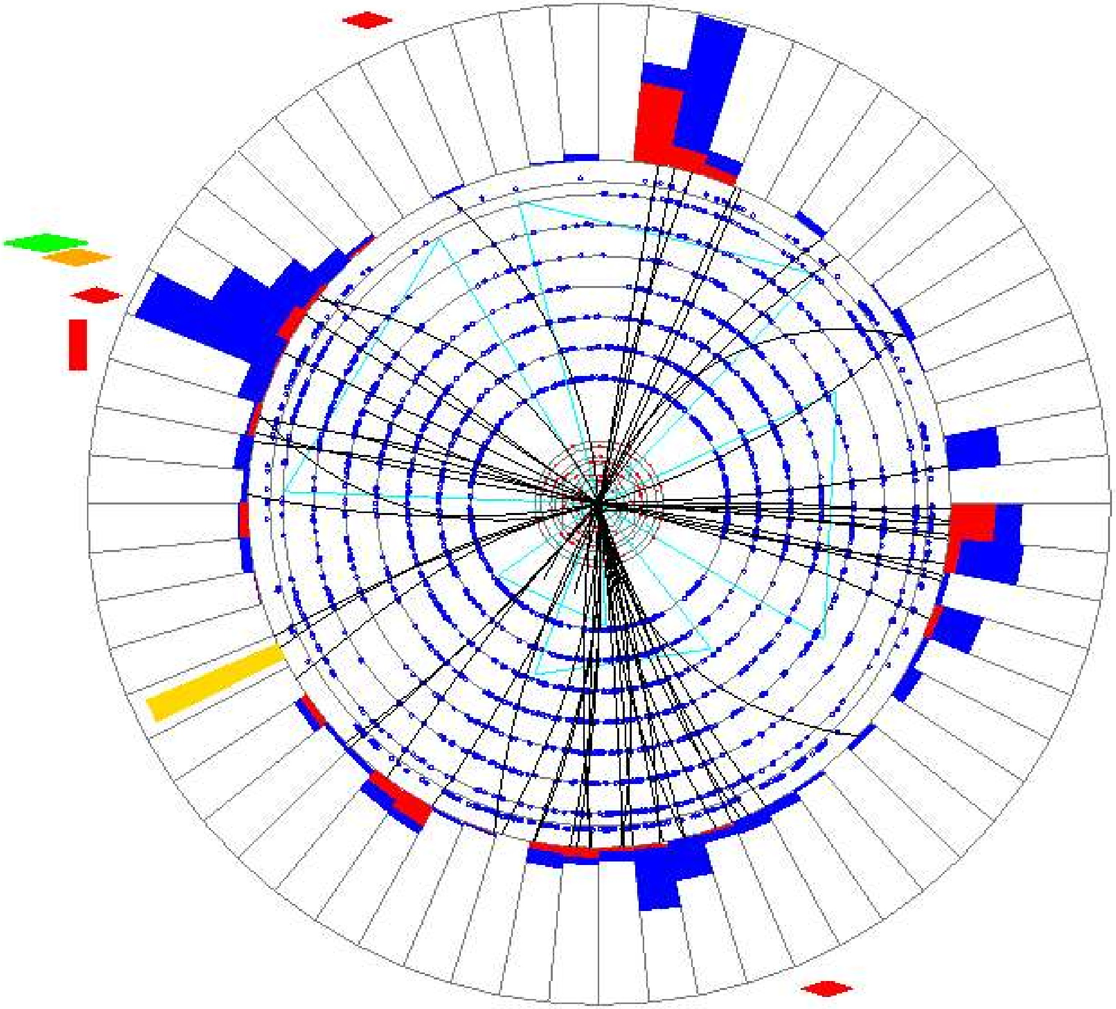
The Search for Neutral Higgs Bosons at DØ

Andy Haas
Columbia University
DØ Experiment

Columbia University
March 30, 2005



The Final Piece of the Standard Model

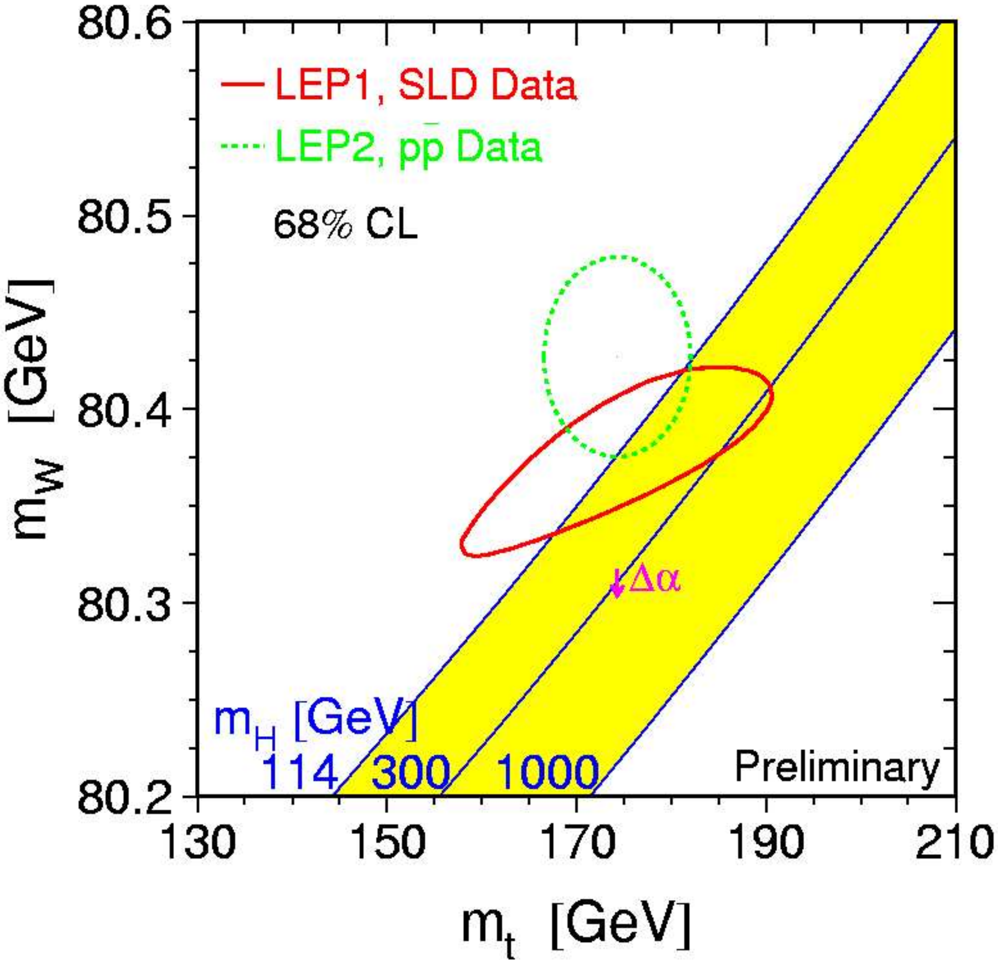
- The Higgs boson is needed for the Standard Model to be consistent

- Only one parameter of the Higgs is undetermined: its mass...

    - Precision measurements predict:

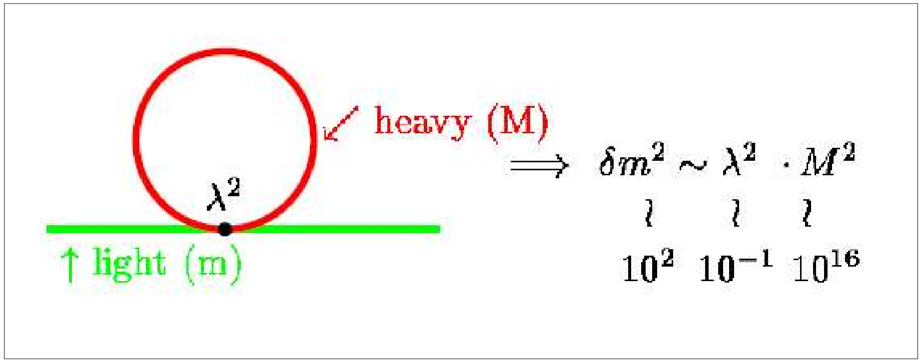
      $114 \text{ GeV} < m_h < 211 \text{ GeV}$

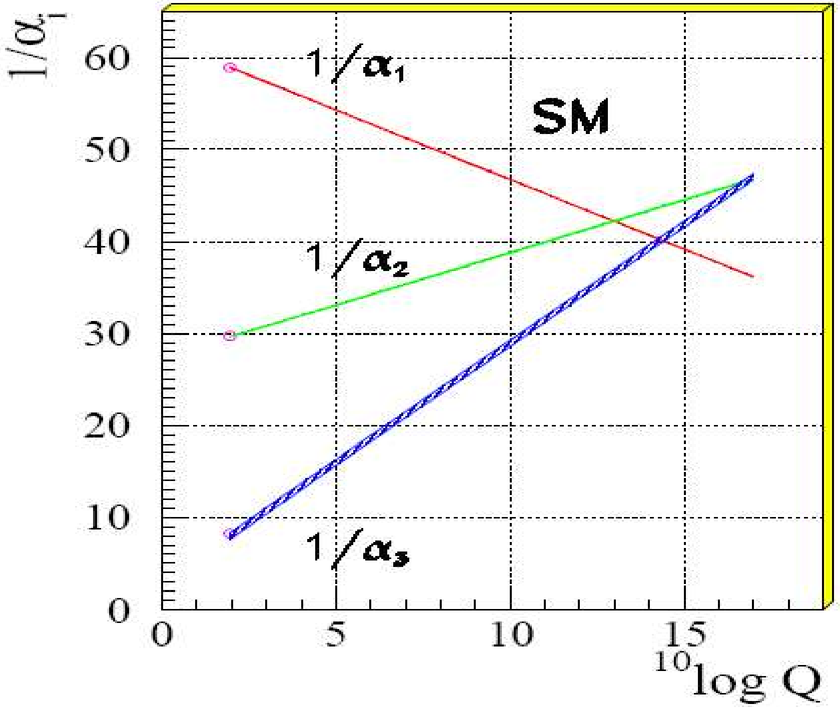
- We're looking hard for the SM Higgs at the Tevatron...



More to the Story?

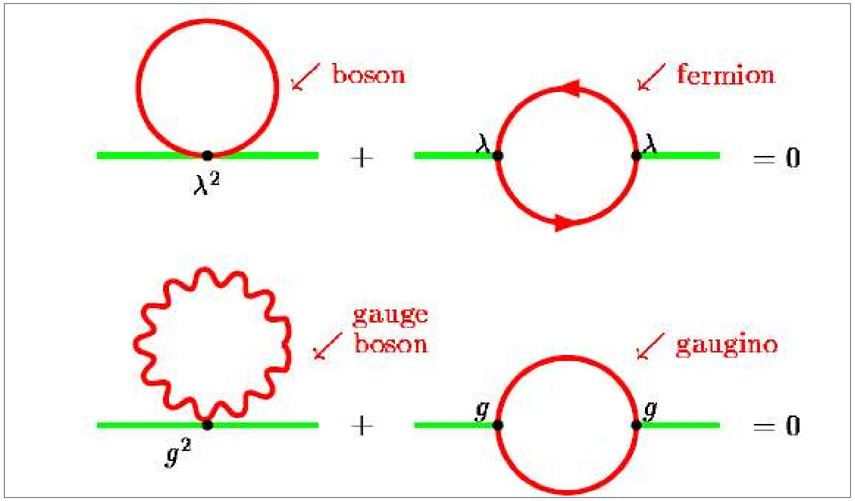
- Hierarchy problem:
  - Why is $m_h << m_{pl}$?
- Hierarchy stability problem:
  - QFT predicts radiative corrections
  - How does m_h stay $<< m_{pl}$?
- Can the gauge couplings unify?
  - Needed for Grand Unified Theories (GUT)
  - Ruled out in the SM !

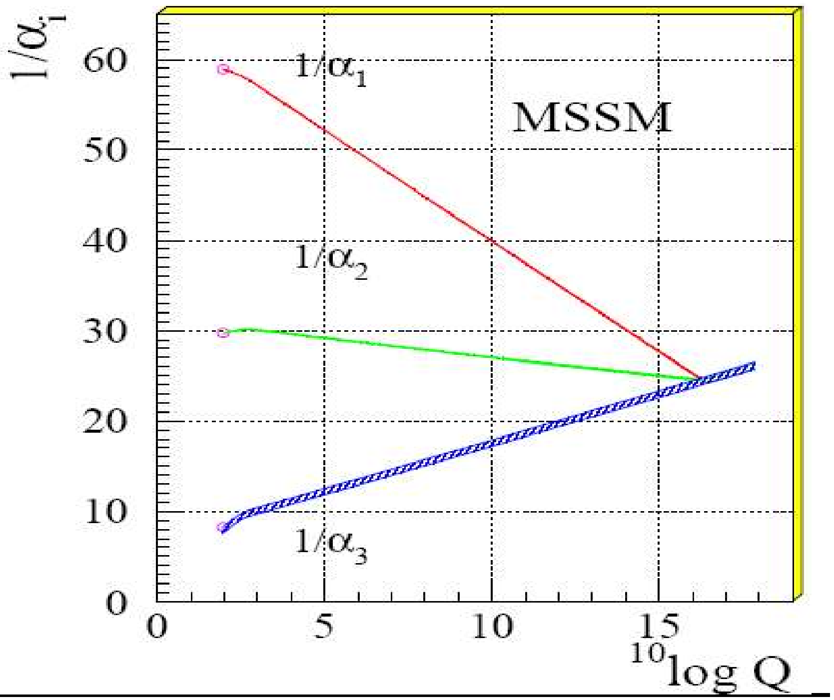




Supersymmetry to the Rescue?

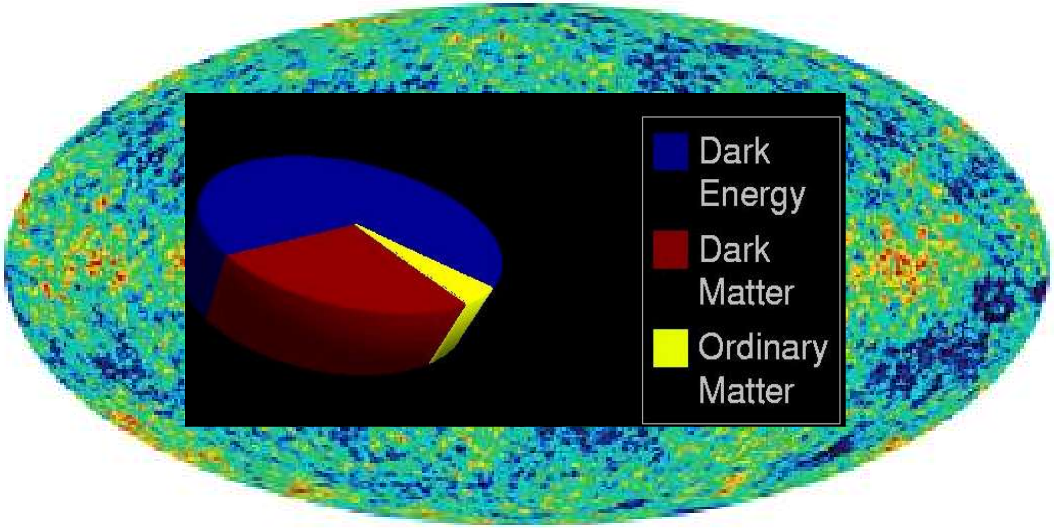
- Boson and fermion loop integral contributions differ by a factor of (-1)
    - A fermion for every boson: scalar field masses are stabilized
- Nobody works at Boselab… Supersymmetry is slightly broken
- Supersymmetry demands an even number of Higgs doublets
    - A light higgs is predicted (m_h<m_Z at tree-level)
    - With 2 Higgs doublets, couplings unify at 10^{16} GeV
        - This is the MSSM

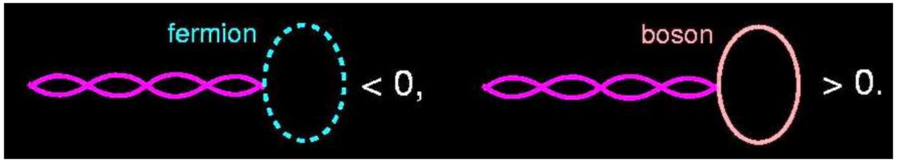


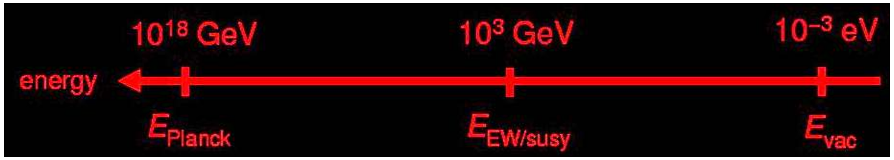


Cosmic Connections

- Supersymmetry has a natural dark matter candidate: the neutralino
  - $\tan\beta$ controls the annihilation rate in the early universe
- The Higgs fields could have CP-asymmetry
  - Extra matter / anti-matter asymmetry of the Universe
- Dark energy is a scalar field, like the Higgs
  - Supersymmetry solves the dark energy problem “halfway”…
- The inflaton is also a scalar field



