# WH→lvbb using DLM and Discriminant Analysis at CDF

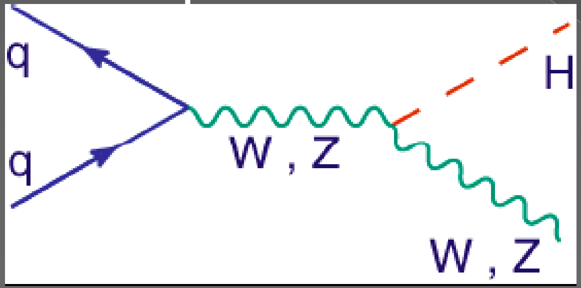
Masakazu Kurata, Shinhong Kim
University of Tsukuba
Kunitaka Kondo
Waseda University

On behalf of the CDF Collaboration

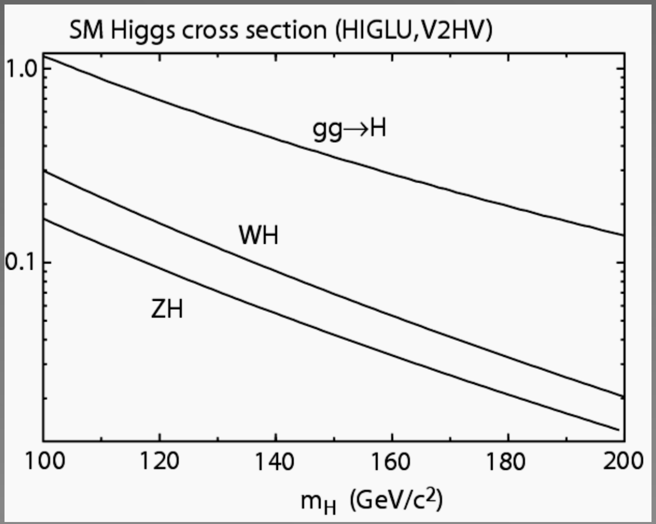
フレーバー物理の新展開研究会, 07/06/2012

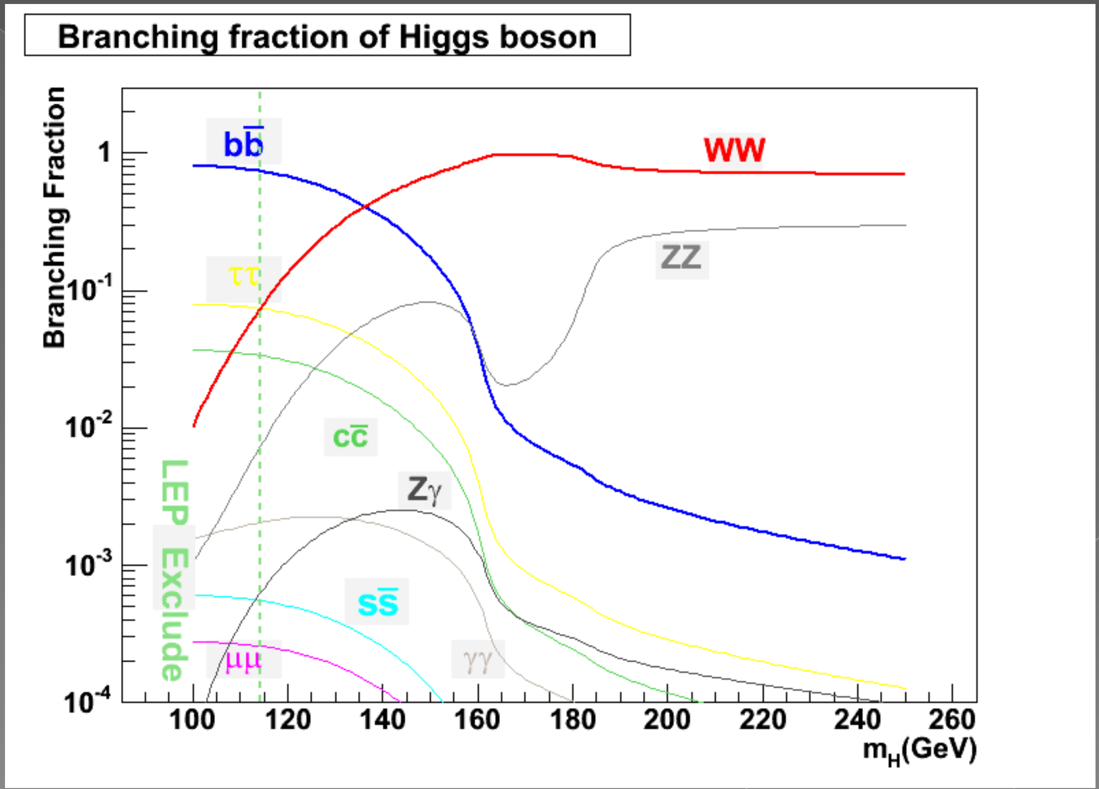
# Introduction

- Higgs associate W→WH→lνbb process is a strong channel at Tevatron
  - It is a golden channel for low mass($m_H$<135GeV) because Higgs dominantly decays into b quark pair(b-tagging is >50% efficient)
  - By requiring one lepton, large QCD background can be suppressed
  - The process is distinct from other backgrounds




Higgs associated



# WH analysis and DLM

- At CDF, previous techniques used to search for WH process:
  - Neural Network
  - Matrix Element
- Main challenge: develop a technique to separate signal from backgrounds
  - We developed a new technique for extracting signal information
  - Establish Dynamical Likelihood Method(DLM)
  - We calculate a discriminant based on DLM to evaluate upper limit on cross section for WH
- DLM was used for top analysis so far
  - Top mass measurement
  - ttbar resonance search

# Event selection and data

- Event selection is based on the standard criteria of WH analysis at CDF
  - Use the events with central lepton or plug electron
    good quality Isolated track
  - Selection cuts: **2jets + 1 lepton + large MET**
    - Jet: 2jets with Et>20GeV and $|\eta_{det}|<2.0$
    - Lepton: Et>20 GeV and require tight lepton selection
      or Isolated track Selection
    - Missing Et>20GeV (for central, Isolated Track), 25 GeV (for plug)
  - b-tag to reject large background events
    - 3 b-tag categories: Double tag(2 categories) and Single tag
- Data and Monte Carlo
  - Data: evaluate expected sensitivity for 7.5 $fb^{-1}$
  - MC: Pythia or MadGraph+Pythia (for EWK)
    Alpgen+Pythia (for W+jets)

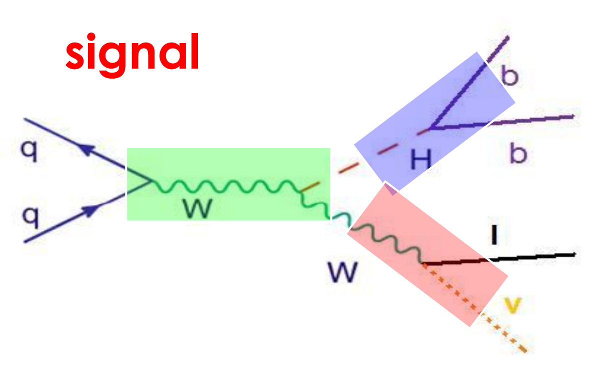
# Dynamical Likelihood Method

- We use DLM method to separate signal and Backgrounds
- DLM is a method to extract signal events from data by evaluating matrix elements as likelihood function.
  - DLM is applicable to processes for which matrix elements can be calculated theoretically : WH, Single top, Wbb
- Formulation:

Likelihood function of DLM:

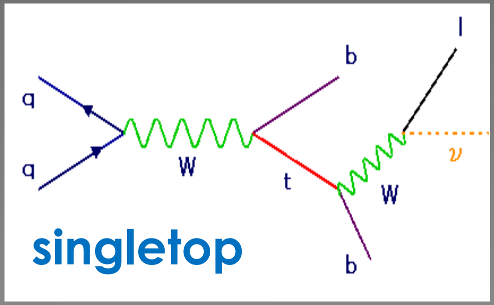
$$L_{path}(\alpha, x \mid y) = N \frac{d\sigma}{d\Phi} w(x \mid y)$$

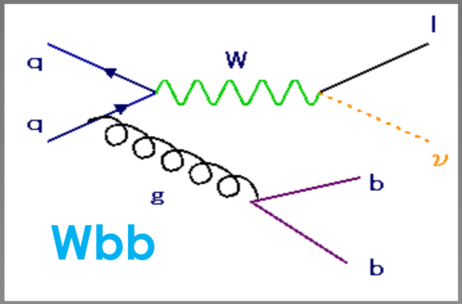
$$\overline{L} = \frac{1}{n_{path}} \sum_{k}^{n_{path}} L_{path}^{(k)}(\alpha, x \mid y)$$




w(x|y): transfer function
⇒ a probability density function for x (parton momenta) when y (observed quantities) is given.

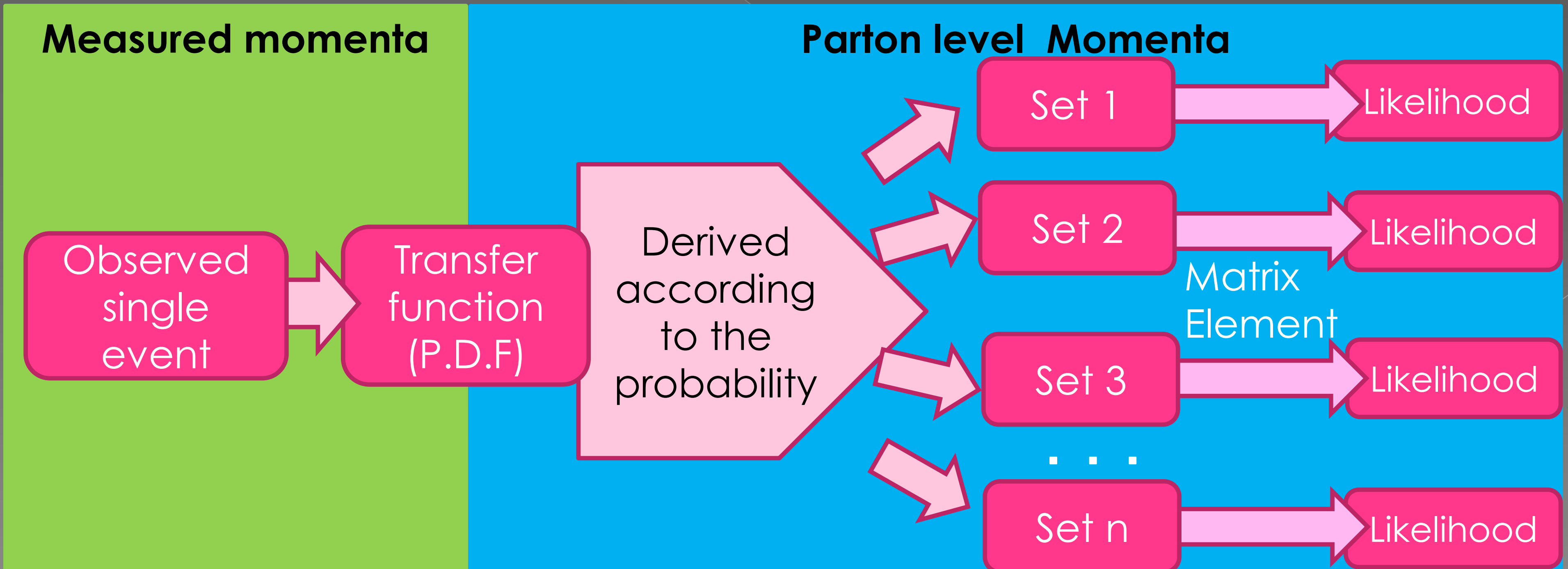
Relationship between parton level and detector measurement



