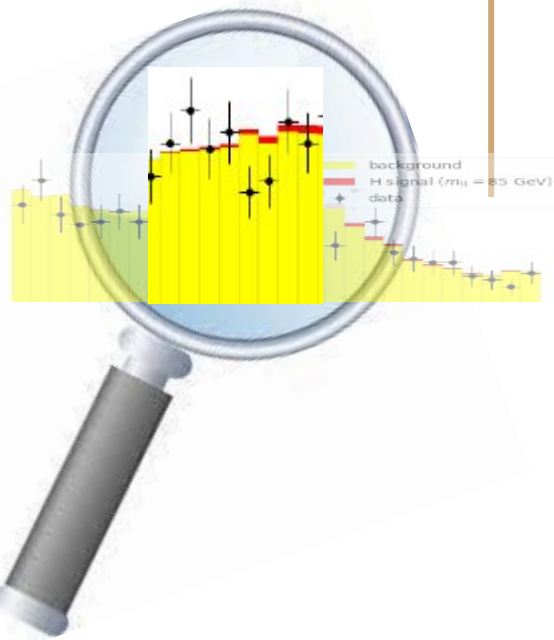


Higgs Search at LEP with the L3 detector

Janík W. Andrejkovic
& Jonas Kunath

23.05.2017

Spring Semester 2017 of "*Statistical Methods and Analysis Techniques in Experimental Physics*"
Taught by: Christoph Grab and Mauro Donegà
Assistants: Matthieu Marionneau and Luca Perrozzi

ETH zürich

Outline

What is the signal we are looking for?

Our Analysis

Step 1	Select signal candidates
Step 2	Perform log likelihood ratio analysis
Step 3	Set confidence levels

- ❏ cut-flow approach
- ❏ BDT optimized cut

- ❏ 1D analysis → single discriminating variable
- ❏ 2D analysis using MVA technique for second variable

- ❏ CL_s method

Outline

What is the signal we are looking for?

Our Analysis

Step 1	Select signal candidates
Step 2	Perform log likelihood ratio analysis
Step 3	Set confidence levels

- ❏ cut-flow approach
- ❏ BDT optimized cut

- ❏ 1D analysis → single discriminating variable
- ❏ 2D analysis using MVA technique for second variable

- ❏ CL_s method

Outline

What is the signal we are looking for?

Our Analysis

Step 1	Select signal candidates
Step 2	Perform log likelihood ratio analysis
Step 3	Set confidence levels

- ❏ cut-flow approach
- ❏ BDT optimized cut

- ❏ 1D analysis → single discriminating variable
- ❏ 2D analysis using MVA technique for second variable

- ❏ CL_s method

Outline

What is the signal we are looking for?

Our Analysis

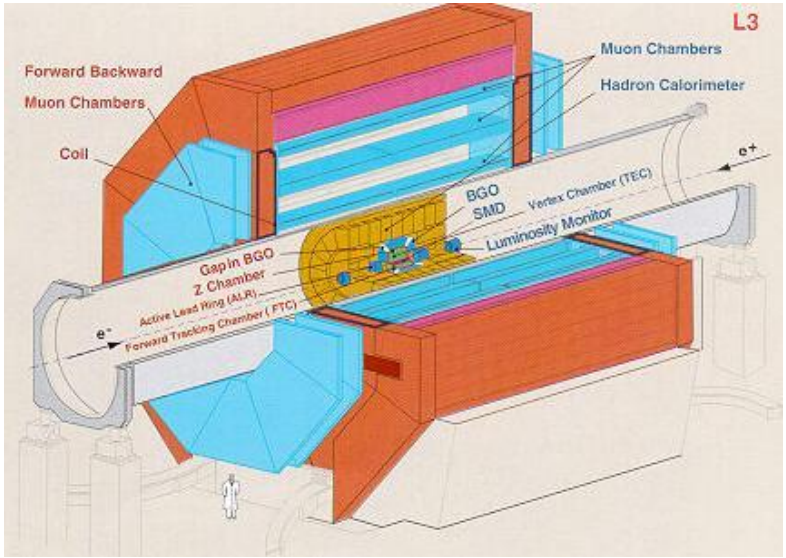
Step 1	Select signal candidates
Step 2	Perform log likelihood ratio analysis
Step 3	Set confidence levels

- ❏ cut-flow approach
- ❏ BDT optimized cut

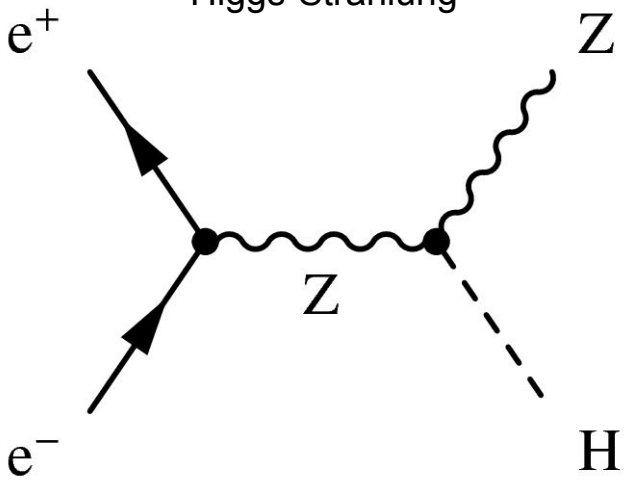
- ❏ 1D analysis → single discriminating variable
- ❏ 2D analysis using MVA technique for second variable

- ❏ CL_s method

Higgs search with L3 detector



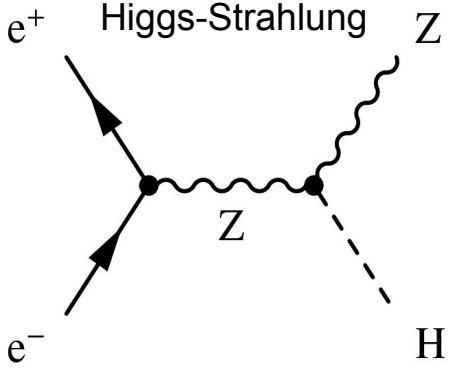
(Production) Process of interest:
Higgs-Strahlung

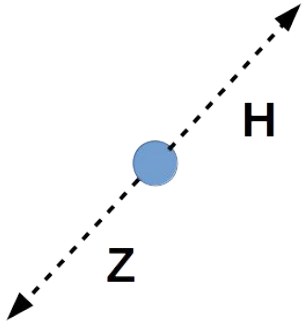


We do not see H and Z in the detector but their decay products

Final State of interest

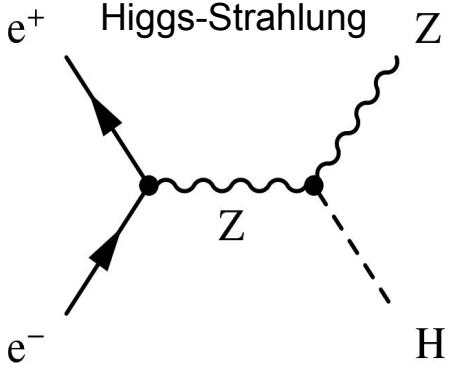
(Production) Process of interest:

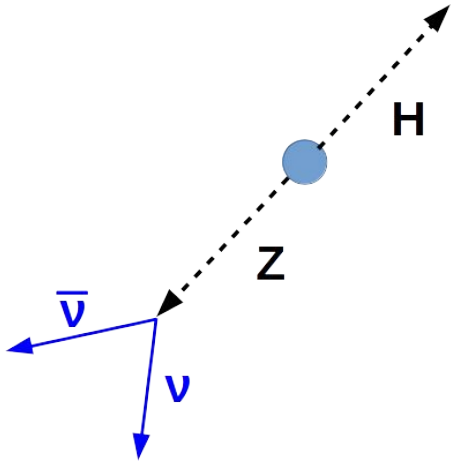




Final State of interest

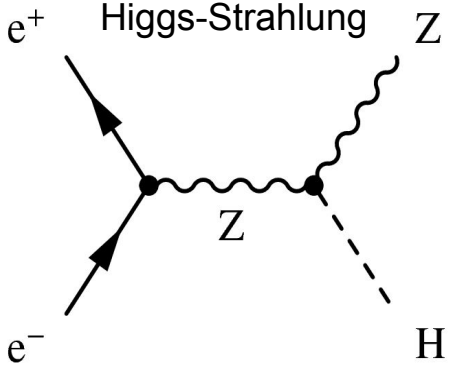
(Production) Process of interest:

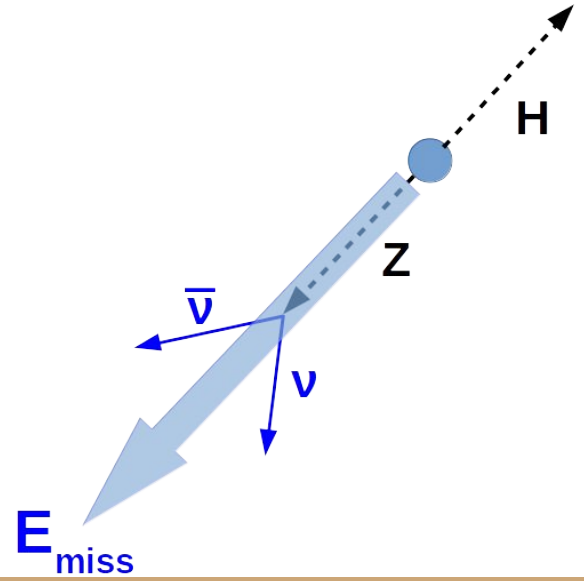




Final State of interest

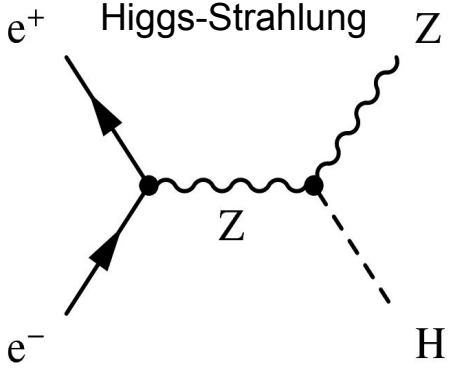
(Production) Process of interest:

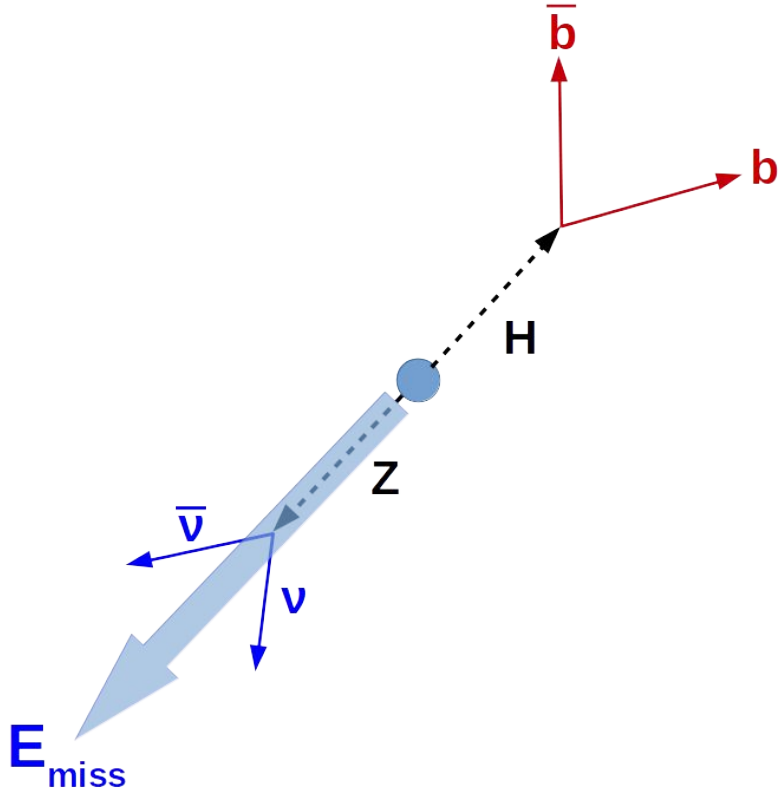




Final State of interest

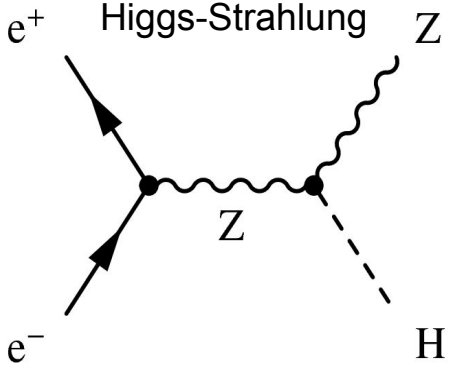
(Production) Process of interest:

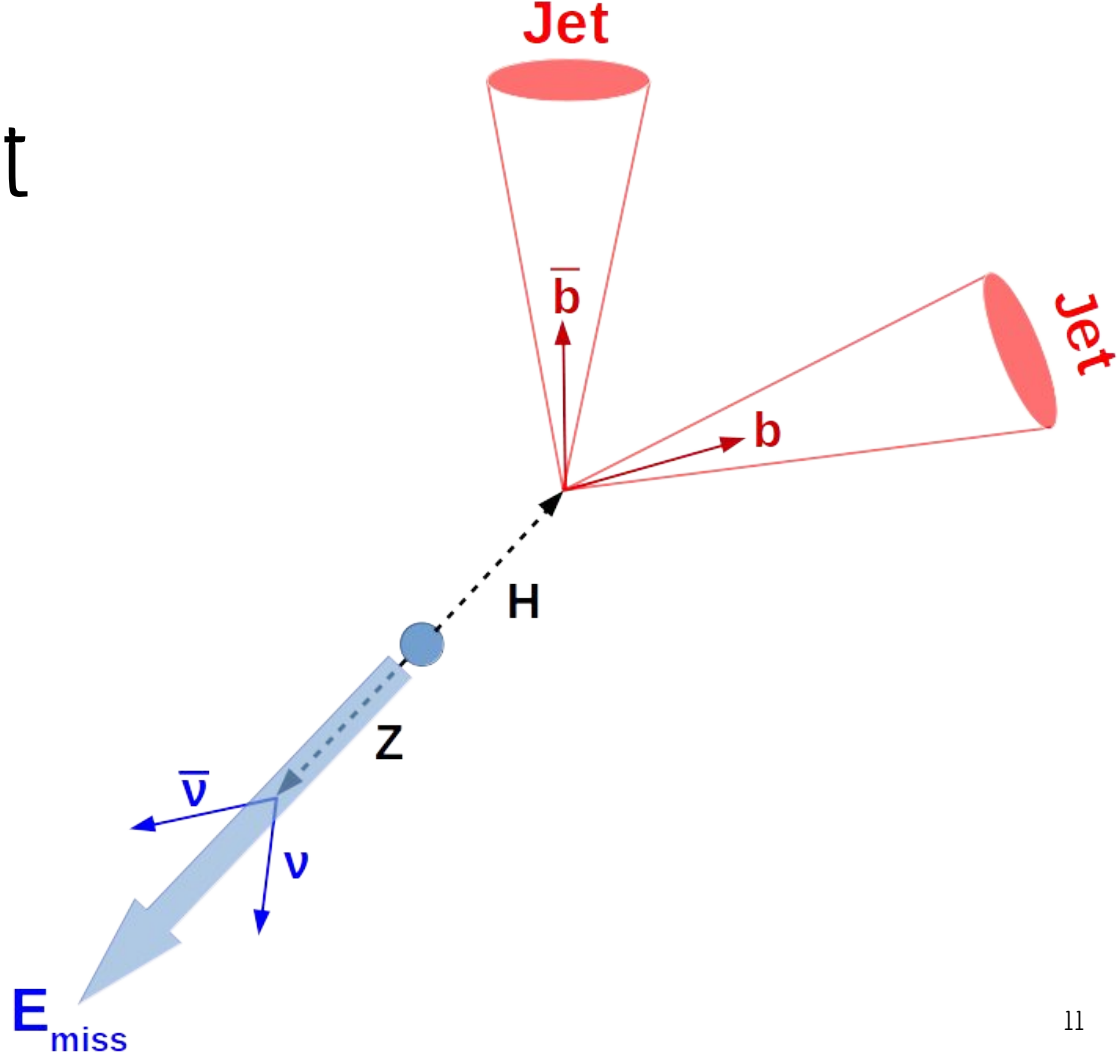




Final State of interest

(Production) Process of interest:





Final State of interest

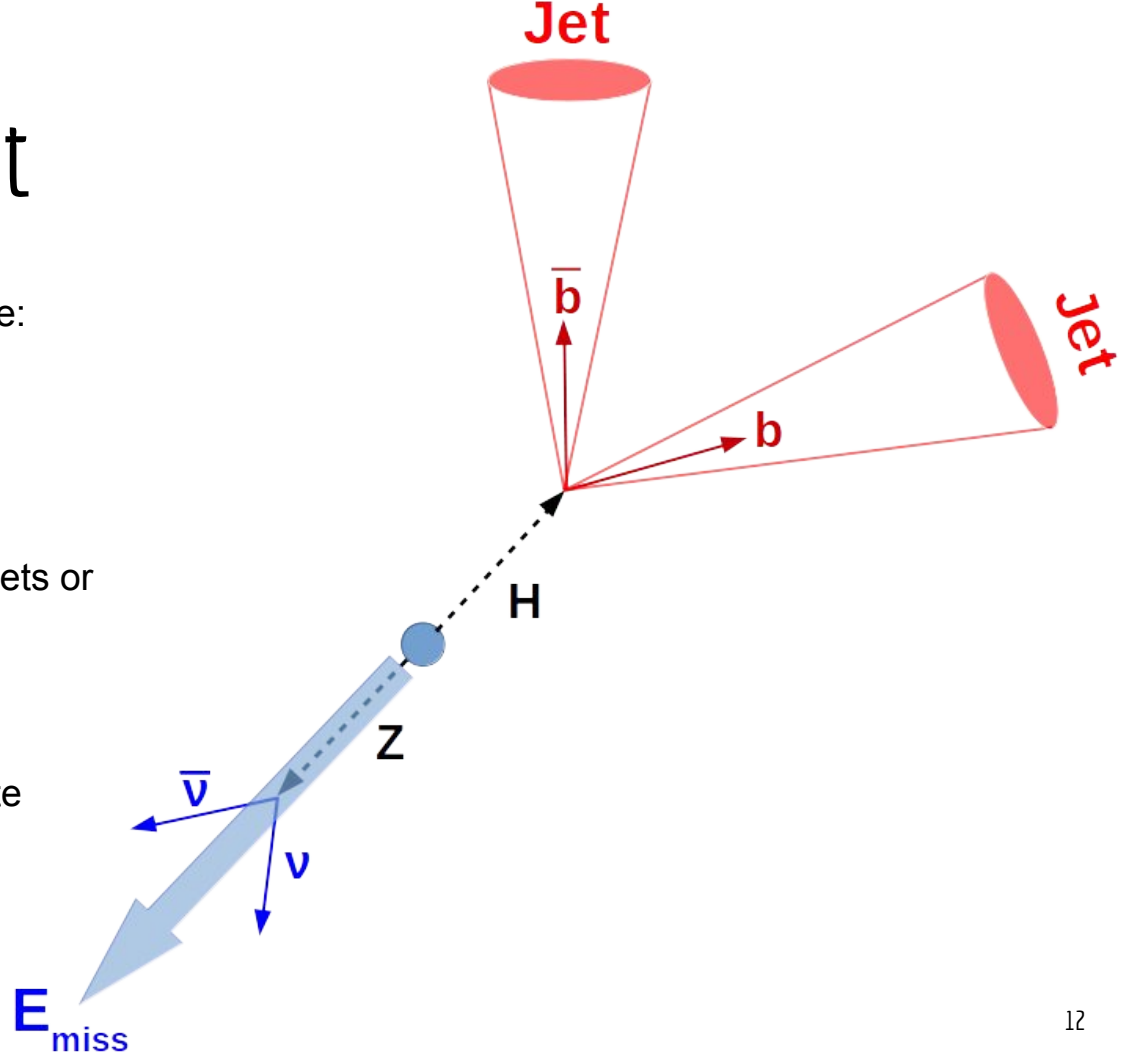
Good variables to characterize the final state:

- missing energy/mass
- btag of the jets, i.e. probability that jet comes from hadronization of b quark

