

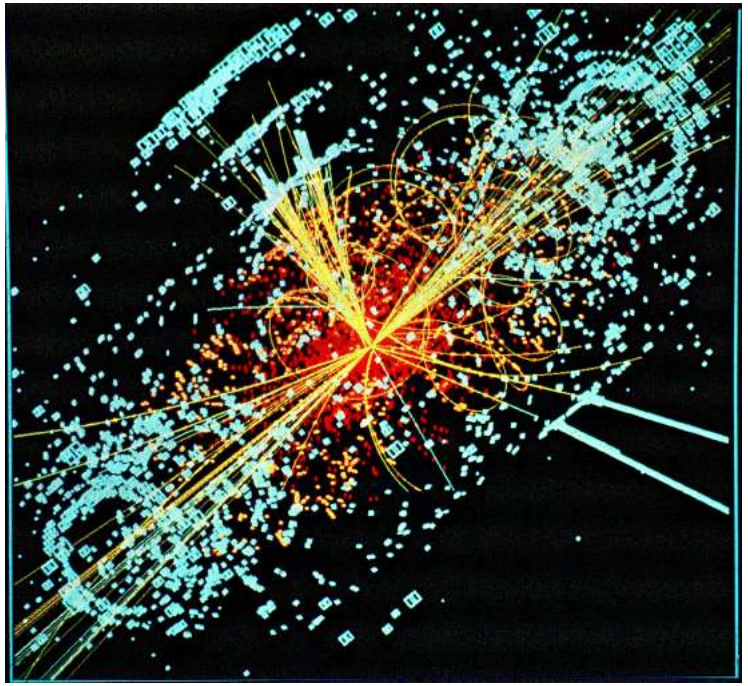
Investigation of Low-Data Higgs Boson Discovery at the LHC

Cheyne M. Scoby
Department of Physics
UC Santa Barbara

Advisor: Prof. David Stuart

June 7, 2006
3:00-4:00 PM
Senior Thesis Presentation

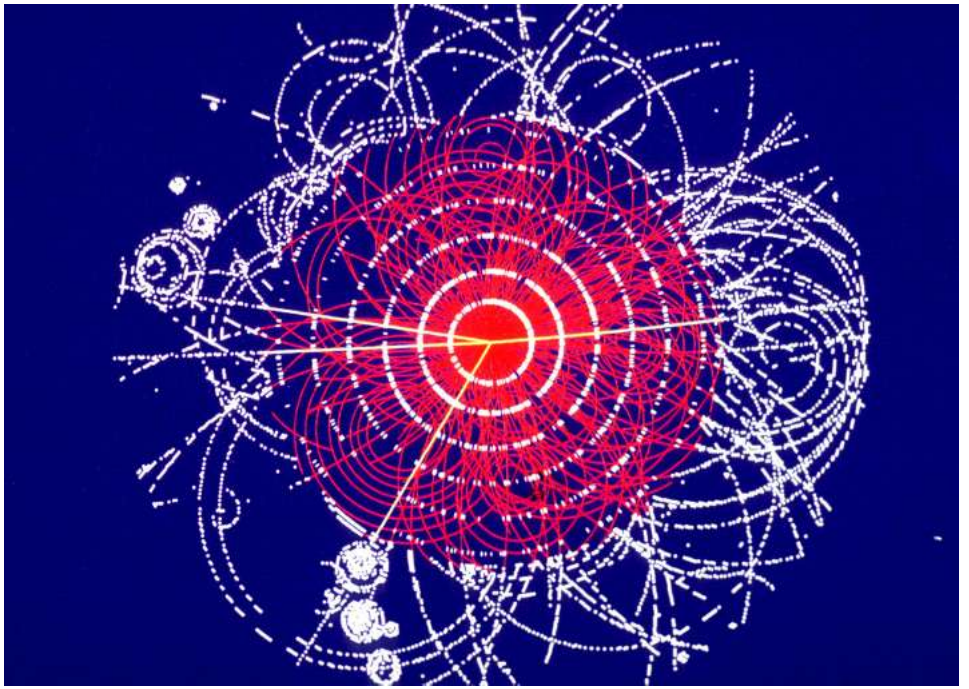
Outline



- The Standard Model
- The Higgs boson
  - A little theory
  - Previous searches
  - Production in a hadron collider
- The Large Hadron Collider
- The Compact Muon Solenoid
  - Detecting the Higgs in CMS

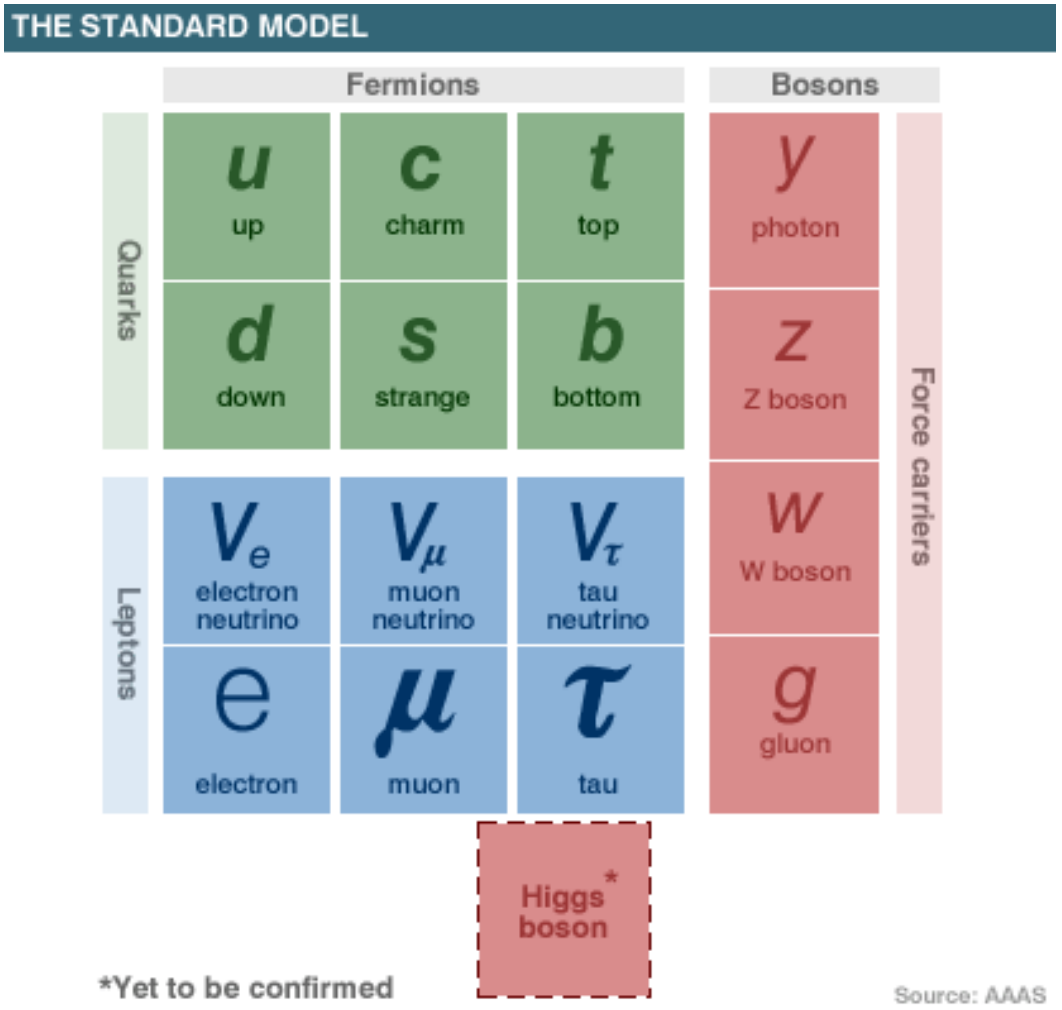
More outline

- Monte Carlo event simulation
  - Cross-section calculations and PDFs
  - PYTHIA and data generation
- Analysis
  - Full reconstruction
  - Single Z kinematics
  - Three-lepton kinematics
- Complications
- Conclusion



The Standard Model (SM)

Our complete understanding of fundamental particles and their interactions.



- This is why the Higgs boson is one of the big questions that the next generation of particle physics experiments must answer.

The Higgs boson

- The Higgs boson is the only Standard Model (SM) predicted particle that has not been observed.
- Explains why the weak vector bosons (Z,W) have mass, while the photon has zero-mass. (Spontaneous electroweak symmetry breaking). The charged fermions (quarks, leptons) also become massive.
- Previous searches have excluded the Higgs mass to greater than 115 GeV/c^2. For EW theory to hold, the mass must be less than 1000 GeV/c^2.
- Too heavy to observe at today's most energetic accelerator (proton-antiproton at 1.98 TeV)
- NB: the Higgs mass is NOT KNOWN!

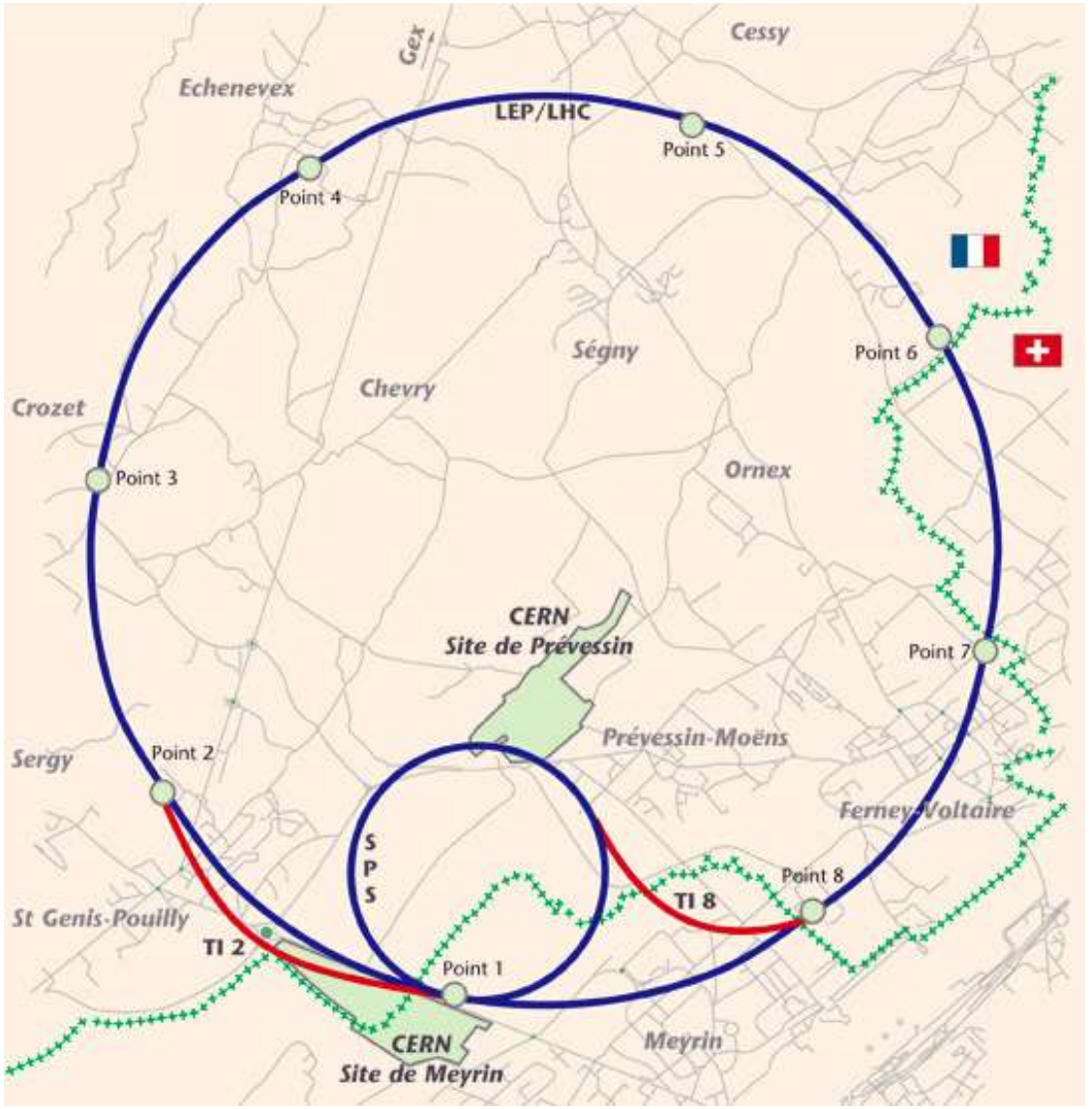
Large Hadron Collider

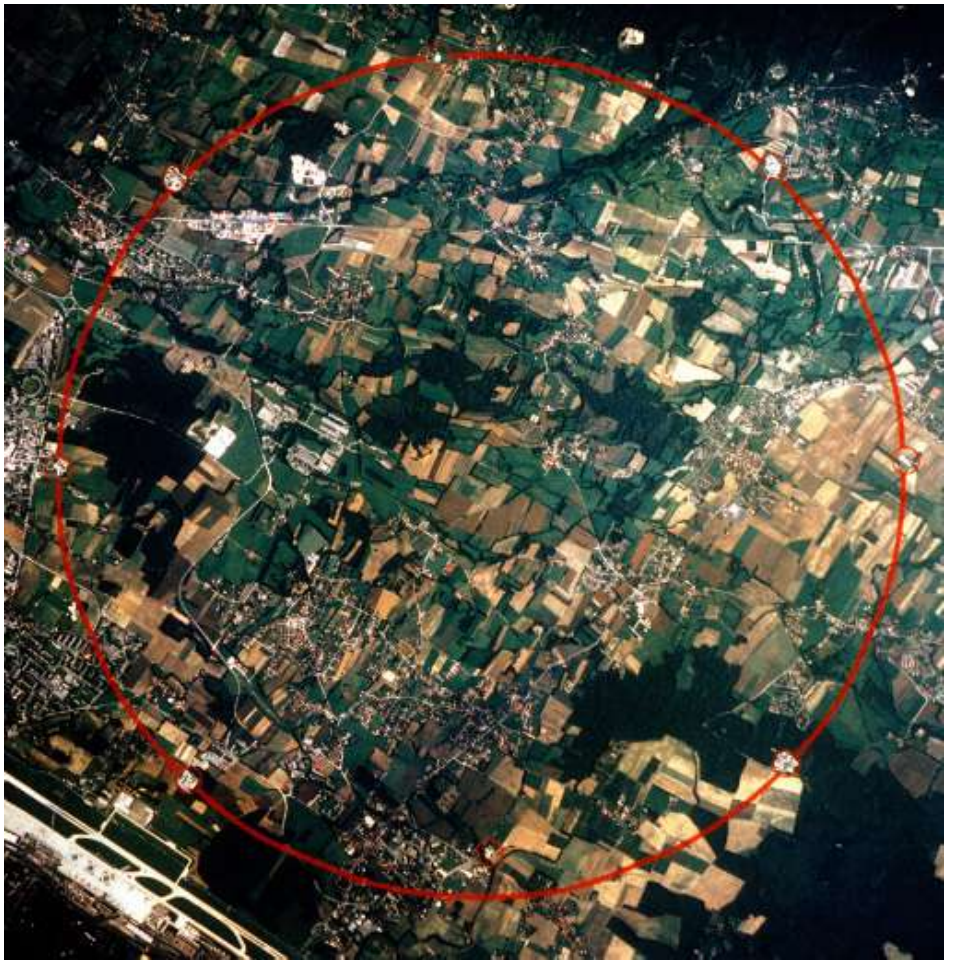
- New accelerator at CERN laboratory in Geneva, Switzerland, buried 100 m below the French-Swiss countryside.
- Proton-proton collider with CM energy of 14 TeV. (Protons at 0.99999999*c*)





- The 27 km circumference tunnel is lined with 1230 superconducting dipole magnets.

