Search for the SM Higgs boson in the di-tau final state at Tevatron

Pierluigi Totaro

University of Trieste

On behalf of the

CDF and DØ collaborations

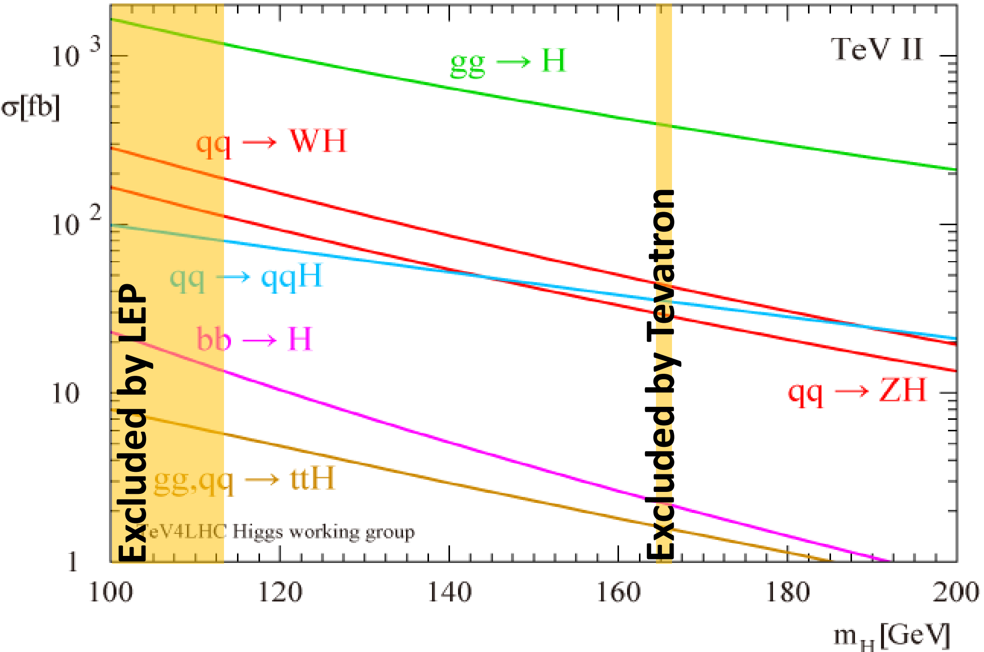
35th International Conference on High Energy Physics

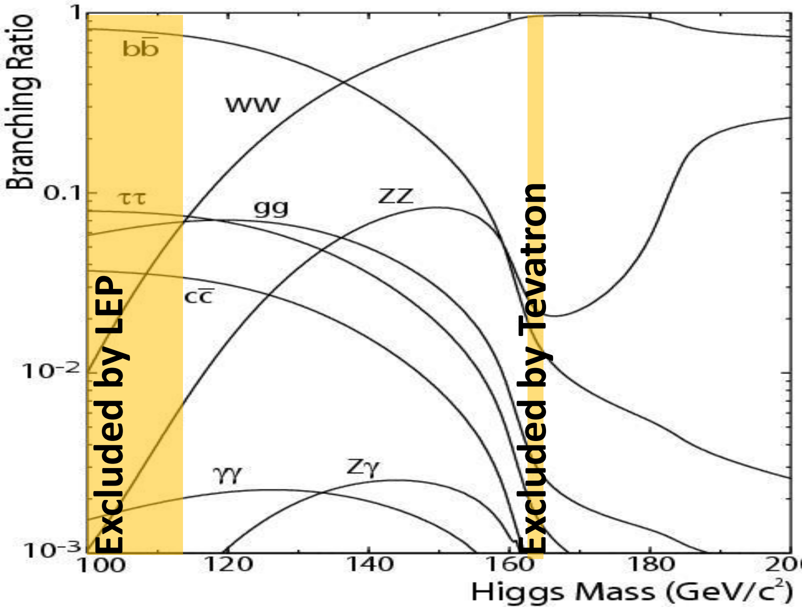
Paris, July 23rd 2010

Outline

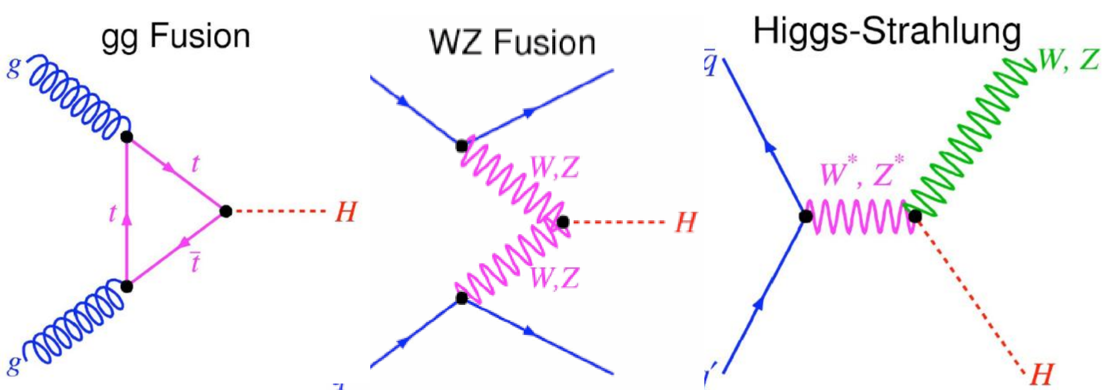
- Standard Model Higgs production and decay at Tevatron
  - Low mass searches
- Motivation for $H \to \tau\tau$ searches
- Analysis strategies for CDF and DØ experiments
- Results: CDF 2.3 fb^{-1} DØ 4.9 fb^{-1}
- Conclusions

Higgs production and decay at Tevatron





Primary production modes:



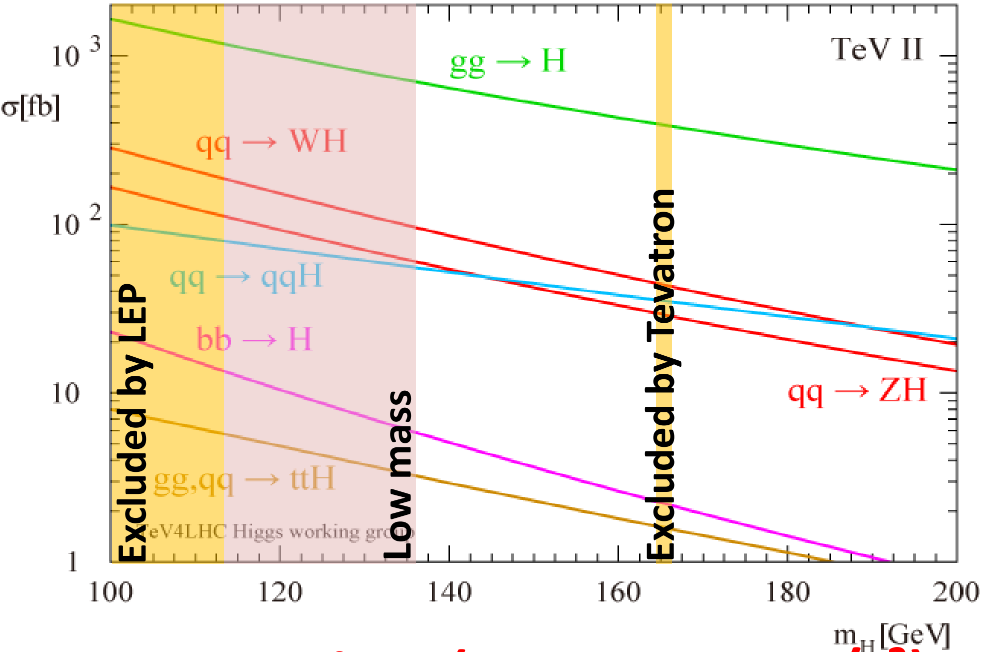
1.8 – 0.2 pb   0.1 – 0.02 pb   0.5 – 0.03 pb

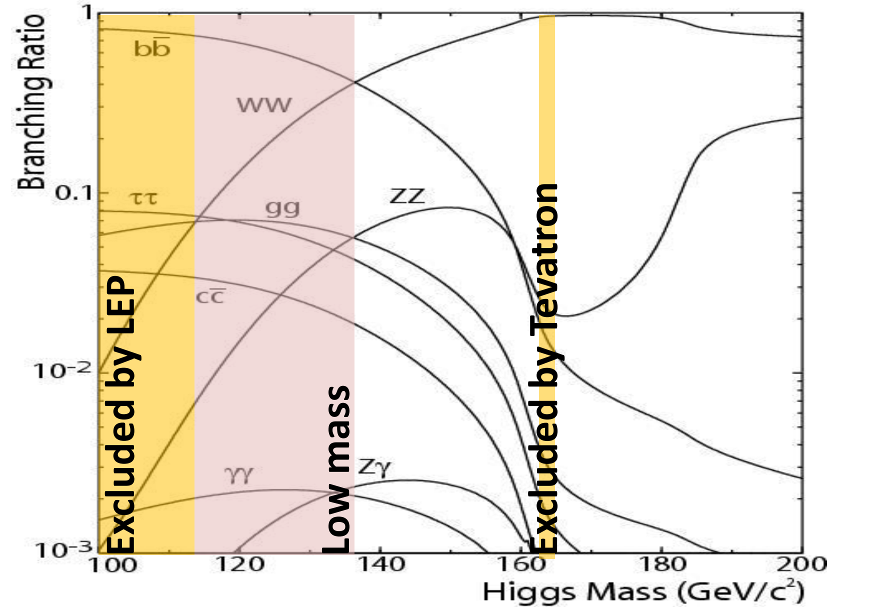
Principal decay modes:

H→bb for M_H<135 GeV/c^2

H → WW* for M_H>135 GeV/c^2

Low Mass Higgs searches at Tevatron





Low mass Higgs (M_H < 135 GeV/c^2)

