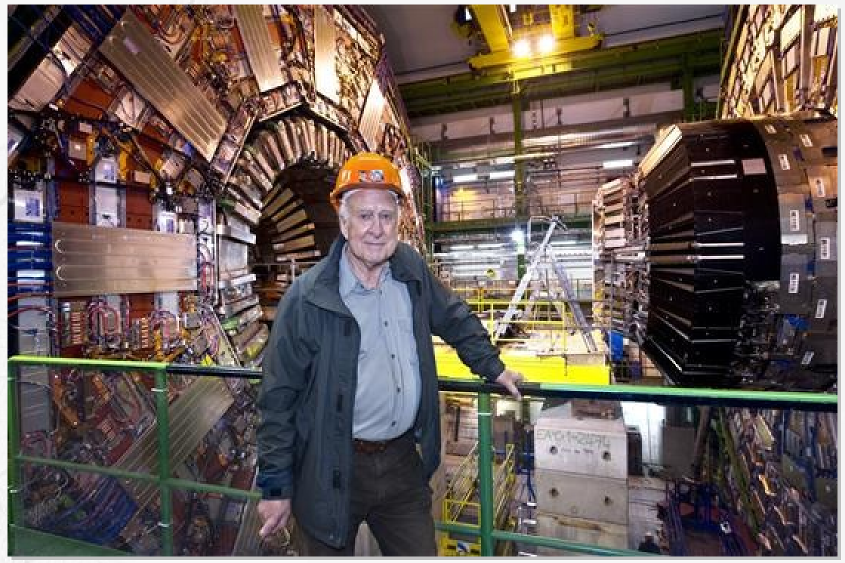
Search for the high mass Higgs with CMS

Alexey Drozdetskiy
for CMS collaboration

With thanks to the collaboration for results of combined efforts presented here.
And to ASPEN'10 organizers for giving the opportunity to make the following presentation

- We kind of seen Higgs at CMS already:

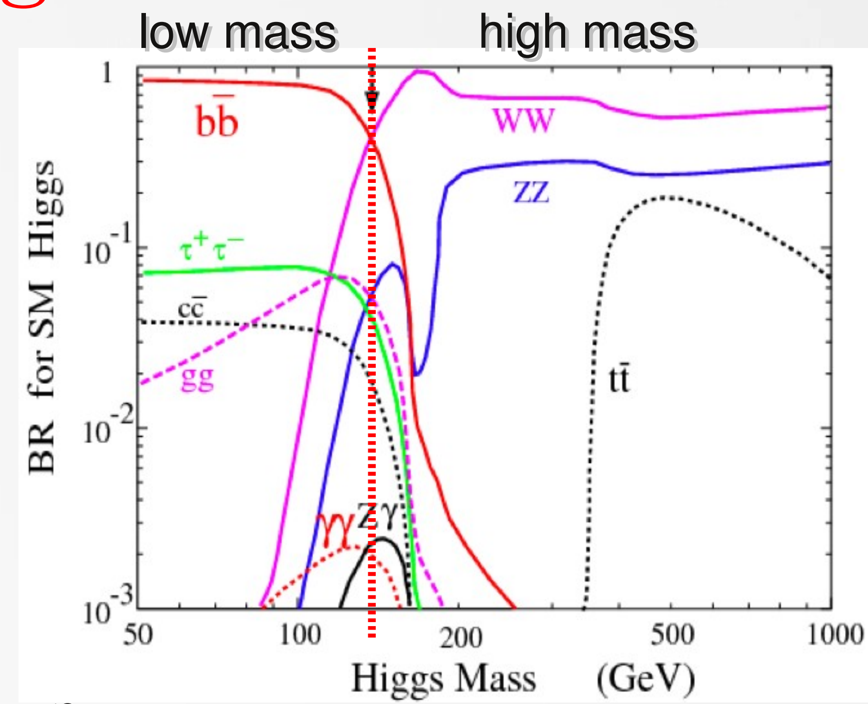


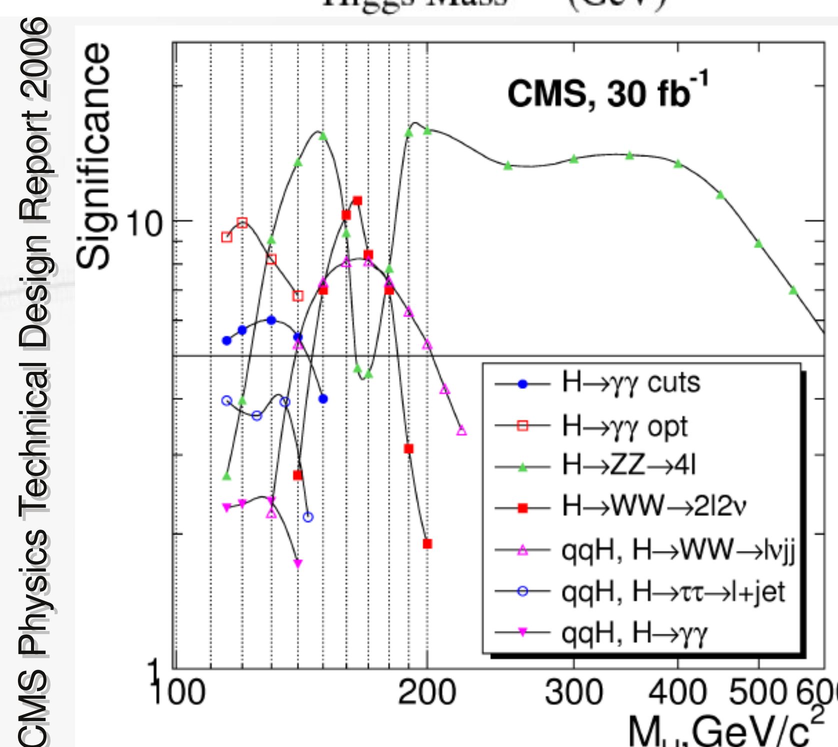
  - ...earlier we excluded the possibility that Higgs mass is less than 114
    - He himself says that with Tevatron's t-quark mass measurements, most likely Higgs mass is about 117
    - We will see about that...

Let's come back now to the Higgs boson, and scenarios where it has (relatively) high mass