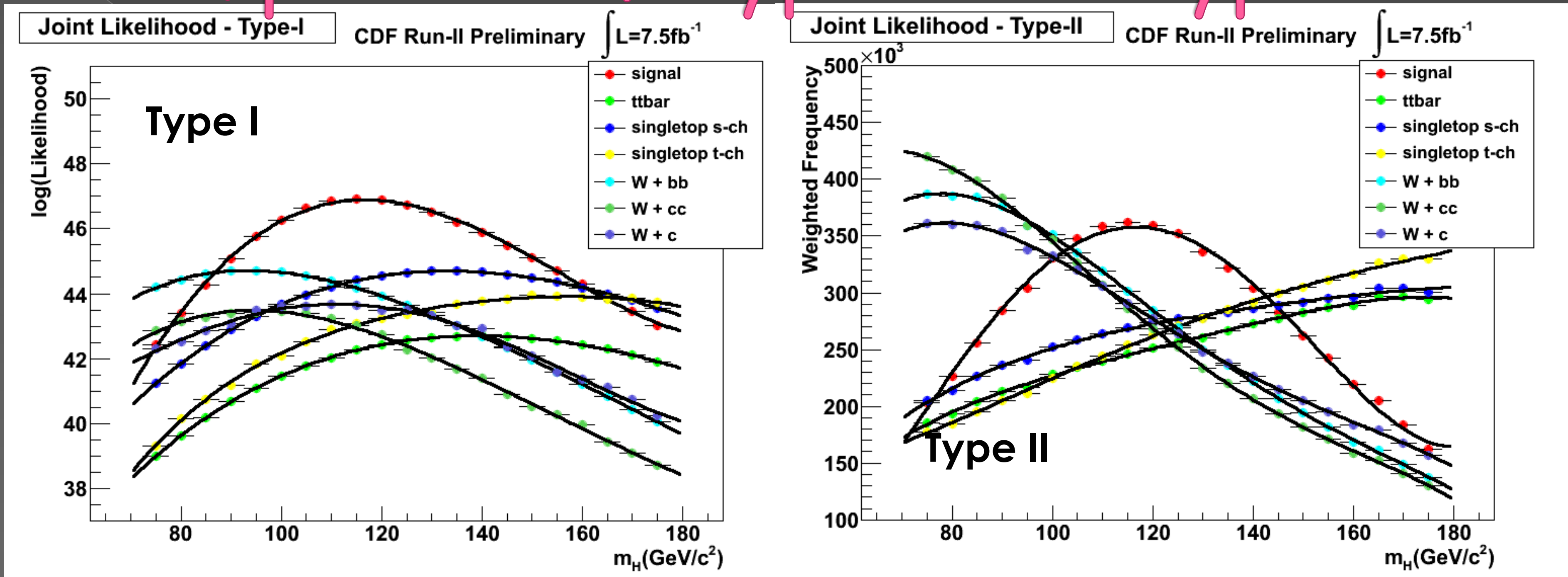



# Summary of DLM

- Basic idea of DLM
  - Given the single event, Parton level Momenta set can be obtained randomly according to the probability of Transfer Functions(P.D.F)
  - We can calculate Matrix Element using parton level momenta set.
  - We can treat statistically by accumulating the results of many parton level momenta sets
  - Likelihood of each momenta set is given by $L_{path}(\alpha, x \mid y) = N \frac{d\sigma}{d\Phi} w(x \mid y)$



# Comparison of Type I and Type II




Mean value of Event ensemble O(1000)

$$L_{joint} = \frac{1}{n_{event}} \sum_{i}^{n_{event}} \overline{L_{eve}} \;\; @each\, Higgs\, mass$$

For the better discrimination of signal and backgrounds,

- Shape information should be obtained from Type II
  - Because the signal shape is quite different from backgrounds
- Absolute value information should be obtained from Type I
  - Basically, signal likelihood is higher than backgrounds at whole Higgs mass range

seem good

# Input variable candidates for MVA

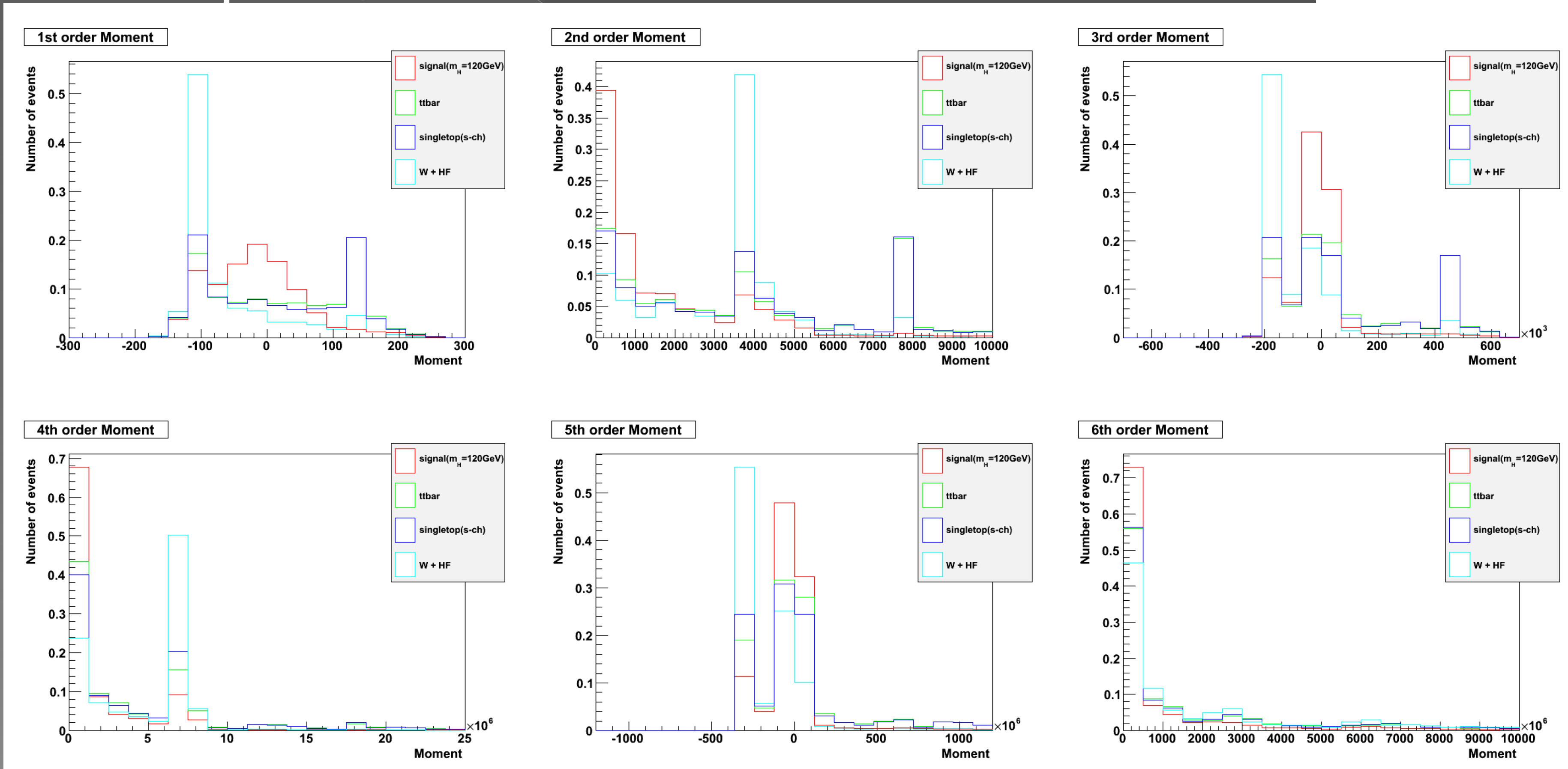
- **Extract the signal feature with 2 points**
  - Higgs mass dependence shape
    - Evaluate Localized Moment – use up to $6^{th}$ order

      $\int (m_H - m_{H0})^n \cdot f_{event}(m_H) dm_H$ nth-order moment
      - As $m_{H0}$, expected mass of signal from DLM is used
  - Absolute value of likelihood
    - Several kinds of likelihood can be obtained from DLM result with signal Matrix Element
      - Maximum likelihood on Higgs mass dependence
      - Likelihood@ DLM expected Higgs mass
      - Higgs part likelihood @ maximum likelihood of overall Matrix Element
      - W part liielihood @ maximum likelihood of overall Matrix element
    - Likelihoods as a result of DLM with Wbb & singletop(s-ch) Matrix Element
      - All the parameters(mass, width, etc) in the Matrix Element are set

# Localized moment around signal

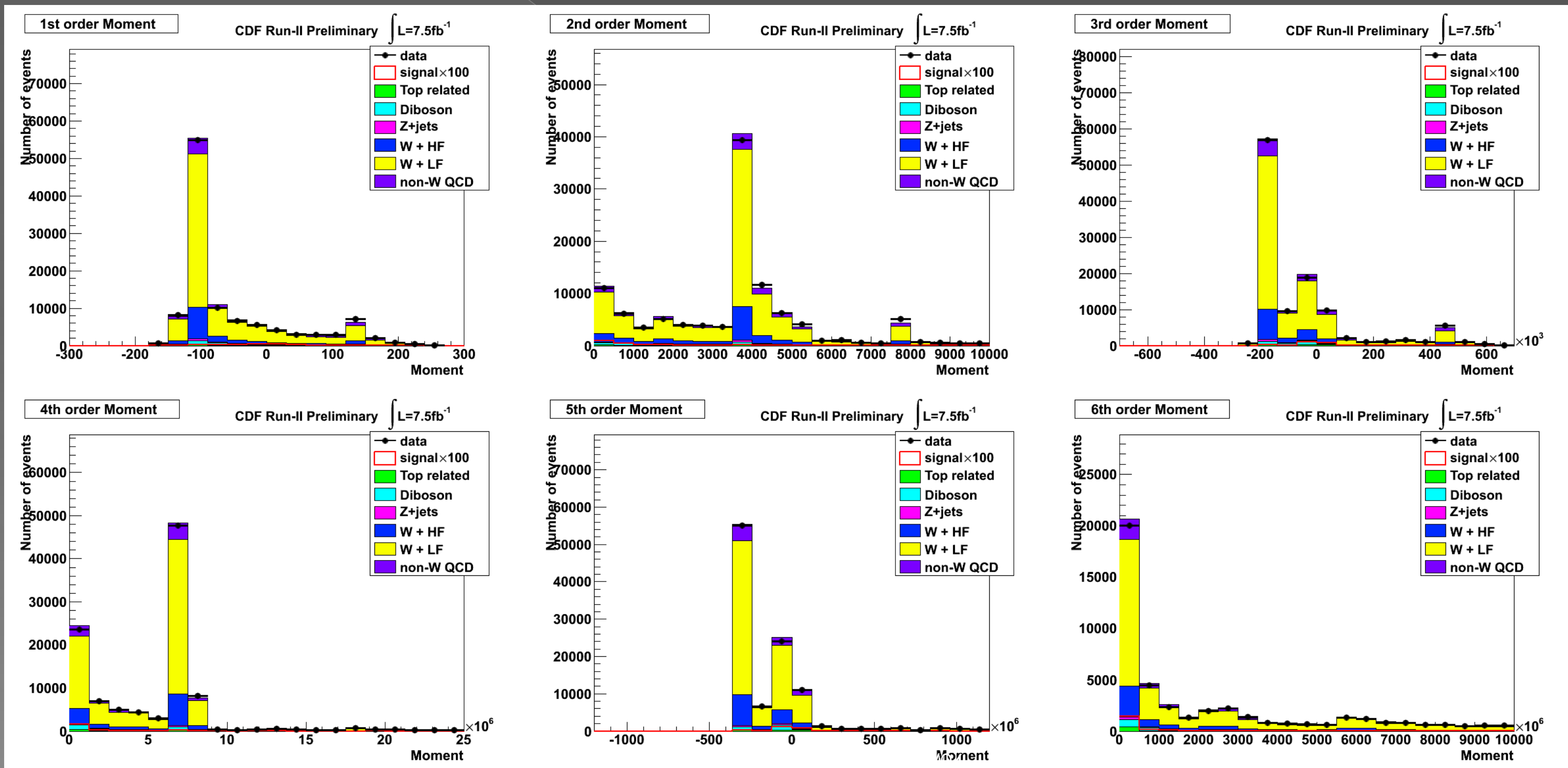
- Localized moment around signal expected mass for each process
  - $\int (m_H - m_{H0})^n \cdot f_{event}(m_H) dm_H$ nth-order moment
  - Up to 6$^{th}$ order