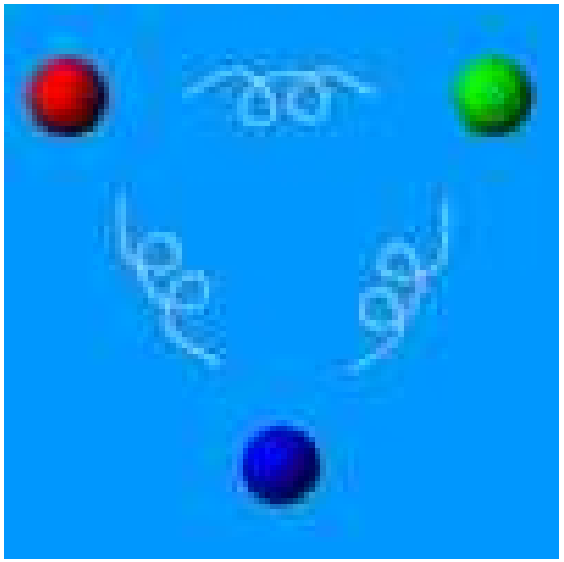




Higgs Production at the LHC

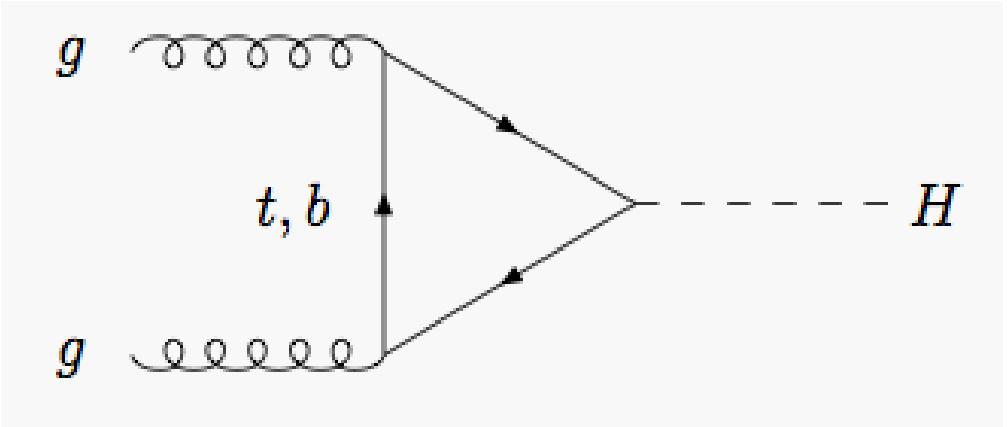
- How do you get a Higgs particle from two protons?
- Protons are composite particles, made of quarks held together by gluons. These components are what interact!

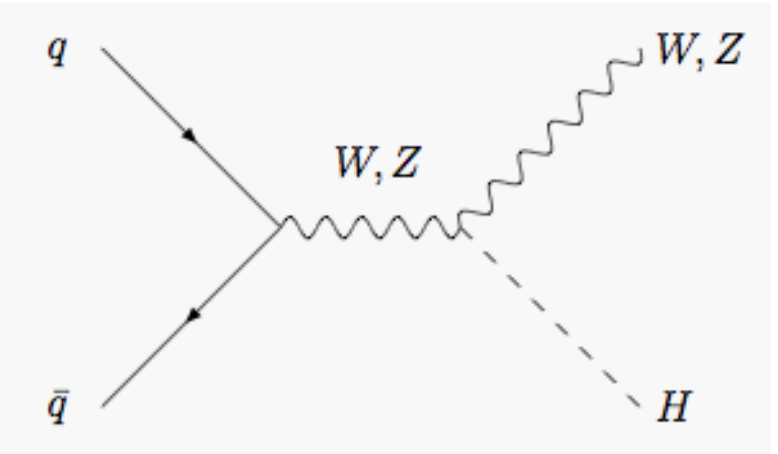


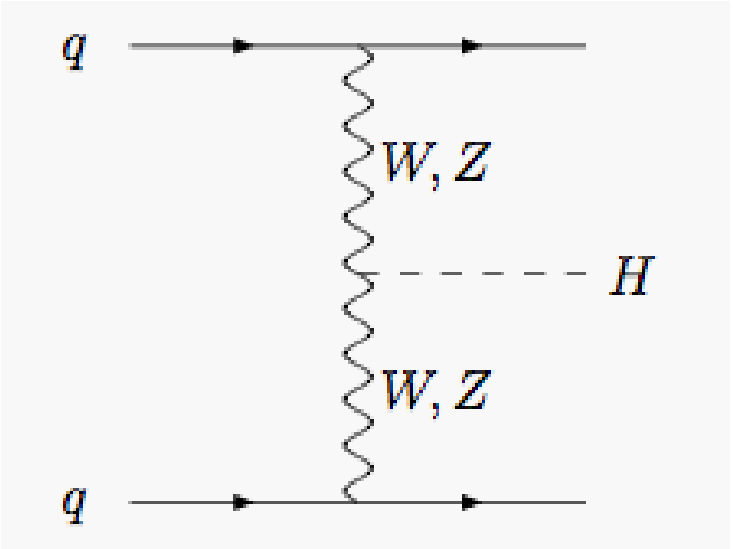


Feynman Diagrams for pp→H

- Each diagram represents an amplitude for Higgs production, which corresponds to a production rate, or *cross section*.



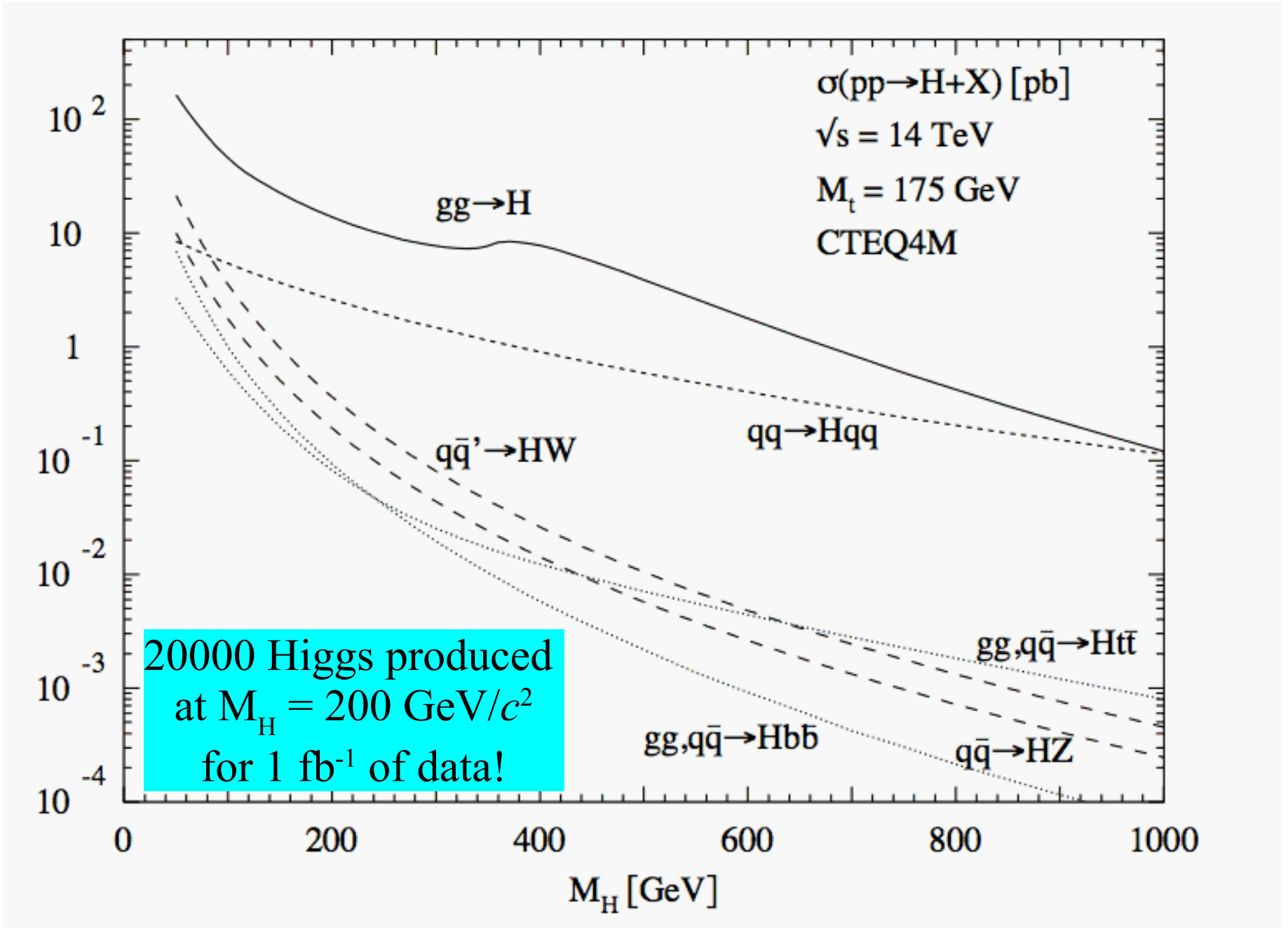




- These are the 5 biggest contributions to the total Higgs cross section at the LHC.

Calculated Cross Section

By NLO Monte Carlo, see M. Spira, Fortsch. Phys. 46 3 (1998)

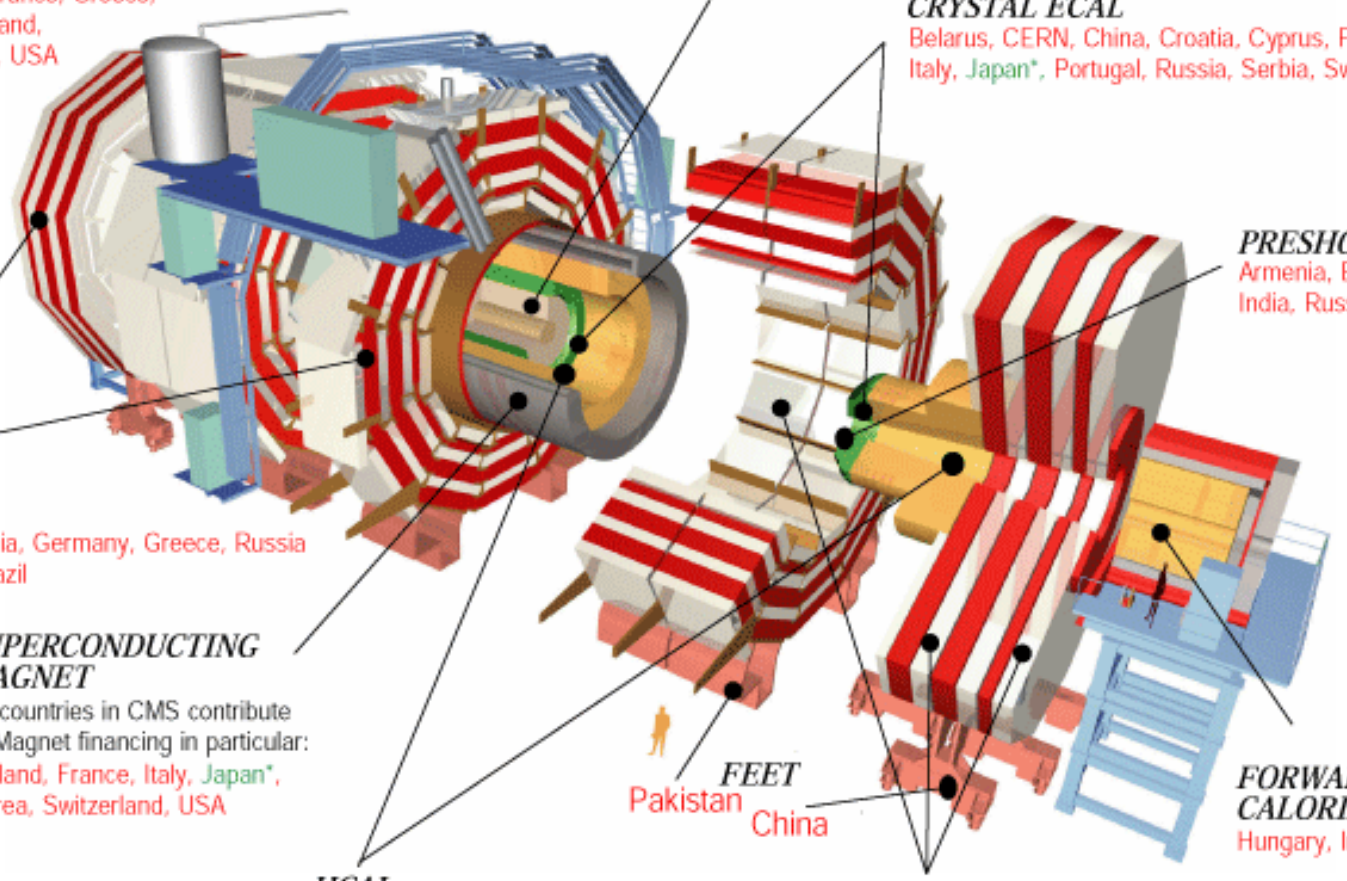


Detecting particles

- For each particle produced, our goal is to measure charge, energy, and momentum as precisely as possible. This allows particle identification.
- For particles that decay before they hit the detector, you can add the energy-momentum four-vector to *reconstruct* a heavier particle.
- Our detector: The Compact Muon Solenoid
  - It's compact (by particle physics standards)
  - It has a good muon-measuring system (*spectrometer*)
  - It has a solenoid inside (to measure charges)

CMS Collaboration

36 Nations, 160 Institutions, 2008 Scientists and Engineers (November 2003)



TRIGGER & DATA ACQUISITION
Austria, CERN, Finland, France, Greece, Hungary, Italy, Korea, Poland, Portugal, Switzerland, UK, USA

TRACKER
Austria, Belgium, CERN, Finland, France, New Zealand Germany, Italy, Japan*, Switzerland, UK, USA

CRYSTAL ECAL
Belarus, CERN, China, Croatia, Cyprus, France, Ireland Italy, Japan*, Portugal, Russia, Serbia, Switzerland, UK, USA

PRESHOWER
Armenia, Belarus, CERN, Greece, India, Russia, Taipei, Uzbekistan

RETURN YOKE
Barrel: Czech Rep., Estonia, Germany, Greece, Russia
Endcap: Japan*, USA, Brazil

SUPERCONDUCTING MAGNET
All countries in CMS contribute to Magnet financing in particular: Finland, France, Italy, Japan*, Korea, Switzerland, USA

FEET
Pakistan
China

FORWARD CALORIMETER
Hungary, Iran, Russia, Turkey, USA

HCAL
Barrel: Bulgaria, India, Spain*, USA
Endcap: Belarus, Bulgaria, Russia, Ukraine
HO: India