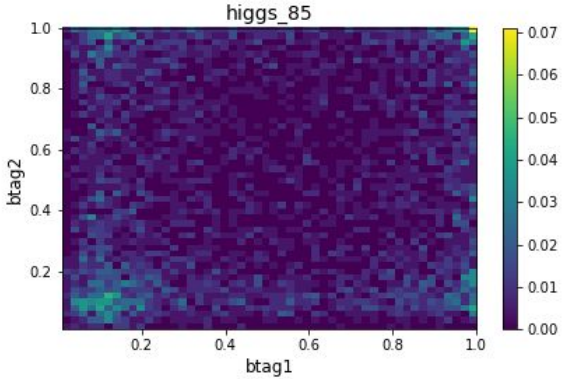
- angular variables, i.e. angle between jets or direction where E_{miss} is pointing.
- …

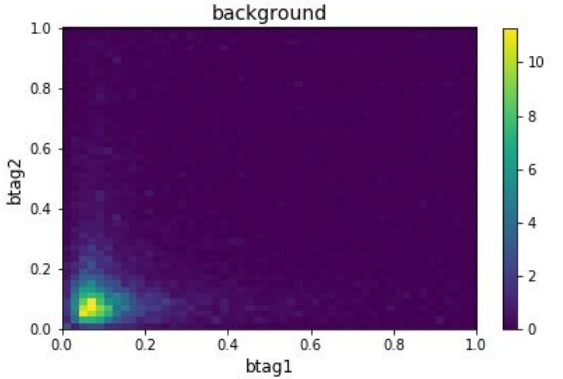
Events without Higgs dominate the final state category → **background**

→ **Apply cuts to enhance signal events**
→ **Obtain Higgs candidates**



Cut-flow approach

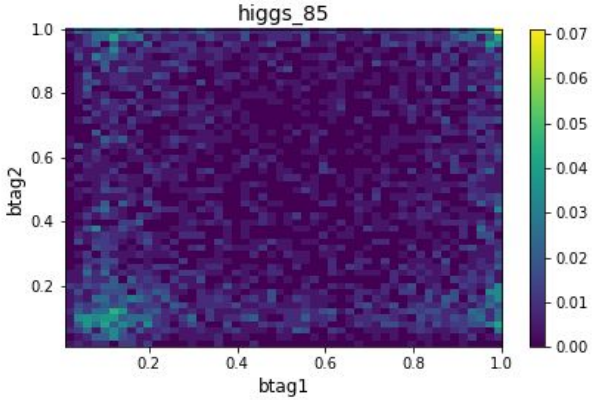


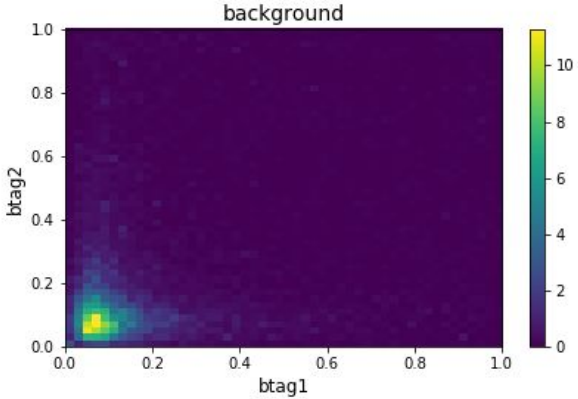


variable	cut applied	data events after cut (85 GeV model)
btag1 and btag2	both < 0.18	274 (/641)
mmis	< 65 GeV	230
mvis	> m_H + 5 GeV	143
fmvis	> m_H + 5 GeV	105
mvissc	> m_H + 5 GeV	97
ucsdbt0	< 1.4	41

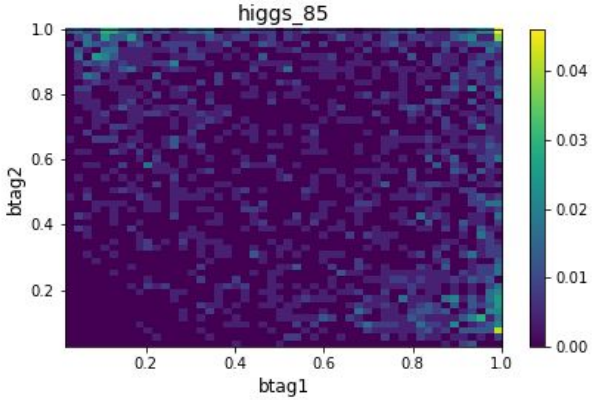
m_H: Higgs-mass in the signal hypothesis

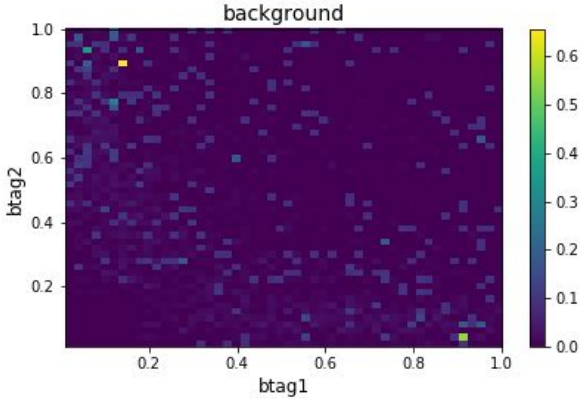
before cut:





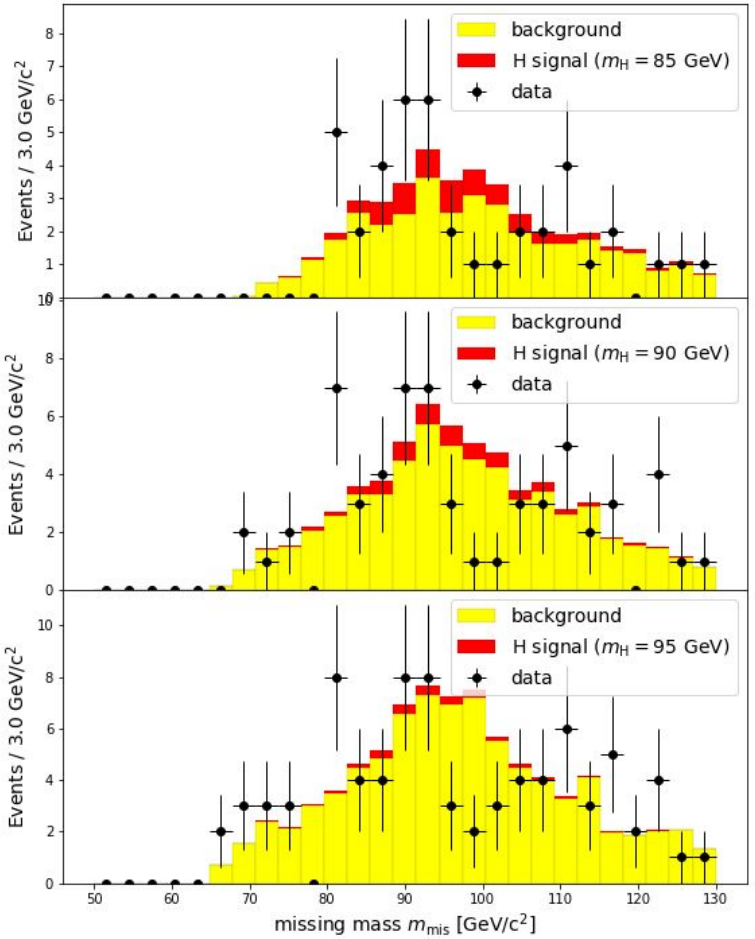
after cut:





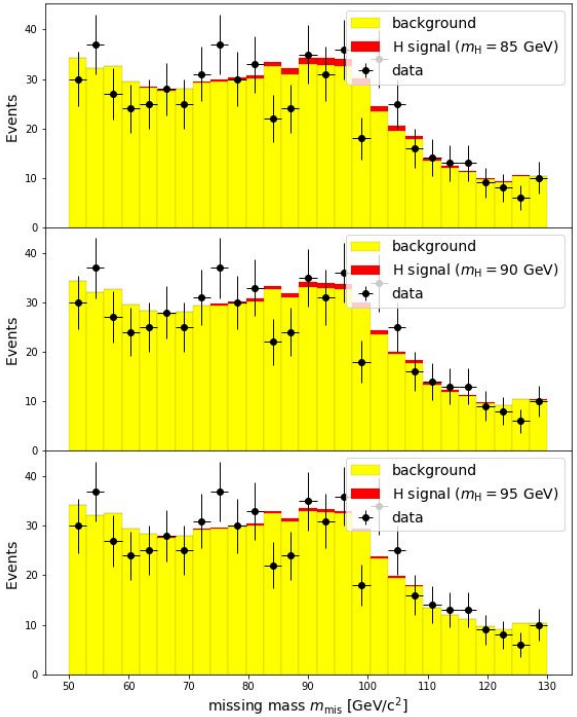
Cut removes all events with btag1 > 0.18 or btag2 > 0.18

