

LHC ATLAS実験におけるbクォークの湯川結合の測定

CiRfSE

2017.1.24

Fumiaki Ito



2017/1/24

CiRfSE

1

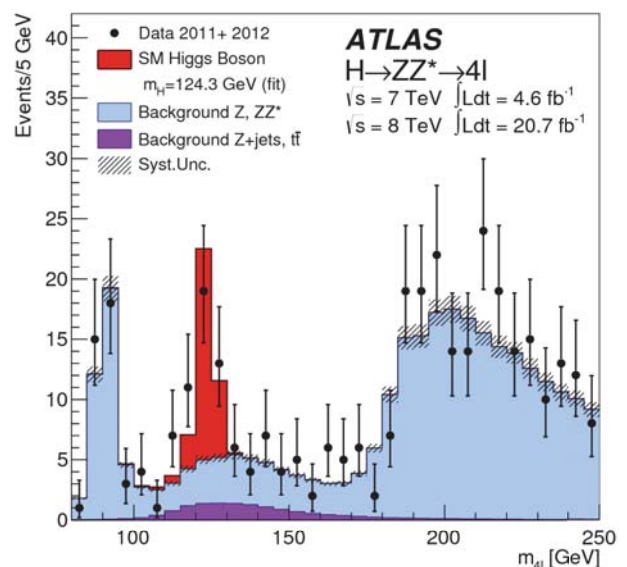
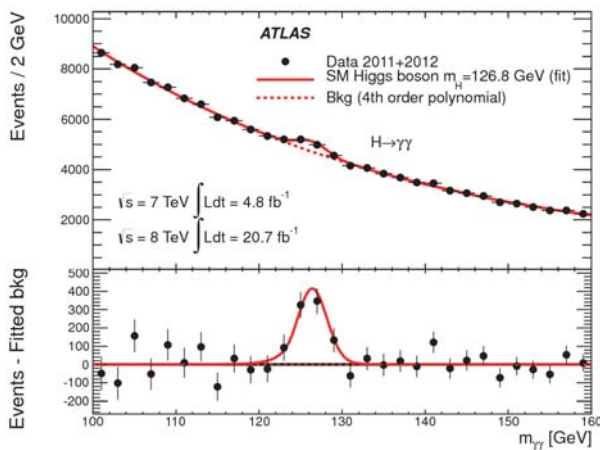


導入



なぜ大規模で、多くの資源と人材を要する実験を行っているのか?? (始められたのか?)

Higgs



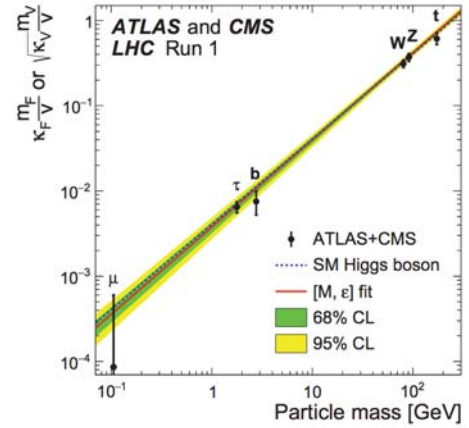
Higgs粒子発見

標準模型は大成功！！

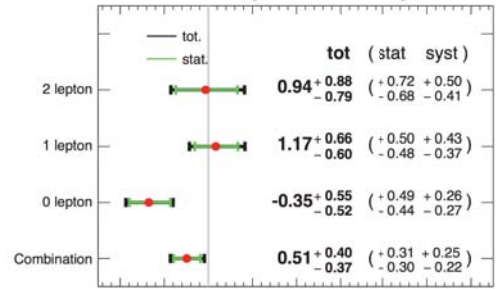
SM Higgsかどうかの検証

BSM探索

- ヒッグス粒子は物質に質量を与える粒子で、フェルミオンとの湯川結合定数は質量に比例
- 図のように比例関係を確認→標準模型の実証
- 比例関係のずれは→新たな物理の証拠
- 標準模型でLHCで近々発見可能な結合
➤ 残りbクォークとtクォーク



ATLAS $\sqrt{s}=7$ TeV, $\int \mathcal{L} dt=4.7$ fb $^{-1}$; $\sqrt{s}=8$ TeV, $\int \mathcal{L} dt=20.3$ fb $^{-1}$



VH associated production

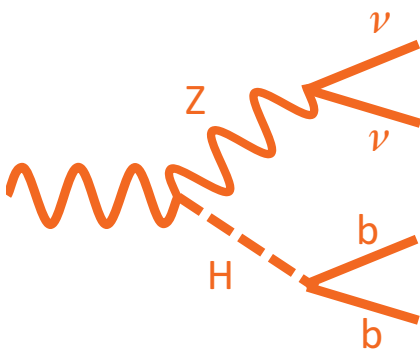
終状態の荷電レプトン数でチャンネルを分ける
W/Zボソンとの随伴生成過程でそれぞれ1/0,2レプトンチャンネルがある

最も感度のよい過程

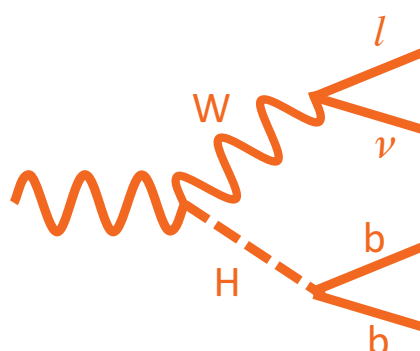
0-lepton chanel

1-lepton chanel

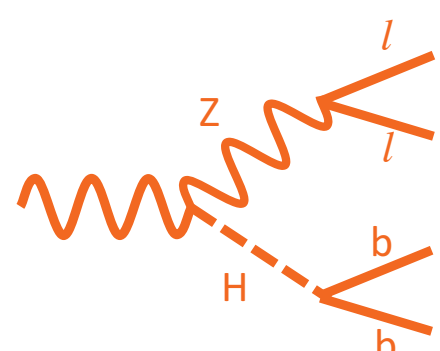
2-lepton chanel



MET
b jet × 2

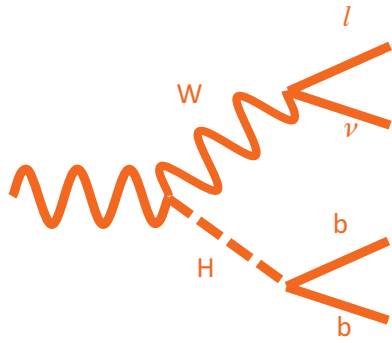


lepton(e,μ)
MET
b jet × 2



lepton(e,μ) × 2
b jet × 2

信号の終状態を定義(Wボソン 随伴生成)



荷電レプトン数やジェット数、消失横運動量(MET)、ベクターボソンの横運動量(p_{TV})、フレーバー同定などにより信号事象を選択
(Single electron trigger または MET trigger, ジェット = 2 or 3, レプトン = 1, ボトムクォークジェット = 2, MET)



