale and resolution, p_T distributions), others only normalization (lepton reconstruction efficiencies and momentum calibration, background normalizations, theoretical cross-sections and luminosity)
  - Systematic uncertainties are treated as nuisance parameters

Systematics Profiling

- Nuisance parameters tend to be correlated, but not 100%, among backgrounds
  - Can affect rates, shapes, or both (in any distribution), and often asymmetric and non-gaussian

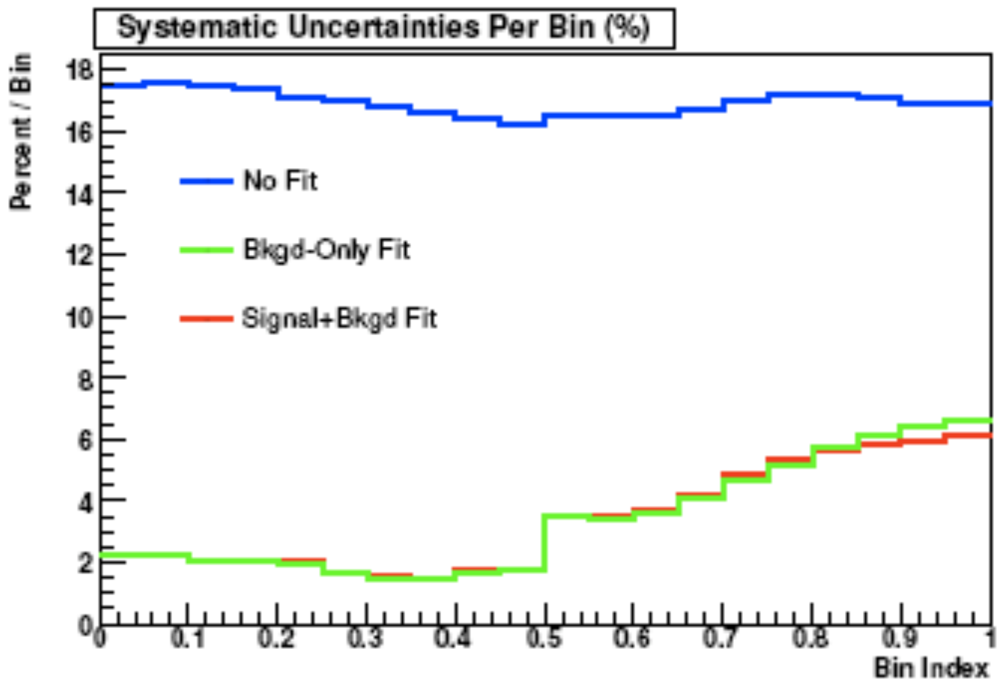


Toy Example (W. Fisher)