1D analysis

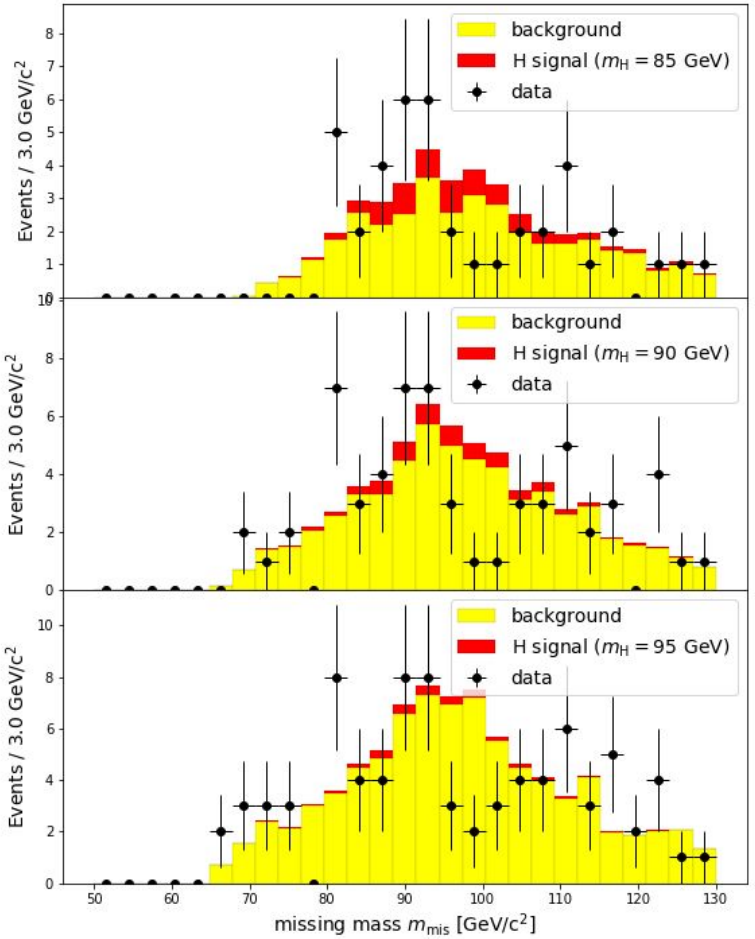


after applying the cuts

1D analysis



without selection cuts



after applying the cuts

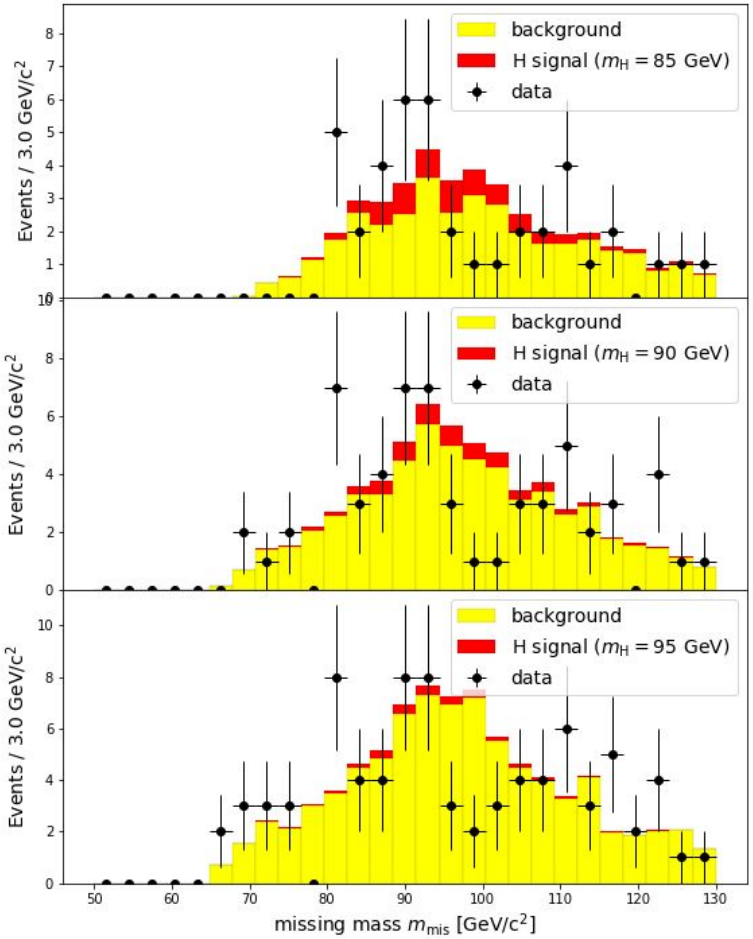
The log-likelihood ratio is defined as:

$$-2\ln(Q(m_H)) = 2s_{tot} - 2\sum_{i=1}^{N_{bins}} N_i \ln\left(1 + \frac{s_i(m_H)}{b_i}\right).$$

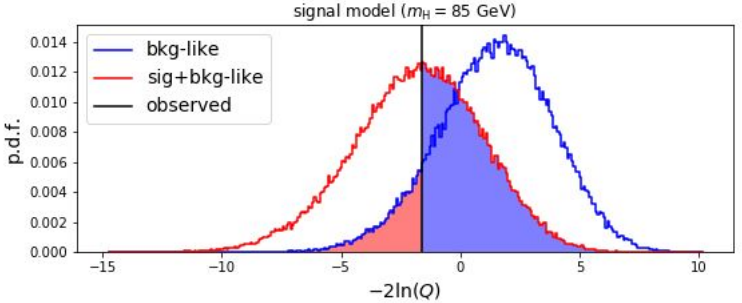
⇒ rebin if $b_i = 0$
⇒ bins below 65 GeV empty because of cut

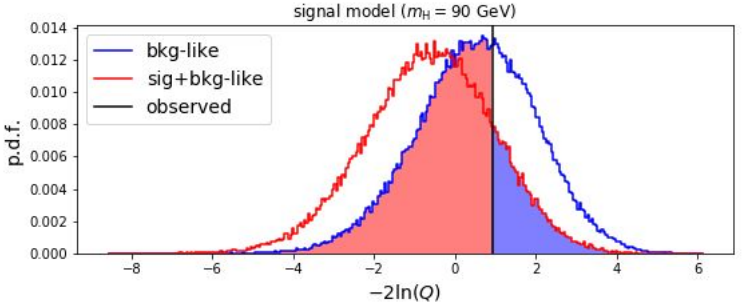
test b-hypo → $N_i \sim b_i$
test s+b-hypo → $N_i \sim b_i + s_i$

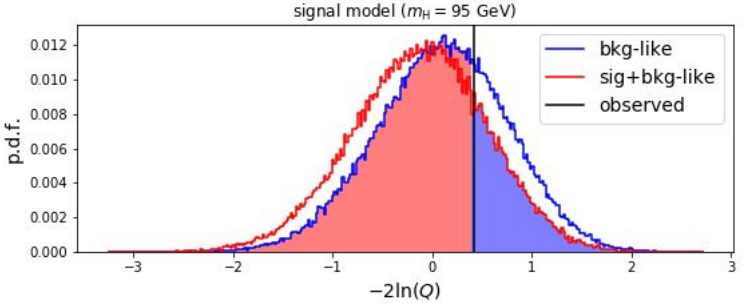
generate toys ⇒ -2ln(Q) distribution



after applying the cuts

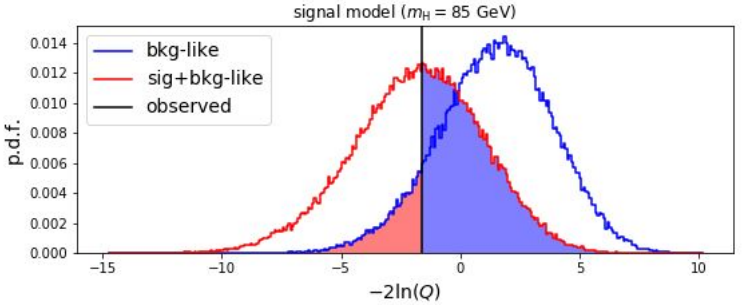


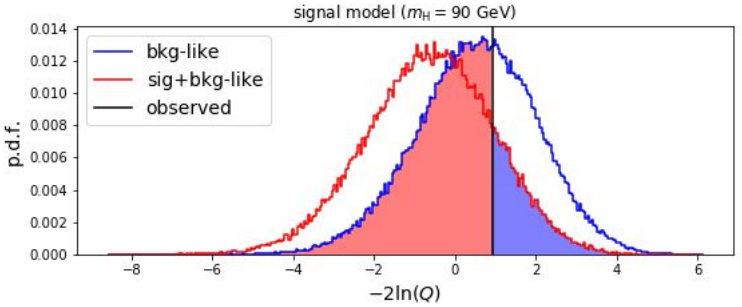


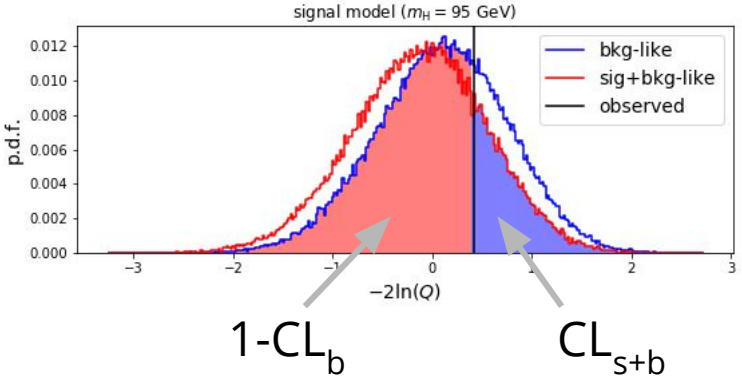


probability density functions of:

$$-2\ln(Q(m_H)) = 2s_{tot} - 2\sum_{i=1}^{N_{bins}} N_i \ln\left(1 + \frac{s_i(m_H)}{b_i}\right).$$







probability density functions of:

$$-2\ln(Q(m_H)) = 2s_{tot} - 2\sum_{i=1}^{N_{bins}} N_i \ln\left(1 + \frac{s_i(m_H)}{b_i}\right).$$

Confidence level:

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{CL_{s+b}}{1-(1-CL_b)}$$

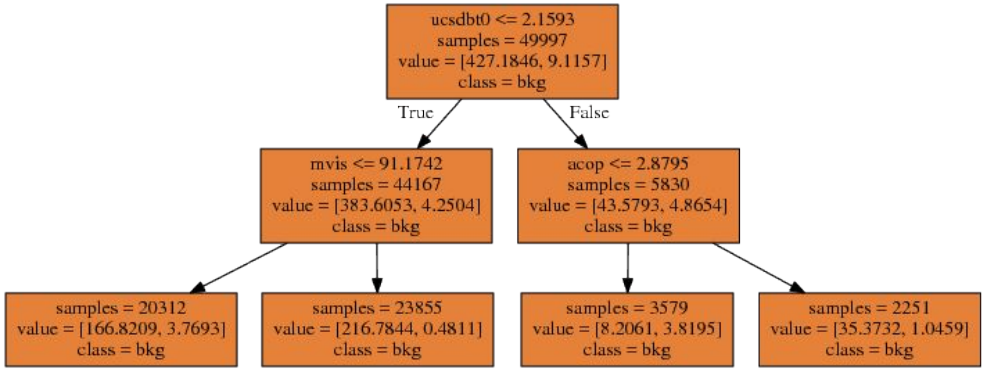
⇒ reject signal hypothesis with 1-CL$_s$

m$_H$	1-CL$_s$	# data events
85 GeV	44.8 %	41 (/641)
90 GeV	59.5 %	60
95 GeV	33.2 %	81