背景事象の理解

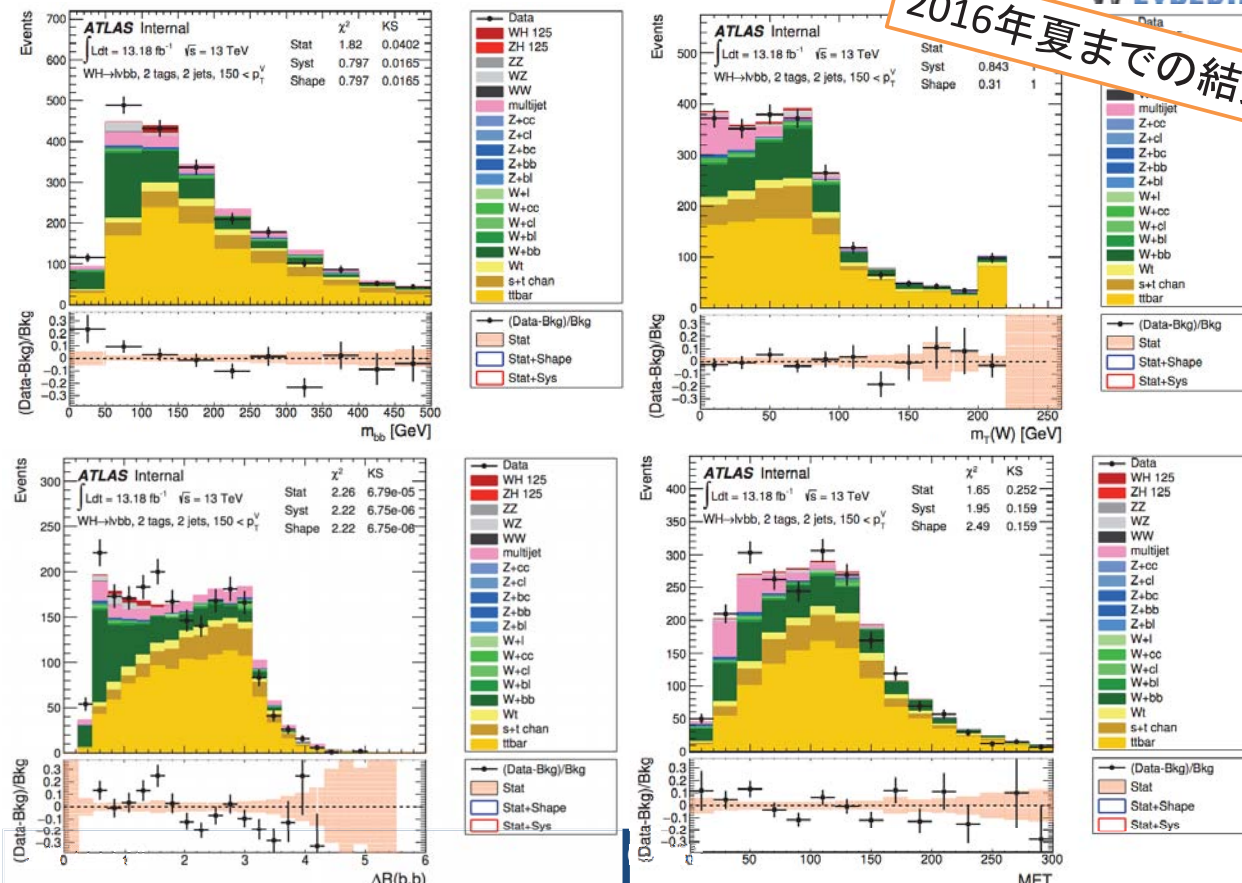


多変量解析(MVA)を用いて信号と背景事象の分離を最適化する

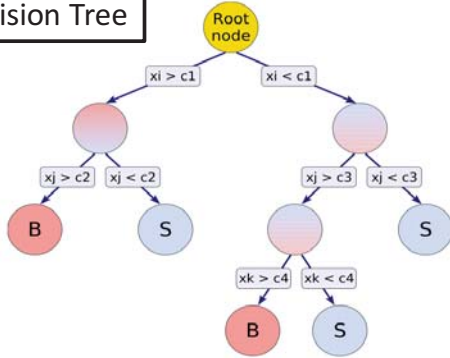


尤度関数(likelihood)を定義し、Fittingを行いsignificanceを計算する

2016年夏までの結果



Decision Tree



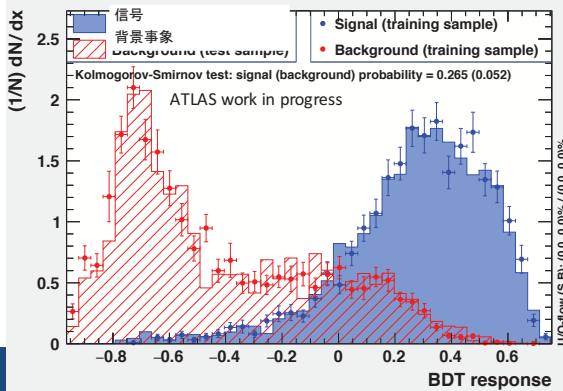
一変数による分離よりも有利である

10程の変数を使って解析する

人間が最適な分離を行うことは困難



MCシミュレーションデータを用いたBDTの学習結果とテスト結果



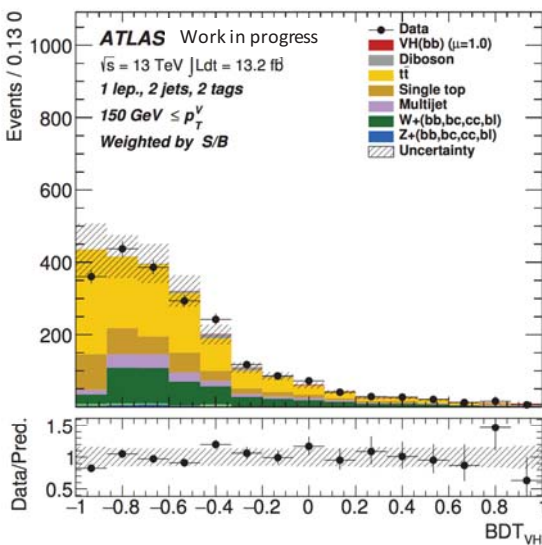
Boosted Decision Tree (BDT)

選別と学習速度のバランスがよい
最適化はBDTに肩代わりさせる

出力結果は信号と背景事象を分離させる

信号存在確率の比較

BDT output 分布



まずビン幅の最適化を行う

予想される信号存在確率(significance)を計算する

$$Z_{\text{exp}} = \sum_i \sqrt{2(S_i + B_i) \ln \left(1 + \frac{S_i}{B_i} \right) - 2S_i}$$

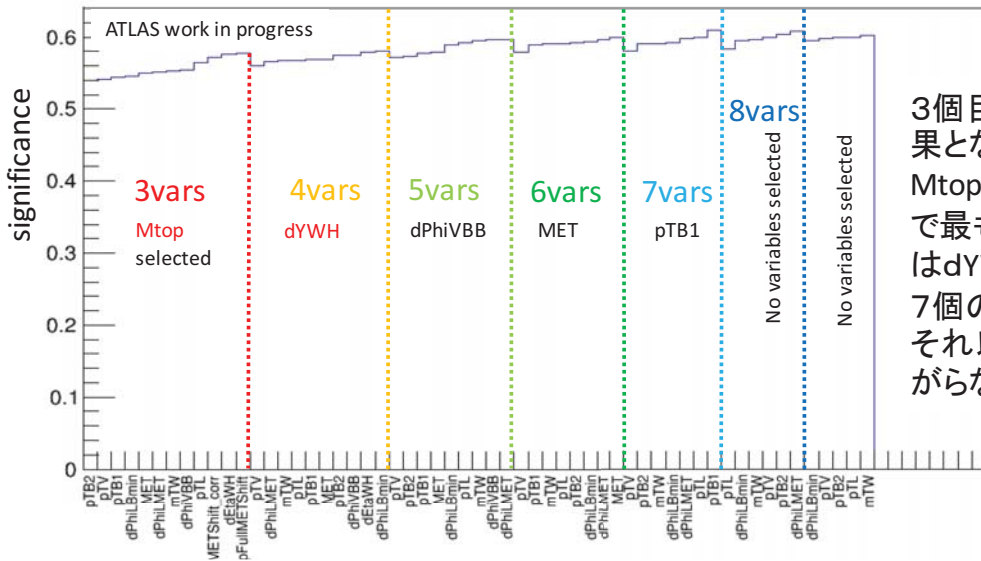
S_i : ビン毎の信号数

b_i : ビン毎の背景事象数

この信号存在確率を比較しBDTに使用する
変数の順位付けを行う

選ばれた変数 (選ばれた順番に記述)

mBB, dRBB, **Mtop**, **dYWH**, dPhiVBB, MET, pTB1 (, mBBJ, pTJ3)