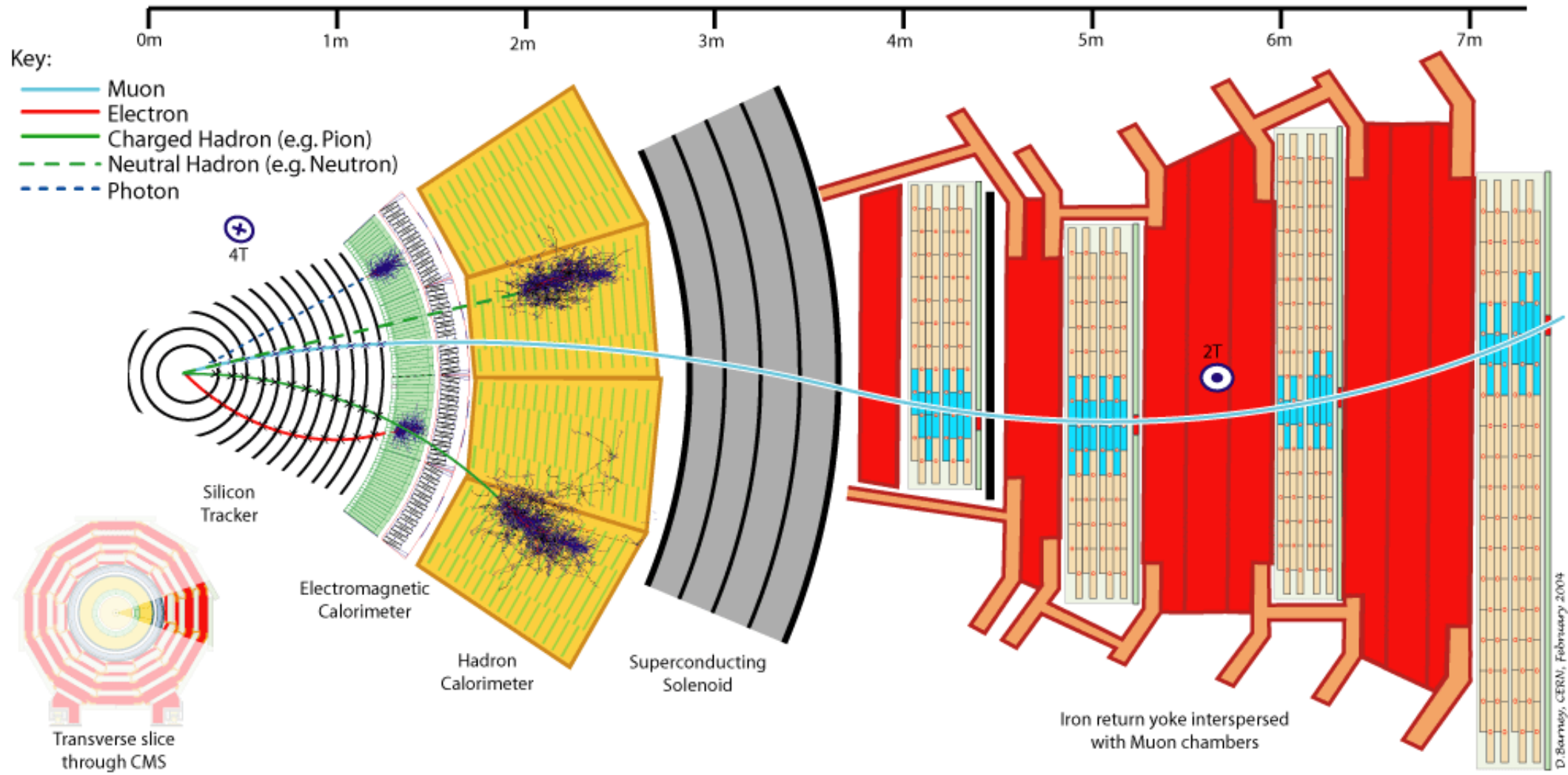
MUON CHAMBERS
Barrel: Austria, Bulgaria, CERN, China, Germany, Hungary, Italy, Spain,
Endcap: Belarus, Bulgaria, China, Korea, Pakistan, Russia, USA

* Only through industrial contracts

Total weight	**:**	**12500 T**
Overall diameter	**:**	**15.0 m**
Overall length	**:**	**21.5 m**
Magnetic field	**:**	**4 Tesla**

- The Compact Muon Solenoid is one of two pp collider detector experiments at the LHC.

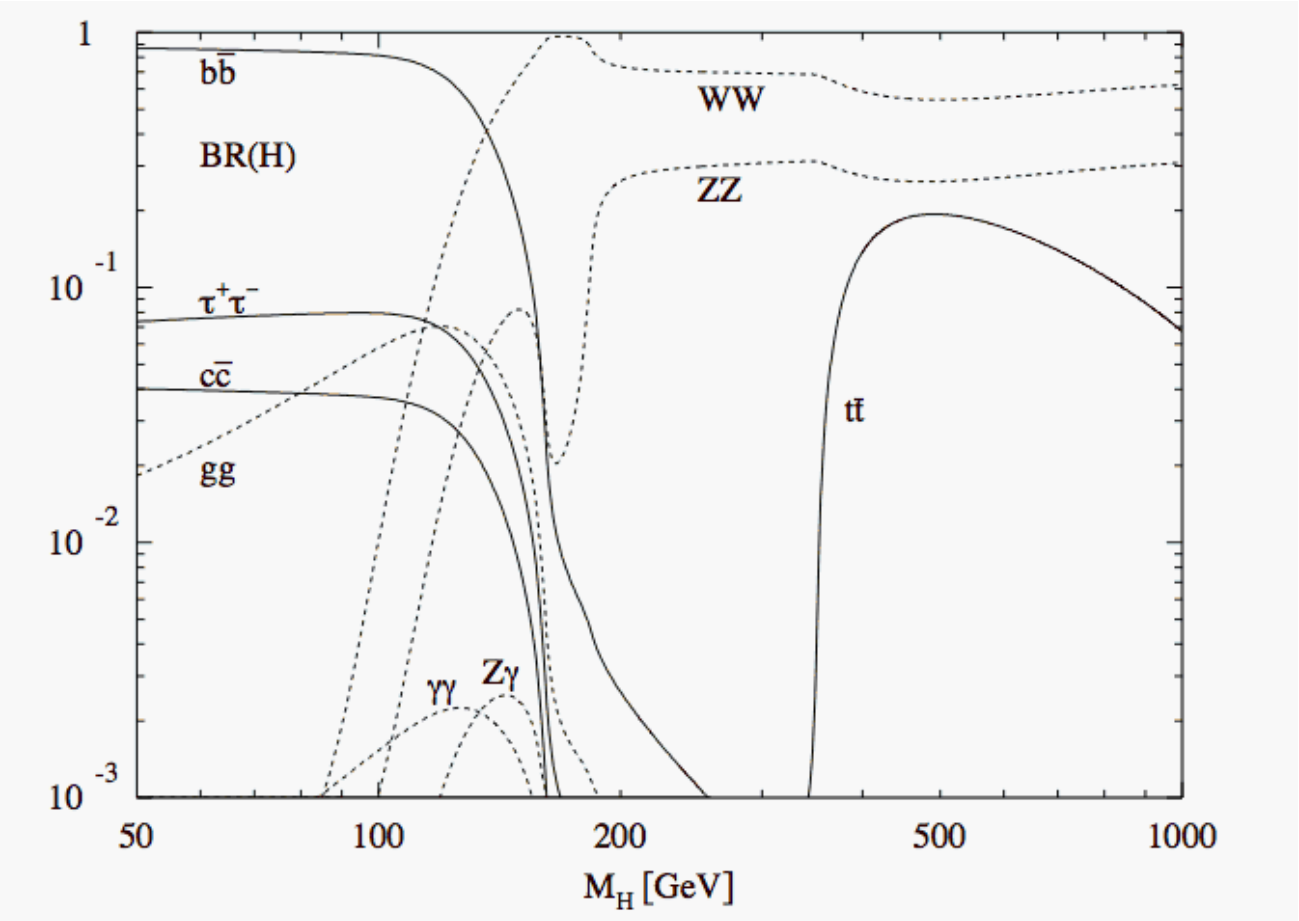
The Compact Muon Solenoid



- CMS detector measures charge, energy, and momentum of particles produced in collisions.

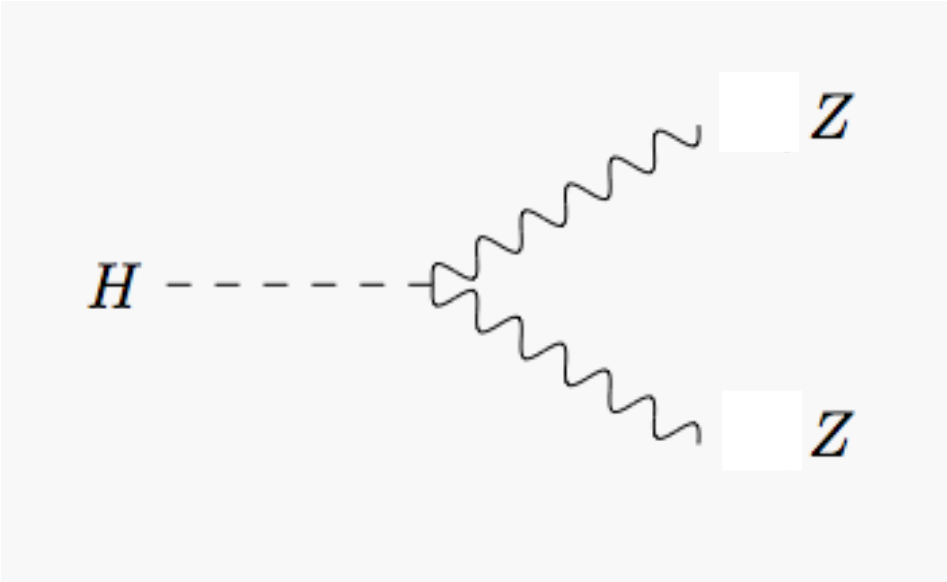
Heavy particles decay!

- Heavy particles always have a finite lifetime, as long as they can conserve four-momentum and a few quantum numbers through decays.
- There are lots of SM Higgs decay modes, each with a probabilistic decay rate (*branching ratio*)



Detecting the Higgs

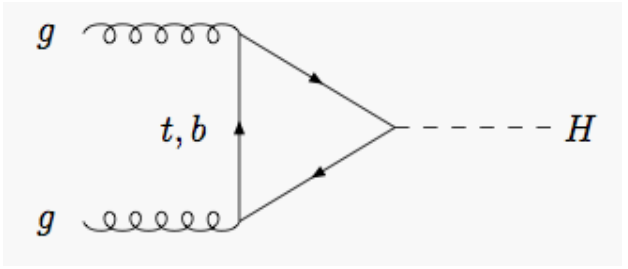
- Look for Higgs decaying to ZZ
  - large branching ratio
  - low SM background
    - $qq \rightarrow Z$, $qq' \rightarrow ZZ$, $qq' \rightarrow WZ$
    - Not including jets faking leptons



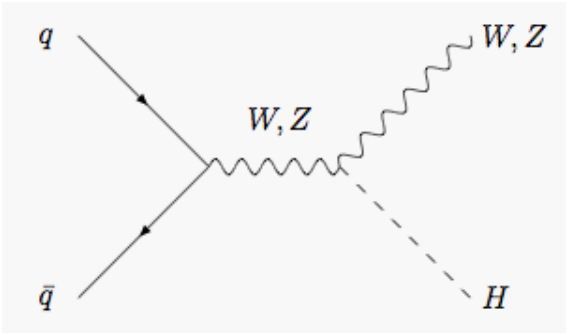
Choose to reconstruct a Z from 2 leptons (e^+e^- or $\mu^+\mu^-$). This is the easiest signal in CMS, and the so-called "Golden mode" for Higgs discovery at the LHC.

Distribution of final states

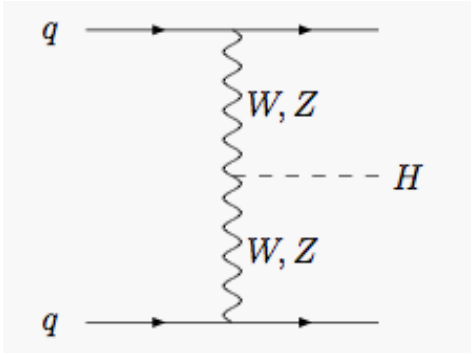
- Can be derived through a relativistic quantum mechanics calculation, gives rate for Higgs to be produced, $\propto \int |M|^2 \, d^4p$. (dp taken over phase space)



+



+



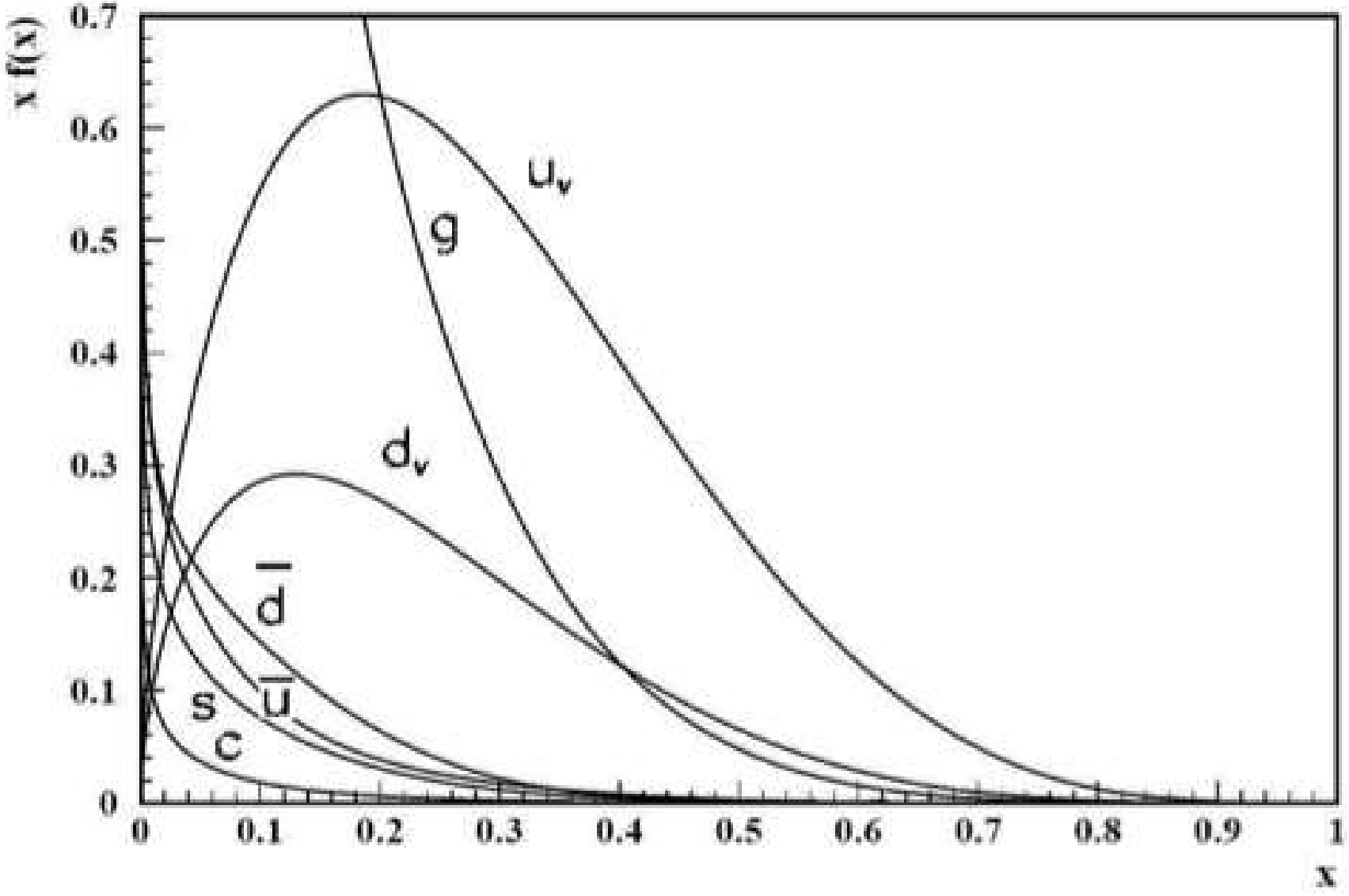
+ (higher order diagrams)

$= \boldsymbol{M}$

- But, a complication: the initial state momenta are not exactly known!