High mass SM Higgs at CMS

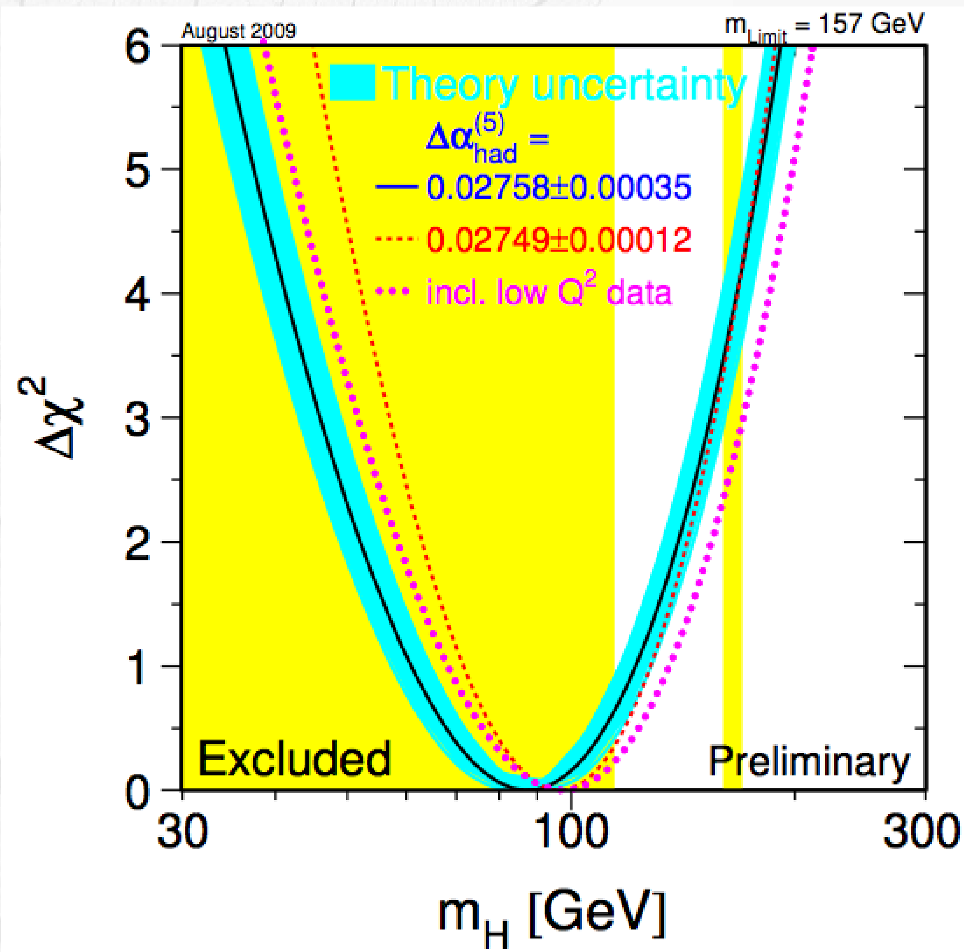
- By high mass Higgs we mean
  - H → WW, H → ZZ (m>130 GeV or so)
    - These are channels which take over from the low Higgs mass dominant channels like bb, 2gamma
- HZZ and HWW are dominant
  - Large branching ratio in wide mass range (**HWW**)
    - Enough expected events to set limits – WW is the leading one to set limits.
  - Signature is clean in **HZZ**
    - This one is 'golden" for discovery for the wide range of possible masses
  - Production:
    - inclusive
    - VBF (~20%) – include for several important reasons we outline below

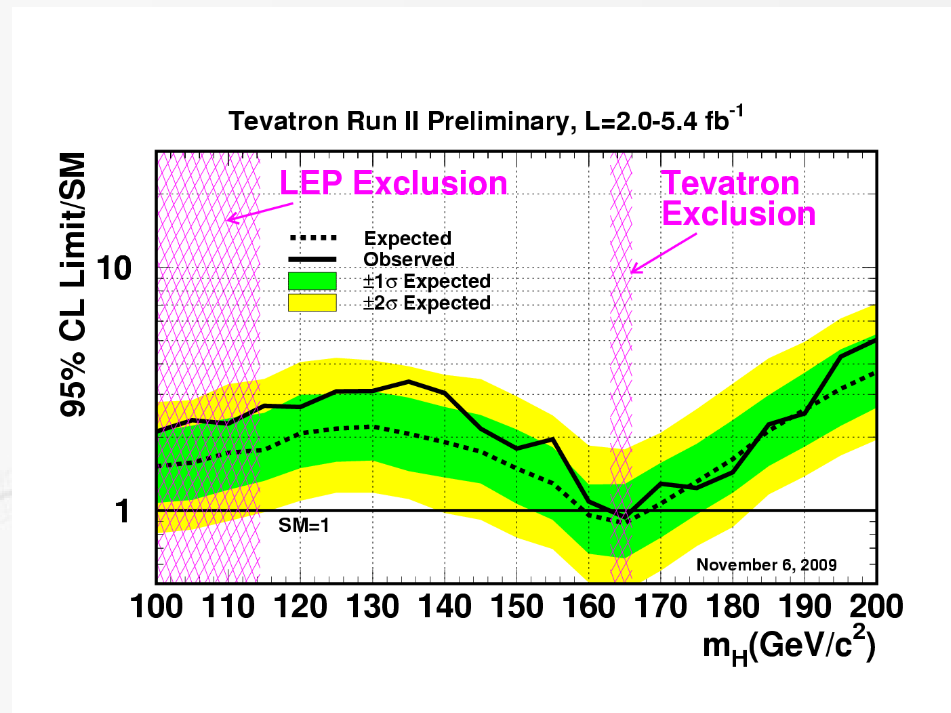




Context

- Theoretical constraints
- LEP limits
- EWK precision measurement constraints
- Tevatron limits: current and future (running experiments)

