

LHC-ATLAS 実験の $H \rightarrow ZZ^* 4\ell$ チャンネルにおける
多変数解析手法による Higgs 粒子の探索

東京大学大学院
理学系研究科物理学専攻
田中研究室

崔原碩

平成 25 年 1 月 7 日

概要

スイス、ジュネーブの欧州原子核研究機構 (CERN) にある Large Hadron Collider (LHC) は 2010 年から重心系エネルギー 7 TeV で本格的に稼働し、2012 年からは 8 TeV で実験を行った。LHC では、新しい物理の探求の幕開けのため、ヒッグス粒子を発見し標準理論を完成させることと標準理論を越えた物理、たとえば、超対称性粒子 (SUSY) を発見することが期待されてきた。そしてデータの解析においてヒッグス粒子や SUSY の探索を精力的に行い、2012 年 7 月にヒッグス粒子と思われる新粒子 (質量約 126 GeV) を発見した。

本研究では、LHC 実験の検出器の中で ATLAS (A Toroidal LHC ApparatuS) 検出器という汎用型粒子検出器で 2012 年 3 月から 8 月までに取得した重心系エネルギー 8 TeV のデータ、積分ルミノシティ 13 fb^{-1} のデータを用いて、軽い質量、110~180 GeV の標準理論のヒッグス粒子の探索を $H \rightarrow ZZ^* \rightarrow 4\ell$ ($\ell = e, \mu$) チャンネルにおいて行った。このチャンネルの特徴は、他のチャンネルと比較するとバックグラウンドが少ないことであり、終状態が同じでヒッグス粒子を介さない $ZZ \rightarrow 4\ell$ が主なバックグラウンドプロセスとなる。

本研究では、カットベースでシグナルと区別がつかないこのバックグラウンドを、レプトンやそれらから構成される Z ボソン候補の角度を用いた多変数解析手法 (Multi-Variate Analysis, MVA) を適用することでその違いを見だし、排除してより高い感度の測定を行うことを目標としている。

最終的には、質量 125 GeV における期待される発見感度にカットベースと比較して MVA では約 1.7% の改善があった。実データではより高い感度の結果には至らなかったが、これはバックグラウンドの排除力に極めて大きい改善がなかったために、統計量の少ない実データにおいて 120~130 GeV の間でイベントの減少が観測されたためである。

実データからヒッグス粒子質量 124 GeV に最大のイベント超過が観測され、その大きさは 3.8σ であった。ヒッグス粒子の発見までには至らなかったが 3σ 以上の兆候を観測した。また、130~165 GeV の領域において標準理論ヒッグス粒子の存在を 95% 信頼度で棄却した。

目次

1	序論	7
1.1	標準理論	7
1.2	Large Hadron Collider (LHC) 実験	7
1.3	Higgs 粒子	9
1.3.1	Higgs 粒子の生成過程	9
1.3.2	Higgs 粒子の崩壊過程	10
1.4	$H \rightarrow ZZ^* \rightarrow 4\ell$ 解析	11
1.5	本論文の構成	11
2	ATLAS 検出器	13
2.1	Inner Detector	14
2.2	Calorimeter	16
2.2.1	EM calorimeter	16
2.2.2	Hadron calorimeter	17
2.2.3	Forward calorimeter(FCAL)	18
2.3	Muon Spectrometer	20
2.4	Magnet System	21
2.5	Trigger	22
3	Data と Monte Carlo sample	24
3.1	実験の Data	24
3.2	Background Process	25
3.3	Monte Carlo Simulation	25
3.4	MC Samples	26
3.4.1	Background Samples	26
3.4.2	Signal Samples	27
4	解析	30
4.1	Object Definition	30
4.1.1	Electron Reconstruction	30
4.1.2	Electron Selection	33
4.1.3	Muon Reconstruction	34
4.1.4	Muon Selection	37
4.2	Event Selection	39
4.3	Background Estimation	43
4.3.1	$Z+X_\mu X_\mu$ background	44
4.3.2	$Z+X_e X_e$ background	48
4.3.3	Cut based 解析の結果	56
4.4	Multivariate Analysis	57
4.4.1	TMVA と BDT	57
4.4.2	MVA の input variable	57

4.4.3	MVA の結果	60
4.5	Exclusion Limit と p-value	63
4.5.1	Profile Likelihood	63
4.5.2	Test static	63
4.5.3	Discovery of signal	64
4.5.4	Confidence Level	64
5	結果と考察	66
5.1	解析結果	66
5.2	考察	69
6	まとめ	72
A	BDT method	73
B	MVA results	75

目 次

1	標準理論の粒子	7
2	LHC 加速器	8
3	LHC の各衝突点に置かれている測定器	8
4	Higgs 粒子の生成過程	9
5	Higgs 粒子の生成断面積	9
6	Higgs 粒子の崩壊分岐比	10
7	ATLAS 検出器の全体構造	13
8	Inner Detector の全体図 1	14
9	Pixel Detector の全体図	14
10	SCT module	15
11	Inner Detector の全体図 2	15
12	TRT	16
13	Calorimeter	16
14	EMCal の構造 (barrel)	17
15	Accordion 型 EMCal(barrel)	17
16	TileCal の構造 (barrel)	18
17	TileCal(barrel)	18
18	EndCap Calorimeter	19
19	Forward Calorimeter	19
20	Muon Spectrometer	20
21	Monitored Drift Tube	21
22	MDT alliment	21
23	Solenoid magnet	21
24	Toroidal Magnet (Barrel)	22
25	Toroidal Magnet (End Cap)	22
26	Endcap の中央 ($z=1050\text{cm}$) における x - y 平面の磁力線	22
27	Trigger の流れ	23
28	Luminosity of ATLAS in 2012	24
29	Peak Luminosity & Average Pileup	24
30	Sliding Window Clustering Algorithm	31
31	Electron Fake	32
32	Muon ID	35
33	Muon Fake	36
34	Heavy Flavor Background	36
35	宇宙線事象	37
36	Signal の M_{12} , M_{34} vs $M_{4\ell}$	41
37	Normalized Isolation	42
38	IP significance	42
39	p_T of Leading & Sub-leading pair's lepton	43
40	method in Control region	43

41	$Z+X_\mu X_\mu$ Background Fit1	47
42	$Z+X_\mu X_\mu$ Background Fit2	47
43	lepton p_T distribution at $Z+X_e X_e$ (truth)	49
44	lepton p_T distribution at $Z+X_e X_e$ (Reconstruction)	50
45	$Z+X_e X_e$ Mass distribution at Control region1	53
46	Fakrate & Background Shape	54
47	$Z+X_e X_e$ $M_{4\ell}$ Mass distribution	55
48	$M_{4\ell}$ distribution at cut base analysis	56
49	各 channel の $M_{4\ell}$ distribution	56
50	Production decay angle	58
51	MVA input variable(Angle)	59
52	MVA input variable(Maxx)	60
53	MVA response	61
54	Significance vs BDT response	62
55	MVA results(MC)	62
56	Confidence Level	65
57	BDT applied MC	66
58	BDT applied data	68
59	P0 and Cls	68
60	BDT response including Z +jet	71
61	IP Iso Correlation	71
62	Decision Tree	73
63	MVA input variable(Mass, removed Data- 4μ)	75
64	MVA input variable (Angle, removed Data- 4μ)	76
65	P_{Sig} distribution	77

表 目 次

3	muon の ID Hit Selection	38
4	Event Selection の CutFlow	39
5	Trigger	39
6	$M_{4\ell}$ に応じる M_{34} の minimum threshold	40
7	各 background の transfer factor	45
8	CR2 での Z +light jet の event	46
9	各 background の control region での Fit 結果と MC の数	46
10	各 background の signal region での event 数	46
11	$Z+X_e X_e$ の Control region における event 数	49
12	event in $Z+X_e X_e$ Control Reagion	51
13	Efficiency of $Z(ee)+X_e X_e$	52
14	Efficiency of $Z(\mu\mu)+X_e X_e$	52
15	各 channel での BG 及び data の event 数	57
16	BDT results	62
17	BDT applied expected Significance	66
18	Cut base Limits	67
19	BDT applied Limits	67
20	BDT applied observed p-value	67
21	BDT applied observed significance	67
22	Signal candidate information	77

1 序論

1.1 標準理論

現在、素粒子物理学において最も矛盾なく世界を記述する理論として標準理論という理論体系がある。標準理論は、物質を構成する spin 1/2 を持つ fermion と、素粒子間の力を媒介する spin 1 の gauge 粒子から出来ている。

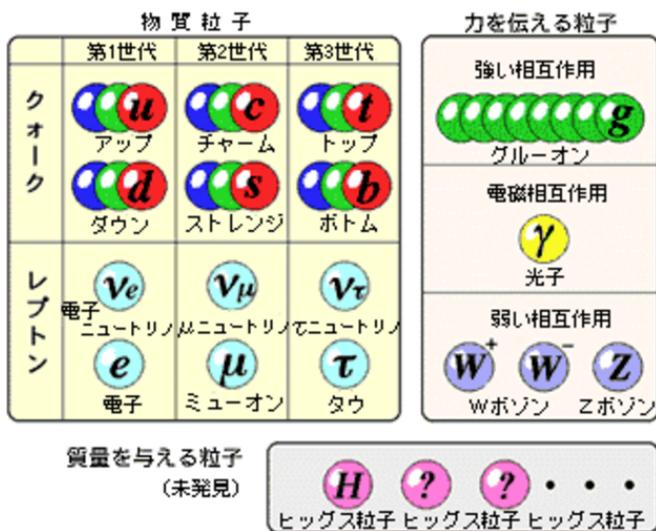


図 1: 標準理論の粒子。標準理論で预言された新粒子の中 Higgs 粒子だけが未発見状態だったが、2012 年 7 月にその候補が発見された。

Gauge 粒子には photon(γ)、gluon(g)、W,Z 粒子があり (1)、これらはそれぞれ電磁相互作用、強い相互作用、弱い相互作用を媒介する。さらに、この中で電磁相互作用と弱い相互作用は電弱相互作用として統一され、このとき重い gauge 粒子を理論に取り入れて繰り込み可能な理論を構築する為に spin 0 の Higgs 粒子が導入された。

この中で唯一見つからない粒子が Higgs 粒子であり、Fermi 国立加速器研究所の Tevatron や欧州原子核研究機構 (以下 CERN) の Large Electron-Positron Collider(以下 LEP) などでの探索が行われ、LHC 実験の開始前には LEP での結果からその存在領域は 114 GeV 以上とされていた。LHC でもその探索が精力的に行われ、2012 年 7 月、ようやくその粒子の候補を 126 GeV 付近に発見した。

この論文では 110~180 GeV の軽い質量の標準理論の Higgs 粒子を、 $ZZ^* \rightarrow 4\ell (\ell = e \text{ or } \mu)$ で探索した。この研究の目的は 1.4 節にまとめる。

1.2 Large Hadron Collider (LHC) 実験

LHC は CERN の地下 100m の場所にある、周長約 27km の大型陽子陽子衝突型加速器である。図 2 に、LHC の全体図を示す。LHC は 2010 年から重心系エネルギー 7 TeV で本格的に稼働し、2012 年からは 8 TeV で実験を行っている。Hadron Collider であるため、synchrotron radiation によるエネルギー損失が少ない特徴を持つ。陽子ビームは、 1.1×10^{11} 個ずつ bunch 化され、bunch crossing rate は 40MHz に達する。LHC の設計時における design 上の性能は、重心系 14 TeV、luminosity

は最大で $10^{34} \text{ cm}^2\text{sec}^{-1}$ に達し、このときの一回の bunch crossing あたりに起こる陽子・陽子の非弾性衝突回数は最大 23 回である。

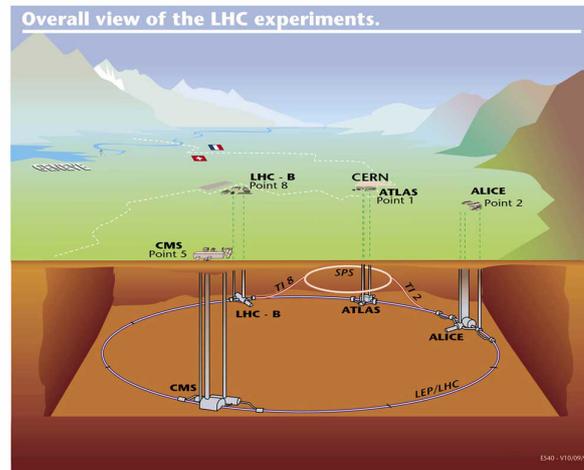


図 2: LHC 加速器の全体図。地下 100 m

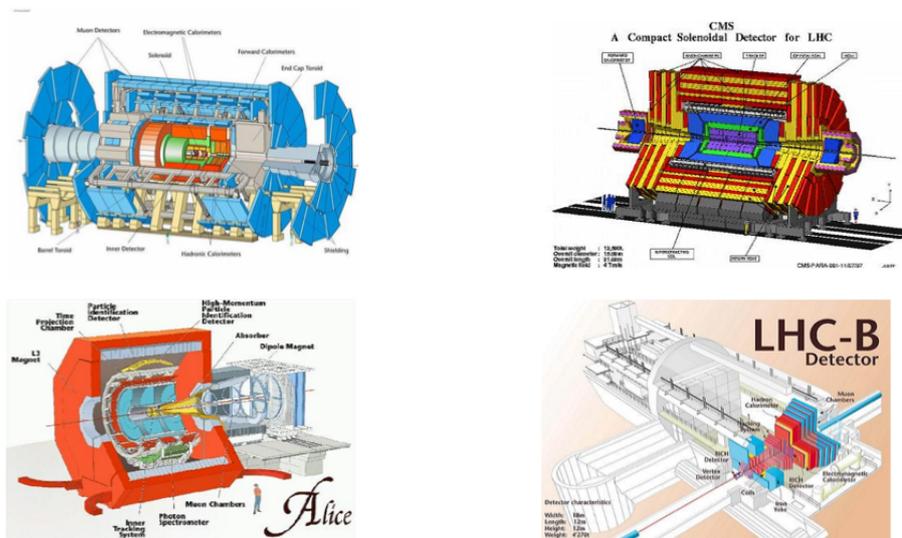


図 3: LHC の各衝突点に置かれている測定器。ATLAS(左上), CMS(右上), ALICE(左下), LHC-B(右下)

ビーム衝突点は4つ存在し、それぞれの衝突点に図3に示すような、

- ATLAS(A Troidal LHC ApparatuS)
- CMS(Compact Muon Solenoid)
- ALICE(A Large Ion Collider Experiment)
- LHC-B(LHC-Beauty)

の測定器が置かれている。本研究では、ATLAS 測定器を用いた実験で取得した data を用いている。使用された data についての詳細は 3 章で説明する。

1.3 Higgs 粒子

ATLAS 実験では、標準理論の精密測定から始まって様々な物理を検証する。その中でも最も重要な目的は Higgs 粒子と標準理論を超えた新しい物理の探索である。ここでは ATLAS 実験における Higgs 粒子の探索について、生成・崩壊過程の視点から説明する。

1.3.1 Higgs 粒子の生成過程

LHC での Higgs 粒子の生成過程は、主に以下の 4 つである (図 4,5)。

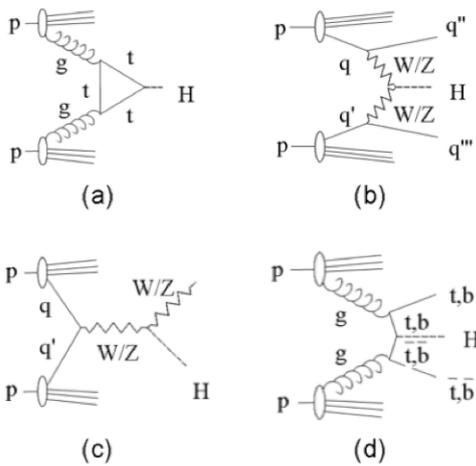


図 4: Higgs 粒子の主な生成過程。(a) gluon-gluon fusion, (b) vector boson fusion, (c) W/Z 随伴生成, (d) t/b 随伴生成

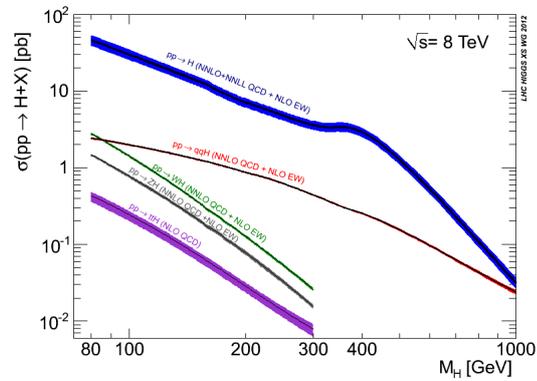


図 5: Higgs 粒子の生成断面積。上から各線が図 4 の (a),(b),(c)(WH と ZH),(d) に対応する。100 GeV から 1 TeV の間では、ggF 過程によって最も生成される。

- $gg \rightarrow H$ (gluon-gluon fusion-以下 ggF)

4 つの過程の中で最も大きい生成断面積を持つ。ggF は Higgs 粒子が fermion と結合する過程で起こる反応であるが、結合定数が質量に比例するため質量がもっとも大きい t-quark の loop を介して生成される過程が大きくなる。しかし、Higgs 粒子の崩壊で作られた粒子以外に大きい transverse momentum (p_T) を持つ粒子がなく、background の選別が非常に難しい過程でもある。
- $qq' \rightarrow q'q'H$ (vector boson fusion-以下 VBF)

ggF の次に生成断面積が大きい過程である。この過程は ggF と比べて断面積は小さいが、散乱角の小さい 2 本の jet を放出する大きい特徴を持つため、background の排除が容易となり感度の高い解析を行う事が出来る。また、ggF は fermion との湯川結合を仮定しているが、VBF は W/Z と結合して生成される。Higgs 粒子は標準理論において W/Z との結合を前提として存在するため、この過程は湯川結合の存在を仮定しなくても存在する。

- $q\bar{q}\rightarrow(W/Z)H$ (W/Z 随伴生成)
quark-反 quark 対消滅により W/Z が生成され、この gauge boson から Higgs 粒子が放出される過程である。これから終状態に gauge boson から崩壊した粒子が観測される特徴を持つ。
- $gg\rightarrow(t\bar{t}/b\bar{b})H$ (t/b 随伴生成)
gluon から対生成された t/b-quark が対消滅し、Higgs 粒子が生成される。終状態に t/b-quark による jet を伴う。

1.3.2 Higgs 粒子の崩壊過程

Higgs 粒子の崩壊過程はその質量に依存している。その分岐比を図 6 に示す。この質量依存性から、質量の領域によって探索に有利な過程が存在する。

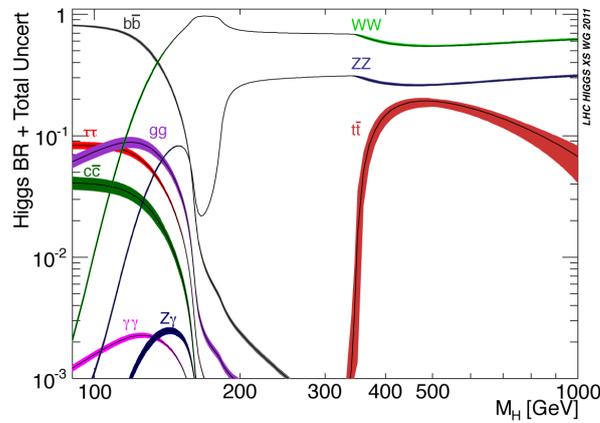


図 6: Higgs 粒子の崩壊分岐比

- $114 \text{ GeV} < M_H < 130 \text{ GeV}$

$$H \rightarrow \gamma\gamma$$

この質量領域では $b\bar{b}$ 、 $\tau^+\tau^-$ 、 $c\bar{c}$ であるが、background が大きいため測定が難しくなる。そのために、分岐比は小さいがこの崩壊過程が主な探索過程となる。 γ は質量を持たないため Higgs 粒子と直接結合することは出来ないため、t/b-quark、W boson の loop を介して光子を放出する。

- $130 \text{ GeV} < M_H < 180 \text{ GeV}$

$$H \rightarrow ZZ^* \rightarrow 4l, H \rightarrow WW^* \rightarrow l\nu l\nu$$

ZZ^* の場合、終状態が 4 つの lepton というまれな事象を見るので background が少ないが、片方が off-shell になる。 WW^* の場合は、neutrino の不変質量を再構成できないので、missing transverse energy (E_T) の精密な測定が必要となる。本研究では、この ZZ^* を用いて解析を行っており、次節で詳しく説明する。

- $180 \text{ GeV} < M_H < 700 \text{ GeV}$

$$H \rightarrow ZZ \rightarrow 4\ell$$

この領域に置ける ZZ^* への分岐比は大きく、2つの Z は共に on-Shell である。従って4つの lepton に2つの Z の mass(約 91 GeV) を要求するので background を取り除きやすくなる。