Parton Distribution Functions

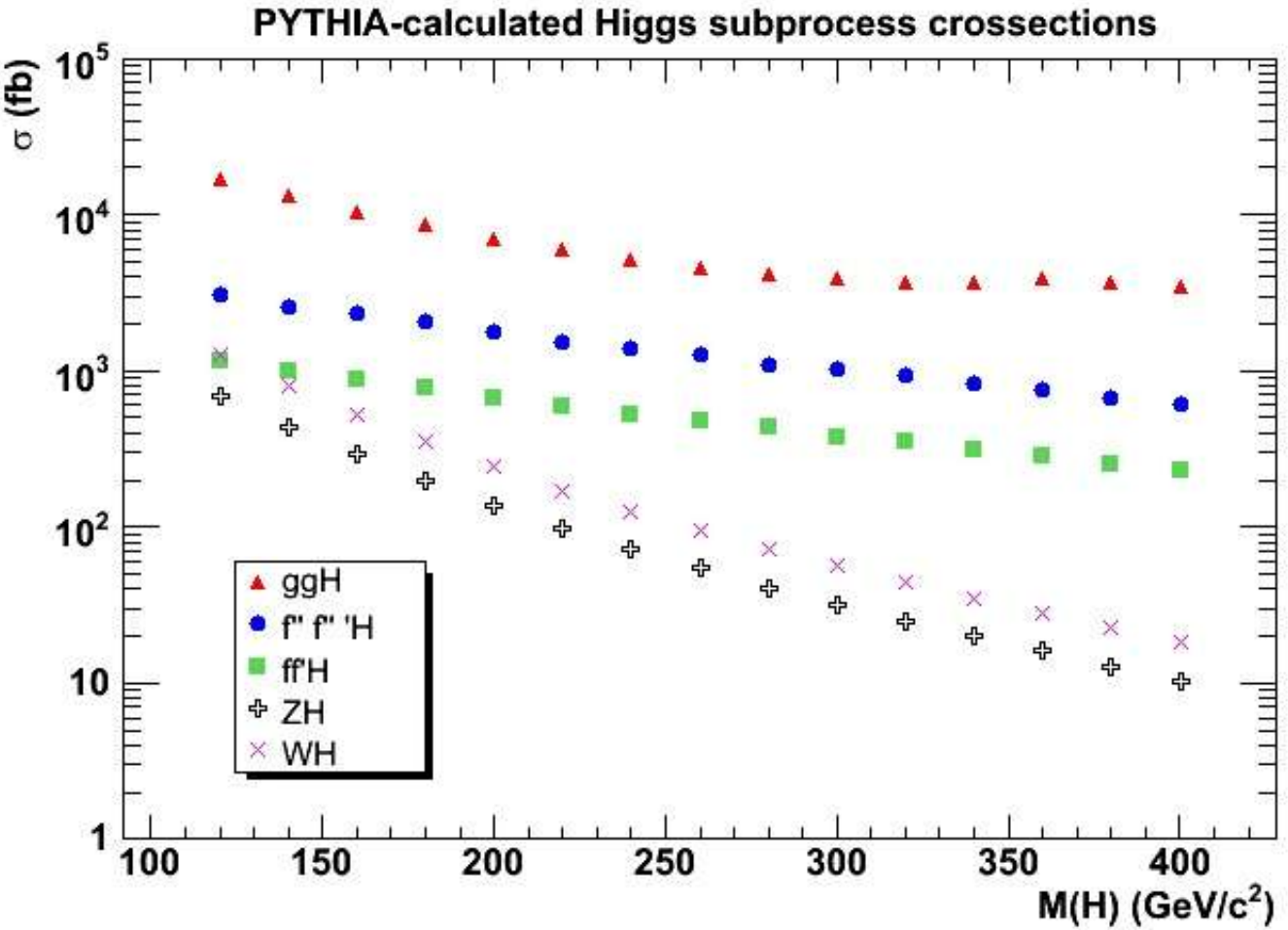
- Initial state of a quark or gluon INSIDE proton is not calculable. Each has some probability (f) to carry a fraction (x) of the total proton momentum.



Monte Carlo Event Generation

- To do the phase space integral, we must randomly sample over this distribution of initial states.
- PYTHIA 6.401 was used as a Monte Carlo event generator to generate 1 fb^{-1} of $H \rightarrow ZZ$ signal and SM background events from pp collisions at the LHC.
- Generated signal for $M_H = (140, 160, \ldots, 400\ GeV/c^2)$.
- The events are writing as massive lists of four-vectors, each having thousands of particles.
- Actually generated 10 times as much data, then scaled it down, to improve sample statistics
- During the generation, a trigger of $E_T > 15$ GeV on leptons was used to save disk space.

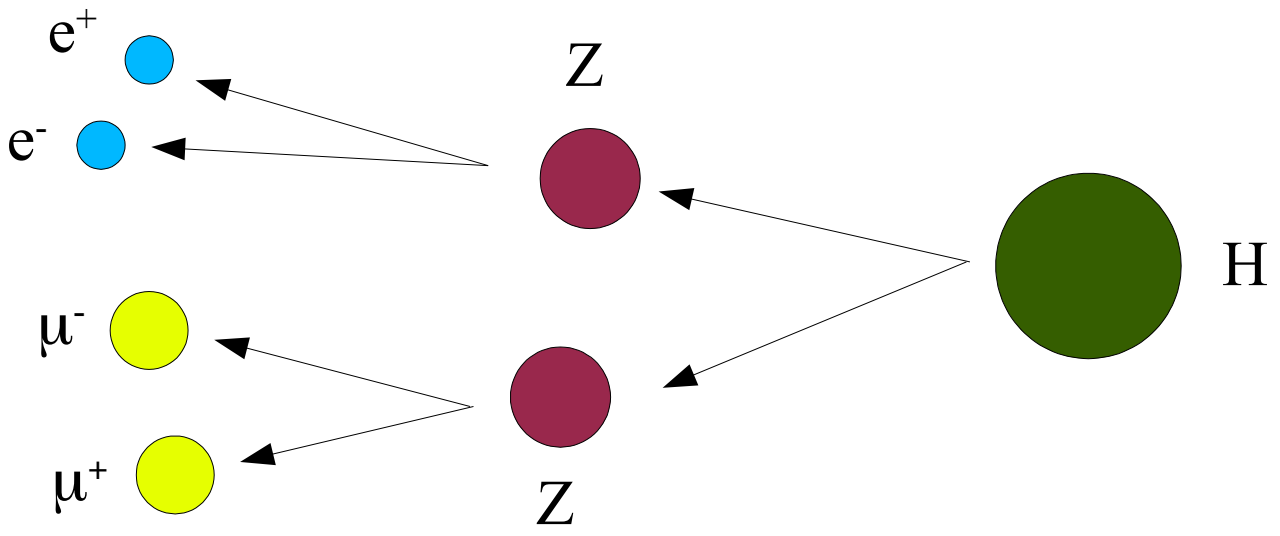
Cross sections by PYTHIA



- The computed cross sections are different than the previous plot. PYTHIA is leading order (LO) and tends to ignore higher-order diagrams.
- Now, we have events... on to the analysis.

Full Reconstruction

- Higgs candidates (ZZ pairs) were reconstructed from events with 2 Z.
- A Z candidate is taken as a dilepton with invariant mass closest to 91 GeV/c^2, but neither Z is required to be on-shell.
- Acceptance cuts of $|\eta| < 2.5$ (*pseudorapidity*) and $E_T > 25$ GeV were put on the leptons for detection and triggering.



Estimated number of signal events

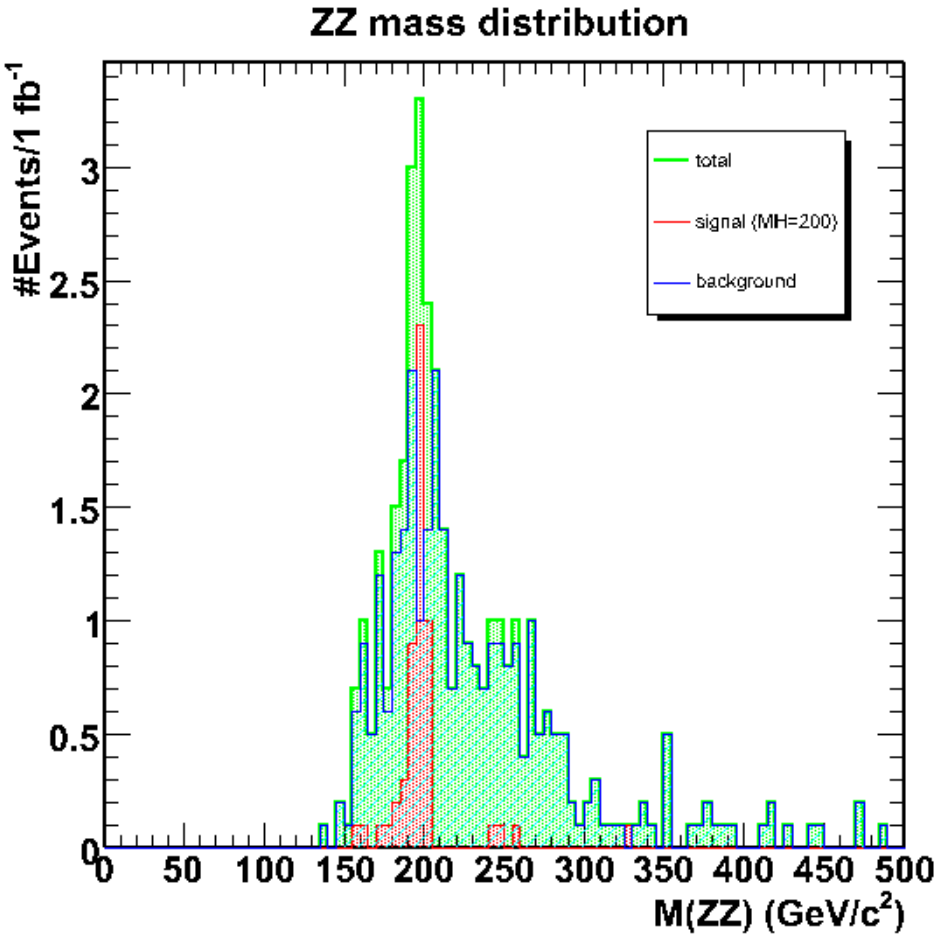
- Given the luminosity, branching ratios, cross section, and acceptances and efficiencies, an estimate of the observed number of events can be made. (Check the Monte Carlo.)

$$\#\text{ of events} = (1\ \text{fb}^{-1}) \cdot \sigma(pp \to H) \cdot \text{Br}(H \to ZZ) \cdot \varepsilon \cdot A$$
$$[\text{Br}(Z \to e^+e^-) \cdot \text{Br}(Z \to e^+e^-) + 2 \cdot \text{Br}(Z \to e^+e^-) \cdot \text{Br}(Z \to \mu^+\mu^-) + \text{Br}(Z \to \mu^+\mu^-) \cdot \text{Br}(Z \to \mu^+\mu^-)]$$

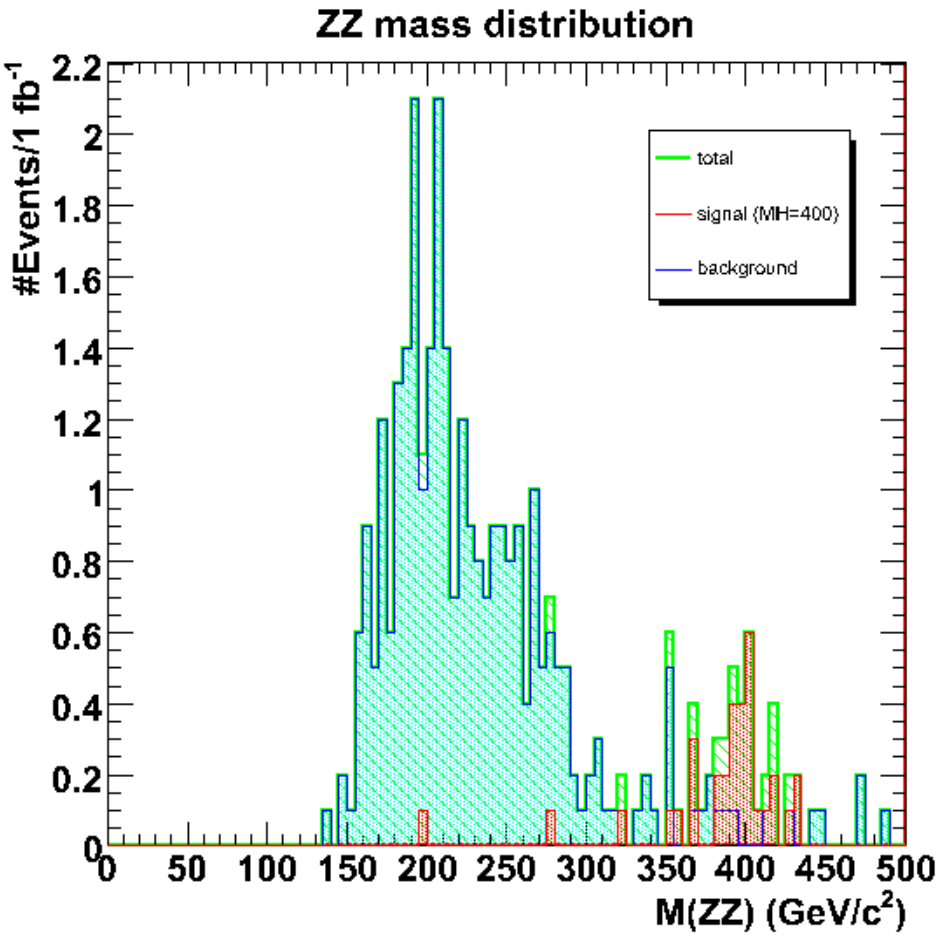
$$= (1\ \text{fb}^{-1}) \cdot [\sigma(pp \to H) \cdot \text{Br}(H \to ZZ)] \cdot (0.00135)$$

- $\text{Br}(Z \to \mu^+\mu^-) = \text{Br}(Z \to e^+e^-) = 3.5\%$.
- The acceptances (A) for full reconstruction are 30%.
- The efficiency (ε) is assumed to be 100% (More later...)
- $\sigma(pp \to H) \cdot \text{Br}(H \to ZZ)$ depends on the Higgs mass.

