MSSM Higgs Searches at CDF

Thomas Wright
University of Michigan

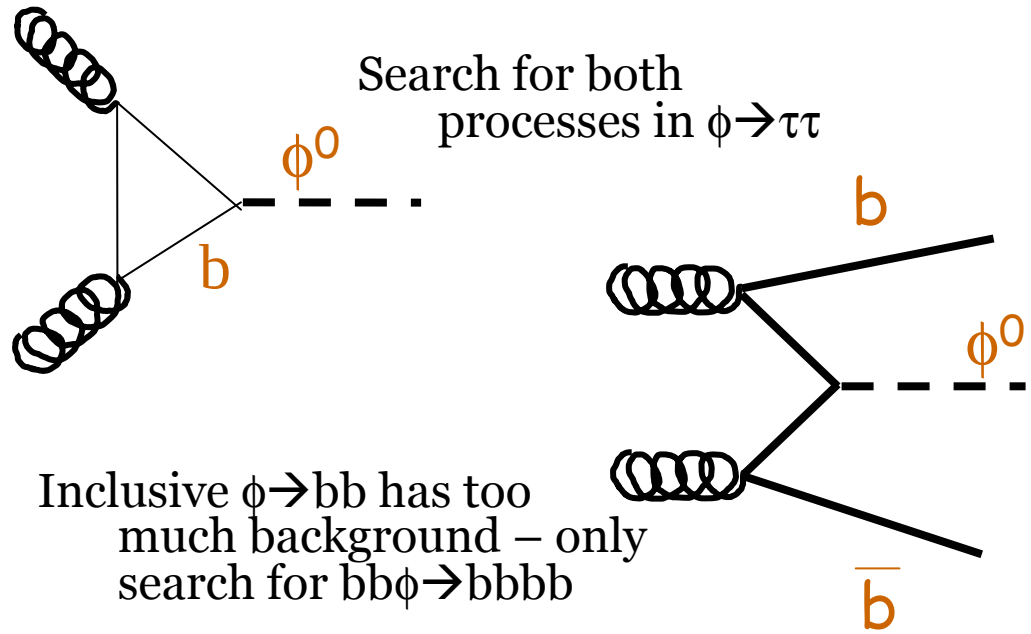
Fermilab Joint Experimental-Theoretical Seminar
September 28, 2007

Outline

- Data samples and the CDF detector
- Overview of the Higgs sector of the MSSM
- Search for associated Higgs production in the 3b channel
  - First CDF results shown at Lepton-Photon
  - Background templates and fitting
  - MSSM interpretation, scenario dependence, and width effects
- Search for inclusive Higgs production in the $\tau\tau$ channel
  - Featured in the world's finest newspapers and blogs
  - Updated results with more data as of today!
- Conclusion and outlook

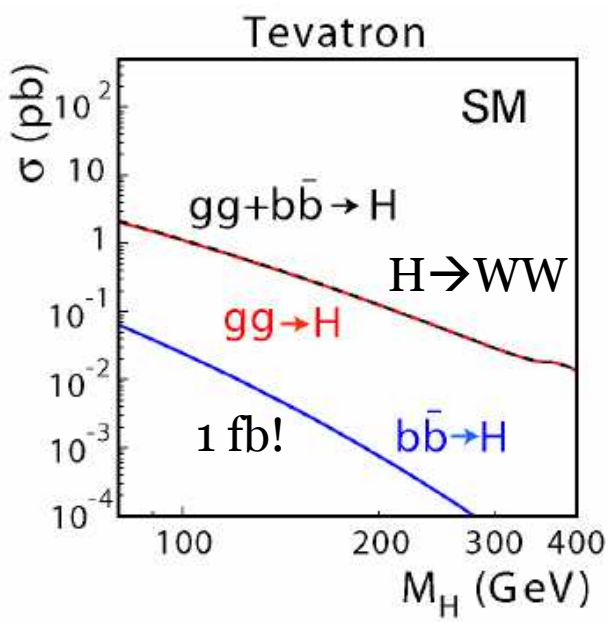
Higgs in the MSSM

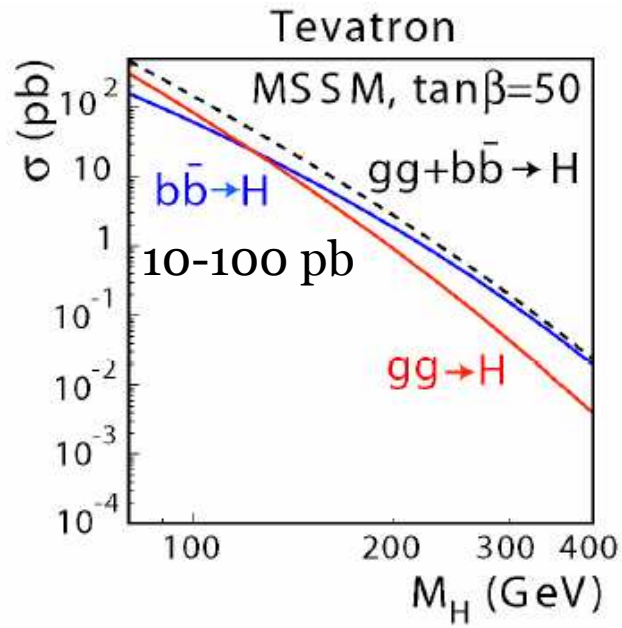
- Five scalars: h, H, A, H^+, H^-
  - Separate couplings for up-type and down-type fermions
- Properties of the Higgs sector largely determined by:
  - m_A : mass of pseudoscalar
  - $\tan\beta$: ratio of down-type to up-type couplings
- More true for $\phi \rightarrow \tau\tau$ than $\phi \rightarrow bb$ because of loop effects
- If m_A is large, then h looks like a standard model Higgs
- If m_A is small and $\tan\beta$ is large
  - Production via b quarks can be greatly enhanced (by factor $\sim\tan^2\beta$)
  - Decays to bb (~90%) and $\tau\tau$ (~10%) dominate
  - One of the scalars h or H is nearly degenerate with A (effectively doubles cross section)



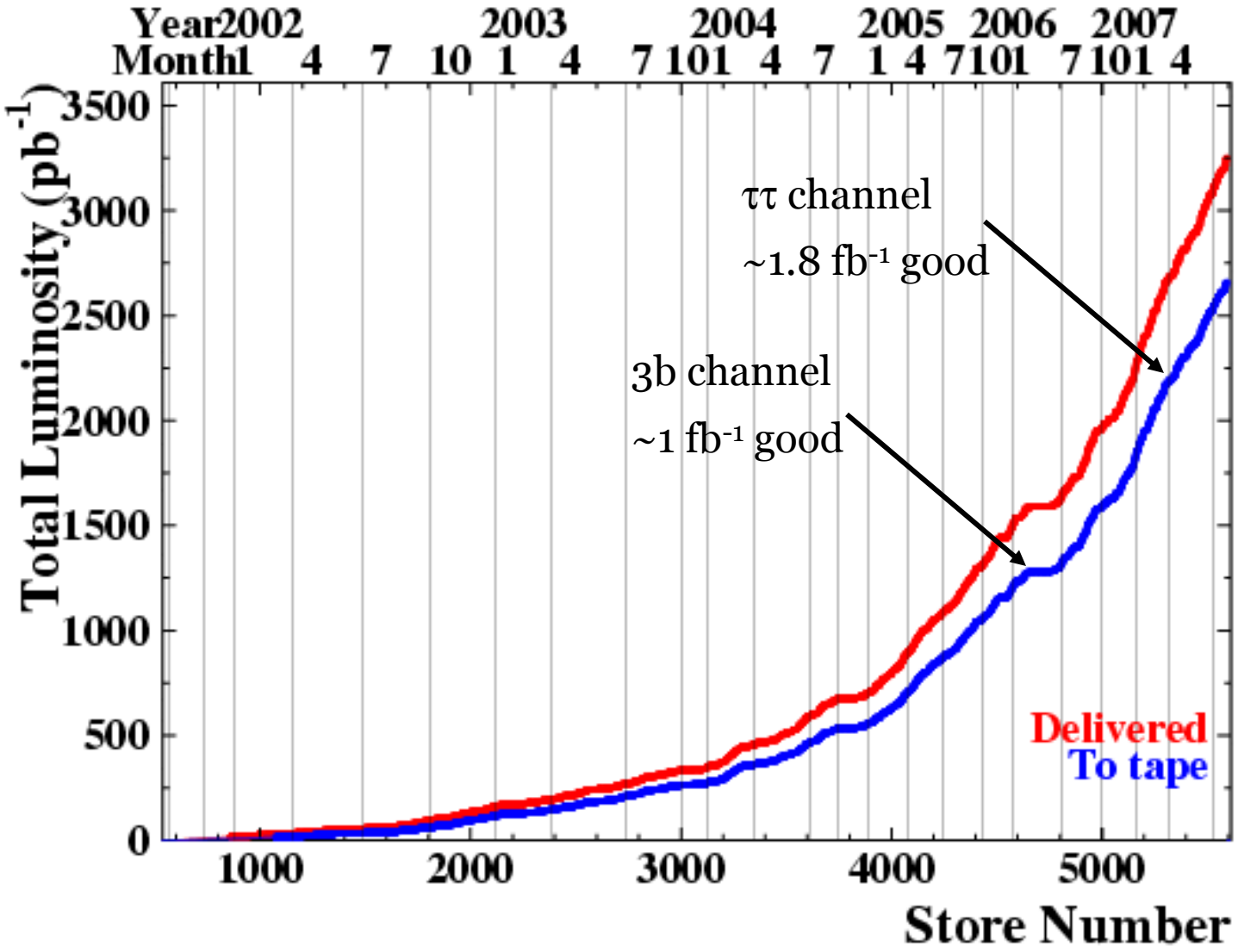
Search for both processes in $\phi \rightarrow \tau\tau$

Inclusive $\phi \rightarrow bb$ has too much background – only search for $bb\phi \rightarrow bbbb$