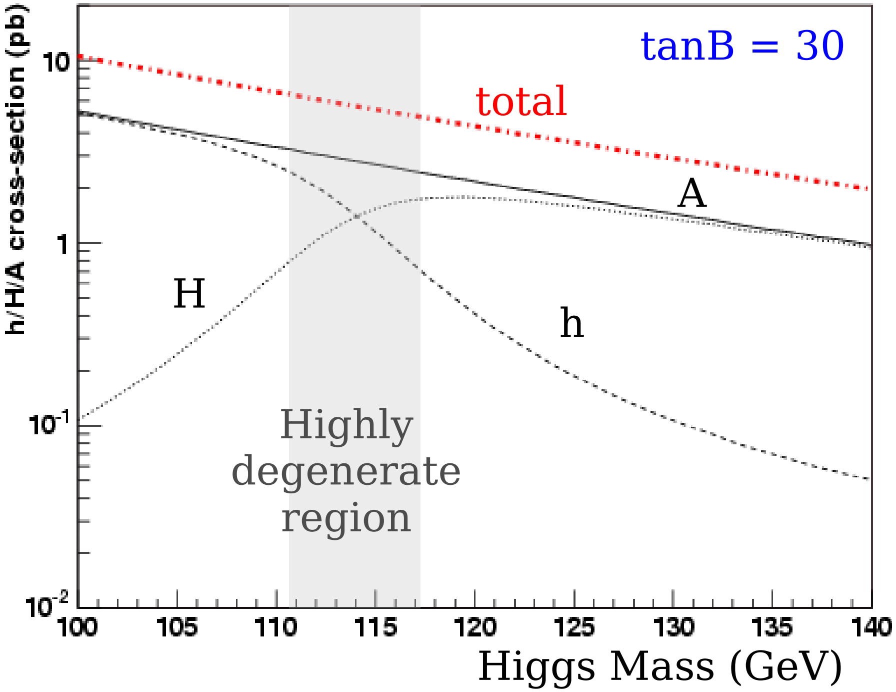


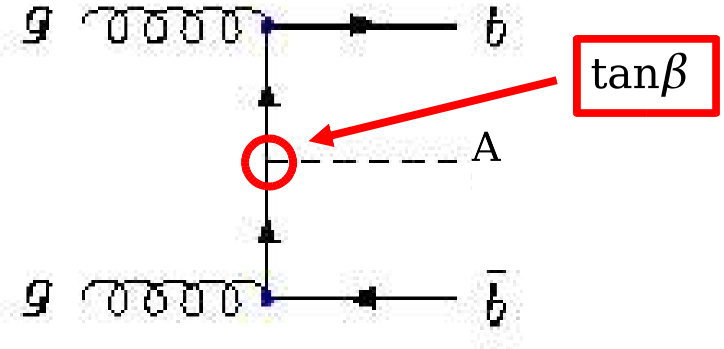


Higgseses in Supersymmetry

- Higgs fields come in pairs (H_u and H_d)
- 5 Higgs bosons : h, H, A, H+, H-
- $\tan\beta = H_u / H_d$
  - At $\tan\beta = \sim 50$, coupling of A to b = A to t
  - Cross-sections for bbh grow like $\tan^2\beta$!
- h is predicted to be light, <~135 GeV
  - LEP limits on Higgs mass -- 92 GeV (not the usual 114 GeV like in the SM...)
  - LEP limits are much looser (~50 GeV) if CP-violation is allowed in the Higgs sector... this talk assumes CP-conserving Higgses !
- At high $\tan\beta$ (>~20), the A is degenerate with the h,H:

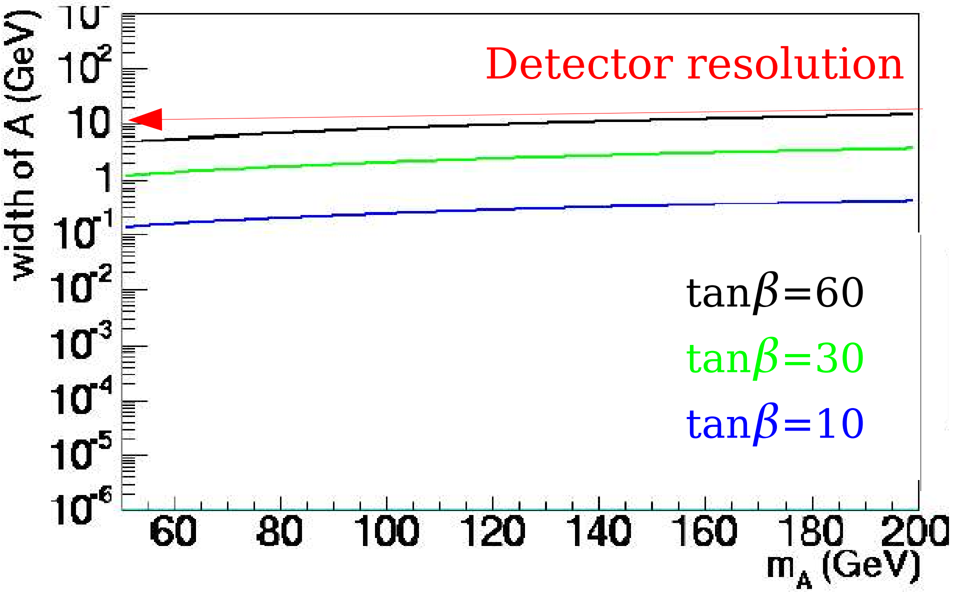
  cs(A) =~ cs(h/H)
  width(A) =~ width(h/H)
  m(A) =~ m(h/H)
  BR(A→bb)=~BR(h/H→bb) =~ 90%

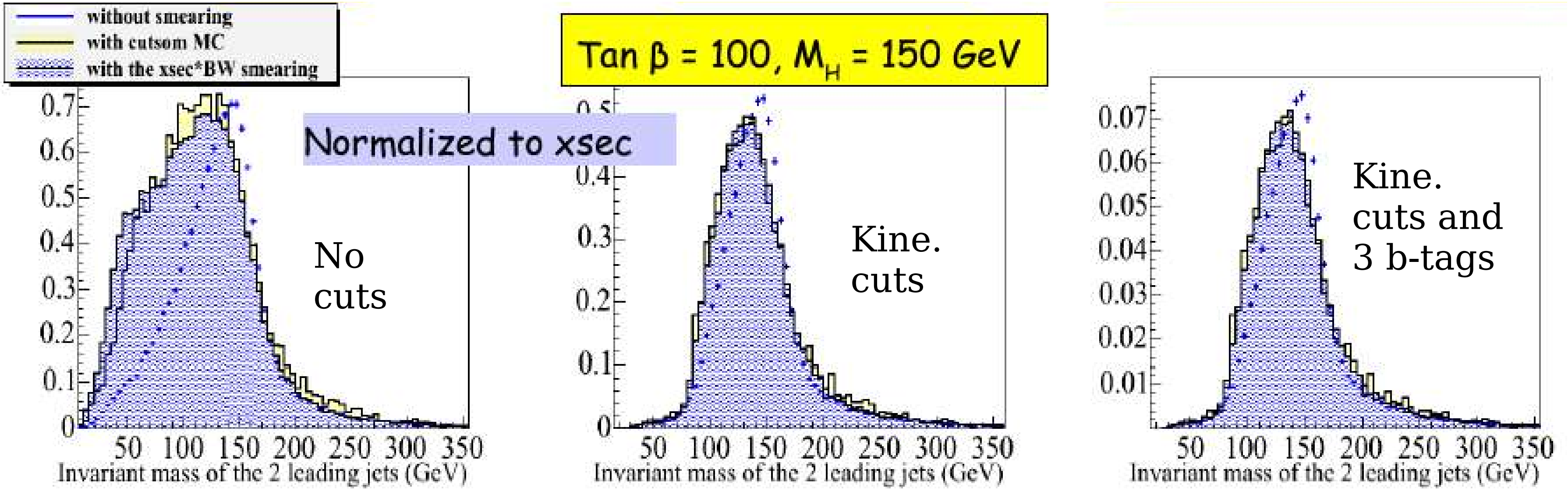




Higgs Widths

- Higgs bosons get wider, proportional to tan$^2\beta$
- Widths are modeled with BW shape, weighted by cross-section
- Not a large effect on sensitivity, for reasonable tanβ values, after considering kinematic criteria





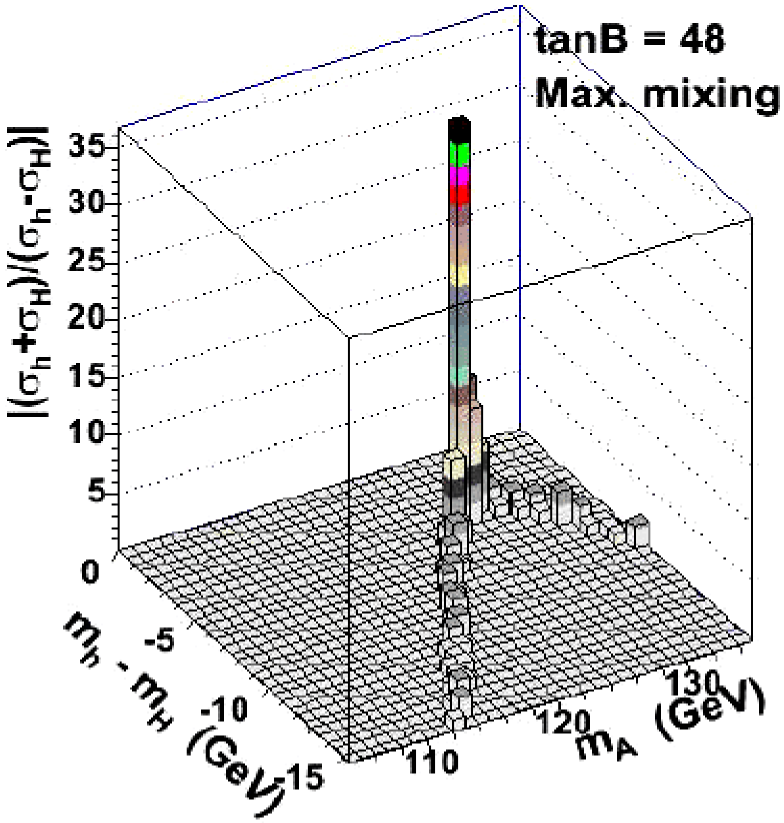
No cuts

Kine. cuts

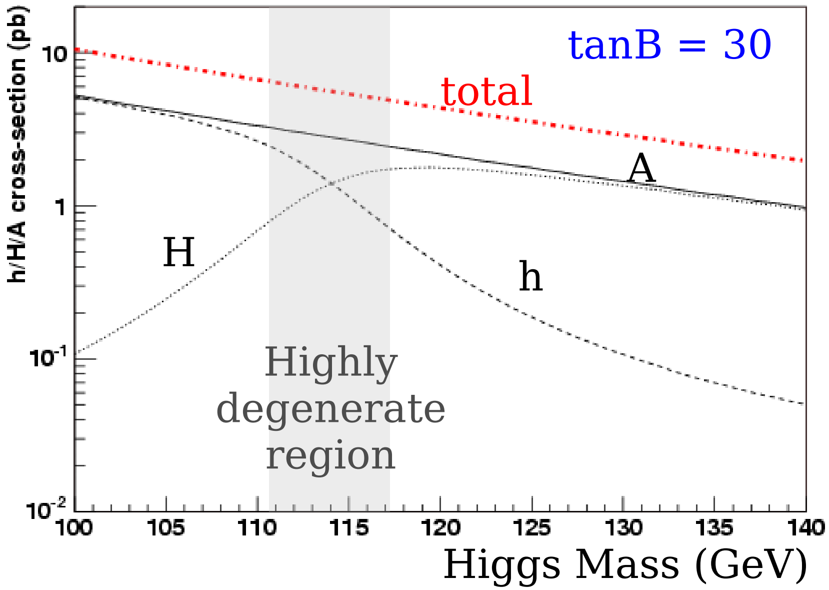
Kine. cuts and 3 b-tags

Mass-splitting of the Higgses

- What about when all 3 Higgs are nearly degenerate?
- What matters is the shape and height of the invariant mass peak
- When cross-sections are significant, the mass-splitting is small compared to the detector resolution

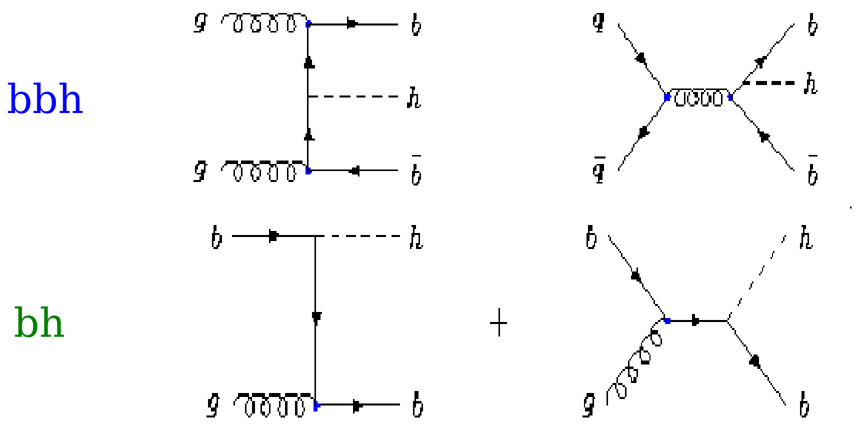


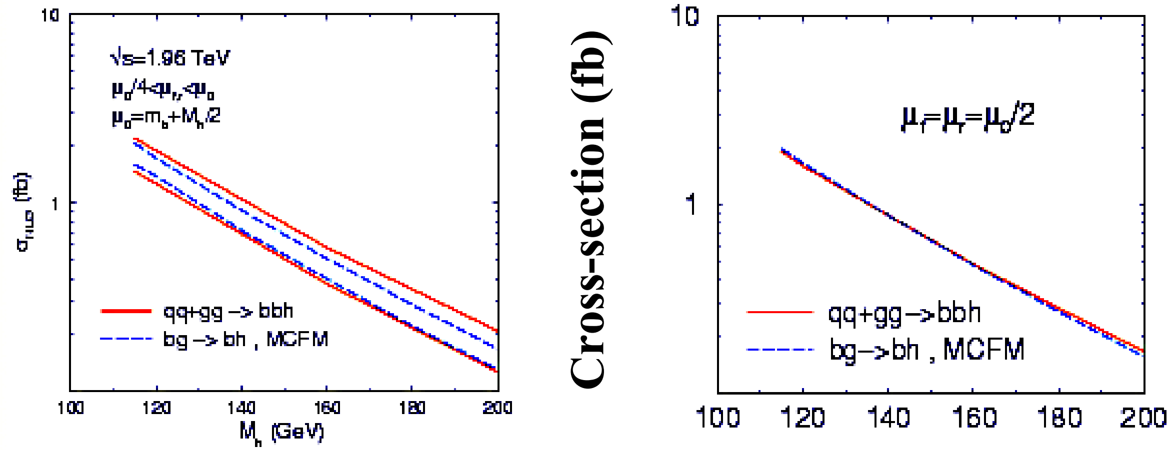
Mass splitting, weighted by cross-section



Signal at NLO in QCD

- There are two ways to calculate the signal production at a ppbar collider:
  - ppbar->bbh (4-flavor scheme)
  - ppbar->bh (5-flavor scheme) : use b-PDF
  - Both methods now agree at NLO
  - Uncertainty from PDF are comparable in size
- D0 has performed a measurement to test the predictions of the b-quark PDF calculations
  - Z+b / Z+j



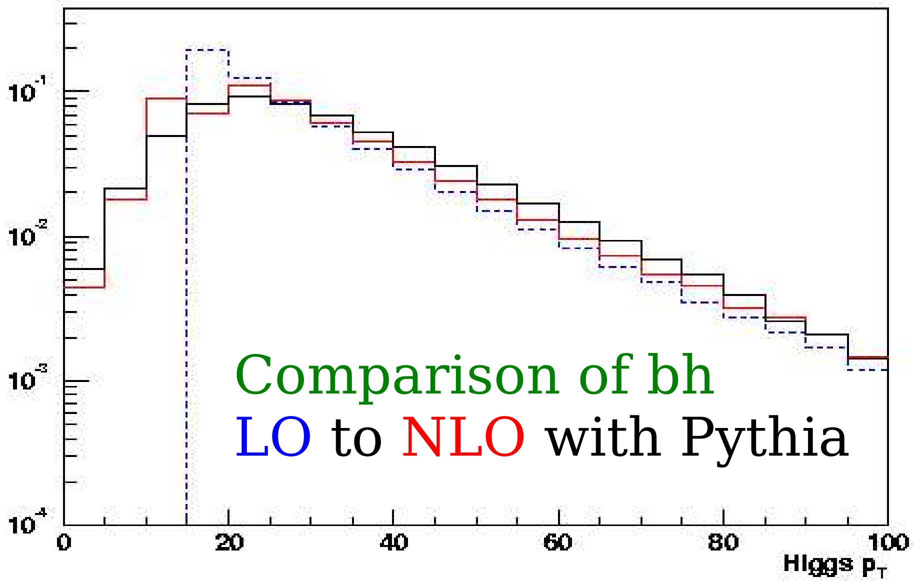


Z+b is a probe of b-quark PDF ⇒
σ(Z+b) / σ(Z+j) = 0.023 ± 0.005
= 0.018 (CTEQ6 prediction)