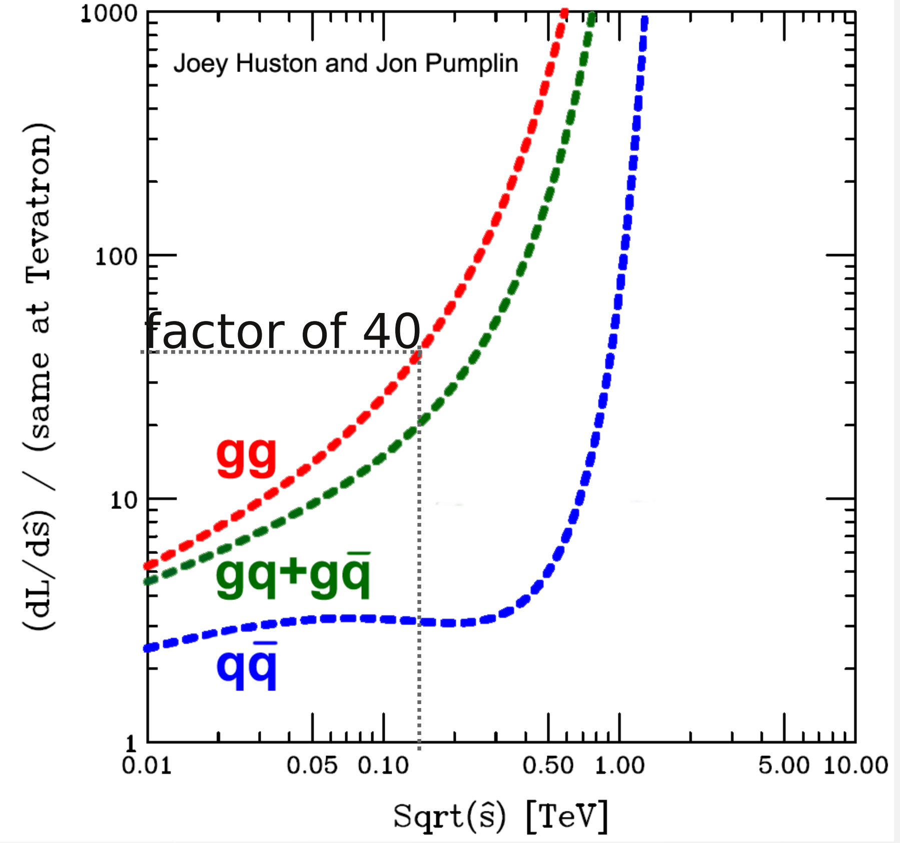




LHC potential with low luminosity

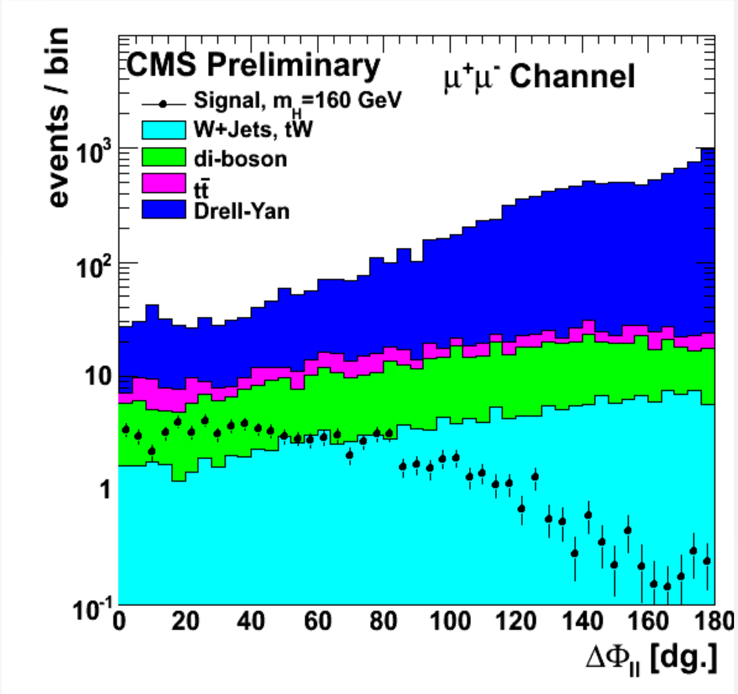
- Main question is – can we "compete" with Tevatron and/or contribute in a meaningful way?
  - Tevatron: ~8fb-1 per experiment, 2 TeV
  - LHC:
    - ~200 pb-1 per experiment, 10 TeV (Chamonix 2009)
    - ~500pb-1 per experiment, 7(?) TeV (current duscussions)
  - As one can see – we would need to gain a factor 40 to be compatible
    - Main factor is: larger cross sections
    - Roughly: we have factor 40 or more for m>150GeV
  - IMPORTANT NOTE:
    - For some channels (like HZZ and HWW) – signal is gg-dominated and main backgrounds are qq-dominated (ZZ and WW)
  - Acceptance, reconstruction efficiency and resolution are more or less similar

10 TeV LHC vs Tevatron



H → WW: analysis

- Production: H → WW → lνl'ν'
  - Final states considered are with l=e,mu. ll' = 2e, 2mu, e+mu
- Event selection
  - Event with two well reconstructed isolated leptons selected
    - PT>20,10
    - If more than two such leptons → rejecting (against WZ, ZZ)
  - Electron-ID cuts applied to reduce W+jets contamination
  - Single lepton triggers used (~90%)
    - Only marginal gain from double lepton triggers
  - Jet VETO: ET>30, |η|<2.5 → reject (tt)
  - Missing ET > 30 GeV (reject Z, QCD)
  - m(ll')>12 (to reject bb)

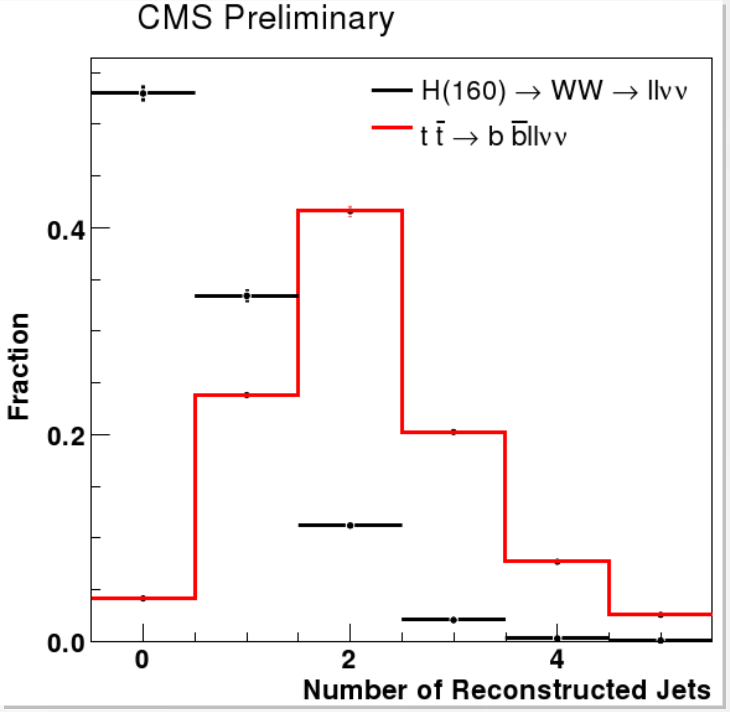


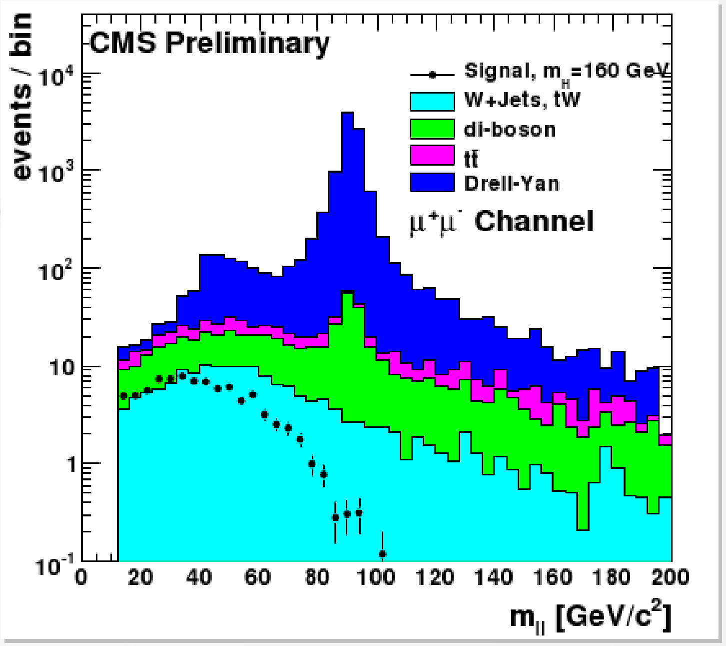
Used to dicriminate against WW (and DY)

- PLUS (for optimization):
  - Opening angle (small for leptons from WW from scalar H): against WW continuum
  - Upper cut on m(l+l-): against Z, WW
  - MET (upper and lower bound)
  - PT of both leptons