3個目の変数で最もよい結果となったのはMtop、次にMtopを加え4個目の変数で最もよい結果となったのはdYWH(3 jetと合わせて)。7個の変数を選んだ時点でそれ以上significanceが上がらなくなる。

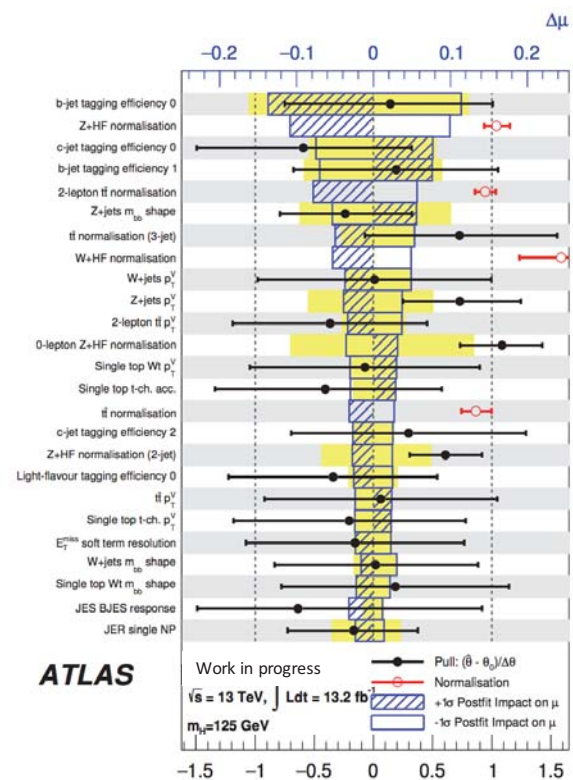
最大予想信号存在確率: 0.61σ

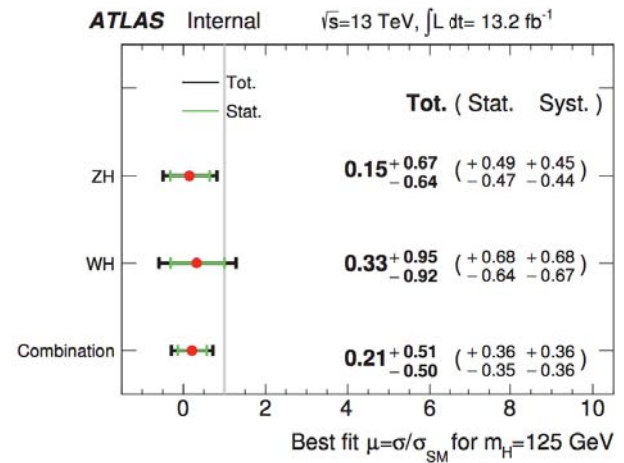
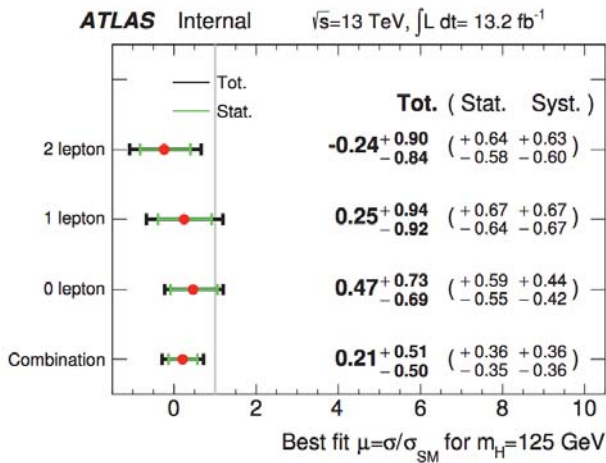
どの変数を使うかの決定を行う

*3.21 -fb(2015年)での統計誤差のみを考慮した値
(この時点で8 %ほどの向上が得られた)

Signal strength and NP ranking

| Setup | Impact on error |
|-------------------------|-----------------|
| Total | +0.508 |
| DataStat | -0.499 |
| FullSyst | +0.361 |
| Floating normalizations | -0.346 |
| All normalizations | +0.358 |
| All but normalizations | -0.36 |
| Jets MET | +0.0985 |
| BTag | -0.15 |
| Leptons | +0.126 |
| Luminosity | -0.171 |
| Diboson | +0.315 |
| Zjets | -0.32 |
| Wjets | +0.0504 |
| Model tbar | -0.0462 |
| Model Single Top | +0.162 |
| Model Multi Jet | -0.19 |
| Signal Systematics | +0.00997 |
| MC stat | -0.0105 |
| $\hat{\mu}$ | +0.0173 |
| | -0.011 |
| | +0.0225 |
| | -0.0217 |
| | +0.118 |
| | -0.179 |
| | +0.0968 |
| | -0.136 |
| | +0.0903 |
| | -0.145 |
| | +0.0421 |
| | -0.0309 |
| | +0.0155 |
| | -0.0165 |
| | +0.0813 |
| | -0.0275 |
| | +0.208 |
| | -0.215 |
| $\hat{\mu}$ | 0.210 |





信号強度は予想よりもだいぶ低く出ている。

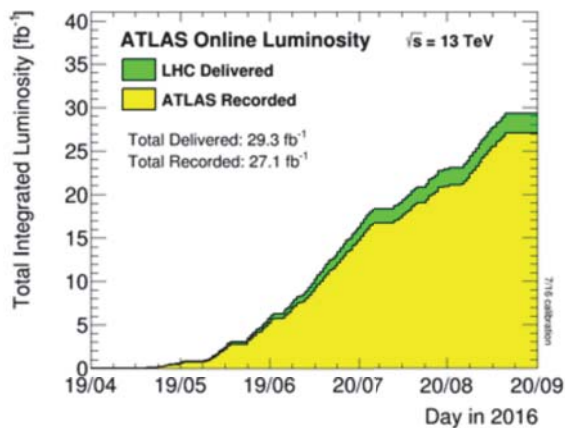
- Validation 用に同じ解析手法でダイボソンの信号強度も観測

WW/WZ : 0.91 ± 0.17
Significance : 3.0

データの統計量的(37 fb $^{-1}$)には今年の春の結果でRun2のみで3 σ に到達(expected)

MCの統計量を増やすことで現状の解析から10 %程度の感度向上が見込まれる
(また様々な改良を行っている)

まとめ



LHC実験の大きな目標の一つであったHiggs粒子が見つかった現在、その性質を調べることは重要なテーマである

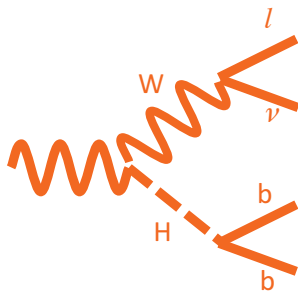
Higgs粒子とb-quarkとの湯川結合の測定は其中でも重要なテーマの一つ

最新の公式結果では信号強度はSMと比較しかなり下振れしている

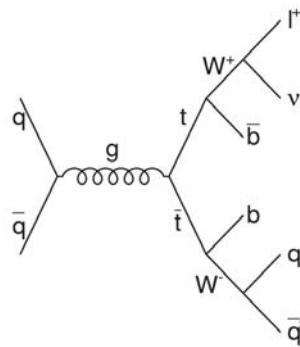
今春の結果に期待！

主な背景事象(1 lepton)

信号(Wボソン随伴生成)

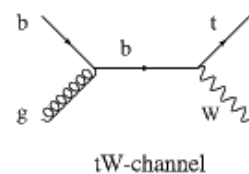
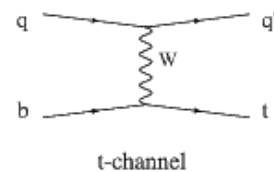