1.4 $H \rightarrow ZZ^* \rightarrow 4\ell$ 解析

本研究では、ATLAS 実験で2012年3月から8月まで取得した重心系エネルギー 8 TeV、 13 fb^{-1} の data を用いて $H \rightarrow ZZ^* \rightarrow 4\ell$ channel における軽い質量 110~180 GeV を持つ Higgs 粒子の探索を行った。実際、2012年7月、ATLAS と CMS 実験は標準理論 Higgs 粒子の探索に置いて、126 GeV 付近に新しい粒子を発見したが、この発見は主に $H \rightarrow \gamma\gamma$ と $H \rightarrow ZZ^* \rightarrow 4\ell$ channel を通して event 超過を観測する事で行われた。

この ATLAS 実験の $H \rightarrow ZZ^* \rightarrow 4\ell$ channel の結果では、重心系エネルギー $\sqrt{s} = 7 \text{ TeV}$ の 4.8 fb^{-1} の data と $\sqrt{s} = 8 \text{ TeV}$ の 5.8 fb^{-1} の data が使用された。本研究では、7 TeV の data は用いないが、8 TeV の data を 13 fb^{-1} まで利用しており、使用 data 量はこの発見の時より若干 data 量が多い。

この channel は、前説でも述べたように、終状態に4つの lepton を要求するため、background が少ない channel である。その中でも主な background は Higgs 粒子を介さない $ZZ^* \rightarrow 4\ell$ で、終状態が signal と同じである。4つの lepton から再構成される $M_{4\ell}$ の分布において、signal は細い peak (mass resolution 2~3 GeV) を作るが、それ以外の区別は難しい。

本研究では、上記の発見時に使われた各変数に制限を加えていく解析手法(以下、Cut base)では区別しにくいこの background を多変数解析方法(Multii-Variate Analysis、以下 MVA)を用いて排除する事で発見感度の改善を目指した。MVA では、多変数を同時に、かつそれらの変数の相関まで考慮して signal と background を分離するために、一般に cut base の手法より改善が期待される。本研究では lepton やそれらから構成される Z boson 候補の角度情報を MVA に取り入れる事で発見感度の改善を行った。

1.5 本論文の構成

本論文は全6章で、第2~6章は下のように構成する。

第2章 ATLAS 検出器

ATLAS 測定器の概要と、検出器の構成要素や性能などについて説明している。

第3章 Data と Monte-carlo sample

本研究で使った data や background process について説明する。

第4章 解析

本研究の解析方法とその結果について述べる。解析で使われる selection や多変数解析手法などについて説明する。

第5章 結果と考察

解析結果を用いて、新粒子の探索及び多変数解析手法の結果について考察する。

第6章 まとめ
全体のまとめ

2 ATLAS 検出器

ATLAS 検出器は、直径約 25 m、長さ約 44 m、総重量は約 7000 t に及ぶ汎用検出器である。検出器の構造は図 7 で示すように、各粒子の透過性から内側から Inner Detector、ElectroMagnetic Calorimeter、Hadron Calorimeter、Muon Spectrometer の順で構成されている。磁石は、calorimeter を囲むような形の Toroidal 磁石を使っていて、これは ATLAS 検出器の大きな特徴でもある。

座標系は、beam 軸を z 軸とし、LHC のリングの中心方向を x 軸、これらに垂直な軸を y 軸としている円筒座標系を使う。z の方向によって、A side($z > 0$) と C side($z < 0$) と定義される。x-y 平面上での ϕ は x 軸からの角度であり、左周りを正の方向とする。また、衝突点を原点として beam 軸と成す角度を θ とし、pseudo rapidity η と $\eta = -\ln(\tan(\theta/2))$ のような関係にある。

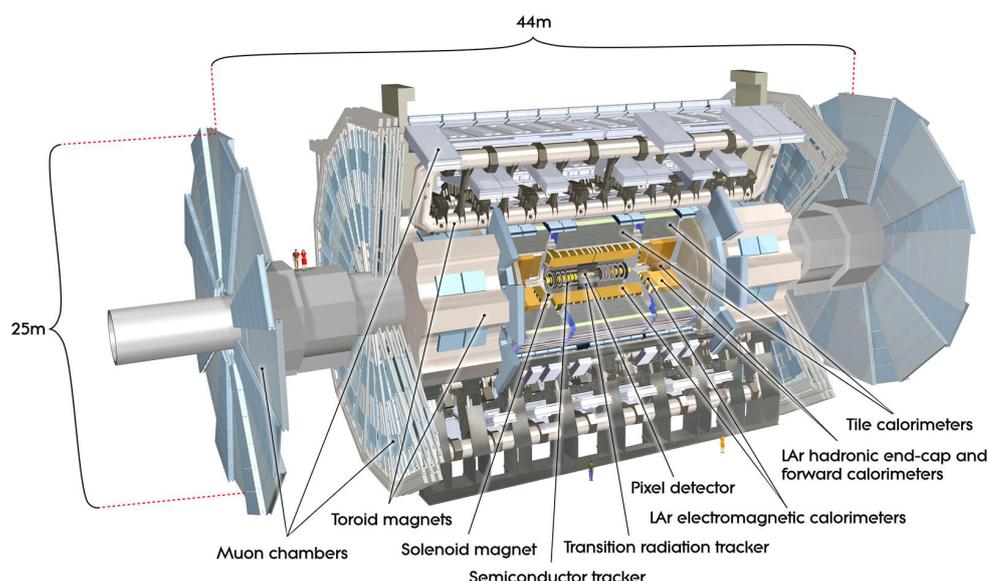


図 7: ATLAS 検出器の全体構造

全体の形は beam 軸を囲んでいる円筒型をしていて、大きく Barrel 部 ($|\eta| < 1.3$)、End-Cap 部 ($1.3 < |\eta|$) に分ける事が出来て、End-Cap はさらに Forward 部 ($3 < |\eta|$) に分けることがある。検出器に主に要求される性能は、

- 大きい pseudo rapidity の範囲において粒子の同定が可能
- LHC の高い luminosity に対応できる高速応答
- 耐放射線性
- Inner detector による荷電粒子の精密な運動量測定
- EM calorimeter による高精度の electron と photon の同定、hadron calorimeter による jet と missing E_T の高精度測定
- Muon Spectrometer による高精度の muon p_T 測定

などがある。

2.1 Inner Detector

Inner detector は、ATLAS 検出器で最も beam 衝突点に近い場所に設置され、2T の磁場を作る超伝導 Solenoid 磁石の内部にある。内側から、Pixel detector、SemiConductor Tracker(SCT)、Transition Radiation Tracker(TRT) の3つから構成されている。(図 8)

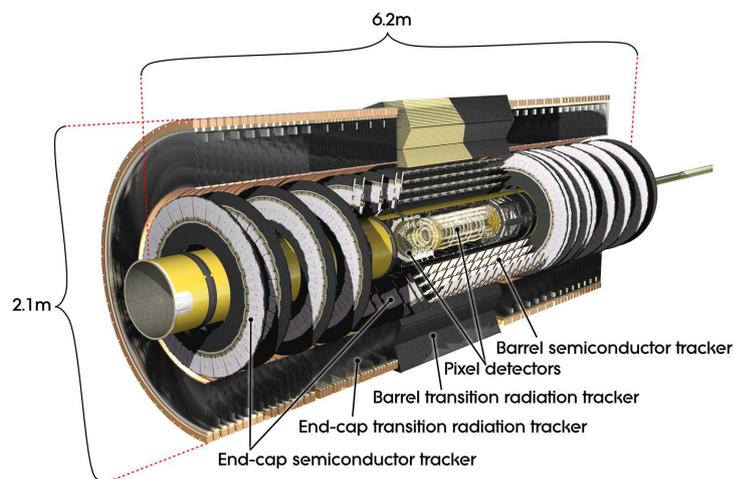


図 8: Inner Detector の全体図。

- Pixel Detector

最も beam pipe に近い半導体検出器である。非常に高い位置分解能を持ち、 $r-\phi$ 方向で $12\mu\text{m}$ 、 z 方向で $110\mu\text{m}$ の精度で測定が可能で、2次元の位置情報を与える。Barrel 部では3層に配置され、End-cap 部には A,C side を合わせて6枚の disk 状に配置されている。最も内側に配置されているため、b-hadron のような長寿命 hadron の同定に用いられる。特に、最も内側は b-layer と呼ばれる。(図 9)

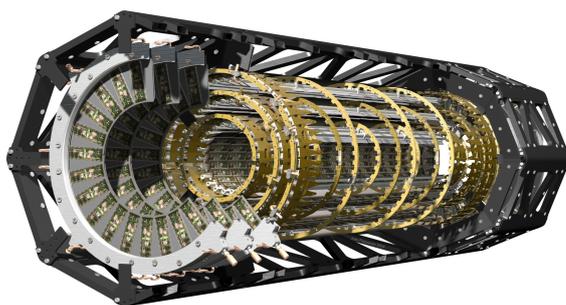


図 9: Pixel Detector の全体図。

- SemiConductor Tracker(SCT)

pixel 検出器の外側にある。図 10 のように、半導体が $80\mu\text{m}$ で並んだ strip 型の sensor を、 40 mrad ずらして2枚を合わせることで pixel と同様に2次元の位置情報を与える。Barrel 部

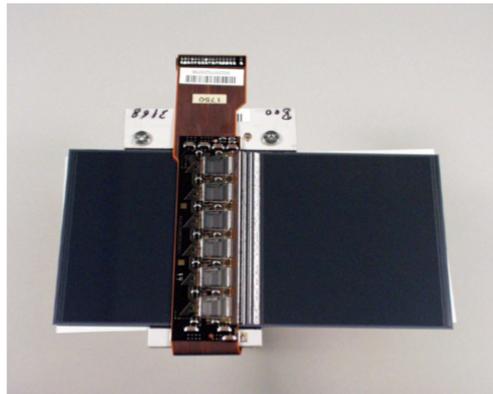


図 10: SCT module の写真。

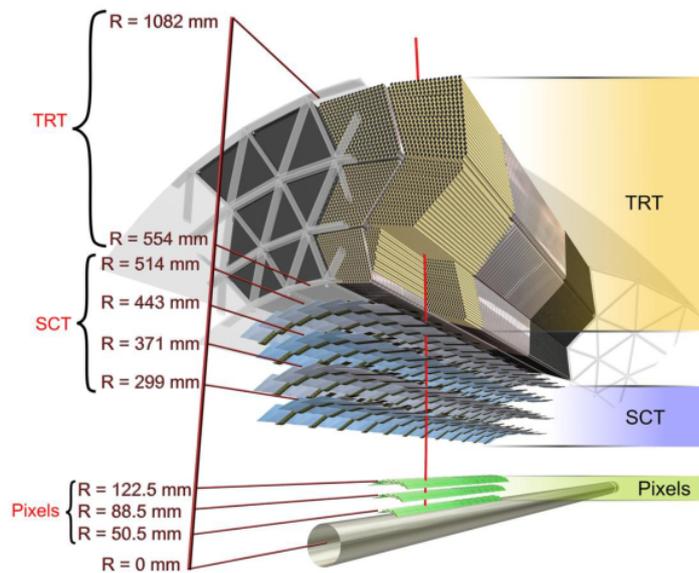


図 11: Inner Detector の barrel 方向の構造。

には円筒型に4層(図 11)が存在し、End-cap には両 side を合わせて 18 枚の disk(図 8)を作っている。位置分解能は r - ϕ 方向で $17\mu\text{m}$ である。

- Transition Radiation Tracker(TRT)

TRT は SCT の外側に位置している gas 検出器である。分解能は $200\mu\text{m}$ で pixel や SCT より悪いが、strow 型の検出器を 36 層積層して連続的な飛跡の測定が出来、Track の reconstruction を可能とする。位置情報は r - ϕ 方向の情報のみを持ち、分解能は一本の strow tube あたり $130\mu\text{m}$ を持つ。(図 12)



図 12: TRT の全体図

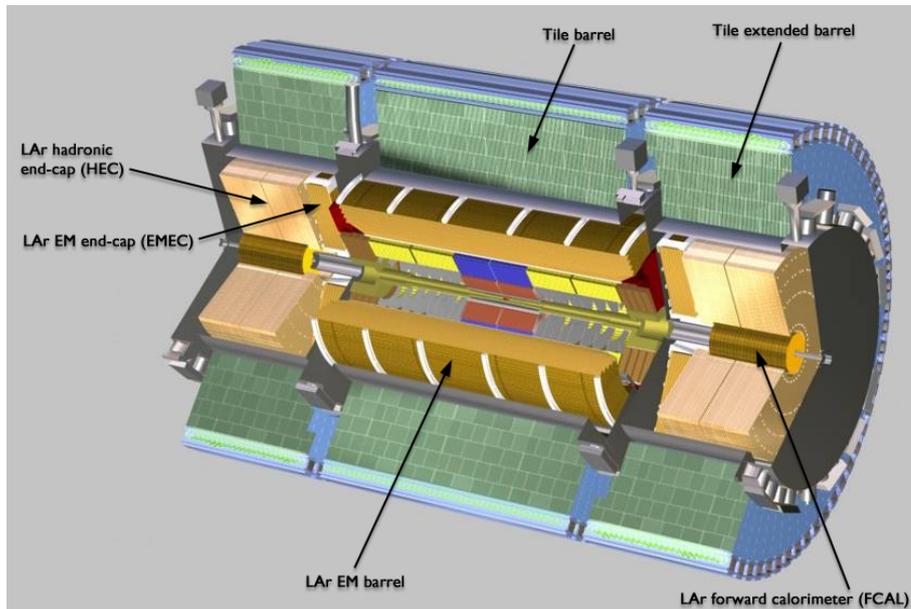


図 13: Calorimeter の全体図。内側が EM、外側が Hadron calorimeter となっている。

2.2 Calorimeter

Calorimeter は、electron や photon、jet などの粒子のエネルギーと角度を測定する。全体は 4 種類の測定機で構成され、electron と photon の測定に用いられる ElectroMagnetic calorimeter (EM calorimeter) と、hadron の測定に用いられる hadron calorimeter に分ける事ができる。いずれも sampling calorimeter の方式を取っている。(図 13)

粒子のエネルギーという非常に重要な測定を担うためになるべく全立体角を囲むようにしており、設置される位置によって Barrel 部、End-cap 部、Forward 部に分かれる。

2.2.1 EM calorimeter

ATLAS の EM calorimeter の大きな特徴は図 15 で示すような accordion 型であることにあり、吸収層と検出層としてそれぞれ鉛と liquid argon を使う。理想的に入射粒子の進行方向に垂直であ

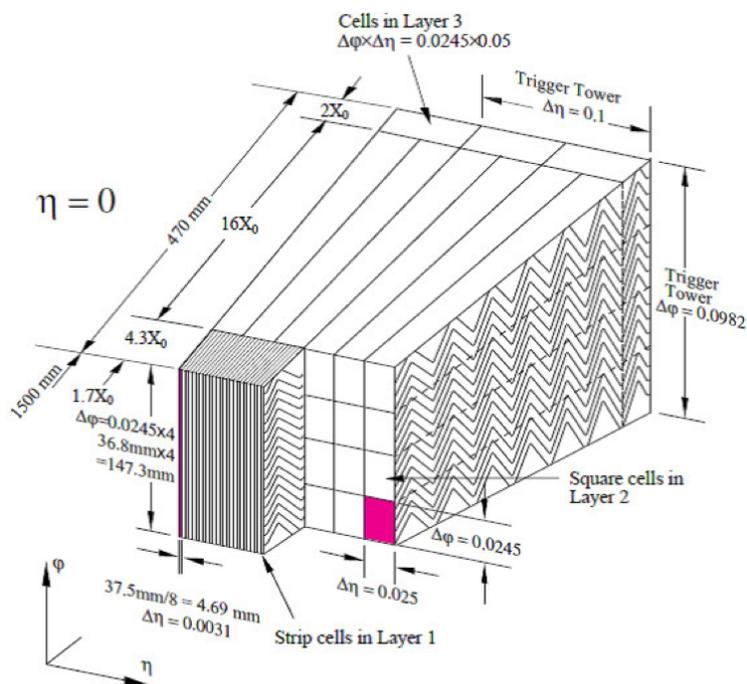


図 14: Barrel 部の EMCal の構造



図 15: Accordion 型 EMCal の実物

るのが良いが、一般に全立体角を覆っていると配線などの問題で不感領域が生じてしまう。しかし、accordion 型にすることによりこのような不感領域を最大限排除し、どの角度からの入射でも同じような物質層を通過するようにすることが出来る。

全体の物質量は、EM shower のエネルギーを全て吸収するために十分な約 $22 X_0$ (radiation length) である。そのエネルギー分解能は、

$$\frac{\sigma_E}{\sigma} = \frac{(10.1 \pm 0.1)\%}{\sqrt{E}} \oplus (0.17 \pm 0.04)\% \quad (1)$$

で表される。第一項は統計から来る項であり、第 2 項は calibration の精度や温度の揺らぎによる定数項である [3]。例えば、エネルギーが 100 GeV の場合、約 1% の分解能で測定が出来る。構造は 3 層からなっており、それぞれ Layer1,2,3 と呼ばれ、以下のような特徴を持つ。(図 14)

- Layer1: EM shower の広がりが小さいので、 η と ϕ の (特に η) 精密測定が可能である。
- Layer2: η と ϕ の精密測定も可能であるが、主にエネルギーの測定を担っている。
- Layer3: Shower の奥行きの違いを利用して、EM shower と hadron shower の区別が可能になる。つまり、この layer に deposit が多かった場合、e や γ ではないと判断する。

2.2.2 Hadron calorimeter

EM calorimeter の外側に配置され、Jet のエネルギーや方角を測定するのに用いられる。Hadron calorimeter は、主に 2 種類に分けられる。

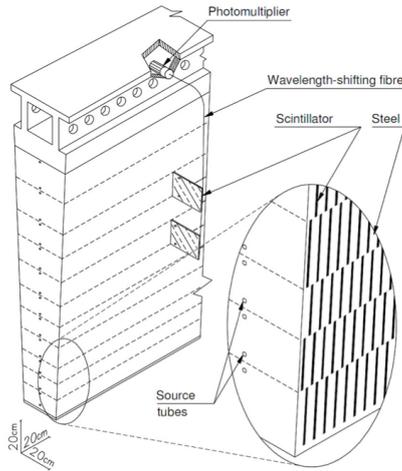


図 16: Barrel 部の Tile Cal 構造



図 17: Barrel 部の Tile Cal の実物

- Tile Calorimeter

$|\eta| < 1.7$ に配置される (図 16、17)。Tile Calorimeter は吸収体として鉄を用い、タイル上の plastic scintillator からなる。 ϕ 方向に 64 分割されており、granularity は $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$ となっている。Tile からの信号は波長変換 fiber を用いて、光電子増倍管で測定している。Test beam によるエネルギー分解能は、

$$\frac{\sigma_E}{\sigma} = \frac{(56.4 \pm 0.4)\%}{\sqrt{E}} \oplus (5.5 \pm 0.1)\% \quad (\text{Barrel}) \quad (2)$$

である。

- End-Cap LAr Calorimeter

放射線強度がより強いいため、銅の吸収体と LAr からなる sampling calorimeter を用いている。EM calorimeter とは違い、銅の吸収体は 25~ 50 mm の平板である。 $1.5 < |\eta| < 3.2$ までの領域を囲んでいる。Test beam によるエネルギー分解能は、

$$\frac{\sigma_E}{\sigma} = \frac{(70.6 \pm 1.5)\%}{\sqrt{E}} \oplus (5.8 \pm 0.2)\% \quad (\text{EndCap}) \quad (3)$$

である。(図 18)

2.2.3 Forward calorimeter(FCAL)

Accordion 型の EM calorimeter は、barrel と end-cap 部を覆っている。しかし、前方領域 ($3.1 < |\eta| < 4.9$) は beam 軸に近く、放射線強度が最も強いので、図 19 の左で示すように銅板に穴を変えた中に棒状の銅を入れ、その隙間に LAr が満たされた特殊な構造をした LAr calorimeter が使われている。FCAL も 3 層に分かれており、この銅を用いたものは最深部に近い FCAL1(図 19 右) に配置