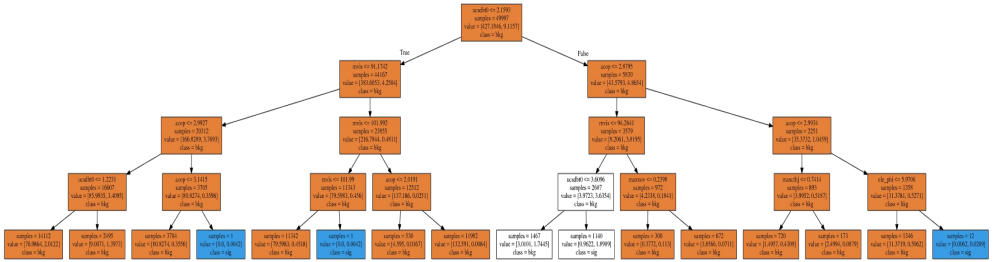
Optimized Cuts from BDT (Boosted Decision Tree)

bkg vs sig discrimination
binary classification problem:

each event characterized by set of kinematical variables → features



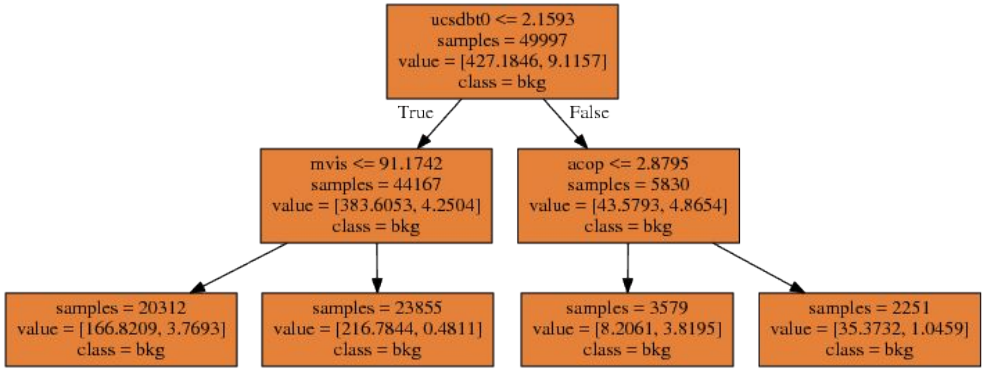
decision tree scans feature space and applies cuts where best separation between classes is achieved



Optimized Cuts from BDT (Boosted Decision Tree)

bkg vs sig discrimination
binary classification problem:

each event characterized by set of kinematical variables → features



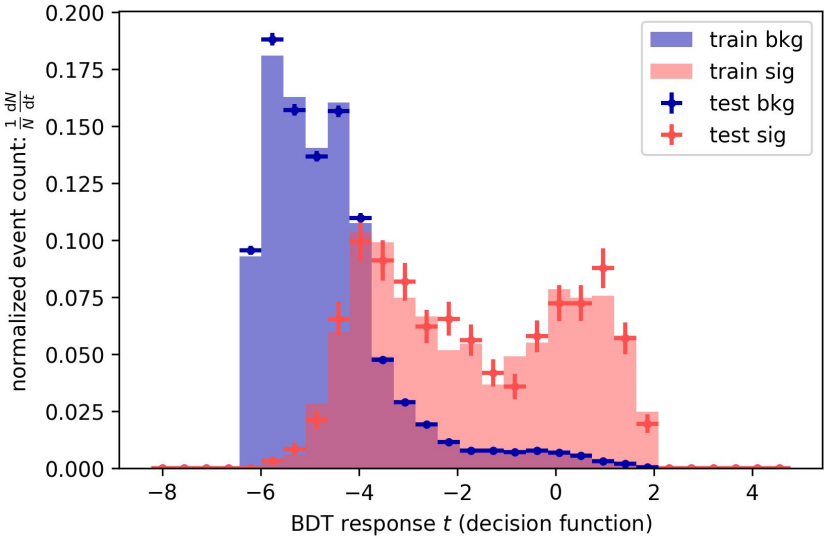
decision tree scans feature space and applies cuts where best separation between classes is achieved

Gradient Boosting → take many shallow trees (weak learners) each trying to correct for the mistakes of the previous one

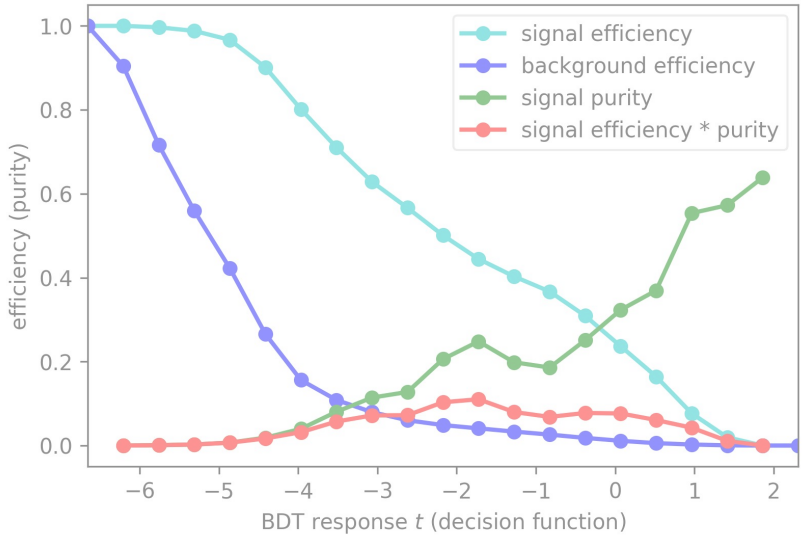
→ BDT builds decision function for sig vs bkg classification

BDT for m_H = 85 GeV

Generalization performance

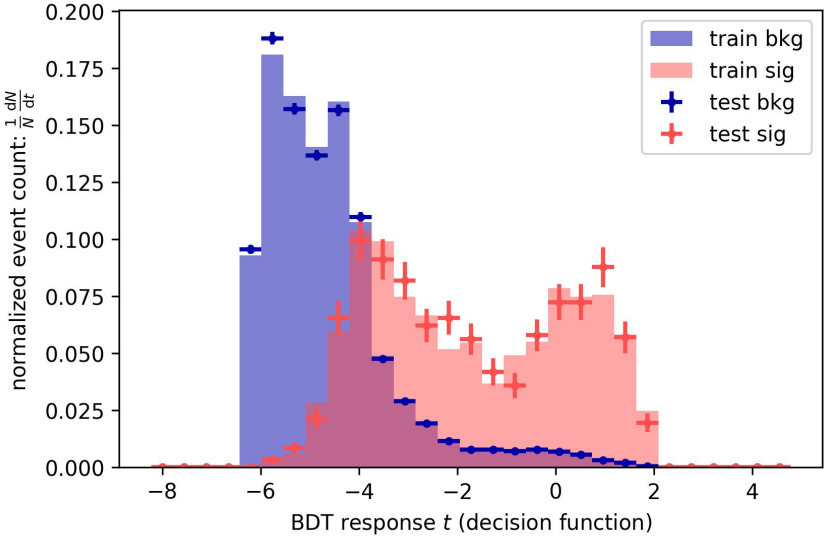


Use BDT response to cut
- but where?

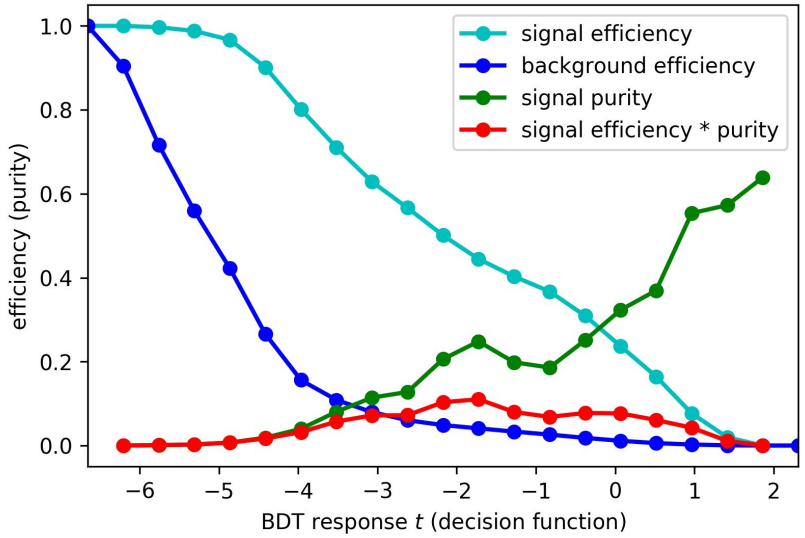


BDT for m_H = 85 GeV

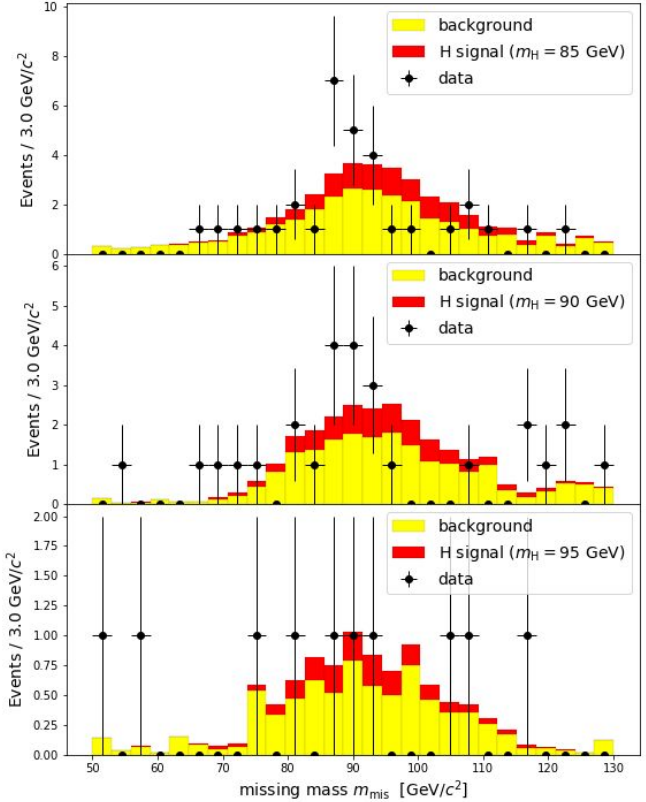
Generalization performance



Use BDT response to cut
- but where?