トップクオーク対生成



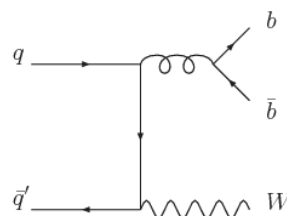
レプトンやMET、ボトムクオークジェット = 2 を要求しても残って来ってしまう。

シングルトップ



MCを使用して見積もる
様々なジェネレータを用いて
誤差を見積もる

Wボソン + ジェット

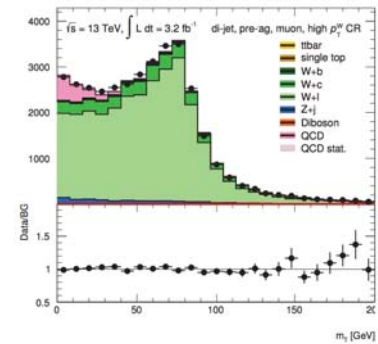
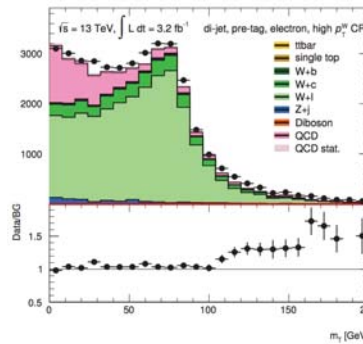
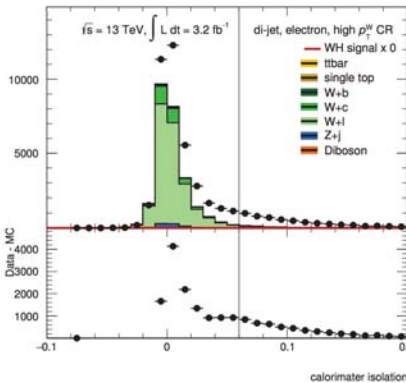


QCDの見積もりには
Fake factor methodを使用

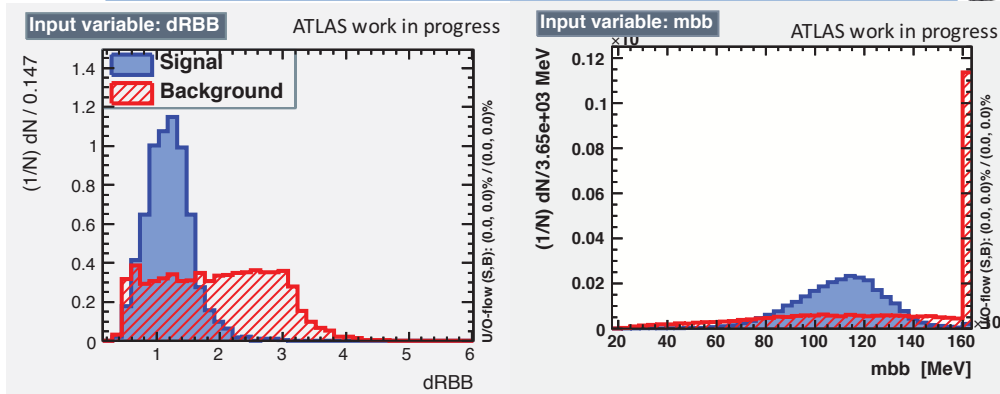
di jet CRを用意
Fake factorをpt, eta(MET) bin 毎に出す
QCD shapeを見積もる

$$f = \frac{N_{dijet}^{Numerator}}{N_{dijet}^{Denominator}}$$

$$N_{WH}^{Numerator} = f \times N_{WH}^{Denominator} = \frac{N_{dijet}^{Numerator}}{N_{dijet}^{Denominator}} \times N_{WH}^{Denominator}$$



どのように変数を選択するか



主な背景事象であるttbarはdRが大きくなる傾向があり、W+jetは小さくなる傾向がある

SMヒッグス粒子であれば、125 GeV付近にピークがある

単一変数において最も信号と背景事象の分離能力の高い、mBB, dRBBという二つの変数をまず選択する

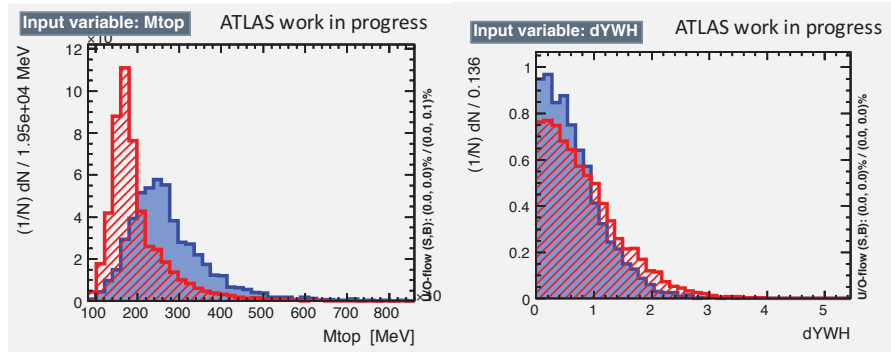
その後一つずつその他の候補となる変数を合わせて多変量解析を行い、最もよい結果となった変数を3番目の変数とする

同じことを4番目以降の変数に対しても行い、予想信号存在確率が向上しなくなるまで繰り返す。

使用する変数の候補

mBB,
dRBB,
dYWH,
Mtop,
pTV,
dPhiLBmin,
dPhiVBB,
pTB1,
pTB2,
dPhiLMET,
pTL,
mTW,
MET,
HT
(only for 3 jets)
(mBBJ)
(pTJ3)

変数の候補はRun1を踏襲
dYWHとMtopは新しく足した候補変数



- dYWH
Wボソンとヒッグスボソンのラピデティー差
- Mtop
Wボソンとbジェット一つの不変質量

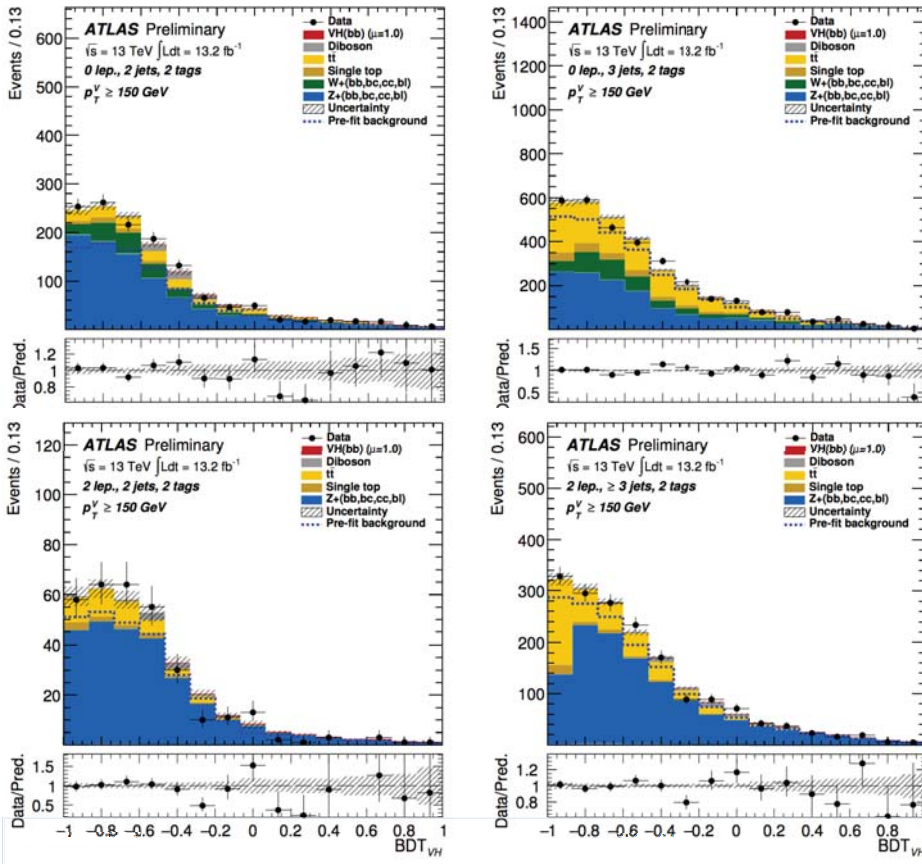
ICHEPで用いた変数群

夏の国際会議(ICHEP)用に使用する変数を決める時点ではMC統計が十分ではなく、ICHEPの時点ではデータ量が増え、さらにtruth taggingという手法を導入することで、統計もさらに稼げるため、より多くの変数が選ばれる方が有利になる。