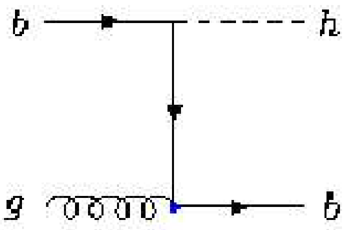
(DØ – hep-ex/0410078)

Monte Carlo Higgs Spectra

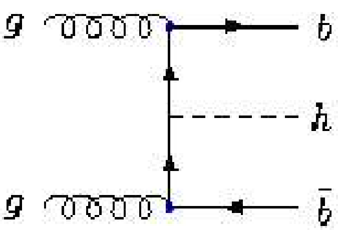
- Our Monte Carlo uses Pythia, which is semi-leading-order
- The events have been re-weighted, to mimic the NLO kinematics as closely as possible
- Comparisons of the bh and bbh processes are in good agreement



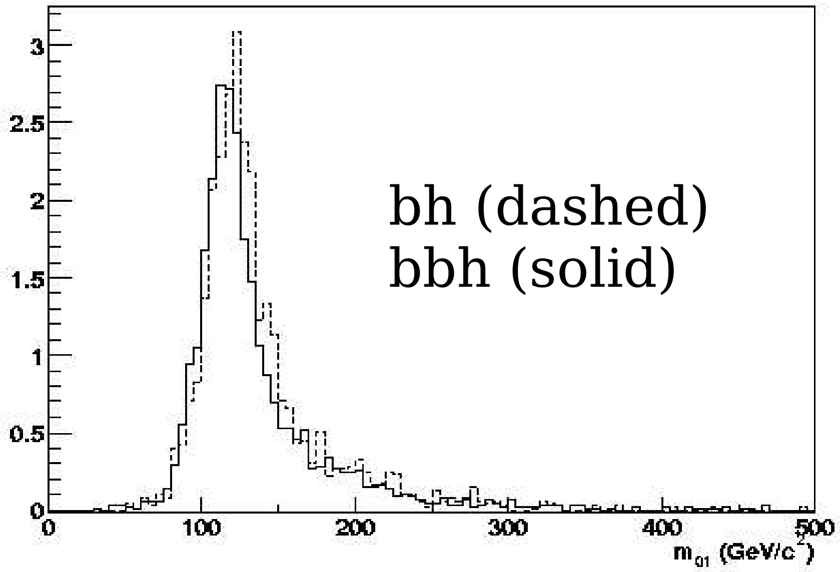
Comparison of bh
LO to NLO with Pythia



bh



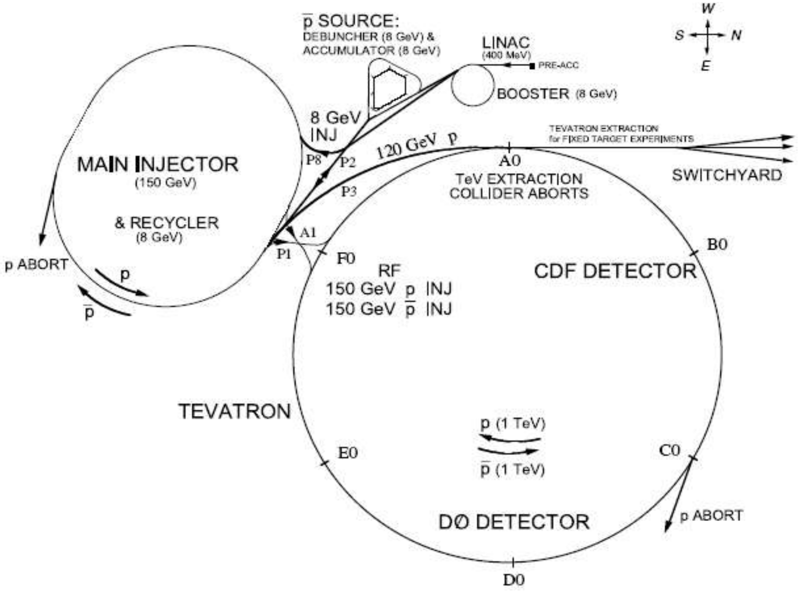
bbh

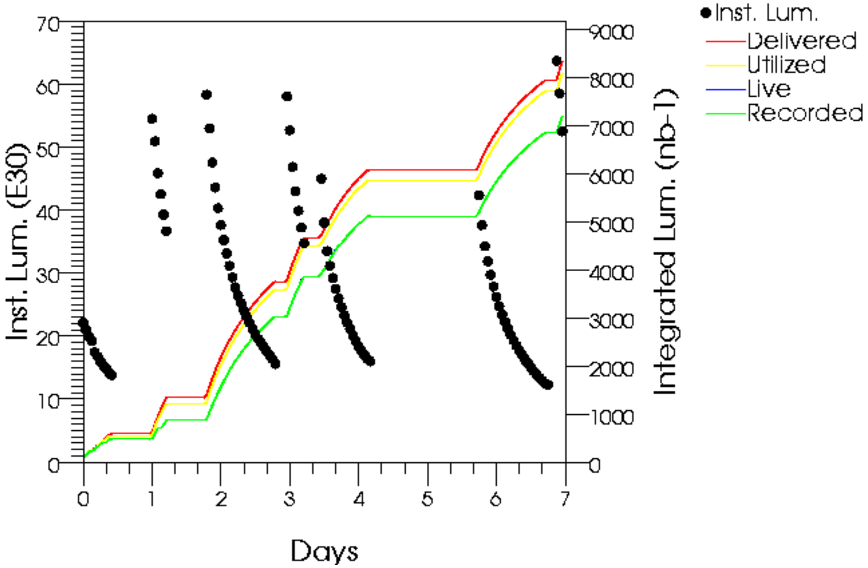


bh (dashed)
bbh (solid)

Fermilab Tevatron Accelerator – Run II

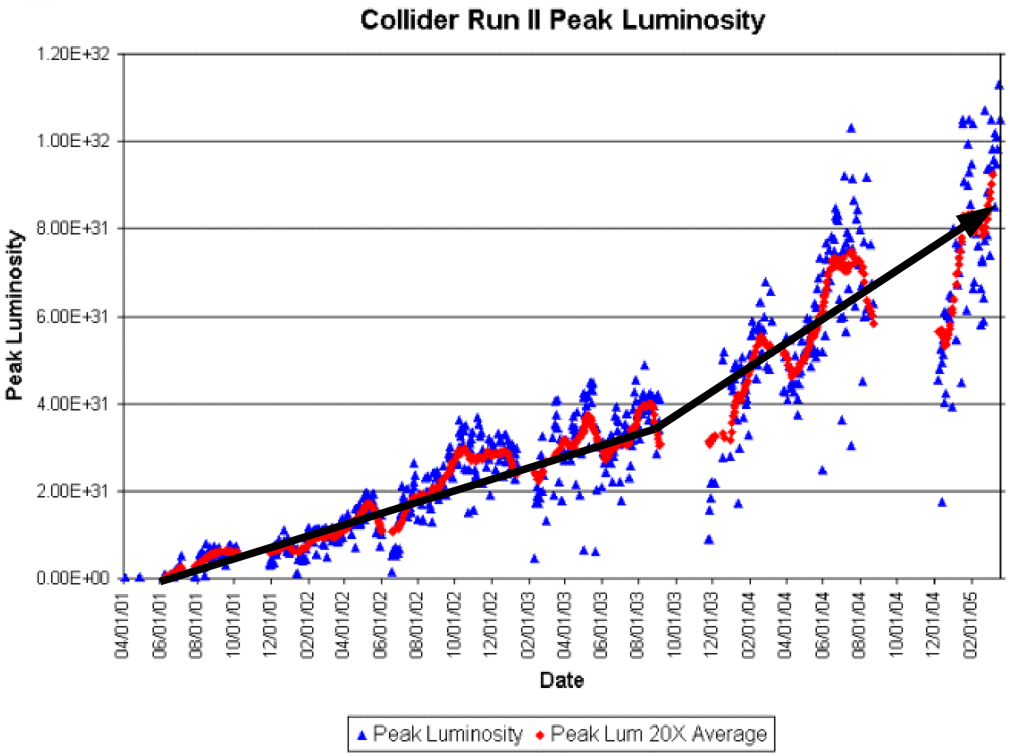
- 1 km radius p-pbar super-syncrotron
  - World's most powerful: 1.96 TeV
  - 36 bunches of p and pbar
  - 396 ns crossing period
  - 25 cm RMS interaction region
- One store every day or so...

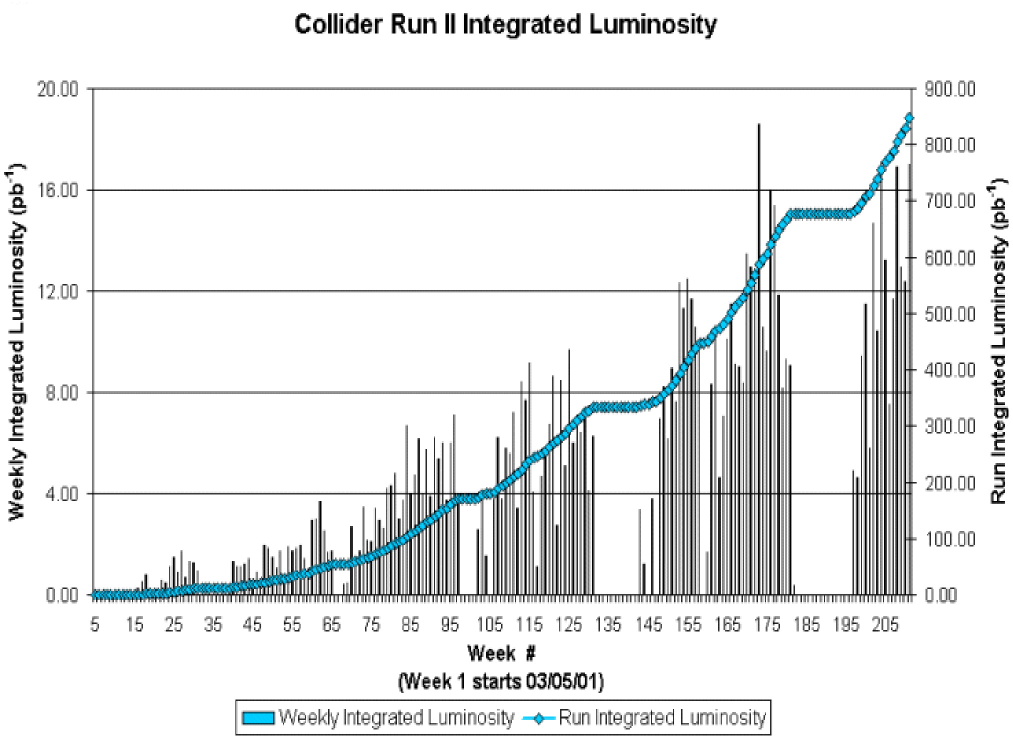




Tevatron Performance

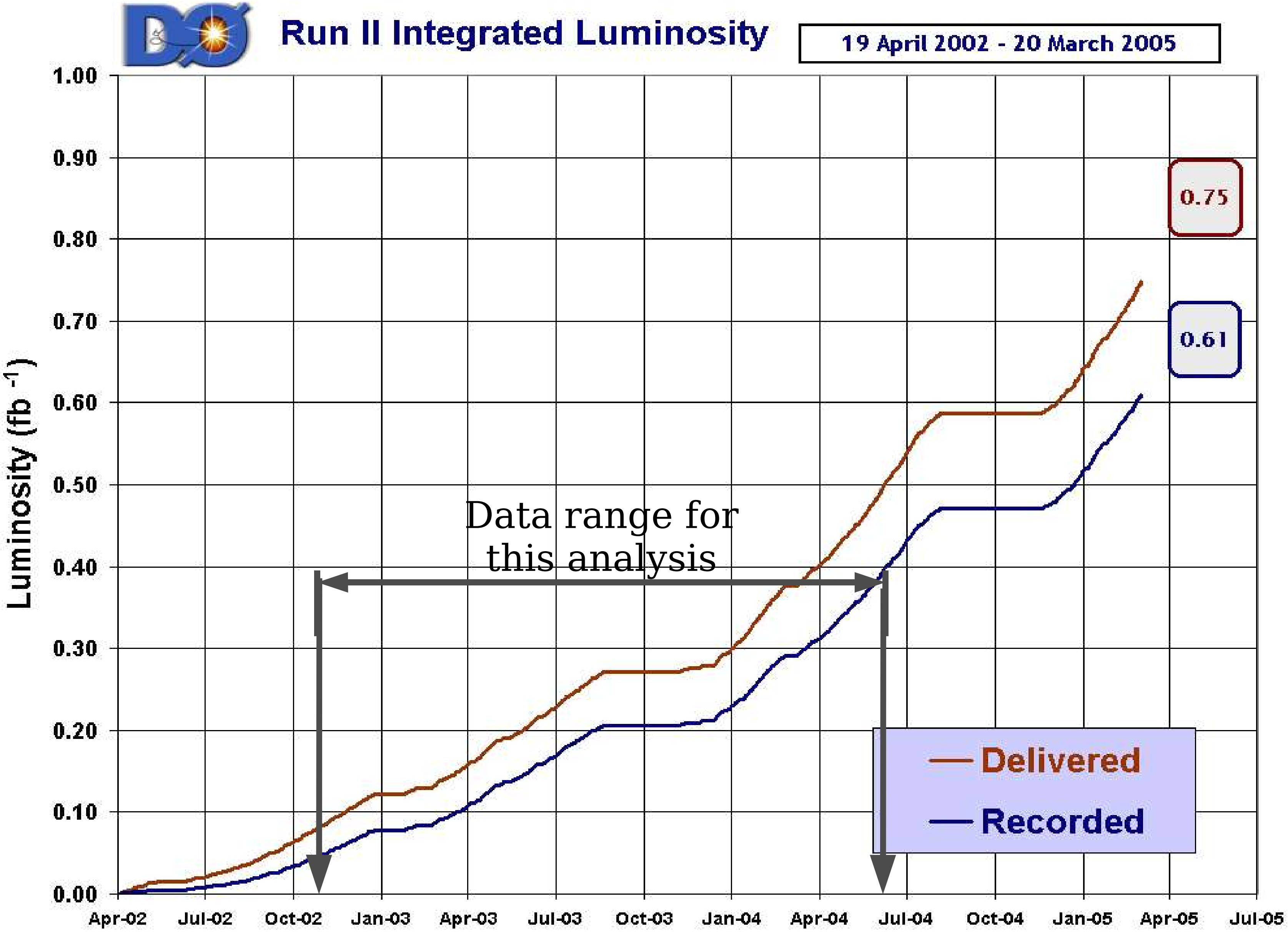
- Run II began 4 years ago
- Instantaneous luminosity has steadily increased
- Last year was excellent, and this year is also right on track
- Almost 1 fb^{-1} delivered so far
- Should collect 4-8 fb^{-1} in the next few years





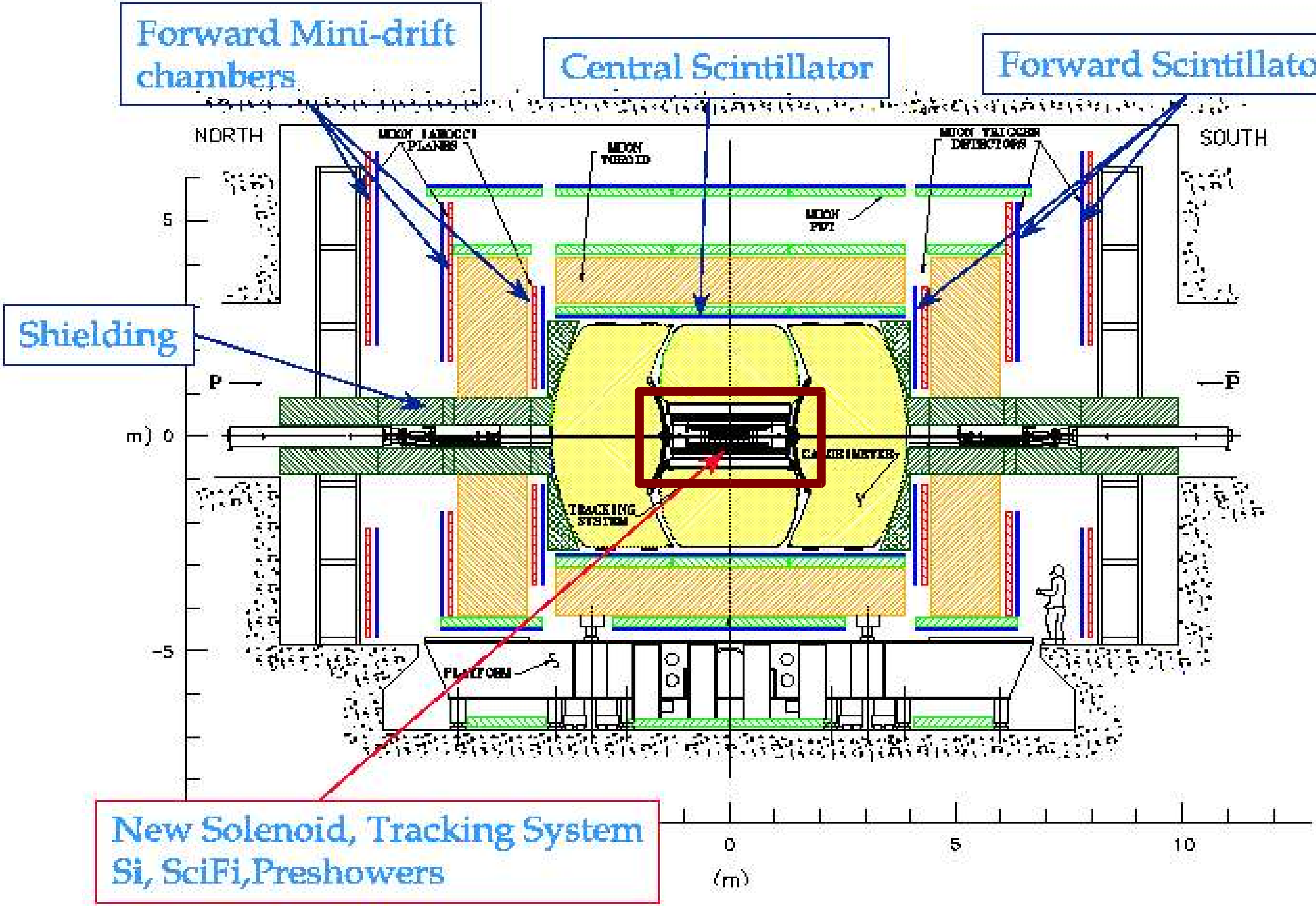
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Current Detector Data / Efficiency