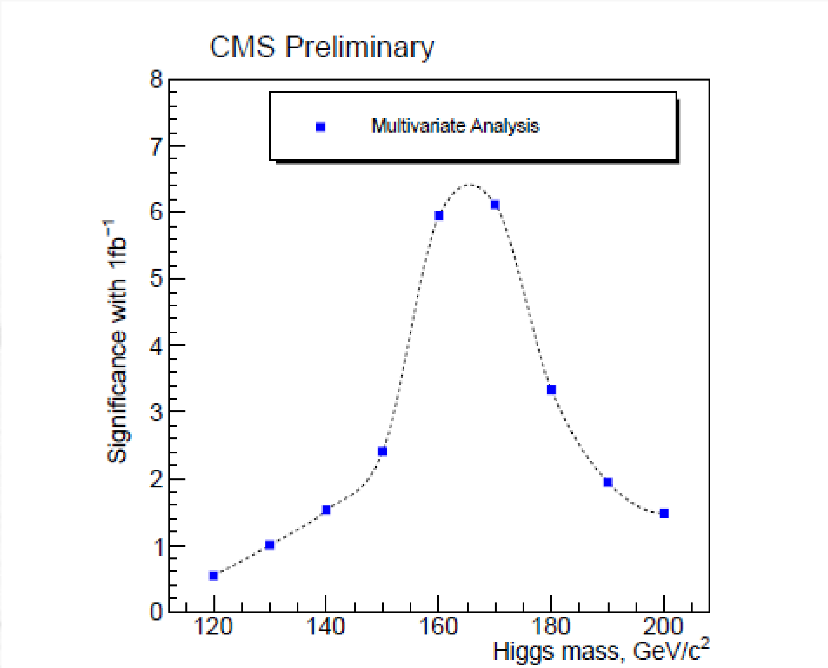
HWW: background control

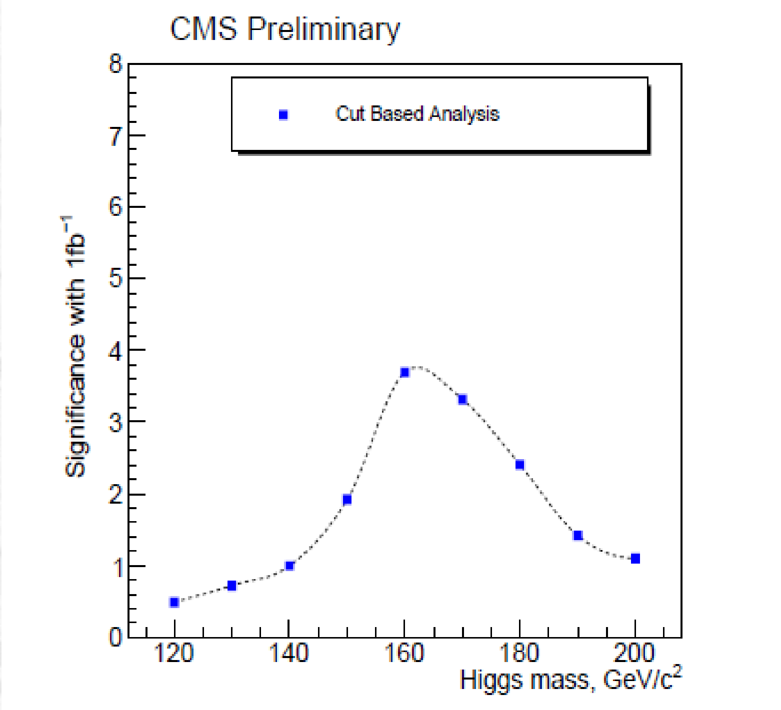
- Main backgrounds are WW, tt, W+jets, DY
  - This is for 0-jets channel (only public so far)
  - Pros for 1 or more jets in the final state
    - Factor of ~2 larger cross section w.r.t. LO
  - Cons:
    - Large tt contribution
- WW
  - Is controlled from small m(ll) region
- TT
  - Is controlled from n(Jets) >= 2
- IMPORTANT ISSUE:
    - We use NLO cross section and re-weighting PT of Higgs according to NLO spectrum. But signal yields/efficiencies we get are calculated for LO process, with no NLO, NNLO jets... So, factor ~2 we are "gaining" (w.r.t. LO) may be significantly smaller...



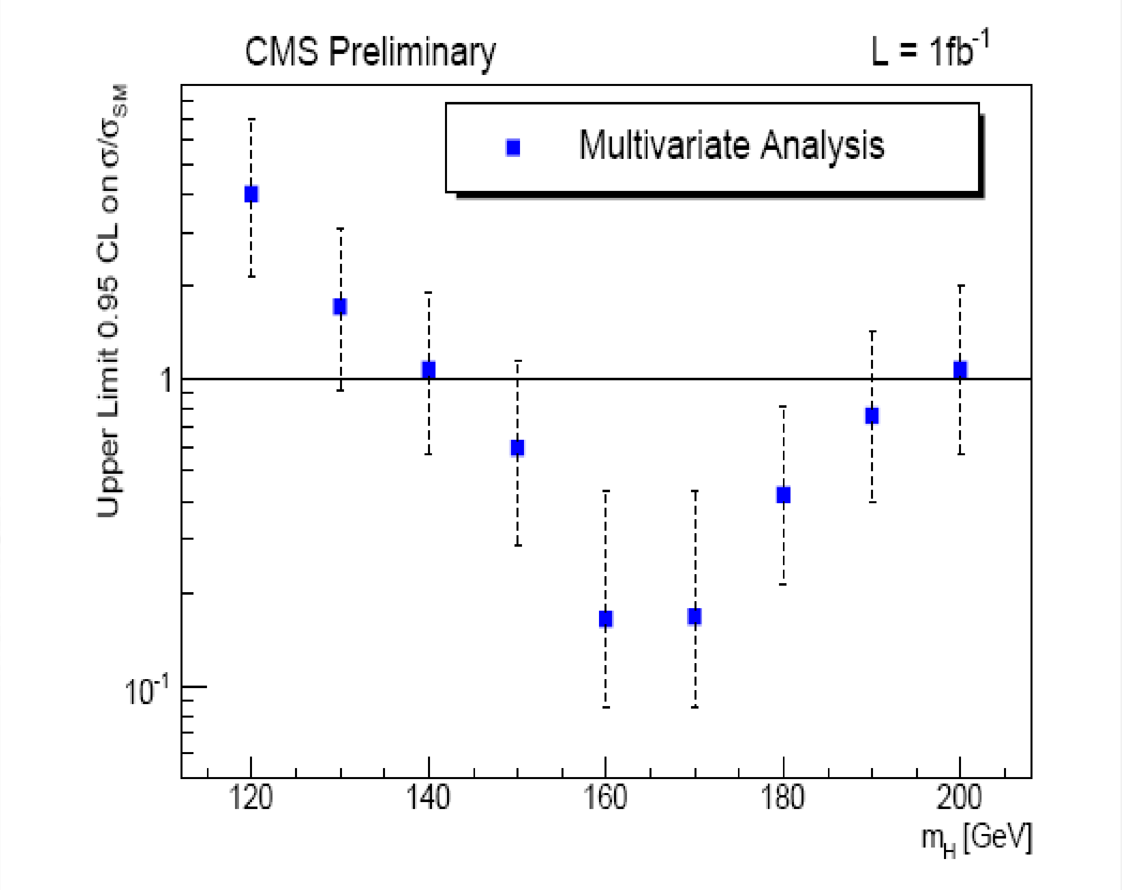


HWW: results





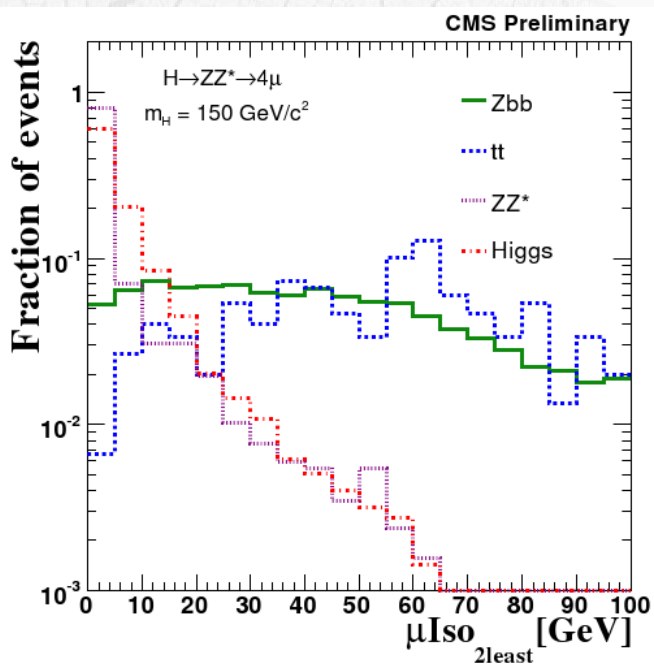
- Multivariate analysis performs way better than cut based
  - Results are for 1fb-1