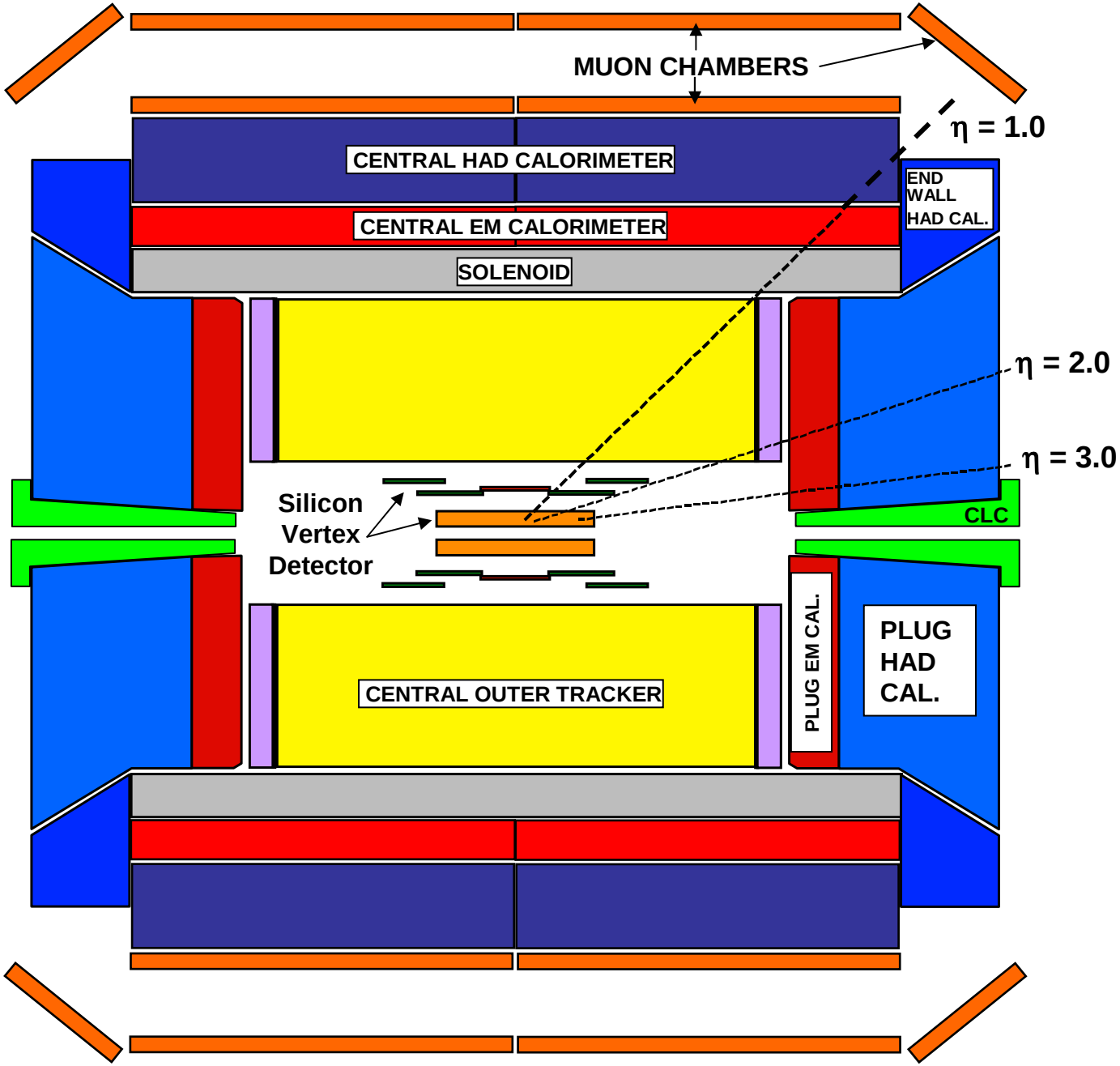




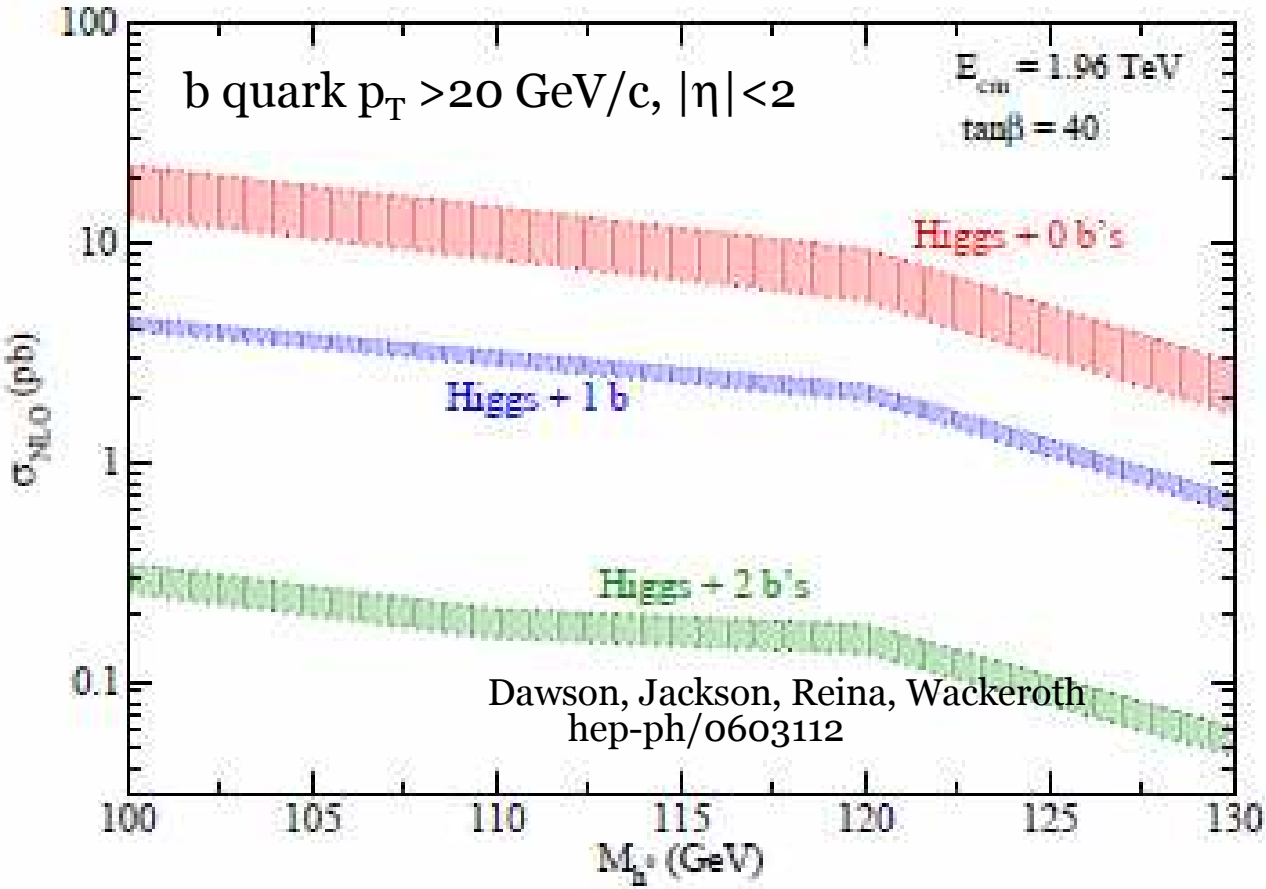
Data Samples

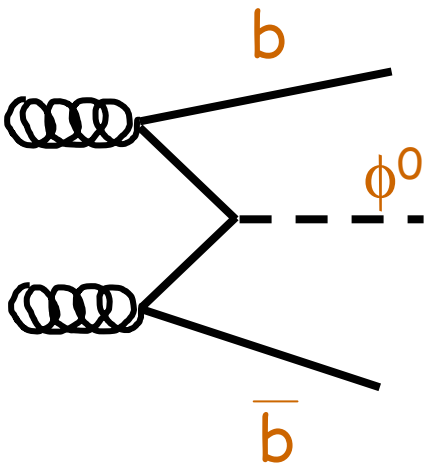


The CDF II Detector



The 3b Channel



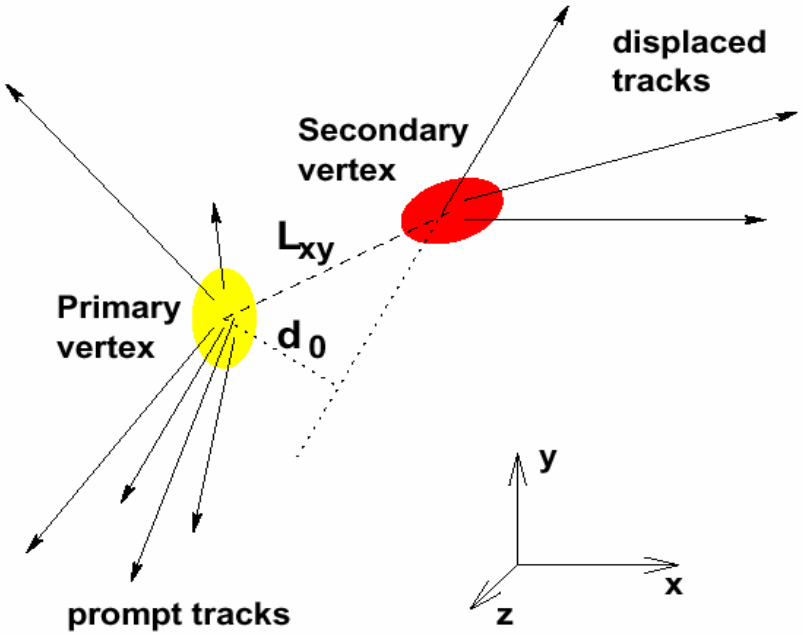


- Search for the bbϕ→bbbb process
- Less cross section when requiring both b's to be high-p_T

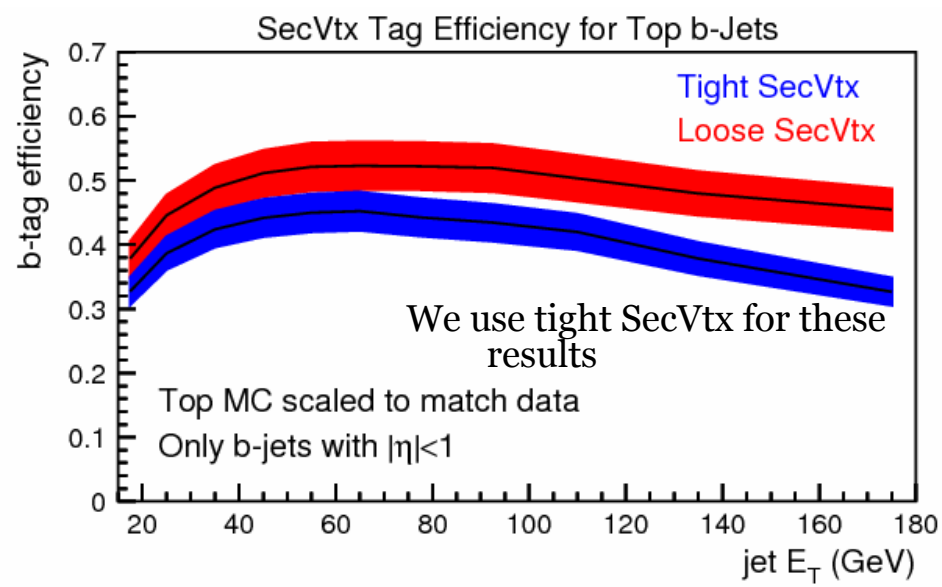
- Look for the Higgs + 1b case
  - Three jets b-tagged with secondary vertices
  - Must be three leading jets, all E_T > 20 GeV (cone radius 0.7), |η|<2

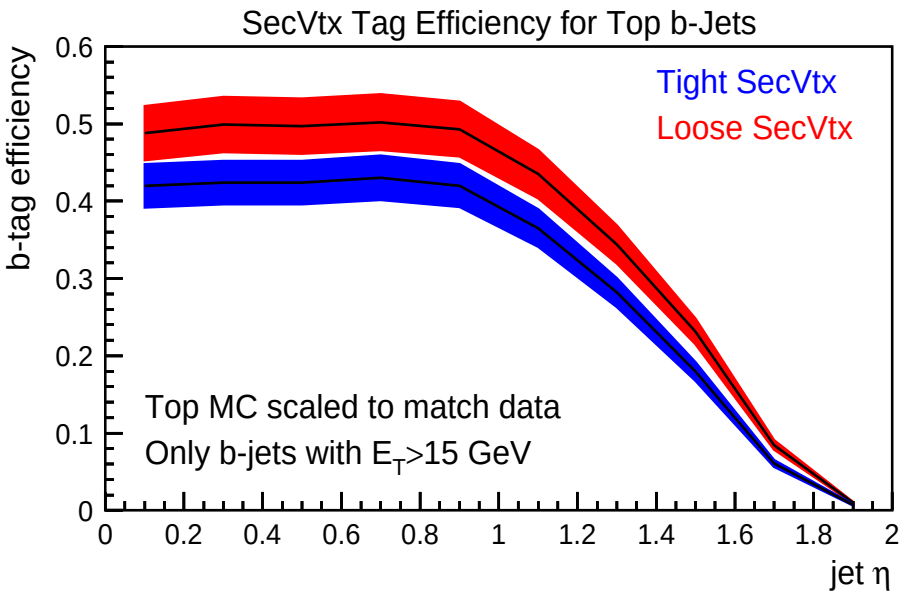
B-Jet Identification (b-tagging)

- B-hadrons are long-lived – search for displaced vertices
- Construct event-by-event primary within beamspot (10-32 μm)
- Fit displaced tracks and cut on L_{xy} significance (σ ~ 200 μm)
- Calibrate performance from data (low-p_T lepton samples)



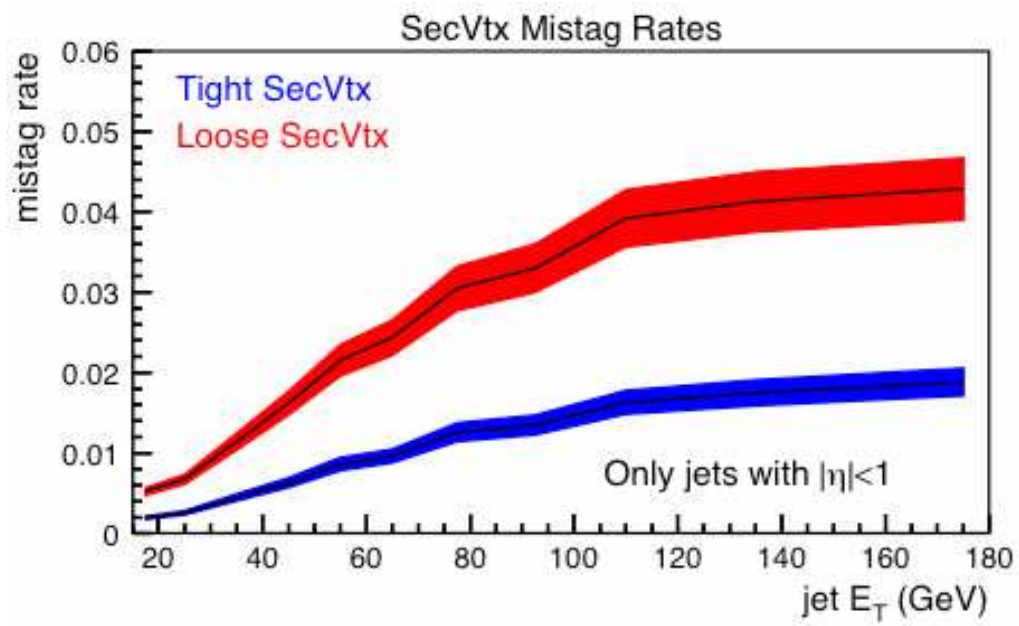
- Charm hadrons have similar tag behavior but lower efficiency



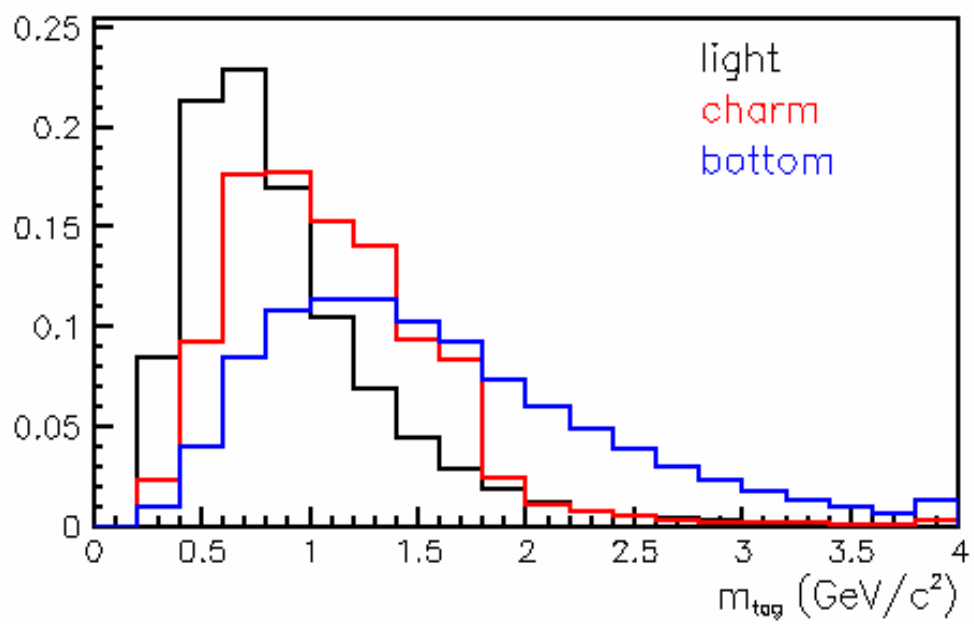


Fake B-Tags (mistags)

- Fake tags of light-flavor jets can result from poorly-reconstructed tracks or interactions with the detector material
- More tracks in a jet above 1 GeV/c means more chances for something to go wrong – fake rate rises dramatically with jet ET



- Can use "tag mass" to deduce the flavor composition of a sample of tagged jets
  - Mass of the tracks forming the secondary vertex
  - B-hadrons are heavy → will have higher m_{tag} spectrum than charm or light jet fakes



3b Channel Roadmap

- Trigger is three jets, two matched to displaced tracks
  - Events are fairly biased even after offline selection
  - Data-driven backgrounds are the natural way to handle this

- Background is QCD multijet production of all possible jet flavors
  - We consider b-jets, c-jets, and "q"-jets (q = light quark or gluon)

- Derive templates for each flavor combination from the data
  - Use Pythia to check/correct for bias

- Look for an excess in the mass of the two leading jets (m_{12})
- Use tag mass (m_{tag}) information to understand flavor composition

- Perform a two-dimensional fit to the data using these templates
  - Tag mass information determines background composition
  - Look for Higgs in the m_{12} distribution

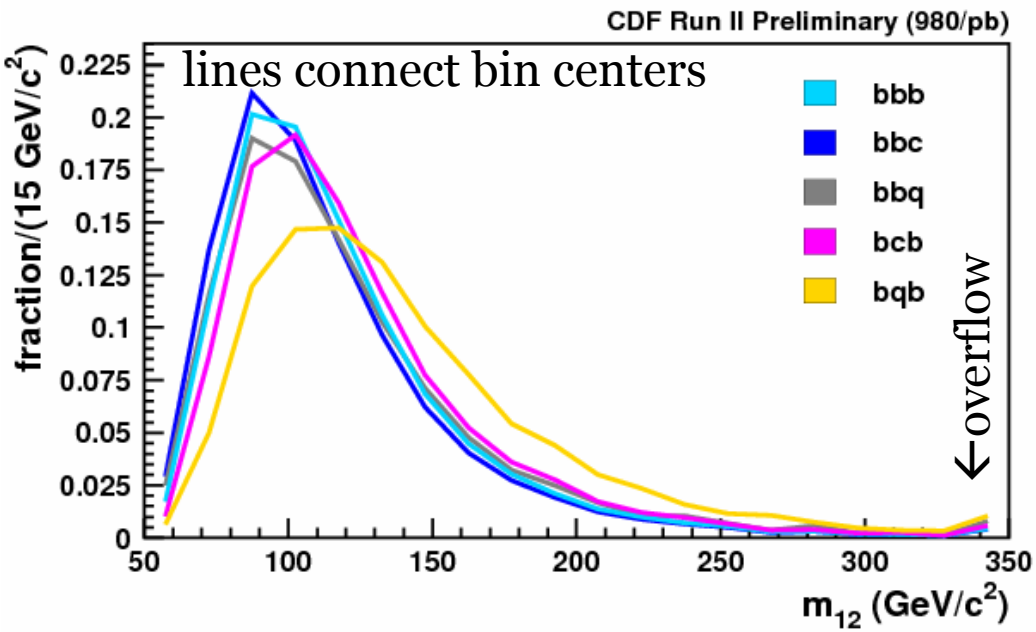
Background Model

- Essentially ~100% of background is QCD multijet production
- Most of it is in five components – two b's plus b/c/q

- Fractions are from fully simulated Pythia samples
  - Do not want to rely on these cross sections
  - Will fit for them in the data

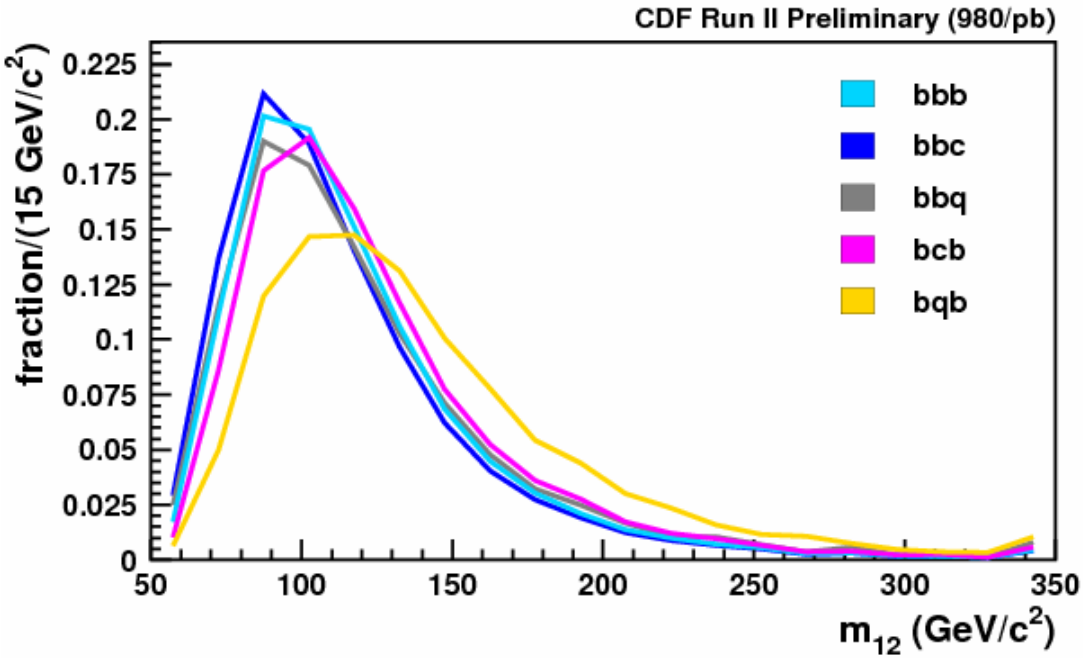
- Shape differences are due to both physics and detector effects
  - Biases from b-tagging
  - Kinematics of heavy quark production processes

jet flavor (12)3	fraction
bbb	0.45
bbc	0.10
bbq	0.10
bcb	0.16
bqb	0.13
other	0.06



