Searches for Low Mass Higgs Boson at the Tevatron

Song-Ming Wang
Academia Sinica

On behalf of the CDF and DØ Collaborations

Les Rencontres de Physique
de la Vallée d’Aoste
La Thuile, Aosta Valley, Italy
2010
Outline

• Introduction
  • Why Higgs?
• Higgs Production at Tevatron
• Low Mass Higgs Search Strategy
• Results from Various Low Mass Searches
• Prospects
• Summary
Why Look for a Higgs Boson?

• In Standard Model (SM), Electromagnetism and Weak interactions are unified under the $SU(2) \times U(1)$ gauge symmetry

• However the symmetry in the theory has to be broken, otherwise:
  • All particles will be massless

• In SM, the Electro-Weak symmetry is broken via the Higgs mechanism
  • Every particle obtain mass by interacting with Higgs field through exchange of Higgs boson
  • SM predicts existence of a Higgs boson but its mass is not predicted

• Higgs boson is the only undiscovered particle in SM, making its discovery one of the most important present goal in Particle Physics
What We know of the SM Higgs Boson

- Lower bound from direct searches at LEP:
  - \( m_H > 114 \text{ GeV/c}^2 \)
- Top and W mass measurements are constraining the Higgs sector

- Latest (LEPEWWG Aug ’09) fits to precision Electroweak data
  - \( m_H = 87^{+35}_{-26} \text{ GeV/c}^2 \)
  - \( m_H < 157 \text{ GeV/c}^2 \) (@ 95% CL)
  - \( m_H < 186 \text{ GeV/c}^2 \) (include LEP limit)
Tevatron: Most Powerful $p$-$p\bar{p}$ Collider!
Since start of Run 2 delivered ~8 fb$^{-1}$ per experiment

CDF/D0 : ~6.6-7 fb$^{-1}$

Expected to deliver ~10-12 fb$^{-1}$ by end of 2011
The Tevatron Experiments

Multipurpose detectors:

- Electron, muon, tau identification
- Jet and missing energy measurement
- Heavy-flavor tagging through displaced vertices and soft leptons
Higgs Boson Production at the Tevatron

**SM Higgs production**

- $gg \rightarrow H$
- $qq \rightarrow WH$
- $qq \rightarrow qqH$
- $bb \rightarrow H$
- $qq \rightarrow ZH$
- $gg,qq \rightarrow ttH$

TeVII

$\sigma [fb]$ vs $m_H [GeV]$

TeV4LHC Higgs working group

**Vector Boson Fusion (VBF)**

- $q \rightarrow q^* W, Z$
- $q' \rightarrow q' W, Z$

**Associated (Higgs-Strahlung)**

- $g \rightarrow g^* H$

**gg Fusion**
• Higgs decays predominantly:
  
  • $H \rightarrow b\bar{b}$ \quad (m_H < 135 GeV)
  
  • $H \rightarrow W^+W^-$ \quad (m_H > 135 GeV)
Higgs Search Strategy

• At Tevatron, Higgs production is very rare process

• Difficult to search, but not impossible. CDF/DØ already probing processes with σ~1 pb (WZ, ZZ, single top).

• Search Strategy:
  • Identify Higgs signal by its unique final state signature
  • Increase signal acceptance
  • Advanced discriminating algorithms
  • Search in as many channels as possible (will present results from various channels)
Analysis Techniques : Identify Higgs Signal

• Look for distinctive signature in final state

High Pt Leptons:

• WH → lνbb, ZH → llbb (l : e, μ, τ)
• Identify charged leptons can greatly suppress multi-jet background
• Extend lepton coverage, using leptons not in detector fiducial region, in forward region
Analysis Techniques : Identify Higgs Signal

• Look for distinctive signature in final state

High Pt Leptons:

• $WH \rightarrow l\nu bb$, $ZH \rightarrow ll bb$ ($l: e, \mu, \tau$)

• Identify charged leptons can greatly suppress multi-jet background

• Extend lepton coverage, using leptons not in detector fiducial region, in forward region
**Analysis Techniques : Identify Higgs Signal**

- Look for distinctive signature in final state

**High Pt Leptons:**
- \( WH \rightarrow l\nu bb, \ ZH \rightarrow llbb \ (l : e, \mu, \tau) \)
- Identify charged leptons can greatly suppress multi-jet background
- Extend lepton coverage, using leptons not in detector fiducial region, in forward region