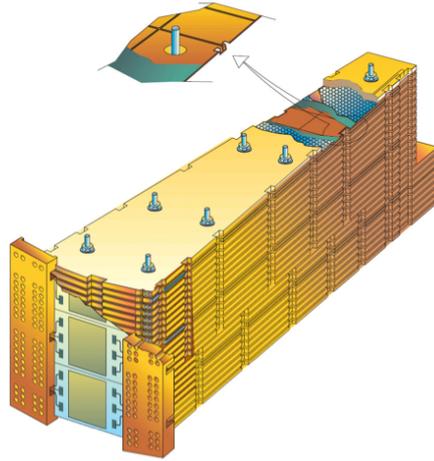


図 18: End Cap Calorimeter

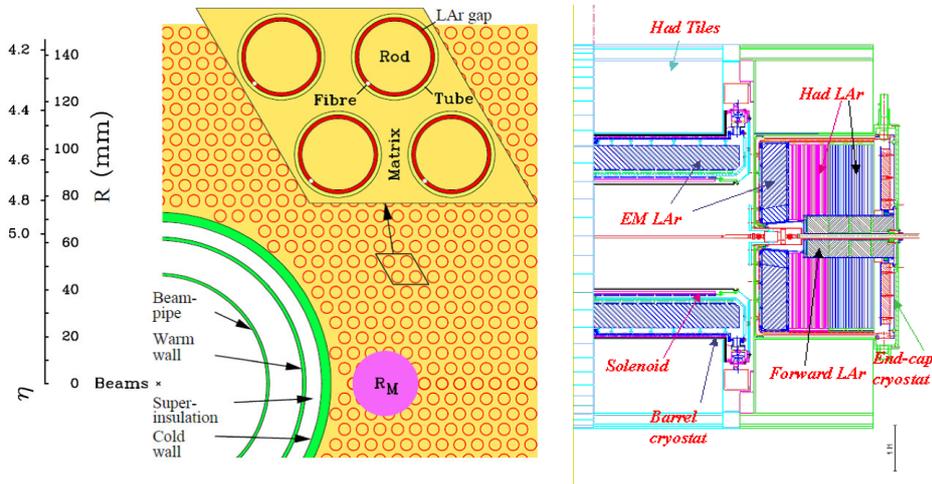


図 19: Forward Calorimeter の構造 (左) と、Calorimeter 全体の位置関係 (右)

されている。残りの 2 層は tungsten が吸収体となっており hadron calorimeter の役割を担う。Test beam に依るエネルギー分解能は、

$$\frac{\sigma_E}{\sigma} = \frac{(94.2 \pm 1.6)\%}{\sqrt{E}} \oplus (7.5 \pm 0.4)\% \quad (\text{Forward}) \quad (4)$$

である。

2.3 Muon Spectrometer

Muon は物質との相互作用が弱く、測定器のほとんどの層を通過してしまう。従って、Muon Spectrometer は検出器の一番外側に置かれている。ATLAS の muon spectrometer は、運動量の精密測定を担う Monitored Drift Tube(MDT) と Cathode Strip Chamber (CSC)、trigger 用の Resistive Plate Chamber (RPC) と Thin Gap Chamber(TGC) がある。図 20 に、Muon Spectrometer の全体像を示す。

- Monitored Drift Tube

MDT は Barrel と End-Cap に設置され、 $|\eta| < 2.7$ という広い範囲を cover し、muon の運動量の測定に用いられる。多くの場合、drift tube が 3 層、または 4 層の multi layer が spacer を挟んだ形をしており、その間隔も 6.5~317 mm まで、module によって異なる。MDT の特徴はその alliment 方法にあり、module 内の配置を確認する為に LED を用いて module の変形が数 μm の精度で monitor することができる。さらに、module 間にも光学 system を用いて monitor しており、運動量精度から要求される $30\mu\text{m}$ より高い精度の alliment が達成されている。

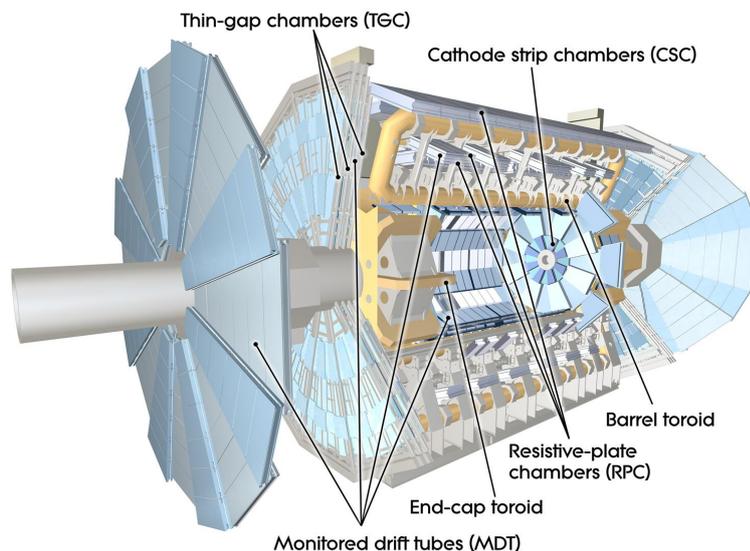


図 20: Muon Spectrometer の全体構造

- Cathode Strip Chamber

CSC は、Forward 部の内側 ($2.0 < |\eta| < 2.7$) に配置されている。beam pipe の近傍なので、強い放射線強度に耐えられるような検出器が必要となる。CSC は全部で 16 chamber が ϕ 方向に交互に配置されている。1 つの chamber は 4 層で出来ていて、それぞれで独立に track の η や ϕ を測定できる構造となっている。

- Resistive Plate Chamber

RPC は、barrel 部分 ($|\eta| < 1.05$) の muon trigger の役割を担う。LHC の衝突間隔である 25 ns よりも高速な読み出しが可能で、2 枚の平衡版を用いた gas 検出器である。RPC の位置分

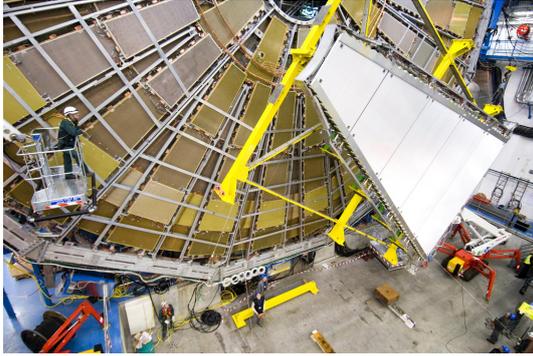


図 21: TGC の後ろに MDT(白) を設置している様子。

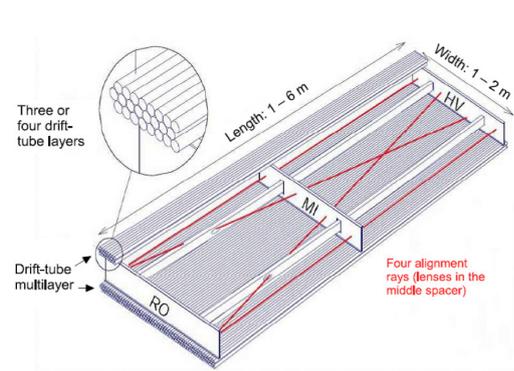


図 22: MDT LED Alignment monitor

解能は 1mm、時間分解能は 2ns である。Magnet system によって曲げられた荷電粒子の飛跡を測定することで、muon の p_T を測定し、trigger を発行する。

- Thin Gap Chamber

End-Cap 部 ($1.05 < |\eta| < 2.7$) を囲んでいる trigger 用検出器である。Drift time が短く、時間分解能が数 ns で優れており、trigger 用の検出器としてすぐれている。3 層構造からなる triplet module と、2 層構造の doublet module がある。

2.4 Magnet System

ATLAS の特徴の一つは、Toroidal Magnet System を採用していることにある。内側の inner detector では solenoid を使っており、全体で 2 種類の磁石を使っており、toroidal magnet の場合は barrel 部に 8 つ、それぞれの end-cap 部に 8 つの toroidal magnet を使っている。

- Solenoid Magnet



図 23: Solenoid magnet

ATLAS では、inner detector での荷電粒子の p_T 測定の為に最大 2T の磁場がかかるように solenoid magnet を使っている。これを持って、x-y 平面内で Lorentz 力を受けて運動量に比例した半径で粒子の軌跡が曲がり、この曲率から運動量を計算することが出来る。中空の磁場

は約 2.0T になっている。(図 23)

- Toroidal Magnet

殆どの粒子は hadron calorimeter まで止まるが、muon は全ての検出層を通過して出てくる。この muon の p_T の精密測定の為に外側にも超伝導の toroidal magnet が用いられている(図 24、25)。これで p_T が小さく、 p_Z が大きい muon に対しても運動量測定が可能となる。Toroidal system の特徴は、solenoid では ϕ 粒子を曲げるのに対して、 η 方向に粒子を曲げる事である。(図 26)

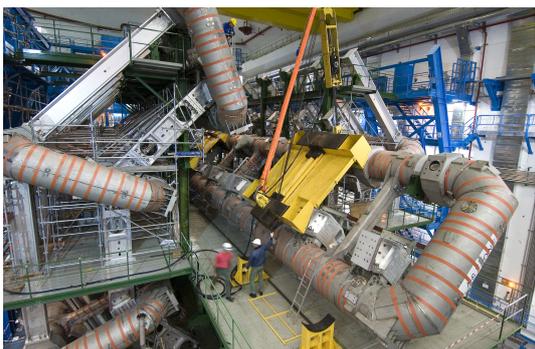


図 24: Toroidal Magnet (Barrel 部)



図 25: Toroidal Magnet (End Cap 部)

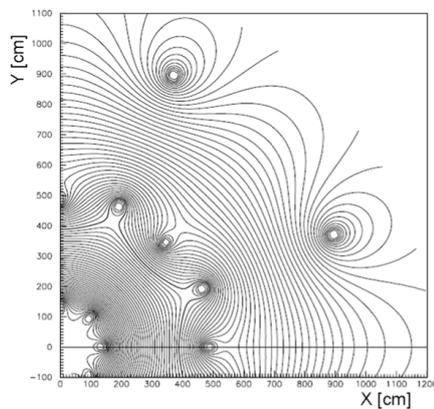


図 26: Endcap の中央 ($z=1050\text{cm}$) における x - y 平面の磁力線

2.5 Trigger

LHC 加速器では、High Luminosity のとき 1 回の bunch crossing あたりに ~ 20 回の p-p 衝突が予想され、bunch の交差は 40 MHz 程度であることを考慮すると約 1GHz の頻度で衝突が起こることになる。これらから発生する多くの事象から真に欲しい事象のみを同定・選別するためには、必要としない事象を記録するために filtering をする機構が必要となり、これを trigger と呼ぶ。最終的な

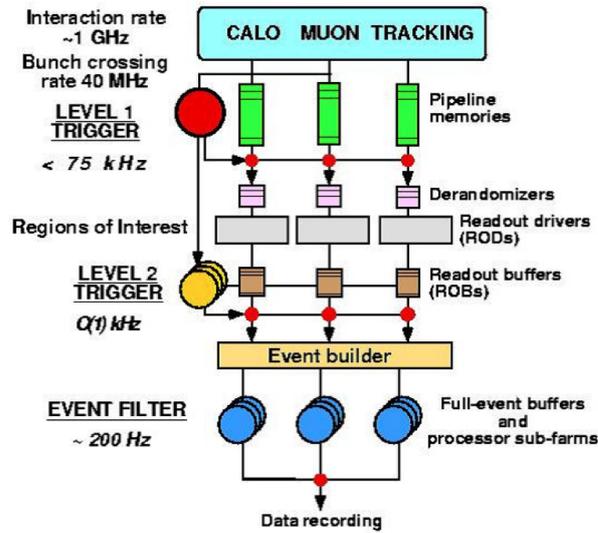


図 27: ATLAS における Trigger system の流れ

event rate は、data storage system の書き込み速度の制限で 200Hz 程度まで抑える必要がある。このために ATLAS 実験の trigger system は 3 段階に別れ、順次 event rate を落として行く。それぞれの段階を、Level 1、Level2、Event Filter と呼ぶ。(図 27)

- Level1 trigger (LVL1 trigger)

Trigger 用の muon 検出器である TGC や RPC から得られる (R, ϕ) 座標、全 Calorimeter のエネルギー情報、運動量を元に選別を行う。この段階では情報量が非常に多く、LVL1 buffer には検出器全ての情報が $2\mu\text{s}$ の間だけ保持されるためその間 trigger の判定を行わなければならない。Calorimeter の LVL1 trigger (L1 Calo) では高い E_T の粒子の識別をする事を目的とする。Muon trigger の場合、 p_T に threshold をかけており、LVL1 には全部で 6 つの trigger が用意されている。この段階で、40MHz だった bunch cross rate は 75kHz まで落とされる。

- Level2 trigger (LVL2 trigger)

この段階では inner detector や MDT などの情報を加えて trigger が掛けられる。ただし、その効率を上げるために LVL1 で用いられた trigger の条件から Regions of Interest (ROI) と呼ばれる領域のみを利用する。これを持って、LVL2 で要求される 40ms 以内での trigger 判定を実現している。ここではエネルギーや運動量の見積もりを正確に行っており、muon の場合は MDT 情報を用いて cut しており、inner detector と MDT の軌跡が一致している可能性も考慮される。この段階での event rate は 1kHz まで落とされる。

- Event Filter

LVL2 を通った event に対し、全ての検出器の情報を用いて trigger の判定を行う。この段階で trigger にかかったものが disk に書き込まれ、event rate は 200~300 Hz まで落とされる。ここで記録されたものが、実際の解析で使われる使用可能な data となる。

3 Data と Monte Carlo sample

3.1 実験の Data

LHC では、2011 年まで重心系エネルギー $\sqrt{s}=7$ TeV で約 4.8 fb^{-1} 、2012 年で 8 TeV で約 21 fb^{-1} の Intergrated luminosity のデータを取得している。図 28 は 2012 年に取得した 8 TeV の全データを日時の関数として表したものである。本研究では 2012 年 3 月から 8 月までに取得した $\sqrt{s}=8$ TeV の 13 fb^{-1} の data を使って解析を行っている。

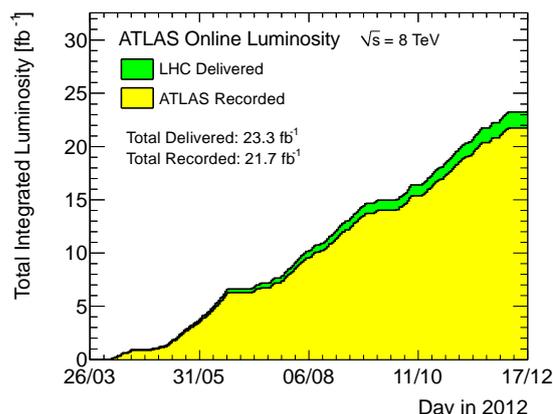


図 28: 2012 年の ATLAS 検出器で蓄積された integrate luminosity。緑が beam を用いて行われた p-p 衝突全体の luminosity で、黄色が実際に記録された luminosity を表す。

また、2012 年の peak luminosity は約 $10^{34} \text{ cm}^2 \cdot \text{s}^{-1}$ で、1 回の bunch crossing あたり発生する平均衝突数 (Pile up) は 20 回程度である。これらを図 29 に示す。

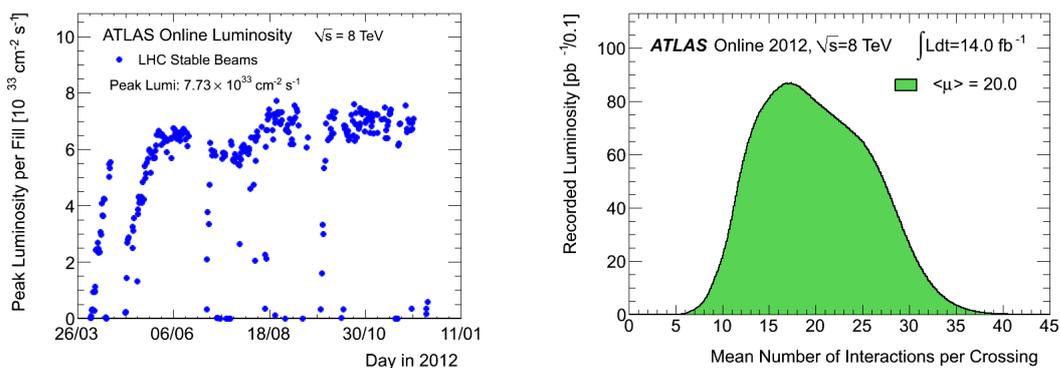


図 29: LHC の $\sqrt{s}=8$ TeV における peak luminosity(左) と平均衝突数 (右)。右の図は 2012 年の 9/17 までの data から得られた者である。

3.2 Background Process

本研究における $H \rightarrow ZZ^* \rightarrow 4\ell$ の background について述べる。 $H \rightarrow ZZ^* \rightarrow 4\ell$ は、基本的に background が非常に少ない channel であるが、以下の三つが主な background となる。

- ZZ^*
Higgs 粒子を介さず、二つの Z (片方は off-shell) が出来て、4 つの lepton となる場合で、終状態が signal と同じであるために排除が困難な (irreducible) background である。その存在領域も、 $H \rightarrow ZZ^* \rightarrow 4\ell$ の探索可能範囲全てに渡って存在するので、これが主な background となる。
- $Z + b\bar{b}$, $Z + \text{jets}$
一つの Z が二つの lepton に崩壊し、2 つの b/c-quark が semi-leptonic 崩壊をして lepton を出す場合や、同じく一つの Z が lepton 対に崩壊し、jet から misidentify された lepton を拾う場合となる。これらは jet が存在することや、長寿命の粒子が崩壊して lepton が得られることを利用して排除が可能である。
- $t\bar{t}$
t が W と b に崩壊し、 W が lepton と neutrino に、b が semi-leptonic 崩壊をして合計 4 つの lepton が得られる場合であり、 $Z + b\bar{b}$ と同様に排除可能な background となる。 $Z + b\bar{b}$ 、 $Z + \text{jets}$ と合わせて、Low mass region でのこれら 3 つの background の量は、 ZZ^* と同程度となる。

これらが 4ℓ channel での主な background となる。本研究は、前者の低領域での Higgs 粒子の探索を行っているので、 $Z + \text{jet}$ 、 $Z + b\bar{b}$ や $t\bar{t}$ の排除が signal の acceptance を上げるのに大事になる。

3.3 Monte Carlo Simulation

一般に素粒子物理における解析では、Monte-Carlo simulation (以下 MC) による解析結果と data の解析結果を比較しながら行う。そのため、実際予想される process を忠実に再現する MC を用意する必要がありこれらを MC samples ともいう。MC sample は、大きく次の 4 つの段階で生成される。

1. Event 生成
2. Detector simulation
3. Digitization
4. Reconstruction

1.~3. は、MC 特有の process で、実際の実験 data と同じ環境を構築して再現する作業である。2. では ATLAS 検出器における粒子反応を Geant4 を用いて再現している。これから、4. では実 data と MC を同じプログラムを用いて event の再構築をすることが可能になる。

これらを再現する為に形成方法によって様々な種類の generator というものが存在する。ここでは、1. で用いられる、本研究で使われた主な MC generator について簡単に説明する [1]。

- PYTHIA・Herwig
 e^+e^- collider や hadron collider の実験から長年に渡って使われてきた Leading Order(LO) の event 生成 program である。この二つの program は衝突による hard scattering process を生成するだけでなく、hadron collider では陽子から実際の parton 同士の衝突までの event の発展、衝突で生じた parton の発展、崩壊して安定した粒子になるまでを取り扱う事が出来る。そのため、次に述べる MC@NLO、Alpgen、PowHeg では、Hard scattering process のみを生成し、parton shower や崩壊については Pythia で行う。
- MC@NLO
 Next to leading order までの QCD の摂動計算で event を発生させ、Herwig program によって parton shower の手法と組み合わせることで適切な MC event を生成する事が出来る event 生成 program である。この枠組みの中で正しい運動量分布を再現するため、負の event weight を持つ event を作ることが特徴の一つである。
 NLO まで考慮するため、 $t\bar{t}$ の process において t-quark の p_T 再現性が Pythia や Herwig による LO の MC event より優れている。
- POWHEG
 基本的に MC@NLO と同様、NLO まで計算された hard scattering process を使用している
 であるが、event weight が正のものしか生成されない。名前もこの特徴から由来したものである (Positive Weight Harrest Emission Generator)。
- Alpgen
 p-p collider である LHC は、hard scattering に parton shower による p_T の小さい jet が伴う event のみならず、hard scattering から十分に離れた p_T の大きい jet が複数生じる event が生成される。このような event は Pythia では再現が困難である。Alpgen は、このような終状態に p_T の大きい複数の jet が存在する event を LO で生成することが可能な program である。Herwig によって生成される parton shower による jet と Alpgen で生成された jet の分離を適切に行うことで、event の double count を防ぐ事が出来る (MLM matching)。

3.4 MC Samples

本研究で利用した MC sample について、signal と background process で使用した MC generator 及びその生成断面積をまとめておく。

3.4.1 Background Samples

process	MC Generator	xsec[fb]	K-factor	Filter
$t\bar{t}$	McAtNloJimmy	1.29×10^5	1.0	Lepton Filter
$Z(\rightarrow \ell^+\ell^-)+\text{jet}$	AlpgenJimmy	9.34×10^5	1.23	$60 < M < 2000$ GeV
$Z(\rightarrow \ell^+\ell^-)+\text{jet}$	AlpgenJimmy	3.65×10^6	1.19	$10 < M < 60$ GeV
$Z(\rightarrow \ell^+\ell^-) + b\bar{b}$	AlpgenJimmy	7.15	1.6	4 ℓ Filter
$Z(\rightarrow \ell^+\ell^-) + b\bar{b}$	AlpgenJimmy	1.40×10^3	1.6	3 ℓ Filter

process	MC Generator	xsec[fb]	K-factor	Filter
$gg \rightarrow ZZ \rightarrow 4\ell$	Jimmy	1.72	1.0	
$qq \rightarrow ZZ \rightarrow 4\ell$	Powheg	4.99	1.0	