Full reconstruction statistics



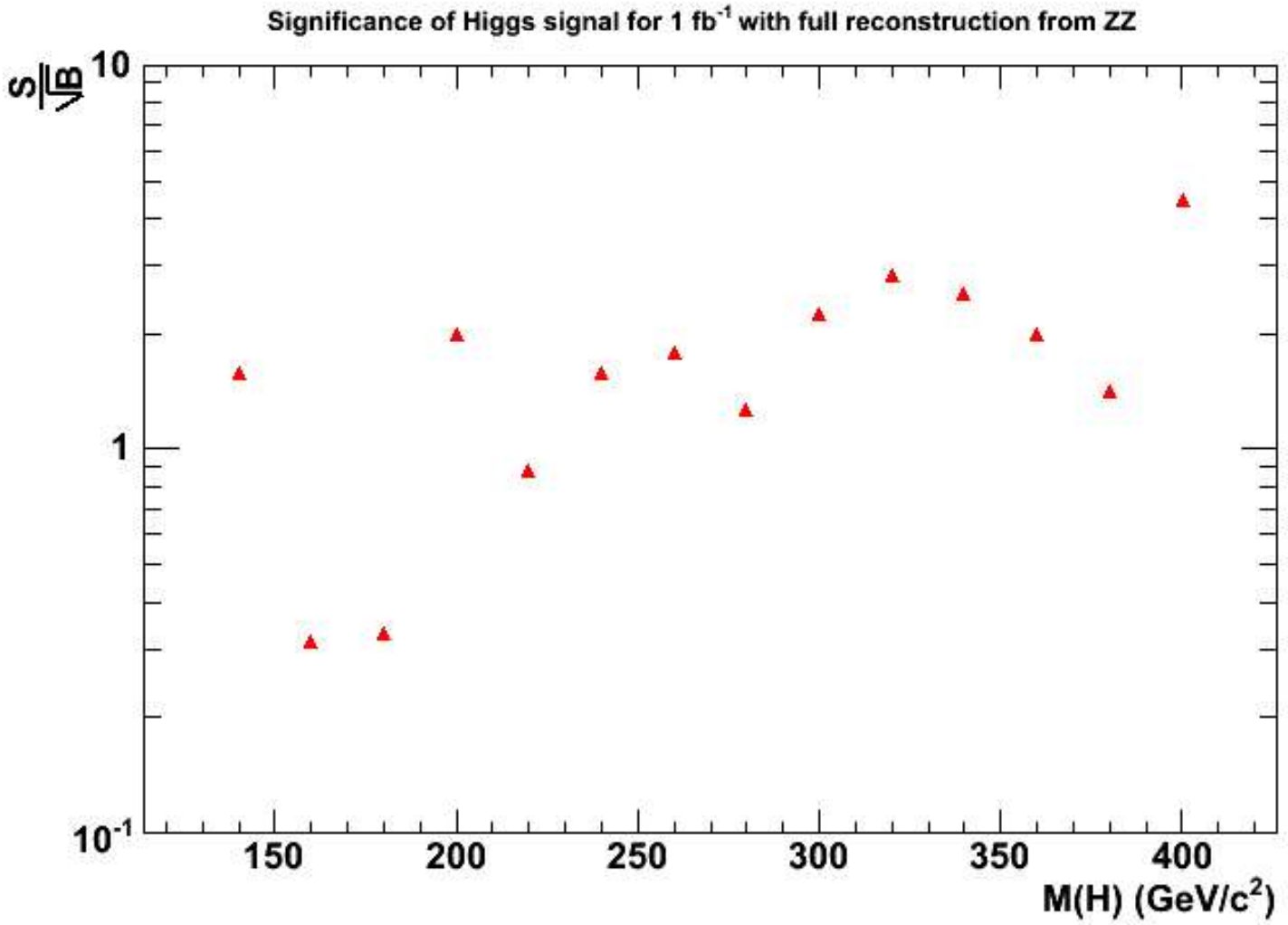
200 GeV/c^2



400 GeV/c^2

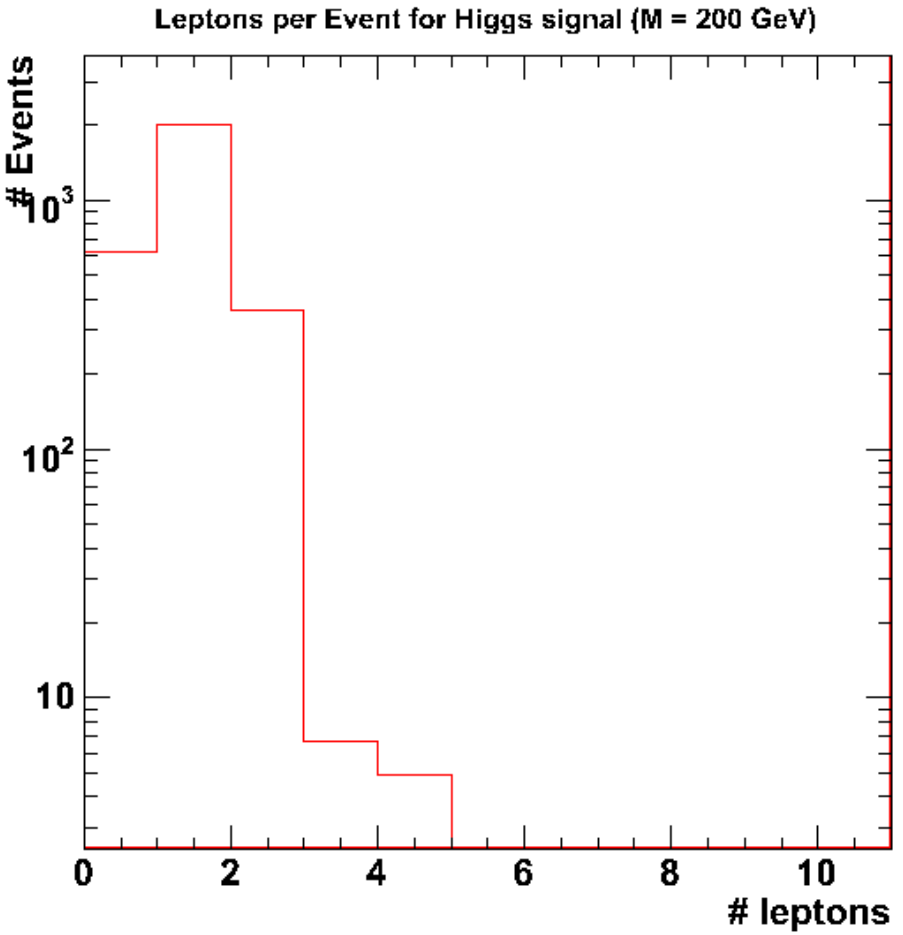
Significance of results

- By comparing the expected SM backgrounds to the observed signal in 10 GeV windows, compute ($S/\sqrt{B}$)
- ($S/\sqrt{B}$) > 5 is needed to convince anybody of discovery.



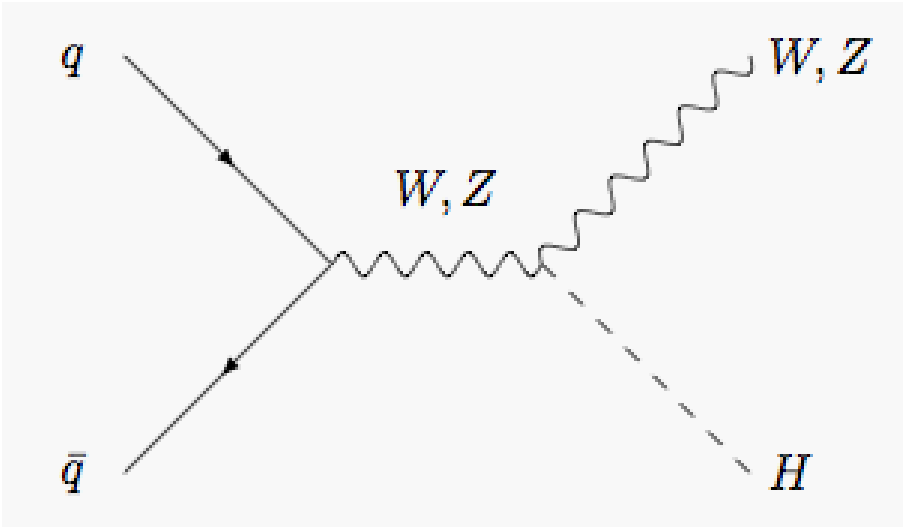
Goal: more signal, less background.

- Generally, full reconstruction will not work with 1 fb^{-1}. For a low-M_H (< 500 GeV/c^2).
- Full reconstruction is too selective for such a small dataset.
- Being less restrictive in how many Zs or leptons (exactly 4 in the full reconstruction) required can help increase the signal.
- There are other ways to reduce bg besides counting leptons.



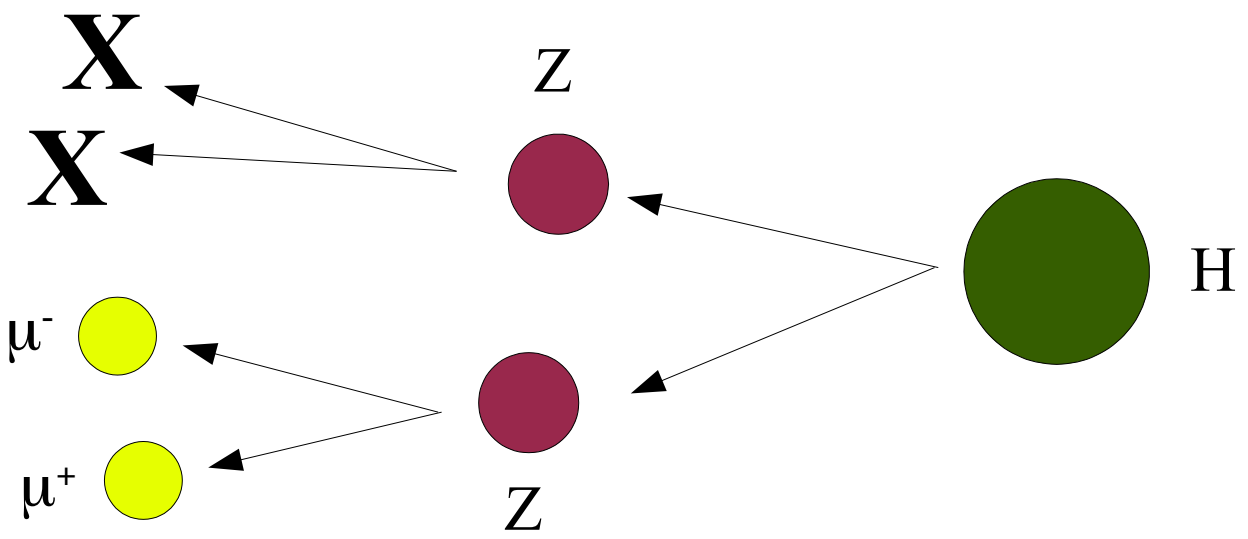
Kinematics and production

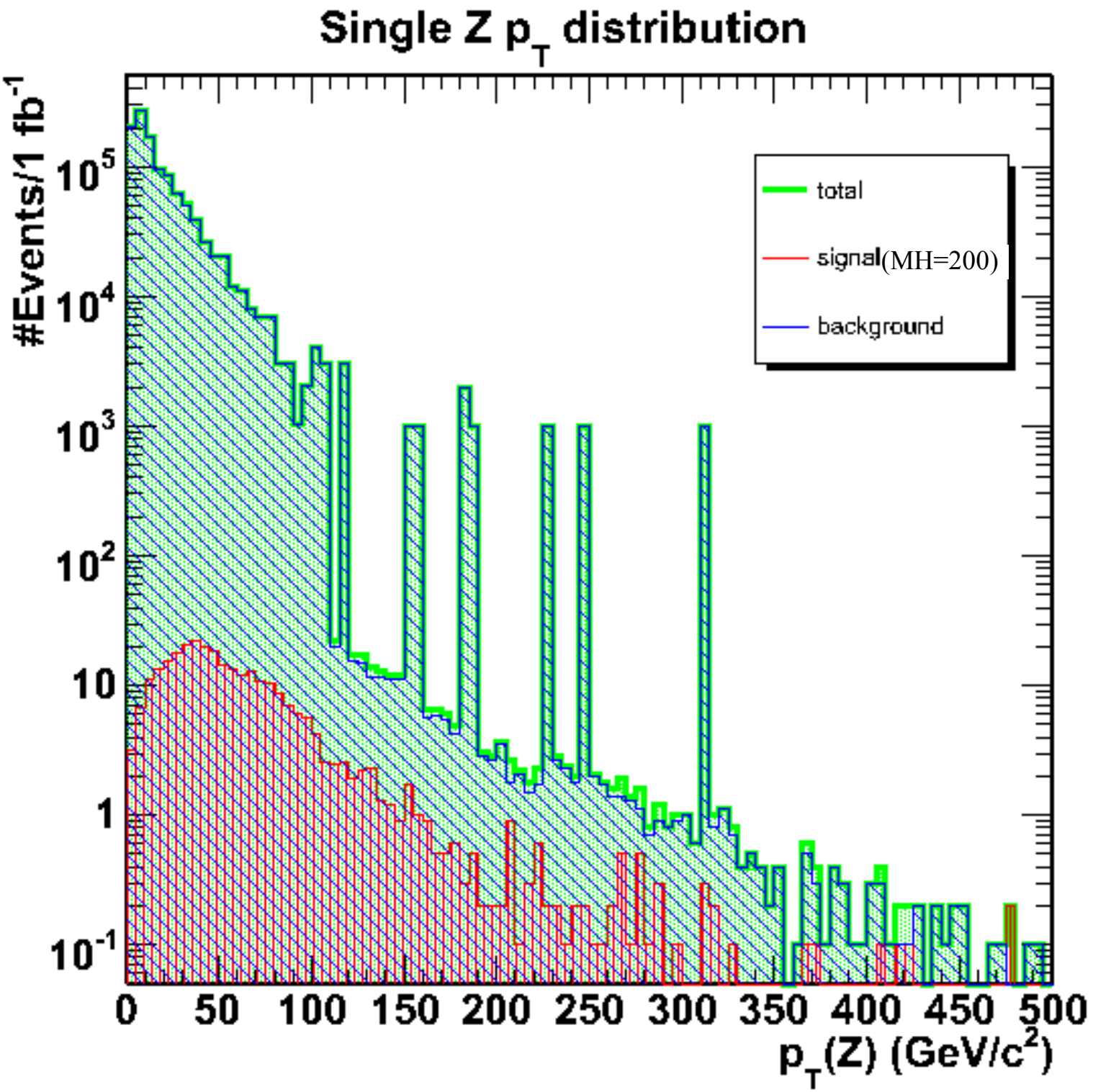
- Explore the kinematic distributions (invariant mass, momentum) of the Zs produced in Higgs decay. They are interesting without reconstructing a Higgs (two-Z) mass.
- The shape and location of a peak in the signal (e.g., the momentum of a single Z that decayed from a Higgs) can be quite different than that of a Z from SM background.
- It is possible to pick up extra Zs and leptons from Higgs production processes (like pp→ZH or WH). The analysis should be sensitive to this, which may help for low M_H.



p_T of a single Z

- Look for one dilepton pair per event with mass closest to on-shell Z mass. (91 GeV/c^2)
- Keep the geometrical and kinematic acceptances on the leptons. ($|\eta| < 2.5$, $E_T > 25$ GeV)
- Plot the dilepton transverse momentum (p_T).

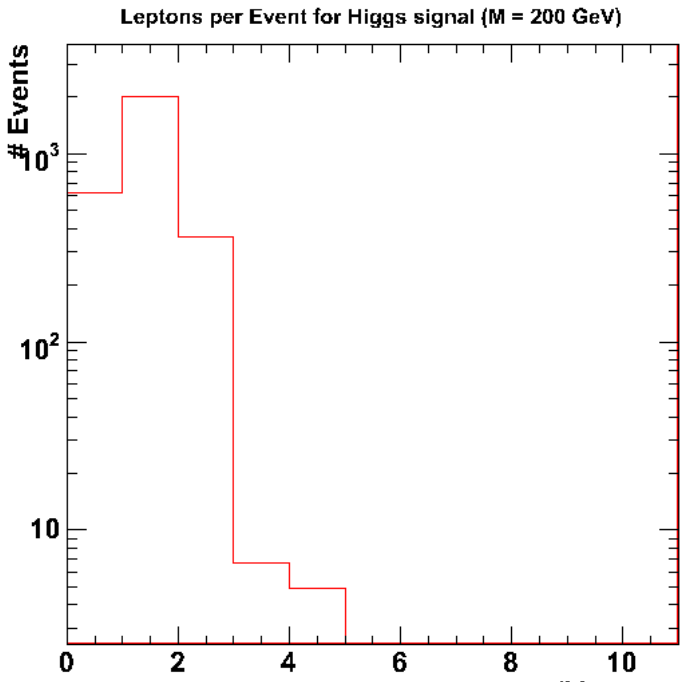


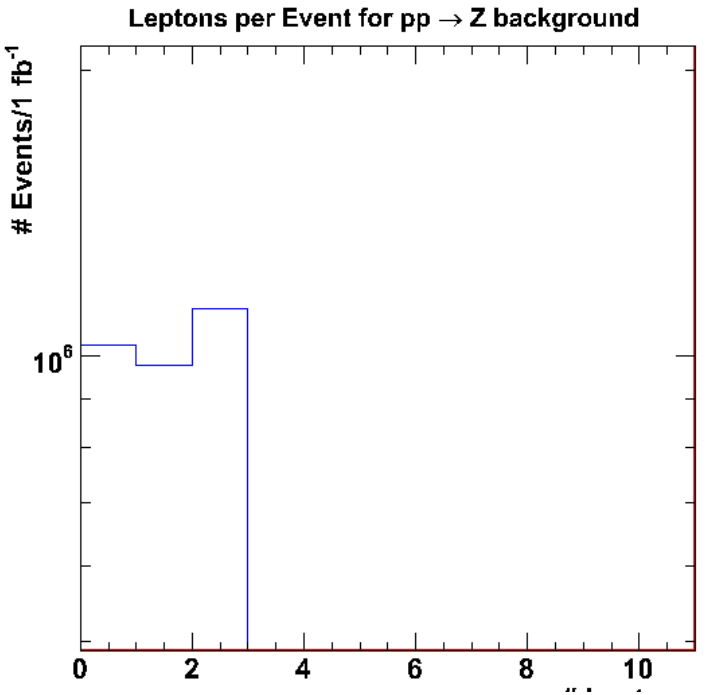

Single Z p_T distribution
#Events/1 fb^{-1}
10^5
10^4
10^3
10^2
10
1
10^{-1}
total
signal(MH=200)
background
0
50
100
150
200
250
300
350
400
450
500
p_T(Z) (GeV/c^2)