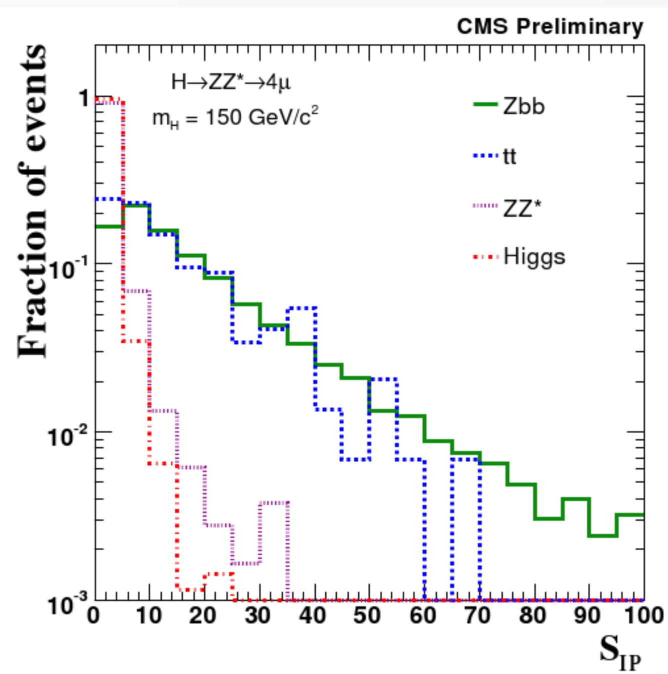


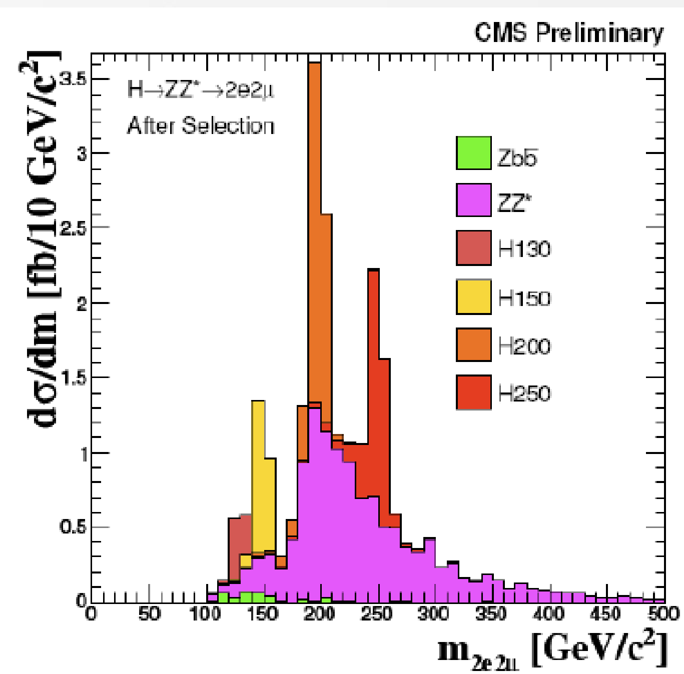
H → ZZ: analysis

- Event selection
  - PT>10,5,5 (skim: against QCD)
  - Single/di-lepton triggers used (100%)
  - Event with two pairs of well reconstructed (isolated) leptons selected
  - Two m(l+l-)>12GeV to fight multi-jet events (QCD, W/Z+jets, tt, Zbb)
  - m(4l) > 100 GeV (not excluded region)
  - Loose track-based isolation (~99% signal)

- PLUS (for optimization):
  - Isolation for all 4 leptons
  - Impact parameter
  - m(4l), m(l+l-), m(l'+l'-)
  - PT of leptons

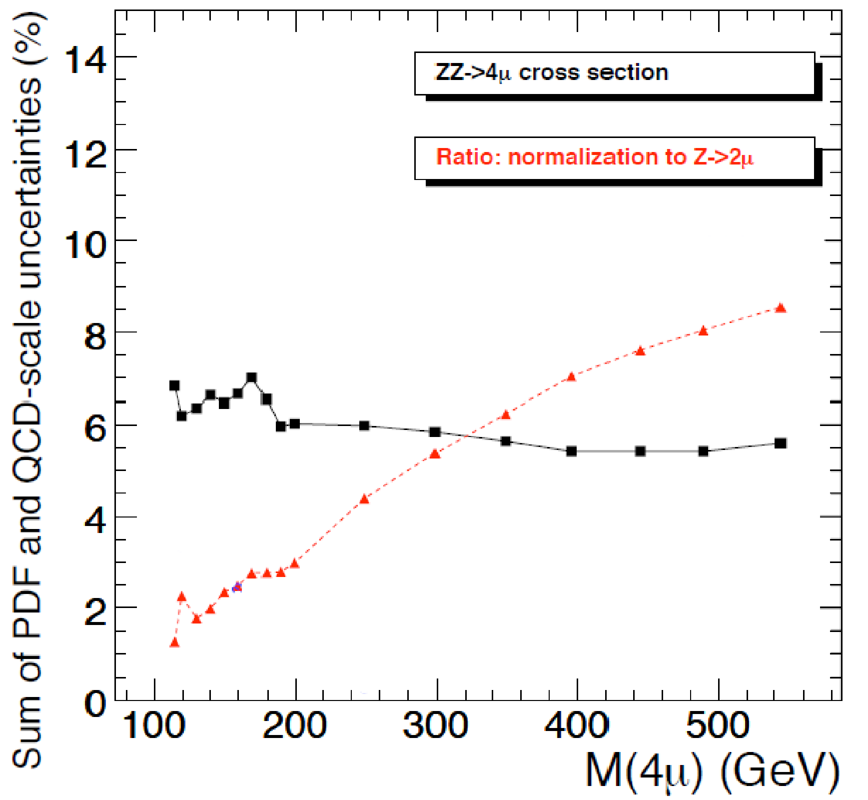






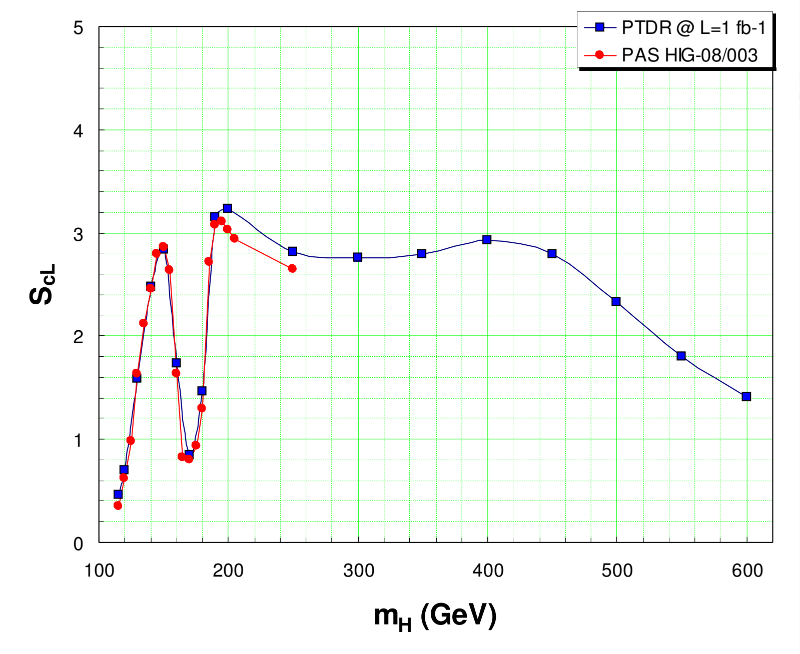
HZZ: background control

- ZZ → 4l – is the main background
  - Control it using ZZ/Z+jets ratio
    - The other method would be through side-bands – but that wouldn't work at the discovery times: too small statistics. On the left is plot from PTDR, for 30fb-1 when statistics was not an issue.
- gg → ZZ (box diagram)
  - Gives about 15% on top of ZZ @NLO
    - With large (~30%) uncertainty
- Zbb
  - Low statistics, hard to come up with data-driven method
    - But it is a minor background
- Other backgrounds:
  - Tt, W/Z+jets, WZ, QCD – should be even smaller



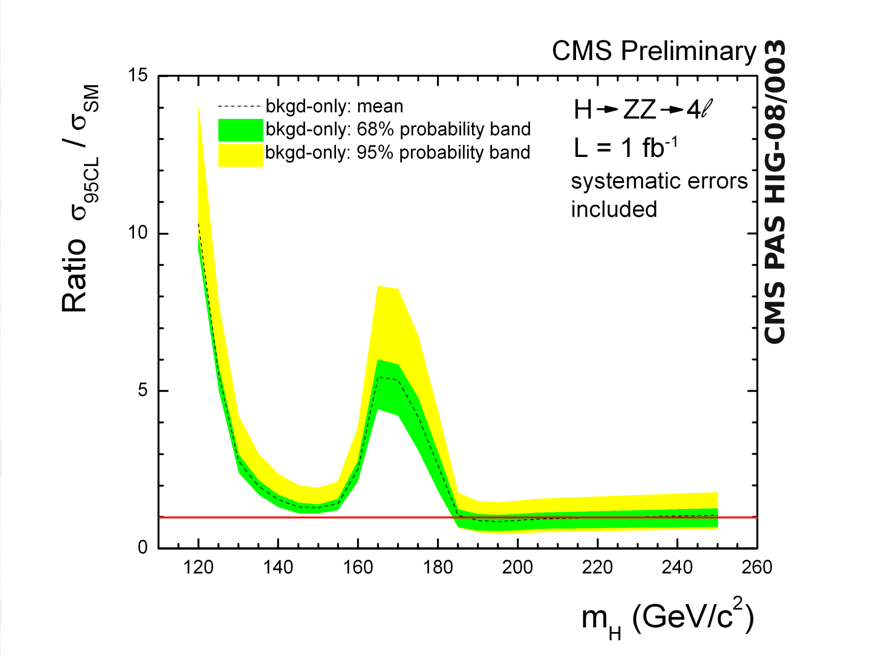
A simple minded estimation for side-bands:
200pb-1 → 3 ZZ events in total which would give approx. 70% uncertainty on prediction from side-bands...