3.4.2 Signal Samples

後述する event 選別後の 4 つの lepton から再構成される不変質量 $M_{4\ell}$ の resolution は、125 GeV の Higgs について 2.0 GeV (4μ)、2.3 GeV ($2e2\mu$ 、 $2\mu2e$)、2.6 GeV ($4e$) である。

process	MC Generator	mass[GeV]	σ [pb]	+Tot%	-Tot%	BR
ggF	Powheg	110	25.04	+15.3	-14.9	4.34×10^{-3}
		115	25.04	+15.3	-14.9	8.65×10^{-3}
		120	21.13	+14.8	-14.8	1.59×10^{-2}
		123	20.15	+14.7	-14.8	2.18×10^{-2}
		124	19.83	+14.7	-14.8	2.40×10^{-2}
		125	19.52	+14.7	-14.7	2.64×10^{-2}
		126	19.22	+14.7	-14.7	2.89×10^{-2}
		127	18.92	+14.6	-14.7	3.15×10^{-2}
		130	18.07	+14.6	-14.6	3.98×10^{-2}
		135	16.79	+14.4	-14.7	5.47×10^{-2}
		140	15.63	+14.3	-14.5	6.87×10^{-2}
		145	14.59	+14.1	-14.4	7.91×10^{-2}
		150	13.65	+14.1	-14.4	8.25×10^{-2}
		155	12.79	+14.1	-14.4	7.34×10^{-2}
		160	11.95	+14.0	-14.4	4.15×10^{-2}
		165	10.89	+13.9	-14.5	2.21×10^{-2}
		170	10.12	+13.9	-14.5	2.36×10^{-2}
175	9.476	+13.7	-14.5	3.23×10^{-2}		
180	8.874	+13.6	-14.5	6.02×10^{-2}		
VBF	Powheg	110	1.809	+2.7	-3.0	4.34×10^{-3}
		115	1.729	+2.7	-3.0	8.65×10^{-3}
		120	1.649	+2.8	-3.0	1.59×10^{-2}
		123	1.608	+2.8	-3.0	2.18×10^{-2}
		124	1.595	+2.9	-3.0	2.40×10^{-2}
		125	1.578	+2.8	-3.0	2.64×10^{-2}
		126	1.568	+2.9	-2.9	2.89×10^{-2}
		127	1.552	+2.9	-2.9	3.15×10^{-2}
130	1.511	+2.8	-2.9	3.98×10^{-2}		
135	1.448	+2.8	-2.9	5.47×10^{-2}		

process	MC Generator	mass[GeV]	σ [pb]	+Tot%	-Tot%	BR
		140	1.389	+2.7	-2.9	6.87×10^{-2}
		145	1.333	+2.8	-2.8	7.91×10^{-2}
		150	1.280	+2.8	-2.9	8.25×10^{-2}
		155	1.231	+2.7	-2.9	7.34×10^{-2}
		160	1.185	+2.7	-2.7	4.15×10^{-2}
		165	1.141	+2.8	-2.7	2.21×10^{-2}
		170	1.098	+2.8	-2.7	2.36×10^{-2}
		175	1.055	+2.8	-2.7	3.23×10^{-2}
		180	1.015	+2.9	-2.7	6.02×10^{-2}
WH	Pythia8	110	1.060	+3.9	-4.4	4.34×10^{-3}
		115	0.9165	+4.0	-4.5	8.65×10^{-3}
		120	0.7966	+3.5	-4.0	1.59×10^{-2}
		123	0.7347	+3.7	-4.1	2.18×10^{-2}
		124	0.7154	+3.7	-4.1	2.40×10^{-2}
		125	0.6966	+3.7	-4.1	2.64×10^{-2}
		126	0.6782	+3.7	-4.1	2.89×10^{-2}
		127	0.6602	+3.7	-4.1	3.15×10^{-2}
		130	0.6095	+3.7	-4.1	3.98×10^{-2}
		135	0.5351	+3.5	-4.1	5.47×10^{-2}
		140	0.4713	+3.6	-4.2	6.87×10^{-2}
		145	0.4164	+3.9	-4.5	7.91×10^{-2}
		150	0.3681	+3.4	-4.0	8.25×10^{-2}
		155	0.3252	+3.7	-4.2	7.34×10^{-2}
160	0.2817	+4.0	-4.5	4.15×10^{-2}		
165	0.2592	+3.8	-4.3	2.21×10^{-2}		
170	0.2329	+4.0	-4.5	2.36×10^{-2}		
175	0.2089	+4.0	-4.5	3.23×10^{-2}		
180	0.1883	+3.7	-4.2	6.02×10^{-2}		
ZH	Pythia8	110	0.5869	+5.4	-5.4	4.34×10^{-3}
		115	0.5117	+5.6	-5.5	8.65×10^{-3}
		120	0.4483	+5.0	-4.9	1.59×10^{-2}
		123	0.4150	+5.1	-5.0	2.18×10^{-2}
		124	0.4044	+5.1	-5.0	2.40×10^{-2}
		125	0.3943	+5.1	-5.0	2.64×10^{-2}
		126	0.3843	+5.1	-5.0	2.89×10^{-2}
		127	0.3746	+5.2	-5.1	3.15×10^{-2}
		130	0.3473	+5.4	-5.3	3.98×10^{-2}
		135	0.3074	+5.4	-5.2	5.47×10^{-2}
		140	0.2728	+5.6	-5.4	6.87×10^{-2}
145	0.2424	+6.0	-5.8	7.91×10^{-2}		

process	MC Generator	mass[GeV]	σ [pb]	+Tot%	-Tot%	BR
		150	0.2159	+5.7	-5.4	8.25×10^{-2}
		155	0.1923	+5.8	-5.5	7.34×10^{-2}
		160	0.1687	+6.4	-5.9	4.15×10^{-2}
		165	0.1561	+6.5	-6.1	2.21×10^{-2}
		170	0.1408	+6.6	-6.2	2.36×10^{-2}
		175	0.1266	+6.6	-6.1	3.23×10^{-2}
		180	0.1137	+6.4	-5.9	6.02×10^{-2}

4 解析

4.1 Object Definition

Object とは、検出器によって得られた情報から再構成され、解析に用いられる粒子のことである。data から得られる object には正しい起源の粒子だけでなく、誤って識別されたもの含まれており、これらを fake 粒子と呼ぶ。

検出器に残された情報を使って object を定義することが object definition で、正しい起源の粒子の選択効率を良くする (悪くする) と fake も同時に減少する (増える)。従って、出来るだけ signal event を残しつつ、background も十分に少なくする事が出来る object definition を採用する必要がある。

$H \rightarrow ZZ^* \rightarrow 4\ell$ は 4 つの lepton という稀な事象を見るので background は少なくなるが、それと同時に signal からくる event も少なくなる。また、生成断面積からそのほとんどが ggF から出来るので、本研究では VBF から来る event を区別せず inclusive な解析を行っている。つまり、jet の再構成は用いておらず、選別する object は electron と muon に限る。

4.1.1 Electron Reconstruction

Electron の再構成には、大きく High p_T electron (standard electron) を対象とする cluster based algorithm, low p_T の electron (soft electron) を対象とする track based algorithm、前方領域の electron (Forward electron) を対象とするものの 3 つの algorithm がある。ここでは、主に本研究で用いている cluster based algorithm と track based algorithm について説明する。

1. Cluster based algorithm

検出器内部の飛跡検出器と EM calorimeter に記録された情報から再構成される。基本的には、Clusterisation、Shower shape、Track Matching の 3 工程から選別される。

(a) Clusterization

ここでは、EM calorimeter の情報を用いて Sliding-window clustering algorithm という手法を使っている。7x7 cell から作られる cluster を EM cluster と定義し、さらに以下の 3 工程に分けられる。(図 30)

- Tower Building
 $\Delta\eta \times \Delta\phi$ で、 0.025×0.025 の grid に領域を分割する。この grid 中にあるすべての検出器に蓄積されたエネルギーの総量で tower を作る。
- Pre-Cluster Finding
Tower を単位とし、 5×5 の大きさで window を決める。この window を上記の grid の単位で、 $\Delta\eta$ と $\Delta\phi$ の順で移動させながら、 E_T がある一定の threshold (EM で 3 GeV) を超えた window に対して pre-cluster を作る。この window の大きさ及び threshold の値は、noise による fake の pre-cluster を最小にして pre-cluster の efficiency を最大とする大きさに optimize されたものである。

Pre-Cluster の位置は、上記の window の中でより狭い 3×3 の範囲で決めており、pre-cluster の初期位置の周りで使うことにより計算する。その後、grid 単位で 2×2 の範囲内に重なっている pre-cluster があるかを確認し、ある場合には E_T の大きいものが選択される。

- Cluster Filling

Pre-cluster finding で探した位置を起点にして、sliding window 中にあるすべての calorimeter 層の cell から cluster を再構成して行く。このとき、起点を中心にして tower を単位とする四角形の形の中にある cell を用いて再構成を行っている。

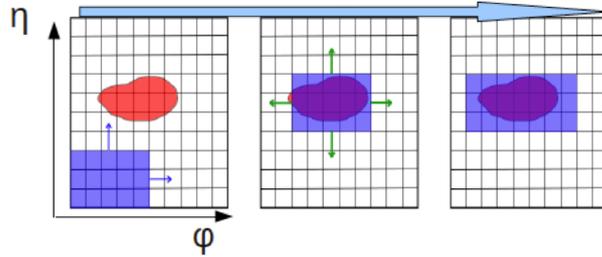


図 30: Sliding Window Clustering Algorithm の簡略図。最初は測定器の情報から tower を作り、window を η , ϕ 方向に移動させながら E_T が閾値を超える、pre-cluster の位置を探す。最後に、その位置を中心として cluster を再構成して行く。

(b) Shower Shapes

得られた shower sampling から、その重心を求める。この重心から shower の形状を決める変数（横/縦への広がり）を求める。EM shower の場合、hadron shower と比べ形状が細く、物質によらず広がりが一定である性質を持つ。

(c) Track Matching

Pixel、SCT 及び TRT の情報から再構成された track の場合、まず track と EM cluster の η と ϕ を比較し、それぞれ 0.2, 0.1 の範囲内で一致することを要求している。続いて、track を calorimeter 内部に引いて各層での深さ、 $\Delta\eta, \Delta\phi$ を計算する。それぞれの track の間の距離 (dR) は第 2 層での情報を用いて計算し、EM cluster に最も近い track を electron track として対応付ける。

TRT の情報だけで構成されている track は、 η 方向を barrel 部と end cap でしか得ることができない。従って、track-matching は ϕ の情報のみを利用して最も近いものを選択する。

2. Track based algorithm

Inner detector で得られた情報から Low p_T の electron や jet 中の electron を構成する。Pre-selection cuts、Clusterisation、Shower shapes、Overlap removal から構成される。

(a) Preselection

electron に対して、以下のような条件を要求する。

- $p_T(\text{track}) > 2 \text{ GeV}$
- $1 < E_T(\text{cluster}) < 5 \text{ GeV}$
- pixel detector hit ≥ 2 、b-layer hit **がある事**
- pixel と SCT の hit ≥ 7

- TRT での high threshold hit >0.08
- (b) Clusterisation
Calorimeter の第 2 層に引いた track から、barrel 部の場合 grid 単位で $\Delta\eta \times \Delta\phi=3 \times 7$ 内に cluster が存在すること、end-cap 部の場合 $\Delta\eta \times \Delta\phi=5 \times 5$ 以内に cluster があることを要求する。また、calorimeter の sampling の層でほぼ全てのエネルギーが再構成される場合は除外するようにしている。
- (c) Shower shapes
Calorimeter の各層から、grid 単位で第 1 層で $\Delta\eta \times \Delta\phi=3 \times 3$ 、第 2 層で 3×3 、第 3 層で 3×1 で、最もエネルギーの高い位置を探す。EM shower の shape に関わる変数は、これら 3 つの点を基準として計算される。
- (d) Overlap removal track が互いに近く、cluster でのエネルギーが同程度の場合、object が重なってしまう場合がある。従って、track が $\Delta\eta < 0.05$ 、 $\Delta\phi < 0.1$ の中に複数存在する場合、 p_T が最も大きいものだけを残し、1 つの track と 1 つの cluster のみを残す。

上記の方法で electron object が選ばれるが、p-p collision を用いる LHC では QCD 事象が多く生成されるので、それでも間違っって選ばれた electron は存在する。これらは、主に π^0 、 γ の conversion、b/c quark の semi-leptonic 崩壊から得られる electron などがある。

- π^\pm と $\pi^0 \rightarrow 2\gamma$
jet の構成要素に π^\pm と π^0 が存在してこれらが近い距離にあった場合、 π^0 が崩壊して前方に 2γ を出し EM cluster の作ることで誤って電子と識別する可能性がある。
- γ conversion
 γ が conversion を通して e^+e^- となり、片方が soft electron で片方がのみが検出された場合 fake になる場合がある。Conversion は、物質との相互作用で起こるため、beam pipe や ID 検出器のある場所で行く。

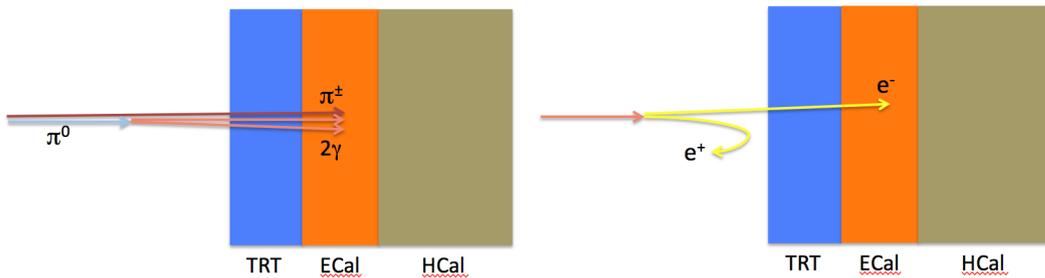


図 31: Electron Fake。左が π^\pm と $\pi^0 \rightarrow 2\gamma$ 、右が γ conversion

- Heavy Flavor

b/c quark が semi-leptonic 崩壊を通して electron を出して fake となる場合である。しかし、electron の周りに jet による寄与があるので、これを用いて落とすことができる (4.2 節を参照)。

4.1.2 Electron Selection

上記のような fake を落としてなるべく質の良い electron を残すために、以下のような条件を要求する。しかし、元から event 数が少ない $H \rightarrow ZZ^* \rightarrow 4\ell$ の場合、electron の selection の段階であまり厳しい選別をしてしまうと統計量が減ってしまう問題がある。そのため、この解析では比較的緩い条件で electron object を定義する。

1. Electron Author

Author は、e の再構成の algorithm を示す tag である。本研究では、cluster base で選別された object と、track-base algorithm と cluster-based の両方を満たす場合を選んでいく。

2. Multilepton ID

2012 年から pileup がより増えたので、pileup 依存性を持つ EM shower の shape や HCal への漏れなどによる cut を緩くし、ID と Calorimeter での track matching を要求して electron の判定を行う electron の ID のことである。この ID の要求を満足させた electron のみを利用している。

3. Kinematics

基本的に、track の情報を用いた electron の同定を行っているので、飛跡検出器が囲む範囲を全て使う必要がある。これから、electron の η は $|\eta| < 2.47$ を要求している。また、なるべく多くの object を同定するために、 E_T の制限を > 7 GeV で低い制限をかけている。

4. Object Quality

Calorimeter に不具合が起きて測定に問題があったと考えられる場合は該当する electron は使わない。具体的には、検出器に不具合 (High voltage の問題等) のあった場所を η と ϕ の二次元 mapping を行い、不具合のあった場所で再構成・識別された電子は排除する。

5. Impact Parameter

3 章で示したように、bunch crossing で起こる平均 pile up の数は約 20 回である。そのため、一つの事象において荷電粒子から再構成される event vertex は複数存在する。Hard process が生じた event vertex からは複数の高い p_T を持つ荷電粒子が生成される。従って再構成された vertex の中で、vertex での $\sum p_T^2$ が最も大きくて再構成された track の数が 3 つ以上あるものを primary vertex と呼ぶ。また、その primary vertex で生成された粒子が二次粒子へ崩壊した点を secondary vertex と呼ぶ。

これから、実際に検出された electron が primary vertex から来たものかを要求する。これは、primary vertex からの距離、つまり impact parameter に制限を掛けて識別を行う。Impact Parameter には x-y 平面上での d_0 、beam 軸方向の z_0 の 2 種類があるが、ここでは track から求められた z_0 が < 10 mm であることを要求している。

6. Overlap Removal

event の中に、内部飛跡検出器から得られた track が全く一致する electron が複数あった場合、calorimeter で得られた cluster の dR が近すぎる electron 同士があった場合は横方向エネルギー (E_T) が高いと評価された方を残し、低い方を除外する。また、electron の muon と飛跡検出器から得られた track 間の距離が $dR < 0.02$ の場合、electron の方を排除する。

4.1.3 Muon Reconstruction

Muon は electron と比べ、質量が約 200 倍程度あるので制動放射の影響を受けにくい。また、calorimeter を Minimum Ionizing Particle(MIP) として通過し、最外層である muon 検出器まで到達する。Muon の再構成及び識別には

- Inner Detector(以下 ID)
- Calorimeter
- Muon Spectrometer(以下 MS)

の三つの情報を使う。再構成の algorithm には STACO と MUID があるが、本研究では基本的に STACO を使っている。これは、ID と MS で検出された飛跡を統計的に平均と取ることによって再演算し、track 情報の精度を上昇させる algorithm である。STACO (STAtistical COmbined) という名前はここから由来している。さらに、その再構成の仕方によって tag が付けられており、それは以下のである。(図 32)

- Stand Alone muons(SA)

MS のみで構成された track を持つ muon を SA muon と言う。Multiple scattering や途中の通過物質でのエネルギー損失を考慮して track を算出する。SA の再構築で muon の同定可能範囲を $|\eta| < 2.7$ までにすることができる。本研究では、Inner Detector の範囲 ($|\eta| < 2.5$) を超える $2.5 < |\eta| < 2.7$ の領域において SA muon を使う。

- ComBined muons(CB)

MS で再構成された track に対応する ID で再構成された track が存在する場合、MS と ID の両方のヒット情報を使って再構成された track を CB muon と言う。具体的には二つの track の parameter を統計的に加えることで CB muon の parameter を計算する。そのため、他の再構成法と比べ運動量の分解能が良い。ID の track 情報は、primary vertex からの muon track の impact parameter の情報も提供してくれる。

- Segment Tagged muons(ST)

ID で再構成した track に対応する MS track は存在しないが MS に hit がある track を持つ muon を ST muon と言う。少なくとも一つの MS hit (segment) と ID track が match することを要求している。当然、track の情報は ID のものとなる。Low p_T muon や MS の構造上、MS に十分なヒットを残さなかった muon を再構成することができる。

- Calorimeter tagged muons(CA)

Track は ID の情報を用いるが、Calorimeter への deposit が、MIP と仮定したときのものと矛盾が無いかを見て muon の判定をする。MS の情報は全く用いない。これらの muon は、配線問題で muon chamber が cover できない $|\eta| \approx 0$ の領域の為に使われる [5]。本研究においては、 $|\eta| < 0.1$ の領域で CA muon を使っている。

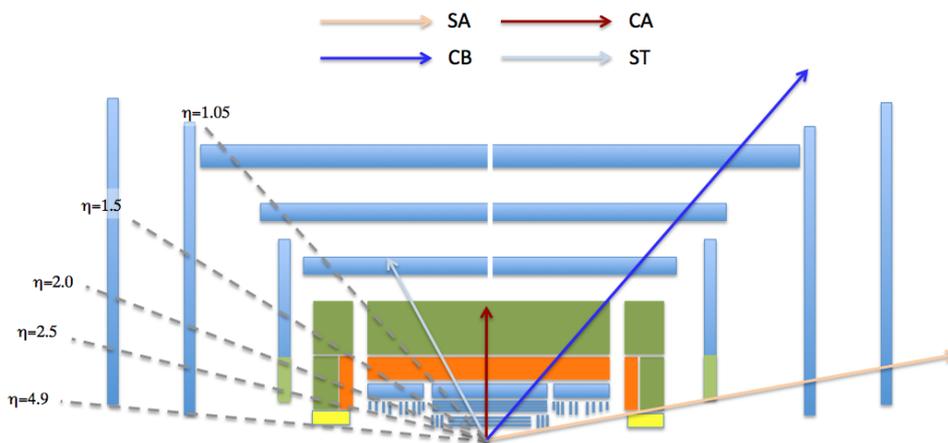


図 32: Muon の同定方法ごとの ID の概略図

上記のように再構成された muon の中には以下のような fake muon が存在することがある。

• π^\pm 起源の fake (図 33)

- Decay on Flight

π^\pm は π^0 とは異なり、その構成要素が $q-\bar{q}$ の関係ではないため、電磁相互作用によって崩壊せず弱い相互作用によってほぼ 100% μ と ν_μ に崩壊する。従って、飛んでいる途中で muon に崩壊し、そのまま検出器で同定されてしまう場合が存在する。