ICHEPではRun1で使用した変数群に、MtopとdYWHを追加することとなる。

| Variable | Name | 1-lepton |
|--|-----------|----------|
| p_T^V | pTV | ✓ |
| E_T^{miss} | MET | ✓ |
| p_T^{jet1} | pTB1 | ✓ |
| p_T^{jet2} | pTB2 | ✓ |
| m_{jj} | mBB | ✓ |
| $\Delta R(\text{jet}_1, \text{jet}_2)$ | dRBB | ✓ |
| $ \Delta\eta(\text{jet}_1, \text{jet}_2) $ | dEtaBB | |
| $\Delta\phi(V, H)$ | dPhiVBB | ✓ |
| $\Delta\eta(V, H)$ | dEtaVBB | |
| $M_{eff}(M_{eff3})$ | HT | |
| $\min(\Delta\phi(\ell, \text{jet}))$ | dPhiLBmin | ✓ |
| m_T^W | mTW | ✓ |
| m_{ll} | mLL | |
| $\Delta Y(W, H)$ | dYWH | ✓ |
| m_{top} | mTop | ✓ |
| Only in 3 Jet Events | | |
| p_T^{jet3} | pTJ3 | ✓ |
| m_{jjj} | mBBJ | ✓ |

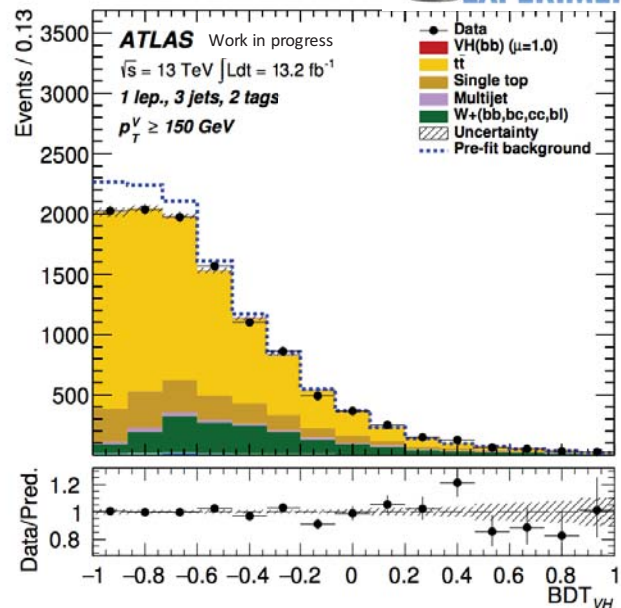
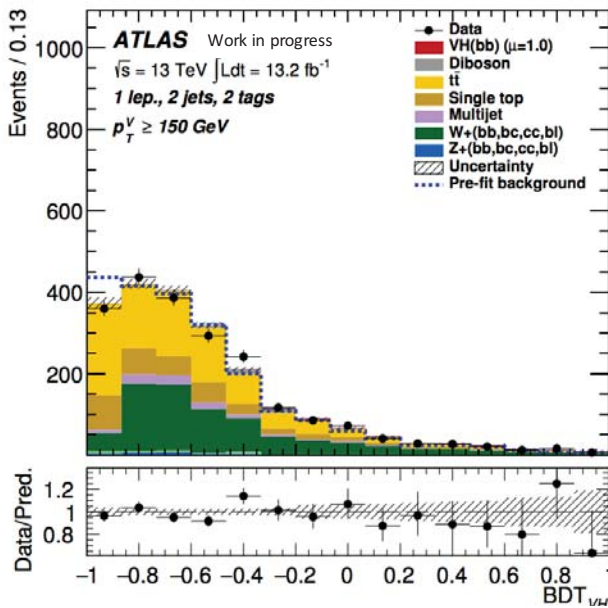
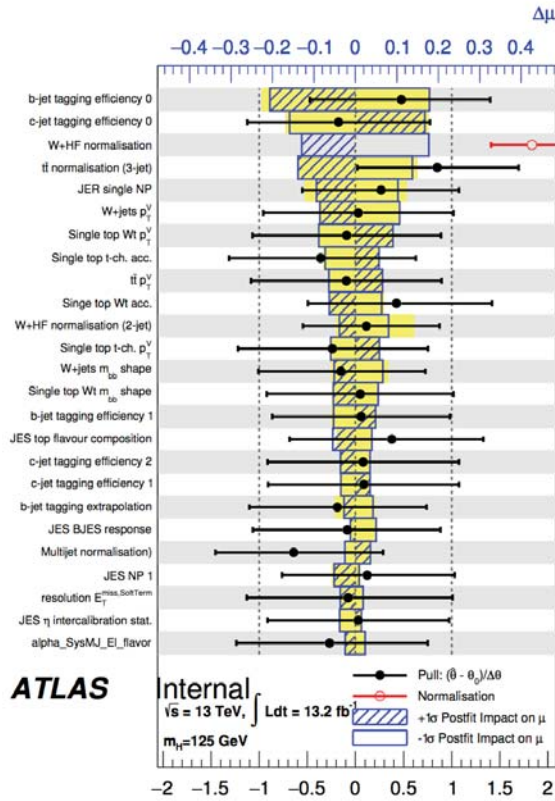


$$\text{Pois}(n|\mu S + B) \left[\prod_{b \in \text{bins}} \frac{\mu v_b^{\text{sig}} + v_b^{\text{bkg}}}{\mu S + B} \right]$$

- n 全実データイベント数
- μ 信号強度
- v_b^{sig} ビン毎の予想信号イベント数
- v_b^{bkg} ビン毎の予想背景事象イベント数
- S 予想信号イベント数
- B 予想背景事象イベント数

v_b^{sig} v_b^{bkg} は nuisance parameters (NP) の関数

NPは基本的には正規分布の確率密度関数を与える



| Channel | Observed Significance | Expected Significance | Observed p-value | Expected p-value |
|----------|-----------------------|-----------------------|------------------|------------------|
| 0-lepton | 1.02 | 1.45 | 0.15 | 0.07 |
| 1-lepton | 0.10 | 1.04 | 0.46 | 0.15 |
| 2-lepton | -0.17 | 1.14 | 0.57 | 0.13 |
| Combined | 0.42 | 1.94 | 0.34 | 0.03 |

ビンを細くすればビン毎に足しあげる信号感度は高くなるが、誤差の影響が大きくなる
➤ ちょうどよいビン幅に設定する必要がある

$$Z = z_s n_s / N_s + z_b n_b / N_b$$

N_s : 信号の全MCイベント数

N_b : 背景事象の全MCイベント数

n_s : 信号のあるビンからあるビンまでのMCイベント数

n_b : 背景事象のあるビンからあるビンまでのMCイベント数

z_s, z_b : フリーパラメータ(10,5 が今回の設定値)

このZ が1以上になるようにビン幅を設定する

AdaBoost(adaptive boosting)

A single decision tree can be sensitive to statistical fluctuations in the training data. The idea of combining many, or a forest of decision trees was proposed to overcome this stability problem.

The original event weights are used for the training of the first tree while subsequent trees are multiplied by a common boost weight give by $\alpha = 1 - \chi$ where χ is the sum of weights of the misclassified events divided by the total sum of weights.