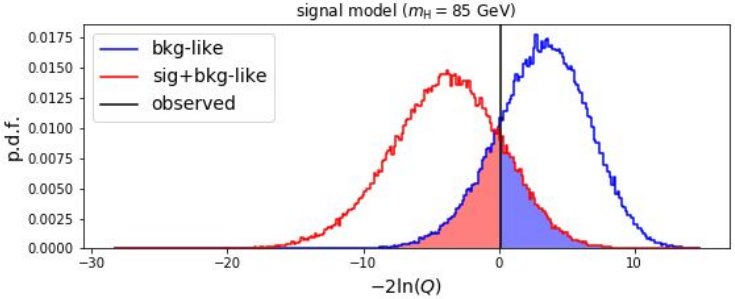


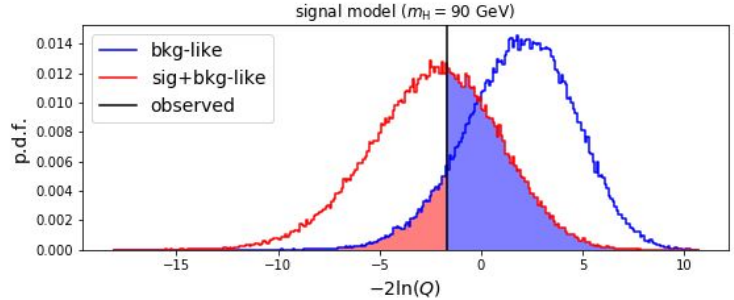
BDT event selection

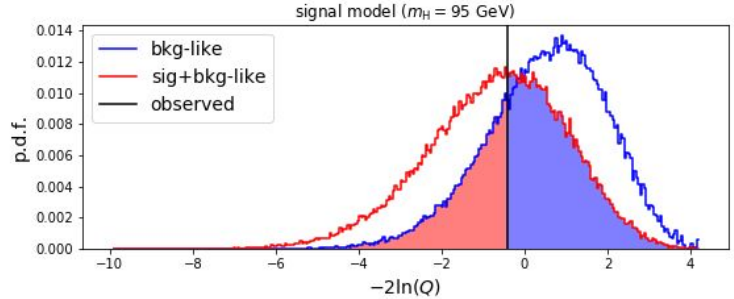


m_H	# data events	purity
85 GeV	27 (/641)	40 %
90 GeV	32	40 %
95 GeV	10	37 %

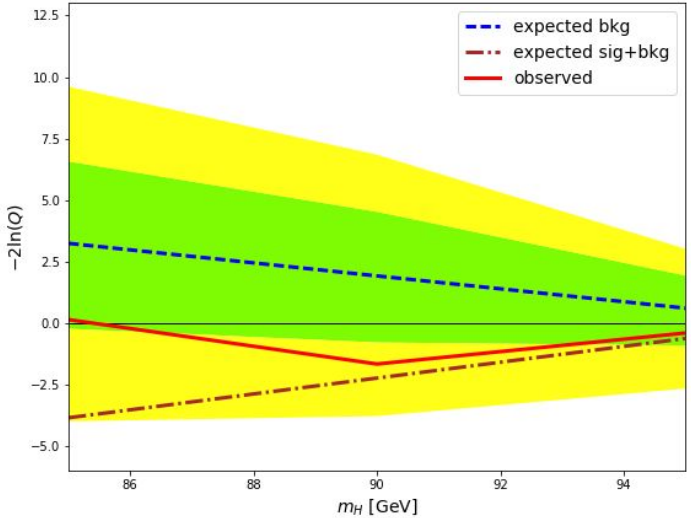
used all (MC) bkg and sig samples for the analysis not just the test set

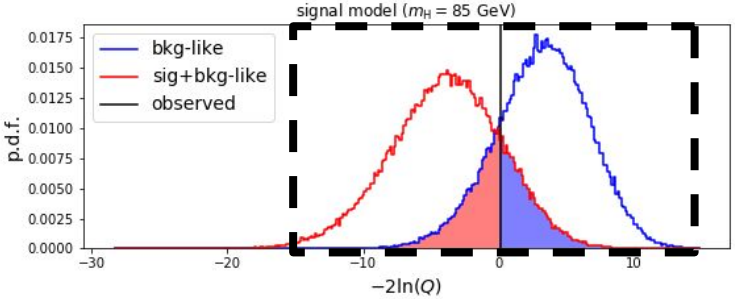


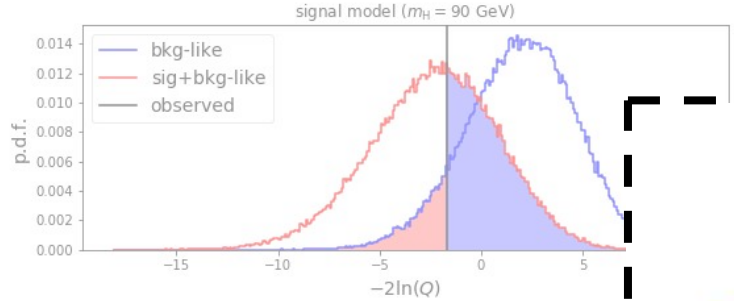


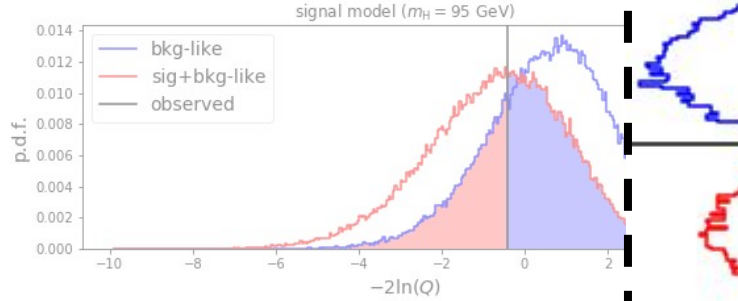


m_H	1-CL_s now	1-CL_s before
85 GeV	78.6 %	44.8 %
90 GeV	50.4 %	59.5 %
95 GeV	37.8 %	33.2 %

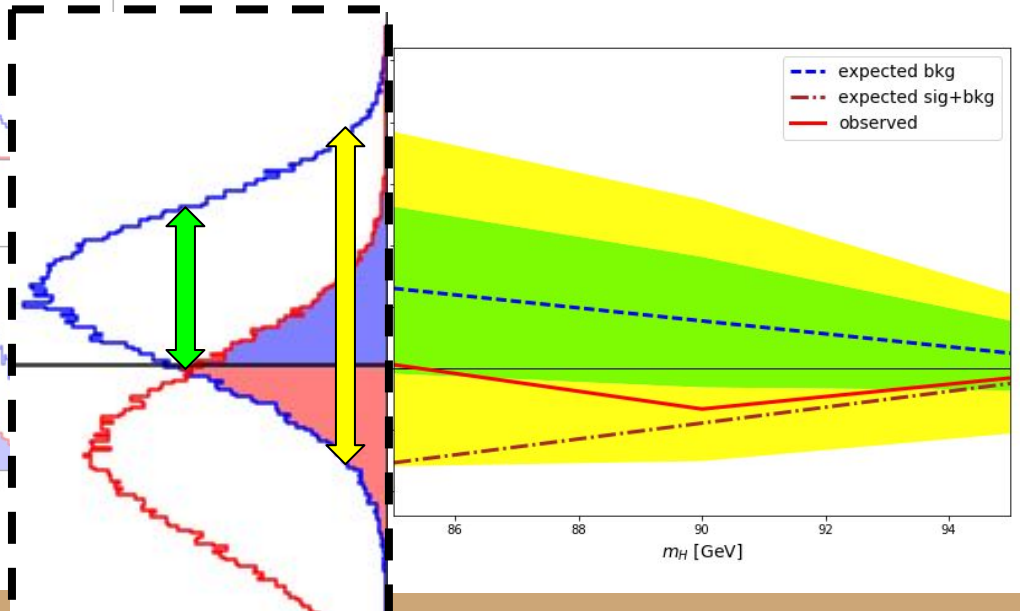








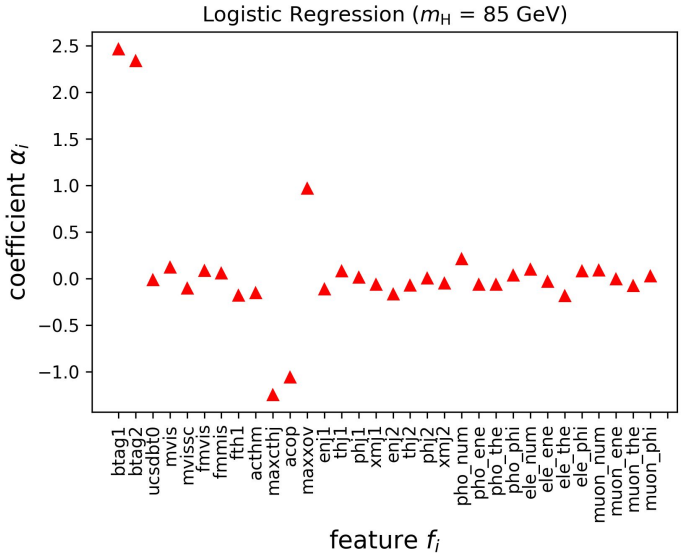
m_H	1-CL_s now	1-CL_s before
85 GeV	78.6 %	44.8 %
90 GeV	50.4 %	59.5 %
95 GeV	37.8 %	33.2 %