- Punch Through

π^\pm が calorimeter で止まる事無く、そのまま素通りしてしまい MS でも観測された場合である。

• Heavy Flaver 起源の fake (図 34)

Electron と同じく、b/c quark の semi-leptonic 崩壊で muon が得られる場合である。同様に、jet を伴うためこの情報を用いて排除することができ、impact parameter で排除することも出来る (4.2 節)。

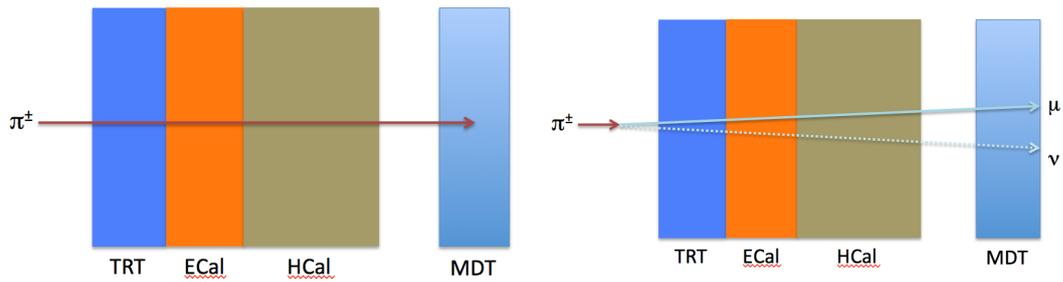


図 33: Muon の punch through(左) と decay in flight(右) の fake event

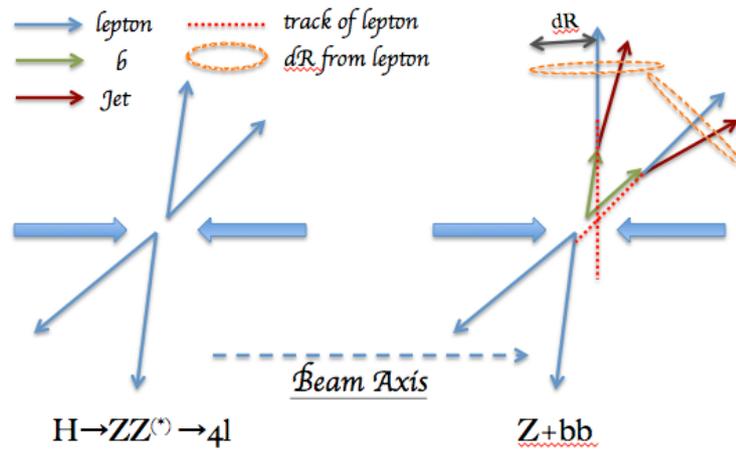


図 34: Heavy Flavor Background の簡略図。右のように、lepton の周りの jet による影響や lepton の track を同定して signal と区別することが出来る。

- 宇宙線起源の muon (図 35)

宇宙から飛んでくる muon が、そのまま検出器を通してまるで中から muon が出来たような軌跡を描く場合である。これは反応とは関係なく起こるので、primary vertex からの impact parameter を要求することで排除可能である。

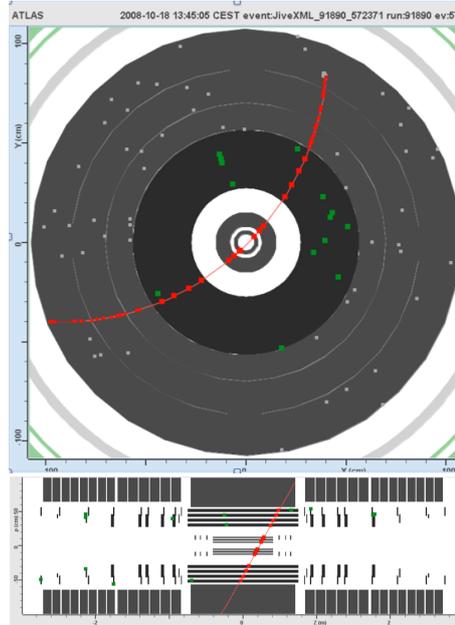


図 35: 宇宙線起源の μ が $Z \rightarrow \mu\mu$ のように見えてしまった event。上が x-y 平面の、下が beam 軸に沿った断面図で見た場合。下を見ると分かるように、中心からのズレを利用して排除することが可能である。

4.1.4 Muon Selection

Electron の場合と同様、本研究では signal の計量を稼ぐために比較的緩い条件で muon object を定義する。

1. μ -ID

4.1.3 の muon の再構成の algorithm (CB,ST,SA,CA) ごとに tag 付けがされており、これらの分類を行っている。

2. Kinematics

electron と同じく、基本的に loose な cut を採用している。さらに、なるべく事象を増やすために検出可能な全ての範囲において検出された muon を使うようにしている。詳しくは、

$$\text{CB+ST: } p_T > 6 \text{ GeV, } |\eta| < 2.5$$

$$\text{CA : } p_T > 15 \text{ GeV, } |\eta| < 0.1$$

SA : $p_T > 6$ GeV、 $2.5 < |\eta| < 2.7$

のように要求している。

3. ID/MS hits

検出器の配置を考慮に入れ、Track に各検出器で反応があったことを要求する。ただし、SA muon の場合は inner detector での track の情報がないため、muon spectrometer からの track 情報で判定を行う。表 3 に、それぞれの muon に対して要求している selection を載せる。

表 3: muon の ID Hit Selection

ID Hit requirements(CB,ST)	
ID Si hit requirement	Expect BLayer hit = false or Number of BLayer hits ≥ 1 No. of pixel hits + No. of crossed dead pixel sensors > 0 No. of SCT hits + No. of crossed dead SCT sensors > 4 No. of pixel holes + No. of SCT holes < 3
TRT hit requirements: $0.1 < \eta \leq 1.9$	Hits + Outliers > 5 & $\frac{Outliers}{Hits+outliers} < 0.9$
TRT hit requirements: $ \eta < 0.1$ or $ \eta \geq 1.9$	if(Hits + Outliers ≥ 5) : $\frac{Outliers}{Hits+outliers} < 0.9$
ID Hit requirements(CA)	
ID Si hit requirement	same as CB, ST
TRT hit requirements: $ \eta < 0.1$	Hits + Outliers < 6 & $\frac{Outliers}{Hits+outliers} < 0.9$
MS Hit requirements(SA)	
CSC&MDT hit requirement	CSC(eta phi)hits >0 & Endcap(em eo) hits >0

4. Cosmic Cut

宇宙線起源の muon を排除するために x-y 平面上での primary vertex からの impact parameter を要求する。値は $|d_0| < 1$ mm で要求している。ただし、SA muon の場合は MS のみからの track 情報となっているので、impact parameter の情報が悪いので要求しない。

5. z_0 cut

Beam 軸方向の impact parameter に、 $|z_0| < 10$ mm を要求する。同じく、SA muon に関しては要求しない。

6. Overlap removal

ID からの track 情報を持つ CB もしくは ST muon と CA muon との track 間の距離が $dR < 0.1$ の場合、CA muon を排除する。また、SA muon と ST muon が $dR < 0.2$ であるとき、SA muon の方を排除する。

4.2 Event Selection

質の良い Object が選別できたら、それらを用いて event を再構成する。H \rightarrow ZZ* \rightarrow 4 ℓ では、最終的に本研究では 4 つ以上の lepton を要求している。lepton の数で event の選別が行われた後は、表 4 の順番でさらに event を絞っている。

この event 選択後の 4 つの lepton から再構成される不変質量 $M_{4\ell}$ の分布を使って結果 (p_0 や limit) を算出する。

表 4: Event Selection の CutFlow

	Cut menu
1	Trigger
2	Charge Selection
3	Kinematic Selection
4	Lepton Separation
5	Zmass Window
6	Normalized Isolation
7	Impact Parameter Significance

1. Trigger

事象の選別の為に、Event Filter(EF) level での trigger を用いて lepton の p_T で trigger をかける。また、EF level で track 周りの p_T の情報を用いて isolation cut を下している。表 5 に、解析で用いた trigger の種類を示す。

trigger の名前には L1、L2 での algorithm や thresholdなどを反映した以下のような規則がある [8]。

$$[TriggerLevel] _ [Triggertype][Threshold][additional\ information]$$

表 5: 解析で用いた trigger の種類。lepton の p_T を用いて threshold をかけており、その値が trigger の名前の tag になっている。例えば、e24vhi_medum1 だと、24 GeV 以上の election があつたときになる trigger である。

Channel	Single-lepton trigger	Di-lepton Trigger
4e	EF_e24vhi_medum1	EF_2e12Tvh_medium1
4 μ	EF_mu24i_tight	EF_2mu13,EF_mu18_mu8_EFFS
2e2 μ	EF_e24vhi_medum1, EF_mu24i_tight	EF_2mu13,EF_mu18_mu8_EFFS

2. Charge Selection

Object の charge が、same flavor, opposite sign(SFOS) を満足している pair が二つ以上あるかを確認する。Z は電氣的に中性であるため、Z が崩壊して出来た lepton の対は電荷量が保存されることから、粒子-反粒子関係で無ければならない。

3. Kinematic Selection

Z が崩壊して出来る lepton には高い p_T を持つことが期待される。従って、あまりにも低い p_T の lepton だけで構成された event が残らないよう、 p_T の threshold cut を要求する。本研究では、3 つの leading lepton に対して、 $p_{T1} > 20$ GeV、 $p_{T2} > 15$ GeV、 $p_{T3} > 10$ GeV を要求している。

4. Lepton Separation

Z が崩壊して 4 つの lepton のみになるので、4 つの lepton が互いに十分 isolated されることが要求される。ここでは、同種粒子間で $\Delta R > 0.1$ 、異種粒子間で $\Delta R > 0.2$ を要求している。

5. Z Mass Window

$H \rightarrow ZZ^* \rightarrow 4\ell$ では、片方が off-shell となっているため、選別を分けねばならない。本研究では、選ばれた 4 つ以上の lepton の組の中で、最も $M_Z = 91$ GeV に近い不変質量を持つ対を leading pair (M_{12})、残った pair の中で最も energetic な方を secondary pair (M_{34}) として選別している。これらにかかる mass window は、 $50 \text{ GeV} < M_{12} < 106 \text{ GeV}$ 、 $M_{\text{threshold}} < M_{34} < 115 \text{ GeV}$ のような選別をしている。 Z の取り得る mass は Higgs 粒子の不変質量 (M_H) に影響されるため、 $M_{\text{threshold}}$ は $M_{4\ell}$ による関数である (図 36)。詳しい値は、表 6 に載せる。各 mass の間では、直線近似で threshold 値を出している。

表 6: $M_{4\ell}$ に応じる M_{34} の minimum threshold

$M_{4\ell}$ (GeV)	≤ 140	160	165	180	≥ 190
threshold (GeV)	17.5	30	35	40	50

6. Normalized Isolation

ZZ^* が 4 つの lepton に崩壊するので lepton に伴われる jet は生じない。従って、 Z から来る lepton は周りに余分な情報が少ない、質のいい lepton であることが要求される。これは、lepton の track の周りで検出された p_T の和と lepton の p_T の比 ($\sum p_T / p_T$)、calorimeter での lepton の cluster を除いた周りの E_T の和と lepton の E_T の比 ($\sum E_T / p_T$) が十分に小さい事を要求することで確認することが出来る。これらをそれぞれ track isolation と calorimetric isolation と呼ぶ。図 37 に、4 lepton 候補全てを、electron と muon に分けてこれらの分布を示す。

この解析では、lepton の track から $dR < 0.2$ の範囲の中の track isolation に < 0.15 、electron の calorimetric isolation に < 0.2 、muon の calorimetric isolation に < 0.3 であることを要求している。

7. Impact Parameter(IP) Significance

この channel では、Higgs 粒子が直ちに 2 つの Z に、 Z は直ちに 4 つの lepton へ崩壊する。従って、lepton の track の軌跡を描くと原理的には一点に集まるはずである。言い換えれば、primary vertex から近い軌跡を持つことになる。

しかし、 b のように比較的長い寿命の粒子が崩壊して得られた lepton は、崩壊まで約数 mm

を走ってから崩壊するので、signal からの event に比べると大きい impact parameter (primary vertex からの距離) を持つ。この特性を用いてその測定誤差まで考慮に入れた impact parameter significance(d_0/σ_{d_0}) で選別を行うことで、 $Z+b\bar{b}$ のような background を排除することができる。(図 38)

本研究では、electron に <6.5 、muon に <3.5 を要求している。electron は、磁場の中での加速運動により制動放射を起し、軌跡が変わってしまうのでより緩い制限をかけている。

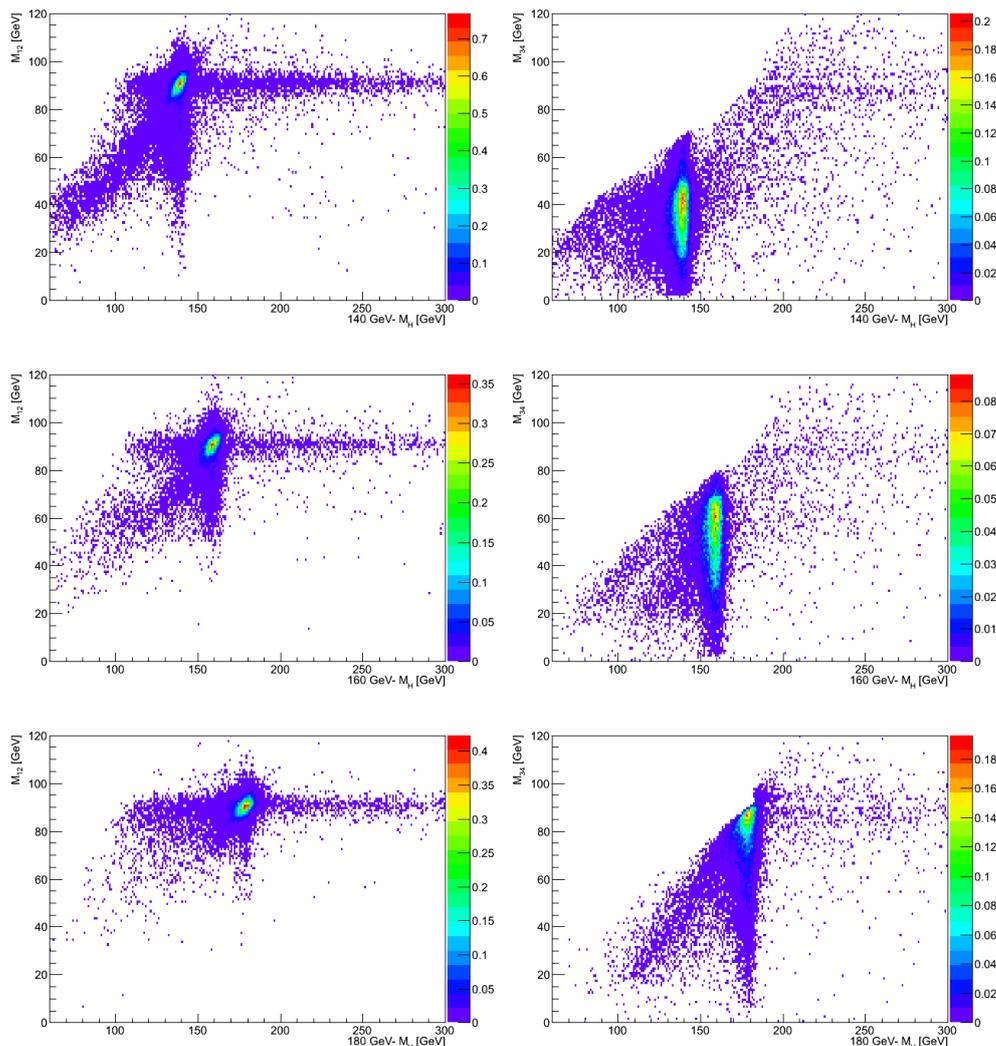


図 36: 上から $M_H = 140, 160, 180$ の時の signal の M_{12} (左) と M_{34} (右) vs $M_{4\ell}$ の分布図。 Z mass window cut の基準は、reconstruct された M_H の値から決めている

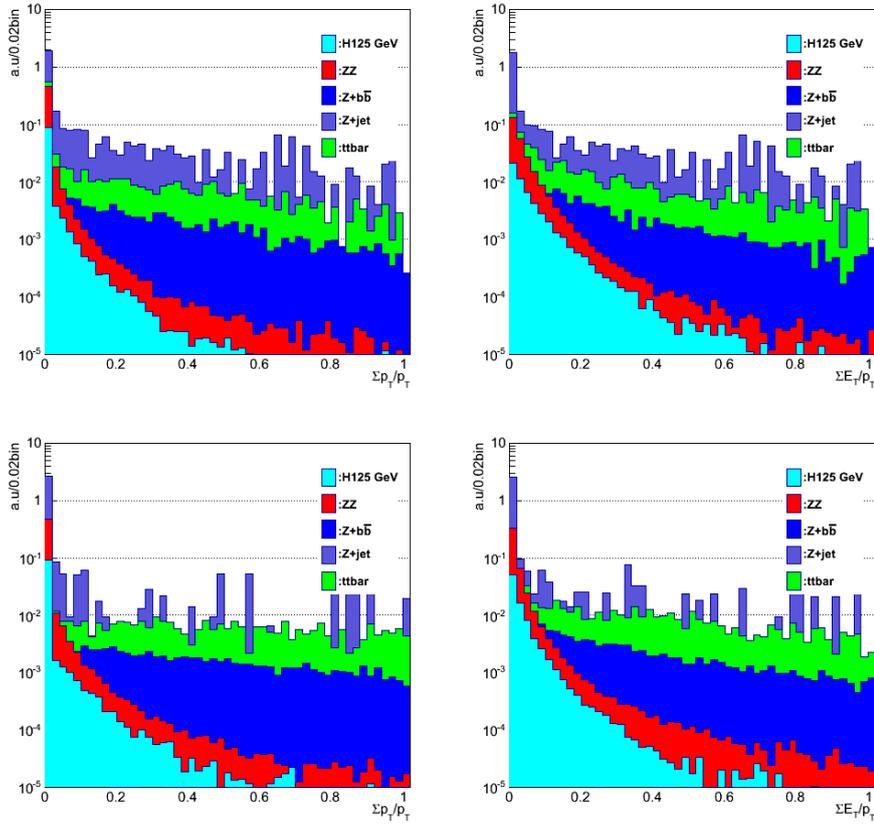


図 37: 4 lepton の normalized isolation の分布。左が $\sum p_T/p_T$ 、右が $\sum E_T/p_T$ 、上が e 、下が μ 。 $tt̄$ (緑) や Z +jet(青) の分布が、 ZZ^* (赤) や signal(シアン) と比べて大きい値を示している。

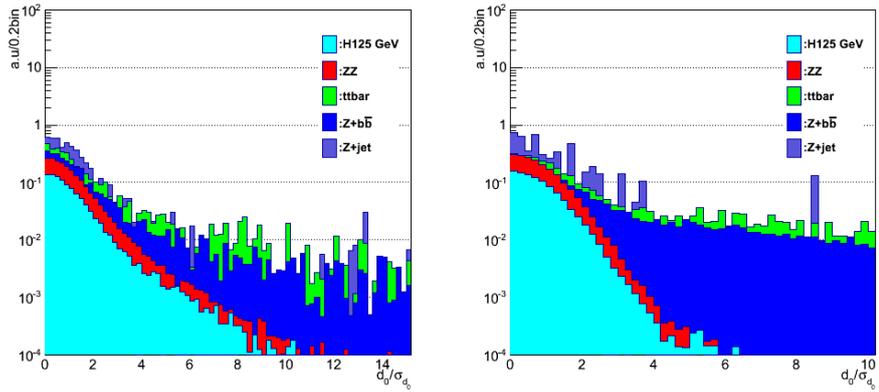


図 38: Lepton の IP significance の分布。左が e 、右が μ 。 $tt̄$ (緑) や Z +jet(青) の分布が、 ZZ^* (赤) や signal(シアン) と比べて大きい値を示している。また、electron は制動放射のため、muon と比べ大きい値を持つことも確認できる。

4.3 Background Estimation

本研究の探索領域である $110 \text{ GeV} < M_{4\ell} < 180 \text{ GeV}$ のでは、片方の Z が off-shell となるため比較的に小さい p_T の値を持つ lepton の存在が予想される。従って、 $Z+b\bar{b}$ 、 $Z+\text{jet}$ や $t\bar{t}$ による background が予想される (図 39)。