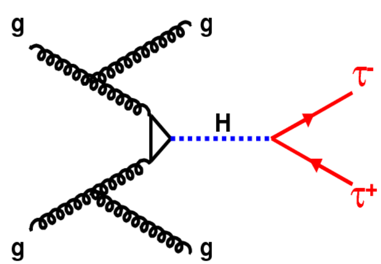
1) gg→H→bb — Very hard: overwhelmed by multijet background
2) WH →lνbb,ZH →llbb, ZH →ννbb — good event selection handles: lepton and b-quark tagging
3) H→ττ — Low branching fraction but more unique event signature

H→ττ searches: motivation

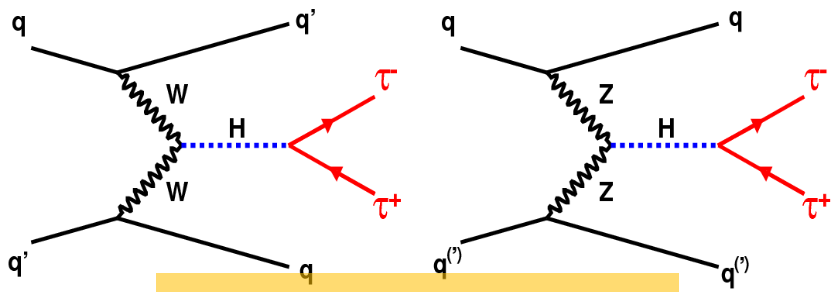
H→ττ branching ratio is small(<10%)

BUT

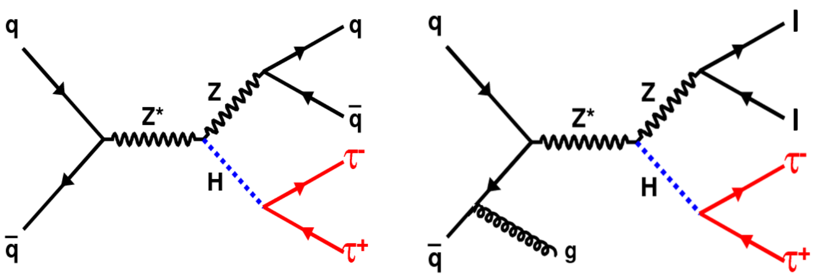
1)Different channels can be studied simultaneously
2)Direct production and VBF become accessible
3)Hadronic W/Z decays in the associated production can be included



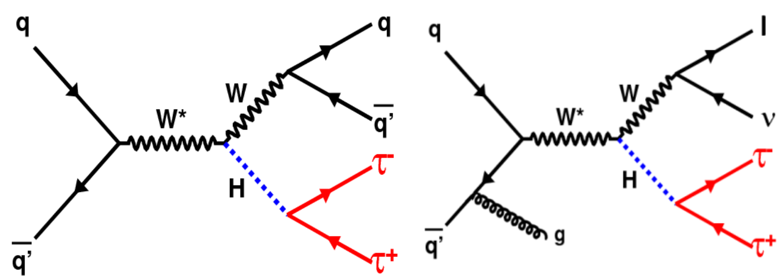
GLUON FUSION



VECTOR BOSON FUSION



Z BOSON ASSOCIATED PRODUCTION



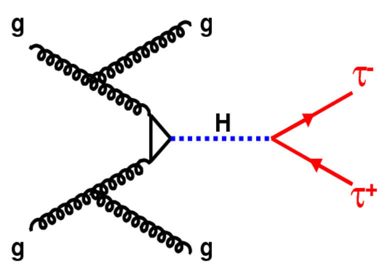
W BOSON ASSOCIATED PRODUCTION

H→ττ searches: motivation

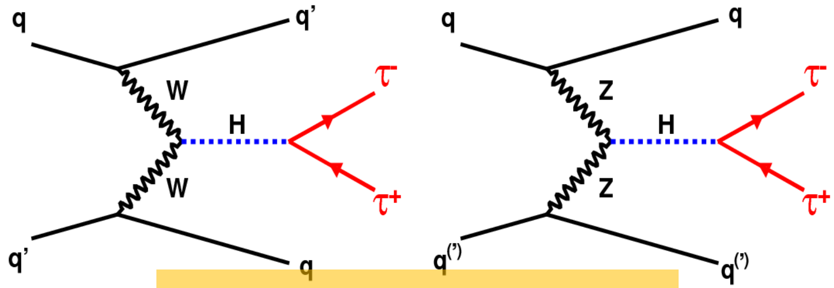
H→ττ branching ratio is small(<10%)

BUT

1)Different channels can be studied simultaneously
2)Direct production and VBF become accessible
3)Hadronic W/Z decays in the associated production can be included

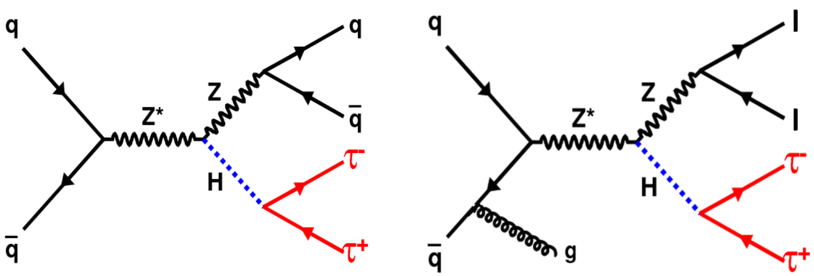


GLUON FUSION

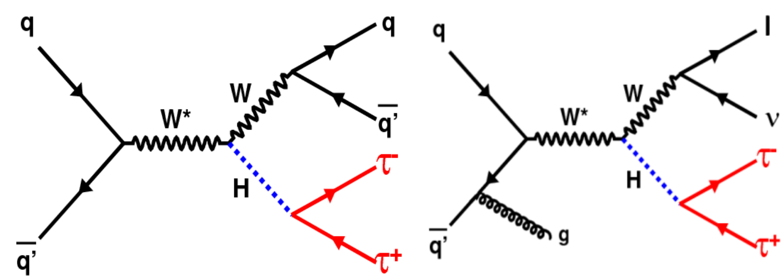


VECTOR BOSON FUSION

Signature search:
- **2 taus**
- **≥ 1 jets**



Z BOSON ASSOCIATED PRODUCTION



W BOSON ASSOCIATED PRODUCTION

What about tau leptons?

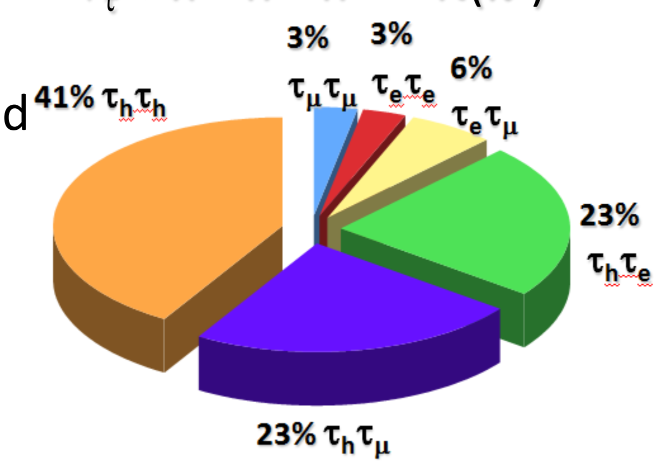
- Heavy particles: 1.78 GeV/c^2
- Short lived: mean lifetime 291 ps ($c\tau$=87 μm) → Detectable only through their decay products
- Decay modes:
  - $\tau \rightarrow \nu_\tau \nu_e e$ (B.R.~17%)
  - $\tau \rightarrow \nu_\tau \nu_\mu \mu$ (B.R.~17%)

  } Look for isolated electrons or muons

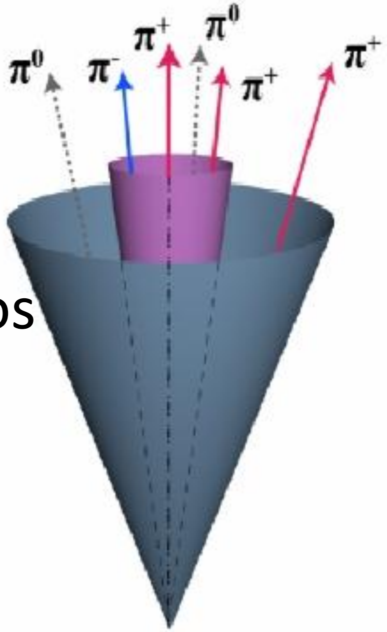
  - $\tau \rightarrow \nu_\tau X_h$ (B.R.~65%) → Hadronic decays:

  1-prongs $\tau^\pm \rightarrow \nu_\tau + h^\pm + N(\pi^0)$

  3-prongs $\tau^\pm \rightarrow \nu_\tau + h^\pm h^\pm h^\pm + N(\pi^0)$
- Di-tau decay combinations:
  - Hadronic+hadronic: 42 % large multijet background
  - Leptonic+hadronic: 46 % **golden channel**
  - ee/$\mu\mu$: 6 % large irreducible Drell-Yan background
  - $e\mu/\mu e$: 6% clean signature but less events