Signal
ttbar
Stop(s-ch)
Wbb



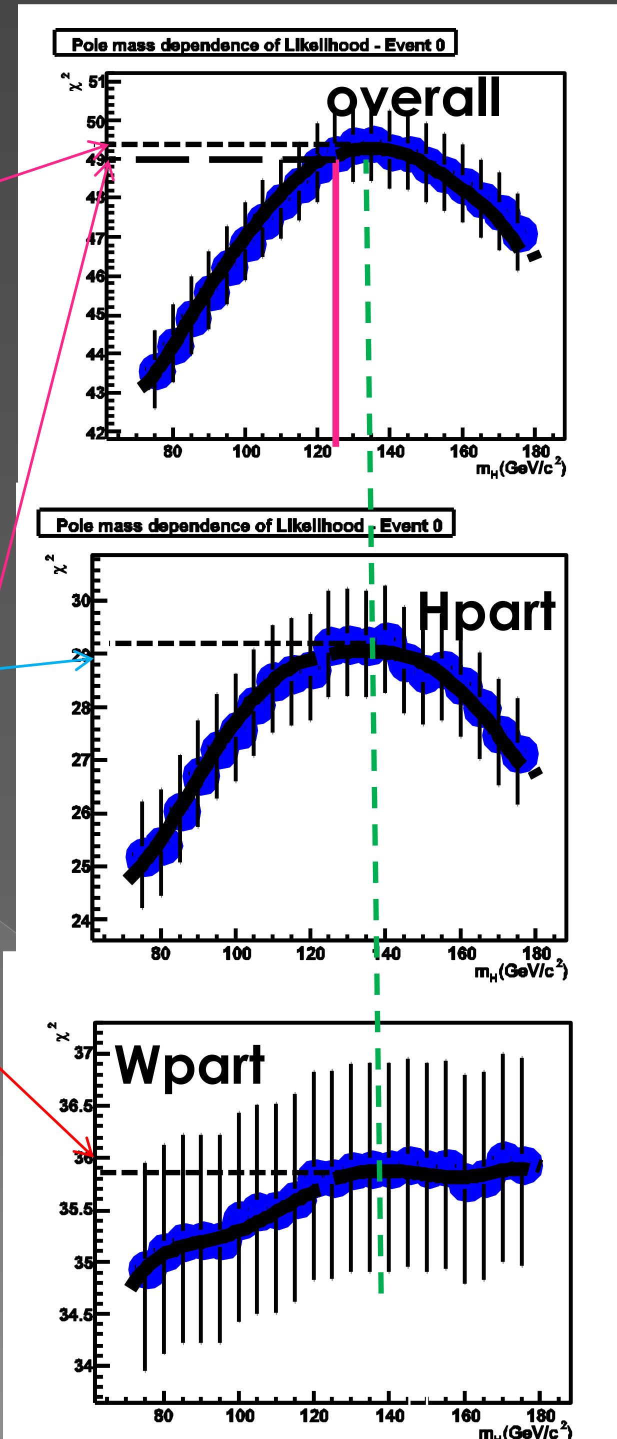
# Localized moment

- Distribution check with pretag events
  - Up to 6$^{th}$ order
  - 3$^{rd}$, 4$^{th}$ and 5$^{th}$ order moments have strong separation power
    - These moments are adopted as the input variables by most of the discriminants



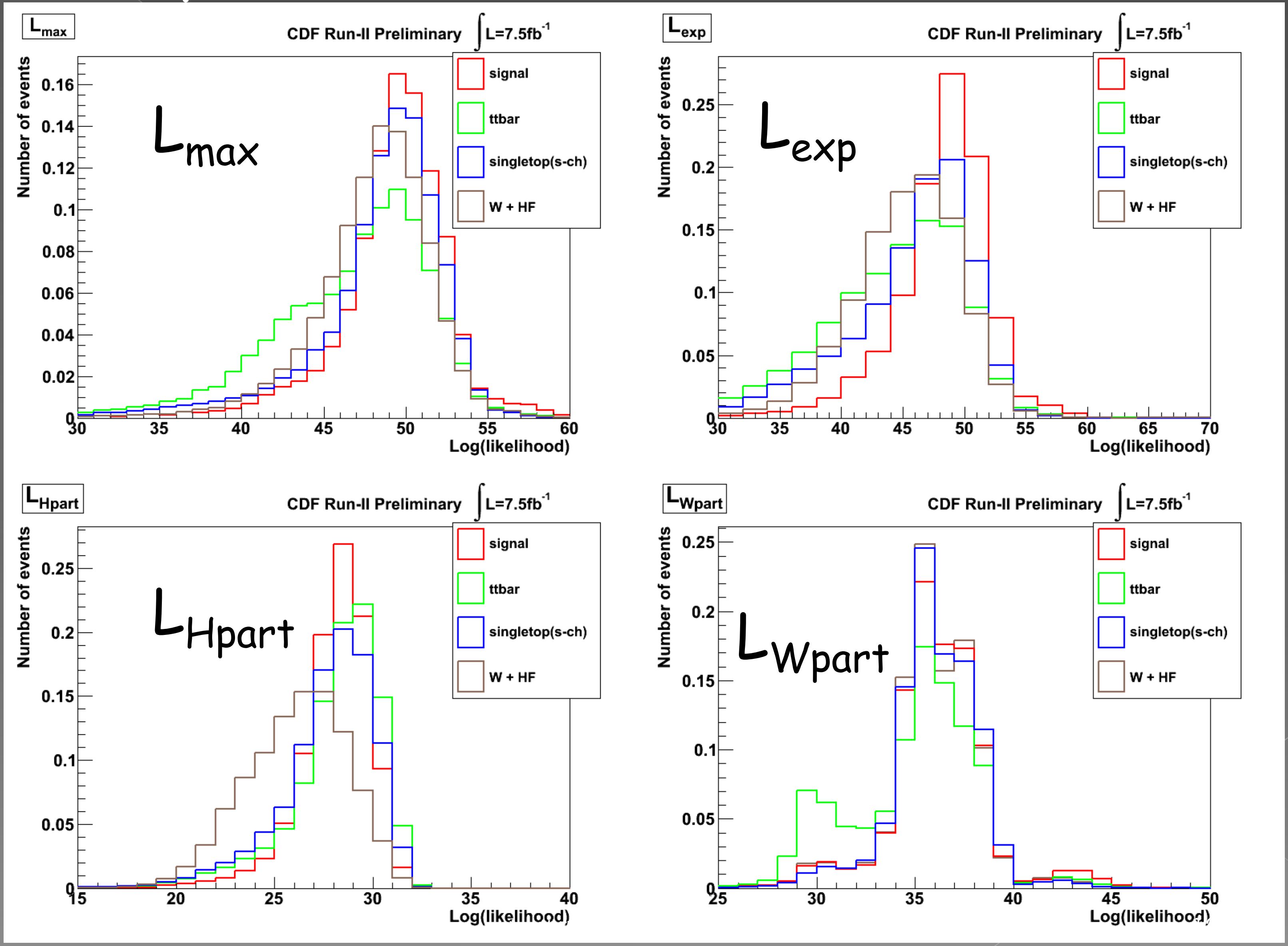
# Likelihood used

- Maximum likelihood: $L_{max}$
  - Use them event-by-event
- Expectation value of likelihood
  - Wbb : $L_{Wbb}$
  - singletop(s-ch): $L_{stop}$
- Higgs part likelihood : $L_{Hpart}$ @maximum likelihood in overall
  - vertex of Higgs decay & T.F.
- W part likelihood: $L_{Wpart}$ @maximum likelihood in overall
  - vertex of W decay & T.F.
- Likelihood @ DLM expected mass: $L_{exp}$
  - Expected mass means the result of event ensemble