Data range for this analysis

The DØ Detector for Run II


Forward Mini-drift chambers

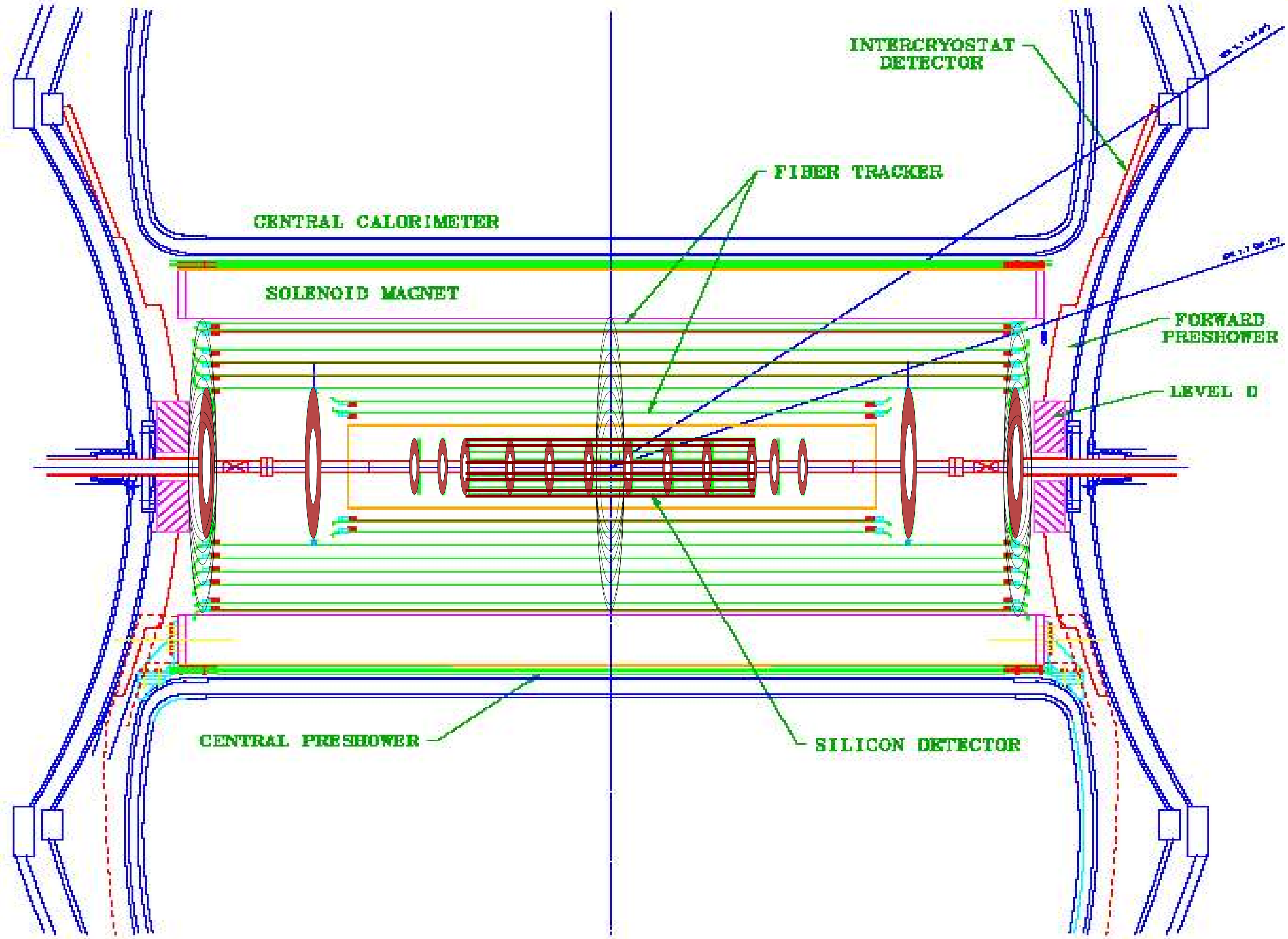
Central Scintillator

Forward Scintillator

Shielding

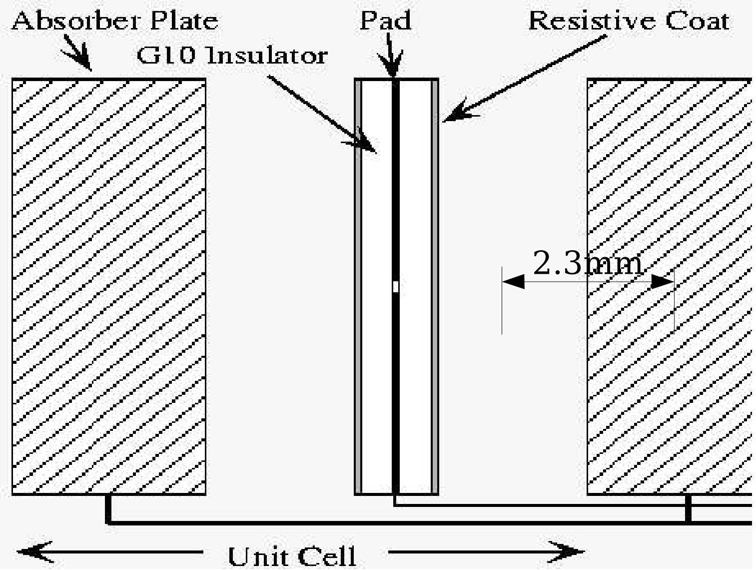
New Solenoid, Tracking System
Si, SciFi,Preshowers

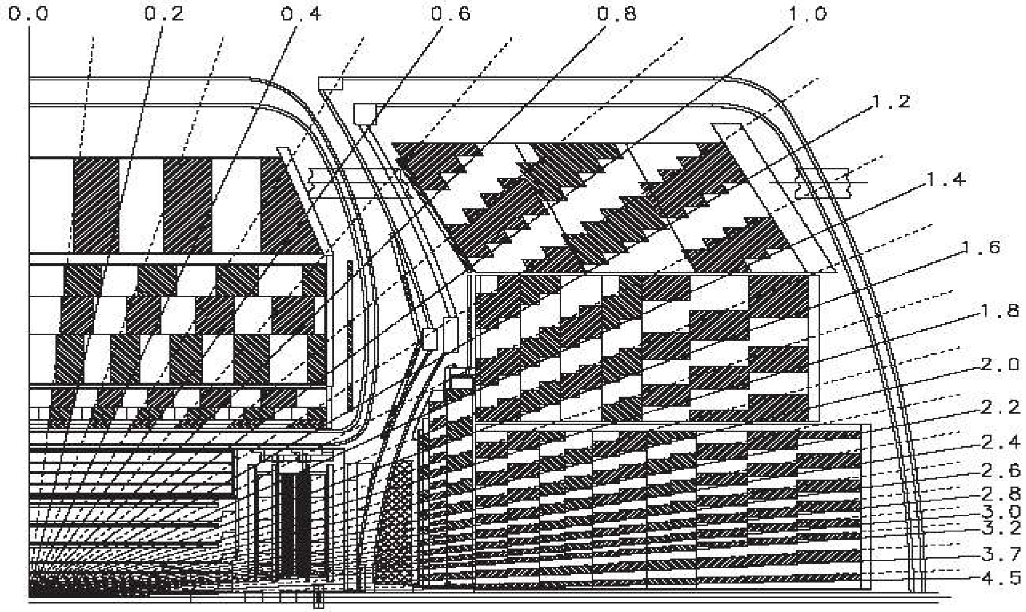
Tracking Systems



Calorimeters

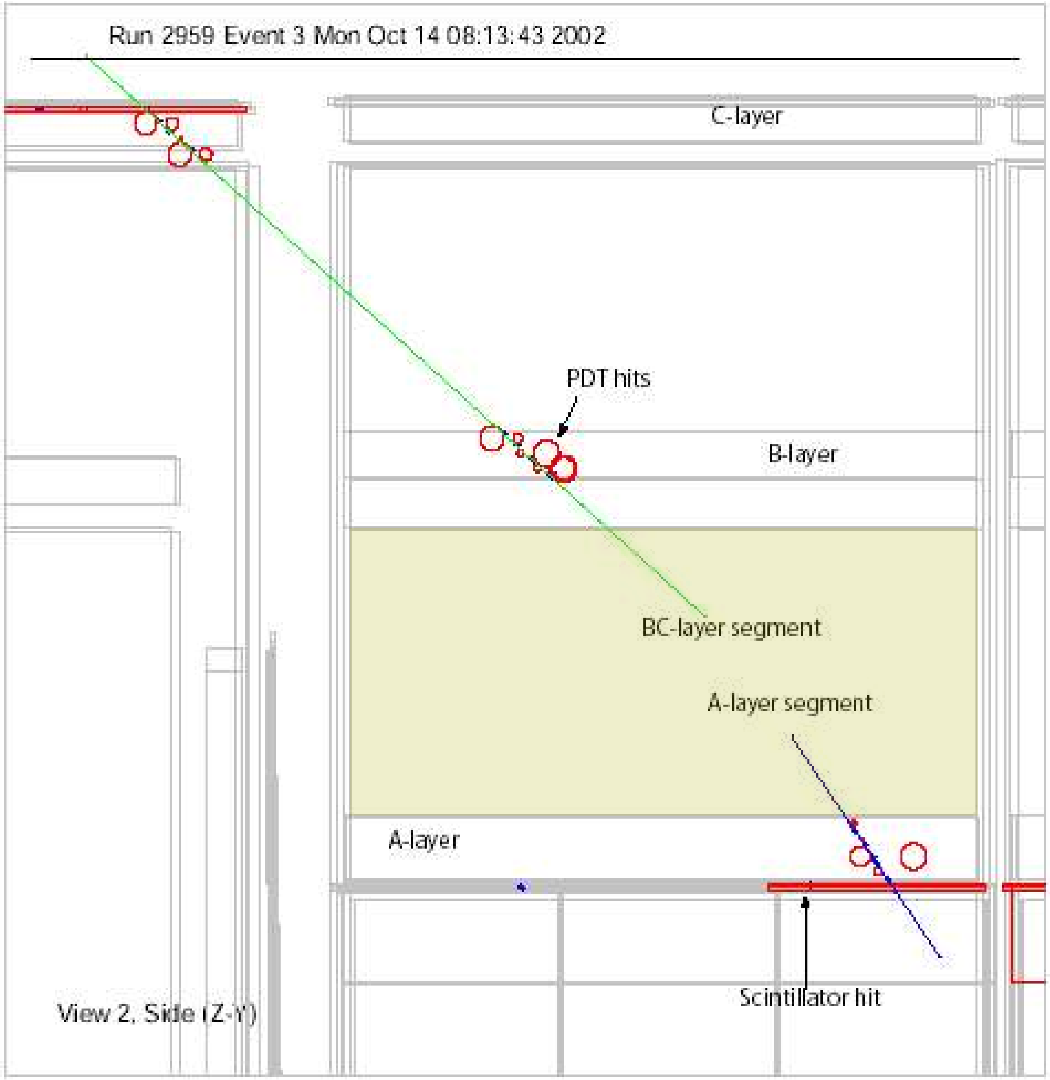
- Showers are induced by sheets of depleted Ur
- Cells measure the ionization in liquid Ar gaps
- Cells are arranged in a projective geometry
- Highly granular (.1x.1)(η-ϕ)
- Jets are almost completely contained: >20 hadronic interaction lengths





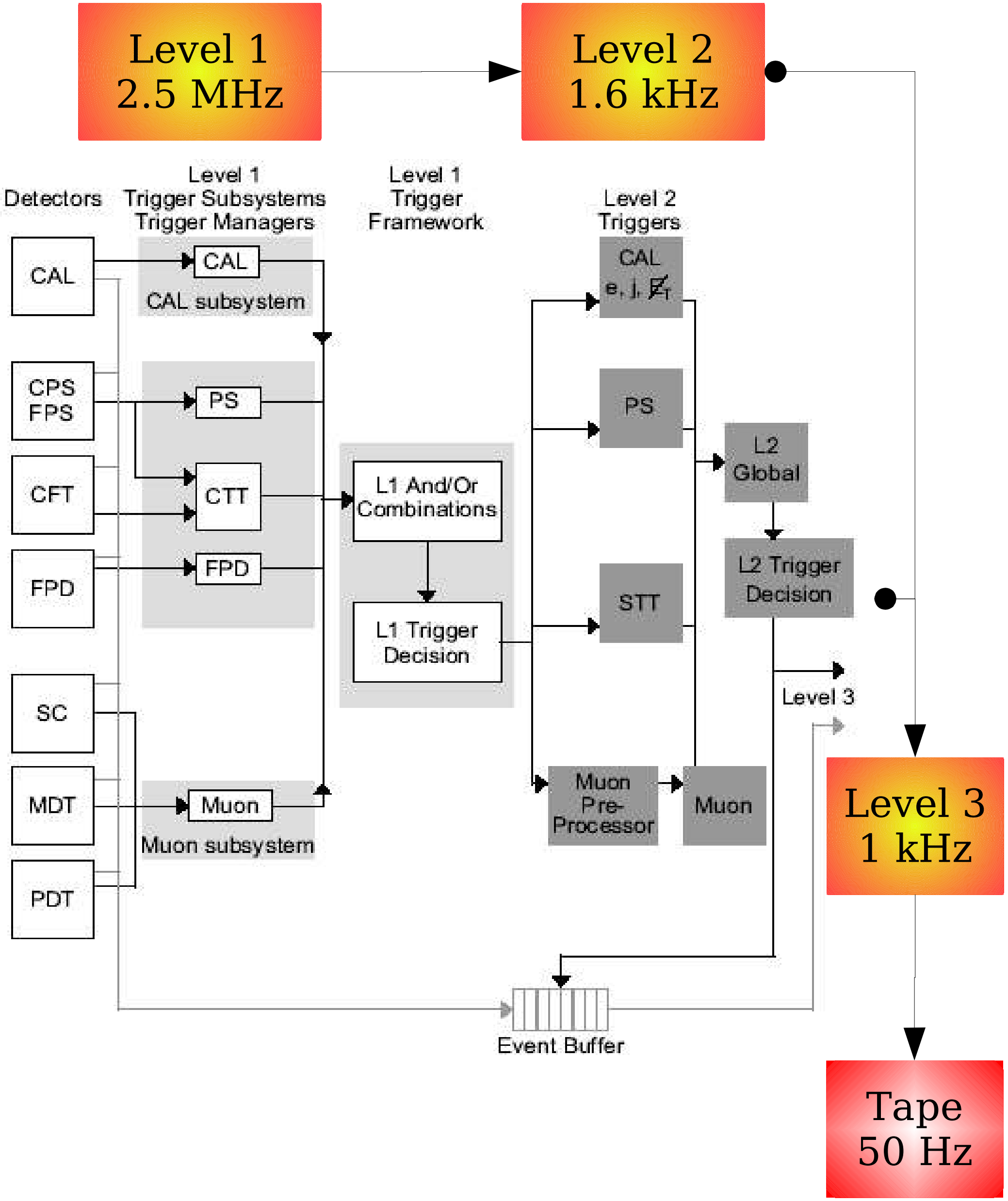
Muon System

- Used in this analysis for:
  - b-tagging studies…
  - jet energy corrections
- A big toroidal Iron magnet: ~2000 Ton and 1.9 Tesla
- Only muons with |p|>3 GeV penetrate the magnet
  - Very low background thanks to heavy shielding
- 3 drift-tube layers and scintillating panels determine the position and timing of muons
- Coverage out to |eta|<2.0
- A-layers and scintillators create the best muon trigger ever built