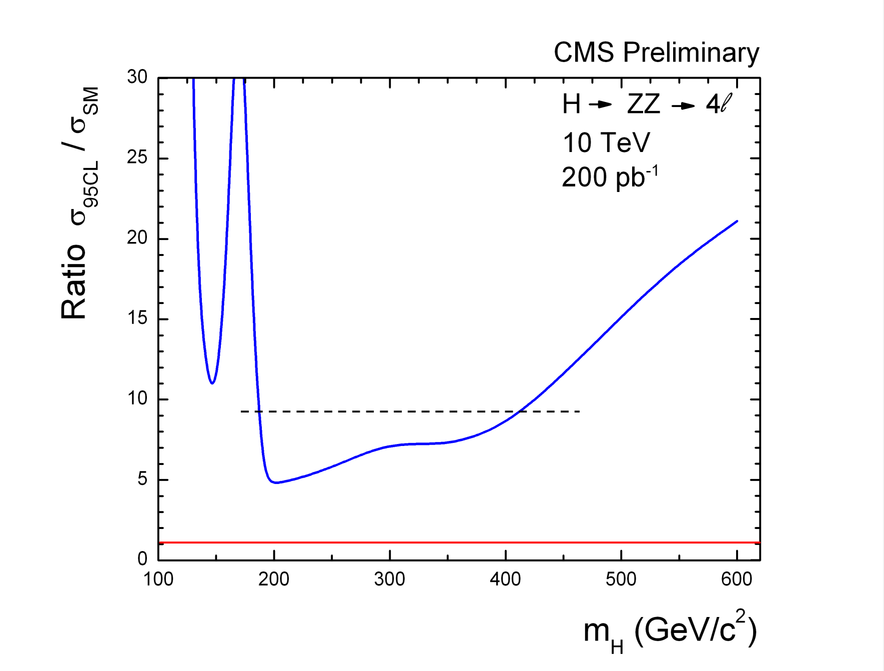
HZZ: results



CMS PTDR-2 scaled down to 1fb-1
CMS PAS HIG08/003 optimized
for mH=150GeV

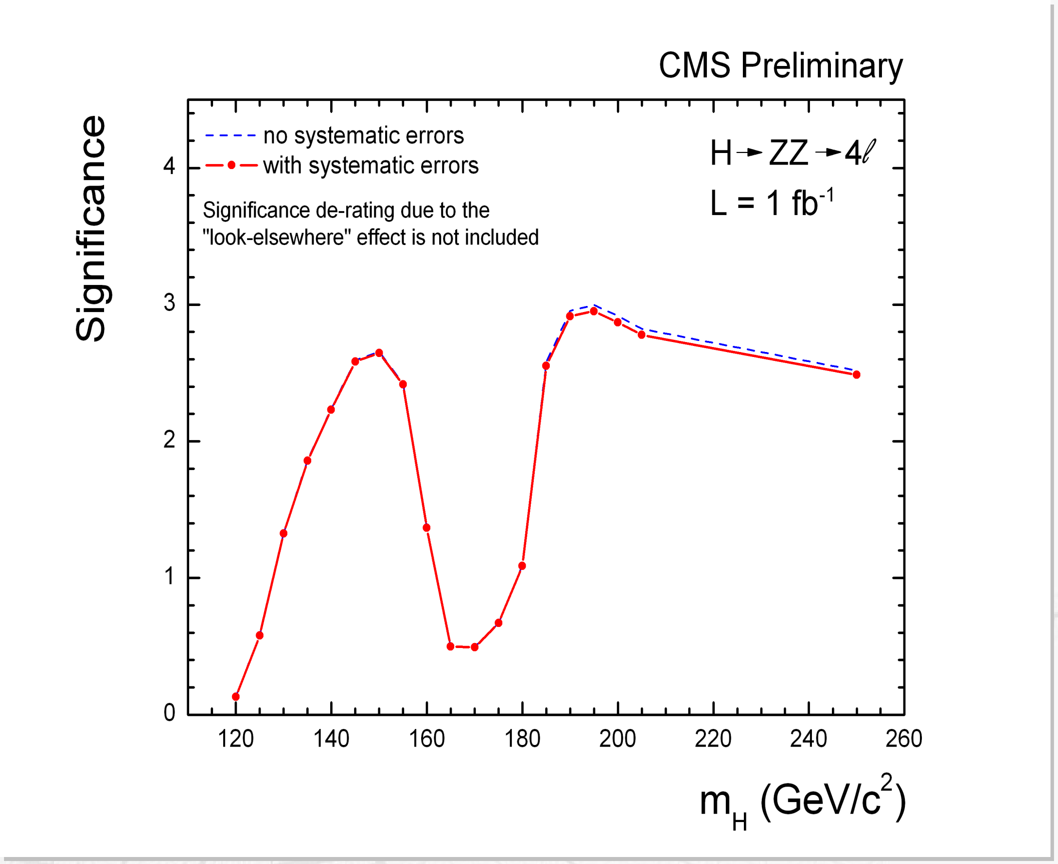
- Scale down PTDR signal and background events to 10 TeV and 200 pb-1:
  - ~3 bkgd events in the full spectrum
  - 0.7 signal events in the best m~200 GeV
- Unofficial projected exclusion limits
  - R<9 in the range 180-400
  - Meaningful in the context of 4 generations
  - High mass range is better than Tevatron

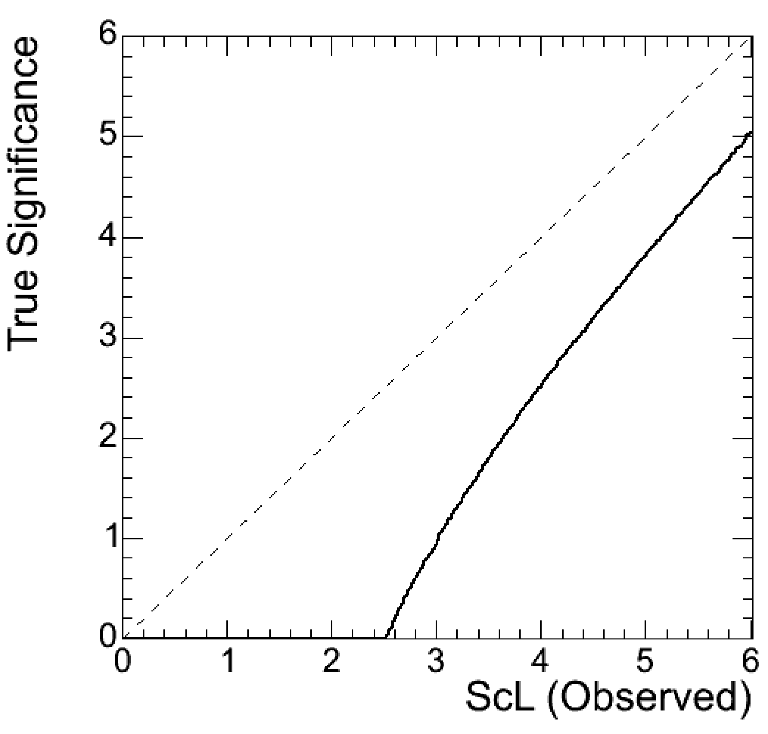




HZZ: results

– Plus: **Look Elsewhere Effect!**
  - Results for “evidence” reach of 2-3 sigma for ~1fb-1 decrease to the level of statistically insignificant bump...