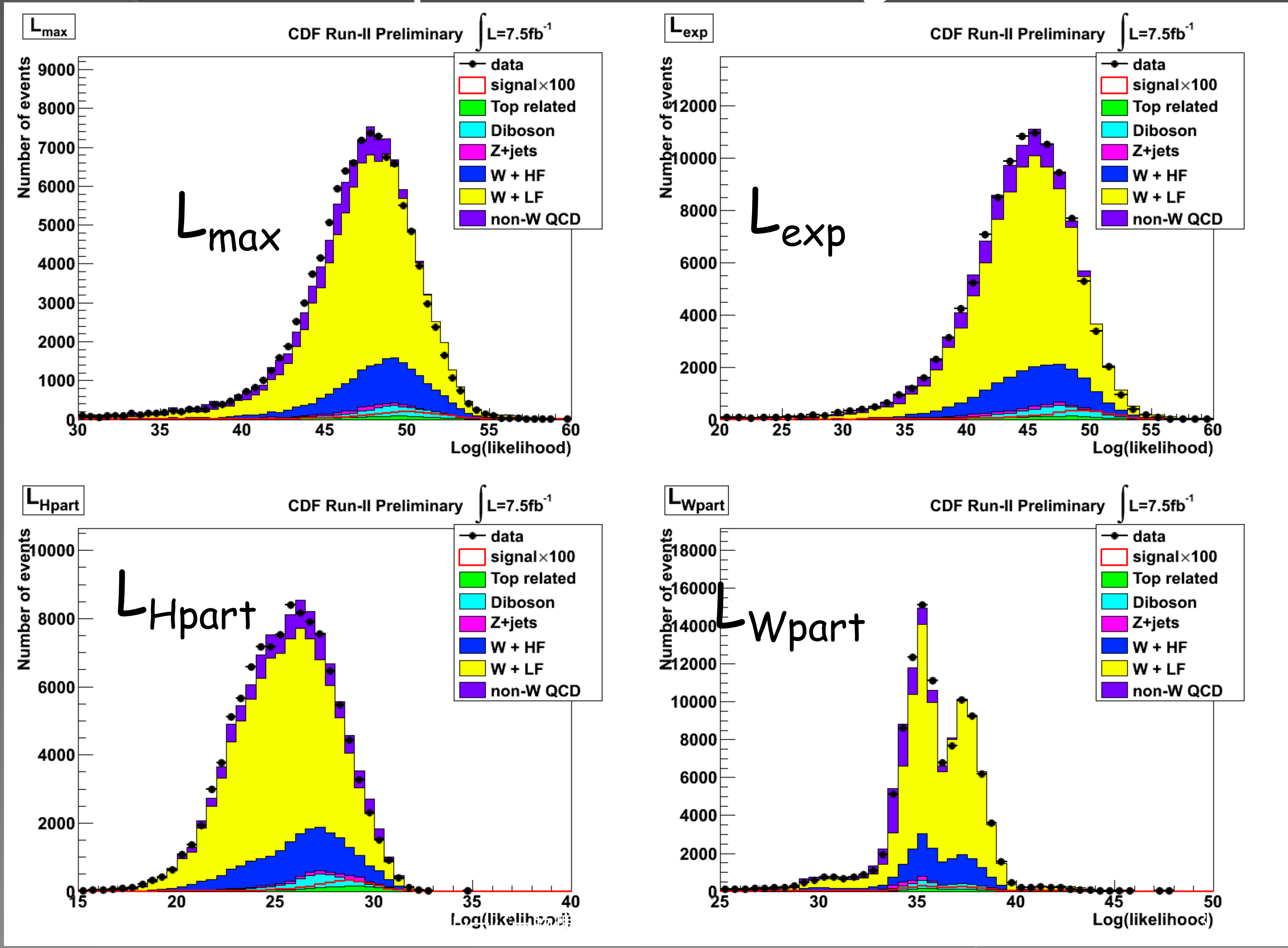
# Likelihood – from signal M.E

- Comparison of the distribution



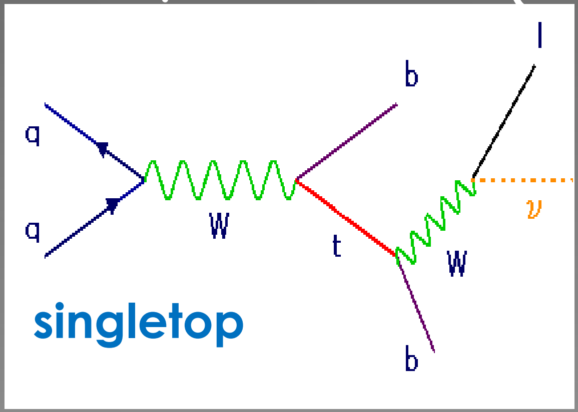
# Likelihoods – from signal M.E

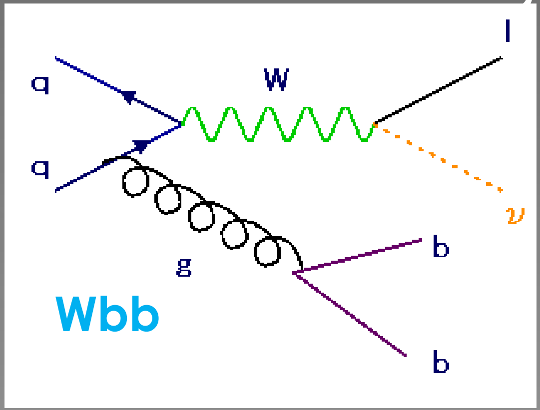
- For the sample before b-tag



# DLM with Singletop & Wbb bases

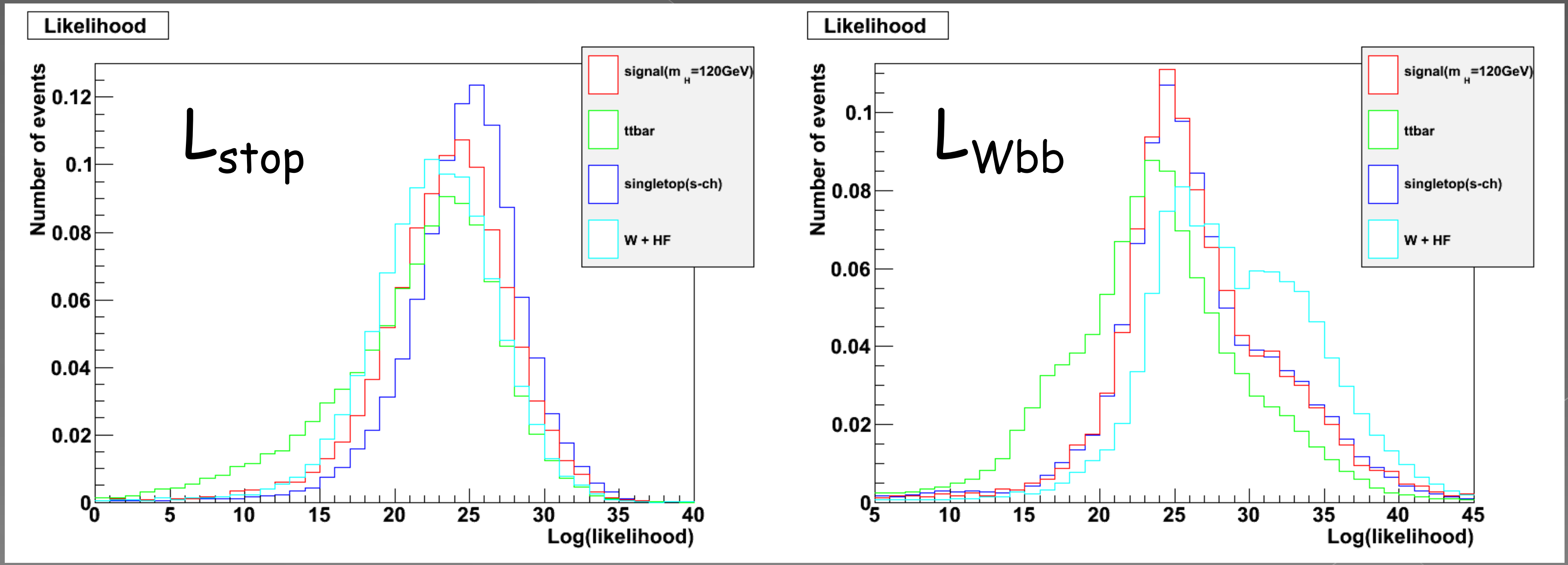
- To improve performance, DLM is used with background matrix elements:
  - Singletop (sch, leading order)
  - Wbb(leading order)
- Optimize transfer functions for each background process using MC
- Same procedure when performing DLM with signal ME($1.0 \times 10^6$paths)
- Background processes don't have any dependence of certain variable. So they only have the expectation value of likelihood ( this means the integration of phase space)



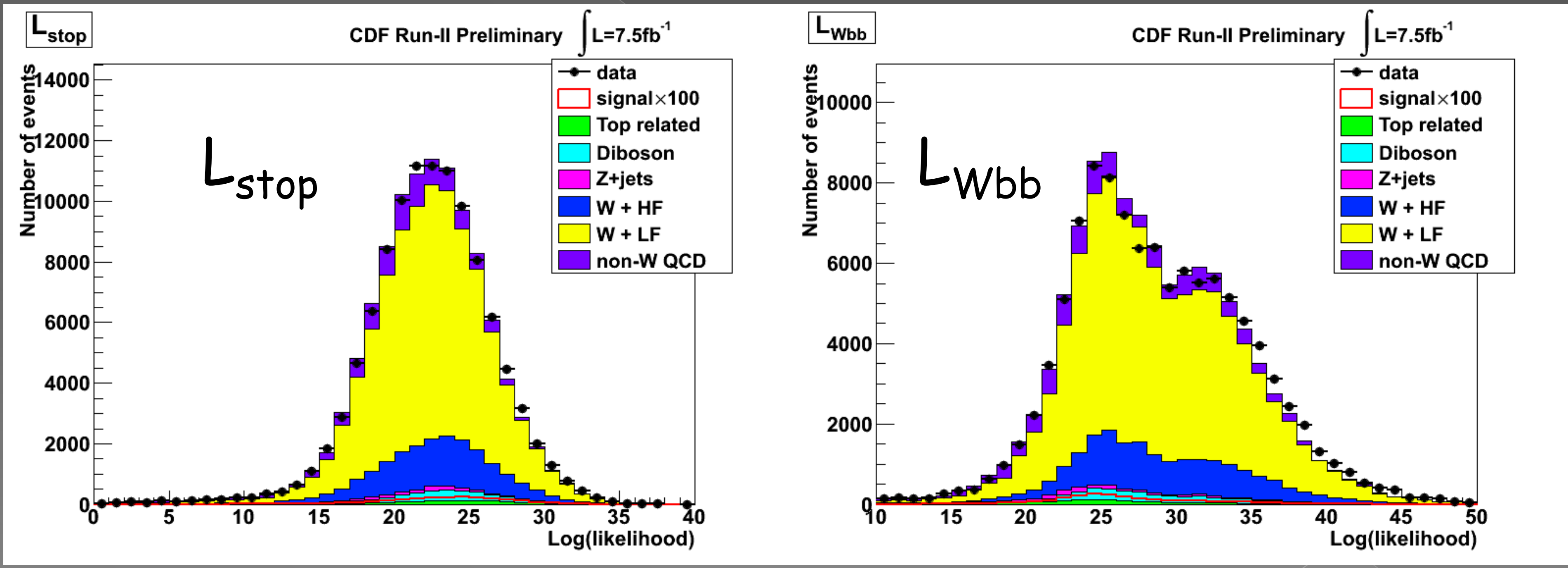
# Likelihood - from Wbb & Stop M.E

- Comparison of the distribution
  - Perform DLM using Wbb(leading order) or singletop(s-ch) Matrix Elements



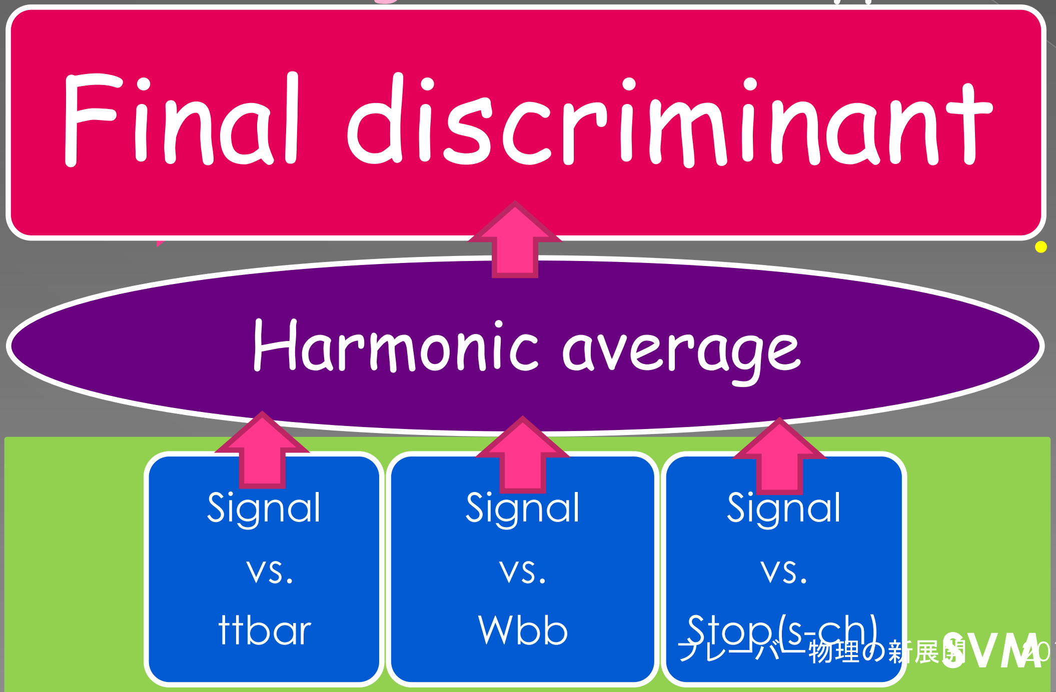
# Likelihood – from Wbb & Stop M.E

- DLM result of the sample before b-tag
  - Perform DLM using Wbb(leading order) or singletop(s-ch) Matrix Elements



# Forming discriminant

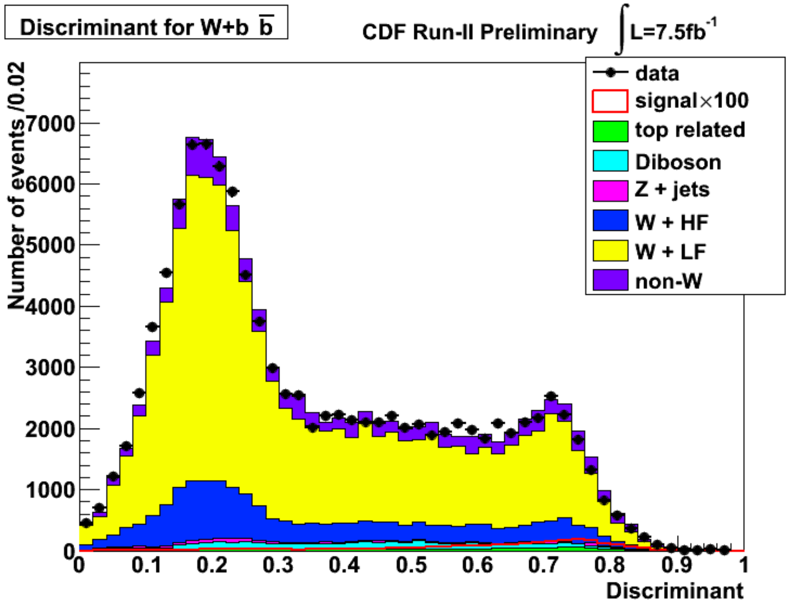
- **Using Support Vector Machine to separate signal from background**
  - 3-type discriminants are made using SVM
    - Signal vs. ttbar ($d_{ttbar}$)
    - Signal vs. Wbb ($d_{Wbb}$)
    - Signal vs. singletop(s-ch) ($d_{stop}$)
  - Final discriminant is obtained by calculating harmonic average of those 3-type discriminants:

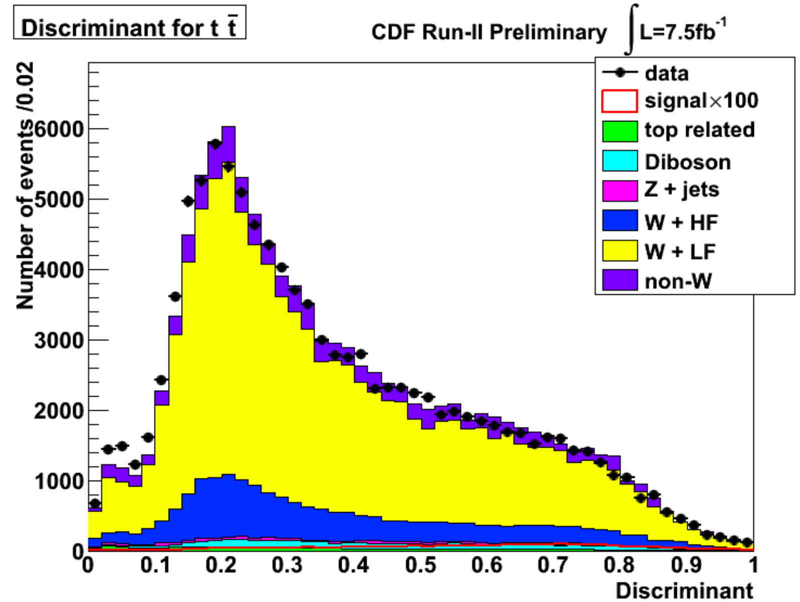


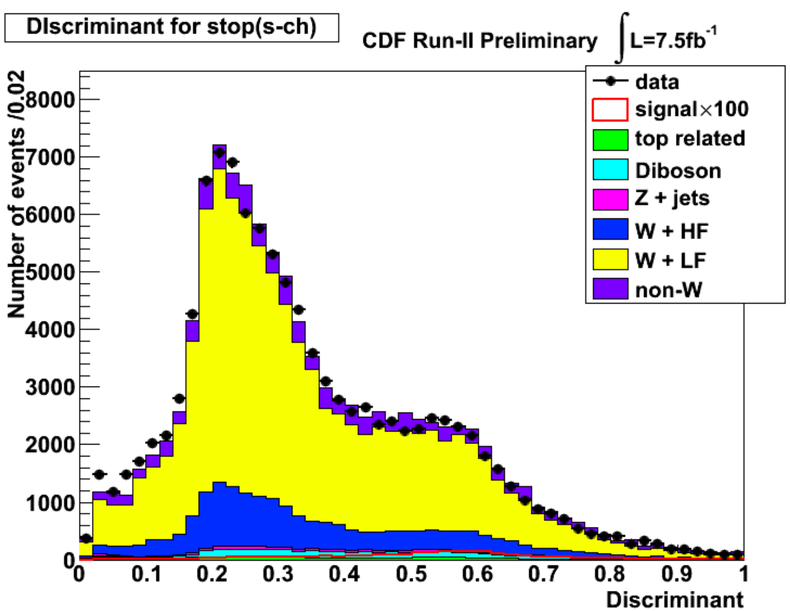

$$d_{final} = \frac{3}{\frac{1}{d_{ttbar}} + \frac{1}{d_{Wbb}} + \frac{1}{d_{stop}}}$$

- Signal event requires high scores of these 3-type discriminant
  - Harmonic average can impose strict condition
  - Harmonic average obtains maximum S/N

# Apply to the sample before b-tag



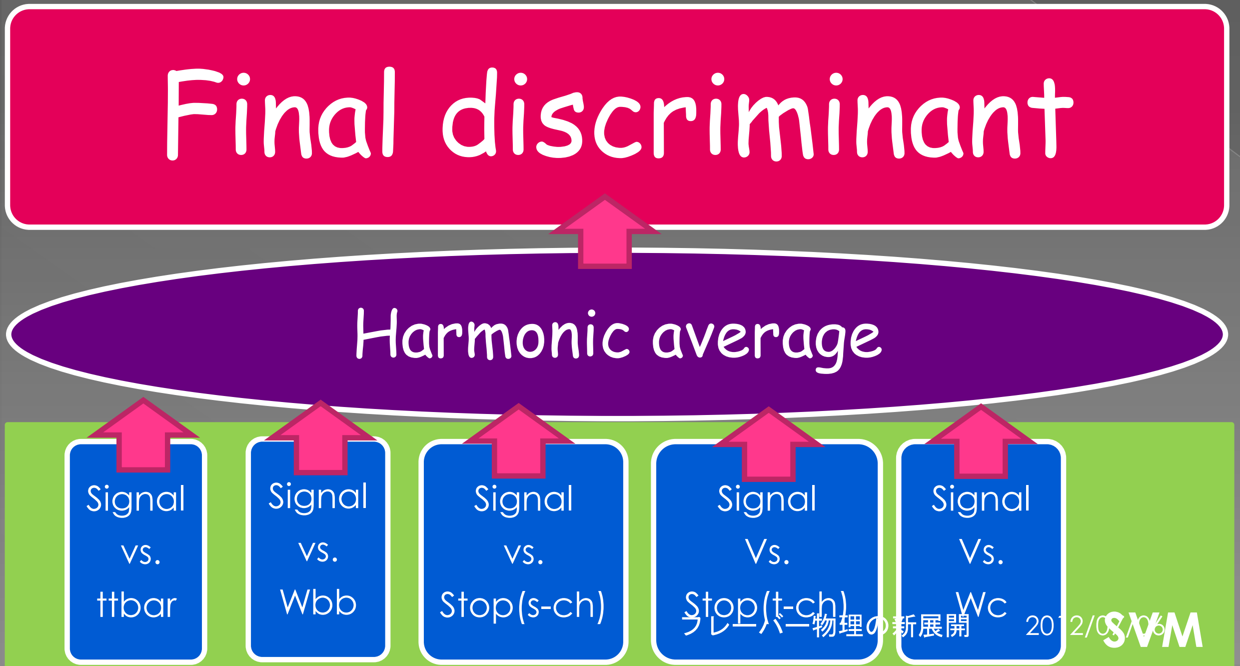




Output result after training
Apply 2SECVTX tag SVM discriminant to pretag

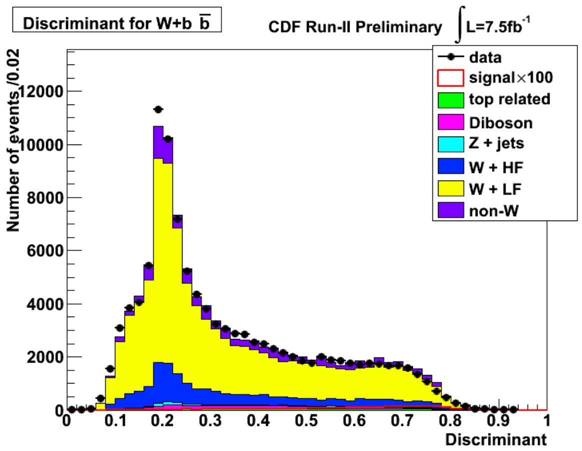
# Making discriminant for single tag case

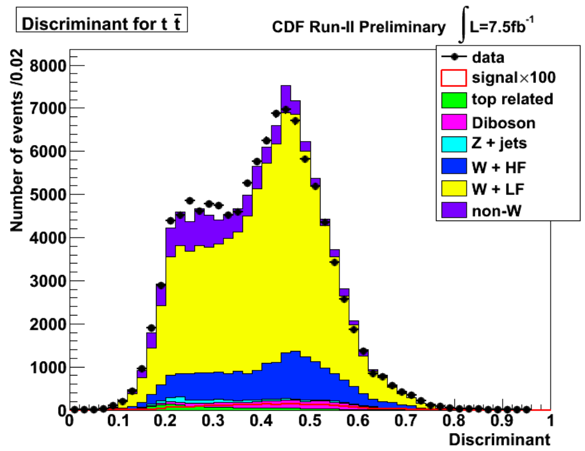
- Almost same method as double tag case
  - 5-type discriminants are made using SVM
    - Signal vs. ttbar ($d_{ttbar}$)
    - Signal vs. Wbb ($d_{Wbb}$)
    - Signal vs. singletop(s-ch) ($d_{stopsch}$)
    - Signal vs. singletop(t-ch) ($d_{stoptch}$)
    - Signal vs. Wc ($d_{Wc}$)
  - harmonic average is used as final discriminanat:

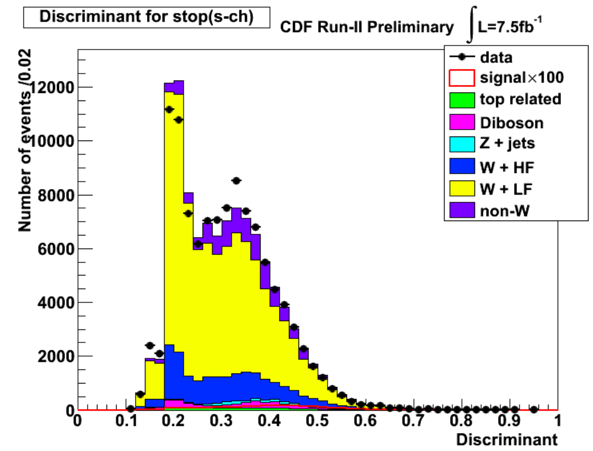
$$d_{final} = \frac{5}{\frac{1}{d_{ttbar}} + \frac{1}{d_{Wbb}} + \frac{1}{d_{stop}} + \frac{1}{d_{stoptch}} + \frac{1}{d_{Wc}}}$$