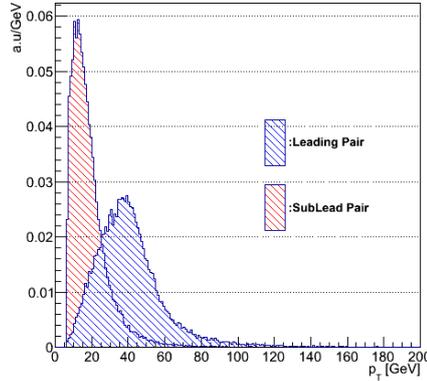


図 39: Signal の Leading と Sub-leading pair の lepton の p_T 分布

この中で、subleading lepton pair が muon である場合 ($Z+X_\mu X_\mu$)、 $Z+b\bar{b}$ 、 $t\bar{t}$ が主な background となり、electron の場合 ($Z+X_e X_e$) $pp \rightarrow Z+\text{jet} \rightarrow 4\ell$ が主な background となる。Electron と muon に対する background の性質はそれぞれ違うため、本研究ではこれらを分け、それぞれに対して background の感度を増やした control region を生成し、control region の data から background の量を評価し、形は MC に基づいて行っている (図 40)。

ただし、 ZZ^* background のみは signal region で MC から評価されたものをそのまま使っている。

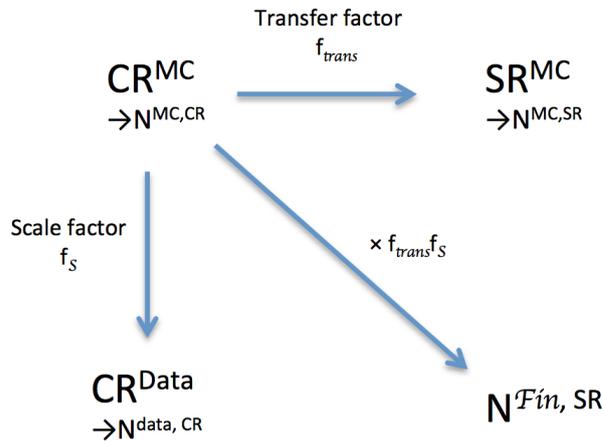


図 40: Control region からの signal region での background 評価方法の概略図。 ZZ^* 以外の background に対しては、MC だけで control region から signal region への transfer factor を、control region での data と MC の差から scale factor を計算し、最終的に signal region での background の event 数を計算する。

4.3.1 $Z+X_\mu X_\mu$ background

$Z+X_\mu X_\mu$ の評価は、 $b\bar{b}$ や light jets からの寄与を大きくした二つの control region (CR1,CR2) を作って行った。これらの control region から signal region (SR) に移行する係数 (transfer factor) と、CR で fit を行って MC を data に補正する係数 (scale factor) から SR での event を見積もる。

1. Control Region の生成

- Heavy Flavor quarks (CR1)

$b\bar{b}$ からの影響を増やすためには、signal region での cut を変える必要がある。 b の semi-leptonic decay から来る lepton は jet を伴うこと、また b の崩壊までの飛行距離によって impact parameter が大きくなることから、isolation の cut や IP significance の cut を緩くすれば良い。

本研究では、secondary pair にかける isolation の cut を排除し、二つの lepton のどれかは IP significance cut を満足しないような条件を加えて control region を作った。こうすることで signal region での主な background である ZZ^* からの寄与や、 π などから来る μ を排除することができ、 $b\bar{b}$ の event を選別して増幅させることができる。この control region での data 及び MC の分布と、fit の様子を図 41 に示す。 $Z+b\bar{b}$ の fit 関数は、Crystal-Ball resolution function で convolute された、Breight-Wigner 関数を使っている。また、 $t\bar{t}$ の場合は 2 次までの Chebychev 多項式を用いている。これら二つの background の評価は、この CR1 で評価している。

- Z +light Jet (CR2)

上記の方法は heavy quark からの寄与を調べるために行われたが、jet 中の π の punch through などからの background も予想されるため、それらを見積もる必要がある。従って、上記の control region から、secondary pair が impact parameter ではなく track isolation selection を必ず外すよう要求する control region を作った。このような要求をすると、 $Z+b\bar{b}$ や $t\bar{t}$ だけでなく、light jet からの event も同時に増幅させることが出来る。

Fit の関数は基本的に heavy Flavor の control region での方法と同じ関数を用いている。CR2 では、 Z +jet の event のみを評価する。

2. Transfer factor

Control region から signal region への transfer factor の評価には、MC を使ってそれぞれの control region と signal region での event の比を transfer factor として使っている。つまり、

$$\frac{N_{CR1}}{N_{SR}} \quad (Z + b\bar{b}, t\bar{t}) \quad (5)$$

$$\frac{N_{CR2}}{N_{SR}} \quad (Z + \text{light jet}) \quad (6)$$

から transfer factor を計算している。これらの transfer factor を表 7 に載せる。

表 7: 各 background の transfer factor

transfer factor (%)	
$Z+b\bar{b}$	3.14 ± 0.33
$Z+\text{light jet}$	2.04 ± 1.6
$t\bar{t}$	0.21 ± 0.11

3. Background の評価

$Z+b\bar{b}$ や $t\bar{t}$ の CR1 での event 数は data への fit から評価している。これから、CR1 で評価された各 background と data 差を scale factor として使っている。そして、上記の transfer factor から、signal region での event 数を求めている。

$$N_{SR} = N_{fit} \cdot R_{CR}^{channel} \cdot f_{trans} \quad (7)$$

$Z+\text{light jet}$ の場合、CR2 での event には $Z+b\bar{b}$ と $t\bar{t}$ もかなり含まれているので、まずこれらの分を引かなければならない。このために、CR2 で評価された $Z+b\bar{b}$ の MC event に CR1 で求められた scale factor を掛けて、CR2 での $Z+b\bar{b}$ を評価する。そして、CR2 で data への fit を行い、Crystall-ball によって評価された event から、 $Z+b\bar{b}$ の event 数を引くことで、 $Z+\text{light jet}$ の event 数を評価する。

それから CR1 でと同じく、MC によって評価された $Z+\text{jet}$ の event と data から評価された量に scale し直し、上記の transfer factor から signal region での $Z+\text{light jet}$ の event を見積もっている。

各 channel での background event 数は、CR1,2 で MC から求めた各 process ごとに 4μ と $2e2\mu$ からの寄与 ($R_{CR}^{channel}$) を比を使って評価している。最終的に評価された background を表 8、9、10 に載せる。

表 8: CR2 での Z +light jet の event

Estimated CR	Fit Yield	$Z+b\bar{b}$ MC	$Z+b\bar{b}$ MC (scaled)	Z +light jet (calculated)	Z +light jet MC
CR1	72.21	38.35	72.21		
CR2	108.15	46.44	87.44	20.71	34.39

表 9: 各 background の control region での Fit 結果と MC の数

	Estimated CR	Fit Yield	MC
$t\bar{t}$	CR1	72.71	72.29
$Z+b\bar{b}$	CR1	72.21	38.35
Z +light jet	CR2	20.71	34.39

表 10: 各 background の signal region での event 数

	$t\bar{t}$	$Z+b\bar{b}$	Z +light jet
4μ	0.069 ± 0.042	1.211 ± 0.321	0.187 ± 0.101
$2e2\mu$	0.084 ± 0.051	1.057 ± 0.281	0.297 ± 0.275

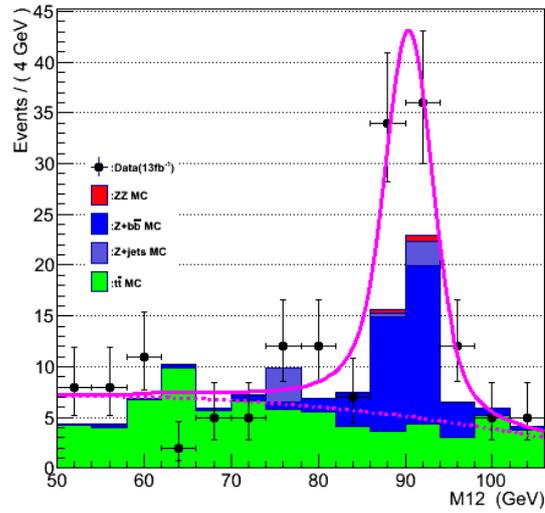


図 41: CR1 での M_{12} の分布。standard selection から、sub-leading pair の μ に対して、isolation cut を外し、IP significance を必ず外すように selection を要求している。 $t\bar{t}$ や Z +jet の MC にそれぞれに行った fit の結果から関数の形を決め、それを data に fit して得られた event 数で $t\bar{t}$ や Z +heavy flavor jet の background を評価している。

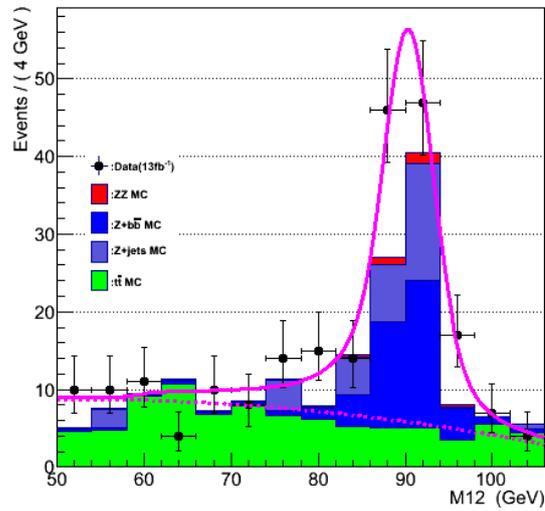


図 42: CR2 での M_{12} の分布。standard selection から、sub-leading pair の μ に対して、calorimetric isolation, IP を外し、track isolation を必ず外すように selection を要求している。 $t\bar{t}$ や Z +jet の MC にそれぞれに行った fit の結果から関数の形を決め、それを data に fit して得られた event 数で Z +light jet の background を評価している。

4.3.2 $Z+X_e X_e$ background

$Z+X_e X_e$ の評価では、 Z +jet、 Z +bb、 $t\bar{t}$ をすべて含んだ評価を行う。Control region は electron の ID を緩くして fake electron を増やすため、sub-leading pair の isolation や IP significance を排除することで作っている (以下 eeCR1)。 $Z+X_\mu X_\mu$ の場合とは違って IP significance を失敗することを要求していないので、 ZZ^* の event も含む解析となる。そこで、 Z +inclusive jet や $t\bar{t}$ のみで評価を行うために、truth の情報や、electron が reconstruction されるときの情報を用いて truth と reconstruction での category を作って、それぞれの category について event や efficiency を求めている。

ただし、 Z +jet の場合、統計量不足で形が揃わないため、形を作るための control region (以下 eeCR2) を作って、これから $Z+X_e X_e$ の signal region での background の shape を作った。

最終的には、 ZZ^* 以外の background をこの方法で評価した。

この方法を流れを説明する。まず、CR での data の event 数を electron の reconstruction 情報で分類 (大文字の EE など) した category で求める (N_{IJ}^{Data})。次に、MC を MC truth の情報を使って (構成比は data と MC で同じと仮定している) electron の起源 (小文字の ee など) で分類する。この構成比を P_{ij}^{IJ} と書く (表 11、12)。

さらに、MC を使って、electron の各起源に対する reconstruction+idenfitication efficiency を求める (ϵ_{Ii})。以上より、ee は ZZ 由来であるため、これを除いて最終的な background を算出する。このようにすると、data からの background の量を見積もる際に ZZ はもちろん、signal が含まれて評価されてしまう可能性を排除できる。(図 43)

1. Control Region の生成

- Relaxed Electron ID

Signal region での standard analysis は、muiltilepton requirement による方法で electron の ID を判別している。Control region ではこれらの要求をいくつか排除することで緩くし、fake の electron を増やして background を評価しており

- Silicon detector で、7 個以上の hit がある
- Pixel detector での hit がある
- HCAL への E_T leakage と EM calorimeter の E_T の比 (EtHad/Etcl)、EM shower の shape ($R_{\eta,w_{\eta 2}}$)
- Calorimeter での shower の幅 (wstot)
- Calorimeter の back layer で観測されたエネルギーと全体のエネルギーの比 (f3)
- Calorimeter の middle layer に引かれた track からの phi と cluster phi の差 ($\Delta\phi$)
- Crack region における TRT での high threshold hit と low hit の比を見て (charged hadron 排除)

のように要求している [6],[7]。

- Truth/Reconstruction Categorization

Truth の category では、MC を base に選別された electron を electron(e)、semi-leptonic decay からの electron(b)、photon conversion からの electron(c)、light jet からの electron(f) などに分類する。これは、electron の truth type と、その起源となる粒子の type を見ることで分類を行っている。

Reconstruction category では、検出器で測定された variable を持って、これらが electron like(E) か、fake like(F) かを見ていて、その方法は、

Electron(E) : $f_1 > 0.1, \text{has BLayer (if expected), if } (|\eta| < 2.0) \text{ rTRTCut}(\eta \text{ dependant})$
 else ($R_\phi > 0.9$)
Fakes(F) : others

のようになる。

それぞれの CR の category における event 数を表 11 と表 12 に、reconstruction category での分布を図 44 に載せる。category の名前は、subleading pair の lepton の p_T の大きい順になっている。

表 11: $Z+X_e X_e$ の control region での各 reconstruction category ごとの event 数

	$Z(ee)+X_e X_e$		$Z(\mu\mu)+X_e X_e$	
	Data	MC	Data	MC
EE	82	77.8 ± 8.7	73	67.9 ± 7.1
EF	52	49.1 ± 9.2	64	63.0 ± 11.3
FE	36	26.1 ± 7.0	28	29.2 ± 5.7
FF	40	67.2 ± 12.8	45	76.5 ± 14.6
Total	210	220.2 ± 19.3	210	236.6 ± 20.6

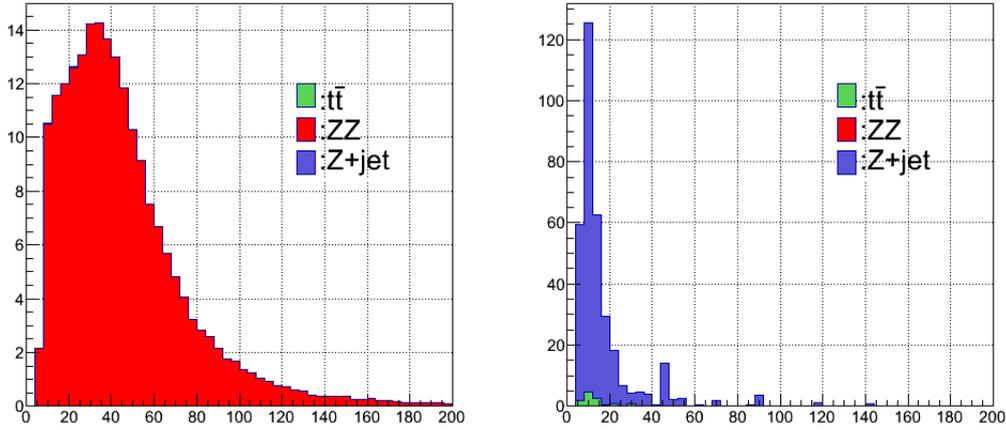


図 43: $4e, 2\mu 2e$ を合わせた、各 truth category における subleading pair electron の p_T distribution. 左が ee, 右が $\mu\mu$ を示す。

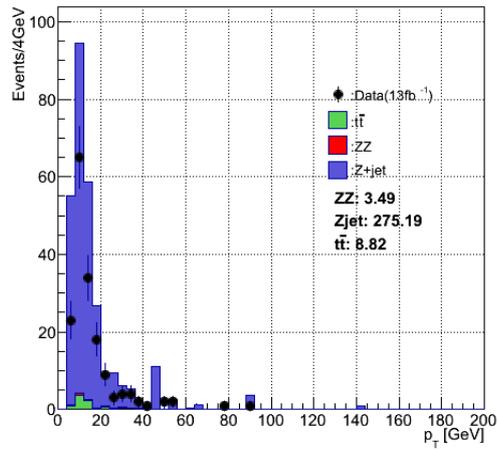
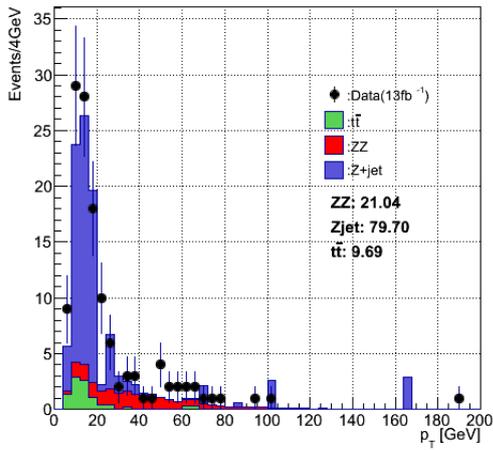
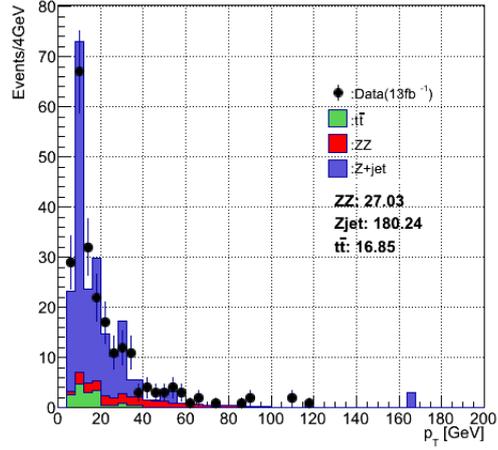
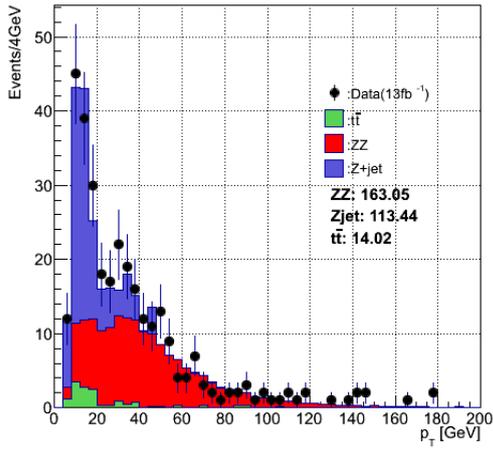


図 44: $4e, 2\mu 2e$ を合わせた、各 reconstruction category における subleading pair electron の p_T distribution。左上から右方向に EE, EF, FE, FF での分布を示す。

表 12: CR1 での purity の計算に使われた、category ごとの MC event の分布。
 $Z^{(ee)}+X_e X_e$

	ee	eb	ec	ef	be	bb	bc	bf	ce	cb	cc	cf	fe	fb	fc	ff
EE	37.3	0.7	1.1	2.3	0.1	4.9	0.4	1.2	0.5	1.8	0.0	11.1	5.5	0.3	3.1	6.9
EF	5.7	0.1	0.3	13.7	0.0	0.6	0.3	5.4	0.1	0.4	0.0	13.6	0.0	0.2	0.3	8.0
FE	4.5	0.0	0.1	0.0	0.0	1.4	0.0	0.2	0.2	1.1	0.0	6.2	0.5	5.5	5.5	1.1
FF	0.6	0.0	0.0	0.1	0.0	0.1	0.0	4.0	0.0	0.0	0.0	3.4	0.0	4.6	1.5	52.9

	ee	eb	ec	ef	be	bb	bc	bf	ce	cb	cc	cf	fe	fb	fc	ff
EE	41.8	0.1	0.7	0.1	0.0	3.4	0.6	5.4	0.6	3.6	0.0	3.1	0.1	0.9	3.5	4.0
EF	6.1	0.0	0.3	0.5	0.0	0.7	0.2	6.7	0.1	0.4	0.1	23.1	0.0	0.1	0.1	22.2
FE	5.2	0.0	0.1	0.0	0.0	0.8	0.1	0.1	0.2	1.0	1.3	1.5	0.2	5.4	1.9	11.4
FF	0.7	0.0	0.1	0.1	0.0	0.1	0.0	0.6	0.0	0.1	0.0	6.3	0.0	0.3	3.5	64.7

2. Transfer factor

Transfer factor は、truth と reconstruction の各 category から、signal region になるように cut をかける前後での event の数で efficiency を出している。ここに、各 electron の category ごとの efficiency を表 13 と表 14 に載せる。

表 13: $Z(ee)+X_e X_e$ の control region での各 category から signal region への cut efficiency

	e	b	c	f
Leading Electron				
E	0.98 ± 0.02	0.13 ± 0.09	0.50 ± 0.10	0.22 ± 0.08
F	0.96 ± 0.08	0.64 ± 0.20	0.60 ± 0.15	0.28 ± 0.05
Sub-Leading Electron				
E	0.951 ± 0.03	0.16 ± 0.09	0.33 ± 0.15	0.01 ± 0.02
F	0.929 ± 0.10	0.72 ± 0.18	0.09 ± 0.18	0.05 ± 0.02

表 14: $Z(\mu\mu)+X_e X_e$ の control region での各 category から signal region への cut efficiency

	e	b	c	f
Leading Electron				
E	0.98 ± 0.02	0.13 ± 0.09	0.50 ± 0.09	0.14 ± 0.06
F	0.96 ± 0.06	0.64 ± 0.26	0.61 ± 0.15	0.21 ± 0.04
Sub-Leading Electron				
E	0.95 ± 0.03	0.17 ± 0.09	0.07 ± 0.09	0.12 ± 0.07
F	0.94 ± 0.09	0.17 ± 0.30	0.08 ± 0.13	0.02 ± 0.01