Hadronic tau identification



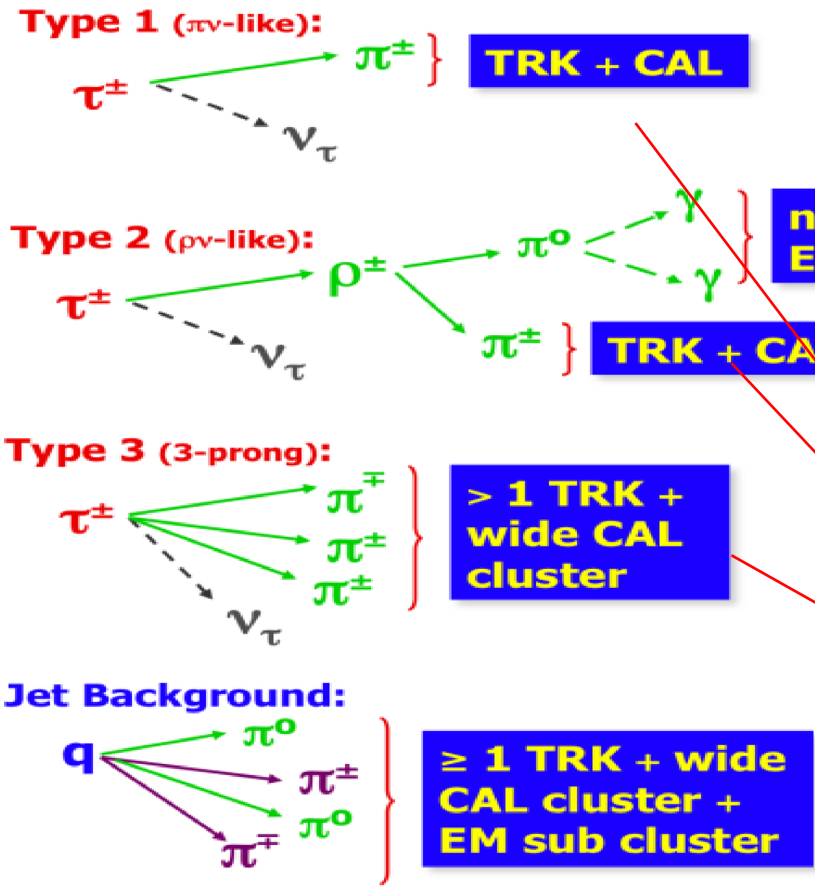
Very challenging task

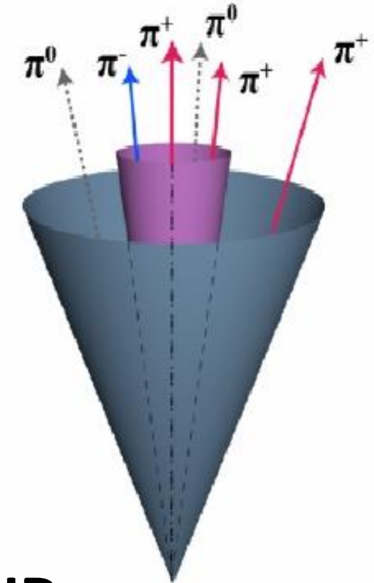
The signature: narrow calorimeter clusters with low track multiplicities
quark/gluon jets can easily lead to fakes

Reconstruction: - very difficult due to undetected neutrinos
- Cluster tau decay hadrons in cones

Identification (ID):
- based on **calorimeter and track isolation requirements**
- Multivariate selections are better than rectangular cuts to exploit correlations and provide a good τ-jet separation:
D0: Neural Networks (NN)
CDF: Boosted Decision Trees (BDTs)
- best performances achieved by considering different tau decay modes separately

Hadronic tau identification





DØ: Neural Network ID

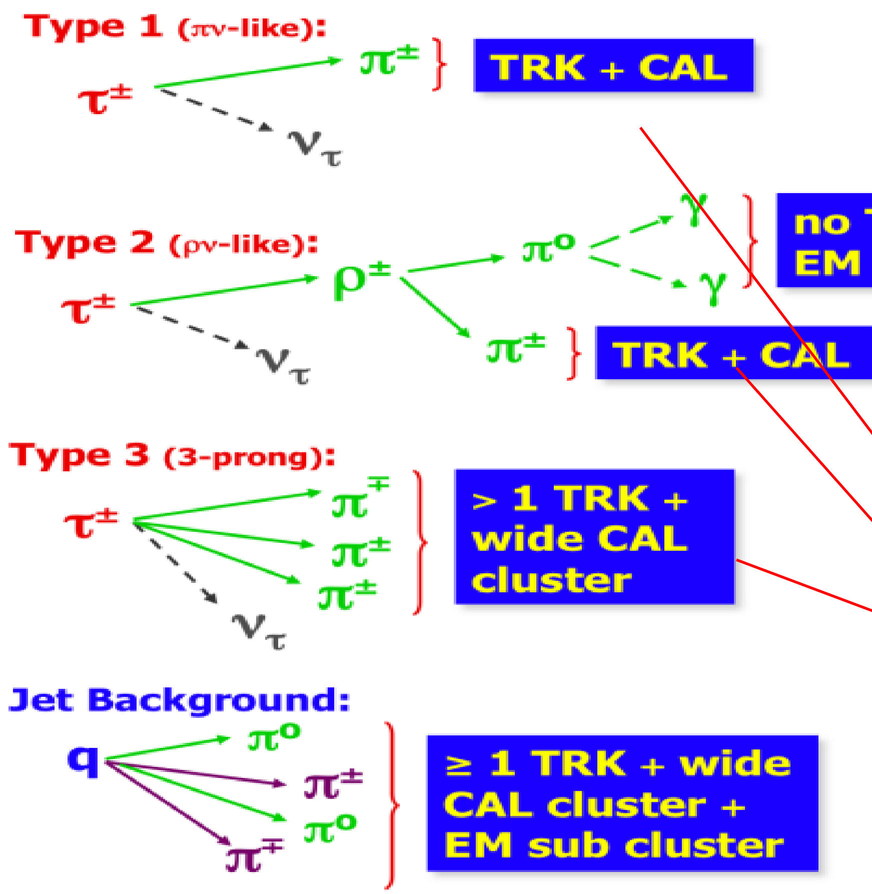
Tau categories:

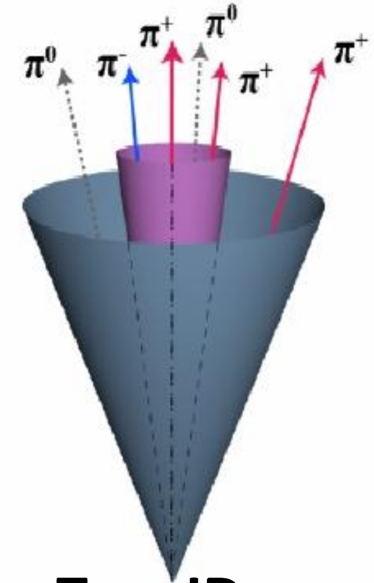
-1 prong $(\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm})$

-1 prong + neutrals $(\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm} + (N)\pi^{0}$'s

-3 prongs $(\tau^{\pm} \rightarrow \nu_{\tau} + 3\pi^{\pm} + (N)\pi^{0}$'s)

Hadronic tau identification





CDF: Boosted Decision Tree ID

Tau categories:

-1 prong ($\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm}$ +(N)π^{0}'s)

-3 prongs ($\tau^{\pm} \rightarrow \nu_{\tau} + 3\pi^{\pm}$ +(N)π^{0}'s)

Different energy subsamples trained separately

Strategies for the analysis 1

SIGNATURE SEARCH: similar approaches for CDF and DØ

- looking for leptonic+hadronic di-tau decay modes.
- jets in the final state optimize sensitivity for $qqH \rightarrow qq\tau\tau$, $WH \rightarrow qq\tau\tau$ and $ZH \rightarrow qq\tau\tau$. $gg \rightarrow H$ events with jets from initial state radiation (ISR) are also included

CDF

One isolated lepton (e/μ)	$p_T > 10$ GeV/c
One hadronic tau	$p_{TVIS} > 15$ GeV/c
Opposite charges	
$\geq$1calorimeter jet:	
- $E_T > 20$ GeV	
- pseudorapidity: $\lvert\eta\rvert < 2.5$	

DØ

One isolated muon	$p_T > 15$ GeV/c
One 1(3)-prong had. tau	$p_{TVIS} > 15(20)$ GeV/c
Opposite charges	
$\geq$2calorimeter jets:	
- $E_T > 20$ GeV	
- pseudorapidity: $\lvert\eta\rvert < 3.4$	

Strategies for the analysis 2

BACKGROUND ESTIMATION

IRREDUCIBLE PHYSICS CONTRIBUTIONS

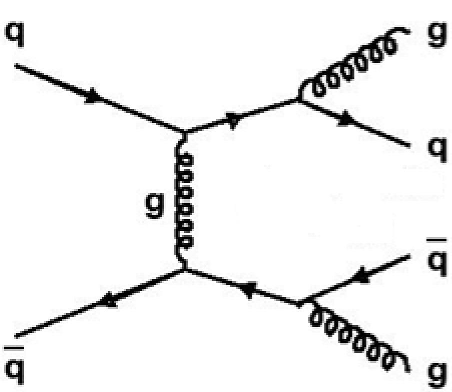
$Z \rightarrow \tau\tau$, top-antitop, dibosons : from MC

BACKGROUND FROM MISIDENTIFIED LEPTONS

W+jets, γ+ jet,multijet: based on MC and data driven techniques

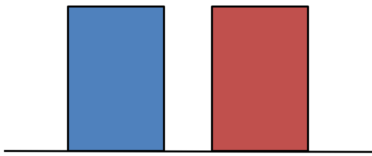
THE CHALLENGE: evaluate jet$\rightarrow\tau$ fake rate.