AdaBoost(adaptive boosting)

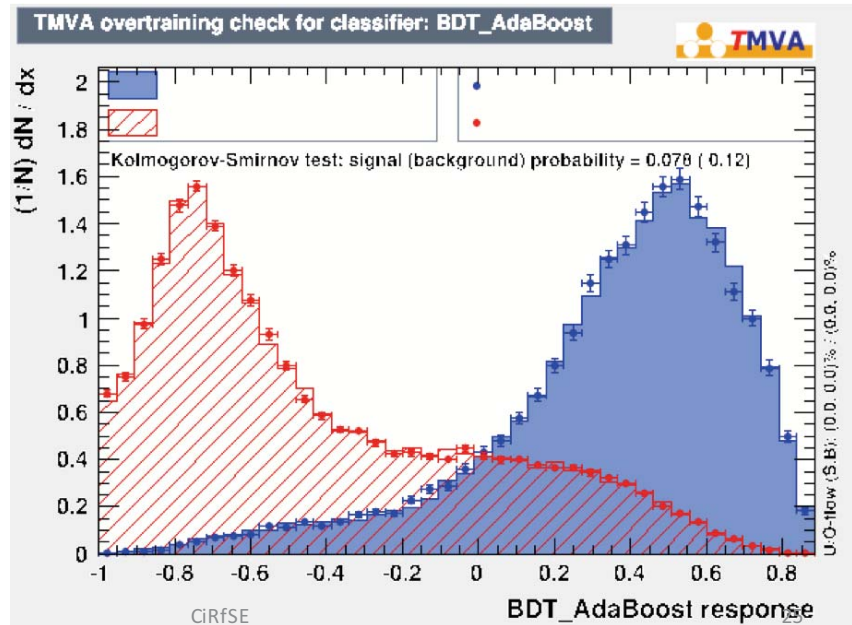
$$BDT_{response} = \frac{1}{N} \sum_i^N \ln(\alpha_i) \cdot h_i$$

N : Num of Trees

α_i : Weights

h_i : $h=1(\text{signal})$

$h=-1(\text{bg})$

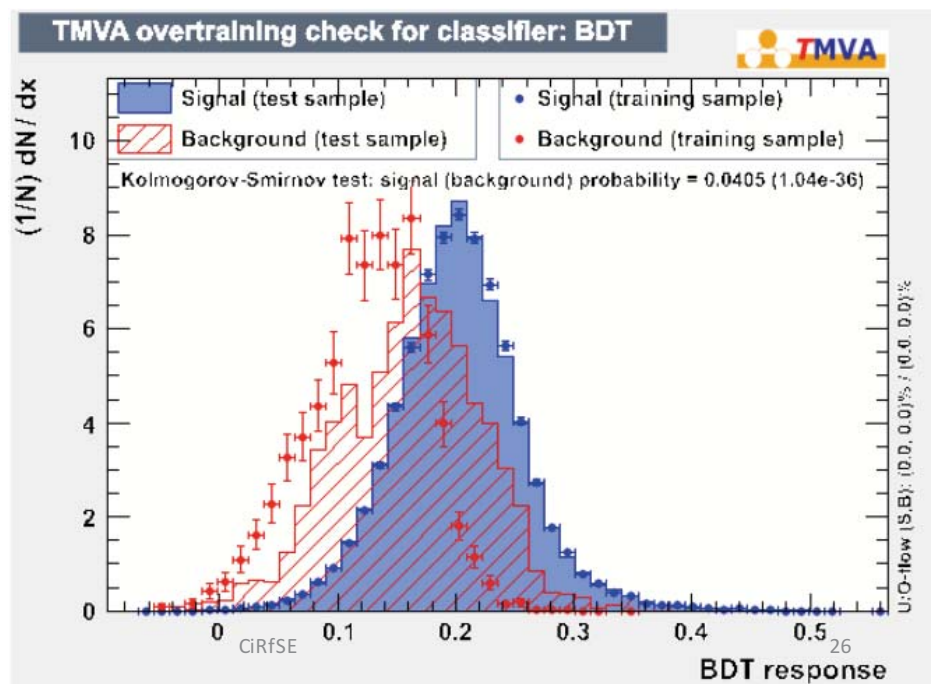


2017/1/24

Overtraining

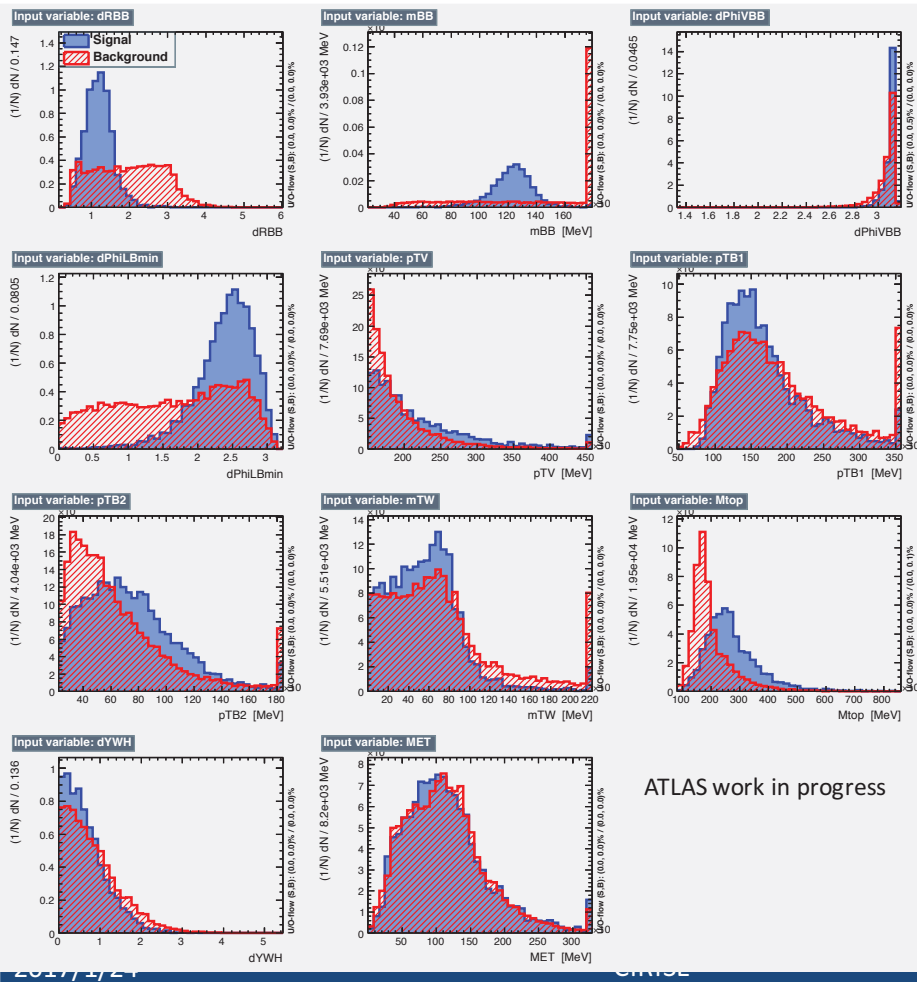
Sensitivity to fluctuations in the training data is known as over-training.

A simple way to measure the amount of over-training is to compare the performance results between the independent training and test samples.