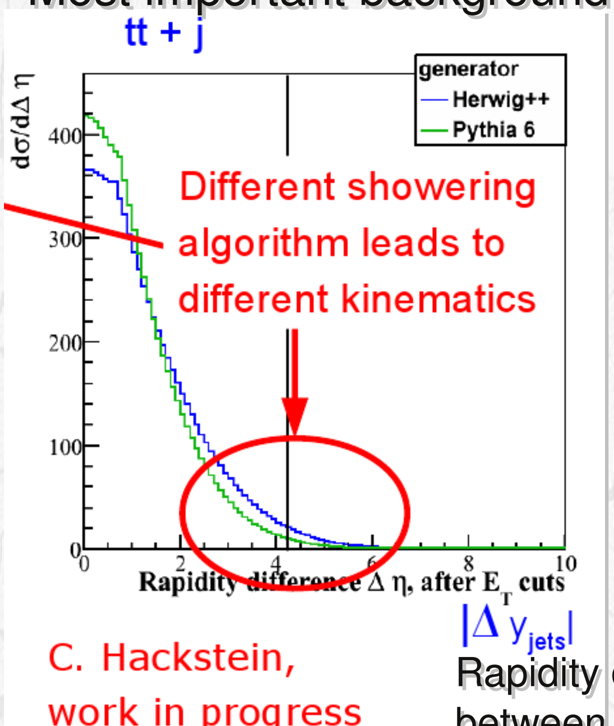




A few remarks on VBF

- It is not a forerunner channel, but it is important
  - Important for disentangling of Higgs couplings
  - Fermiophobic Higgs scenario: VBF becomes the most important mechanism
  - If we do discover SM Higgs: non-resonant VBF WW will confirm that the Higgs does what we expect it to do for WW → WW scattering. If EWK symmetry breaking would occur at ~1TeV scale, non resonant VBF WW may be the first sign of it
  - Problems: large uncertainties in signal and backgrounds → hard/impossible to make exclusions or discoveries...

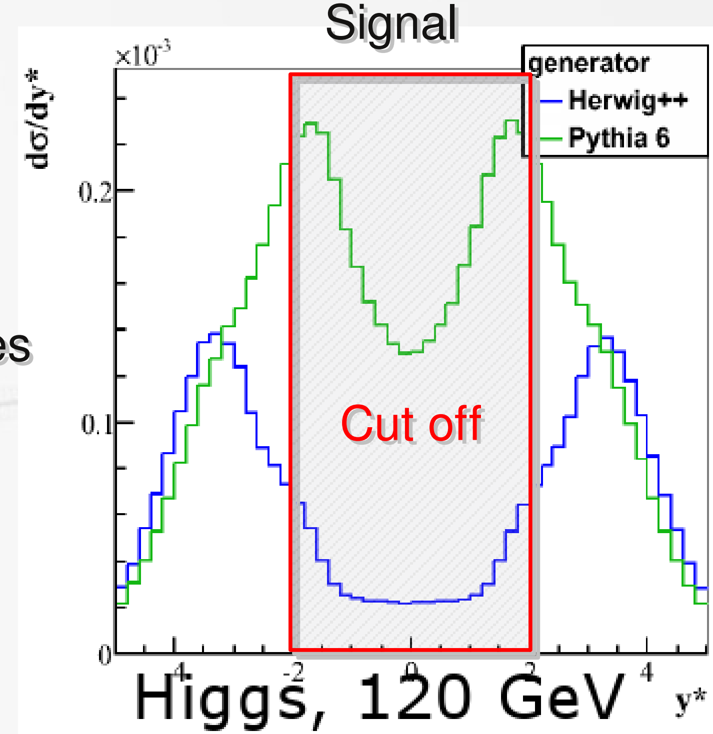
Most important background



$|\Delta y_{jets}|$ Rapidity distance between tag-jets

Additional hadronic activity in EW processes differs extremely for different generators...