To estimate multijet bkg, both CDF and DØ use **same-sign (SS) data**:



jet $\rightarrow$ e/μ
jet $\rightarrow \tau$

Little or no correlation is expected between charges

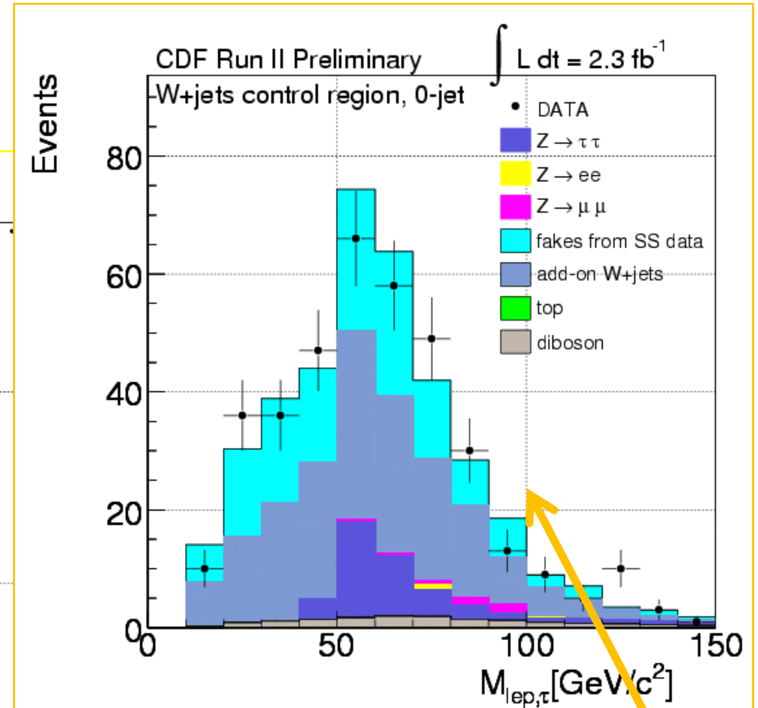


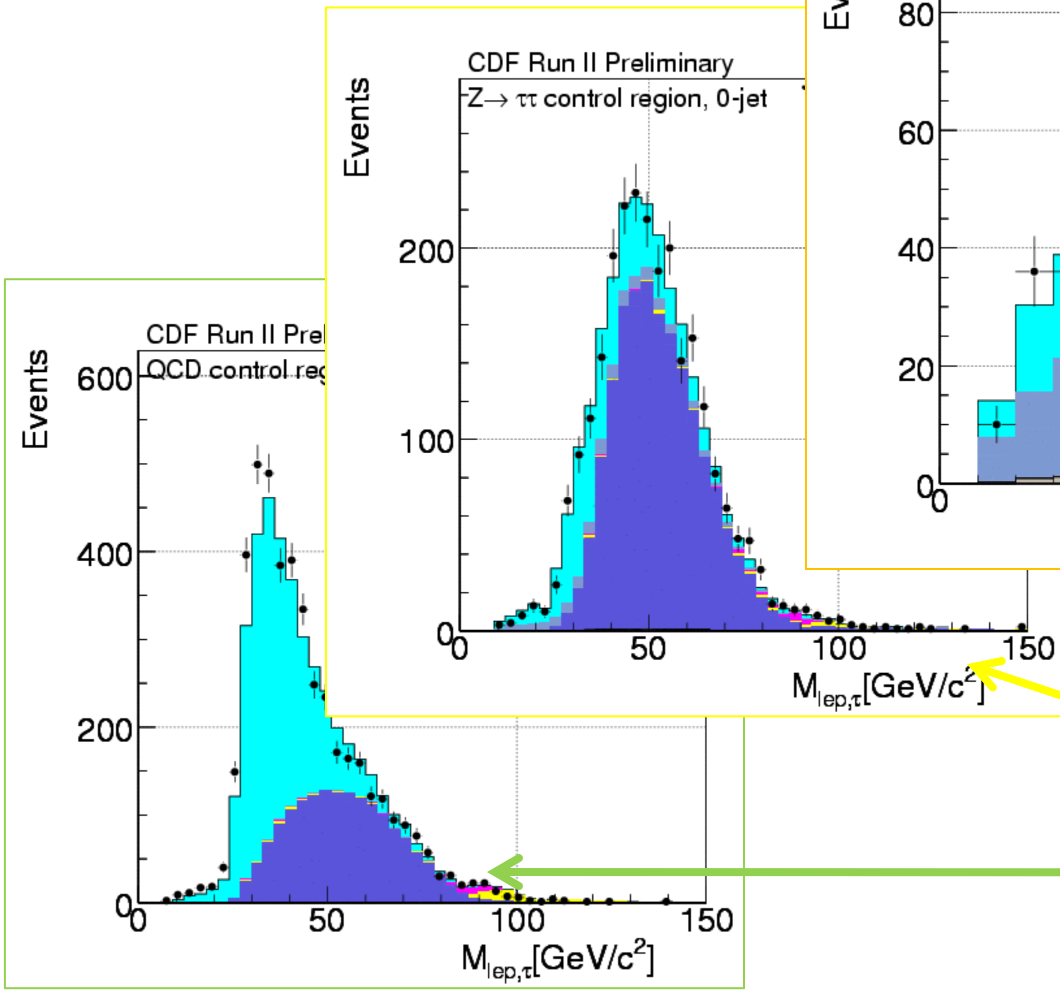
From high-isolation sidebands control regions:

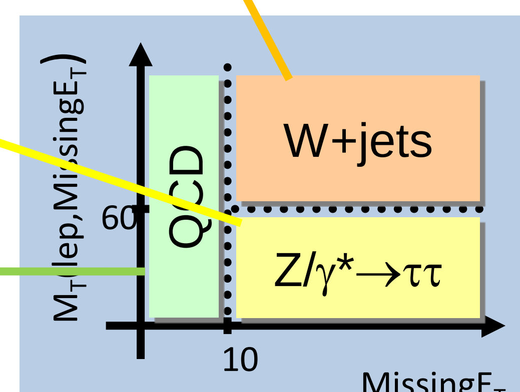
- Normalization factors for SS ~ 1
- Corrections for W+jets OS/SS asymmetries

0-jet control region: background testing

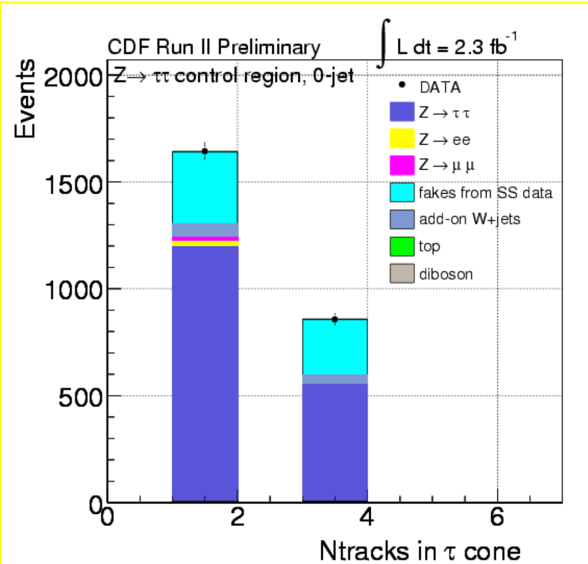
Dilepton invariant mass distribution in 3 control regions