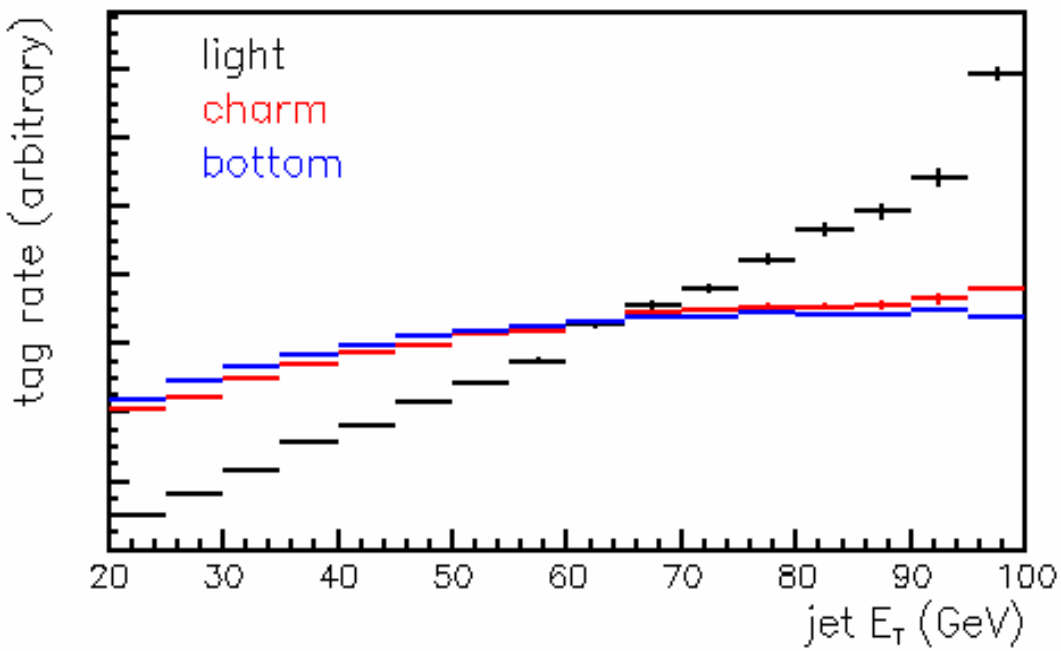
Background Strategy – m_{12}

- Use three-jet events with two leading jets tagged to construct bbb, bbc, bbq, and bcb templates
  - Verified that m_{12} in this sample looks like bbX using Pythia samples
  - Also works for bcb due to similar tag behavior vs jet E_T
  - Third jet flavor doesn't affect m_{12} very much



- For bqb, start with events with a tagged third jet and either of the two leading jets tagged
  - Weight event by light-flavor tag rate of remaining jet
  - Fake tag biases m_{12} considerably

- This is one axis of our fit templates

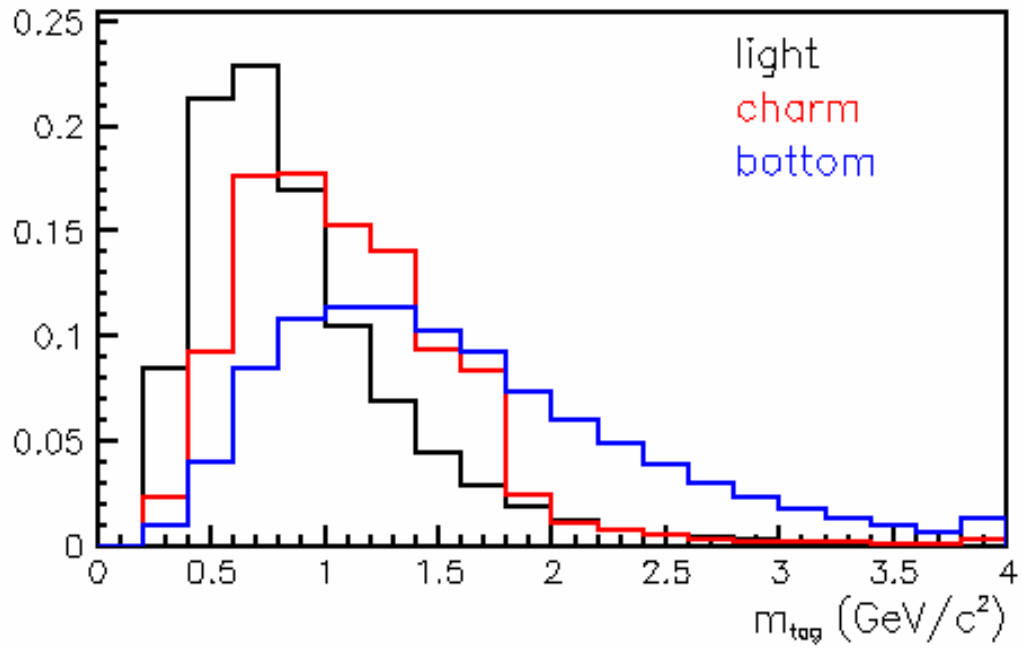


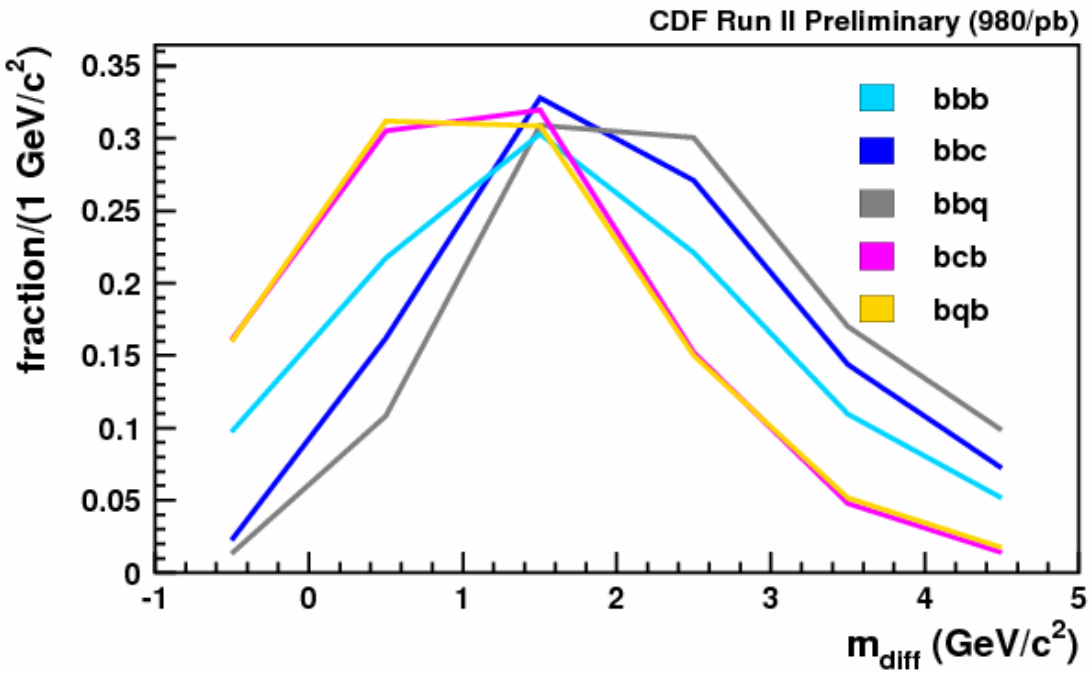
Background Strategy - m_{tag}

- Several components with small but important differences in m_{12} distribution
- Need something more sensitive to flavor composition as an additional fit variable
- Use tag mass information
- Simple combination:

  $m_{diff} = m_{tag}(jet1) + m_{tag}(jet2) - m_{tag}(jet3)$

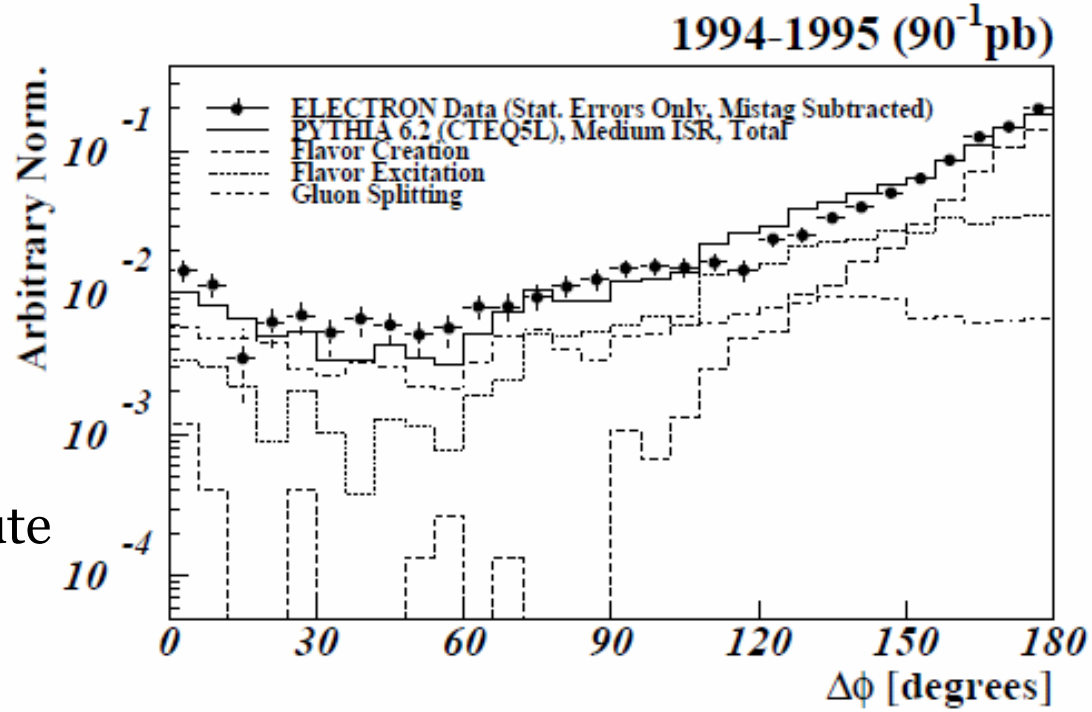
  separates the background components well
  - Firms up the overall background m_{12} shape
- This is the second axis of our background fit templates

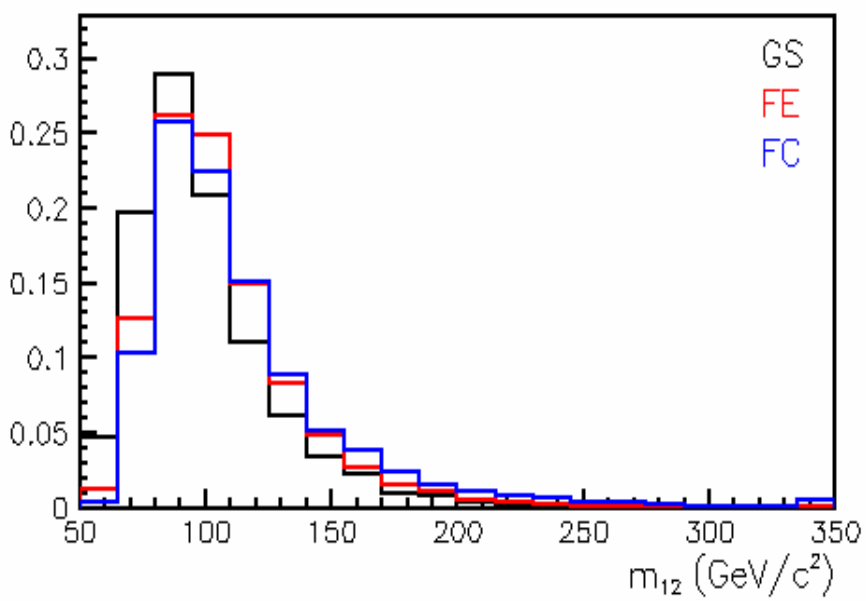


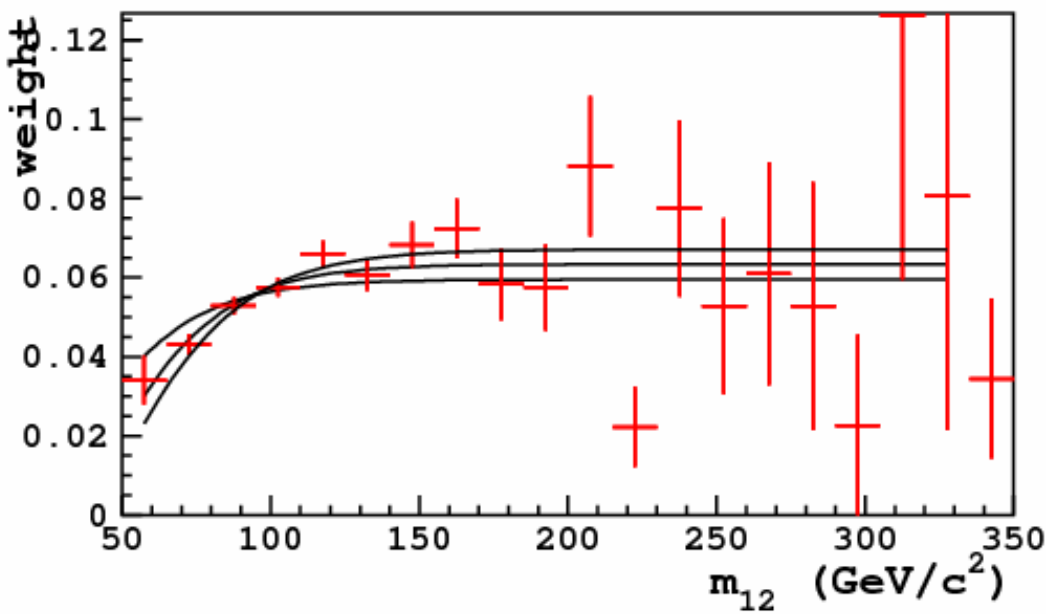


Heavy Flavor Production Kinematics

- Heavy flavor production model affects bbj shapes
  - bbX might not be a perfect model for bbb and bcb
  - Derive correction based on Pythia generator studies
- Run 1 studies indicate gluon splitting production may contribute from 1-3 times default Pythia rate
- Derive correction and systematic errors for bbb and bcb templates

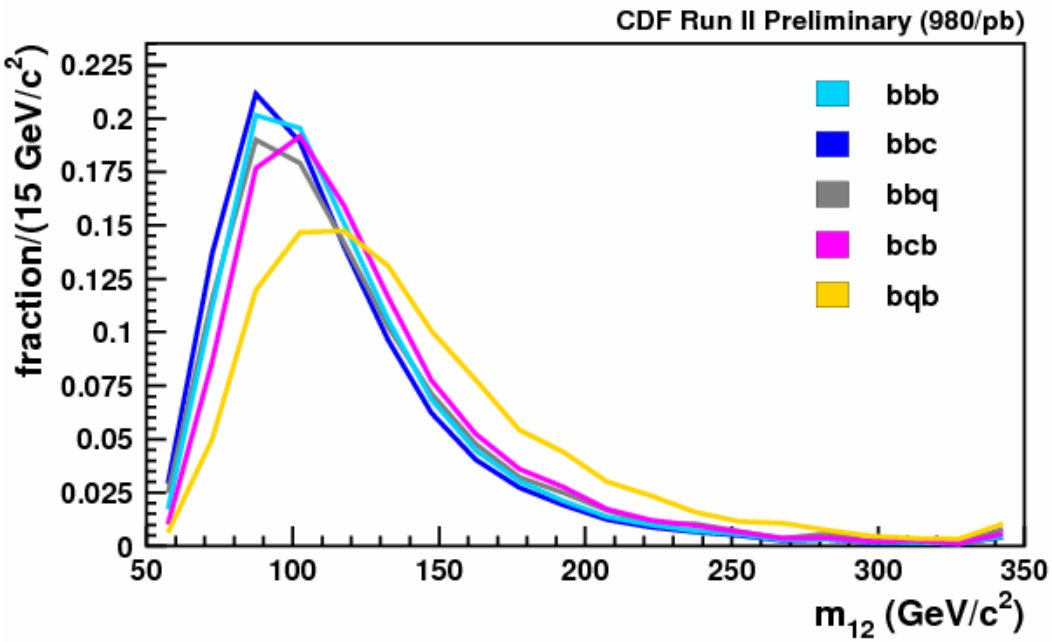




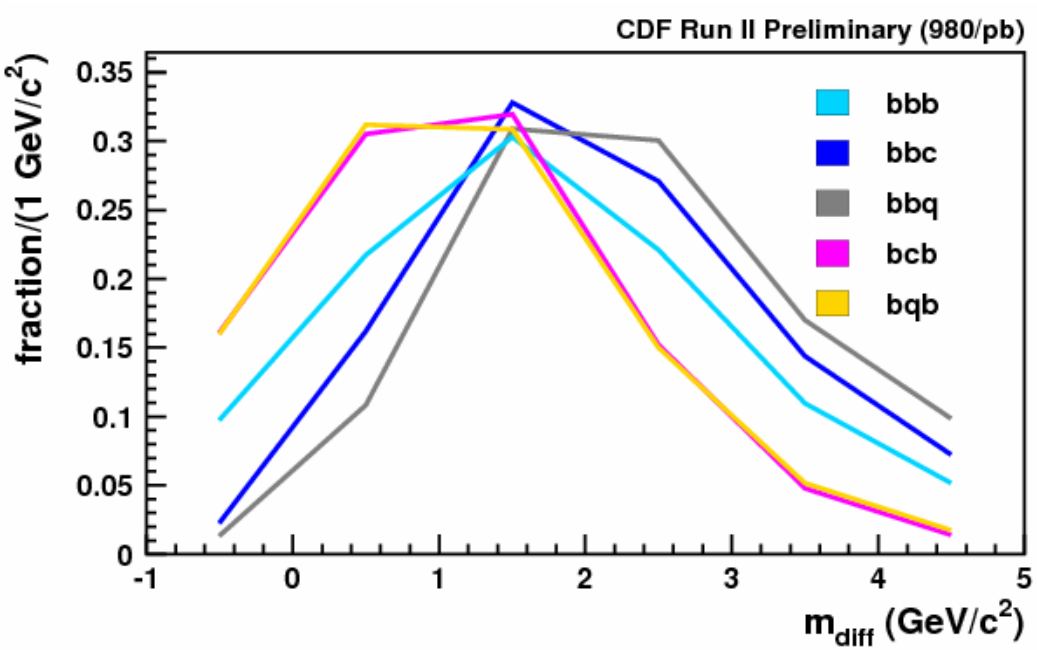


Final Background Shapes from Data

- Starting from double-tags in data, apply all weights and corrections to get final fit templates
  - m_{12} : dijet mass, for Higgs discrimination
  - m_{diff} : tag mass combination, fixes background composition