**Heavy-Flavor Jets:**
- Low mass Higgs decays primarily \( H \rightarrow bb \)
- Tag the b-jets by exploiting long life time of B hadrons
- Develop advanced tagging algorithms, increase tagging efficiency and reduce fakes
Analysis Techniques : Identify Higgs Signal

• Look for distinctive signature in final state

High Pt Leptons:
• WH→lνbb, ZH→l_lbb (l : e,μ,τ)
• Identify charged leptons can greatly suppress multi-jet background
• Extend lepton coverage, using leptons not in detector fiducial region, in forward region

Heavy-Flavor Jets:
• Low mass Higgs decays primarily H→bb
• Tag the b-jets by exploiting long life time of B hadrons
• Develop advanced tagging algorithms, increase tagging efficiency and reduce fakes

Large Missing E_T (MET):
• ZH→ννbb, WH→lνbb : MET from ν
• Require large MET can also greatly reduce multi-jet background
Analysis Techniques: Multivariate Discriminant

- Expected Higgs signal too small for counting experiment search
- Single kinematic distribution does not provide sufficient discriminating power
- Exploit every possible information in an event
- Multivariate discriminant tools:
  - Artificial Neural Network (NN)
  - Boosted Decision Tree (BDT)
  - Matrix Element (ME) probabilities
- Multivariate tools have been successfully applied to rare processes:
  - Single top
  - WW,WZ (in lνqq final state)
Results from Various Low Mass Searches
ZH→l⁺l⁻bb (l = e, µ)

- Low signal statistics, but clean channel
- Fully reconstructible final state, 2 resonances
- Event Selection:
  - Select Z candidate decaying into ee or µµ
  - ≥2 jets, with ≥1 b-tag jet

- Jet energy resolution is crucial in constructing \( M(bb) \) mass for signal/background separation
- CDF, DØ use event information (e.g. MET, Z mass, vector transverse momentum of \( Z_{jj} \)) to correct jets’ energy
Multivariate algorithms to separate signal from backgrounds:

- **CDF**: 2D NN (separate $ZH$ vs $t\bar{t}b$, $ZH$ vs $Z+jets$)
- **DØ**: Train BDT to separate $ZH$ from SM backgrounds

Search performed in several sub-channels:

- Tight, loose lepton ID
- Single, double $b$-tag jets

**ZH$\rightarrow l^+l^-bb$ ($l = e, \mu$)**

- Pre-tag
  - 2-tight muons
  - 1 tight, 1 loose muons

CDF Run II Preliminary (4.1 fb$^{-1}$)

DØ Run II Preliminary (3.1 fb$^{-1}$)
ZH→l^+l^-bb (l = e, μ)

CDF Run II Preliminary (4.1 fb^{-1})

<table>
<thead>
<tr>
<th>Mass=115 GeV/c^2</th>
<th>∫L (fb^{-1})</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>4.1</td>
<td>6.8</td>
<td>5.9</td>
</tr>
<tr>
<td>DØ</td>
<td>4.2</td>
<td>8.0</td>
<td>9.1</td>
</tr>
</tbody>
</table>

- All limits will be presented as ratios to the SM prediction σ×BR
WH→lvbb (l = e, µ)

• Clean signature, relatively larger signal statistics

• Event Selection:
  • 1 high Pt e or µ
  • large missing transverse energy
  • ≥2 jets, ≥1 b-tag jet

• Main backgrounds:
  • W+heavy-flavor jets
  • Mistags (light-flavor jets mis-ID as b-jets)
  • QCD multi-jet (jet faking lepton)
  • Top, Dibosons
WH→lvbb (l = e, µ)

To improve search sensitivity:

• CDF: use ME method to calculate an Event Probability Discriminant

• DØ: train NN on several kinematic variables to separate signal from background

<table>
<thead>
<tr>
<th>Mass=115 GeV/c²</th>
<th>∫L (fb⁻¹)</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>4.8</td>
<td>3.8</td>
<td>3.3</td>
</tr>
<tr>
<td>DØ</td>
<td>5.0</td>
<td>5.1</td>
<td>6.9</td>
</tr>
</tbody>
</table>
• Contributions from:
  • $ZH \rightarrow \nu\nu bb$
  • $WH \rightarrow l\nu bb$ \{ $l$ not identified, or identified as hadronic tau \}