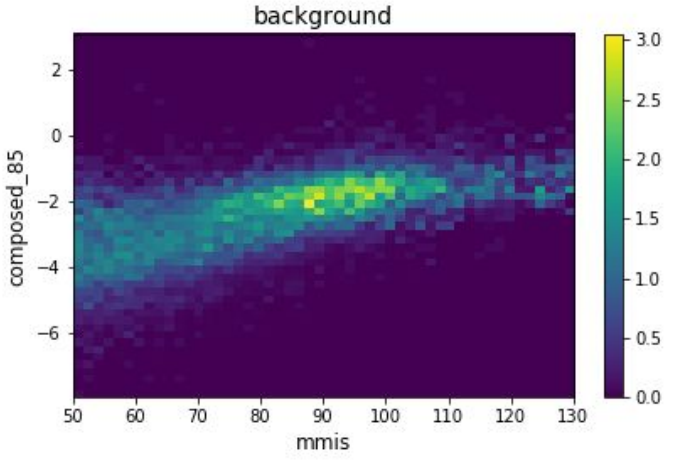
2D MVA analysis

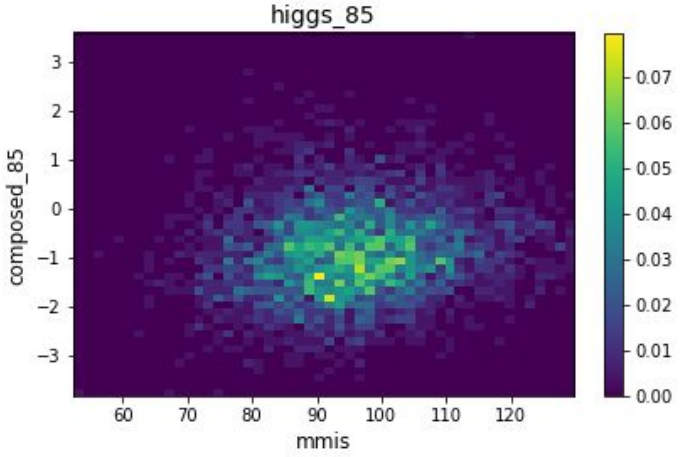
composed variable

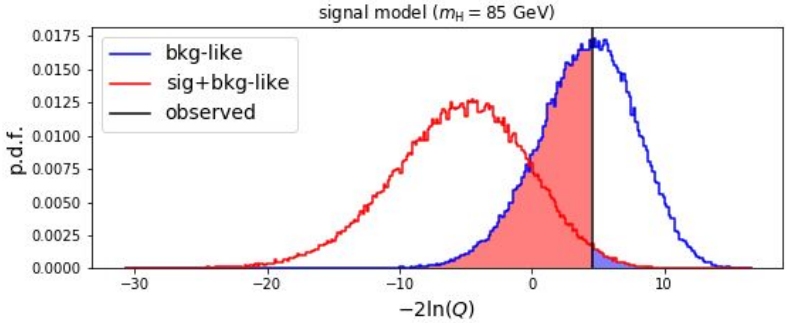
$$c = \sum_i \alpha_i f_i$$

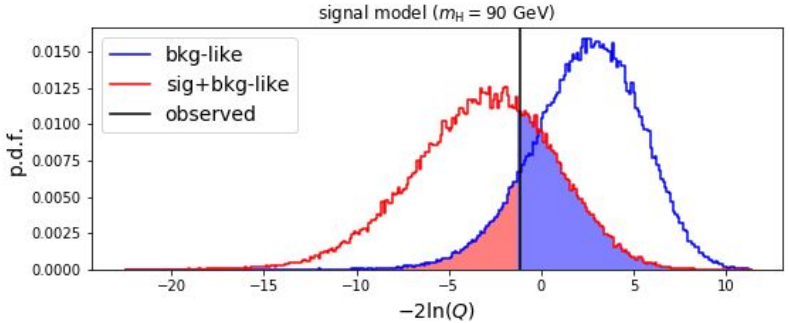


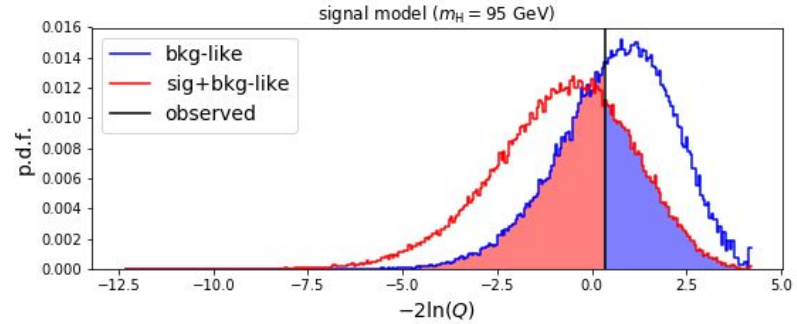
$$-2\ln(Q(m_H)) = 2s_{tot} - 2\sum_{i=1}^{N_{bins}} N_i \ln\left(1 + \frac{s_i(m_H)}{b_i}\right).$$

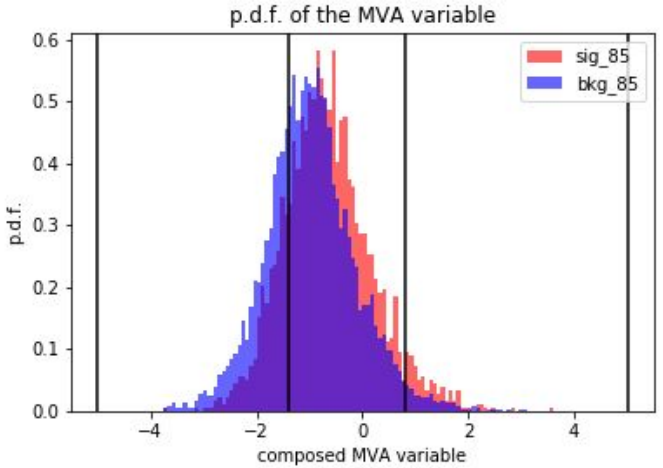






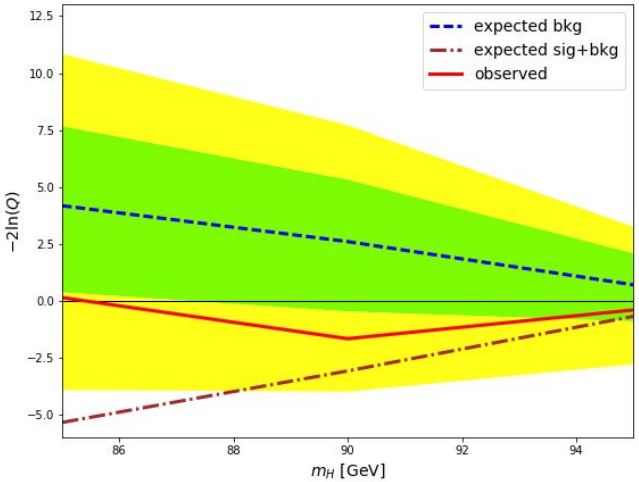






m_H	1-CL_s	# data events
85 GeV	96.1 %	32 (/641)
90 GeV	62.7 %	27
95 GeV	52.1 %	10

Conclusion



Approach		m_H= 85 GeV	m_H= 90 GeV	m_H= 95 GeV
1D Selection cuts by eye		44.8 %	59.5 %	33.2 %
1D BDT cuts		78.6 %	50.4 %	37.8 %
2D BDT and MVA for second variable		96.1 %	62.7 %	52.1 %

References

Higgs @ LEP:

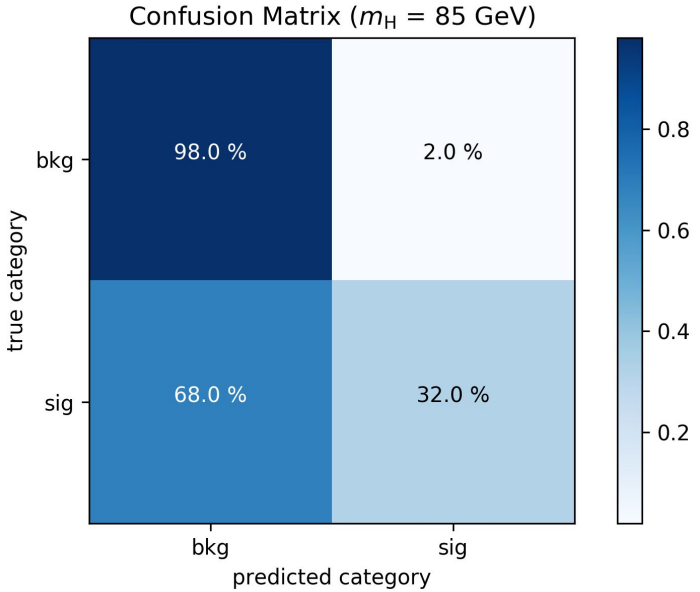
Gross, Eilam, and Amit Klier. "Higgs statistics for pedestrians." *arXiv preprint hep-ex/0211058* (2002).

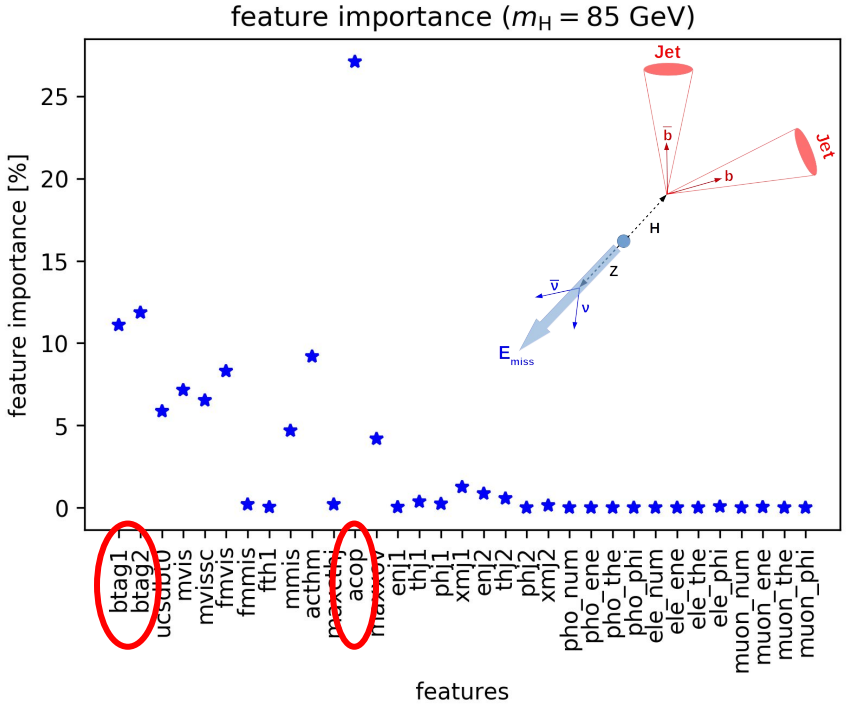
for Higgs, The LEP Working Group, et al. "Search for the standard model Higgs boson at LEP." *Physics Letters B* 565 (2003): 61-75.