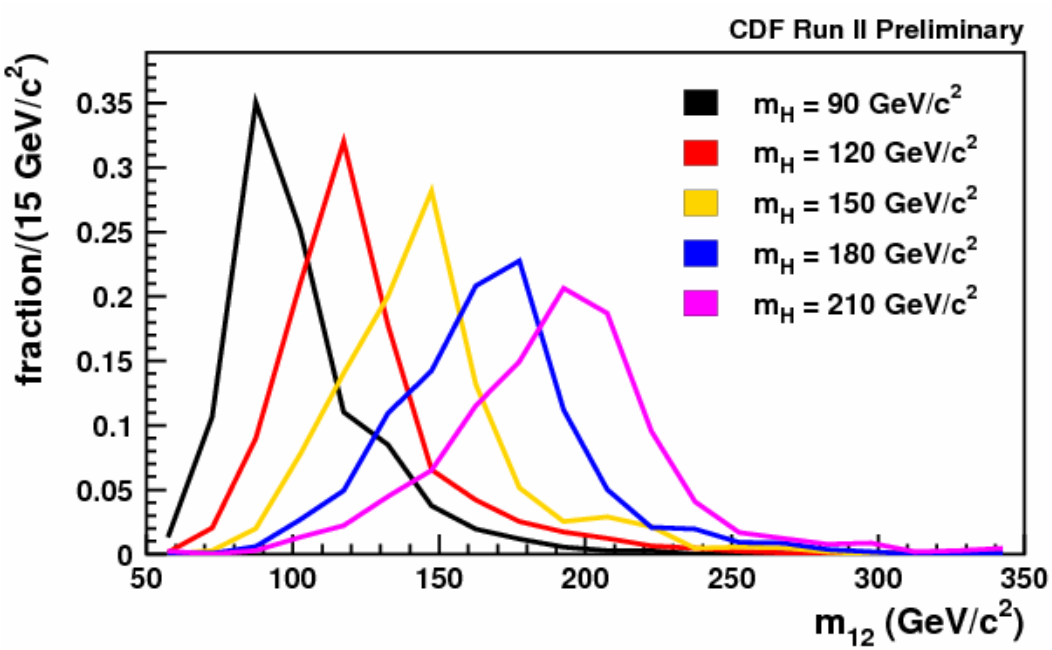
- Templates are actually 2D histograms in both m_{12} and m_{diff}
  - Fit itself is also 2D
  - Only show projections for clarity

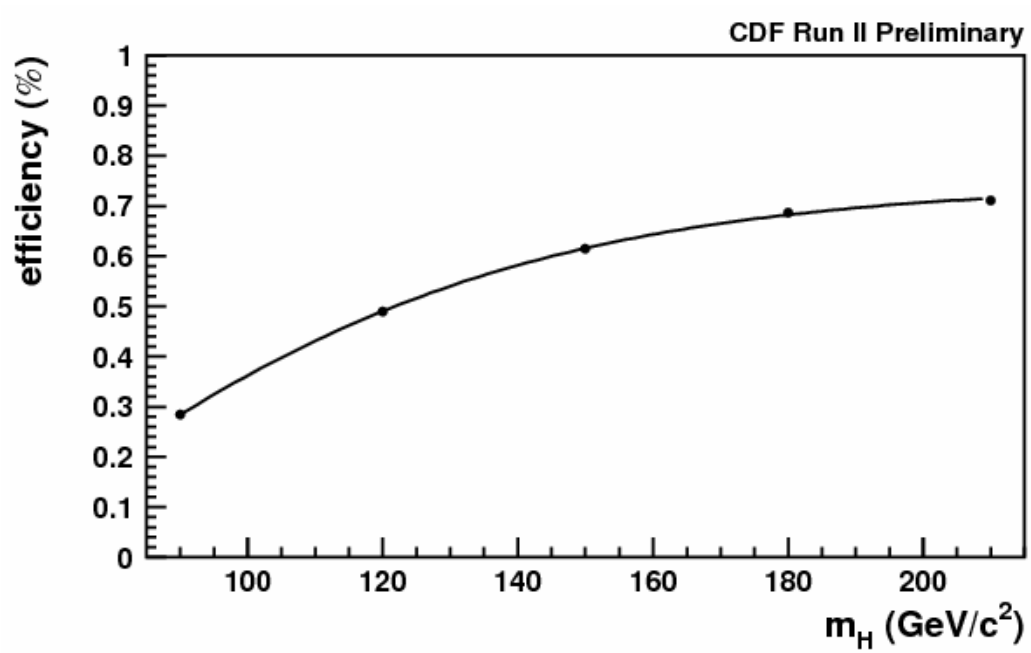


- bbq and bbc are too similar to separate in the fit
  - Use average of the two (bbx)
  - Vary between extremes for a systematic error

Signal Samples

- Generated using Pythia, MSUB=32 (bg→bH)
- SM Higgs, no width
- Cross section we are measuring is the one reported by FeynHiggs
  - From MCFM calculation
  - b quark p_T>15 GeV/c, |η|<2.5
- Efficiency is relative to that denominator





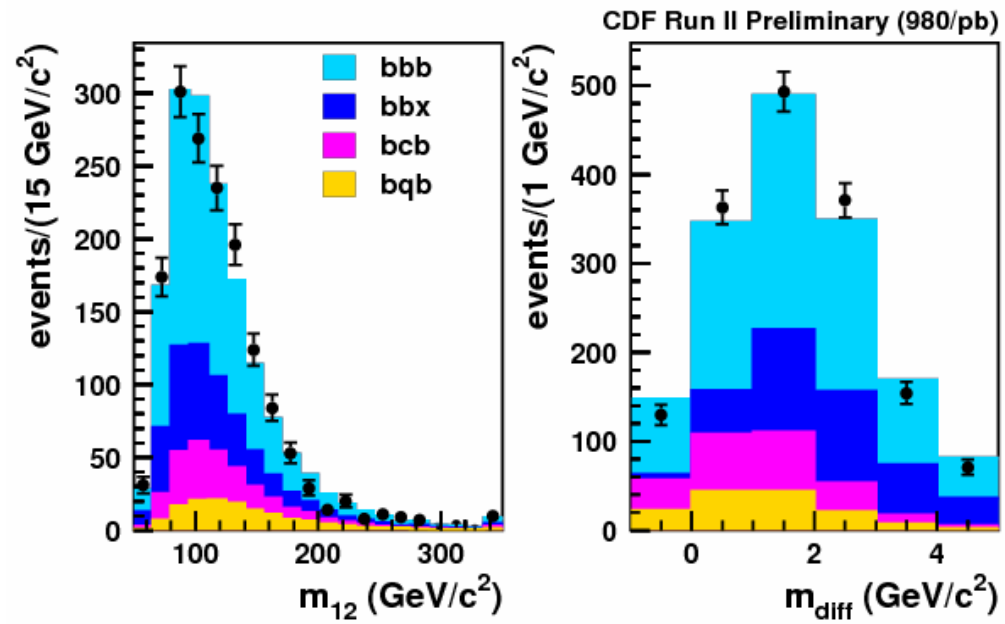
Background-Only Fit

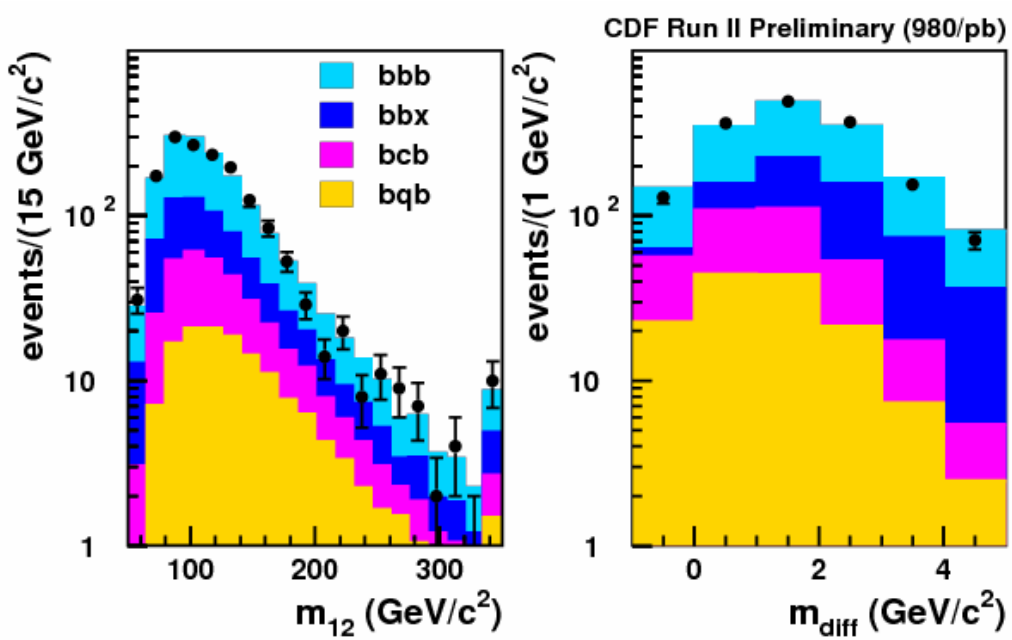
- Simple binned max-likelihood fit
- No systematics

- Total of 1582 triple-tagged events in 1 fb^{-1} of data

- Fit returns
  - `bbb : 870 ± 270 events`
  - `bbq : 360 ± 140 events`
  - `bcb : 210 ± 170 events`
  - `bqb : 140 ±  80 events`
- Pretty close to the Pythia estimates

- Backgrounds describe the data pretty well
  - Some high bins around 140
  - Some low bins at 200





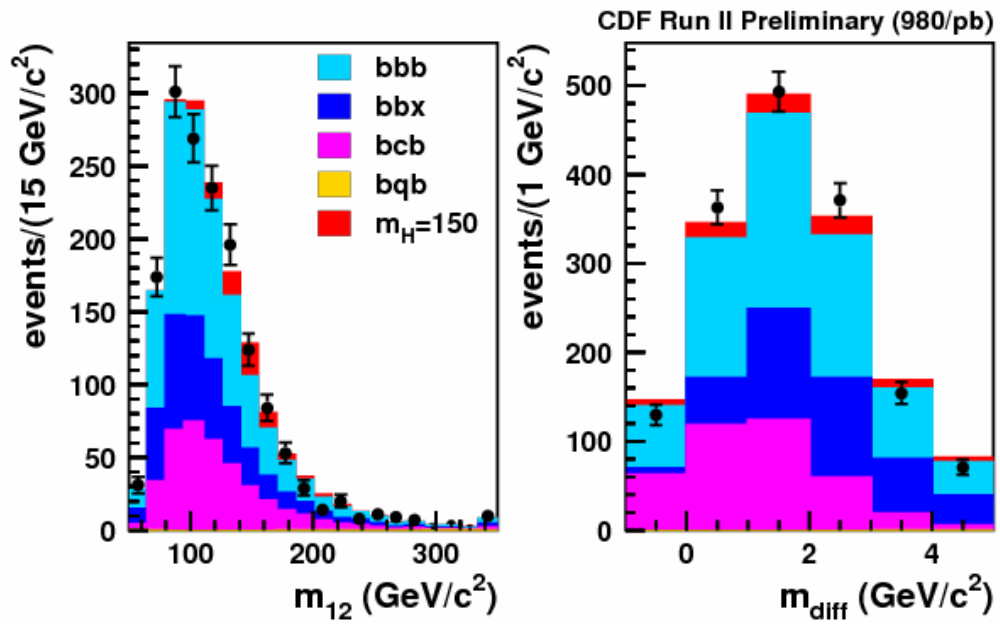
Fit Including Signal

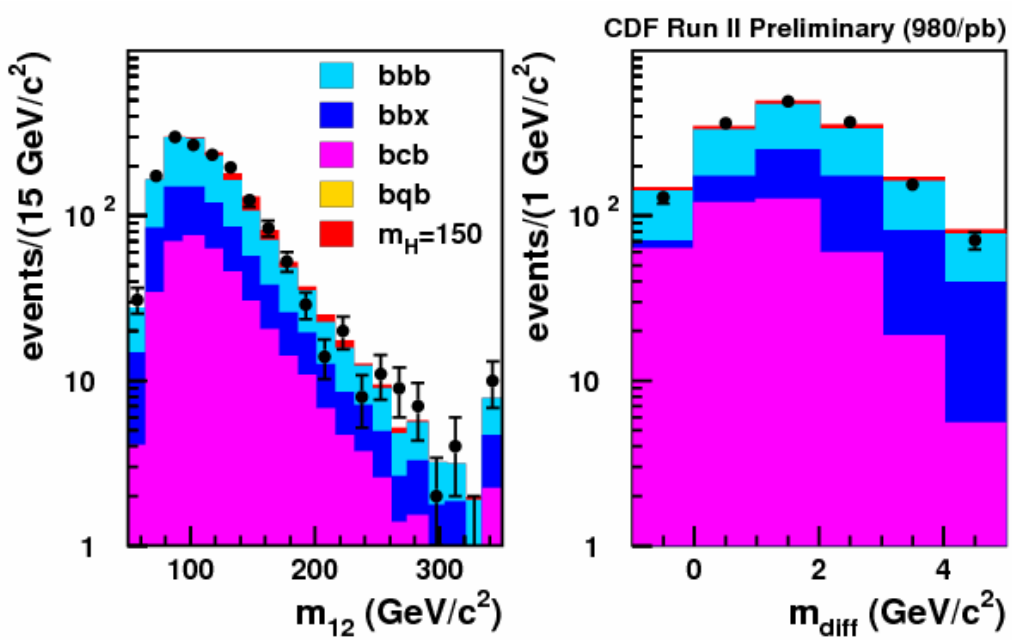
- Repeat simple fit, using backgrounds plus a Higgs signal component (m_H = 150 GeV/c²)

- Fit returns
  - bbb : 740 ± 270 events
  - bbq : 380 ± 150 events
  - bcb : 380 ± 160 events
  - bqb : 0 ± 10 events
  - Higgs: 80 ± 40 events