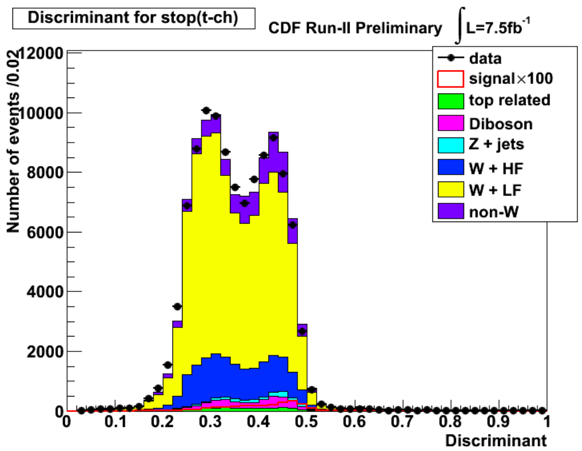


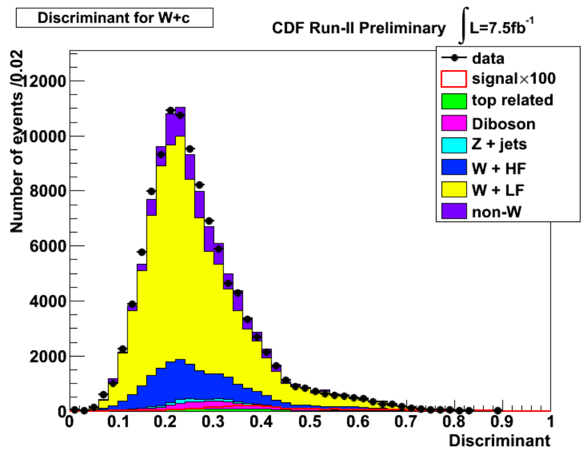
# Apply to the sample before b-tag



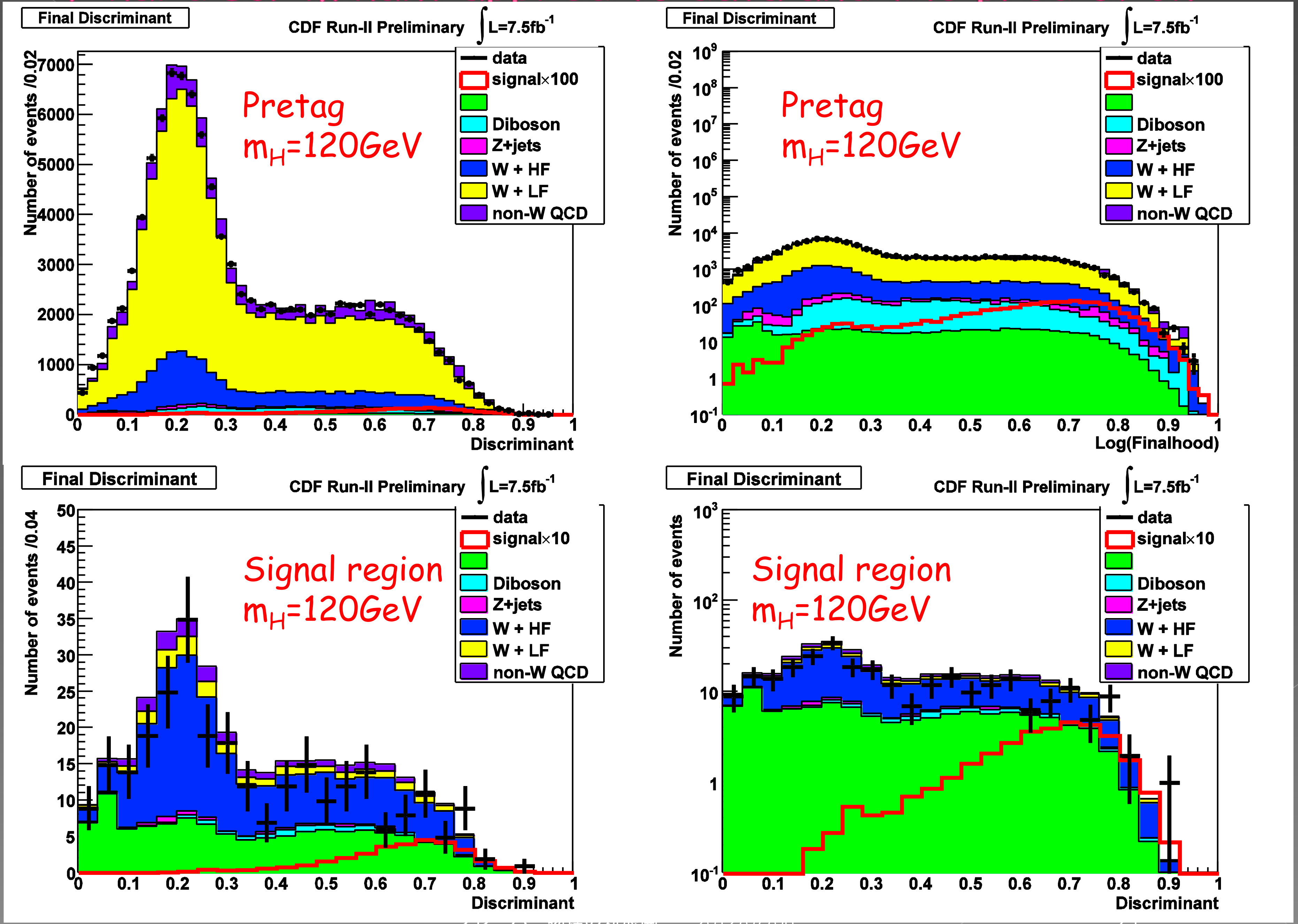



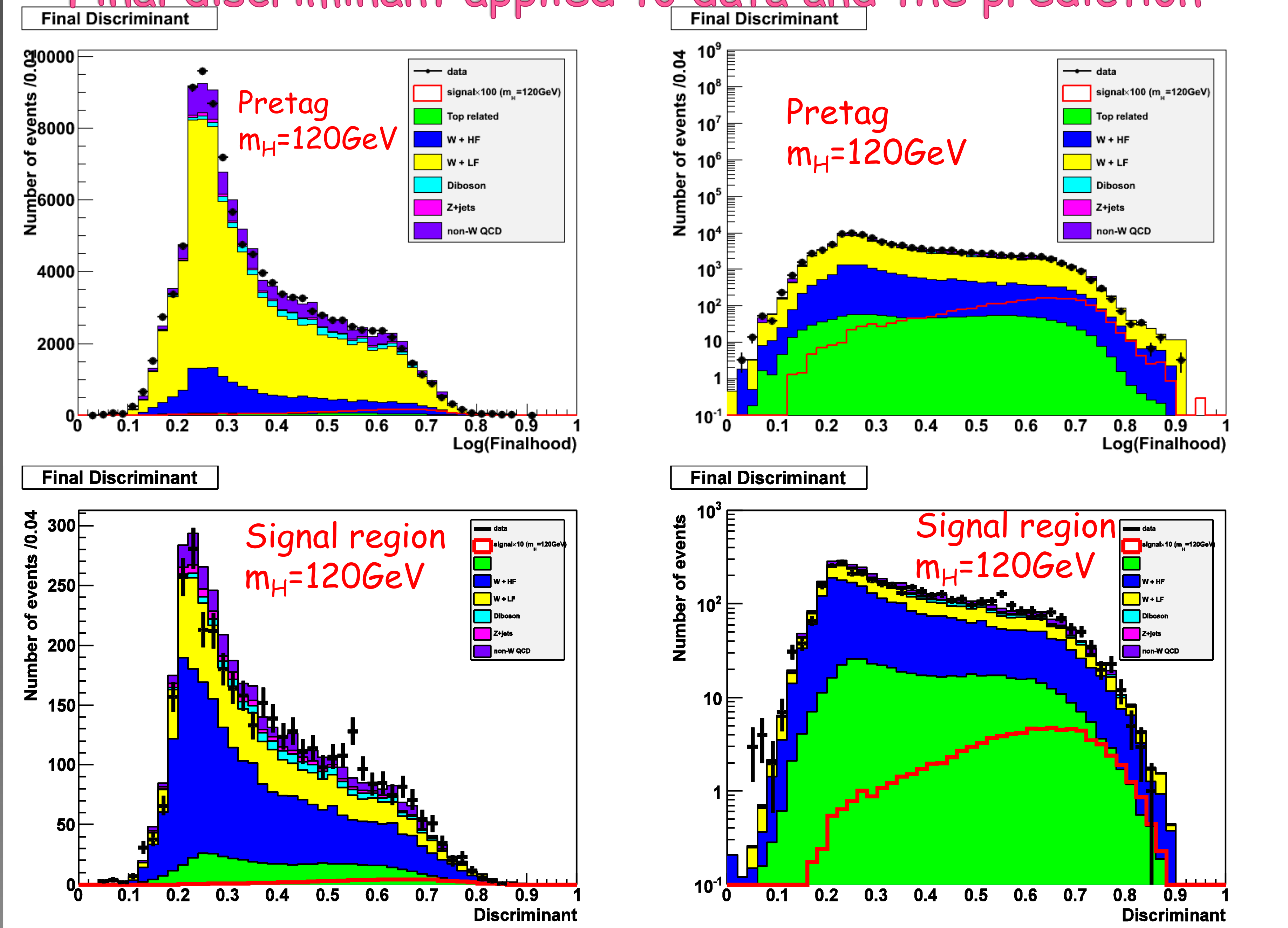




Output result of 1secvtx case
Apply to pretag

# Final discriminant applied to data and the prediction



# Final discriminant applied to data and the prediction



# Systematics summary on signal

| Source | Error (%) | | |
|---|---|---|---|
| | STST | STJP | ST |
| JES | 2.4 | 2.2 | 2.9 |
| ISR/FSR | 6.0 | 4.0 | 3.1 |
| PDF | 1.5 | 1.4 | 1.1 |
| b-tagging | 8.6 | 8.1 | 4.3 |
| Luminosity | 6 | 6 | 6 |
| Lepton ID SF | 2 | 2 | 2 |
| Trigger | $\sim 1$ | $\sim 1$ | $\sim 1$ |

Table 1: Summary of systematic uncertainties on the acceptance in central lepton events

| Source | Error (%) | | |
|---|---|---|---|
| | STST | STJP | ST |
| JES | 2.7 | 3.6 | 2.5 |
| ISR/FSR | 4.4 | 5.9 | 5.5 |
| PDF | 2.7 | 1.7 | 4.1 |
| b-tagging | 8.6 | 8.1 | 4.3 |
| Luminosity | 6 | 6 | 6 |
| Lepton ID SF | 2 | 2 | 2 |
| Trigger | $\sim 1$ | $\sim 1$ | $\sim 1$ |

Table 2: Summary of systematic uncertainties on the acceptance in forward-backward electron events

| | STST | STJP | ST |
|---|---|---|---|
| JES | 2.2 | 3.6 | 2.5 |
| ISR/FSR | 4.0 | 5.9 | 5.2 |
| PDF | 2.8 | 1.2 | 1.2 |
| b-tagging | 8.6 | 8.1 | 4.3 |
| Luminosity | 6 | 6 | 6 |
| Track Reco. | 8.85 | 8.85 | 8.85 |
| Trigger | 2 | 2 | 2 |

Table 3: Summary of systematic uncertainties on the acceptance in Isolated Track events

Main source on signal:

- Luminosity
- b-tagging uncertainty

# Systematics summary on backgrounds

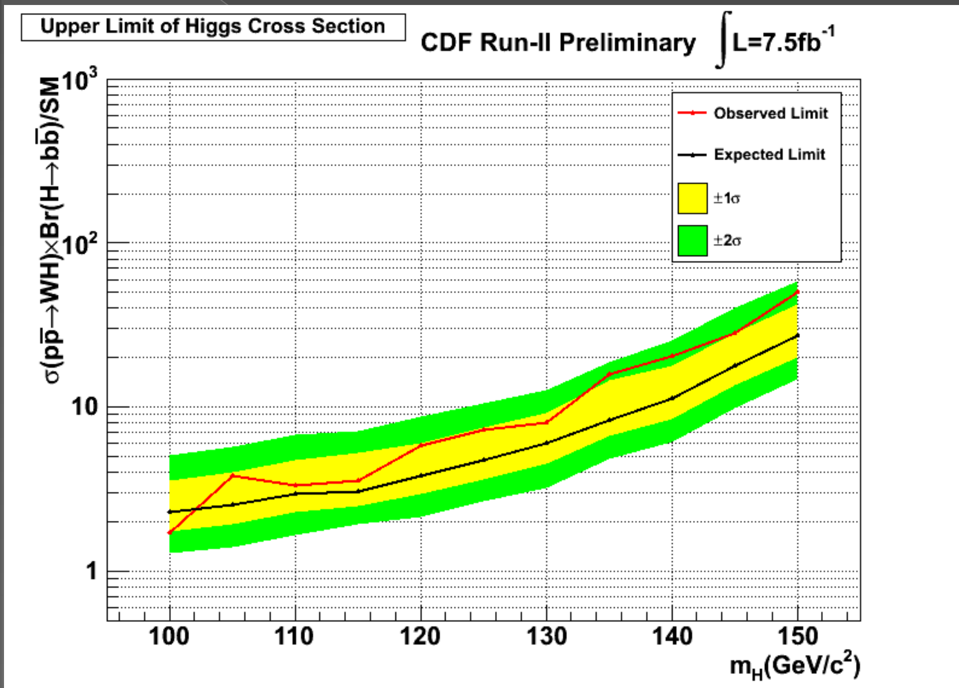
Typical value

| Source | Error (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}$ | singletop | $W+b\bar{b}$ | $W+c\bar{c}$ | Mistag | Diboson | Z+jets | nonW |
| JES | 7 | 2 | 10 | 10 | | 2 | 8 | |
| ISR/FSR | 5 | 3 | | | | | | |
| PDF | 5 | 1 | | | | 3 | 5 | |
| HF fraction | | | 38 | 38 | | | | |
| Mistag rate | | | | | 20 | | | |
| Z+jets cross section | | | | | | 7 | | |
| Fit | | | | | | | | 40 |
| Luminosity | 6 | 6 | | | | 6 | 6 | |
| Trigger(central & plug) | 1 | 1 | | | | 1 | 1 | |
| Trigger(Isolated Track ) | 2 | 2 | | | | 2 | 2 | |
| Lepton ID (central & plug) | 2 | 2 | | | | 2 | 2 | |
| Reconstruction (Isolated Track) | 8.9 | 8.9 | | | | 8.9 | 8.9 | |
| b-tag(STST) | 8.6 | 8.6 | 8.6 | 8.6 | | 8.6 | 8.6 | |
| b-tag(STJP) | 8.1 | 8.1 | 8.1 | 8.1 | | 8.1 | 8.1 | |
| b-tag(ST) | 4.3 | 4.3 | 4.3 | 4,3 | | 4.3 | 4.3 | |

Table 4: Summary of systematic uncertainties on the backgrounds

Main source on backgrounds:
- Heavy Flavor Fraction uncertainty
- nonW(Fake) uncertainty
- B-tagging
- luminosity