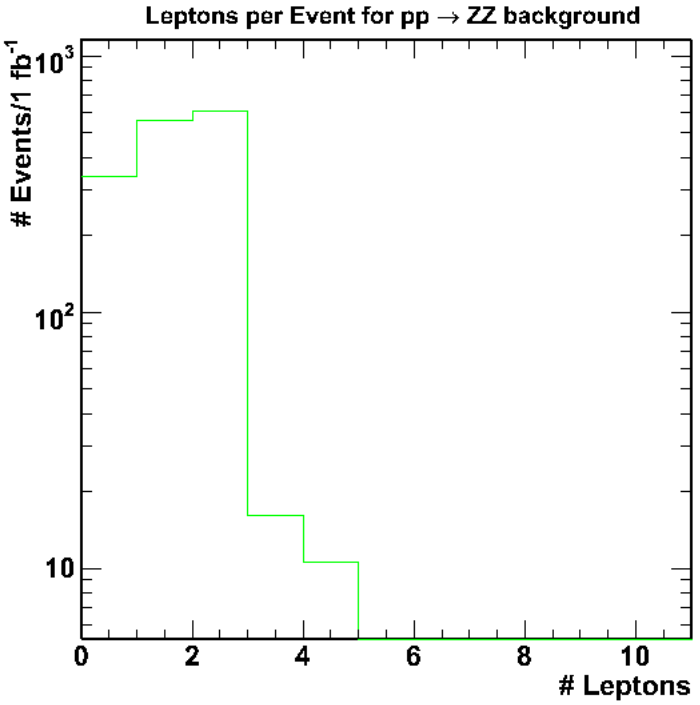
Significance of p_T of a single Z signal

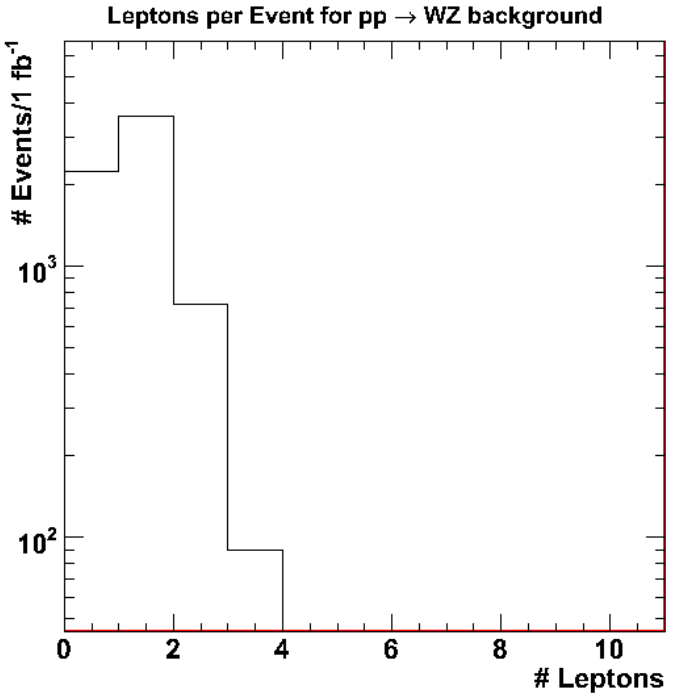
- This method does not reduce the pp→Z background; the cross section for this background is $6x10^8$ fb at the LHC... this is much, much bigger than the signal.
- Signal extraction is pretty much hopeless for any amount of data by looking at $p_T(Z)$.
- How can we beat down the single Z background?

Counting Leptons



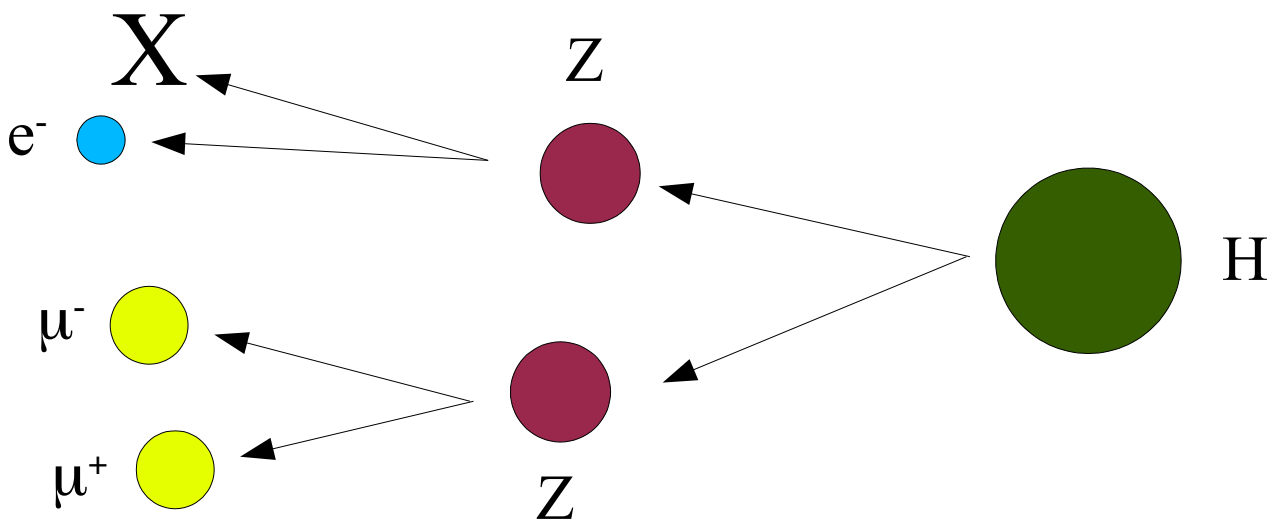


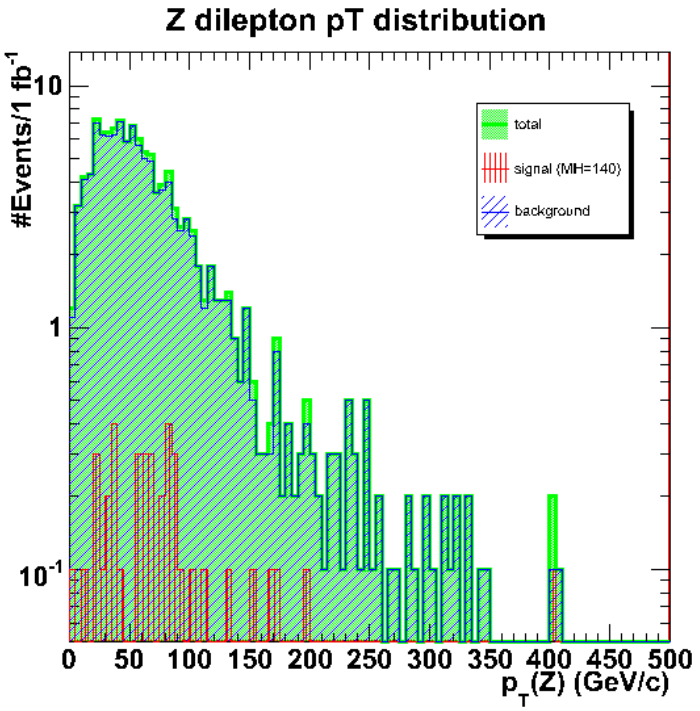




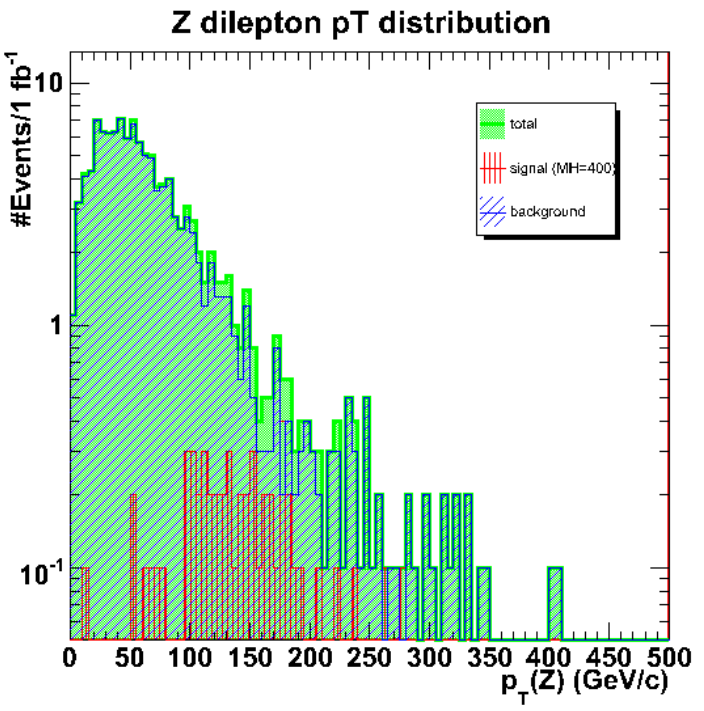
Z (+ lepton) kinematics

- A Z candidate is taken as a dilepton with invariant mass between 81 and 101 GeV/c^2.
- We look for a third lepton in the event that passes the acceptance cuts. ($|\eta| < 2.5$, $E_T > 25$ GeV)
- Plot the p_T of the Z.