- bqb events mostly move into bcb, not into Higgs

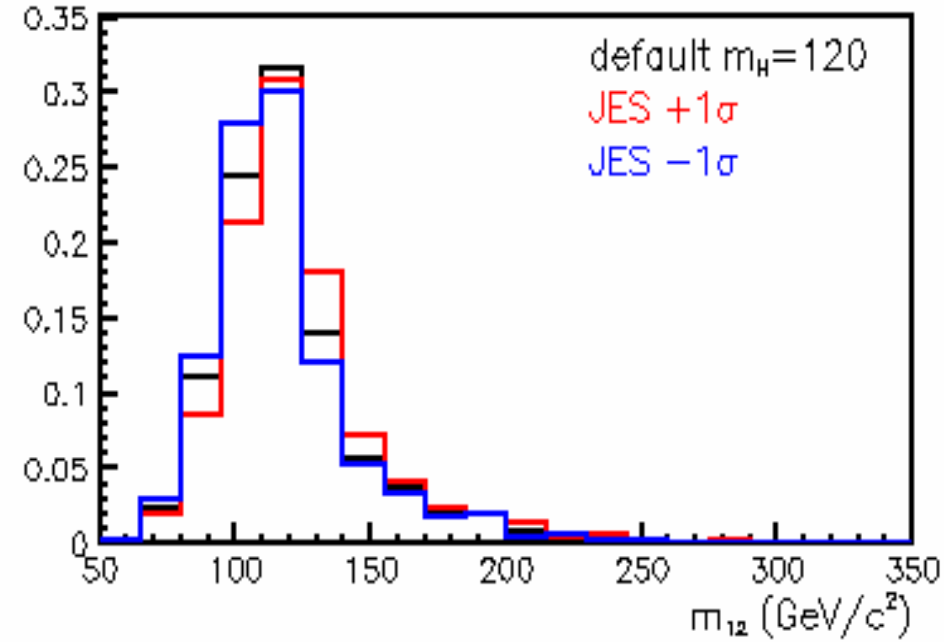
- Errors are statistical only

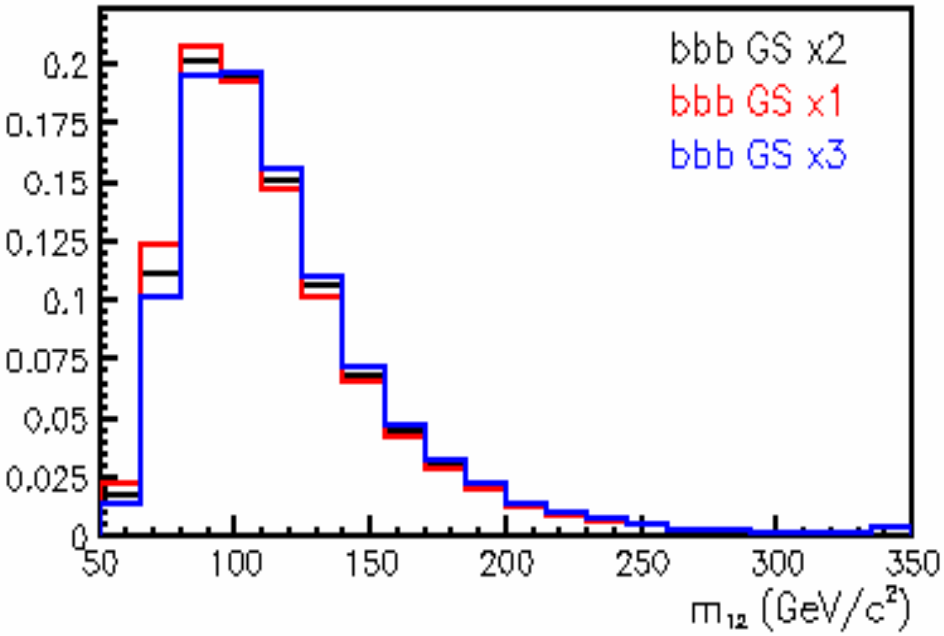




Systematic Errors

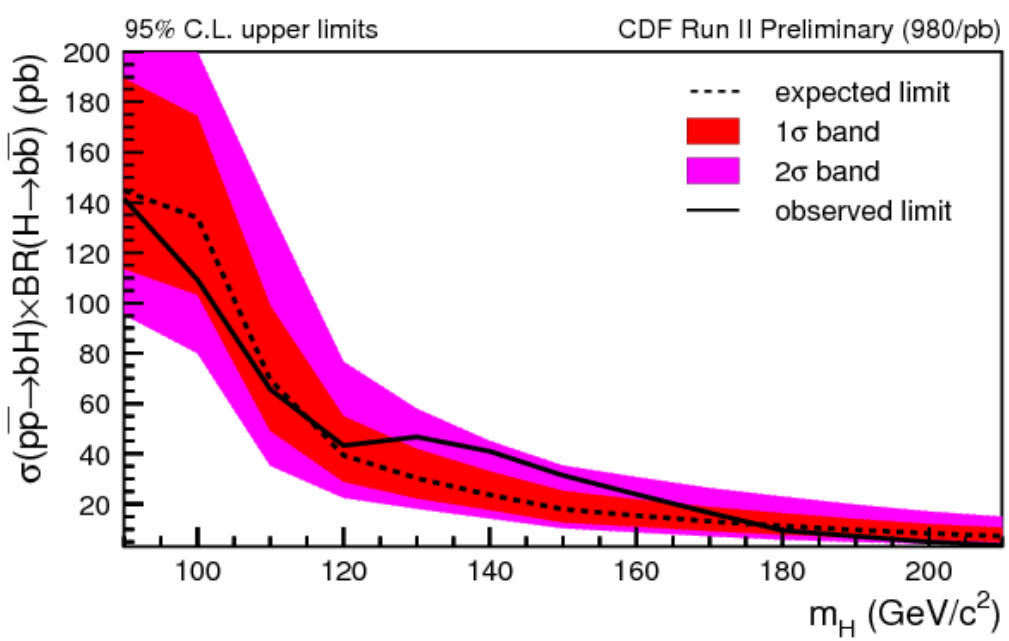
- Acceptance systematics
  - B-tag efficiency: 11-16%
  - PDF's: 8%
- Luminosity 6%
- Shape variations
  - Backgrounds: GS scaling, bbq vs bbc variation
  - Signal: Jet energy scale and SecVtx mass (±3%)
  - These have the most effect on the limits

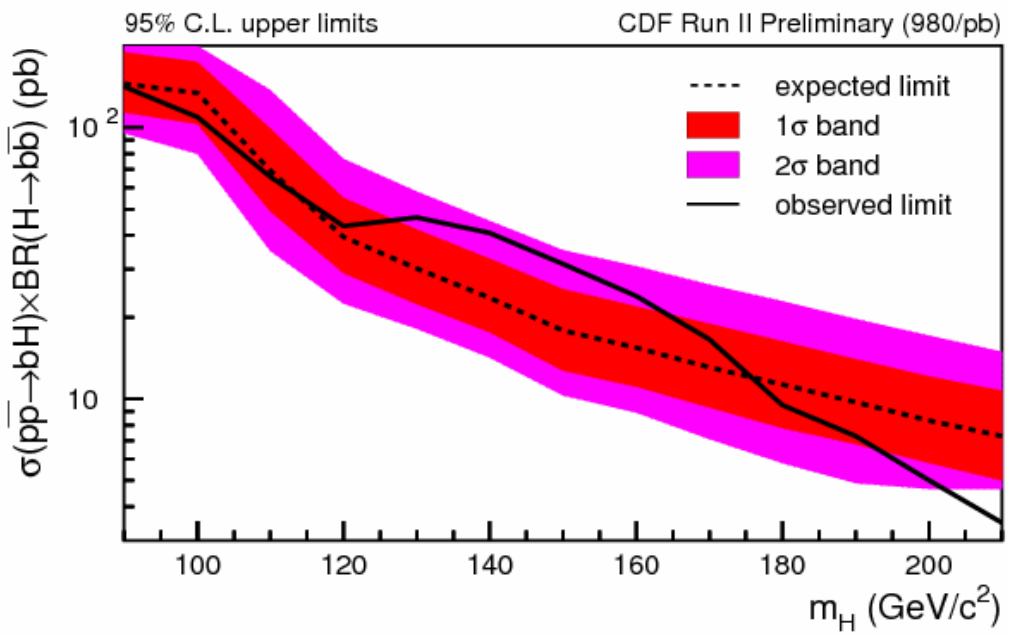




σ x BR Limits

- Use mclimit_csm* program to compute limits using CL_s method
  - Handles shape variations on signal/background templates using histogram interpolation
- Also use interpolation between MC signal samples (steps of 10 GeV/c²)
- Maximum positive deviation from expectation at 140 GeV/c²
  - $1\text{-}CL_b = 6\%$
  - including trials factor of ~5, would expect 30% of the time for the null hypothesis





* http://www.hep.uiuc.edu/home/trj/cdfstats/mclimit_

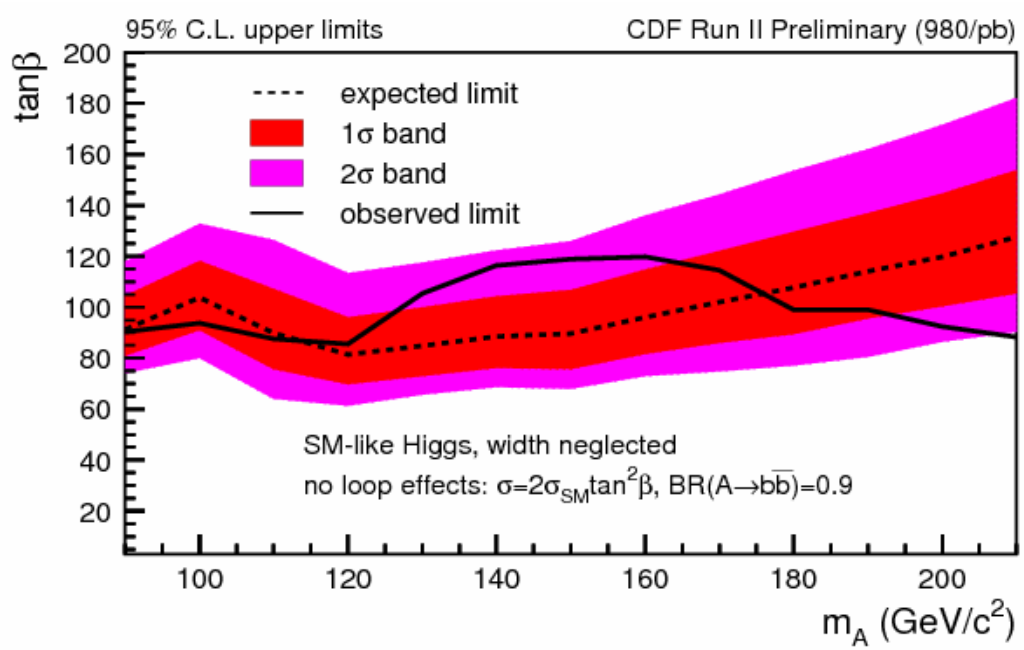
Simple tanβ Limits

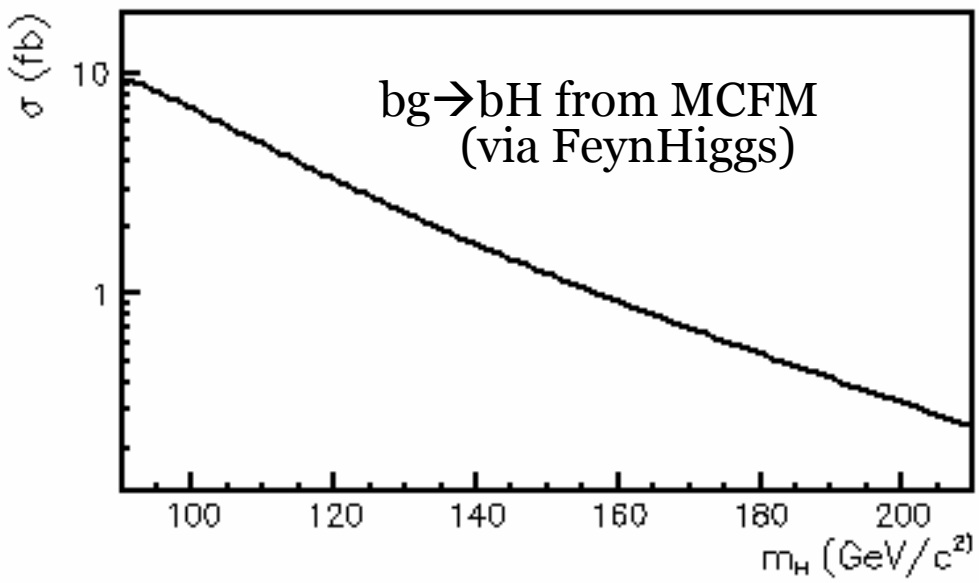
- At tree level,
  $\sigma \times BR = 2\sigma_{SM}\tan^2\beta \times 90\%$

- Get σ_{SM} from FeynHiggs
- Factor of 2 from h/H degeneracy
- Turn σ x BR limits into tanβ limits

- For tanβ limits, need to include systematics on the predicted cross section as well
  - PDFs:
    - was constant 8% on the acceptance
    - change to uncertainty on the total event yield
    - 8%(23%) for m_H=90(210)
  - Additional 8% variation for scale dependence of NLO calculation
- Limits change by ~5-10%

- These limits are not very realistic
  - Do not include loop corrections
  - Do not account for Higgs width





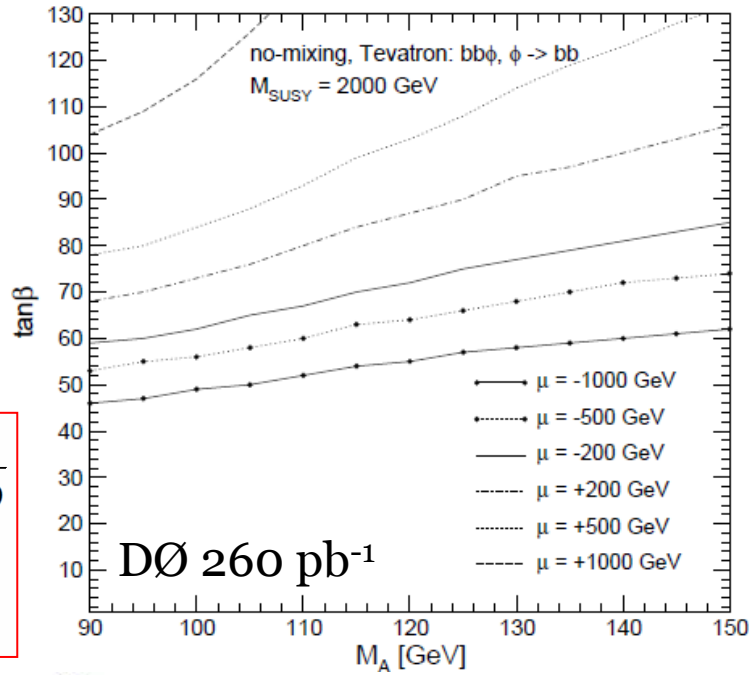
Scenario Dependence

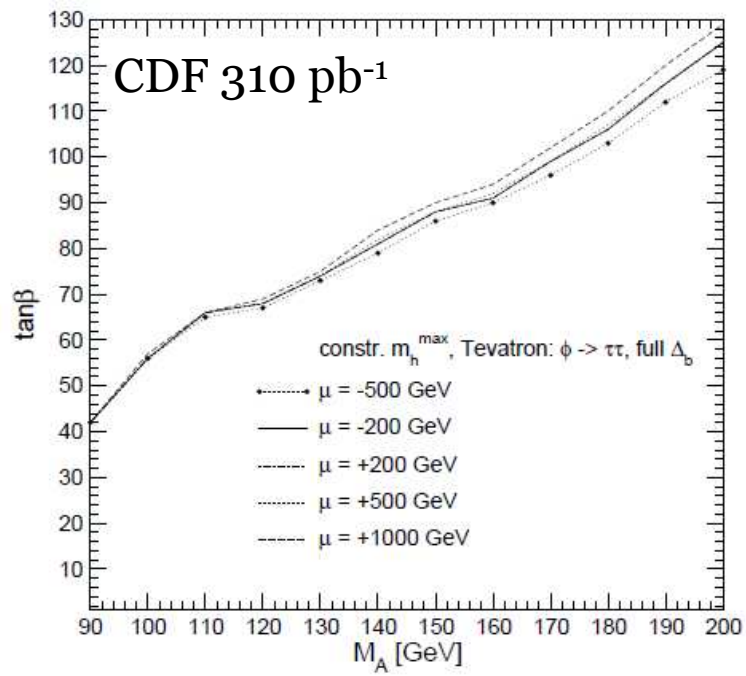
- Higgs properties are largely, but not completely, determined by m_A and tanβ
- Loop corrections introduce dependence on other SUSY parameters
  - M. Carena *et al.*, Eur.Phys.J. C45 (2006) 797-814 (hep-ph/0511023)

$$\sigma(b\bar{b} \to A) \times BR(A \to b\bar{b}) \cong \sigma(b\bar{b} \to A)_{SM} \times \frac{\tan\beta^2}{(1+\Delta_b)^2} \times \frac{9}{(1+\Delta_b)^2 + 9}$$