Trigger System

- Level 1
  - Calorimeter EM and TOT E_T (.2x.2)(η-ϕ)
  - Muon stubs
  - Central high-p_T axial tracks
- Level 2
  - Clustered calorimeter EM and TOT E_T
  - Missing transverse calorimeter E_T
  - Matched muon segments
  - Displaced high-p_T axial tracks
- Level 3 trigger algorithms:
  - Jets (no split/merge)
  - Central stereo tracks
  - Muons matched to central stereo tracks
  - Primary vertex
  - Jet lifetime b-tags

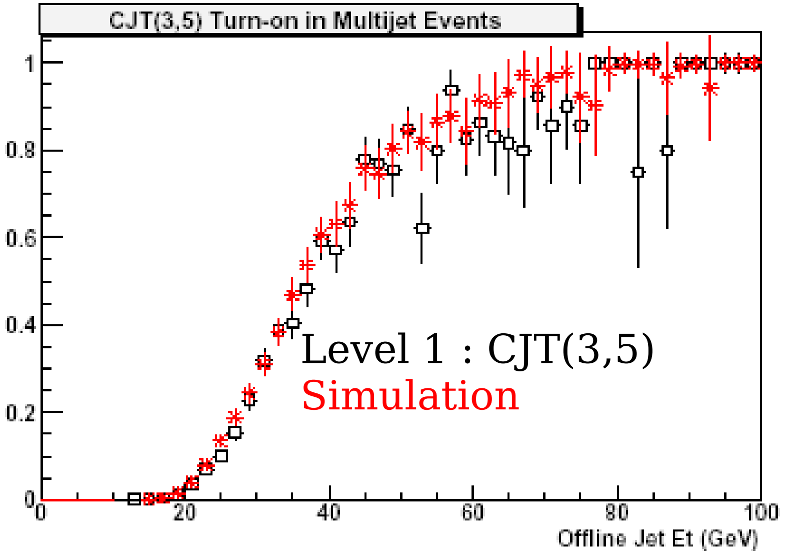

Level 1
2.5 MHz
Level 2
1.6 kHz
Level 3
1 kHz
Tape
50 Hz
Detectors
Level 1 Trigger Subsystems Trigger Managers
Level 1 Trigger Framework
Level 2 Triggers
CAL
CPS FPS
CFT
FPD
SC
MDT
PDT
CAL
CAL subsystem
PS
CTT
FPD
Muon
Muon subsystem
L1 And/Or Combinations
L1 Trigger Decision
CAL e, j, $\not{E}_T$
PS
STT
Muon Pre-Processor
Muon
L2 Global
L2 Trigger Decision
Level 3
Event Buffer

Data / Triggering

- Custom multi-jet triggers
    - Level 1:
      Three towers with TOT E_T > 5 GeV
    - Level 2:
      Three clusters with E_T > 8 GeV
      and H_T >50 GeV
  - Each triggerlist had new Level 3 criteria, as we improved them to remain within bandwidth limits while retaining signal efficiency

  Three jets with E_T > 15 GeV with two above 25 GeV in |eta|< 3.0
  Jets use 0.5 cone algorithm and
  PV Z position to correct jet E_T and eta
  Require a good L3 PV with |z|<35cm

- 260pb^{-1} of good data
  - 87.5 Million events

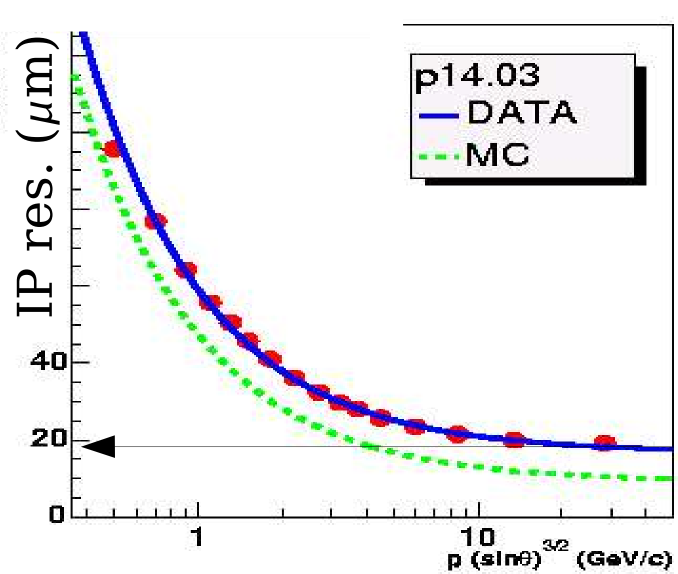


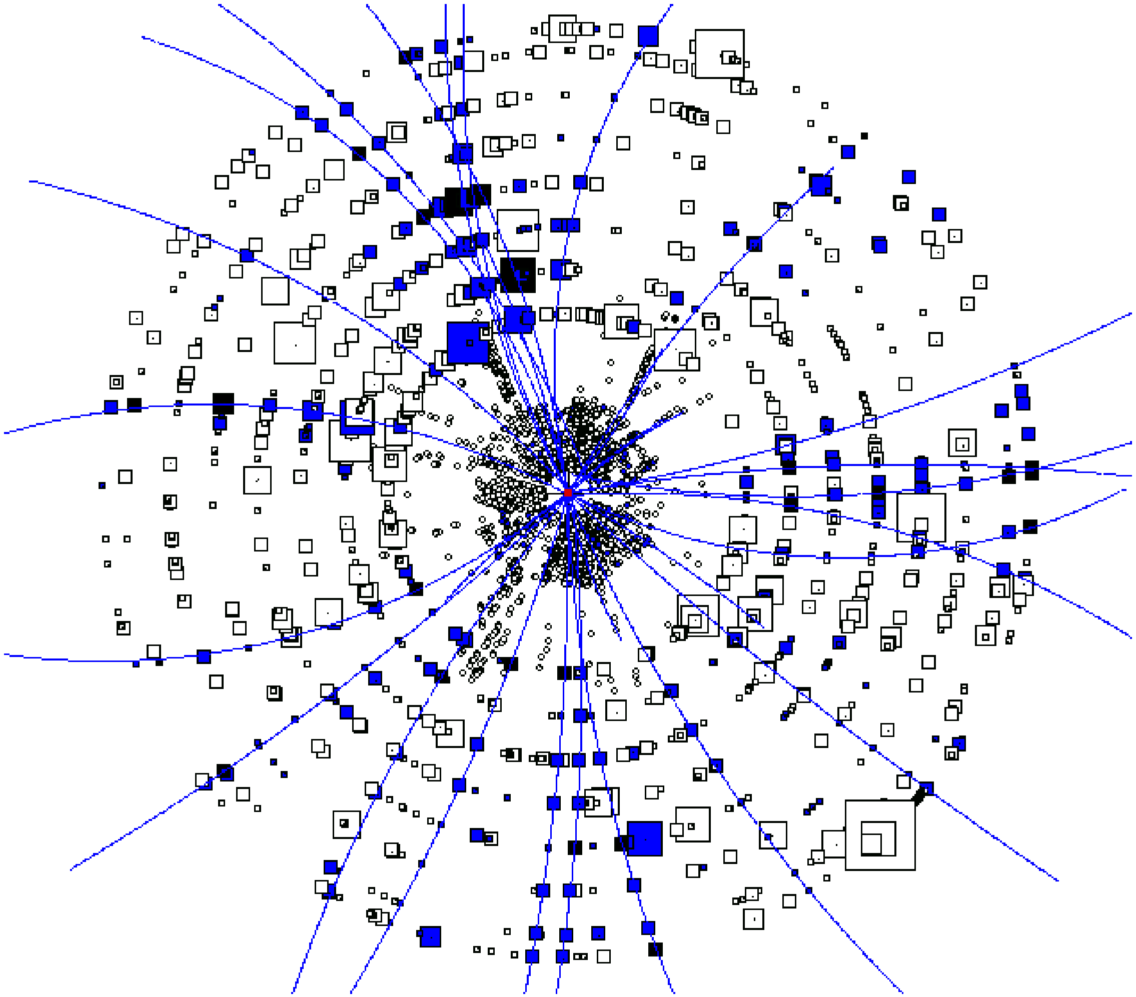
Trigger efficiencies measured using data:
> mu-based unbiased trigger for L1/L2
> special run for L3 study

68-80% efficient for m_A of 90-150 GeV

Track Reconstruction

- Reconstruct detector energy into clusters
- Pattern recognition combines clusters into tracks
- Efficiency for p_T>1GeV is about 95% in the central region, 85% for the forward
- Path resolution is ~20 microns

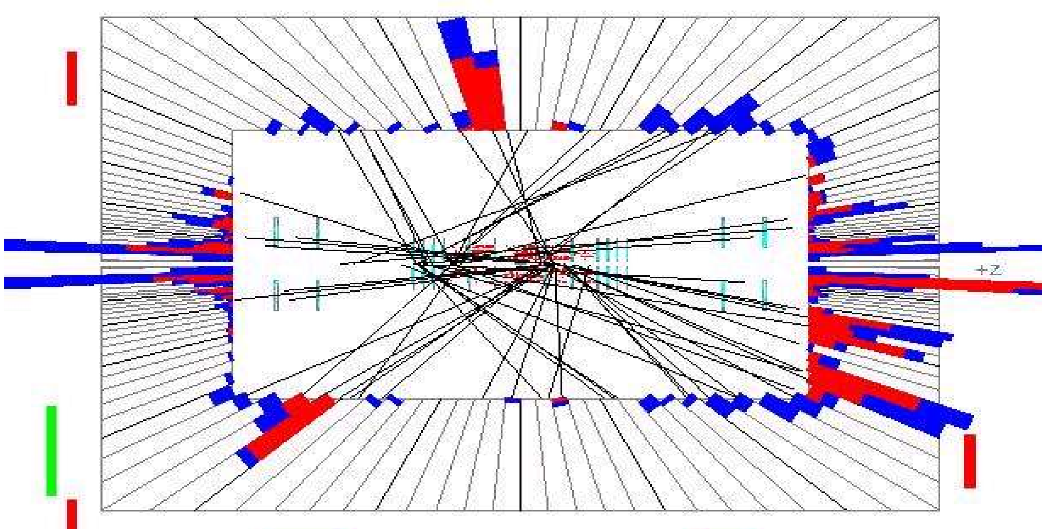


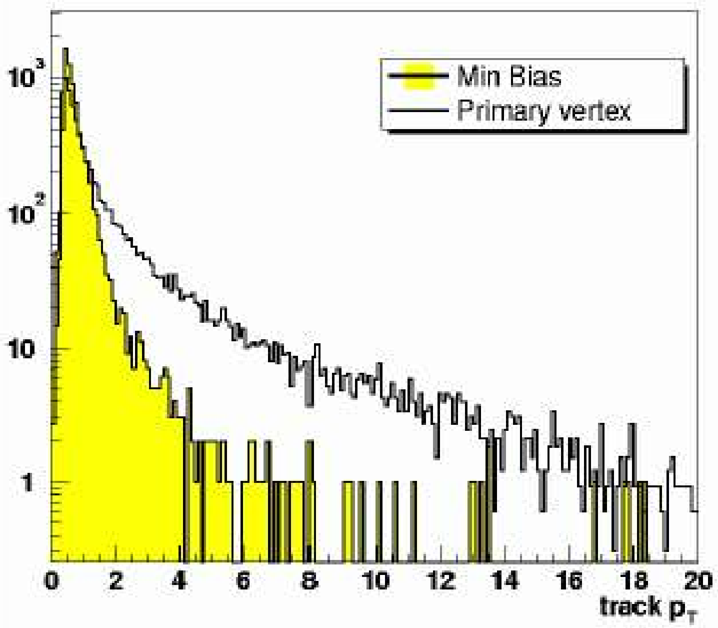


An axial view of a 3-jet event in data, with reconstructed clusters and tracks

Primary Vertex Reconstruction

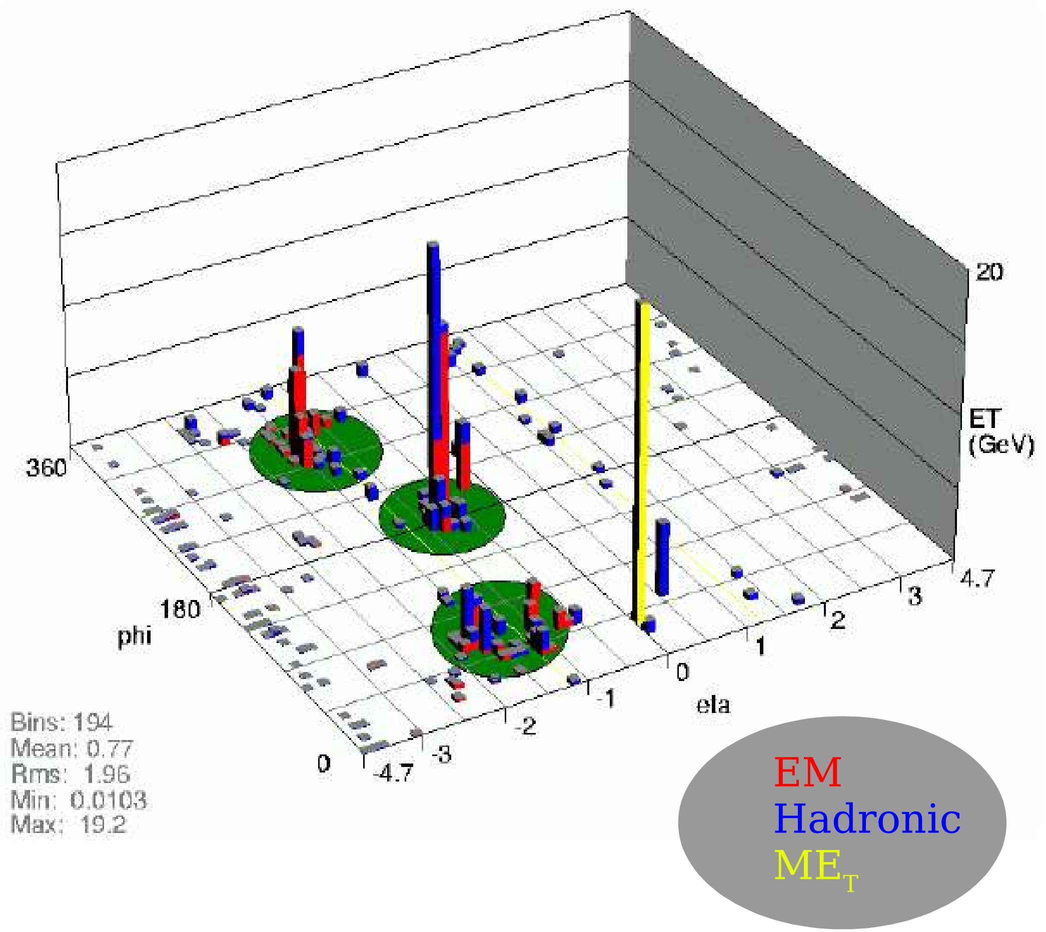
- Common locations of tracks along the beam-line are clustered into “primary vertices”
- There can be more than one interaction in a crossing
  - Very unlikely to have more than one interesting event in a single crossing
  - About 0.4 interactions per crossing on average for this data set - depends on instantaneous luminosity
- The interesting vertex is selected based on the momenta of the vertices' tracks
- This “hard-scatter” primary vertex is used to calculate jet E_T





Jet Reconstruction

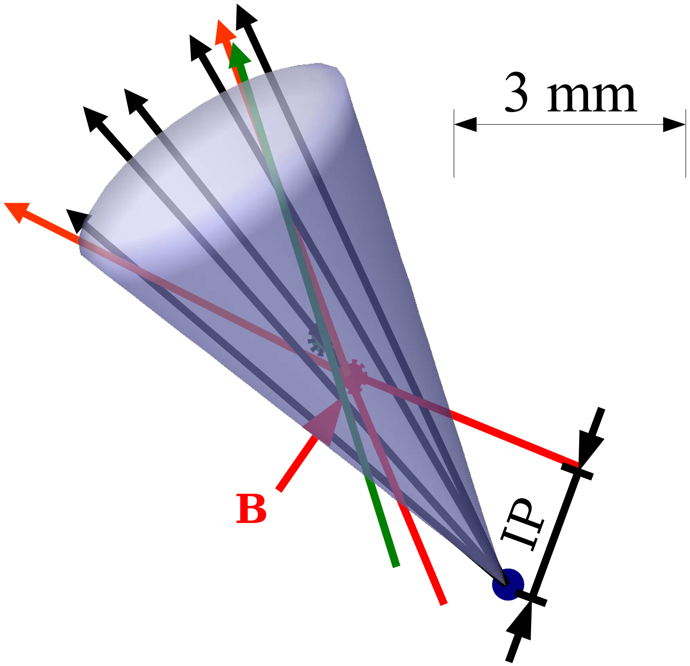
- The E_T in $(0.1\text{x}0.1)(\eta\text{-}\phi)$ towers is summed, and seeds are found
- Energy is clustered in cones of $\Delta R=\text{sqrt}(\Delta\eta^2+\Delta\phi^2)=0.5$ around the seeds
- The midpoints (in η-ϕ) between stable cones are also used as seeds
- The unique cones with $E_T>8$ GeV are merged or split, depending on whether they share more or less than 50% of a jet's E_T
- Jets' E_T are calibrated, using the jet energy scale (JES)
  - accounts for out-of-cone showering and the underlying event, on average



$$E_{jet}^{corrected}=\frac{E_{jet}^{colorimeter}-E_{offset}}{R_{jet}\cdot R_{cone}}$$