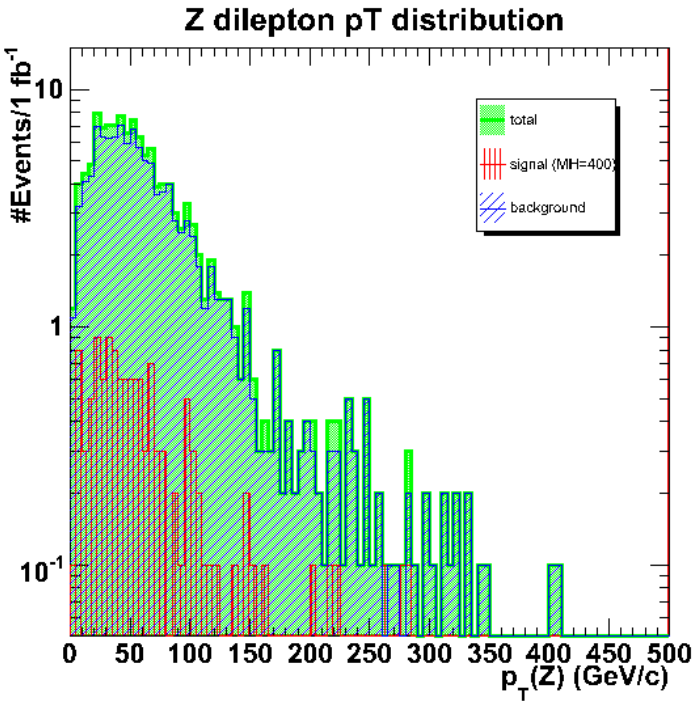


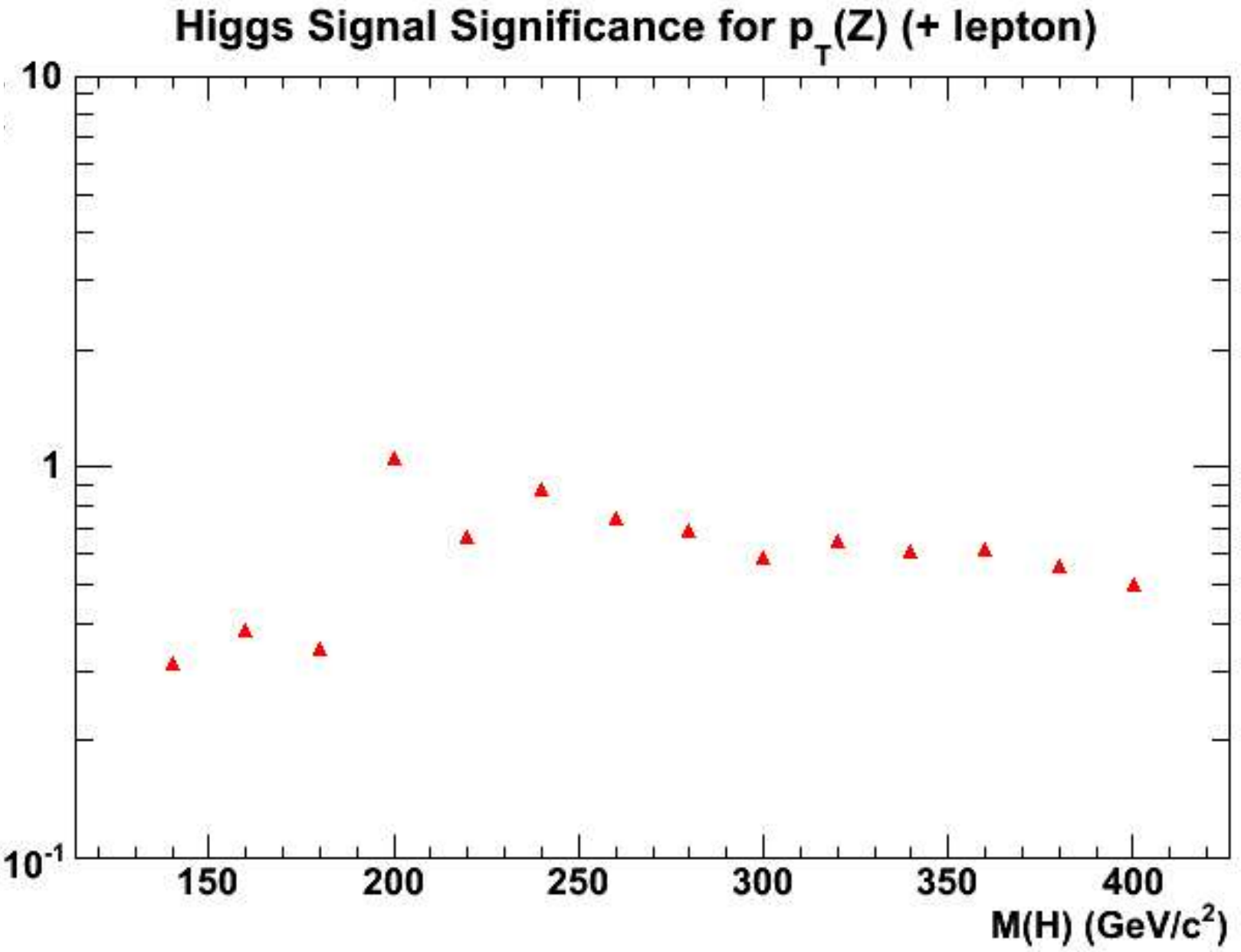
140 GeV/c^2



400 GeV/c^2

200 GeV/c^2

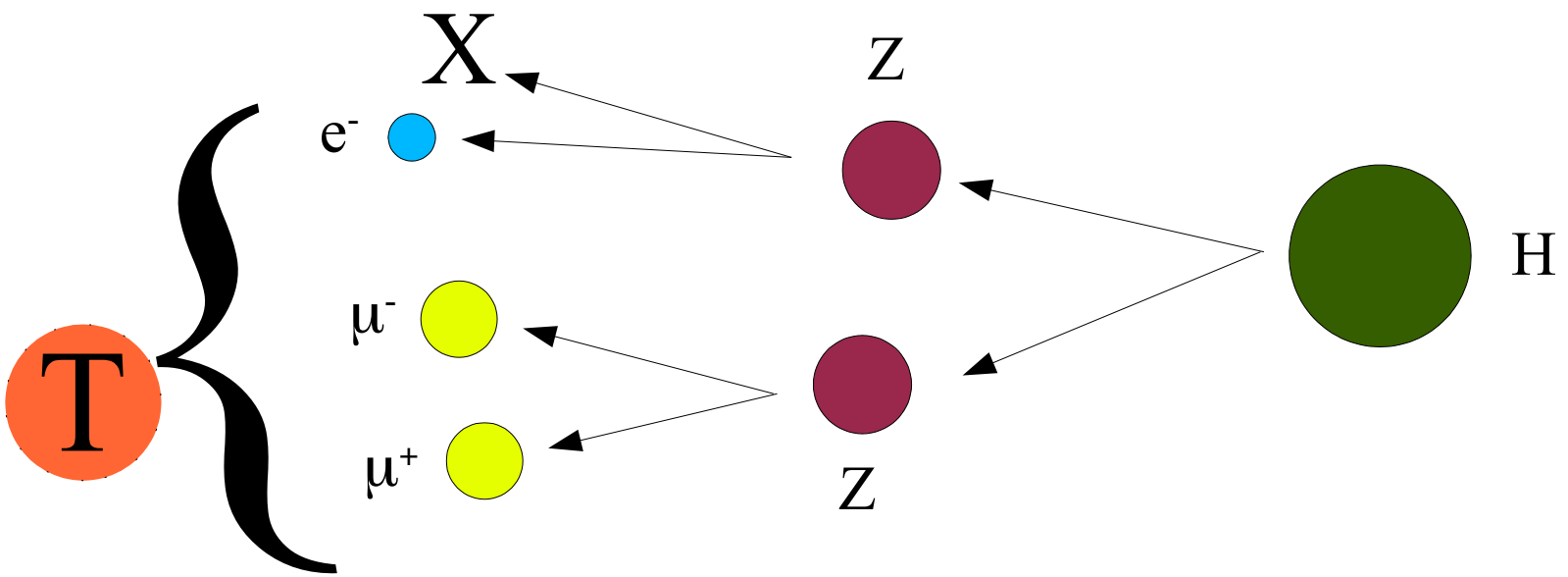


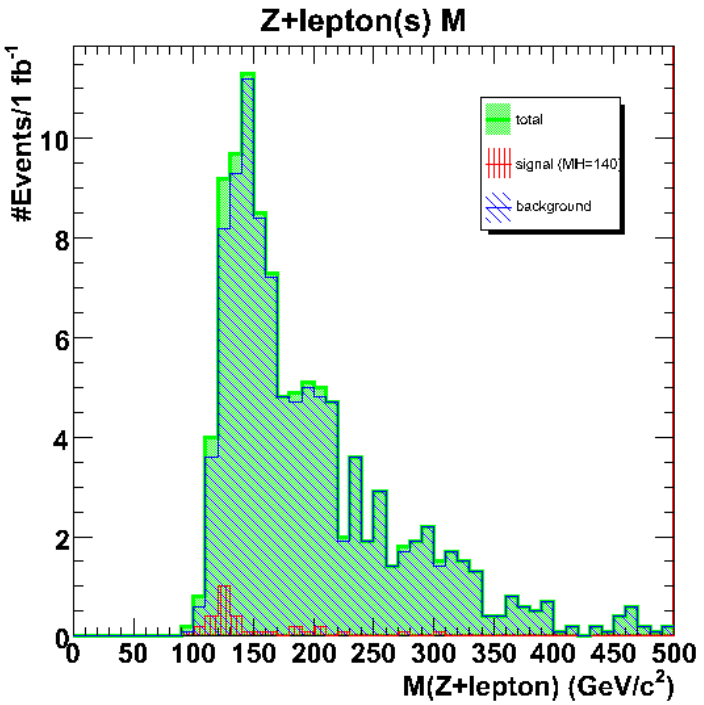


- Significance is higher than full reconstruction for some values of M_H in the low-mid mass range.

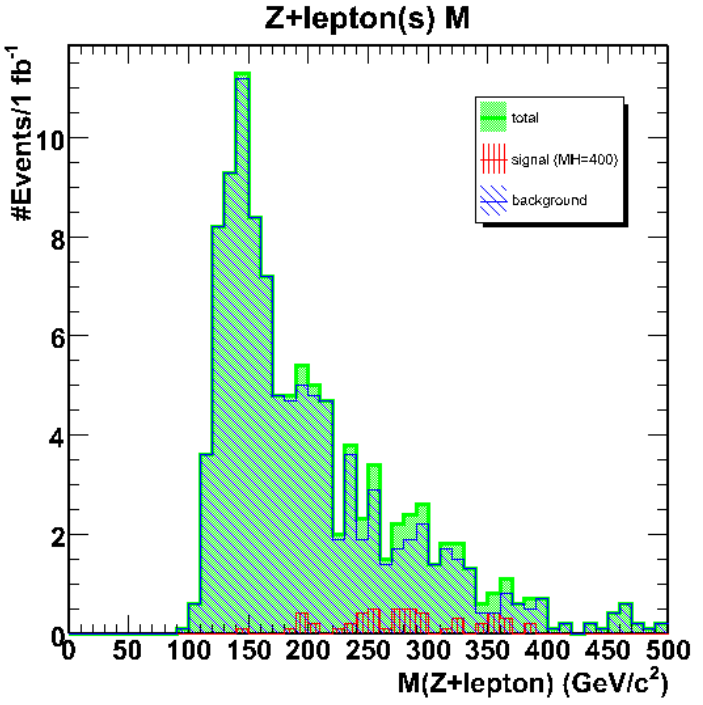
Trilepton (Z+ lepton) kinematics

- A Z candidate is taken as a dilepton with invariant mass between 81 and 101 GeV/c^2.
- We look for a third lepton in the event that passes the acceptance cuts. ($|\eta| < 2.5$, $E_T > 25$ GeV)
- Plot the p_T of the Z+lepton (the *trilepton* p_T).



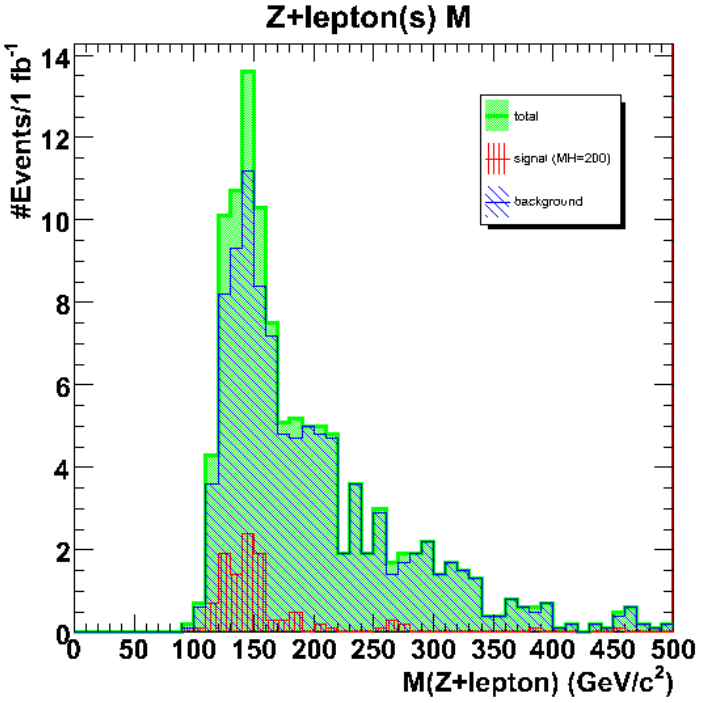


140 GeV/c^2

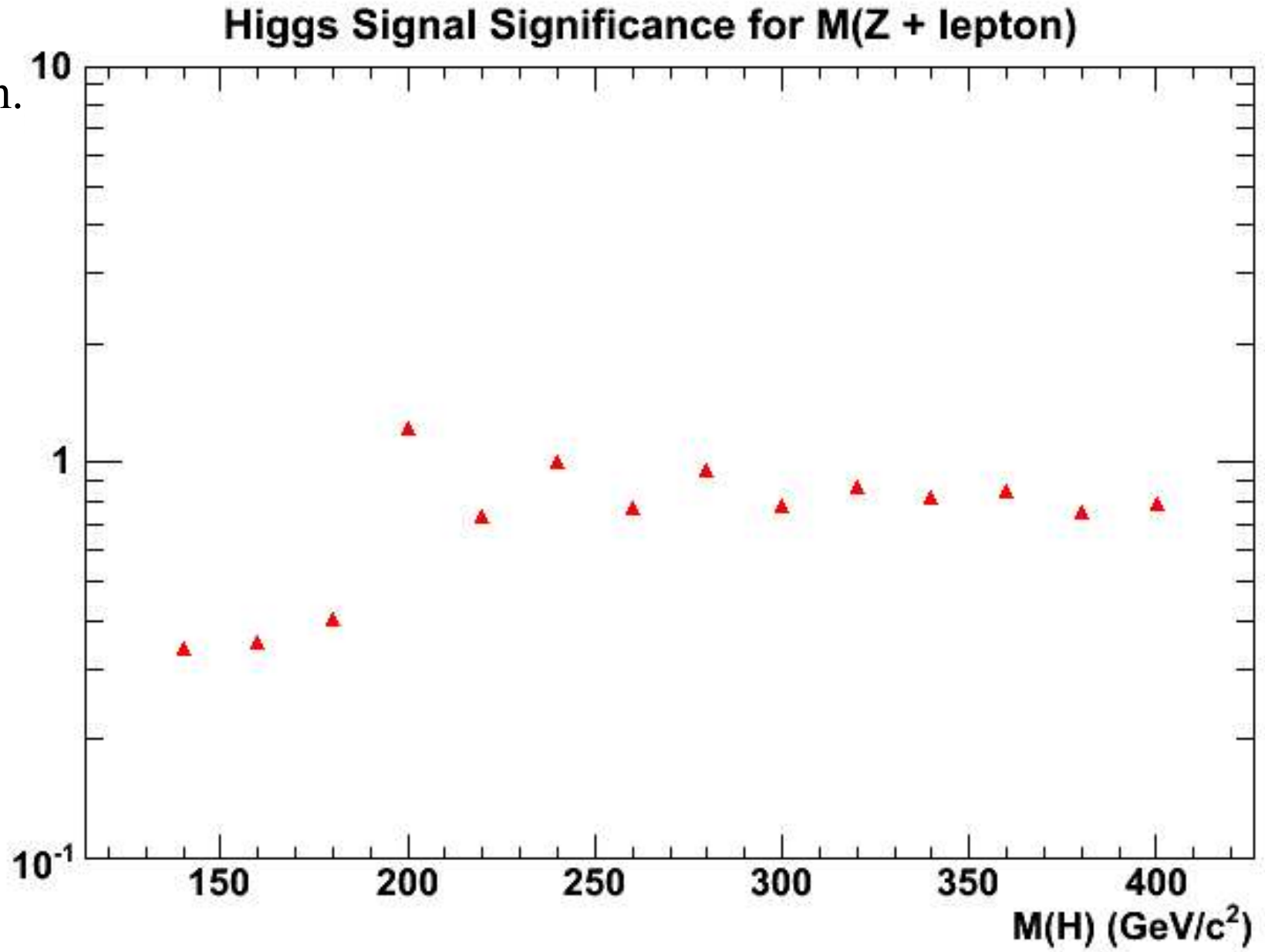


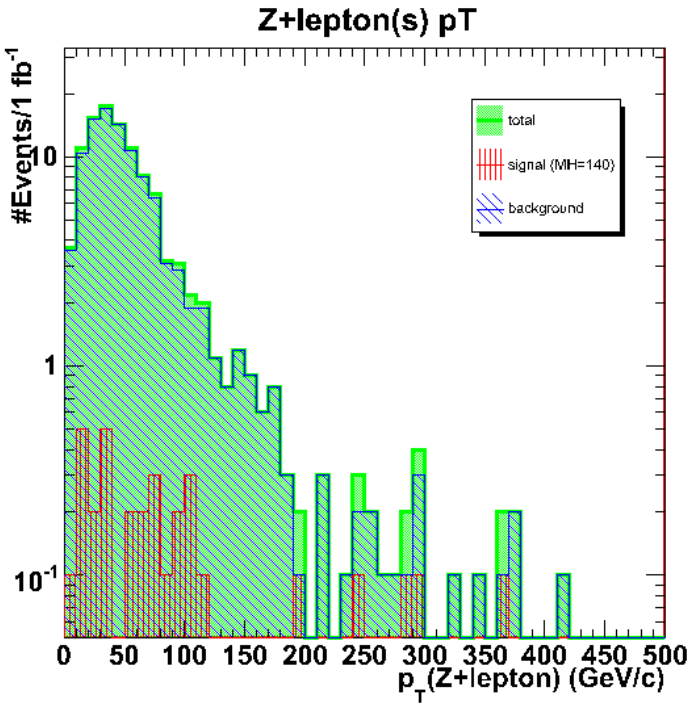
400 GeV/c^2

200 GeV/c^2

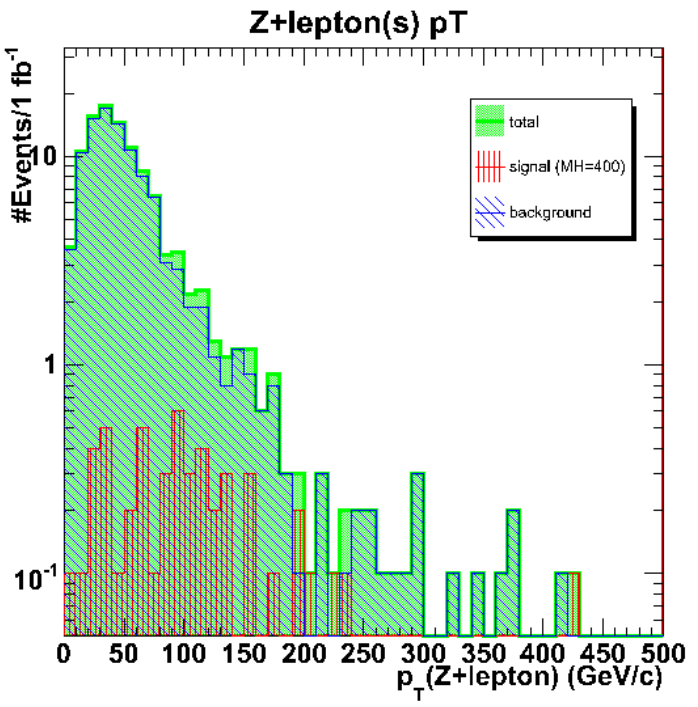


Sign.