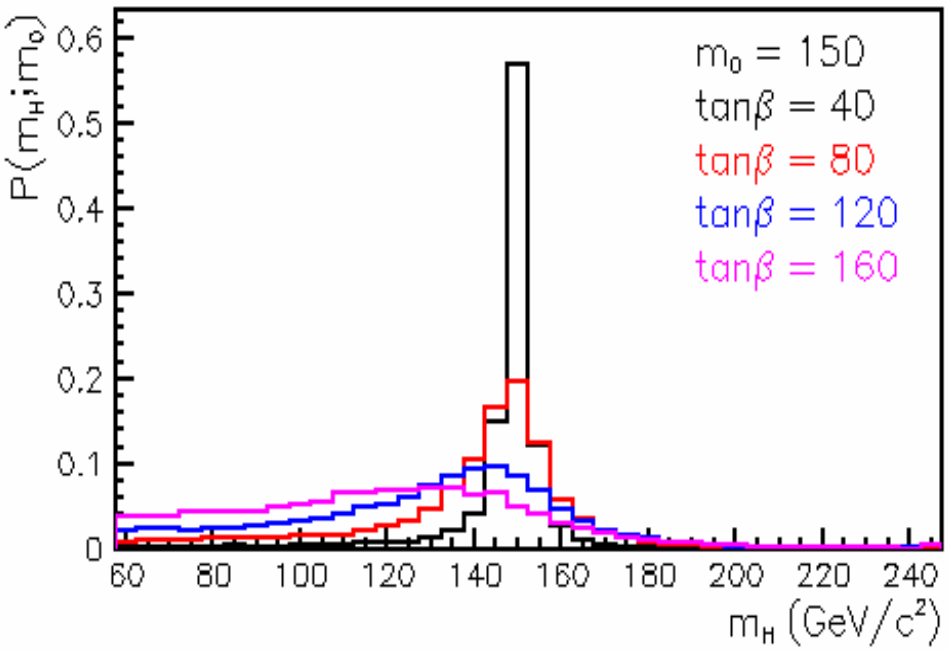
$$\sigma(b\bar{b}, gg \to A) \times BR(A \to \tau\tau) \cong \sigma(b\bar{b}, gg \to A)_{SM} \times \frac{\tan\beta^2}{(1+\Delta_b)^2 + 9}$$

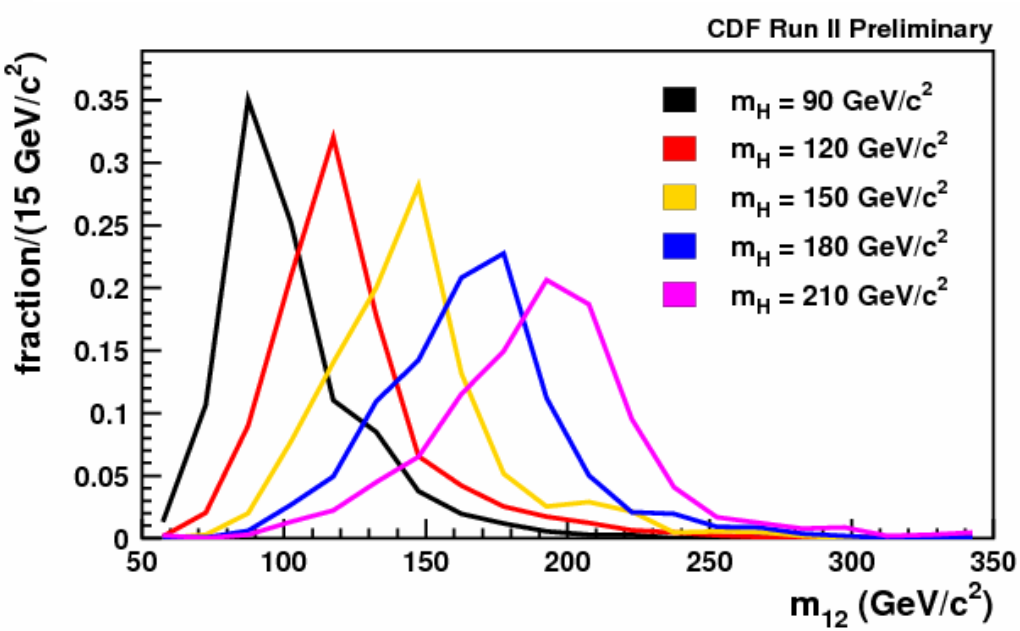
- Δ_b is a function of SUSY parameters
  - Δ_b proportional to μ x tanβ (sign of μ matters)
- Use two "benchmark" scenarios
- For tanβ = 50, μ = -200 GeV
  - m_h^{max}: $\Delta_b = -0.21$
  - constrained m_h^{max}: $\Delta_b = 0$

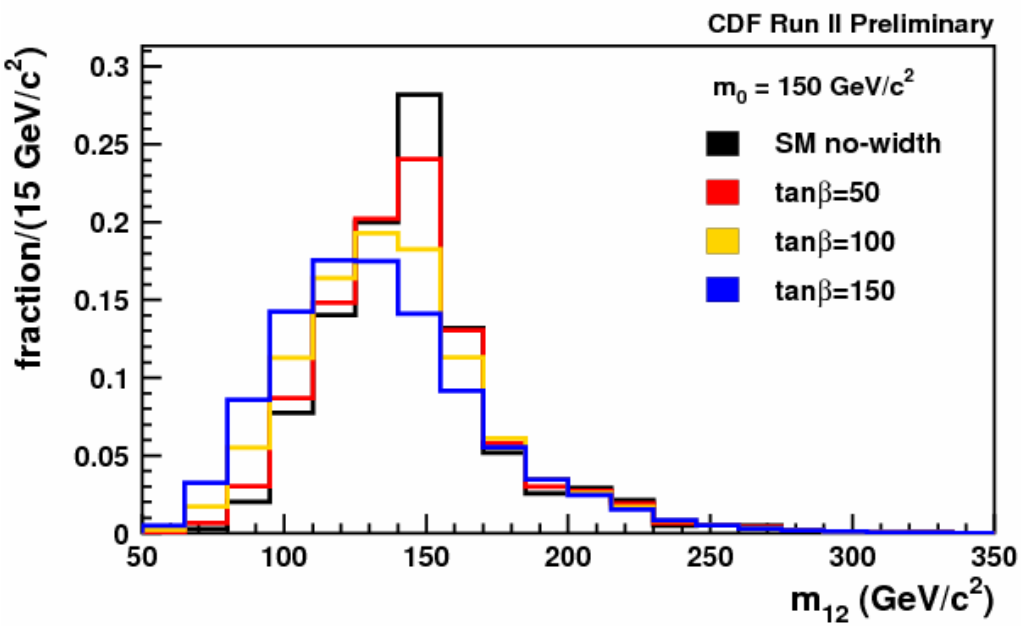




Fit Templates with Width



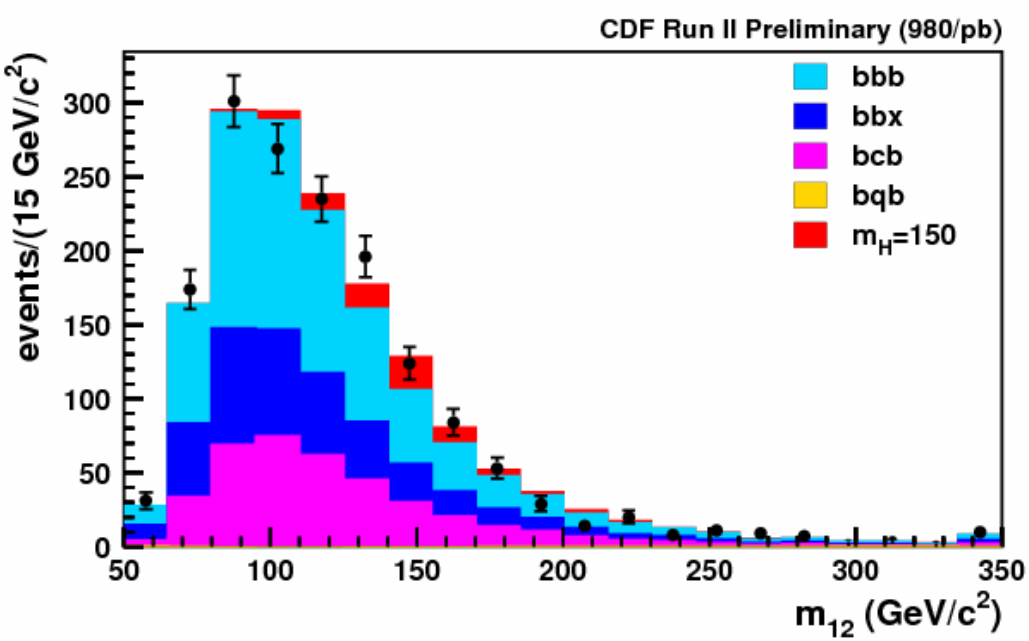


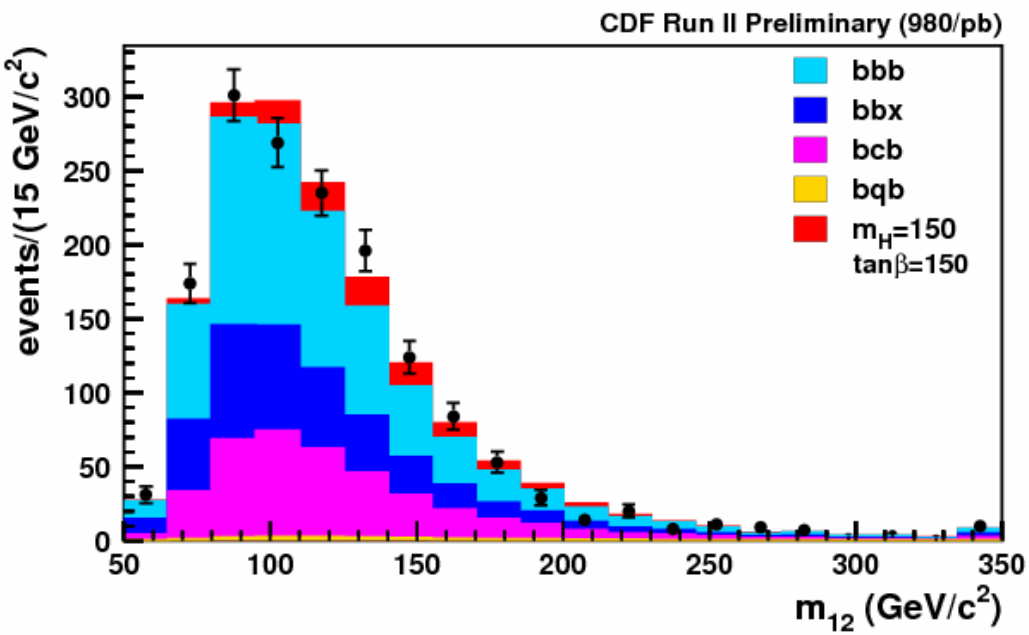


- Can't turn up cross section without increasing Higgs width
- Natural width eventually becomes comparable to experimental resolution
- Fit templates broaden and shift towards lower m_{12} because of falling cross section

Simple Fits with Width

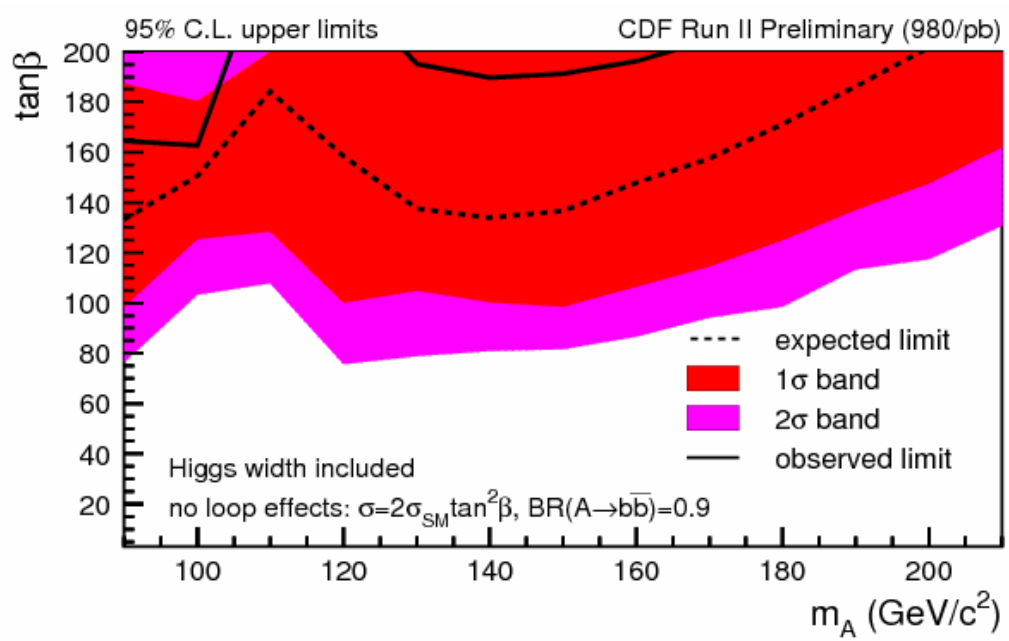
- Both fits are for $m_H = 150$
  - Upper is SM no-width case
  - Lower is for $\tan\beta = 150$
- Number of signal events is similar, but spread over more bins
  - Less significant
- Fit sensitivity goes down, so
  - Increase width
  - Limit goes up
  - Increase width some more
  - Limit goes up some more
  - etc… eventually converges

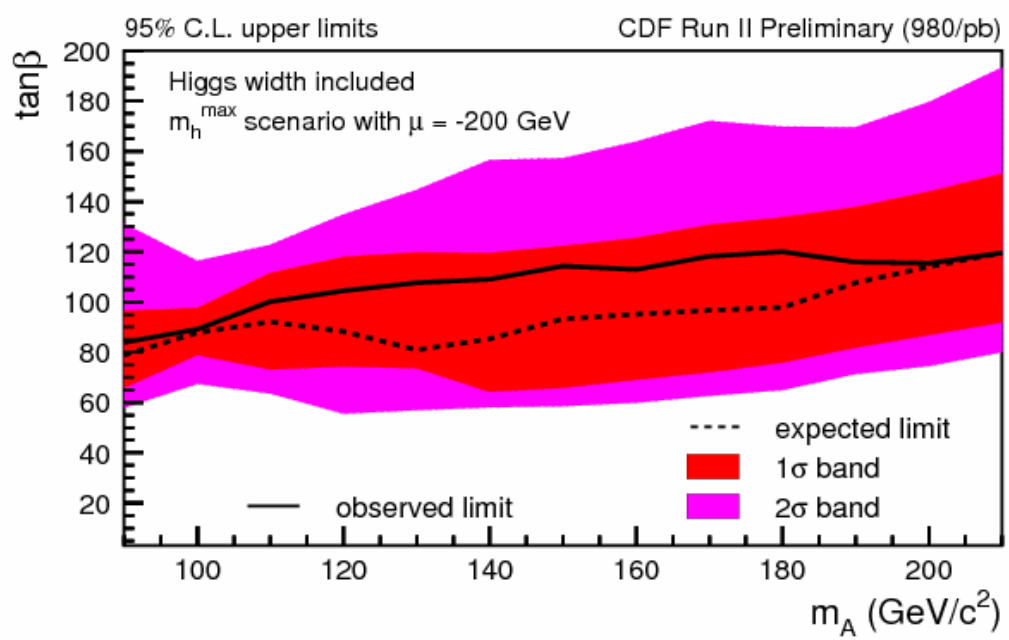




Benchmark tanβ Limits with Width

- Upper plot is for constrained m_h^{max} (Δ_b=0) and with width effect included
  - Limits slide off to high tanβ
- Lower plot is for m_h^{max} scenario with negative μ
  - Enhanced signal production improves the limits
  - Effect is greatest at large tanβ because $\Delta_b \sim$ tanβ
- With more data this problem should eventually solve itself
  - Sensitivity to smaller cross sections means we probe narrower Higgs regime
  - Limits will probably improve faster than $\sqrt{\mathcal{L}}$