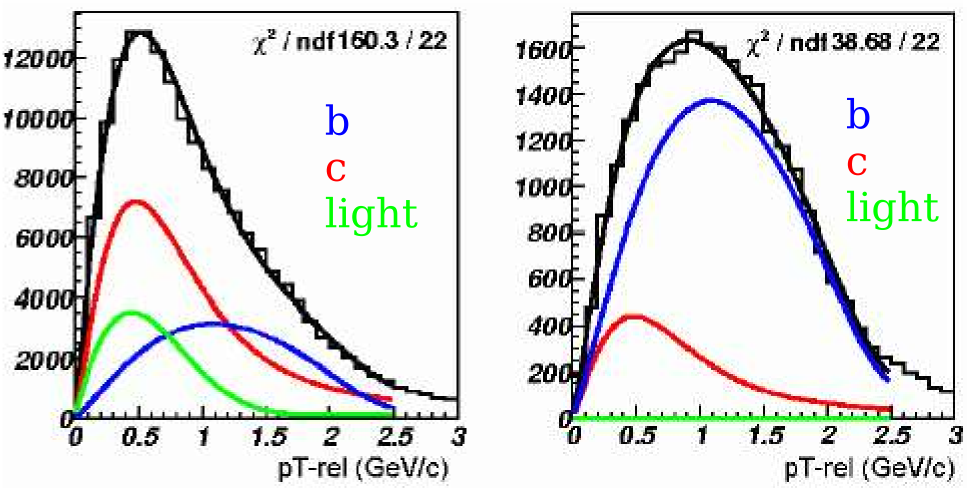
b-Jet Identification

- Identify jets containing bottom hadrons
  - main weapon again the huge QCD background
  - b-hadrons travel ~3 mm before decaying
  - Their mass leads to an opening angle between tracks --- large IP significances
  - Additional large IP significance tracks may arise from daughter decays (charms, tau)
- High-IP-significance tracks are used to find the secondary vertex(es)
  - Jets with a secondary vertex in $\Delta R<0.5$ are “tagged” as b-jets
- Backgrounds consist of:
  - mis-tags: poorly reconstructed tracks which randomly happen to form a vertex
  - long-lived strange / charm decays or photon to e^+e^- conversions
  - “gluon splitting” into nearly collinear charm or bottom pairs

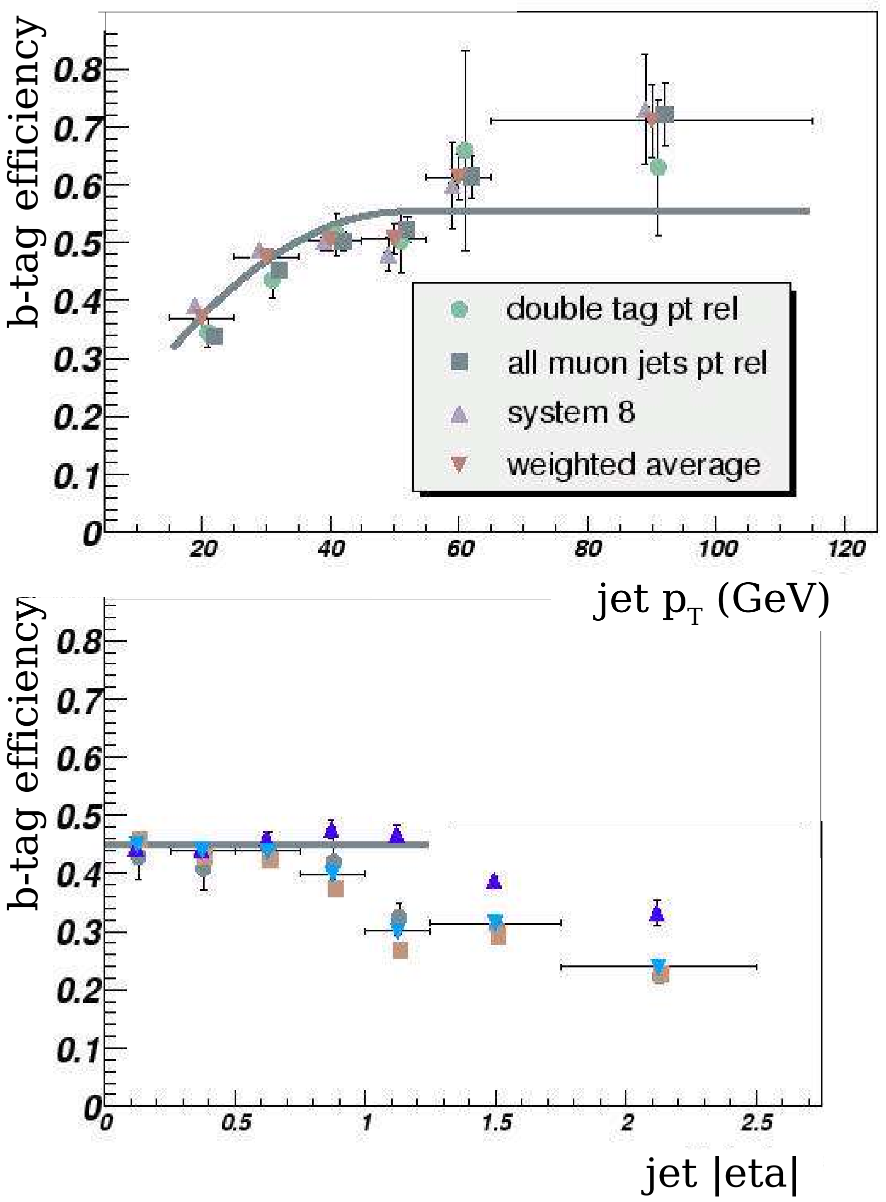


b-Tagging Efficiency

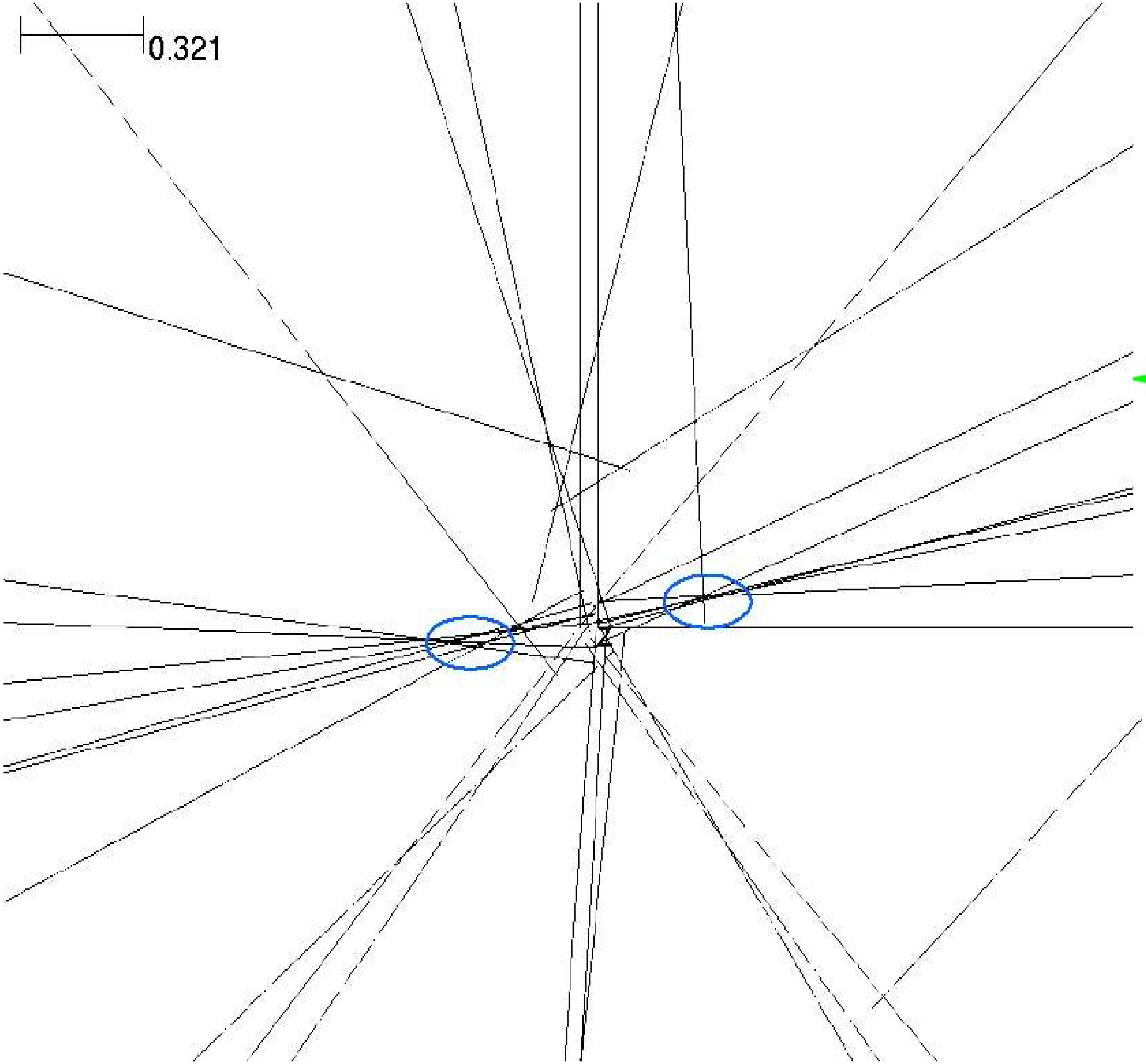
- The b-tagging efficiency is measured in data and compared to simulations
  - Very difficult because no pure b-jet data sample is available
- Use the p_T^{rel} spectrum of muons in jets to extract the b, c, light – jet fractions
- Then apply tagging and measure the flavor fractions again

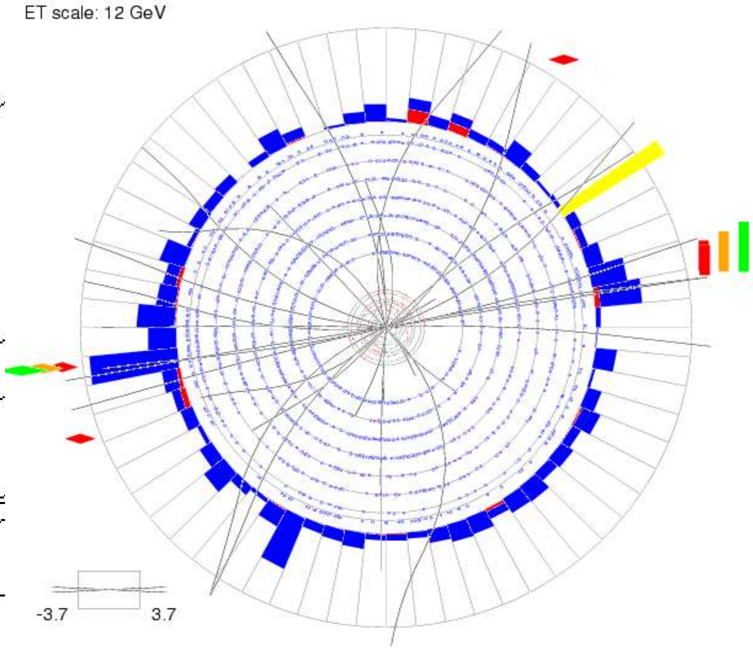


Muon P_T^{rel} spectra, before/after b-tag



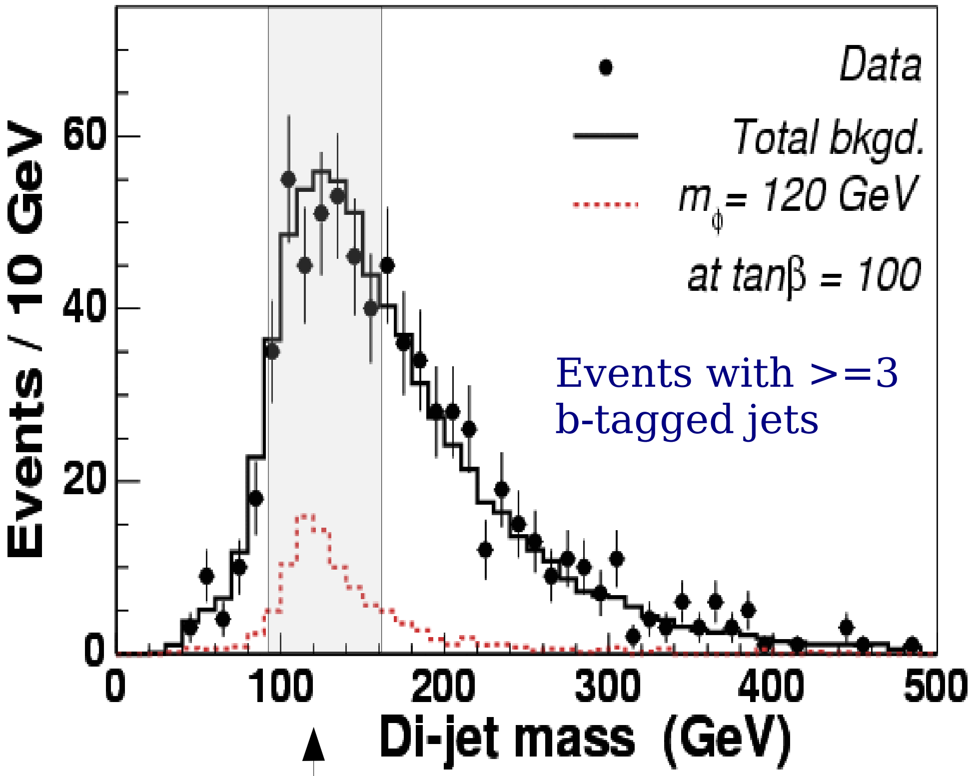
Secondary Vertices in Jets





Analysis Method

- Signal:
  - 3 or more b-tagged jets
  - Invariant mass of leading jets is peaked at m_A
- Backgrounds:
  - Determined from data!
    - Shape estimated from the double b-tagged data sample (taking into account the kinematic bias from requiring a 3rd b-tag)
    - Normalized outside the “signal region”
  - Also modeled in MC as a cross-check
    - “fakes”: all light-quark/gluon jets (measured from data)
    - “heavy flavor”: (ALPGEN) bbj(j), ccj(j), bbcc, cccc, bbbb
    - “other”: tt, Z(→bb)+jets (Pythia)



Events with >=3 b-tagged jets

“Signal window”

Triple b-tagged Background

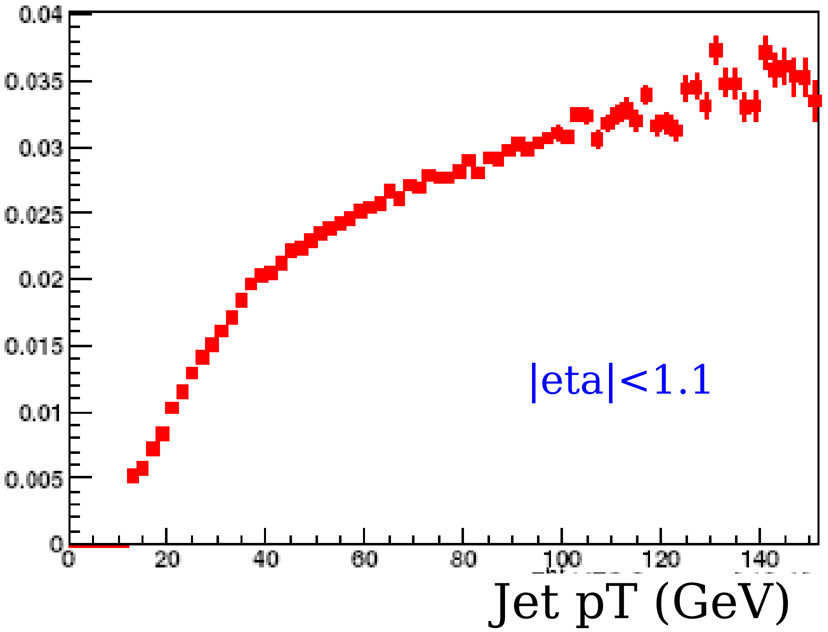
Full multi-jet data sample → Calculate TRF (E_T & 3 regions of η)

Double b-tagged data sample → Apply TRF to un-b-tagged jets

↓

Triple b-tagged background shape → Fit outside the signal region to data → Triple b-tagged background

Tag-rate-function (TRF)



Background is completely determined from data !