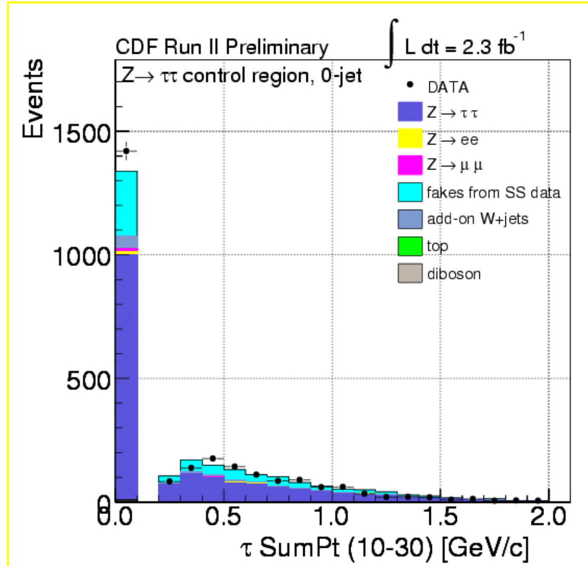


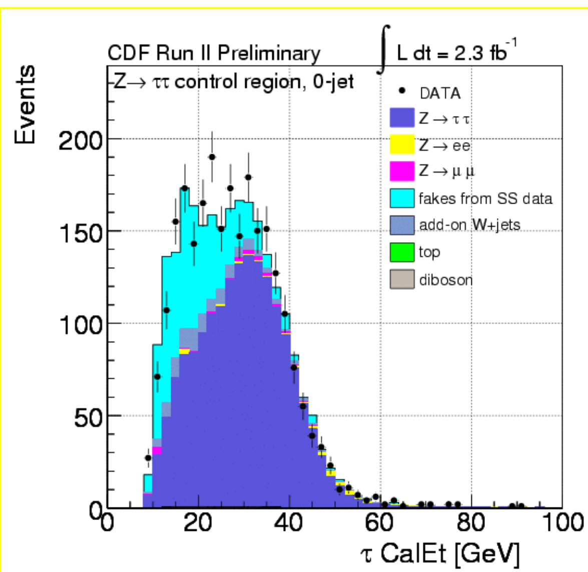


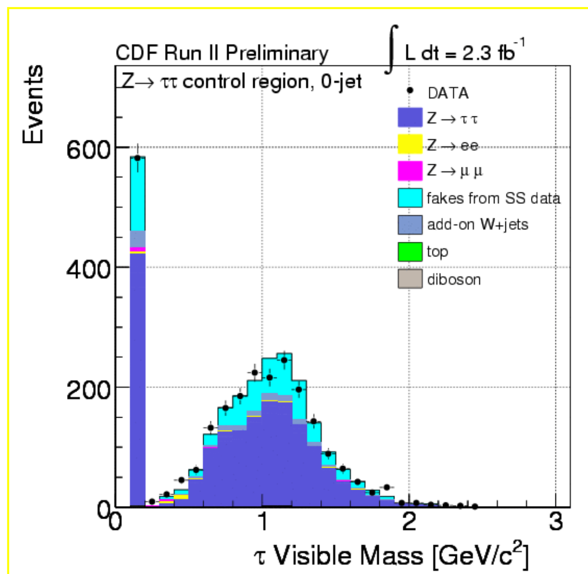


Z→ττ control region: tau ID testing



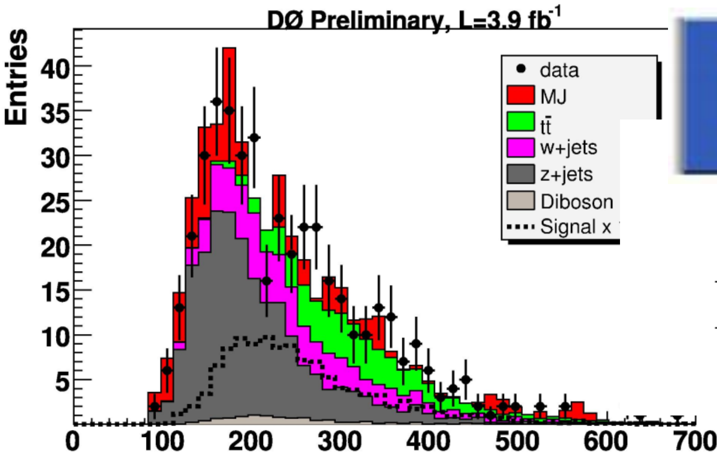




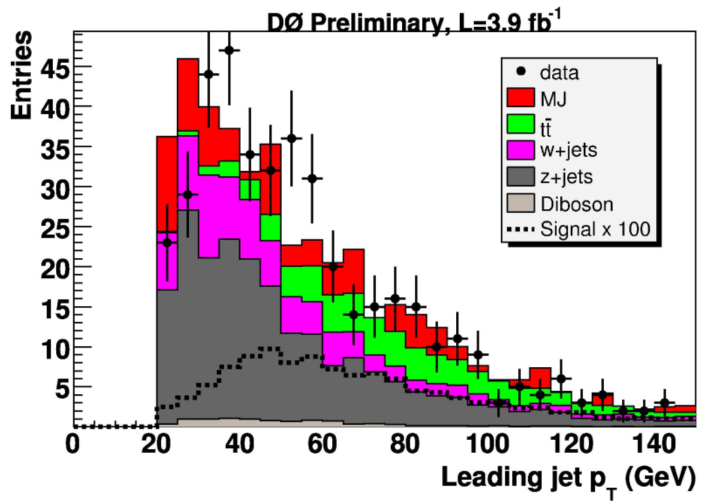


BDT-based tau ID variables

Signal channel: ≥ 2 jets



	Data	Σ Bknd	$t\bar{t}$	W+jets	Z+jets	diboson	multijet
≥2 jets	433	439.9	66.7	81.5	222.7	10.2	80.7



Main background contributions:

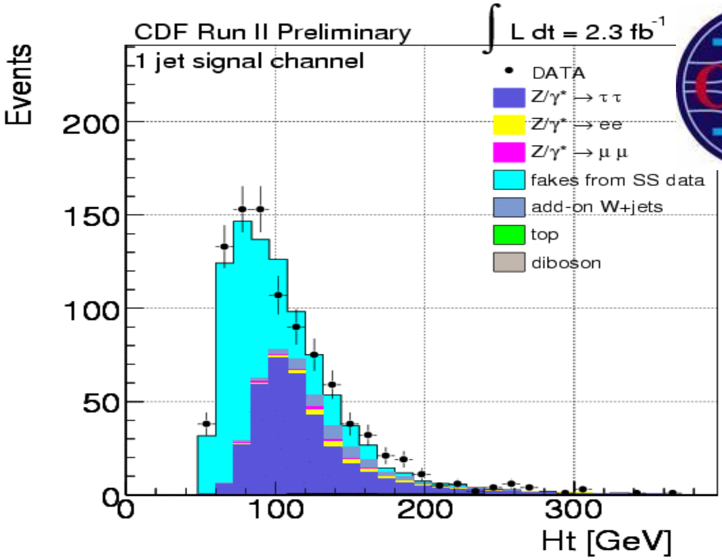
- $Z \to \tau\tau$
- jet→τ fakes in multijet and W+jets

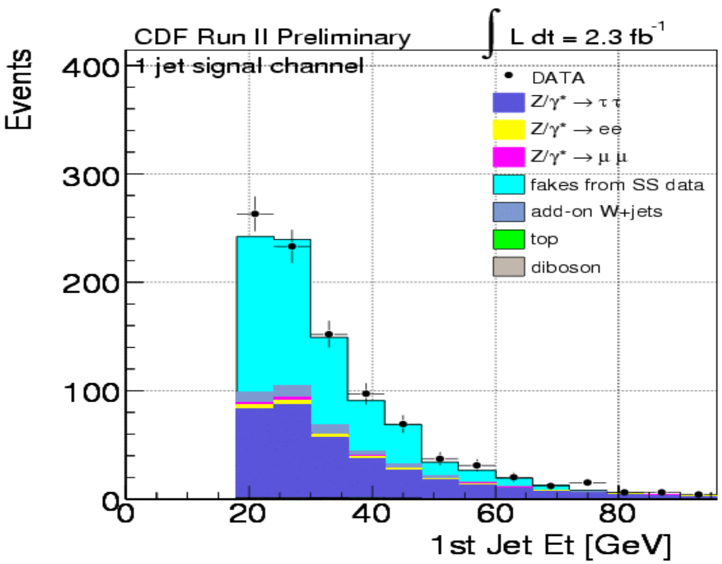
Process	ZH	HZ	HW	VBF	GGF
Event yield	0.11	0.23	0.72	0.12	0.15

- Signal yield: 1.33
- Included also $Z(\to\tau\tau)H(\to qq)$

VBF: Higgs from Vector Boson Fusion
GGH: Higgs from Gluon-Gluon Fusion

Signal channels: 1 jet & ≥ 2 jets





	Data	Σ Bknd	$t\bar{t}$	$Z\rightarrow \tau\tau$	$Z\rightarrow ll$	diboson	jet→ τ fakes
1 jet	965	921.7	4.6	357.9	26.4	3.9	528.8
≥2 jets	166	159.4	16.3	59.3	4.8	0.9	78.1

Main background contributions:

- $Z\rightarrow\tau\tau$

-jet→τ fakes in multijet and W+jets

	HZ	HW	VBF	GGF
1 jet	0.050	0.091	0.070	0.535
≥2 jets	0.099	0.150	0.099	0.129

- Signal yield: 1 jet: 0.746
  ≥2 jets: 0.477

Systematic uncertainties

This search relies on a good jet multiplicity modeling.

Main source of systematics for MC-derived processes:

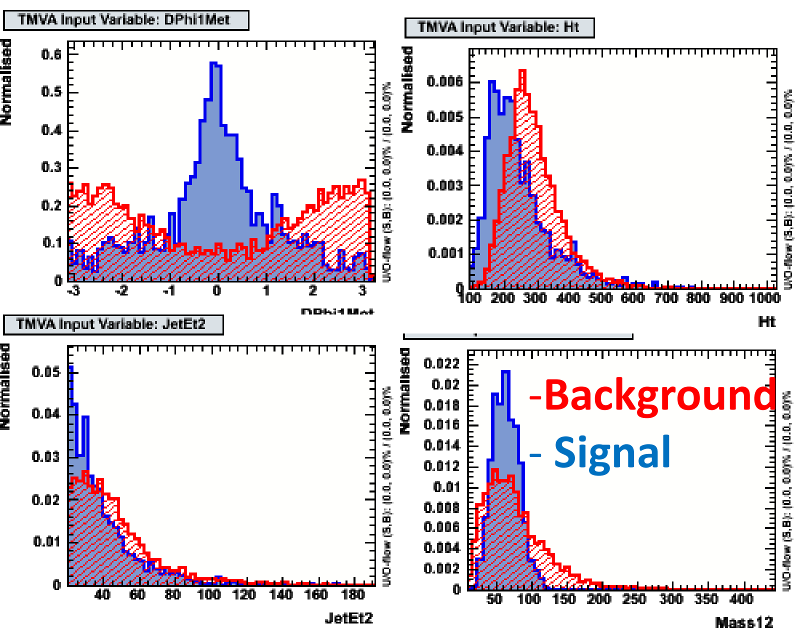
Jet Energy Scale (JES)

Other sources of uncertainty taken into account are:

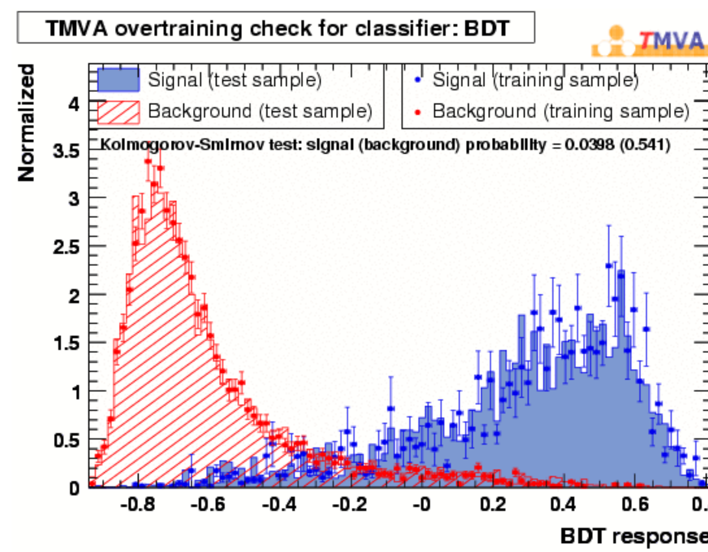
- Cross section and MC acceptances
- Parton Distribution Function (PDF) modeling
- W+JETS and QCD multijet modeling
- Initial State Radiation (ISR)
- Final State Radiation (FSR)
- Tau ID scale factors