# The preliminary result of upper limit




@$m_H$=125GeV
Exp. 4.74 × SM Obs. 7.32 × SM

| $m_H$(GeV) | Obs. | $-2\sigma$ | $-1\sigma$ | median | $+1\sigma$ | $+2\sigma$ |
|---|---|---|---|---|---|---|
| 100 | 1.72 | 1.29 | 1.76 | 2.29 | 3.56 | 5.14 |
| 115 | 3.85 | 1.41 | 1.94 | 2.54 | 4.02 | 5.73 |
| 110 | 3.32 | 1.67 | 2.31 | 2.94 | 4.79 | 6.77 |
| 115 | 3.60 | 1.93 | 2.49 | 3.07 | 5.30 | 7.19 |
| 120 | 5.88 | 2.16 | 2.94 | 3.82 | 6.08 | 8.78 |
| 125 | 7.32 | 2.67 | 3.63 | 4.74 | 7.53 | 10.64 |
| 130 | 8.10 | 3.21 | 4.50 | 6.04 | 9.21 | 12.67 |
| 135 | 15.83 | 4.81 | 6.69 | 8.33 | 14.48 | 18.76 |
| 140 | 20.41 | 6.10 | 8.54 | 11.30 | 17.95 | 25.54 |
| 145 | 28.05 | 9.88 | 13.59 | 17.84 | 28.18 | 40.22 |
| 150 | 57.95 | 14.88 | 20.13 | 27.42 | 42.70 | 58.69 |

Table 1: The numbers of the upper limit of Higgs production cross section

# Summary

- DLM is being established to analyze WH→lvbb process
  - Signal information can be extracted effectively
  - Performance check is OK for Higgs analysis
- Expected upper limit is calculated using Discriminant
  - Discriminant is obtained by Support Vector Machine and integrate into the final discriminant

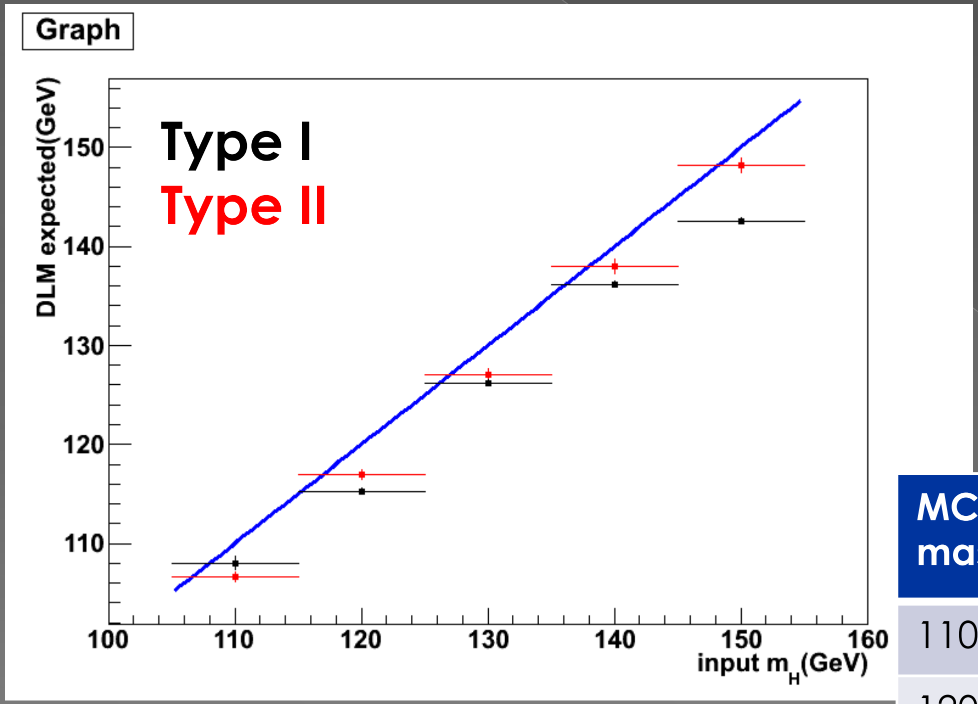
## To do:

- Increase the acceptance
  - Use Full dataset of CDF
  - Use additional lepton events – looser selected lepton events
  - Incorporate same b-tagging as standard analysis
    - Introducing new b-tagger
- Finally, validate background modeling, and calculate observed limit with systematics

# Backups

# Linearity check

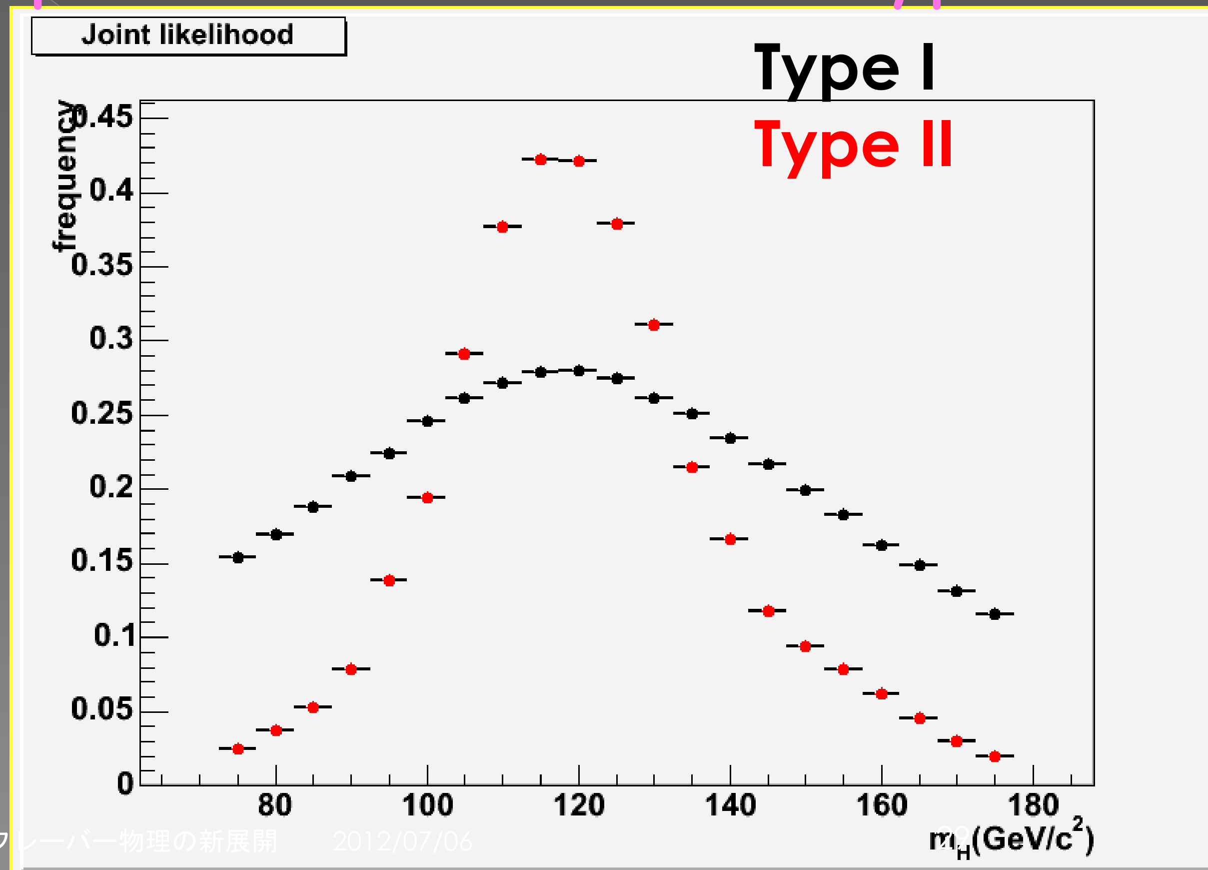
- DLM expected Higgs mass is defined as the maximum likelihood point of the mean of event ensemble
- Linearity is well reserved when using Type II




| MC Higgs mass(GeV/c²) | DLM expected Type I(GeV/c²) | DLM expected Type II(GeV/c²) |
|---|---|---|
| 110 | 108.0±0.7 | 106.6±0.5 |
| 120 | 116.2±0.3 | 117.0±0.5 |
| 130 | 125.2±0.2 | 127.0±0.6 |
| 140 | 136.2±0.2 | 138.0±0.7 |
| 150 | 142.6±0.2 | 148.2±0.7 |

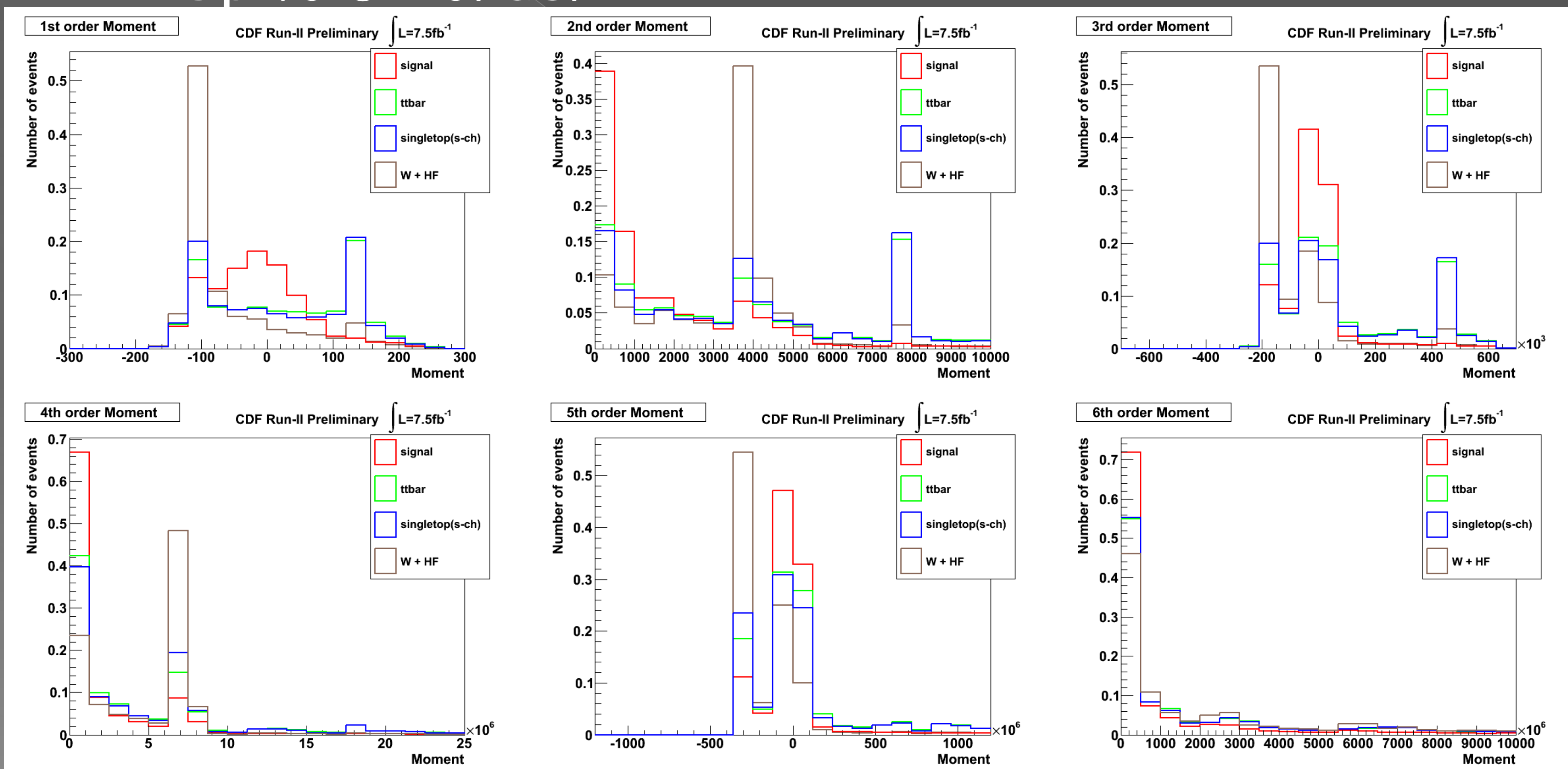
# Shape check with Type I  and Type II

- Comparison of mean value of event ensemble
  - O(10) events
- Normalized to compare the shape
- TypeII has shaper distribution than Type I