• Signature: large MET + heavy-flavor jets

• Large signal statistics but large background from multi-jet processes

• Background:
  • Multi-jet \{ Fake MET due to instrumental effect \}
  • Top, W/Z+HF jets, Dibosons \{ Real MET \}
Missing Transverse Energy + Jets Final State

• Employ Multivariate Discriminant to separate signal from background

<table>
<thead>
<tr>
<th>m=115 GeV/c²</th>
<th>∫L (fb⁻¹)</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>3.6</td>
<td>4.2</td>
<td>6.1</td>
</tr>
<tr>
<td>DØ</td>
<td>5.2</td>
<td>4.6</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Run 248968 Evt 48062268 Fri Jan 23 06:59:26 2009

Leading Jet p_T = 85.6 GeV
Second Jet p_T = 62.3 GeV
DiJetMass = 106.7 GeV
Missing E_T = 128.9 GeV

E_T=62.3 GeV
E_T=85.6 GeV
Mjj=107 GeV
MET=129 GeV

Both jets are b-tagged
qqbb Final State

- **Signal from associated (VH) and vector boson fusion (VBF) productions**
- **Advantage**: Large signal yield, fully reconstructible final state
- **Disadvantage**: Huge multi-jet background
- Event selection: ≥4 jets, 2 b-tagged jets
- Train NN to separate multi-jet from Higgs

Variables used in NN training

<table>
<thead>
<tr>
<th>m=120 GeV/c²</th>
<th>∫L (fb⁻¹)</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>4.0</td>
<td>19.9</td>
<td>10.4</td>
</tr>
</tbody>
</table>
\( \tau + \text{MET} + bb, \, \tau \tau qq \) Final State

\( \tau + \text{MET} + bb \) get contributions from:

- \( \text{WH} \rightarrow \tau \nu bb \)
- \( \text{ZH} \rightarrow \tau \tau bb \) (one \( \tau \) is not identified)

Select events:

- 1 hadronic \( \tau \), MET, \( \geq 2 \) jets, \( \geq 1 \) b-tag
- Limits (\( m = 115 \) GeV, 4 fb\(^{-1}\)):
  - Observe (expect) 14.1 (22.4)

\( \tau \tau qq \) get contributions from:

- \( \text{ZH} \rightarrow \tau \tau qq, \, \text{ZH} \rightarrow q\ell\tau \nu, \, \text{WH} \rightarrow q\ell\tau \nu, \, \text{VBF}(H \rightarrow \tau \tau), \, \text{gg} \rightarrow H+\text{jets} \rightarrow \tau \tau+\text{jets} \)

Select events: 2 \( \tau \) (one hadronic \( \tau \), one decays to \( \mu \nu, \nu \tau \)), \( \geq 2 \) jets
- Limits (\( m = 115 \) GeV, 4.9 fb\(^{-1}\)):
  - Observe (expect) 27.0 (15.9)
\( \gamma\gamma + X \) Final State

- Signal contributions from:
  - \( gg \rightarrow H \rightarrow \gamma\gamma \)
  - WH/ZH, \( H \rightarrow \gamma\gamma \)
  - \( qq \rightarrow qqH \) (VBF), \( H \rightarrow \gamma\gamma \)

- Select events with \( \geq 2 \) photon candidates

- Background:
  - Direct QCD di-photon
  - \( \gamma + \text{jet}, \text{di-jet} : \text{jet fakes as } \gamma \)
  - \( Z/\gamma^* \rightarrow ee, \text{ e mis-identified as } \gamma \)

- Scan in a di-photon mass window to search for a \( H \rightarrow \gamma\gamma \) mass peak

<table>
<thead>
<tr>
<th>( m=120 \text{ GeV/c}^2 )</th>
<th>( \int L \text{ (fb}^{-1}) )</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>5.4</td>
<td>19.4</td>
<td>22.5</td>
</tr>
<tr>
<td>DØ</td>
<td>4.2</td>
<td>17.5</td>
<td>13.1</td>
</tr>
</tbody>
</table>
Combined Search Results From Each Experiment