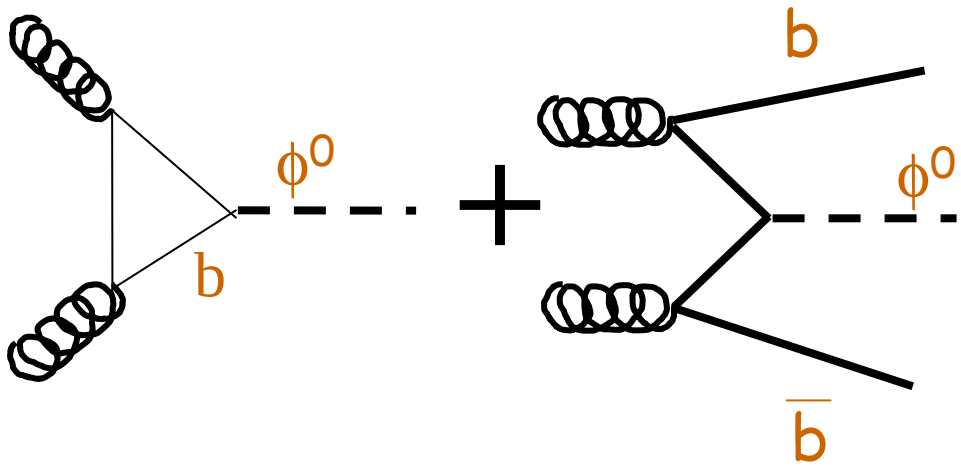


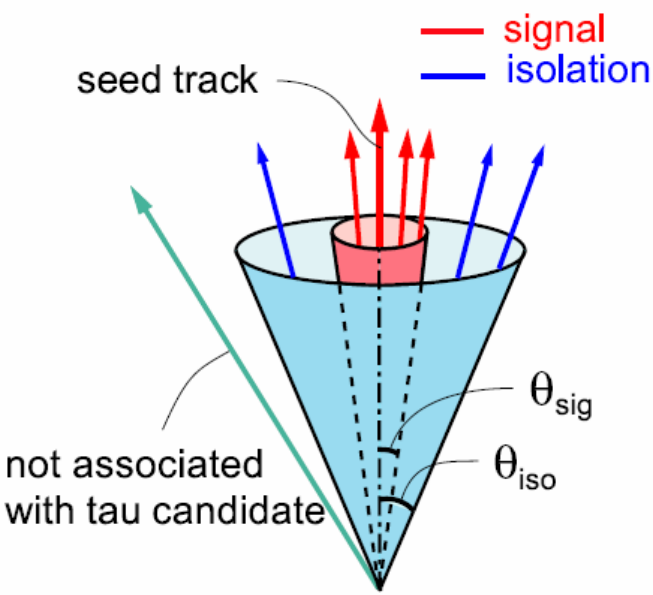
The gg+bb $\rightarrow \phi \rightarrow \tau\tau$ Channel

- High cross section and unique final state (not QCD)
  - Allows inclusive search
- Low sensitivity to Δ_b

- Best signatures are:
  - one τ decay into e or μ and the other hadronic (23+23% BR)
  - one τ decay into e and the other into μ (6% BR)

- Event selection
  - Leptons: $p_T > 10$ GeV/c
  - Hadronic τ: $p_T > 15$ GeV/c, mass < 1.8 GeV/c²
  - Opposite charge
  - Missing-E_T not recoiling against leptons (rejects $W \rightarrow l\nu$)





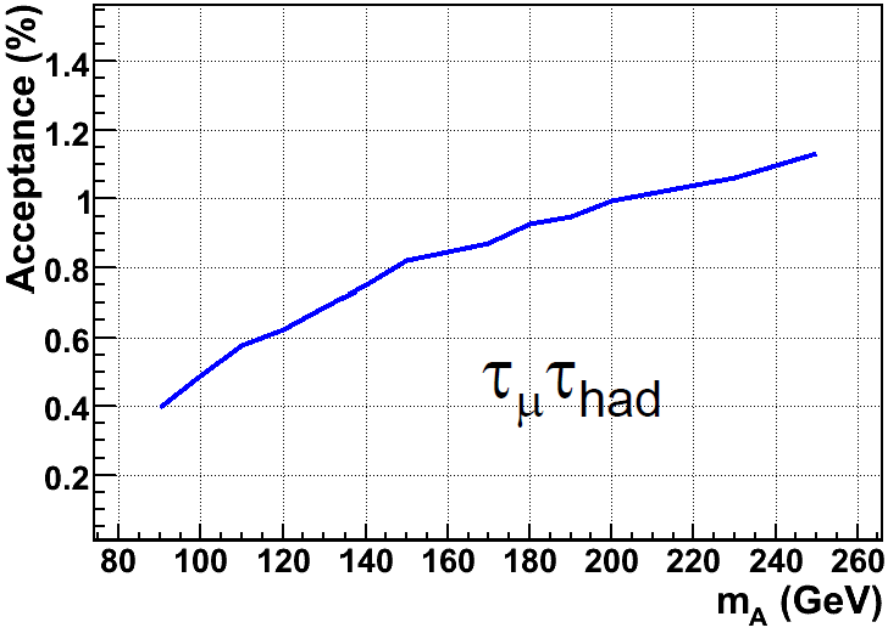
$\phi \rightarrow \tau\tau$ Analysis Inputs

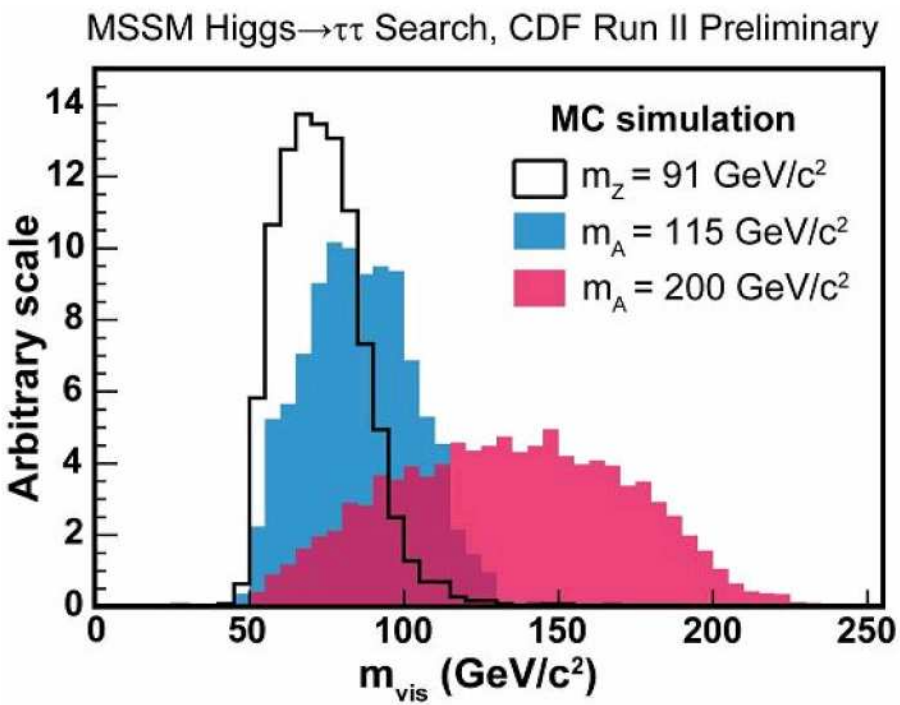
- Acceptance is 0.5-1% for $\tau_\mu \tau_{had}$, similar for $\tau_e \tau_{had}$, and 0.25-0.5% for $\tau_e \tau_\mu$ (including BR's)

- Not possible to reconstruct $\tau\tau$ invariant mass due to multiple neutrinos
- Instead, use "visible mass" m_{vis}
  - Mass of the leptons and hadronic τ's and the net missing-E_T (sum of ν's)

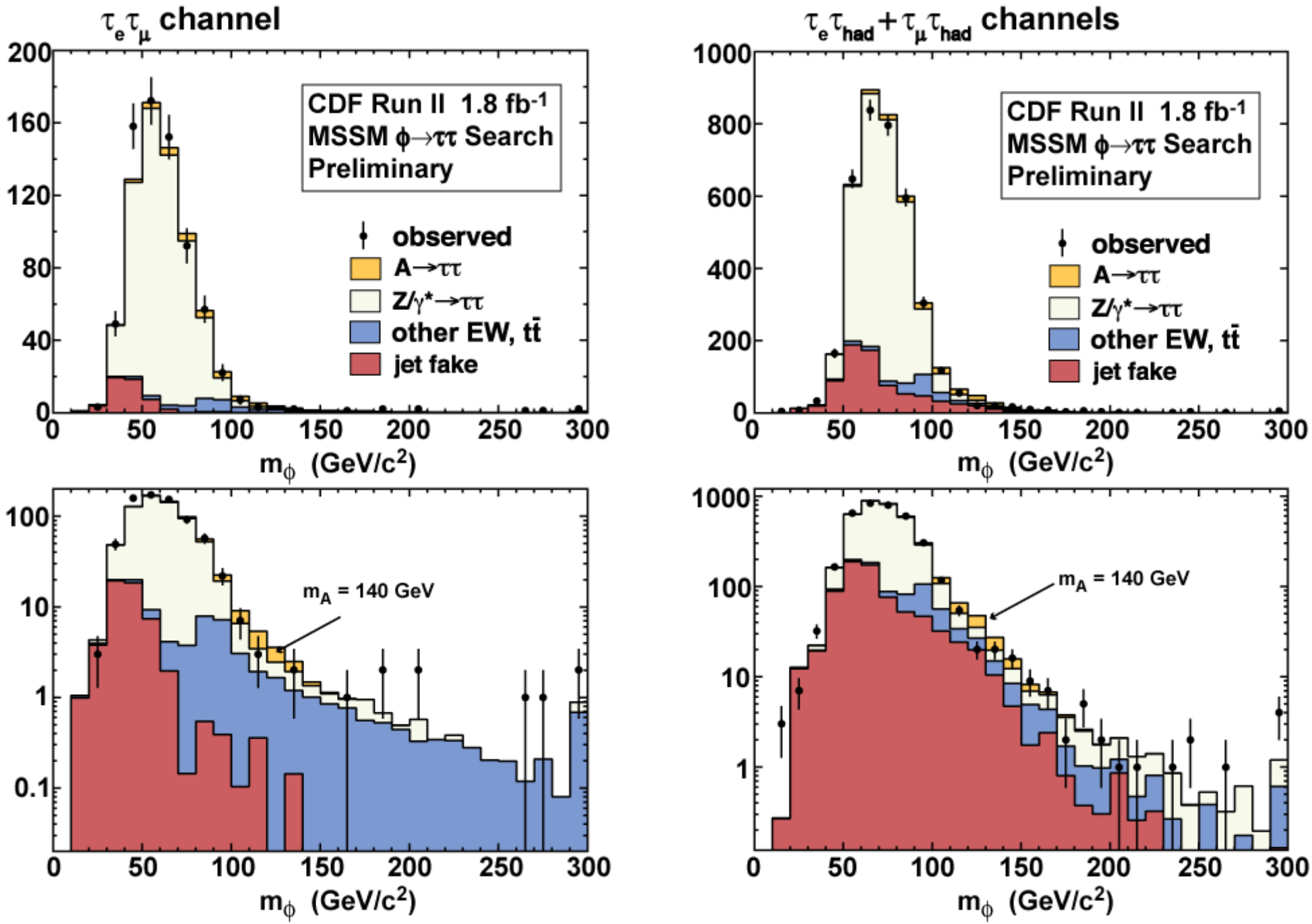
- Advantage: can be computed for any event
- Higgs width effects are not going to matter here

- Main background is $Z \rightarrow \tau\tau$, with small contributions from other EWK processes and QCD jets

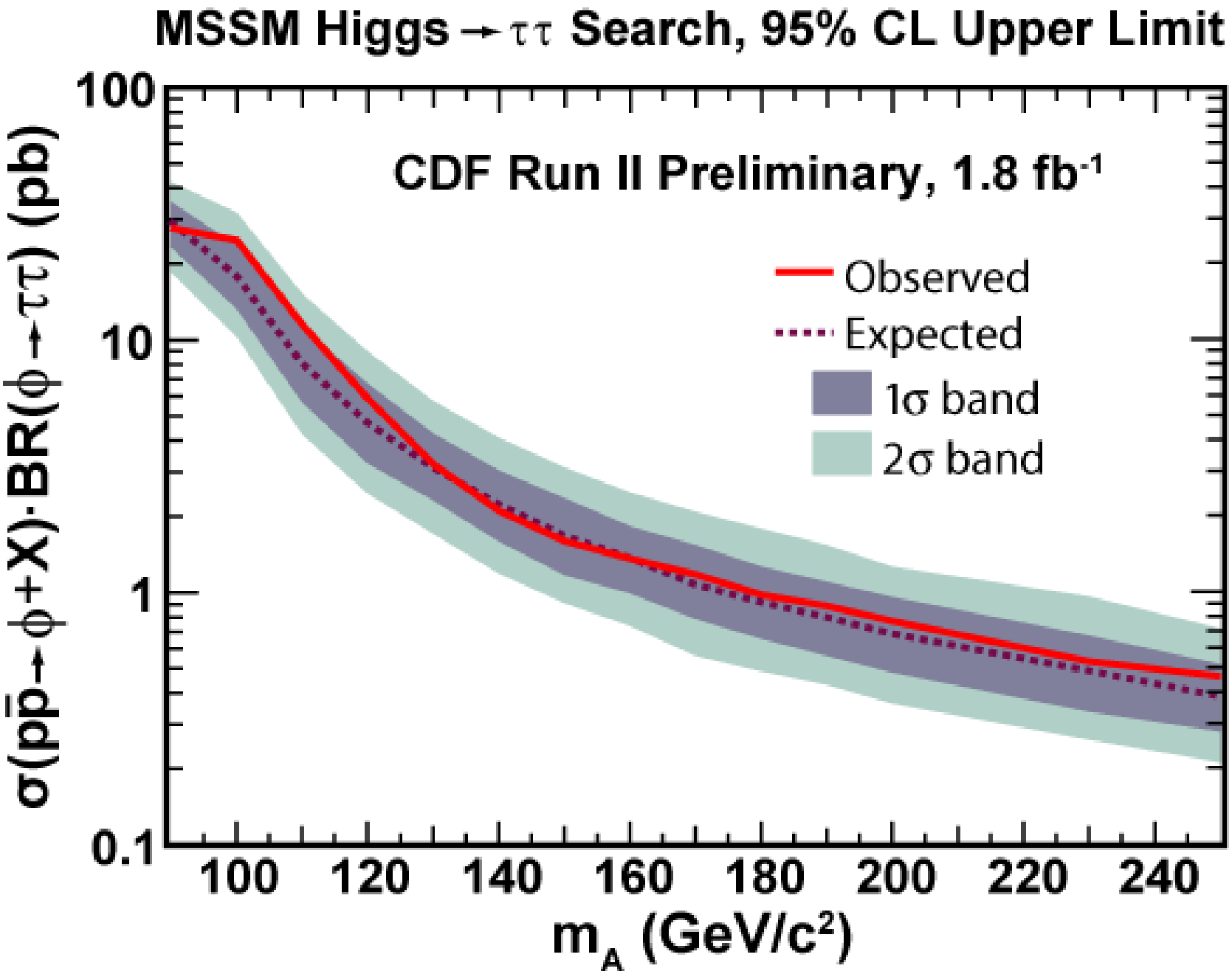




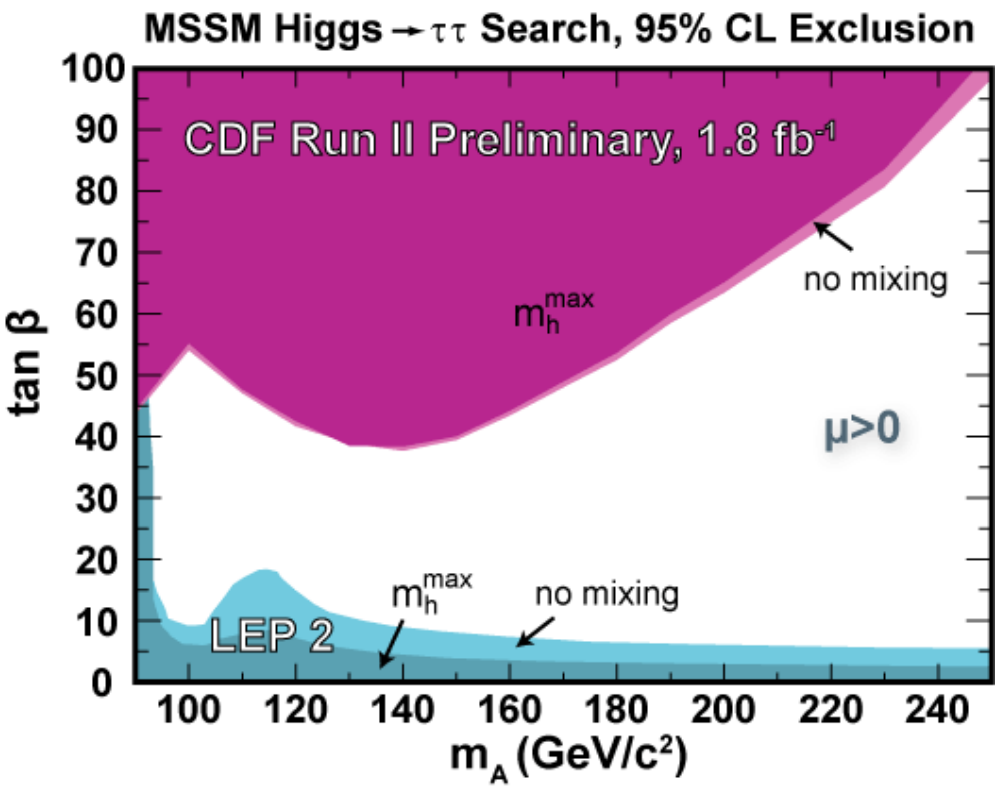
Visible Mass Distributions

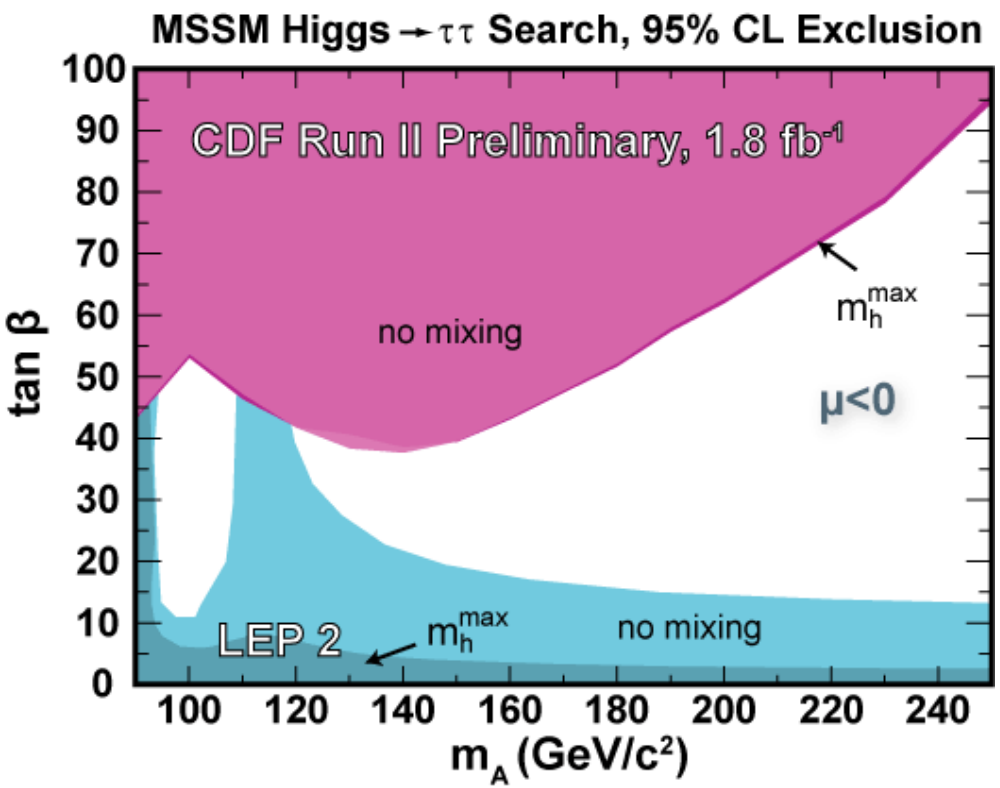


σ x BR Limits (Hot Off the Press!)