# Localized moment around signal

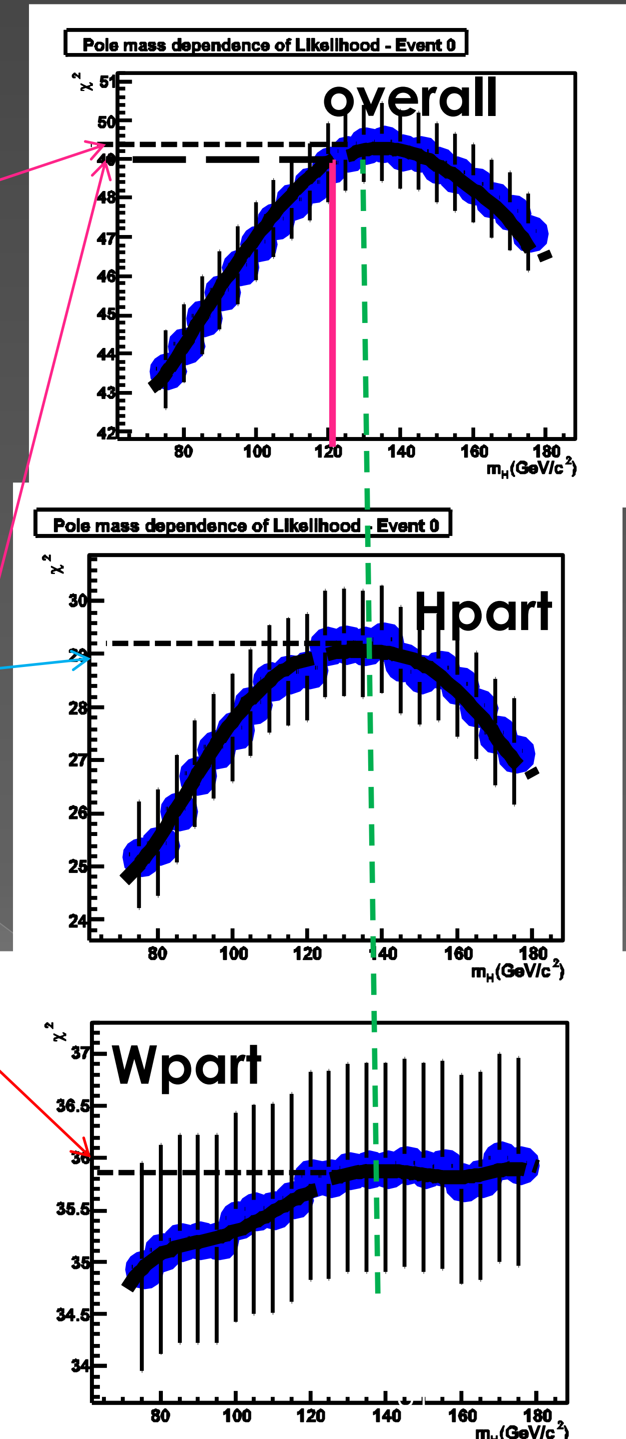
- Localized moment around signal expected mass for each process
  - $\int (m_H - m_{H0})^n \cdot f_{event}(m_H) dm_H$ nth-order moment
  - Up to 6th order



# Likelihood used

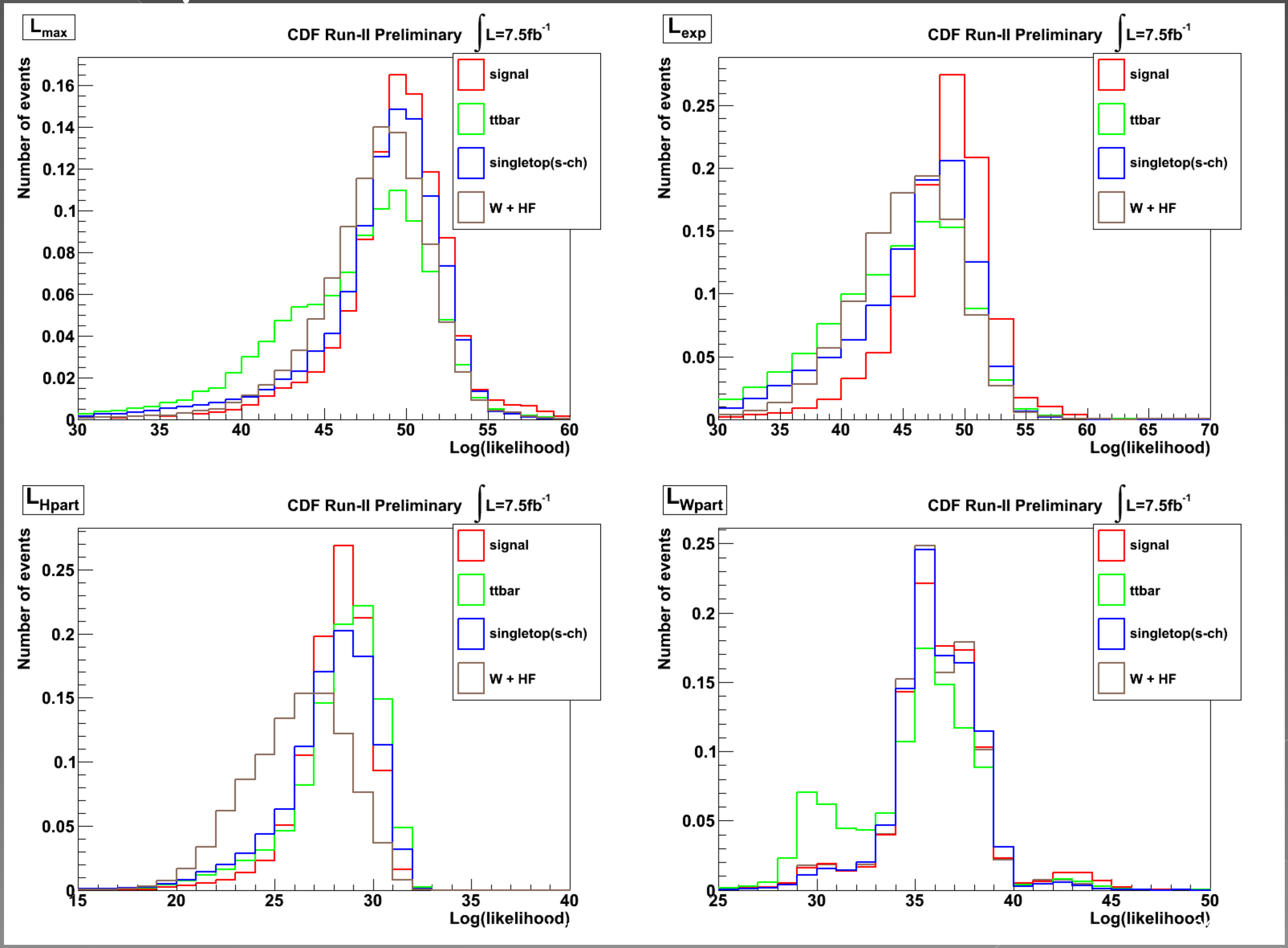
- Maximum likelihood: $L_{max}$
  - Use them event-by-event
- Expectation value of likelihood
  - Wbb : $L_{Wbb}$
  - singletop(s-ch): $L_{stop}$
- Higgs part likelihood : $L_{Hpart}$ @maximum likelihood in overall
  - vertex of Higgs decay & T.F.
- W part likelihood: $L_{Wpart}$ @maximum likelihood in overall
  - vertex of W decay & T.F.
- Likelihood @ DLM expected mass: $L_{exp}$
  - Expected mass means the result of event ensemble

 (e.g. 117.01GeV@mH=120GeV)

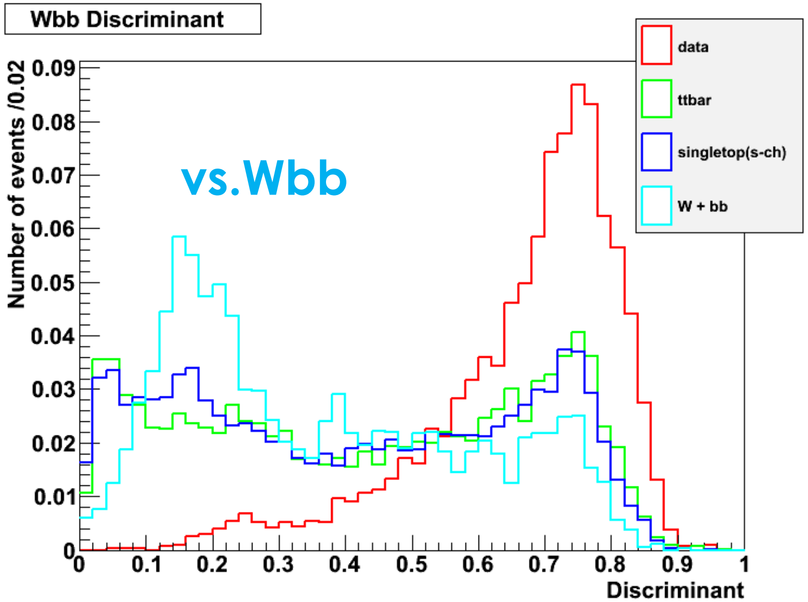


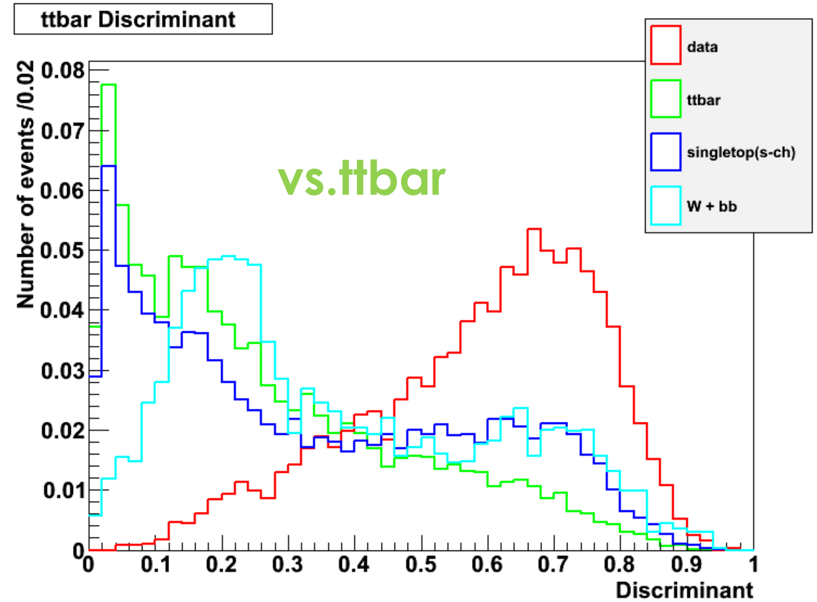
# Likelihood – from signal M.E

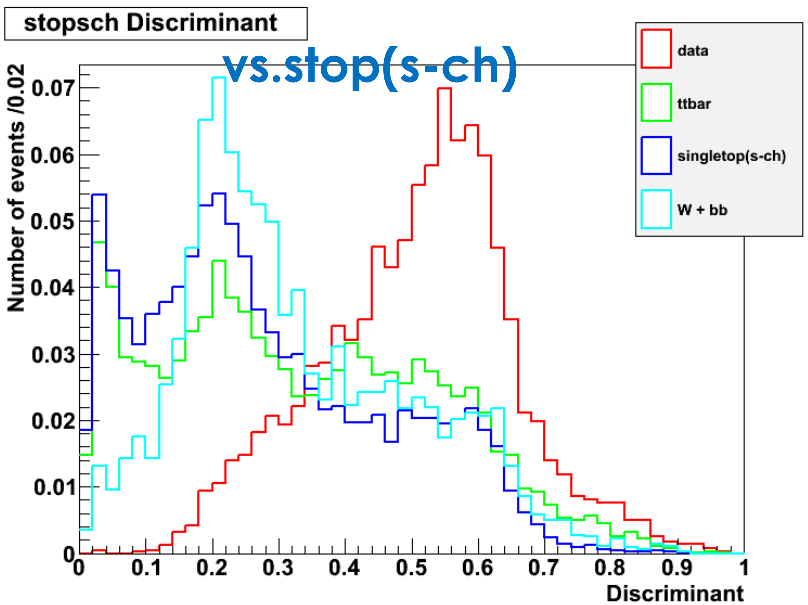
- Comparison of the distribution



# SVM output example




Output result after training
2secvtx tag case

Signal
ttbar
Stop(s-ch)
Wbb

# Final discriminant

- Final discriminant after calculating harmonic average
  - $d_{final} = \frac{3}{\frac{1}{d_{ttbar}} + \frac{1}{d_{Wbb}} + \frac{1}{d_{stop}}}$
  - $m_H$=120GeV
  - 2secvtx tag