Signal vs. Background discrimination

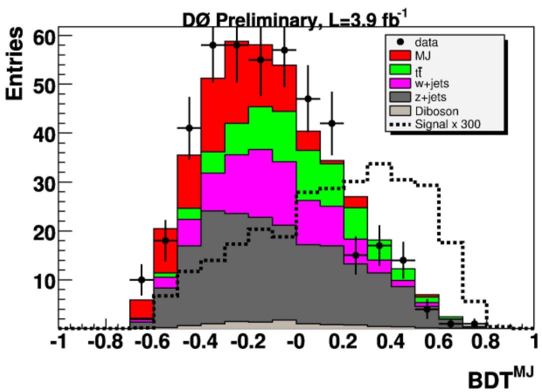
- Good agreement in almost all kinematic distributions
- Expected signal is much smaller than background uncertainties
- S/B is small → counting experiment is not possible.
- Need to exploit all the event information to extract S from B
  Multivariate techniques combine the **discriminating power** of different kinematical and topological distributions

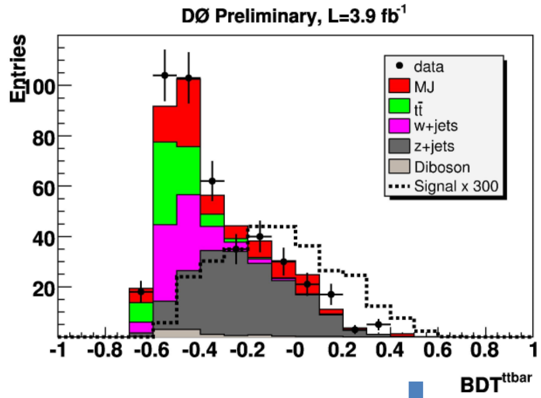


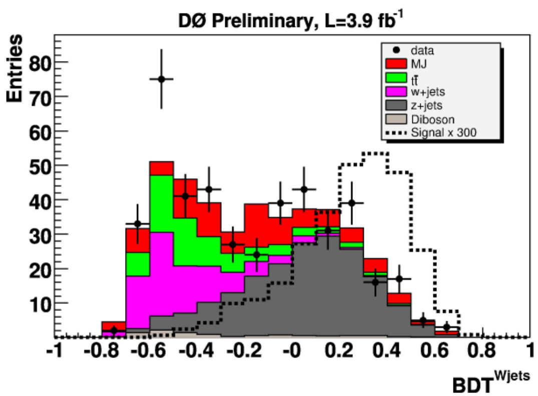
MULTIVARIATE ALGORITHM

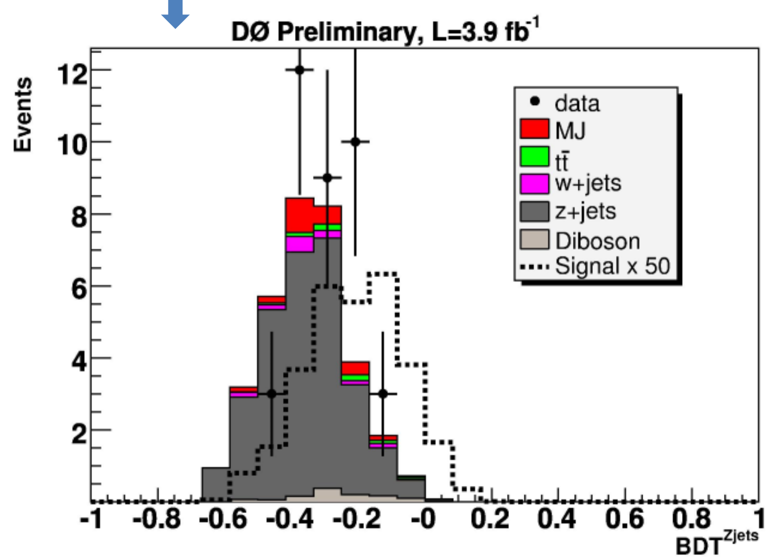


Multivariate discriminants









-Applying a lower cut on:

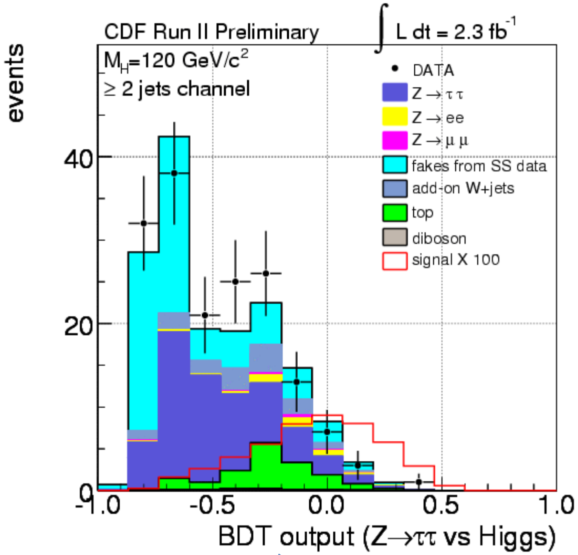
BDT(signal vs. MultiJet),

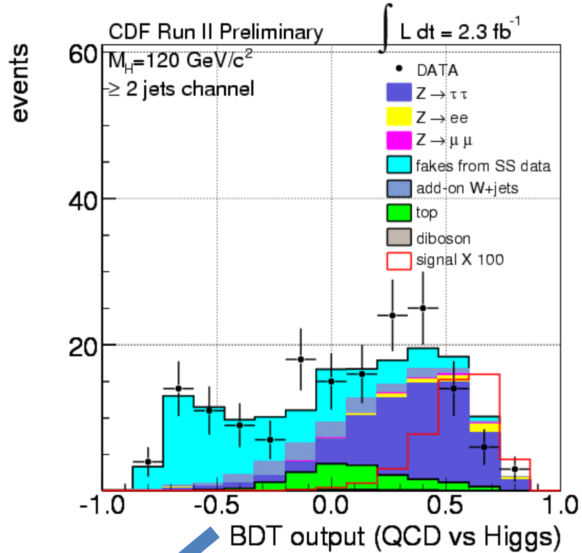
BDT(signal vs. W+JETS),

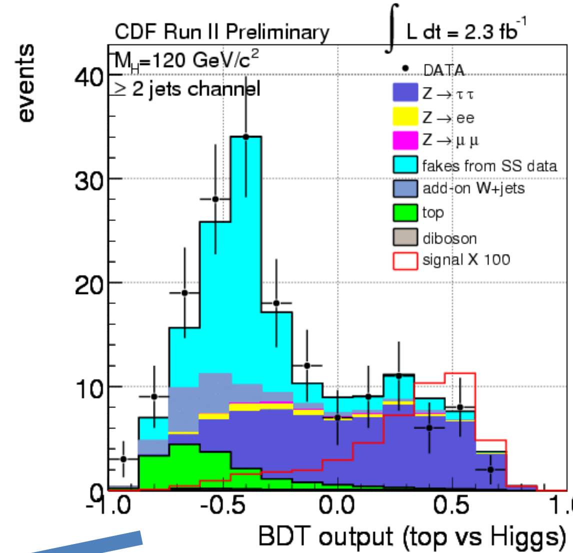
BDT(signal vs. top)

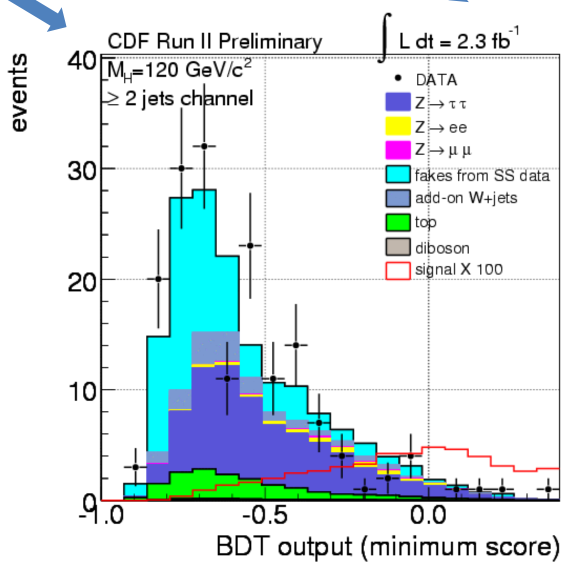
Final discriminant: BDT(signal vs. $Z \rightarrow \tau\tau$)

Multivariate discriminants









-Final discriminant:

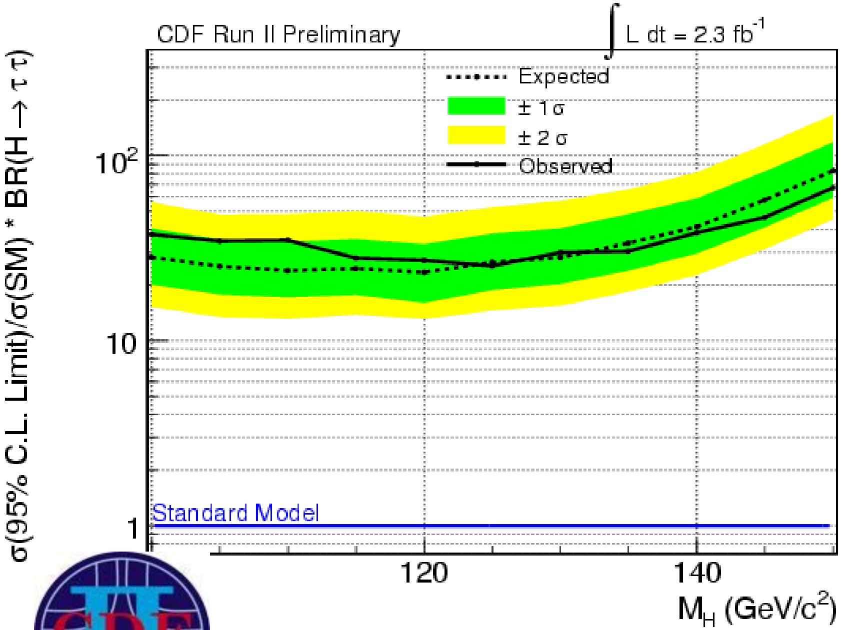
minimum output among

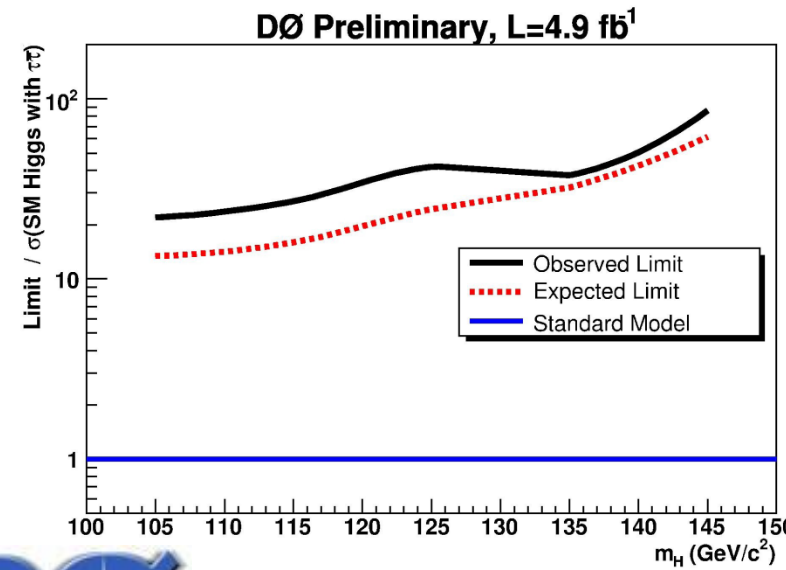
BDT(signal vs. $Z\to\tau\tau$),

BDT(signal vs. multijet)

BDT(signal vs. top)

Results: 95% C.L. upper limit





2.3 fb^{-1}

4.9 fb^{-1}

Mass ranges explored:
100 – 150 GeV/c^2

CDF Expected limits x SM: 23.4 – 82.6
CDF Observed limits x SM: 25.3 – 70.0

Mass ranges explored:
105 – 145 GeV/c^2

DØ Expected limits x SM: 13.4 – 61.4
DØ Observed limits x SM: 21.9 – 86.0

Summary

- Latest results for SM $H \rightarrow \tau\tau$ search at Tevatron presented
    - complementary channel for the low mass region

CDF: 2.3 fb^{-1} exp.(obs.) limit @ M_H = 115 GeV/c^2	24.5(27.9) x SM
DØ: 4.9 fb^{-1} exp.(obs.) limit @ M_H = 115 GeV/c^2	15.9(27.0) x SM

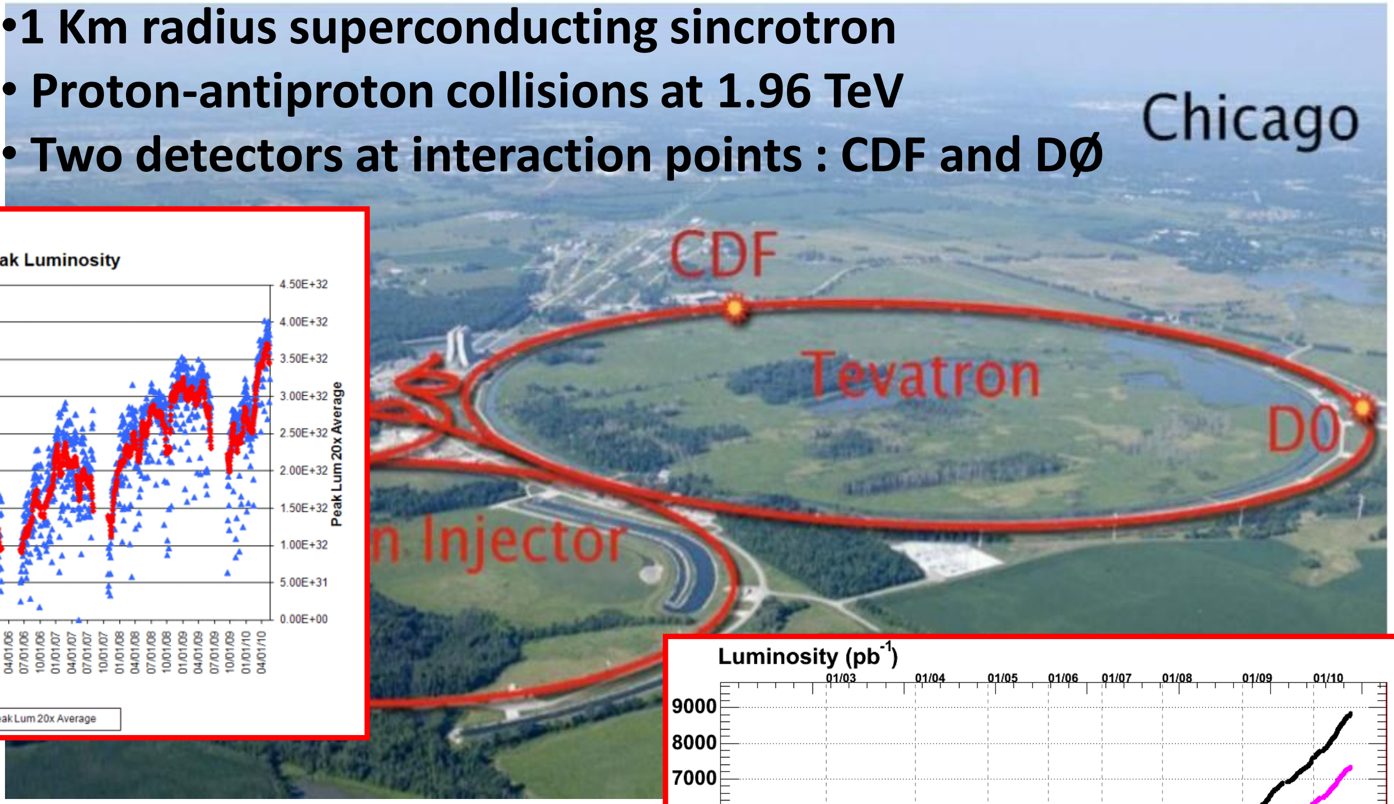
- Many improvements beyond luminosity scaling
    - new tau identification algorithms
    - increased acceptances
    - more sophisticated multivariate method

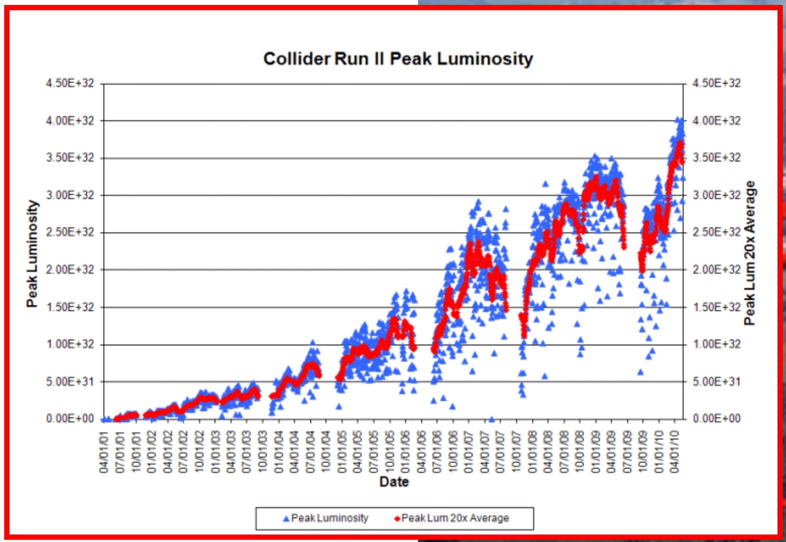
- Now working to add more data and get further improvements!

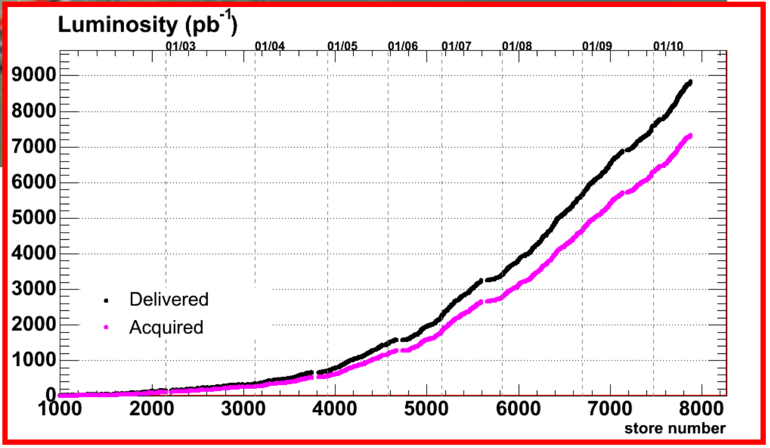
BACK-UP SLIDES

The Tevatron

- 1 Km radius superconducting sincrotron
- Proton-antiproton collisions at 1.96 TeV
- Two detectors at interaction points : CDF and DØ