3. Background の評価

最終的な background の大きさは、data から見積もっている。ただし、efficiency は truth の category を含めて出しているが、data ではそのような categorization は不可能なため、MC の category から得られた purity (P_{ij}^{IJ}) で data の truth categorization を作っている。最終的に signal region で期待される background の量 (N_{SR}^{Data}) は、control region での reconstruction category の data の数に purity と表 13、14 から得られる efficiency (ϵ_{Ii}) を使って、

$$N_{SR}^{Data} = \sum_{(IJ)} \sum_{(ij)} N_{IJ}^{Data} P_{ij}^{IJ} \epsilon_{Ii} \epsilon_{Jj} \quad (IJ = EE, EF, \dots), (ij = eb, ec, \dots) \quad (8)$$

$$P_{ij}^{IJ} = \frac{N_{ij}^{MC, IJ}}{\sum_{(ij)} N_{ij}^{MC, IJ}} \quad (9)$$

から計算され、以下のような結果となった

$$Z(ee)+X_e X_e: 7.2 \pm 0.9$$

$$Z(\mu\mu)+X_e X_e: 2.5 \pm 0.4$$

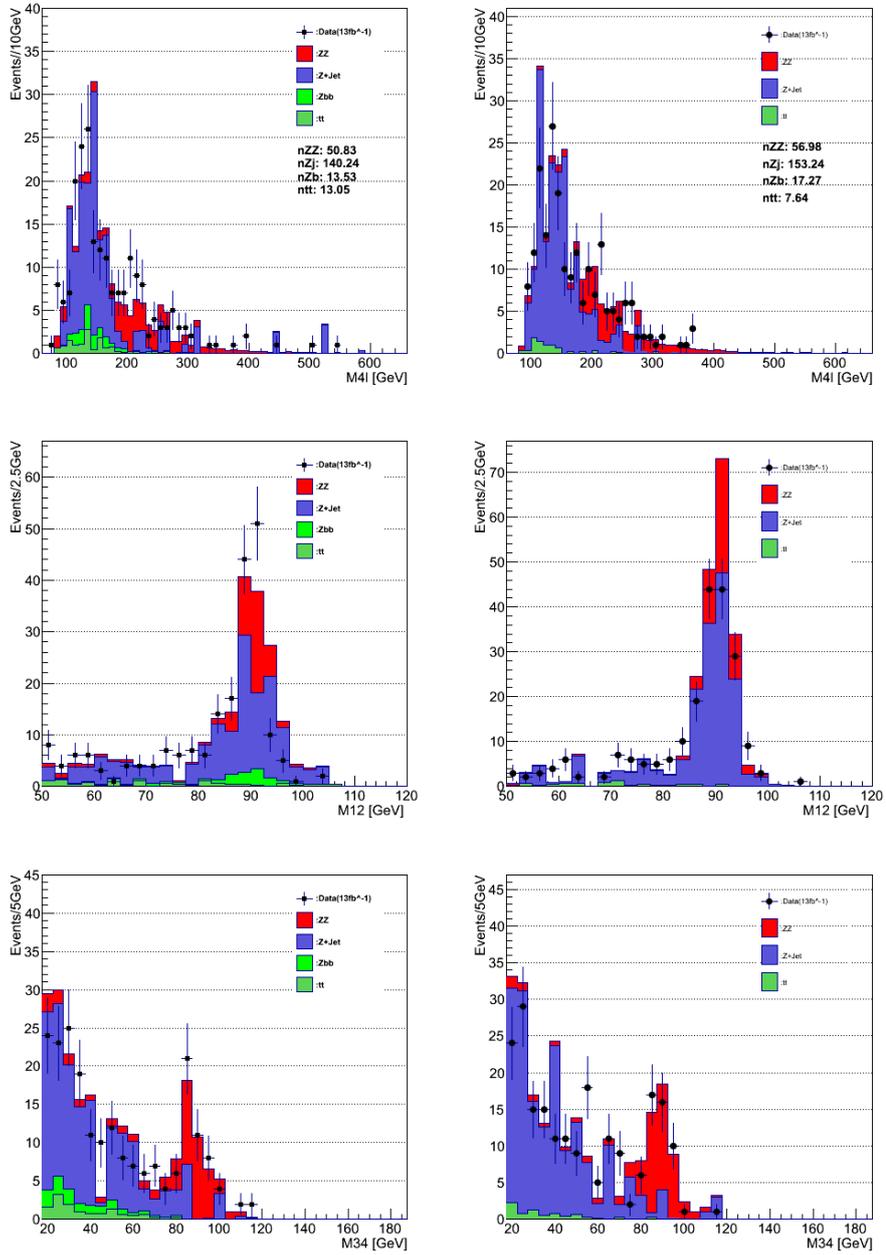


図 45: CR1 での各 channel の M_{4l} 、 M_{12} 、 M_{34} の図。上が $4e$ 、下が $2\mu 2e$ 。Z+jet に spike が見える。

4. Shaping

$Z+X_eX_e$ の background では $Z+jet$ からの寄与が一番大きい。しかし、今回用いた $Z+jet$ の MC sample はその luminosity が最大で 7 fb^{-1} で、実際の data の半分程度の統計量しかない。さらにこの channel の特性上 $Z+jet$ の event もかなり少なくなり、図 45 のように control region を作って back ground の事象を増やして signal region での event 数を評価するだけではその形が整われない。(図 45)

統計量の不足を補完するために、本研究では $Z+X_eX_e$ を BG として持つ $4e$ channel と $2\mu 2e$ channel を合わせて、shape を得るために $Z+e$ の control region (eeCR2) を作った。ここで、 e には eeCR1 でと同じく、electron の ID を緩くし、isolation と IP significance を外している。また、この control region では sub-leading pair は作れないので、それに関わる selection (kinematic selection, sub-leading pair mass window) は外している。

この control region から、electron ID, isolation, IP significance を要求したときの p_T と fakerate との関係を調べ、それを $Z+X_eX_e$ の control region (eeCR1) に反映することで signal region での $Z+jet$ background の shape を出している (図 46)。この際、fakerate は p_T の関数であり、10 GeV ごとに計算し、50 GeV まで p_T の関数として、それ以上の領域からは定数としている。

これはあくまで shape を得るためだけであって、最終的な event 数の評価は eeCR1 から得られた event の数を使った。

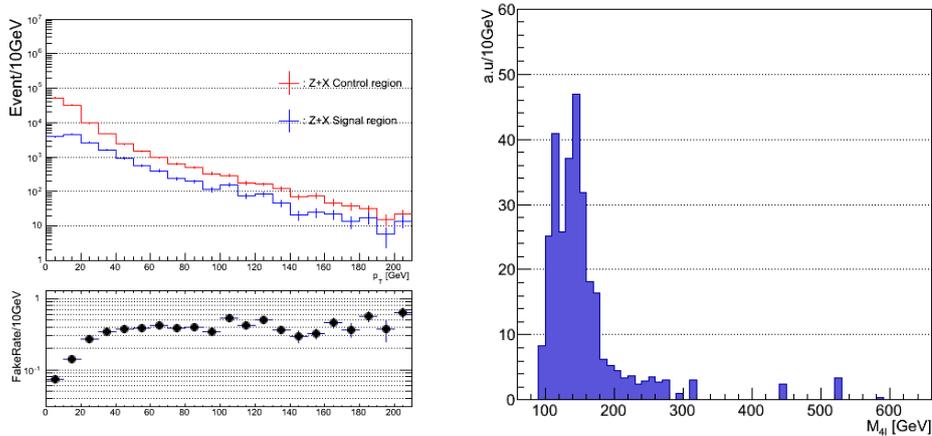


図 46: $Z+e$ CR での p_T 分布と fakerate(左) と、 $Z+X_eX_e$ CR で、 $4e$ と $2e2\mu$ の $Z+jet$ event を合わせた分布。これに左の fakerate を反映して、SR での background の shape を作る。

Shape をする前後での signal region における $M_{4\ell}$ 分布の様子を、図 47 に載せる。

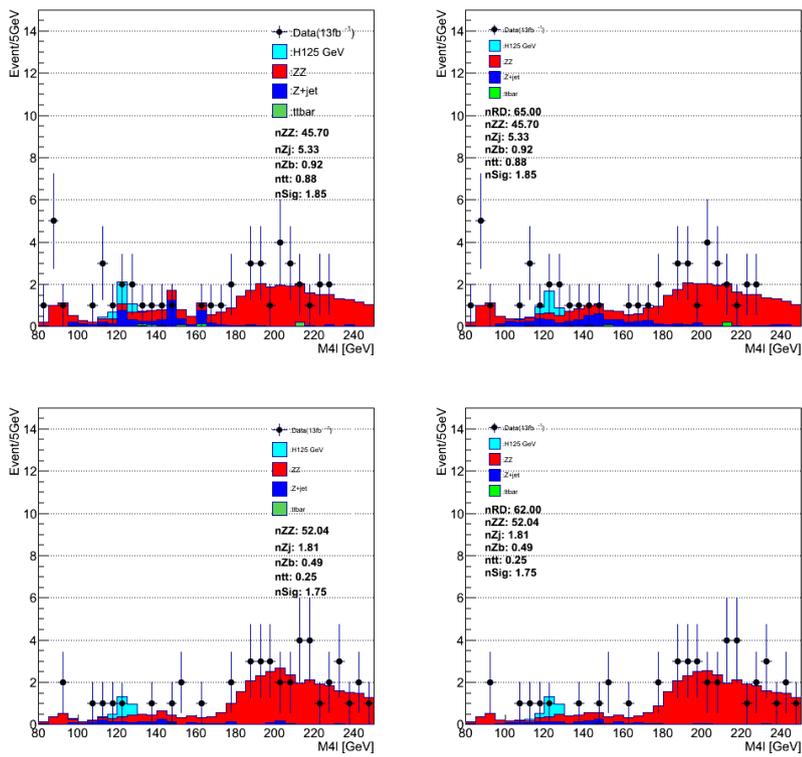
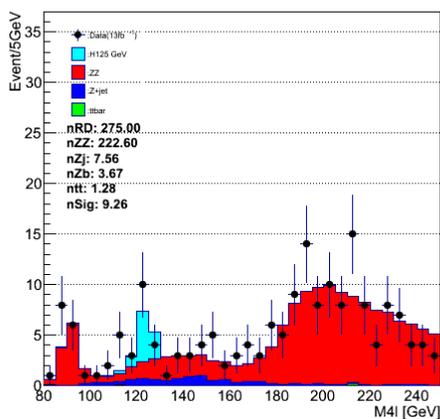


図 47: Shape を使う前後での $4e$ と $2\mu 2e$ の $M_{4\ell}$ の分布。上が $4e$, 下が $2\mu 2e$ で、左が eeCR1 から直接 extrapolate した場合、右が eeCR2 から得た shape を使った場合

4.3.3 Cut based 解析の結果

最終的に得られた data 及び MC の event 数、及びその分布を以下に示す。(図 48、49、表 15)

図 48: 4 channel を全部合わせた $M_{4\ell}$ の分布。



combined	
Luminosity	13 fb^{-1}
data	275
$M_H = 125 \text{ GeV}$	9.3 ± 0.1
BG total	235.1 ± 1.1
ZZ^*	222.6 ± 0.8
$Z + jet$	7.6 ± 0.7
$Z + b\bar{b}$	3.7 ± 0.1
$t\bar{t}$	1.3 ± 0.3

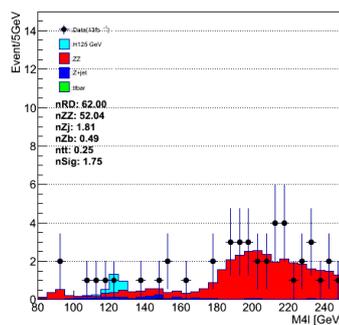
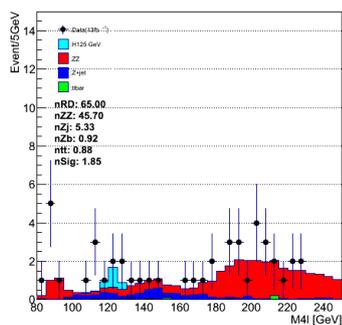
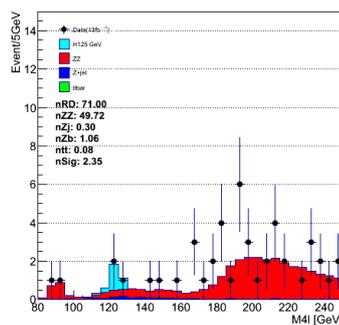
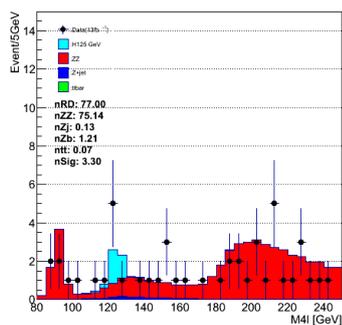


図 49: 各 channel の $M_{4\ell}$ distribution. 左上から右に 4μ 、 $2e2\mu$ 、 $4e$ 、 $2\mu2e$ の分布となる。

表 15: 各 channel における BG と data の low mass ($100 \text{ GeV} < M_{4\ell} < 160 \text{ GeV}$) と high mass ($M_{4\ell} > 160 \text{ GeV}$) の event 数。 $ee\mu\mu$ は $ee\mu\mu$ と $\mu\mu ee$ を合わせている。

	$\mu\mu\mu\mu$		$ee\mu\mu$		$eeee$	
	Low mass	High Mass	Low mass	High Mass	Low mass	High Mass
ZZ^*	8.3	60.4	6.9	91.9	4.1	39.0
$Z, Z\bar{b}b, t\bar{t}$	1.0	0.4	2.3	1.2	4.0	2.5
Data	16	56	14	115	13	45
$M_H = 125 \text{ GeV}$	3.3		4.1		1.9	

4.4 Multivariate Analysis

$H \rightarrow ZZ^* \rightarrow 4\ell$ は、全領域に渡って ZZ^* からの background を持つ。これは ZZ^* から 4ℓ になる過程が signal と全く一緒である故、kinematics に基づいた cut では排除ができない irreducible な background となる。従って本研究は、多変数解析手法 (Multi-Variate Analysis、以下 MVA) という手法を用いて、cut base の selection で残った event から ZZ^* と signal の違いを探し、significance の改善を試みた。

4.4.1 TMVA と BDT

一般に High Energy Physics では多くの物理量 (変数) を使った解析を行う。これらの変数は、その物理的意味を考えて真に欲しい event と background を分ける手段として使うことが出来るが、これらの変数の中で相関関係にあるものが存在する場合、これを用いた解析が有効となる。しかし、高い自由度を持つ解析で各変数の signal と background を分ける効率やそれぞれの相関などを統計的に処理するのは非常に難しい作業となる。このような多くの変数を用いた解析を可能にする tool に TMVA (Tools of MultiVariate Analysis) がある。

TMVA では、MVA を行う際に必要とされる統計的手法 (method) をいくつか用意しており、この中で良く使われる method として Boosted Decision Trees (以下 BDT) という手法がある。本研究においては、この BDT method を使うことで ZZ^* background の排除することで significance に改善があるかを調べた。BDT 手法については、Appendix A で述べるとする。

4.4.2 MVA の input variable

本研究においては、MVA を用いた解析のための input variable として、 $H \rightarrow ZZ^* \rightarrow 4\ell$ の過程における Z の mass と、 Z や lepton の角度分布を用いている。これらの変数を使うことで、 4ℓ と 2 つの Z で構成された全全体の kinematics を含めて解析を行うことが可能となる。ここでの角度は 2 つの production angle (θ^* , ϕ_1) と、3 つの decay angle (θ_1 , θ_2 , ϕ) の 5 つが存在し、Higgs 粒子の spin と parity によって特徴的な分布を示す [11]。

それぞれの角度の定義は、

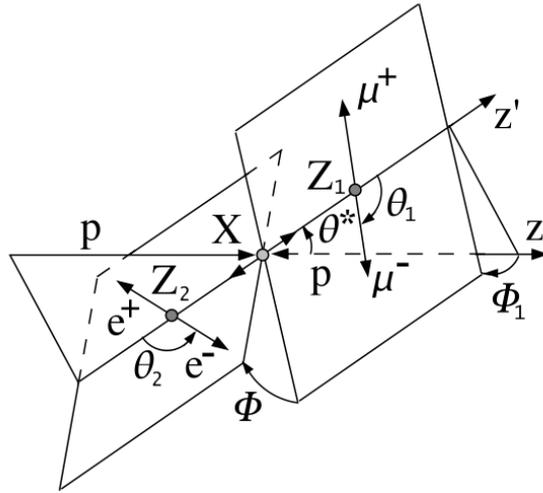


図 50: 角度の定義。本研究では、 Z_1 を primary pair の Z にしている。

- θ_1 と θ_2 は、それぞれの Z の静止系で負の電荷の lepton とその対応する Z の進行方向 vector の間の角度。
 - ϕ は、Higgs 粒子の静止系において、primary Z の進行方向 (z' 軸方向) 周りでそれぞれの Z の崩壊面が成す角度。
 - ϕ_1 は、Higgs 粒子の静止系で見た primary Z の崩壊面と、beam 軸 (z 軸) と z' 軸が作る、 z' 軸周りで角度の差。
 - θ^* は、Higgs 粒子の静止系で見た primary Z の進行方向と beam 軸が成す角度
- のようになる。BDT training の際に渡している各変数の分布を図 51、52 に示す。

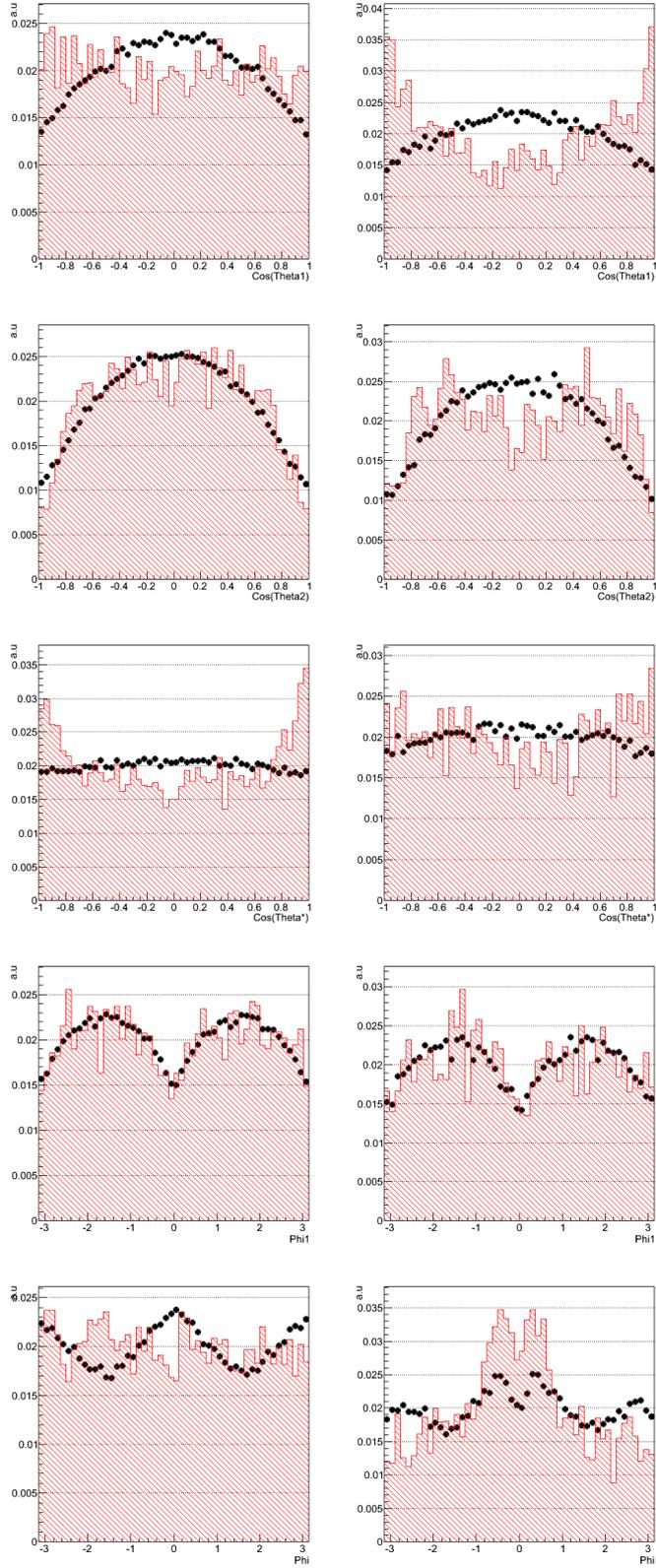


図 51: MVA の MC の ZZ^* (赤線) と signal (点) の input variable の角度分布。左が truth、右が reconstruction で、上から $\cos(\theta_1)$, $\cos(\theta_2)$, $\cos(\theta^*)$, ϕ_1 , ϕ 。truth の場合、 p_T , η , mass window の cut を入れている。