Cross-check Using Double b-tagged Data

TRF is iteratively corrected for “heavy-flavor” in the un-b-tagged data sample

Full multi-jet data sample

Calculate TRF (E_T & 3 regions of η)

Full multi-jet data sample

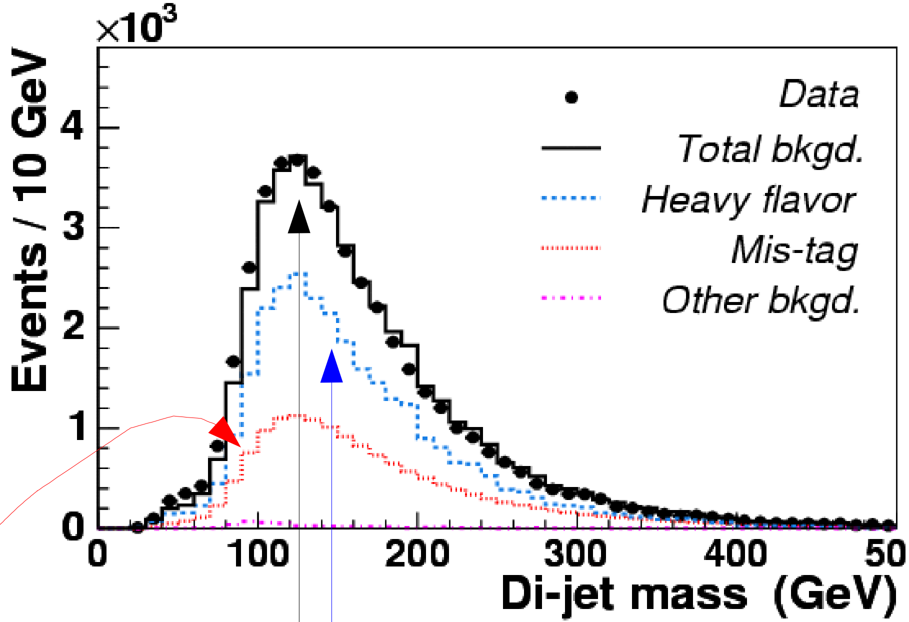
Apply TRF to un-b-tagged jets

Double b-tagged background

Compare

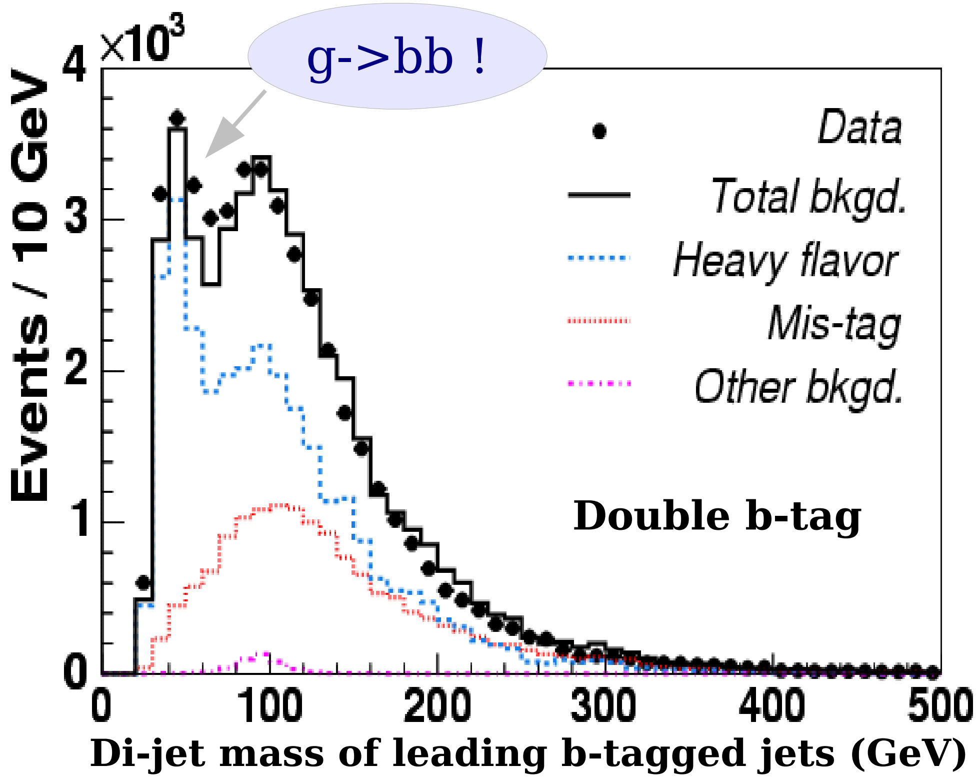
Double b-tagged data sample

Contribution from “heavy-flavor” MC

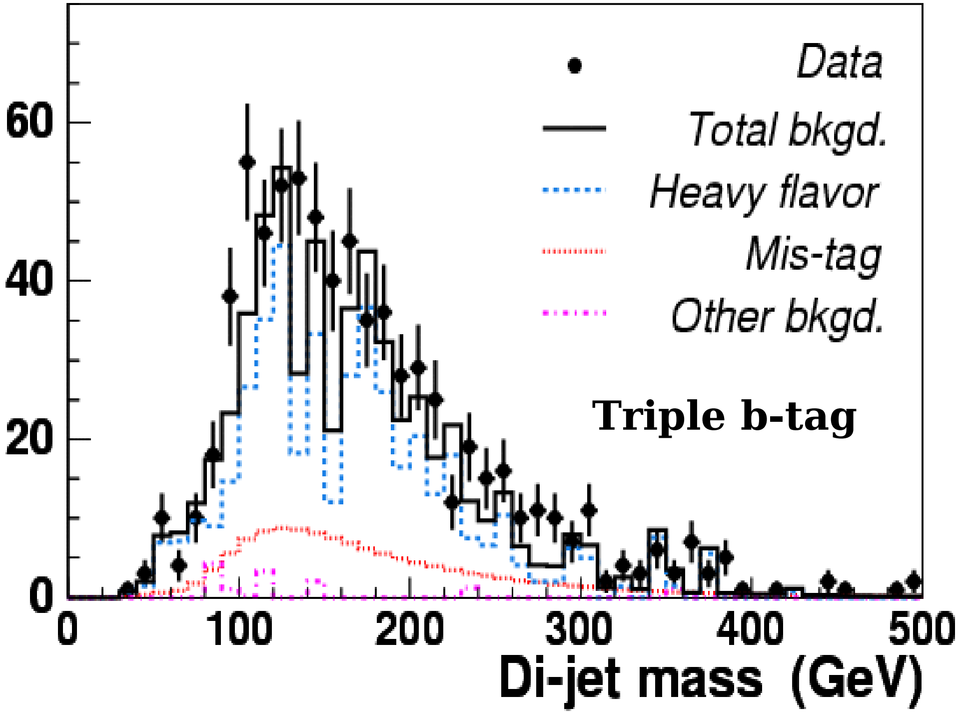


Good agreement in shape.
Normalization of “heavy-flavor” MC agrees to within 10%.

MC Cross-checks



g->bb !

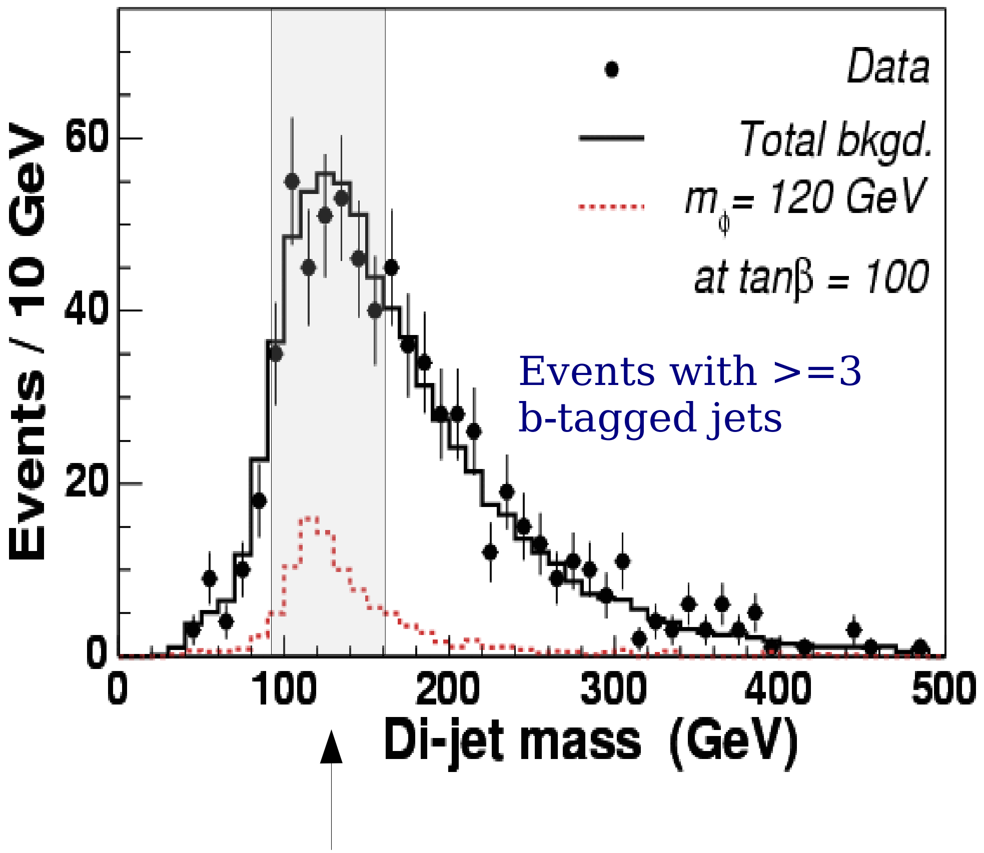


Monte Carlo accurately predicts many other distributions in the double b-tagged data as well.

Monte Carlo predicts a background shape and normalization in good agreement with the triple b-tagged data.

Systematic Uncertainties

- Signal efficiency uncertainties (total = 21%):
  - b-tagging efficiency (15%)
  - Trigger efficiency (9%)
  - Jet energy scale (8%)
  - Integrated luminosity (6.5%)
  - NLO kinematics (5%)
  - Jet reco/ID efficiency (4%)
  - Jet energy resolution (1%)
- Background shape / normalization uncertainties (total = 3%):
  - TRF shape (2%)
  - Statistics outside signal window (1%)
  - Width of signal window (1%)

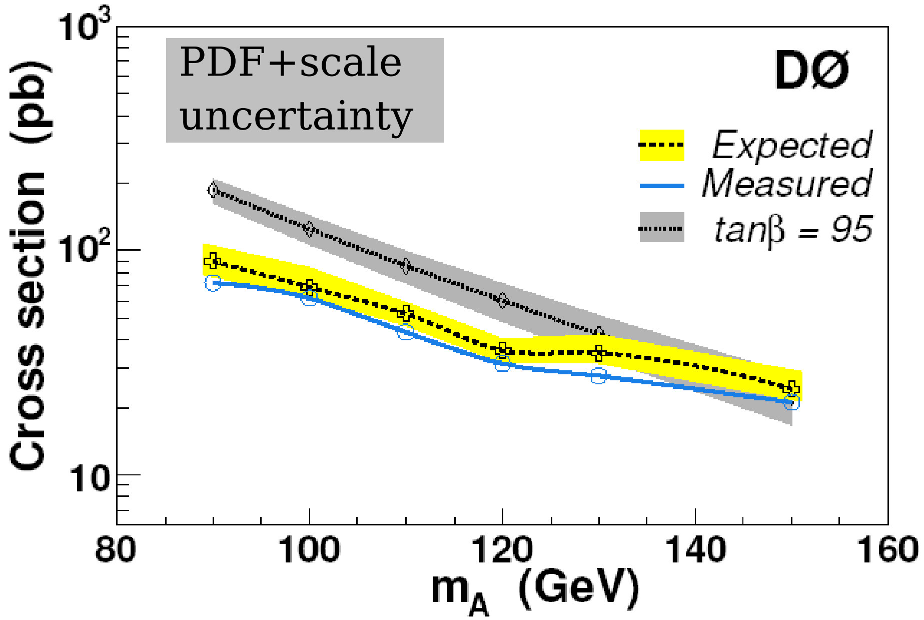


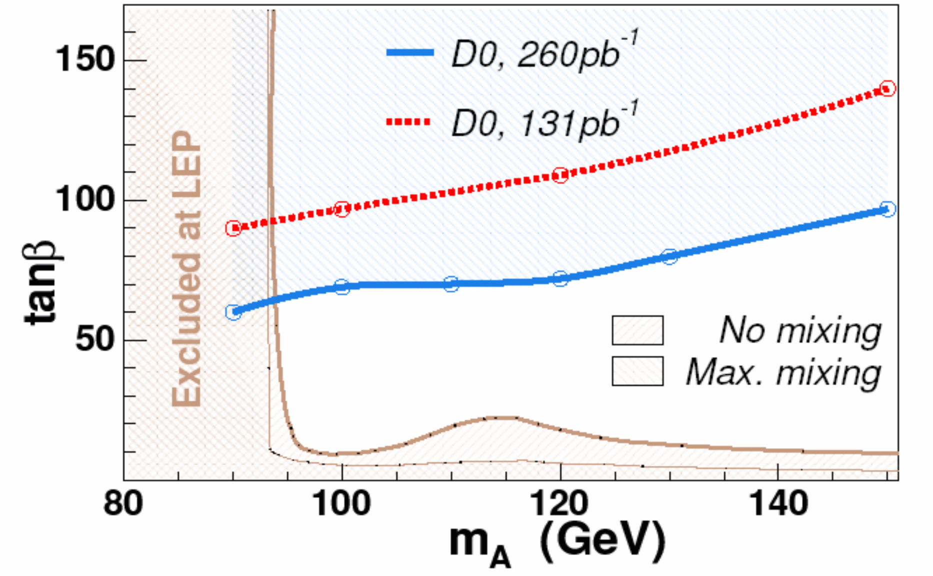
Events with >=3 b-tagged jets

“Signal window”

MSSM Tree-level Limits

- Limits set using CL_S method (TLimit)
  - For each m_A, set a cross-section limit at 95% C.L.
  - Interpret in MSSM at tree-level, assuming $\tan^2\beta$ cross-section enhancement
  - Cross-section uncertainties are calculated by varying the factorization / renormalization scales, and the PDF sets (CTEQ6)
  - Uncertainty on expected limit comes from varying the background by ±1 sigma
  - We got a little lucky
- The current result excludes much more parameter space than our result from last year





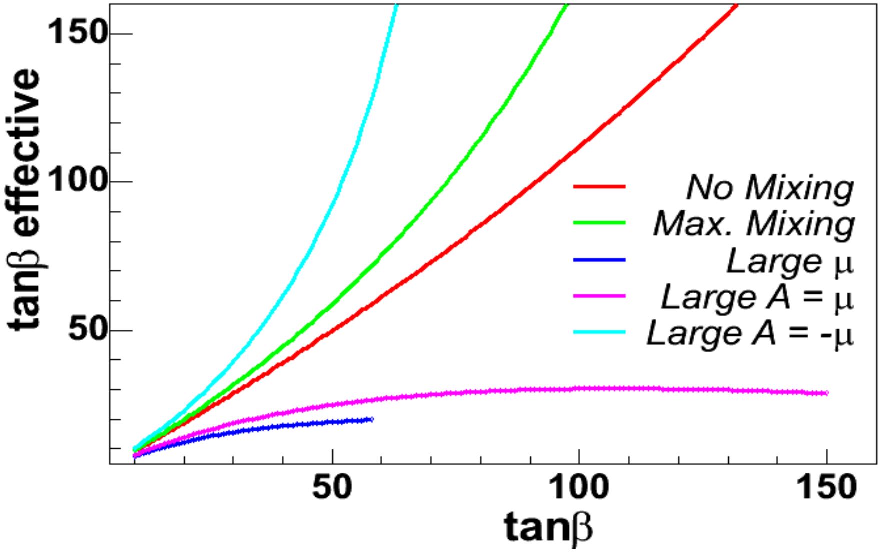
MSSM Beyond Tree-Level

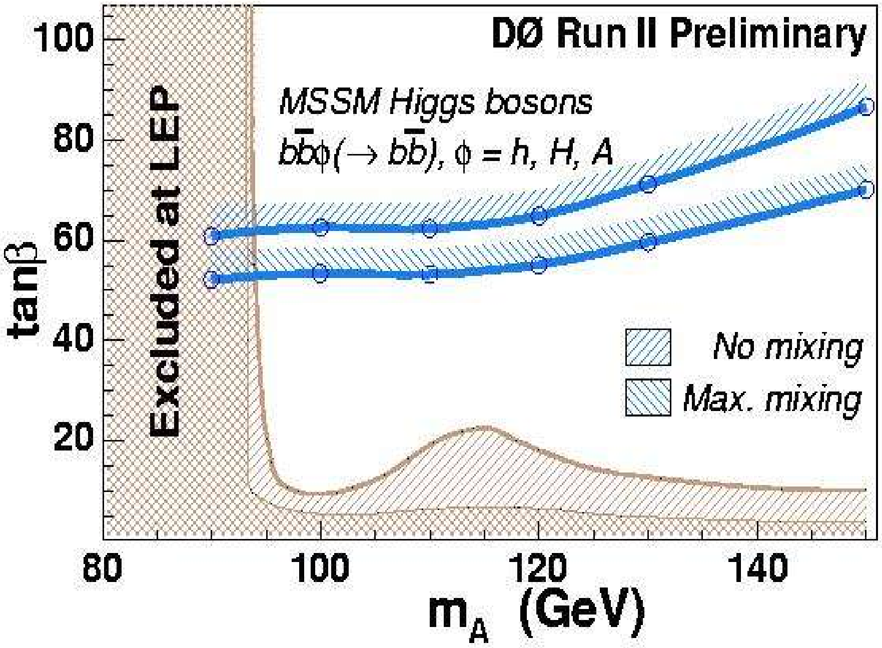
- Supersymmetric loop corrections are calculated using “CPSuperH”
    - Production doesn't scale like $\tan^2\beta$… can be faster or slower depending on the supersymmetric model and its parameters
- We set limits on $\tan\beta$ vs. m_A in two of the “benchmark scenarios”:
    - no mixing and maximal mixing in the stop-quark sector

$$\sigma \times \boldsymbol{BR}_{SUSY} = 2\times\sigma_{SM} \times \frac{\tan\beta^{2}}{(1+\Delta_b)^2} \times \frac{9}{[9+(1+\Delta_b)^2]}$$