2017/1/24

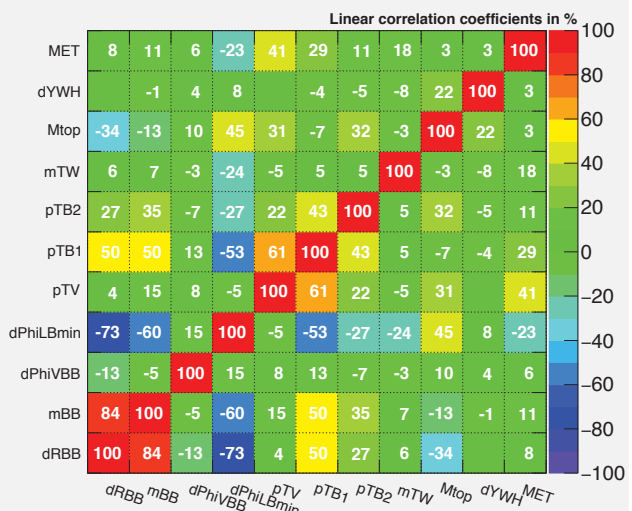
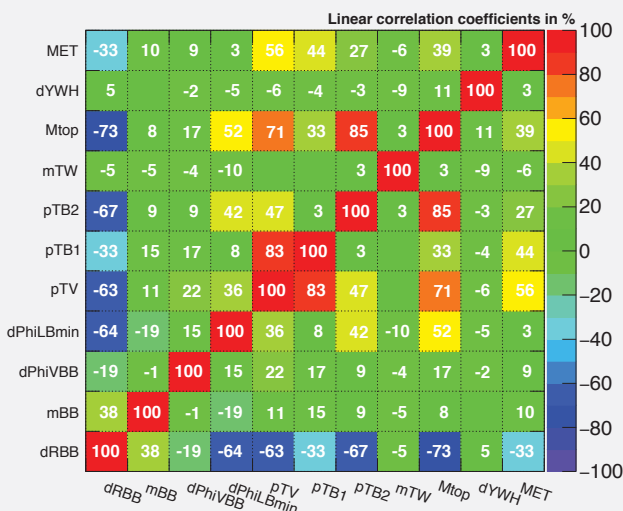
変数分布



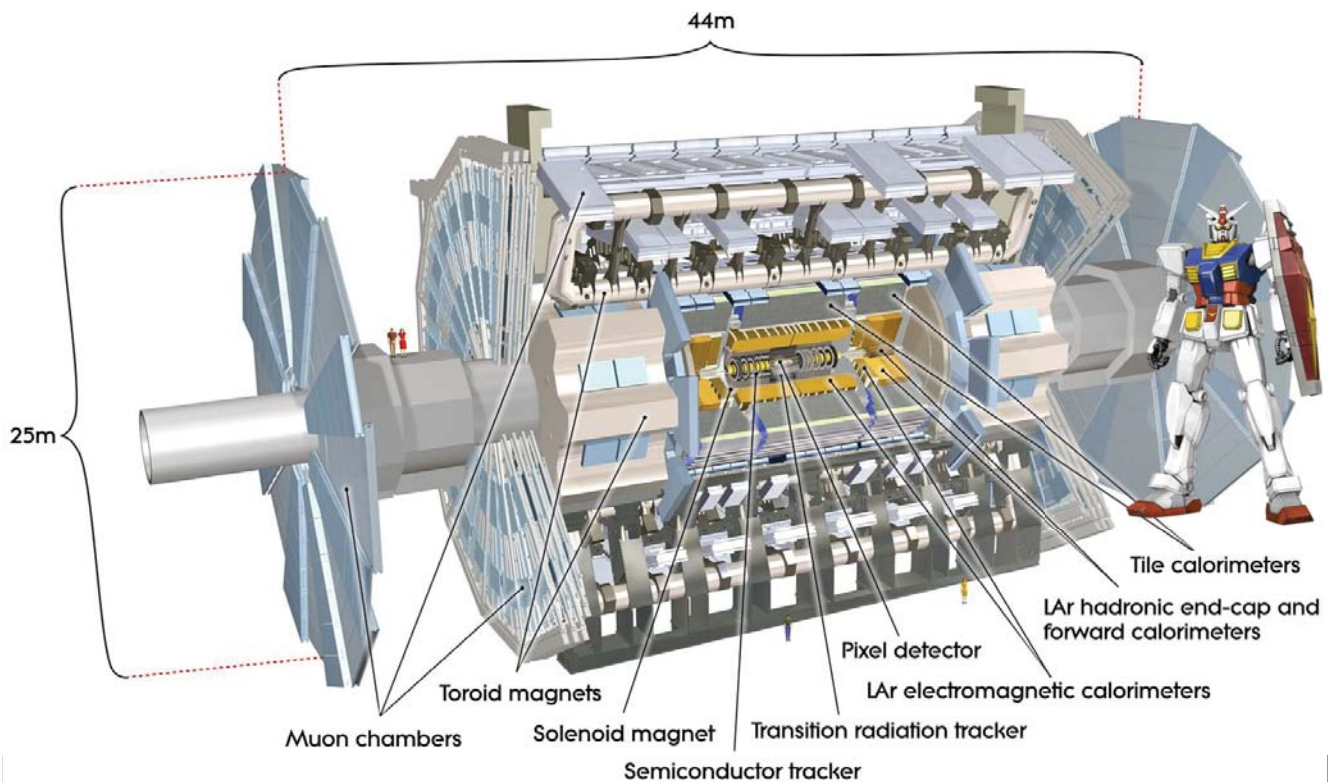
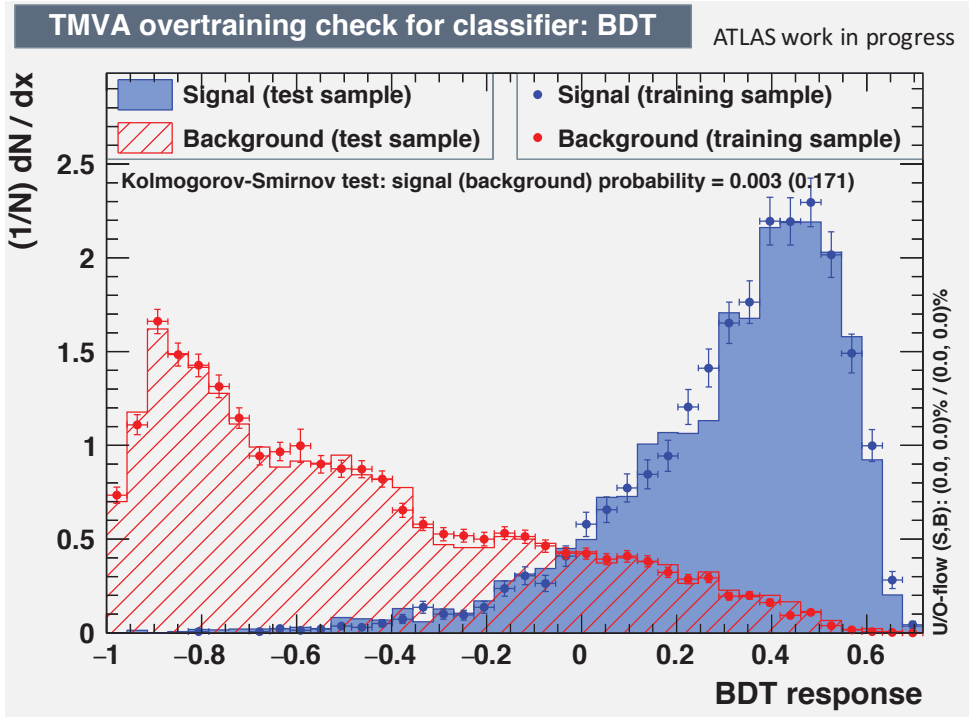
変数同士の1次元相関