Higgs Signal Significance for M(Z + lepton)
10
1
10^{-1}
150
200
250
300
350
400
M(H) (GeV/c^2)

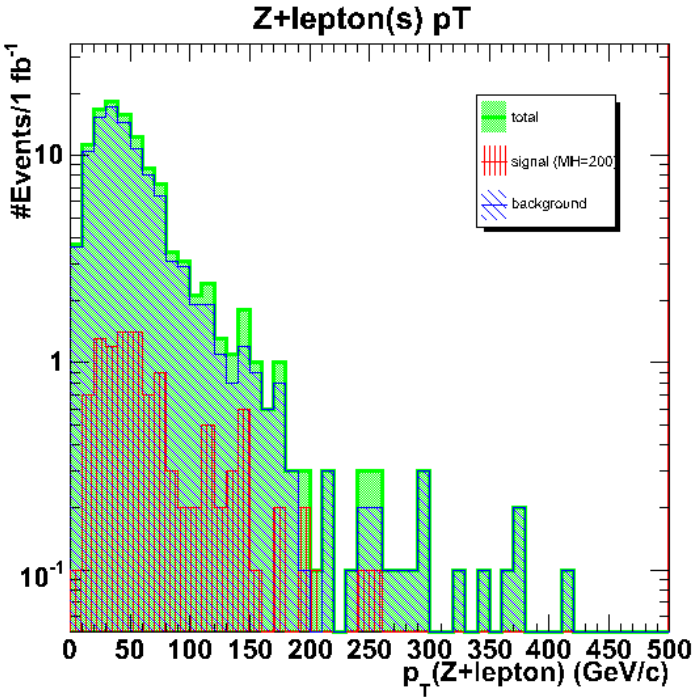


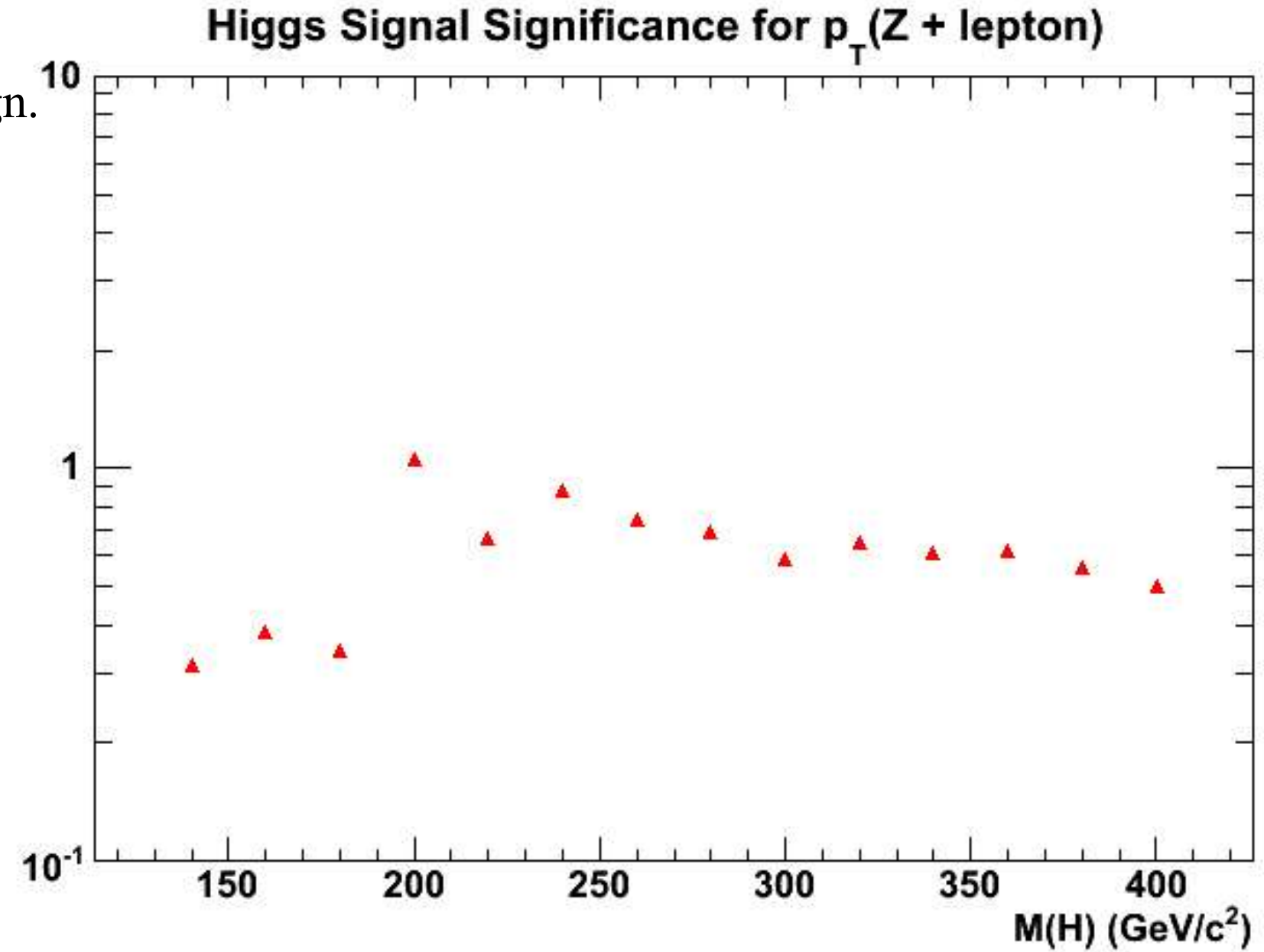
140 GeV/c^2

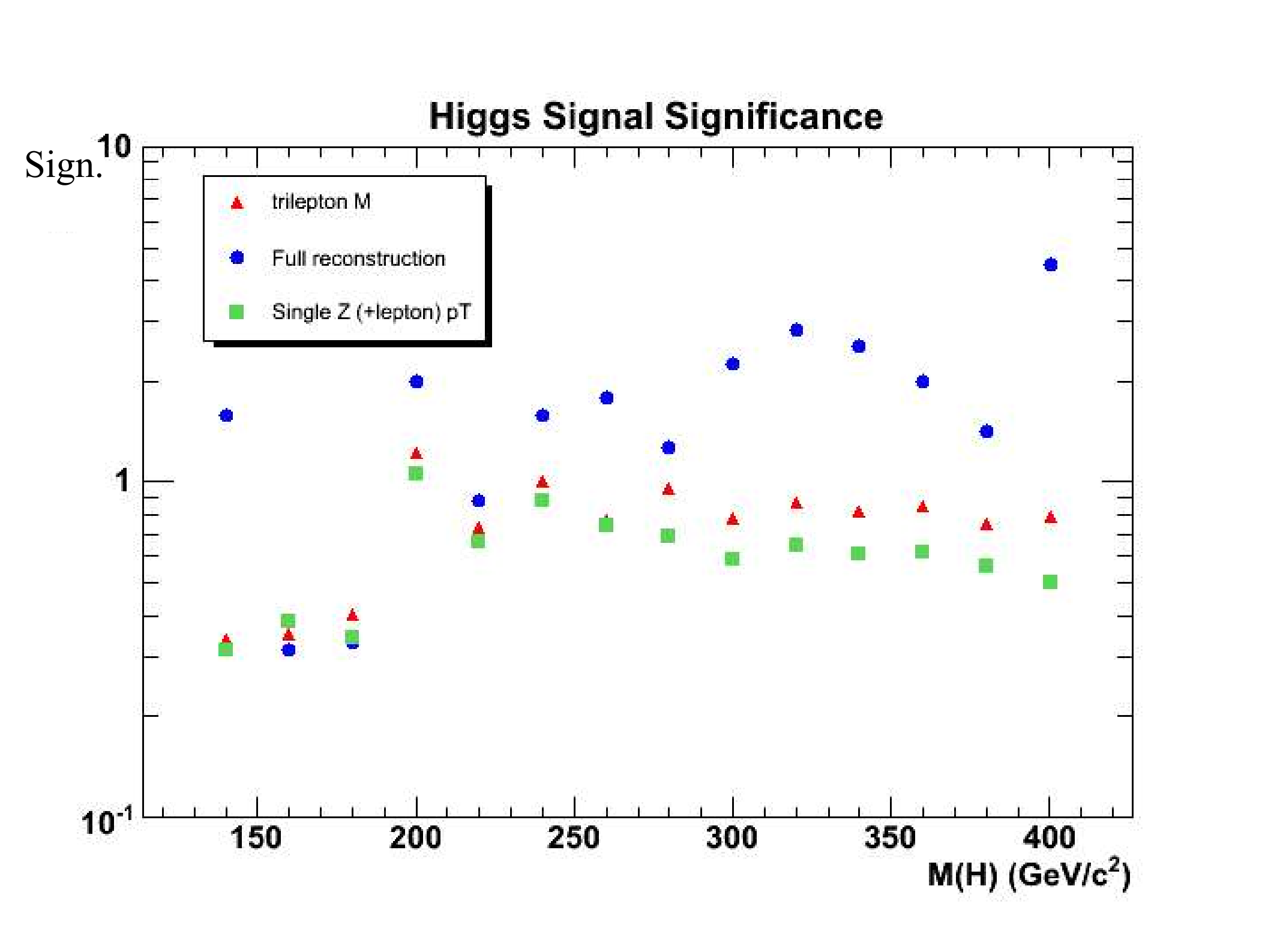


400 GeV/c^2

200 GeV/c^2




Higgs Signal Significance for p_T(Z + lepton)
Sign.
10
1
10^{-1}
150
200
250
300
350
400
M(H) (GeV/c^2)

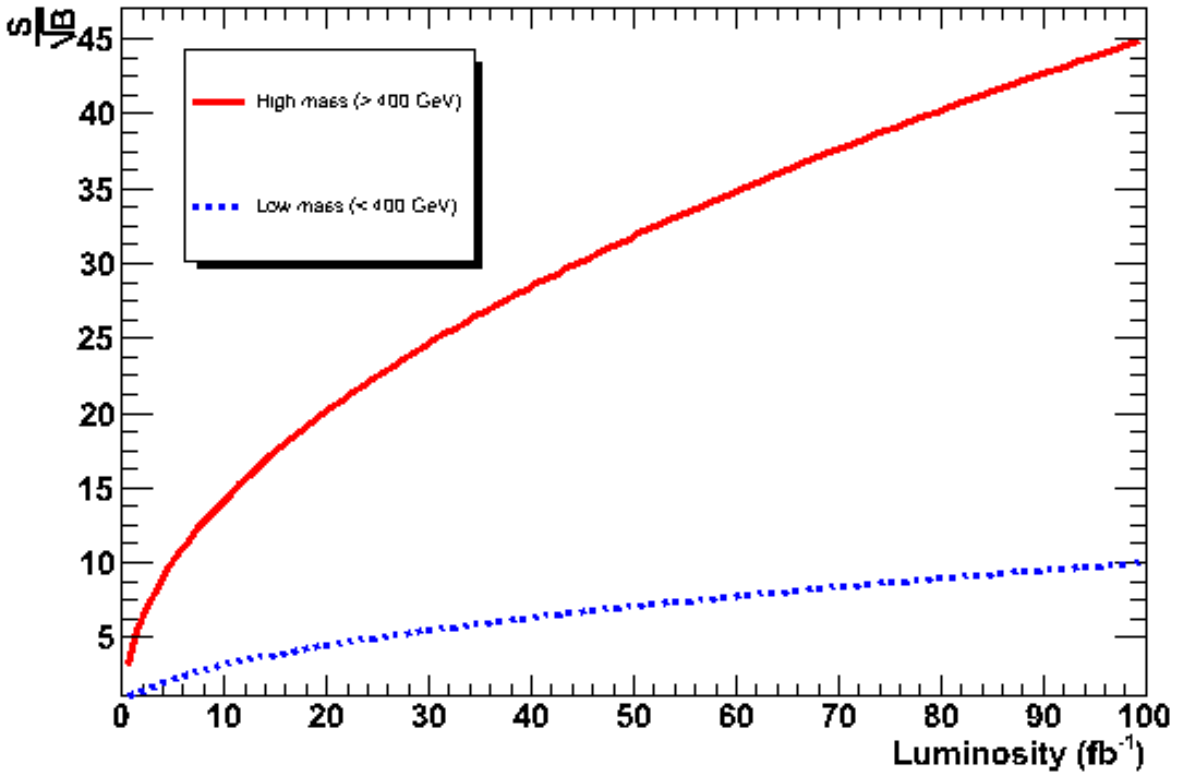

Higgs Signal Significance
Sign.
10
1
10^{-1}
trilepton M
Full reconstruction
Single Z (+lepton) pT
150
200
250
300
350
400
M(H) (GeV/c^2)

What now?

- We have a slight edge doing this kinematic analysis for a lower mass Higgs, but limiting the data set to 1 fb^{-1} is hard.

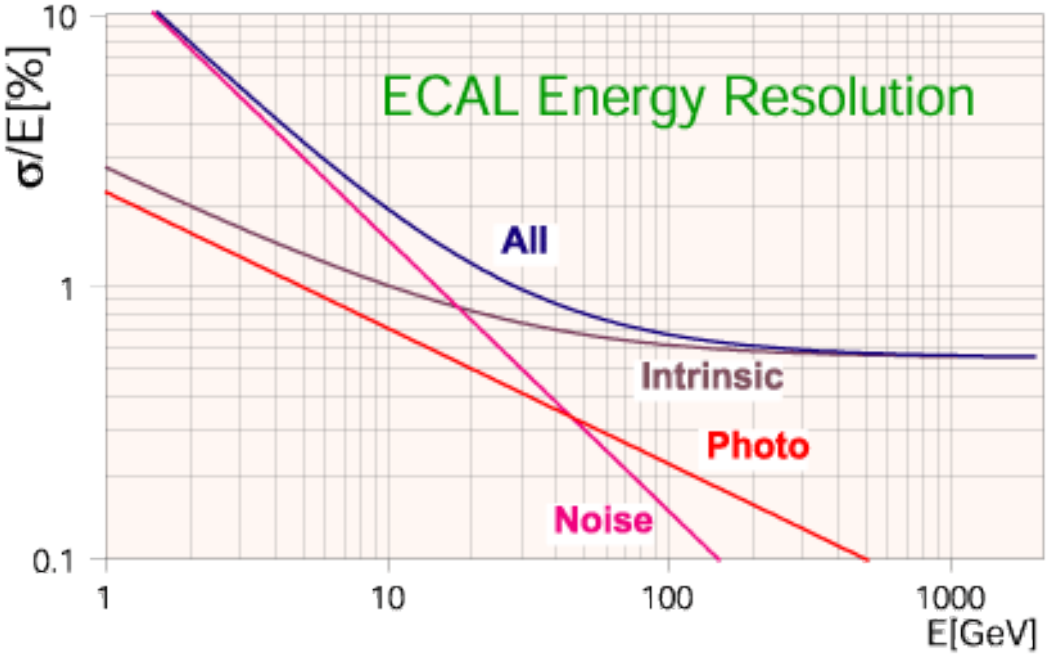
Ok, so we wait...

- $(S/\sqrt{B}) \sim L/\sqrt{L} \sim \sqrt{L}$
- $L = 1.25$ fb^{-1} for $M_H > 400$ GeV/c^2 for $S/\sqrt{B} = 5$.
- $L = 25$ fb^{-1} for $M_H < 400$ GeV/c^2 for $S/\sqrt{B} = 5$.



Complications from real life

- PYTHIA is only LO. Due to QCD effects, the Monte Carlo may be off by factor of 5-10.
- No inclusion for rate of jets faking leptons.
- Detector inefficiencies and resolution ignored.



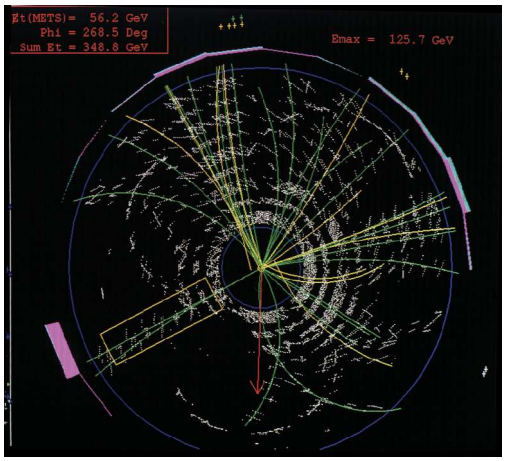
If the low-data analysis is tough now, these considerations will only make it worse. We need a promising start to make it worth the time to explore these complications.

Summary

- By considering Monte Carlo events, I developed an analysis for Higgs discovery in the so-called "Golden mode," H→ZZ that focused on kinematics of Zs and multiple-lepton objects, rather than full reconstruction.
- The purpose was two-fold;
  - To try something interesting that not everybody is doing, i.e. partial reconstruction and kinematics.
  - By doing this, to try to achieve greater sensitivity to the low-mass Higgs signal with a small amount of data. (This is difficult to do with full reconstruction.)

Conclusion

- For high-mass Higgs ($M_H > 400$ GeV/c^2), the full reconstruction seems to be a clear winner, beginning to show a significant signal at an integrated luminosity of 1 fb^{-1}.
- Z + lepton pT distributions gave a slight edge to a significant Higgs discovery for a low-mass higgs (< 400 GeV/c^2), though more data is required
- Using this analysis, we have to wait a year for more data.





Acknowledgements

I would like to thank the professors, students, and post-doctoral researchers in the UCSB HEP CMS analysis group, and especially my advisor, David Stuart for help, support, and insight whenever it was needed. (Which was quite frequently!)

Additionally, I would like to thank my family for always coming to these types of things. Special thanks to my brother, Steven, who was the target for many practice talks.

Additional slides