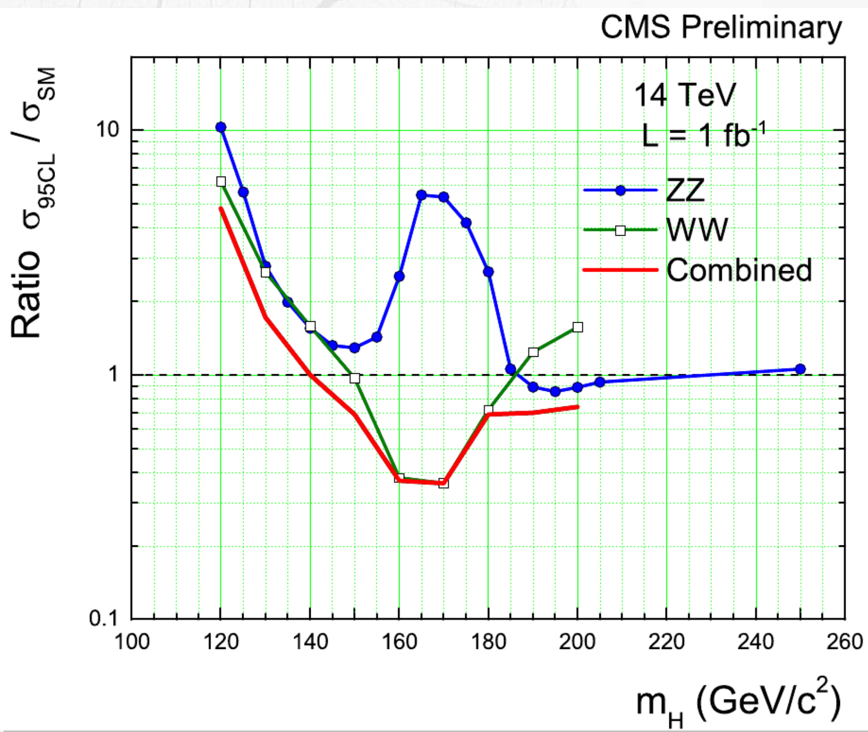
Signal

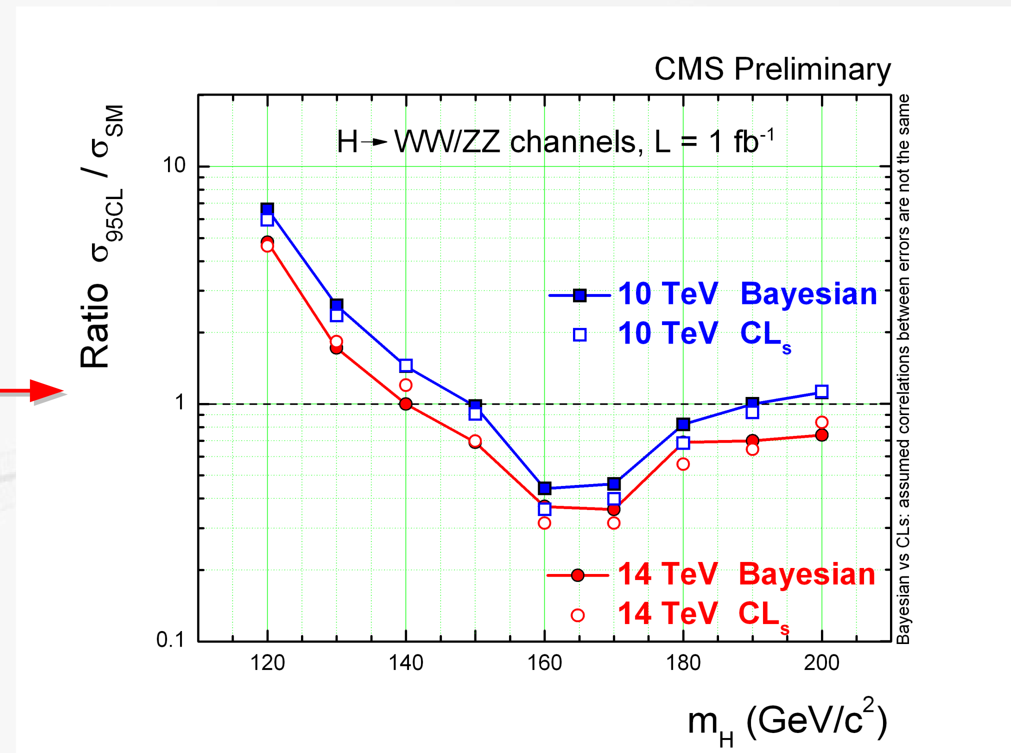


y*=yj-0.5(y1+y2) – this is a measure of how central any additional jet is...

HZZ and HWW combination

- 14 TeV, 1fb-1
- Statistical tools
  - We use Bayesian and modified frequentist (CLs) approaches
  - In both cases we use approximations to do a quick estimation
    - For example assuming 100% correlations between systematic uncertainties (for Bayesian), flat prior for Higgs, etc.



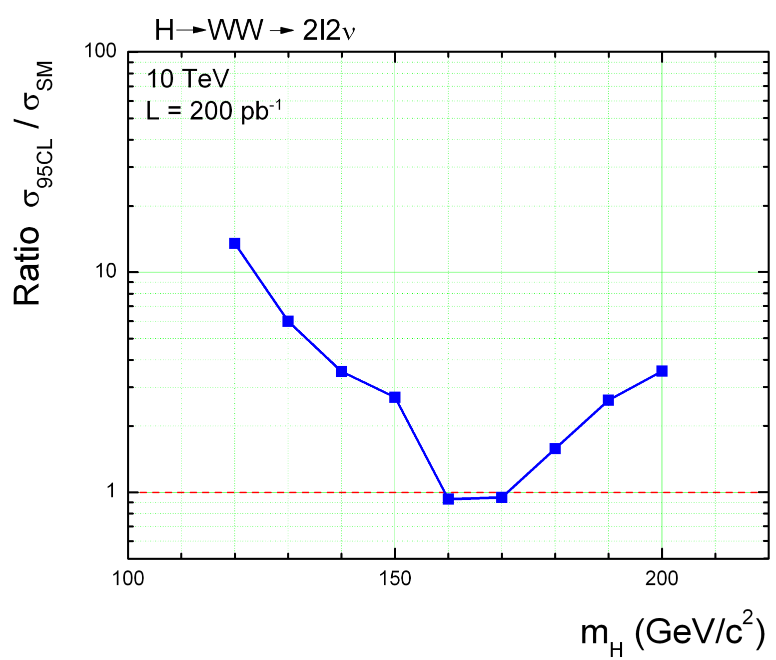


10 TeV, 200 pb-1

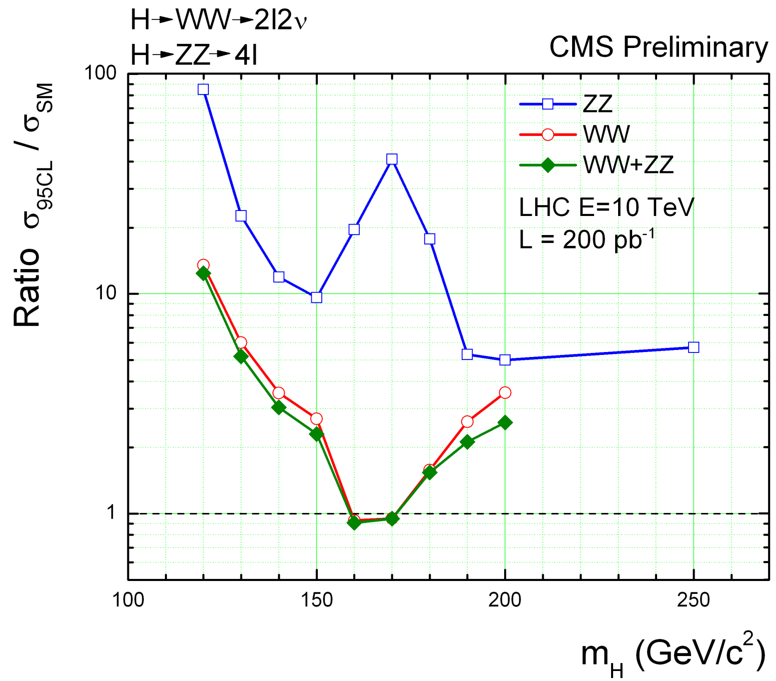
- ...one should scale even further to get to point relevant for 2010 LHC runs. HWW:

  scale PAS HIG-08/006
  signal and bkgd events down
  to 10 TeV and 200 pb-1
  - after a cut on the MVA output for mH=160 GeV
  - 4 bkgd events
  - 7.3 signal events

  Projected exclusion limits
  - r<1 in the range 160-170 GeV
  - r<9: ... - 200+ GeV (meaningful in the context of 4 generations)
  - comparable to or better than CDF/D0 above 150 GeV

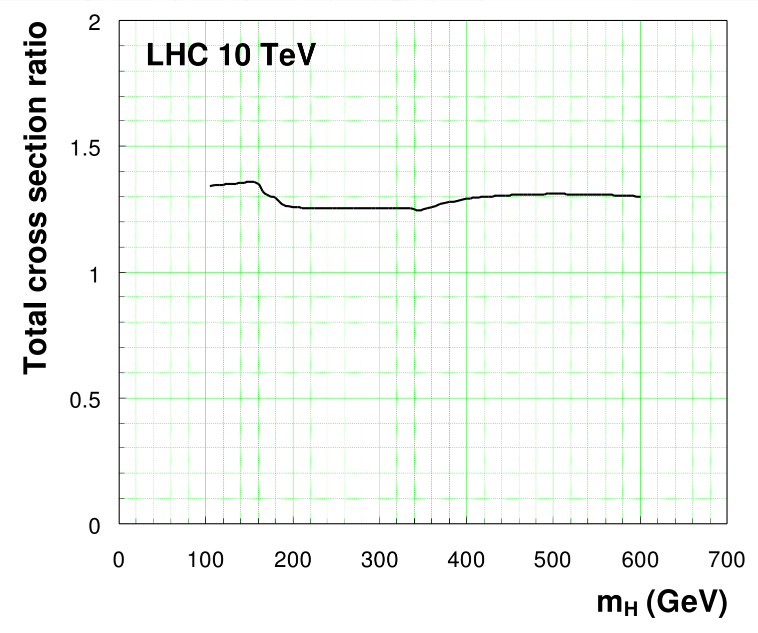


- And then combination: HWW+HZZ:
  - We end up with more or less current Tevatron exclusion interval



To end on a positive note…

- Tevatron experiment use NNLO, CMS – was not so far (for all the results above)
  - Recent calculations using HggTotal by Fabian Stoeckli
  - NNLO contribution to gluon-fusion: ~30% - “for free” :-)



$$ratio = \frac{gg(NNLO)+VBF(NLO)+VH(NLO)+ttH(LO)}{gg(NLO)+VBF(NLO)+VH(NLO)+ttH(LO)}$$

Conclusions

- 200pb-1 @ 10 TeV (500 pb-1 at 7 TeV) will put LHC experiments on a map of Higgs searches
  - We are looking forward to the challenge!