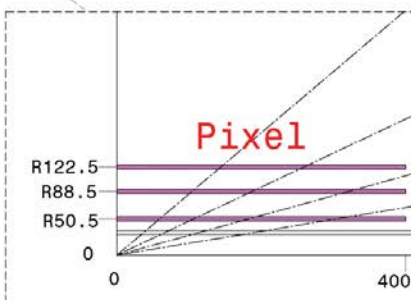
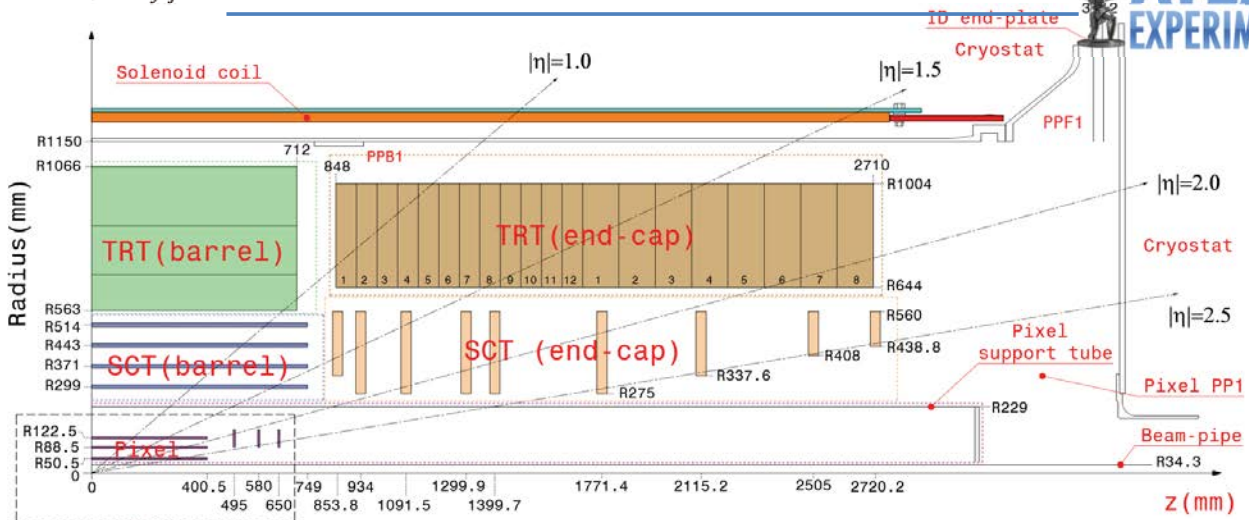
Correlation Matrix (signal)

Correlation Matrix (background)



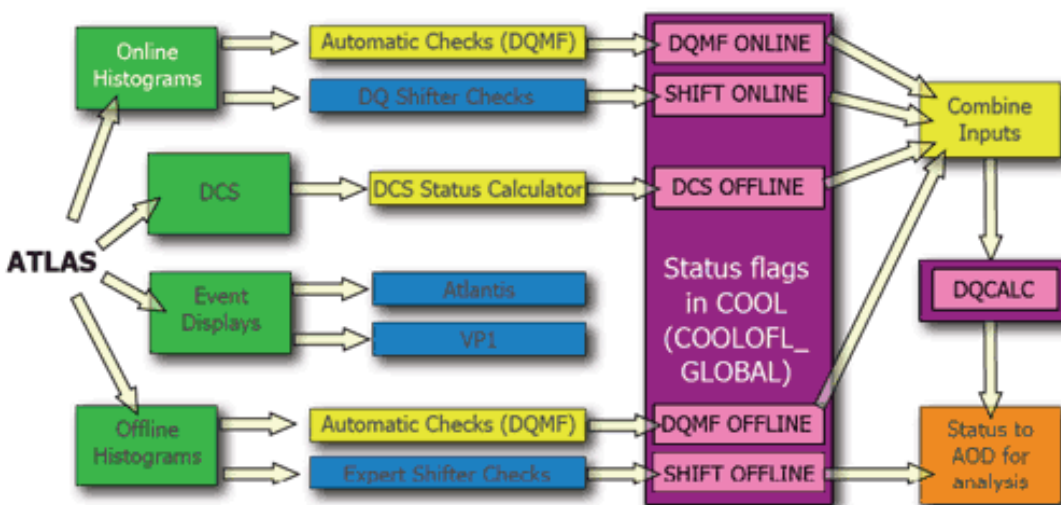
ATLAS work in progress





| Detector component | Required resolution | η coverage |
|-----------------------------|--|----------------------|
| Tracking | $\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$ | ± 2.5 |
| EM calorimetry | $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ | ± 3.2 |
| Hadronic calorimetry (jets) | $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ | ± 3.2 |
| barrel and end-cap forward | $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ | $3.1 < \eta < 4.9$ |
| Muon spectrometer | $\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV | ± 2.7 |

Framework



In DQMF, data quality is assessed by checking histogram content and comparing against reference histograms both online and offline. The results of the checks, along with information from the Detector Control System (voltage, temperatures, etc.), are used to determine the status of each detector



↑ 一方向からだけでは色付きの領域と白い領域を分離できない

| | 物質粒子 | | | 力を伝える粒子 | |
|------------|-------------|-------------|-------------|--|--|
| | 第1世代 | 第2世代 | 第3世代 | | |
| クォーク | アップ | チャーム | トップ | 強い力 グルーオン 電磁力 光子 弱い力 Wボゾン Zボゾン | |
| | ダウン | ストレンジ | ボトム | | |
| レプトン | eニュートリノ | μニュートリノ | τニュートリノ | | |
| | 電子 | ミューオン | タウ | | |
| ヒッグス場に伴う粒子 | | | | ヒッグス粒子 | |

現在の素粒子物理学の枠組み非常に成功している理論体系。これは、自然界に存在する電磁相互作用・弱い相互作用・強い相互作用・重力相互作用の4つの内重力相互作用以外の3つの相互作用を非常に良く記述している。標準理論を構成する粒子は、主に3種類に分けられる。物質を構成する粒子、 γ 相互作用を媒介する粒子、他の素粒子に質量を持たせると言われているヒッグス粒子である。

ラグランジアン密度を

$$L_{QED} = |D_\mu \phi|^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - V(\phi)$$

$$D_\mu = \partial_\mu + ieA_\mu \quad , \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

A_μ は電磁場、 ϕ はスカラー場

U(1)ゲージ変換に対して

$$\begin{cases} \phi(x) \rightarrow \phi'(x) = e^{i\alpha(x)} \phi(x) \\ A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) - \frac{1}{e} \partial_\mu \alpha(x) \end{cases}$$

不変にしようとする、質量項の導入が許されない
しかし、実際にはW/Zは質量を持つ
なぜか??

SU(2)まで考慮したラグランジアン密度

$$L = |D_\mu \phi|^2 - \frac{1}{4} A_{\mu\nu}^a A^{\mu\nu a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - V(\phi) \quad , \quad D_\mu = \partial_\mu + igA_\mu^a \tau^a - i\frac{g'}{2} B_\mu$$

ヒッグス場

$$\begin{aligned} \varphi(x) = \frac{1}{2} \begin{pmatrix} 0 \\ v+h(x) \end{pmatrix} & \quad |D_\mu \phi|^2 \\ & = \frac{1}{2} \begin{pmatrix} 0 \\ v+h(x) \end{pmatrix}' (\partial_\mu + gA_\mu^a \tau^a + \frac{1}{2} g' B_\mu) (\partial_\mu + gA_\mu^b \tau^b + \frac{1}{2} g' B_\mu) \begin{pmatrix} 0 \\ v+h(x) \end{pmatrix} \\ & = \frac{1}{2} (\partial_\mu h(x))^2 + \frac{1}{8} (v+h(x))^2 [g^2 (A_\mu^1)^2 + g^2 (A_\mu^2)^2 + (-gA_\mu^3 + g' B_\mu)^2] \end{aligned}$$

ウィークボソンの場を

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (A_\mu^1 \mp A_\mu^2)$$

$$Z_\mu(x) = \frac{1}{\sqrt{g^2 + g'^2}} (gA_\mu^3 + g'B_\mu)$$

$$L = \frac{1}{2} (\partial_\mu h(x))^2 + (1+h(x)/v) [m_W^2 W + \mu W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu]$$

自発的対称性の破れによって真空期待値 v を持ち、これによりゲージ不変性を維持しながら、ウィークボソンに質量を持たせる