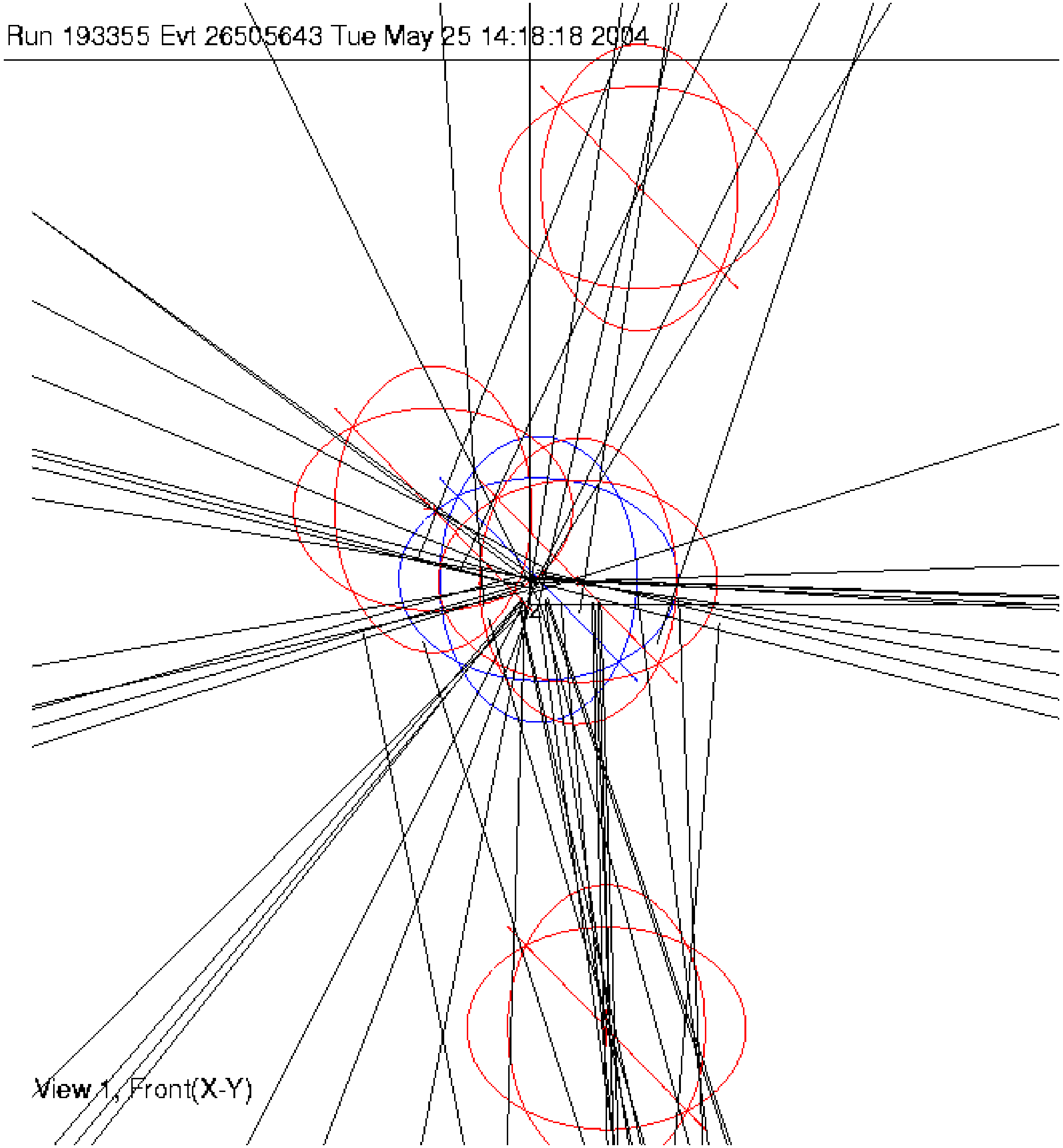
Δ_b *is a function of various SM/SUSY parameters:*

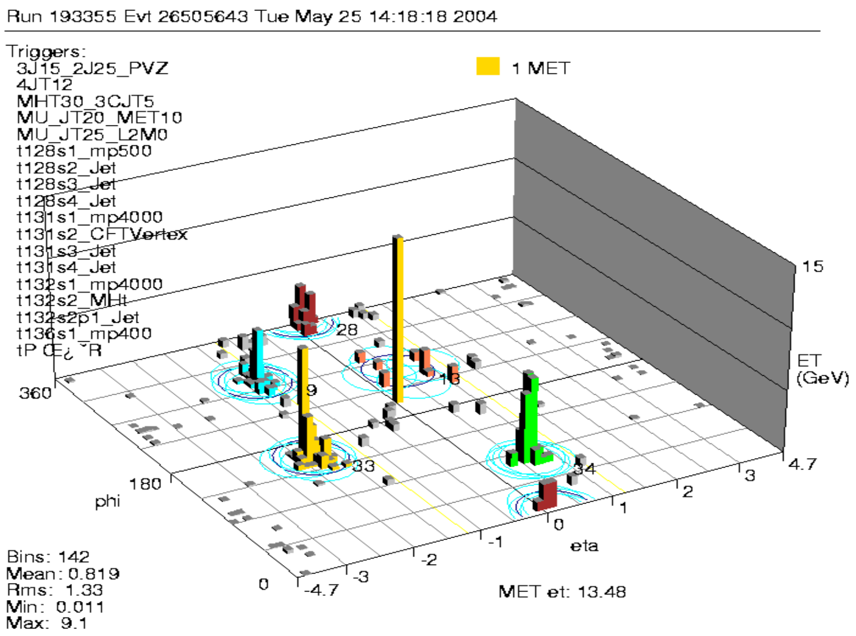
$X_t = A_t - \mu\cot\beta$*,* μ*,* M_g*,* M_q*, etc.*

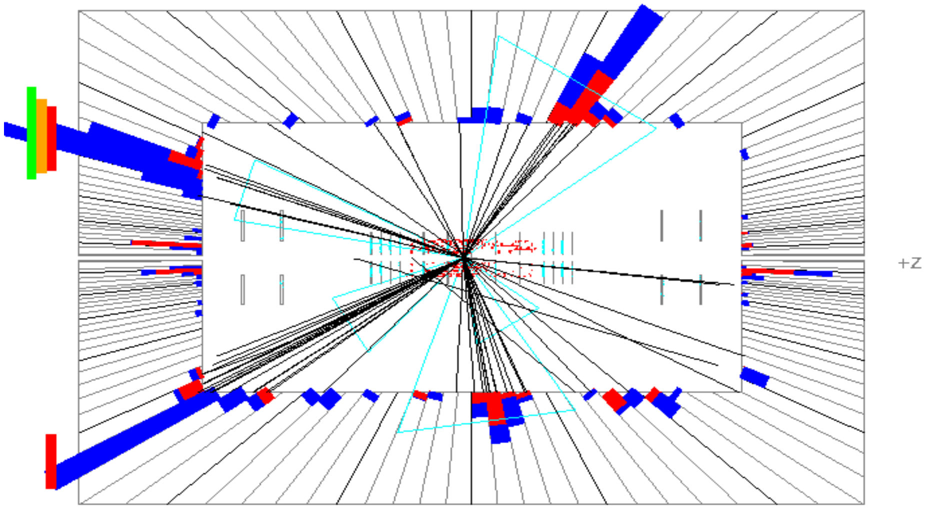




A Quadruple b-tagged Event

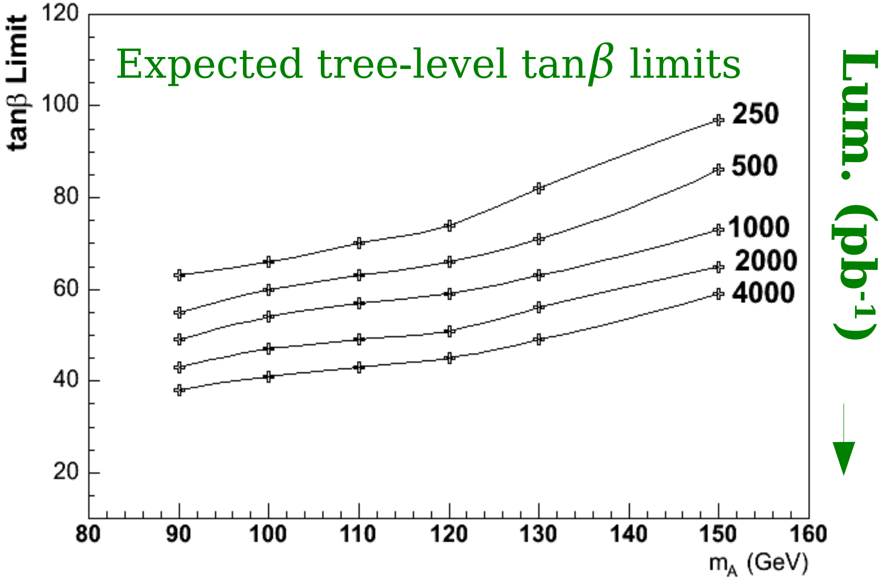




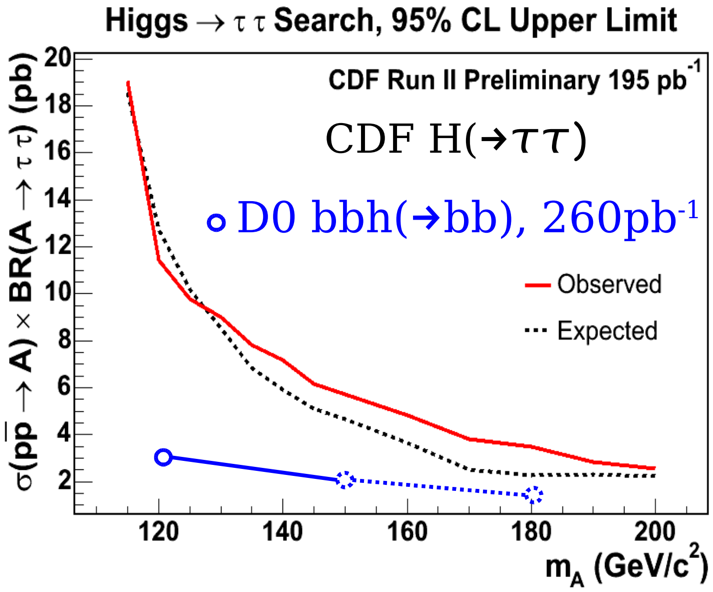


Future of the Analysis

- With no analysis improvements, expect to exclude down to tanβ=40 with 4fb^{-1}
- But we also look forward to a much improved analysis with next year's data set (1fb^{-1})
    - Inclusion of more triggers
    - Better calorimeter and jet energy calibration / resolution
    - Improved b-tagging
    - Use quadruple b-tagged events
    - Neural network?
- Combine with bbH(→$\tau\tau$) ?
- Also keeping an eye on our competitors...
    - Combine with H(→$\tau\tau$), H(→bb) at higher m_H?


Expected tree-level tanβ limits

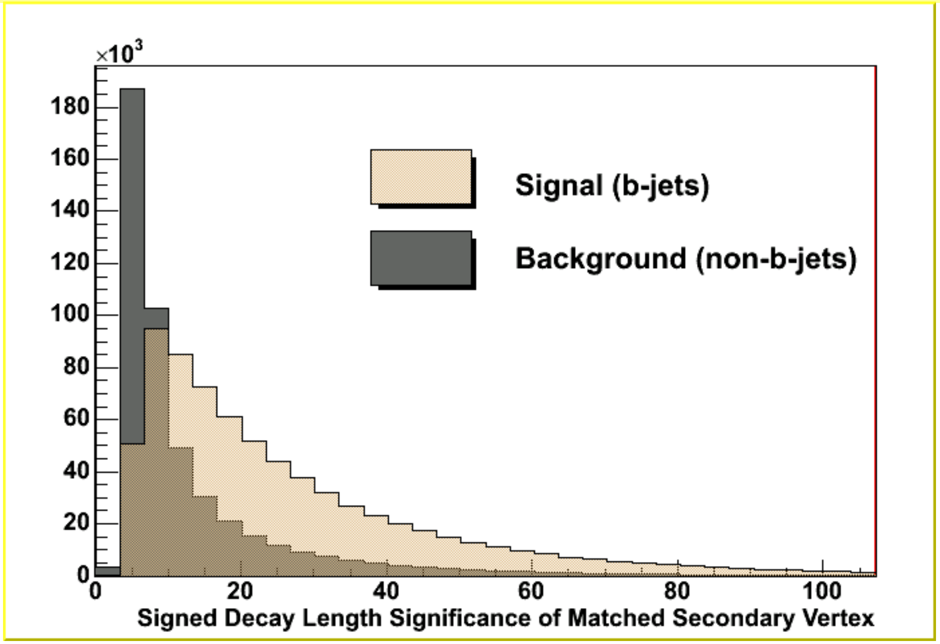
Lum. (pb^{-1})

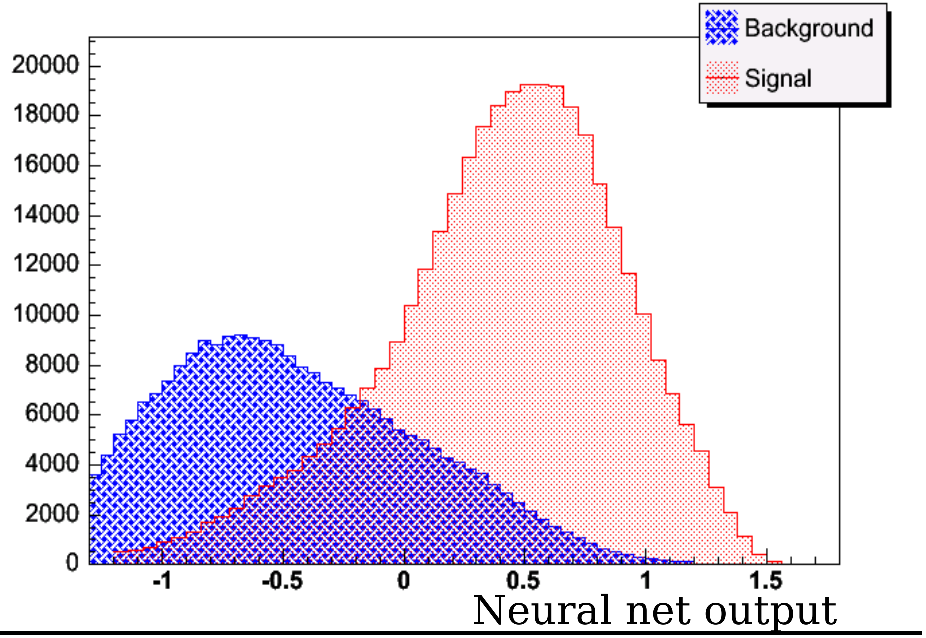

CDF H(→$\tau\tau$)

D0 bbh(→bb), 260pb^{-1}

Improving the b-tagging

- This analysis is very sensitive to the b-tagging efficiency and the background tagging rate
- One obvious way to improve the b-tagging S/B is to use more information about the secondary vertex(es) in each jet
  - number of tracks on vertex
  - the χ^2
  - the “mass”
  - decay length significance
  - number of secondary vertices
  - angles between secondary vertices
  - … combine into neural network?
- Initial studies show large gains in performance are possible





Conclusions

- A search has been completed for neutral Higgs bosons in the MSSM using 260 pb^{-1} of DØ Run II data
  - Good agreement between heavy-flavor multijet simulations and data
  - Methods were developed for estimating the triple b-tagged backgrounds
  - Solid theoretical understanding of the production process in the MSSM
- This is a very challenging analysis!
  - Triggering: optimization and calculation of efficiency
  - Jets: efficiency, energy scale, energy resolution
  - Tracking and primary vertexing
  - Secondary vertexing and b-jet identification
  - Monte Carlo generators and simulations
- We hope to make a discovery, or place impressive limits on $\tan\beta$ vs. m_A in the future using larger data sets and more advanced analysis techniques