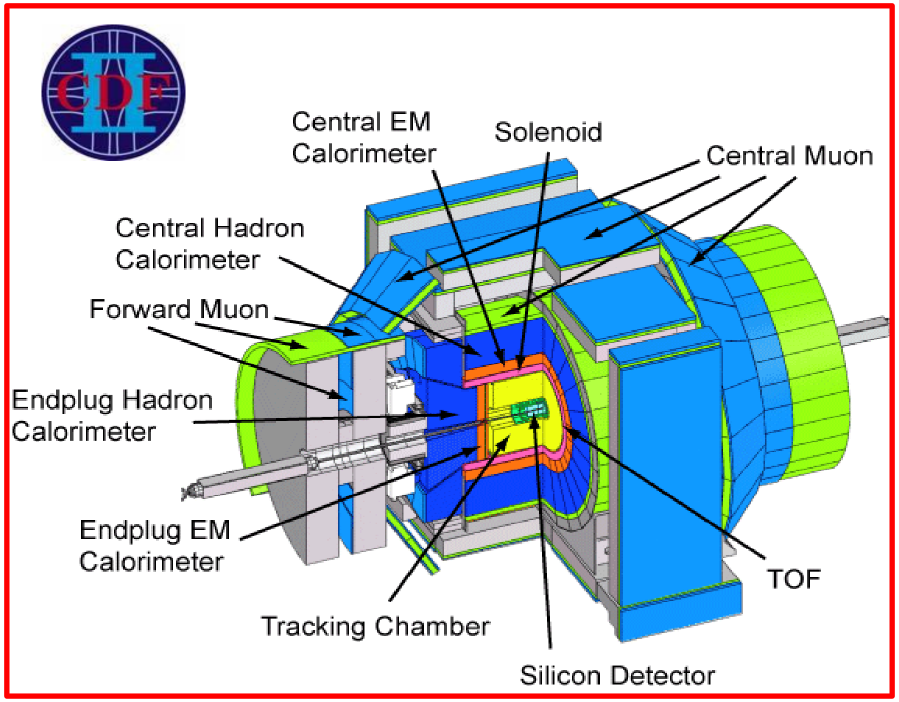






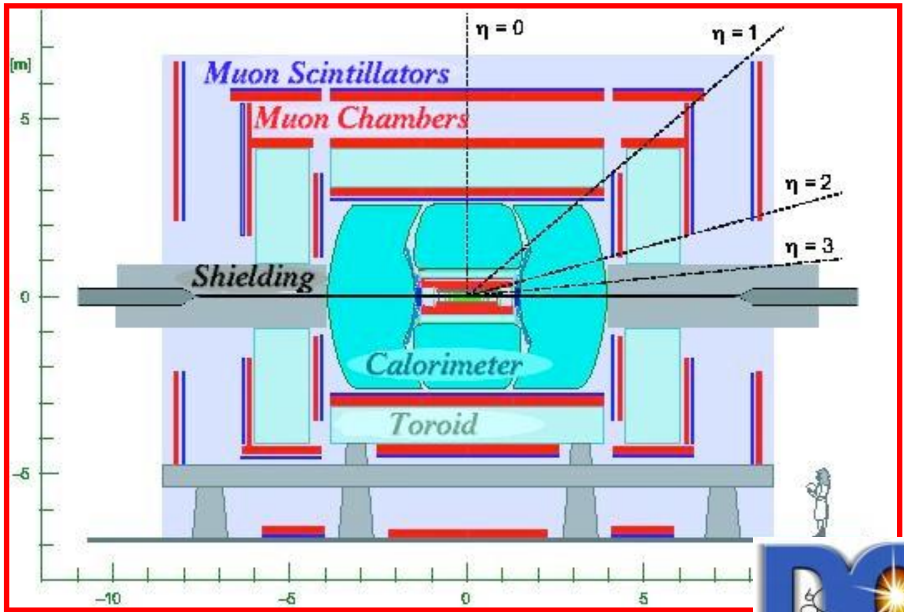
- peak luminosity 4 X 10^{32} $cm^{-2}s^{-1}$;
- weakly integrated lum. ~60 pb^{-1};
-8.8 fb^{-1} delivered per experiment (7.4 fb^{-1} on tape)

CDF and DØ detectors

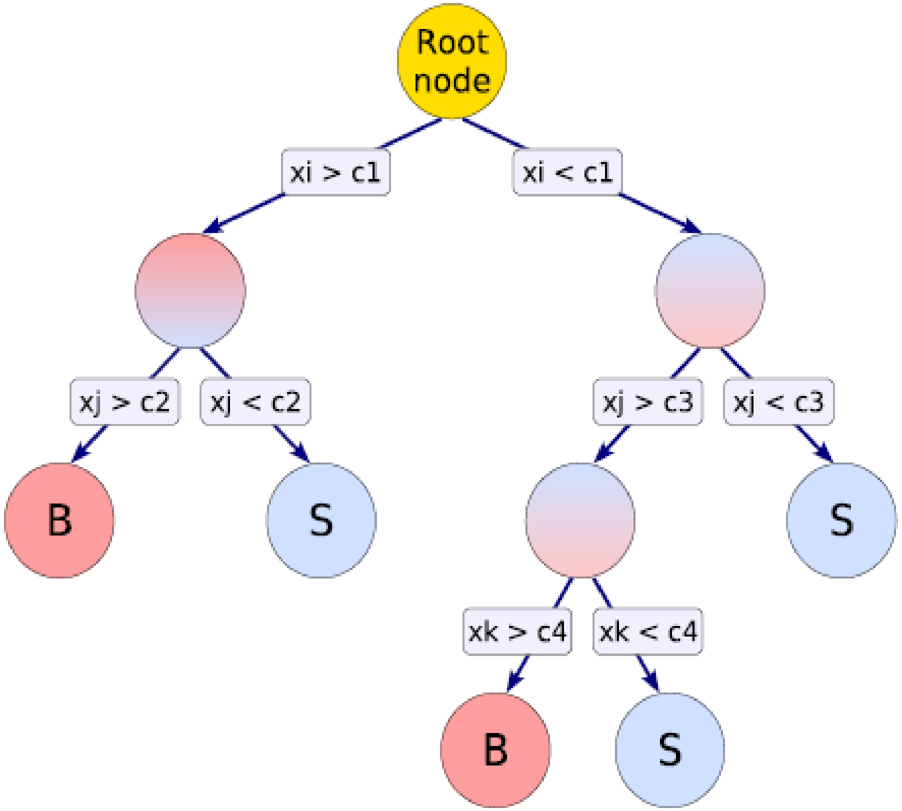


- Silicon Tracking |η|<2-2.5
- Drift cell Tracker 1.4 T, |η|<1.1
- Scintillator Cal. |η|<3.2
- Muons: |η|<1.5

- Silicon tracking |η|<3
- Fiber tracker 1.9 T, |η|<1.7
- LAr/DU calor. |η|<4
- Muons: |η|<2



The Boosted Decision Tree method



A DECISION TREE: a sequence of rooted binary splits

Ingredients : 1) a **training sample** for signal and background

2) a set of **discriminating variables**

At the end of a splitting, leaves are classified as signal-like (event score +1) or background-like (event score -1), accordingly to the purity.

BOOSTING: N trees are created. Events misclassified in the N-th tree, are given an increased weight in the (N+1)th tree.

An event final score is given by the weighted average of different tree outputs

DØ: systematic uncertainties

Source	Uncertainty (%)
Luminosity	6.1
μ ID, track match, iso.	5.0
trigger	5.0
W/Z+light flavor XS	6.0
W/Z+heavy flavor XS	20.0
$t\bar{t}$, single top XS	10.0
diboson XS	7.0
Higgs boson XS	6.0
τ ID NN	8.9
Jet ID/reco eff.	3.0
Jet E resolution.	5.0
JES	7.5
jet p_T	10.0
pdfweight	Shape (currently 3.0)
MJ estimation	17

CDF: systematic uncertainties for background

Systematic uncertainties for the background (%)

Source		$Z/\to ll$	$t\bar{t}$	diboson	fakes from SS	W+jets
JES	(0 jet)	-0.6	-19.0	-0.9		
	(1 jet)	+6.2	-7.7	+7.1		
	($\geq$2 jets)	+14.2	+3.2	+11.7		
Cross section		+2.2	+10.0	+6.0		
PDF		+1.0	+1.0	+1.0		
SS data					+10.0	
W+jets scale	(0 jet)					+5.0
	(1 jet)					+18.0
	($\geq$2 jets)					+30.0
Acc.(DY)		+2.3				
tau ID SF:						
N_{obs}		+2.8	+2.8	+2.8		
N_{SSdata}		-3.3	-3.3	-3.3		
$N_{\mathrm{W+jets}}$		-0.3	-0.3	-0.3		
cross section(DY)		-2.1	-2.1	-2.1		
Acc.(DY)		-2.2	-2.2	-2.2		

CDF: systematic uncertainties for signal

Systematic uncertainties for the signal
1 jet and ≥2 jet channels

Source		ggH	WH	ZH	VBF
JES	(1 jet)	+5.1	-4.8	-5.3	-3.7
	(≥2 jets)	+13.2	5.4	+4.8	-5.2
cross section	(1 jet)	+23.5	+5.0	+5.0	+10.0
	(≥2 jets)	+67.5	+5.0	+5.0	+10.0
PDF		+4.9	+1.2	+0.9	+2.2
ISR	(1 jet)	+13.0	-6.1	-1.7	-2.9
	(≥2 jets)	+15.5	-1.5	+0.1	-2.7
FSR	(1 jet)	-5.0	+4.3	+1.0	+1.7
	(≥2 jets)	-5.2	-2.1	+0.4	-1.1
tau ID SF:					
N_{obs}		+2.8	+2.8	+2.8	+2.8
N_{SSdata}		-3.3	-3.3	-3.3	-3.3
$N_{\mathrm{W+jets}}$		-0.3	-0.3	-0.3	-0.3
cross section(DY)		-2.1	-2.1	-2.1	-2.1
Acc.(DY)		-2.2	-2.2	-2.2	-2.2