<table>
<thead>
<tr>
<th>Mass</th>
<th>CDF</th>
<th>DØ</th>
</tr>
</thead>
<tbody>
<tr>
<td>115  GeV/c²</td>
<td>2.38</td>
<td>2.80</td>
</tr>
</tbody>
</table>

- Limits at low mass also receive contributions from high mass H→WW searches (Maiko Takahashi’s talk)
Tevatron Combined Search Results

<table>
<thead>
<tr>
<th>Mass=115 GeV/c²</th>
<th>Expected Limit</th>
<th>Observed Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF+DØ</td>
<td>1.78</td>
<td>2.70</td>
</tr>
</tbody>
</table>

Tevatron Run II Preliminary, L=2.0-5.4 fb⁻¹

95% CL Limit/SM

LEP Exclusion

Expected ±1σ Expected ±2σ Expected

Observed

m_H (GeV/c²)

SM=1

November 6, 2009
Higgs Boson Search Projection

- Improvements are faster than $1/\sqrt{fLdt}$ (gain from increasing data size)
  - Due to better analysis techniques
- Band indicates possible improvements.
- Tevatron may exclude Higgs (@ 115 GeV) with $fLdt \sim 6-10$ fb$^{-1}$
- With 10 fb$^{-1}$ for each experiment and including all improvement techniques:
  - ~35% chance to observe 3σ evidence in low mass region
Summary

• CDF and DØ are pursuing extensive direct searches for the SM Higgs boson
• Exploring all possible search channels for low mass Higgs boson
• Combined Tevatron sensitivity is below $2 \times \text{SM}$ (for $M(H)=115 \text{ GeV}/c^2$)
• Tevatron is expected to deliver $\sim 10-12 \text{ fb}^{-1}$ by end of 2011
• May have a chance to find evidence of Higgs in the low mass region if it exists there
Back-UP
Tevatron Combined Search Results
Experiments’ Performance

Data Taking Efficiency

- Good Store Eff
- 20 Store Ave (Acquired)
- 20 Store Ave (Good)
- 20 Store Ave (SVX)

Daily Data Taking Efficiency

- Daily Efficiency
- 10 Day Average
- 30 Day Average

19 April 2002 - 7 February 2010
Running through FY10 will yield 7 fb⁻¹ of data for analysis.

Running through FY11 would yield 9-10 fb⁻¹ of data for analysis.

Real data for FY02-FY09
Higgs Boson Search Projection

- With 10 fb\(^{-1}\) for each experiment and including all improvement techniques:
  - ~70% chance to observe 2\(\sigma\) excess in low mass region
  - ~35% chance to observe 3\(\sigma\) evidence in low mass region
Matrix Element Discriminant

- Matrix Element (ME) probability: probability that an observed event is from a particular physics process
  
  \[ P(x_{\text{obs}} | \text{WH}) \]: probability of event from WH production

  \[ x_{\text{obs}} \]: measured quantities of the event

  \[
P(x_{\text{obs}}) = \frac{1}{<\sigma>} \int \frac{d\sigma(y)}{dy} \varepsilon(y) G(x_{\text{obs}}, y) dy
\]

  \[ y \]: true values of the observables

  \[ d\sigma(y)/dy \]: parton level differential cross section

  \[ \varepsilon(y) \]: detector acceptance & efficiency

  \[ 1/<\sigma> \]: normalization constant

  \[ G(x_{\text{obs}}, y) \]: transfer function representing the detector resolution

- The true values of the observables are unknown, therefore we need to integrate over them

- Use probability values to compute likelihood:

  \[
  L = \frac{P_S(x_{\text{obs}})}{P_S(x_{\text{obs}}) + \sum_i f_i \cdot P_i(x_{\text{obs}})}
  \]

  \[ P_S(x_{\text{obs}}) \]: probability of event to be signal

  \[ P_i(x_{\text{obs}}) \]: probability of event to be background \( i \)th

  \[ f_i \]: fraction from background \( i \)th
**Boosted Decision Tree**

- Training a decision tree:
  - Events of equal weight pass through a cascade of cuts
  - Eventually events land on signal (S) or background (B) leaf node
  - Events landed in wrong leaf node (e.g. signal event land in BG node) are given larger weights (boosted)
  - Tree is retrained, may build up 100-1000 trees
- Trained decision trees are used to classify an event:
  - A discriminant value is given to an event based on the weighted sum of all the trees (weight is the boost-weight)