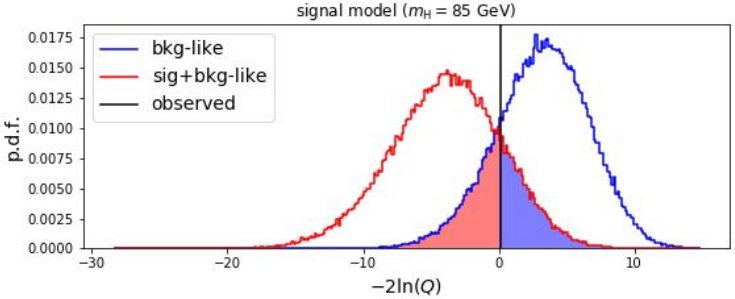
Machine Learning:

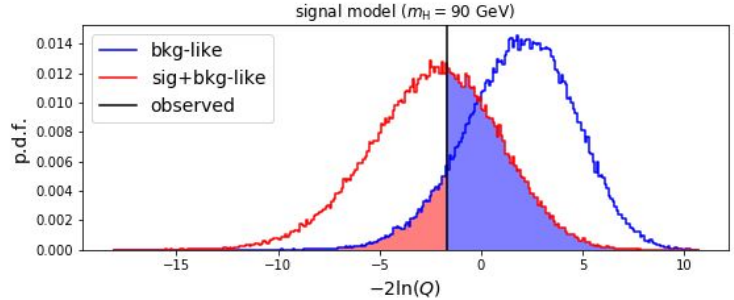
Müller, Andreas C., and Sarah Guido. *Introduction to machine learning with Python*. O'Reilly Media, 2017.

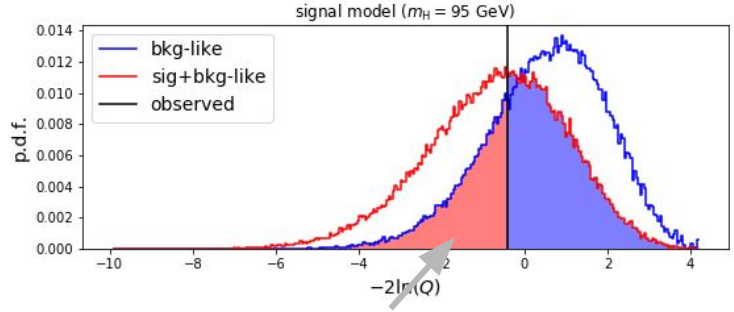
Backup


Confusion Matrix (m_H = 85 GeV)
98.0 %
2.0 %
68.0 %
32.0 %
bkg
sig
true category
bkg
sig
predicted category
0.8
0.6
0.4
0.2

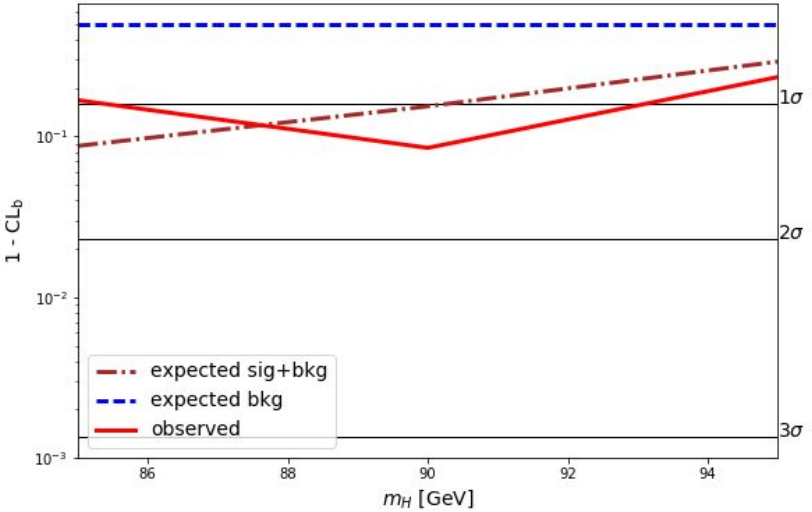

feature importance ($m_H = 85$ GeV)
feature importance [%]
features
btag1
btag2
ucsdbb0
mvis
mvissc
fmvis
fmmis
fth1
mmis
acthm
maxethj
acop
maxeov
enj1
thj1
phj1
xmj1
enj2
thj2
phj2
xmj2
pho_num
pho_ene
pho_the
pho_phi
ele_num
ele_ene
ele_the
ele_phi
muon_num
muon_ene
muon_the
muon_phi
Jet
Jet
H
Z
E miss





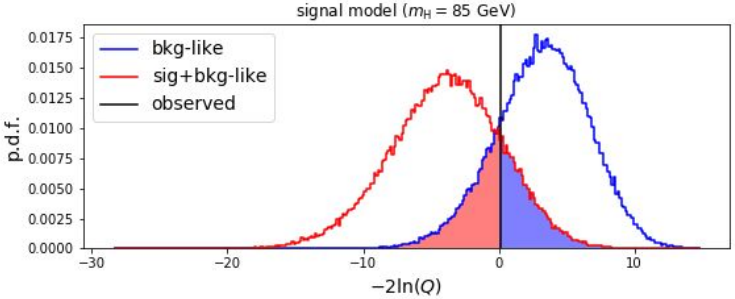


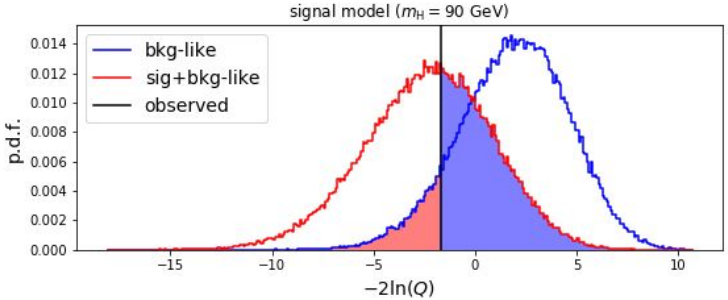
1-CL_b

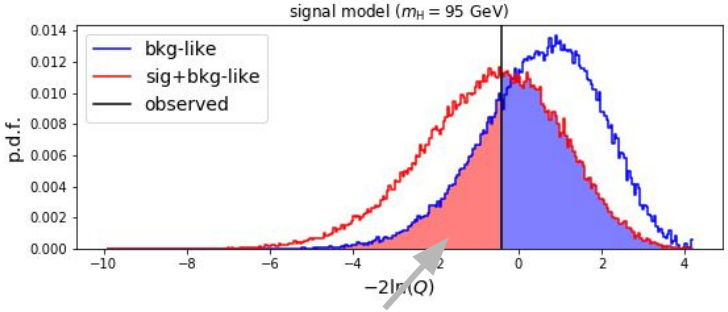


generate toys

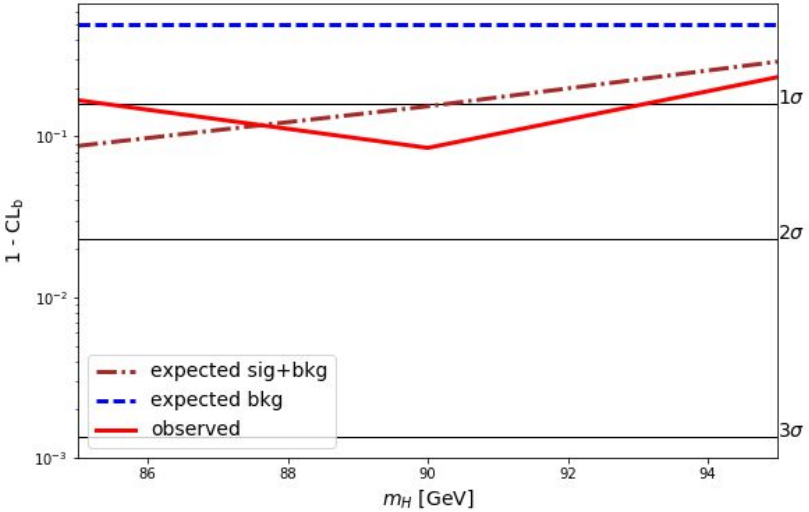
1) according to bkg and calculate 1-CL_b
2) according to sig+bkg and calculate 1-CL_b

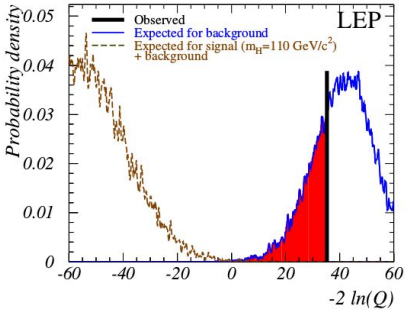

signal model (m_H = 85 GeV)
bkg-like
sig+bkg-like
observed
p.d.f.
−2ln(Q)

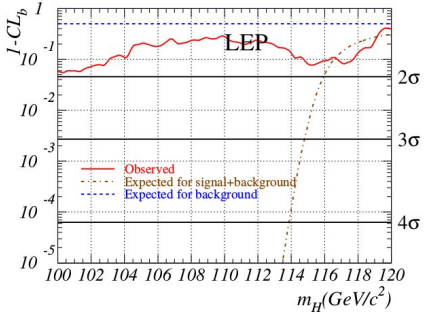

signal model (m_H = 90 GeV)
bkg-like
sig+bkg-like
observed
p.d.f.
−2ln(Q)


signal model (m_H = 95 GeV)
bkg-like
sig+bkg-like
observed
p.d.f.
−2ln(Q)

1-CL_b


1 - CL_b
expected sig+bkg
expected bkg
observed
1σ
2σ
3σ
m_H [GeV]


Observed
Expected for background
Expected for signal (m_H=110 GeV/c^2) + background
LEP
Probability density
-2 ln(Q)


LEP
Observed
Expected for signal+background
Expected for background
2σ
3σ
4σ
m_H(GeV/c^2)