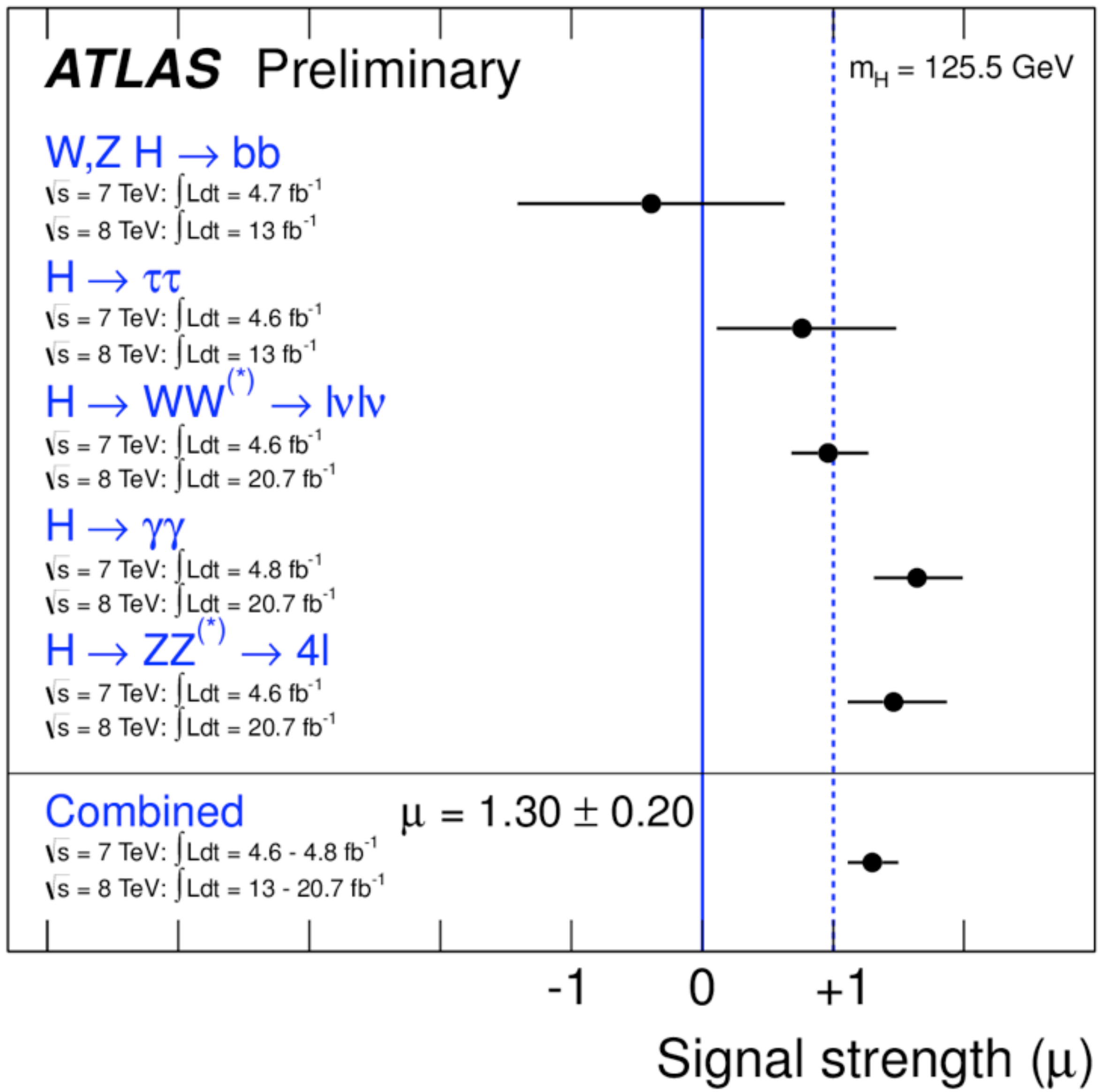


- Generate pseudo-experiments (events in bins according to poisson), then for each experiment vary nuisance parameters
    - Variations in background (& S+B) prediction
        - Compare results to data using log-likelihood ratio
- So we can maximize likelihood ratio as a function of nuisance parameters → constraint them
    - I.e. use full shape of distribution(s) to see which background uncertainties are over/underestimated
        - Of course limited to size of statistical fluctuations
    - Can remove bins with large S/B if needed
        - Mostly important if uncertainties lead to similar shape distortions

- Test example:
  - Data constructed to disagree with background-only hypothesis (wrong estimates for background uncertainties)
  - But to agree with background-only better than signal+ background
    - Improvement quite spectacular (but by construction)



Measure of Difficulty



So, Physics Analysis

- Start from:
  - "*How well* do we understand data *and* the SM?"
    - How confident are we in corrections we apply?
- Given that:
  - Which measurements can we make? What do we need to do to improve our understanding?
- Balance the work!
  - Early, low background searches
  - Detailed understanding/verification of SM predictions

Complementary measurements!

- Increasingly complex searches
  - Tough backgrounds, hard work
  - Don't scorn multivariate and statistical tools

Higgs Drawbacks

- So with the addition of a Higgs boson around 125 GeV particle physics could be "complete"
  - Like Mendeleev's table for chemistry, but not understood. By itself, the Higgs is very unsatisfactory:
    - Why are the couplings to the fermions what they are?
      - Dumb luck (aka landscape)?
    - What is the link to gravity?
    - What about Dark Matter?
    - Why does the Higgs break the symmetry?
    - Why are there 3....?