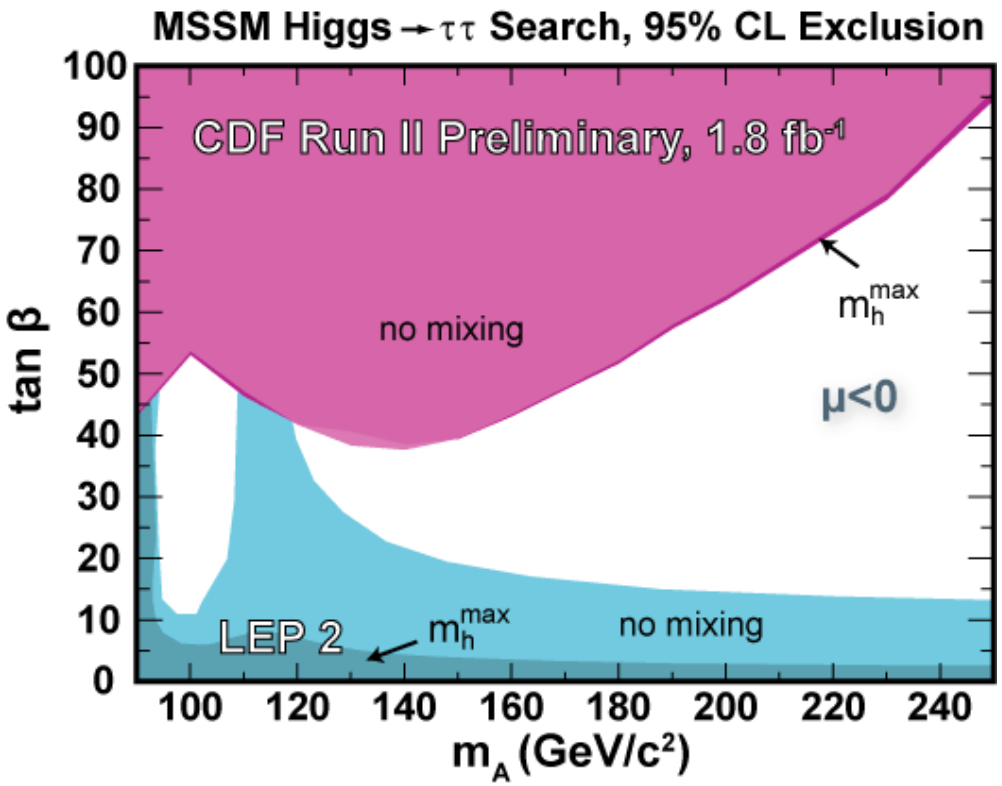
MSSM Interpretation

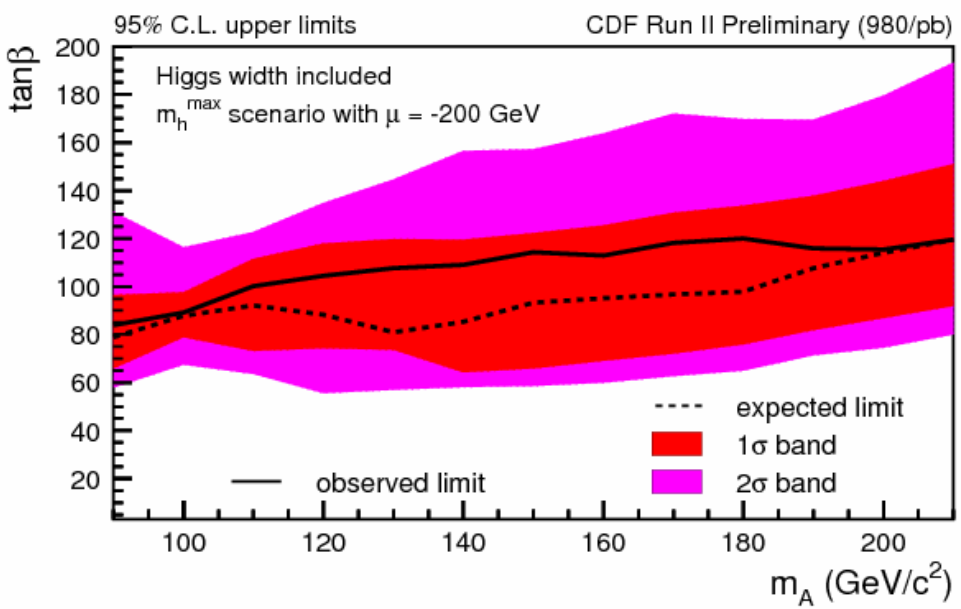




- Benchmark scenarios are:
  - m_h^{max} : $\Delta_b = \pm 0.21$ @ $\tan\beta = 50$
  - no mixing : $\Delta_b = \pm 0.10$ @ $\tan\beta = 50$
- Excluded regions insensitive to different scenarios or flipping the sign of μ – much different than in the 3b channel

Side-by-Side





- Even for m_h^{max} with μ = -200 (best-case for 3b looked at so far), ττ channel is still considerably more sensitive (but twice the data also)
- Going to even more negative μ would bring them closer together

- If we were to eventually observe something, could use the two channels together to measure not only tanβ and m_A but also Δ_b

Summary and Outlook

- CDF is pursuing the MSSM Higgs in both of the preferred high-tanβ decay modes $\phi \rightarrow bb$ and $\phi \rightarrow \tau\tau$

- Result in the 3b channel on 1 fb^{-1} released this summer
  - No significant excess observed

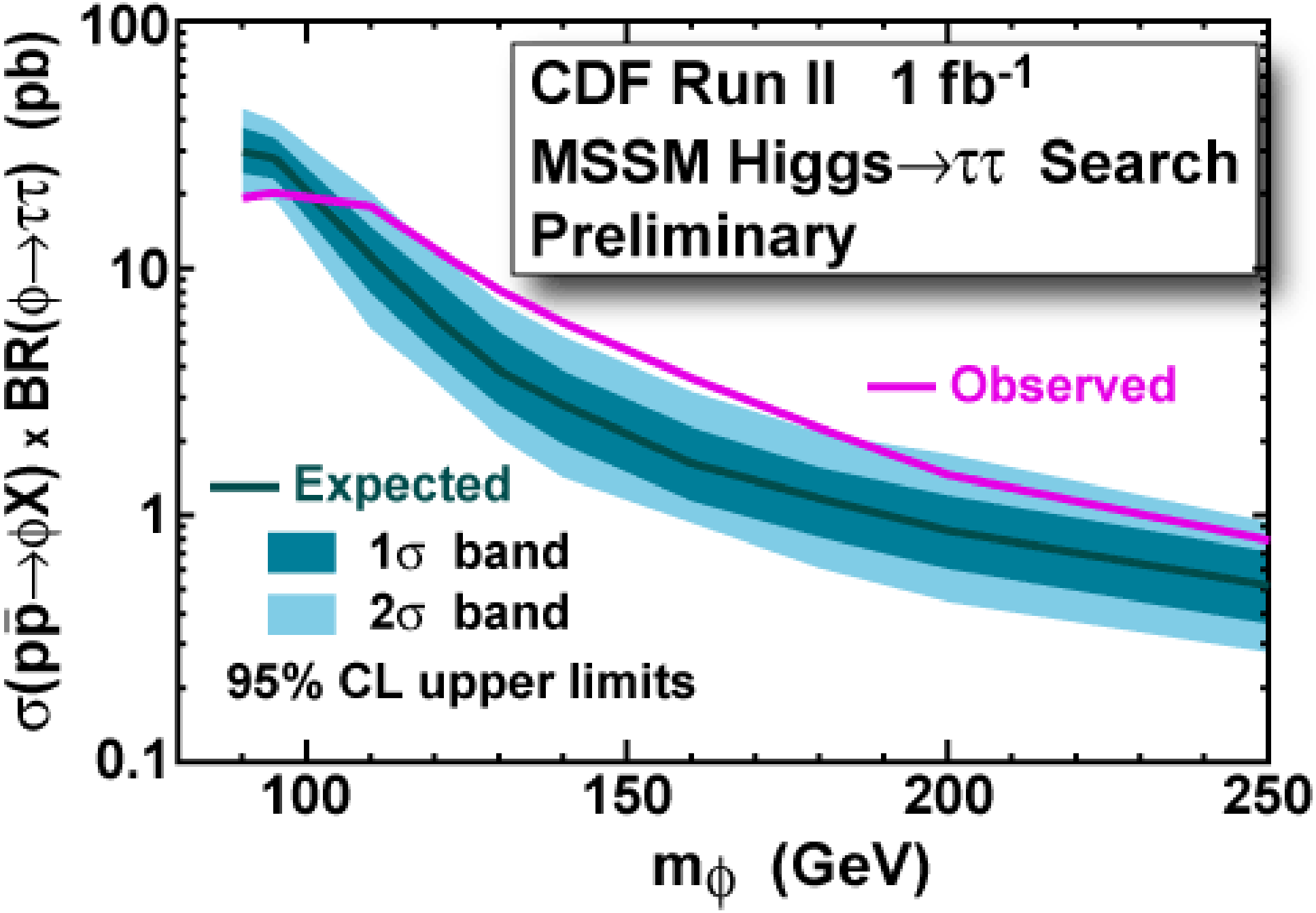
- Brand-new result in the $\tau\tau$ channel on 1.8 fb^{-1} approved today
  - The 2σ excess around 160 GeV/c^2 in the 1 fb^{-1} result went away

- Both channels have improvements in the pipeline
  - 3b
    - Add another 1 fb^{-1} of data already recorded
    - Improve the tag mass variable m_{diff} for better background fit
  - $\tau\tau$
    - Split events into b-tagged and untagged subsamples
    - Switch to a multivariate approach

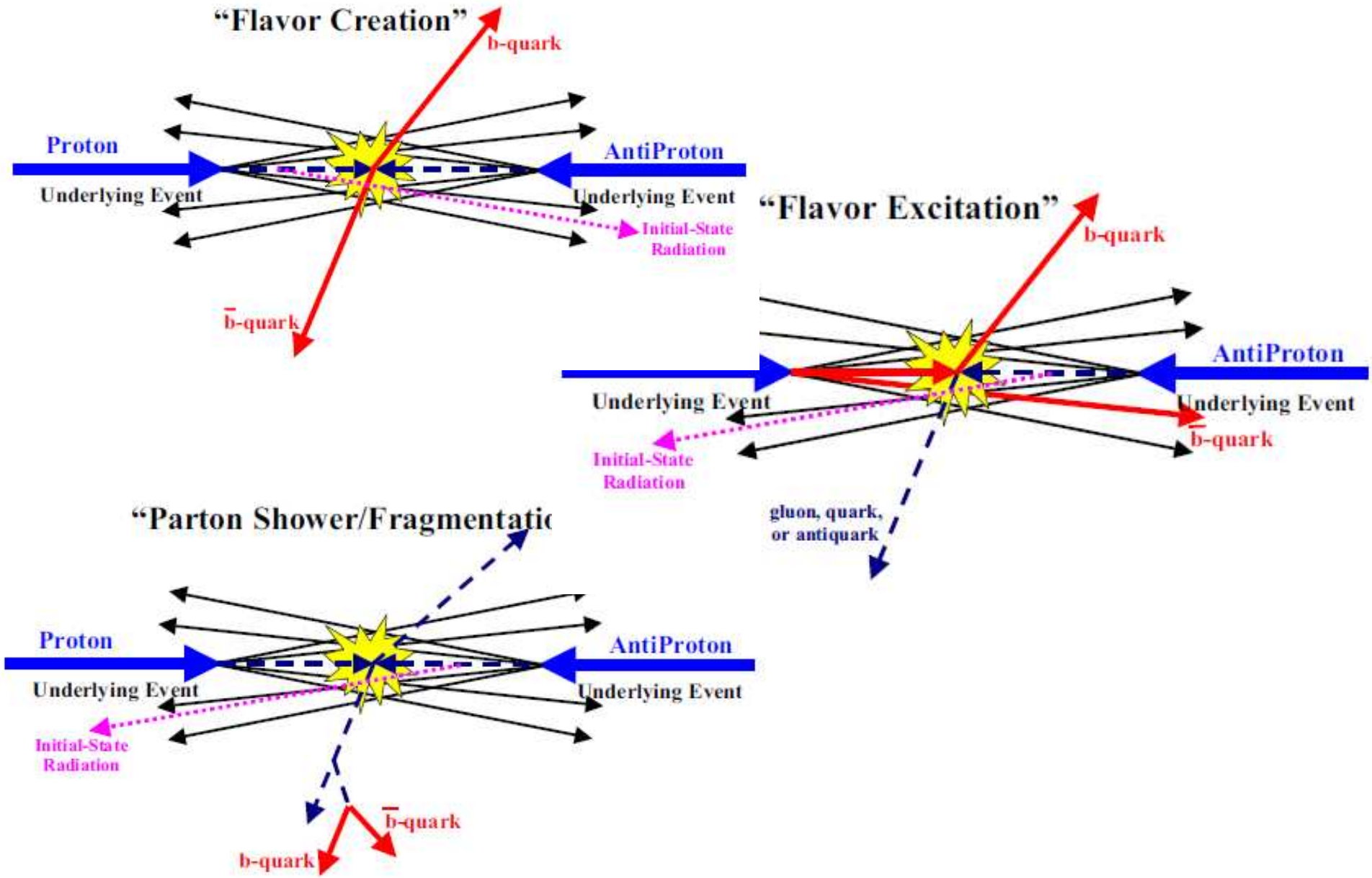
- With 6-8 fb^{-1} we will be able to push even farther

Backup Material

1 fb^{-1} $\phi \rightarrow \tau\tau$ Limits



QCD Heavy Flavor Production



thanks to Rick Field

Pythia Process Calibration

- Calibration using the $\Delta\phi$ distribution found that gluon splitting to bb in the Pythia "High ISR" scenario (our default) is underestimated by factor ~2
- Rest is made up of flavor creation and flavor excitation (hard to differentiate)
  - In our three-jet events the distinction is not important

- We choose a GS scale factor of 2 as the default, and factors of 1 (no correction) and 3 as systematic variations

- GS = gluon splitting
- FE = flavor excitation
- FC = flavor creation

	Electrons			
	Fixed Normalization		Floating Normalization	
PYTHIA High ISR	FC:	43.7%	FC:	66.1%
	FE:	40.7%	FE:	1.7%
	GS:	15.6%	GS:	32.2%
$\chi^2/d.o.f.$	125.7/29		34.1/27	
χ^2 probability	5.26×10^{-14}		0.136	