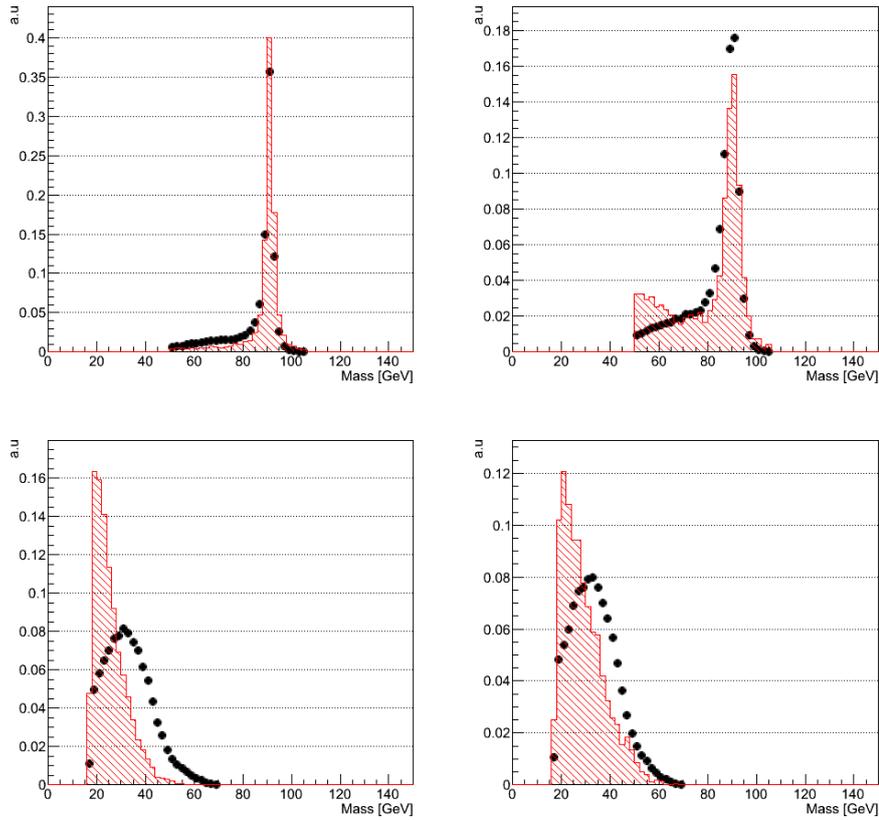


図 52: MVA の MC の ZZ^* (赤線) と signal (点) の input variable の mass 分布。左が truth、右が reconstruction。上から primary pair の mass と secondary pair の mass 分布となる。

4.4.3 MVA の結果

MVA を行うときは、training を行う際の input sample の統計量が必要となる。仮に統計量が足りないもので training を行った場合はその統計的変動に影響され、用いる sample によって同じ training の結果を適用しても違う結果を得る事もある。これを、overtraining と言う (図 53)。Overtraining は、decision tree の node を分ける回数を増やしたり (maxDepth)、boost の為に形成する tree の数 (NTree) を統計量に比べて上げすぎた場合に起こりやすい。本研究においては、Ntree は 50、maxDepth は 3 と低い値を使っている。

$H \rightarrow ZZ^* \rightarrow 4\ell$ はその特性上統計量が少ないため、本研究では 4ℓ の 4 つの channel ($4\mu, 2e2\mu, 4e, 2\mu 2e$) の解析結果を全部合わせて MVA の input として使った。Training を行うときは MC sample を分け、それぞれ training 用とその結果を適用する apply sample に分けて使っている。また、統計量の足りない ZZ^* 以外の BG は、overtrain を引き起こす原因となり得るので本研究では除外して training を行っている。

また統計量を増やす意味で 110~140 GeV の間にある event を用いて training を行っている。これは、あまり低い mass 領域まで入れてしまい $Z \rightarrow 4\ell$ の影響を受ける可能性を排除するためである。

上限の値は、secondary pair の mass window が変わるところから決めた上限である。一方 signal sample の場合、background と比べ特定の質量のところに peak を持つ特徴的な $M_{4\ell}$ の分布になる。本研究では、見ている領域全てにおいて signal が存在する可能性を置いて行っているため、また、統計量を増やすという意味で signal sample も、この領域内にある sample を全て使って training を行っている。

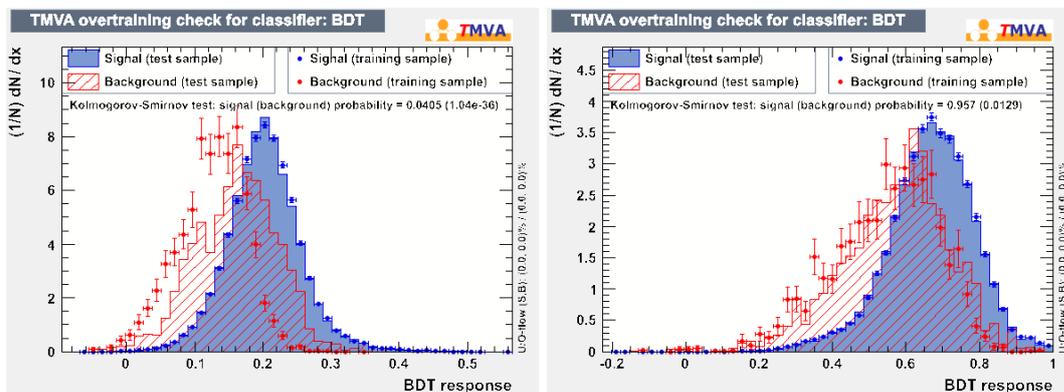


図 53: overtrain された sample の BDT response(左) と実際適用した training の結果。background の統計が足りなく、左の方は test sample (histo, 赤) と training (point, 赤) の結果が違う分布をしていることが分かる。

Training を行った後は、その結果を apply sample に適用して BDT に対する response (式 (20)) を計算する。そして、ある threshold の値以下の event を排除したとき、background と signal の significance を評価し、threshold をかける前後でどれほどの改善があるかを見た。これを基準に、最も significance が高くなったときの threshold で、最終的に background を排除した。

図 55 のように、BDT response の cut 前後で、 ZZ^* を含む background が 20%程度排除されていることが確認できる。これによる significance の変化は、約 2.8%改善される事が確認できた。Significance 及び event 数の変化を、表 16 にまとめる。

表 16: BDT cut を使う前後での変化

BDT response	before		after	
Significance(4ℓ)	2.84 ± 0.08		2.92 ± 0.09	
	overall events		at 122~128 GeV	
	before	after	before	after
Data	275	265	9	8
Signal	9.3 ± 0.1	8.6 ± 0.1	6.27 ± 0.09	5.90 ± 0.08
BG	234.8 ± 1.5	219.2 ± 1.4	3.17 ± 0.24	2.52 ± 0.2
ZZ^*	222.1 ± 1.2	214.9 ± 1.2	2.28 ± 0.12	1.71 ± 0.10
$Z + jet$	7.6 ± 0.9	7.1 ± 0.8	0.5 ± 0.19	0.47 ± 0.17
$Z + b\bar{b}$	3.7 ± 0.1	3.5 ± 0.1	0.32 ± 0.02	0.28 ± 0.02
$t\bar{t}$	1.3 ± 0.3	1.1 ± 0.3	0.06 ± 0.02	0.06 ± 0.02

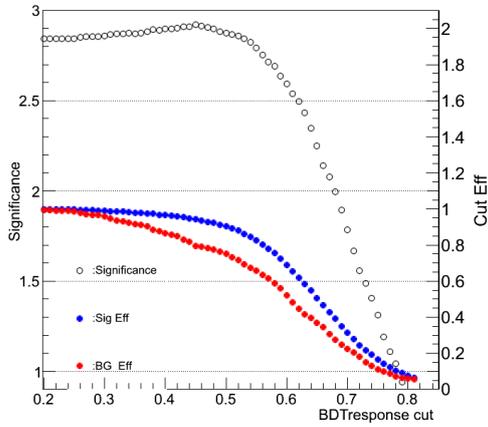


図 54: BDT response に対する significance の変化。青と赤は signal と BG の cut efficiency を示す。

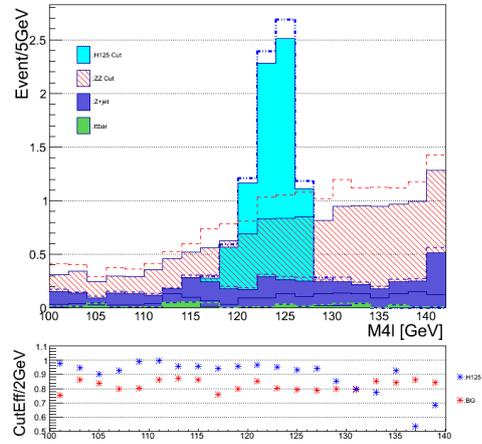


図 55: MC の $M_{4\ell}$ distribution。BDT response による cut の前(点線)と比べ、background は signal 周辺で $\sim 20\%$ 排除されている様子が見られる

4.5 Exclusion Limit と p-value

4.5.1 Profile Likelihood

新しい物理を発見したというのは、統計的手法を使って観測された event が background で説明できる確率がどの程度かで説明される。つまり、既存の知られている物理過程のみを仮定した場合 (background only) と、予想される新しい物理をも含んだ場合 (signal+background) を観測された data と比較することである。

このような比較のための量として p-value という、signal がないときを仮定したときと data から得られた結果がどれほど一致するかを判断する確率が用いられる。Data から得られた p-value がある threshold 以上であればその領域での新しい物理が存在すると判断する。

ここで p-value を判定するとき何を持って判断をするかが必要で、signal の強度 (μ) とその likelihood の比を持って検定を行う。例えば、ある物理量 x に対して測定を行ったとき期待される event 数は、signal を S 、background を B とすると、signal の強度 μ を入れた $\mu S+B$ が全 event 数と考えられる。ここで、 $\mu=0$ が background only、 $\mu=1$ が理論から予想される signal+background の場合となる。

これからある μ に対する確率関数 P を考えると、 P は Poisson 分布に従うことが期待され、また系統誤差や histogram の bin 数、channel 数などを考慮すると、以下のようなになる [15][16]。

$$P(n_b, a_p | \mu, \alpha_p) = \prod_{c=\text{channels}} \prod_{b=\text{bins}} \text{Pois}(n_{cb} | \nu_{cb}) \cdot \prod_{p=\text{Sys}} G(a_p | \alpha_p) \quad (10)$$

ここで、 n は event 数、 c は channel 数、 b は bin 数、 p は systematic の種類を指す。つまり、 n_{cb} は、 c channel の b bin の event 数、のようになる。また、 a_p は α_p (nuisance parameter) を制御する項で、 $G(a_p | \alpha_p)$ は測定された a_p の分布である。 ν は期待される event の数である。これが likelihood となり、 μ と nuisance parameter(θ) の関数である。以後、これを $L(\mu, \theta)$ と表す。

4.5.2 Test static

LHC を用いた実験では、p-value を求めるための検定量として、上記の profile likelihood の比を取った、

$$t(\mu) = -2 \ln \lambda(\mu) \quad (11)$$

を使っている。ここで $\lambda(\mu)$ が、

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} \quad (12)$$

のように表される profile likelihood ratio である。 $\hat{\theta}$ はある特定の μ に対しての nuisance parameter の best fit 結果を意味し (conditional)、 $\hat{\mu}, \hat{\theta}$ は μ と nuisance parameter 全部に自由度を与えたときの fit 結果 (unconditional) を意味する。

この定義から、 $0 \leq \lambda \leq 1$ である。 t_μ が大きくなると、それだけ data と仮定した μ は一致しなくなることを意味する。これから仮定と data (observed) の一致の具合を表す p-value は、

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu) dt_\mu \quad (13)$$

のように定義され、 $t_{\mu,obs}$ は data から評価された test statistic で、 $f(t_\mu|\mu)$ は signal の強度 μ の仮定の下での確率密度分布である。しかし、実際は $\mu \geq 0$ であるはずなので、もしも data の fit 結果が $\hat{\mu} < 0$ だった場合は $\hat{\mu} = 0$ のようにすることが要求される。従って、上記の λ は、

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\ \frac{L(0, \hat{\theta}(\mu))}{L(0, \hat{\theta}(0))} & \hat{\mu} < 0 \end{cases} \quad (14)$$

のように変形をして定義する。これを用いて t_μ を再定義すると、

$$\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu) = \begin{cases} -2\ln\frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\ -2\ln\frac{L(\mu, \hat{\theta}(\mu))}{L(0, \hat{\theta}(0))} & \hat{\mu} < 0 \end{cases} \quad (15)$$

のようになる。これらが、実際の計算で用いられる。

4.5.3 Discovery of signal

新しい物理の発見を調べるときは、前節で書いた p_μ を用いる。 $\mu \geq 0$ の元で $\mu = 0$ 、つまり background のみを仮定した場合が排除されることで、新しい signal の発見に繋がるわけである。このときの \tilde{t}_0 を特に q_0 と書こう。式 (16) から、 $\mu = 0$ のときは、

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (16)$$

で、 $\hat{\mu}$ が負の時は 0 となる。そして p_0 は、13 から

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|0) dq_0 \quad (17)$$

のようになる。

4.5.4 Confidence Level

実験の結果と理論を比べ、どのぐらいの信頼度で理論が排除されるかをみるとき、confidence level (CL) が用いられる。つまり、観測量 (observed) とある理論を仮定した時の確率密度の分布 (expected) と比較し、観測量がどれだけ有意に理論から離れているかを見る。一般に、その判断の基準は CL=5% とされる。

例えば、図 56 のように、event 数に対して signal があるとき (s+b) とないとき (b) の確率分布があったとする。このとき、ある測定量 (observed) を得たとすると、s+b に対して左側検定、b に対して右側検定を行い、それぞれの面積を CL_{s+b} 、 $1-CL_b$ と表す。この面積が 5% 以下のとき、その測定量はその理論から有意に離れているとし、理論は破棄される。実際は、確率密度の分布は式 (10) において、 $\mu = 0$ のときと、 $\mu = 1$ のときがこれに対応する。

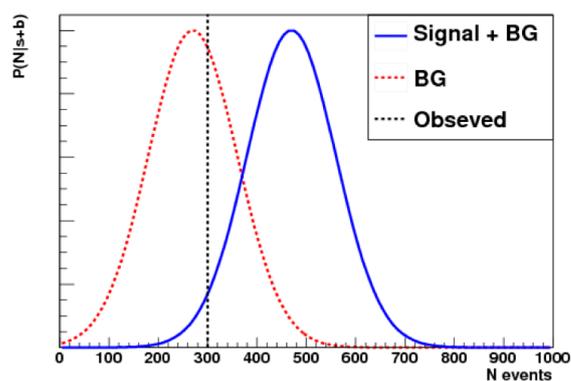


図 56: BG only と signal が存在すると仮定したときの確率分布の事例。実際観測された (observed) 量が仮定した model に対して有意であるかを判断する。

しかし、シグナルに対する感度が少ない測定結果においては、これらでは background の統計的ゆらぎだけでも理論の破棄が起こり得る。そのため、ATLAS では

$$CL_s = \frac{CL_{s+b}}{CL_b} \quad (18)$$

と定義されるものを使う事を推奨している。 $CL_s < 5\%$ を破棄域とすれば、上記のような問題は起こらない。

5 結果と考察

5.1 解析結果

まず、Low mass 領域において MVA による変化を見てみると、図 57 のようになった。

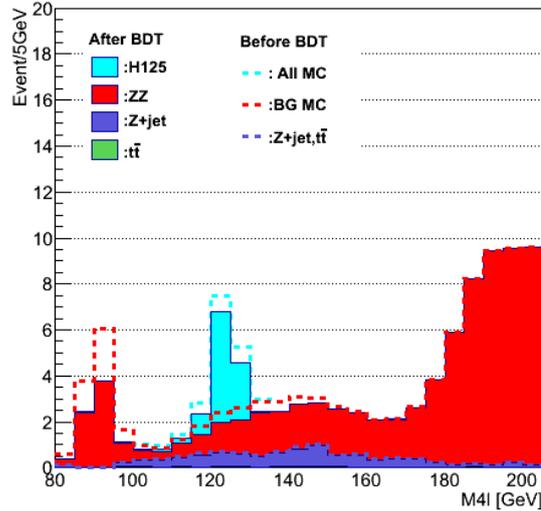


図 57: Cut base (Filled Histogram) と MVA (Dashed Histogram) の MC での $M_{4\ell}$ 分布。Signal 領域での変化は、 ZZ^* 以外の background にはあまり減少が見られない。

Low mass 領域における ZZ^* background の排除はされているが、その多くが Z mass (約 91 GeV) 付近であり、signal の期待される領域での background の減少量は signal の約 2 倍 (4 章、表 12) 程度であった。

まず MC の結果から計算された、local p-value から算出した expected sensitivity の変化を signal 周りで確認してみた。その値を表 17 に載せる。

表 17: BDT 適用前後の signal 予想領域付近で p-value から算出された expected sensitivity の変化

Mass [GeV]	120	123	124	125	126	127	130
Cut	1.82	2.33	2.52	2.68	2.88	3.08	3.65
BDT	1.84	2.33	2.53	2.73	2.91	3.10	3.69

値は若干であるが全体的に改善する傾向にあり、125 GeV 付近で最も大きい改善が見られ、その値は約 1.7% であった。一方 data の変化を見てみると図 58 のようになり、signal 付近 (122~128 GeV) での data が 9 つから 8 つに一つ減ってしまったような結果となった。これから得られた exclusion limit や p_0 の分布は図 59 の通りで、MVA を使う前後での大きい変化は得られなかった。詳しい値を表 18、19、20、21 に載せる。

この結果から見ると、Signal と ZZ^* だけによる BDT training は有効ではあるがそれほど大ききを得る事はできなかった。

表 18: Cut base の生成断面積の比の強度

Mass[GeV]	120	123	124	125	126	127	130	135	140	145	150
exp:	1.11	0.79	0.71	0.65	0.59	0.54	0.42	0.30	0.24	0.21	0.21
obs:	3.49	2.82	2.47	2.14	1.83	1.56	0.87	0.42	0.22	0.28	0.33

表 19: BDT response を適用したときの生成断面積の比の強度

Mass[GeV]	120	123	124	125	126	127	130	135	140	145	150
exp:	1.09	0.79	0.70	0.63	0.58	0.52	0.40	0.29	0.23	0.21	0.21
obs:	3.00	2.57	2.27	2.01	1.84	1.61	0.99	0.49	0.25	0.31	0.34

表 20: Cut base 方法と MVA 方法のそれぞれに対して、質量 125 GeV 付近において観測されたデータがどの程度背景事象の揺らぎで起こりえるのかを表した確率 (p_0 value)

Mass[GeV]	120	123	124	125	126	127	130
Cut	7.66×10^{-4}	1.69×10^{-5}	2.14×10^{-5}	6.46×10^{-5}	4.94×10^{-4}	1.63×10^{-3}	2.83×10^{-2}
BDT	4.74×10^{-3}	1.15×10^{-4}	8.6×10^{-5}	1.35×10^{-4}	2.89×10^{-4}	7.80×10^{-4}	7.42×10^{-3}

表 21: Cut base 方法と MVA 方法のそれぞれに対して、 p_0 value から算出した sensitivity

Mass[GeV]	120	123	124	125	126	127	130
Cut	3.17	4.15	4.09	3.83	3.29	2.94	1.91
BDT	2.59	3.68	3.76	3.64	3.44	3.16	2.44

最終的に MVA を使った 実 data の解析結果から Higgs 粒子質量 124 GeV 付近で最大のイベント超過が観測され、sensitivity は約 3.8σ であった。Higgs 粒子の発見までは至らなかったが 3σ 以上の兆候を観測する事ができた。また、Low Mass 領域の 130~165 GeV の領域において、標準理論 Higgs 粒子の存在を 95% confidence level で棄却した。

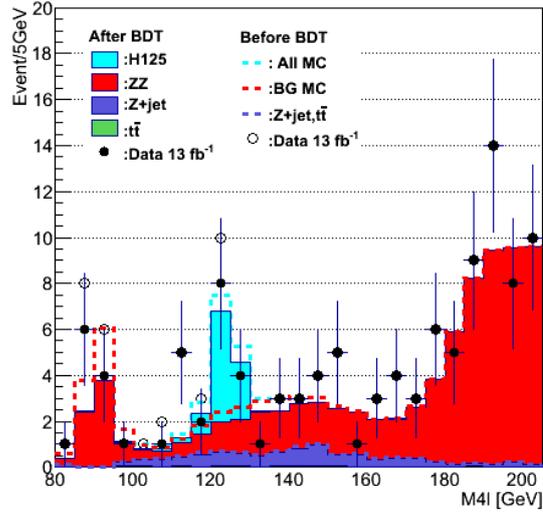


図 58: Cut base (Filled Histogram) と MVA (Dashed Histogram) の MC と data の $M_{4\ell}$ 分布。Signal 領域において、data 点が減少していることが見える。

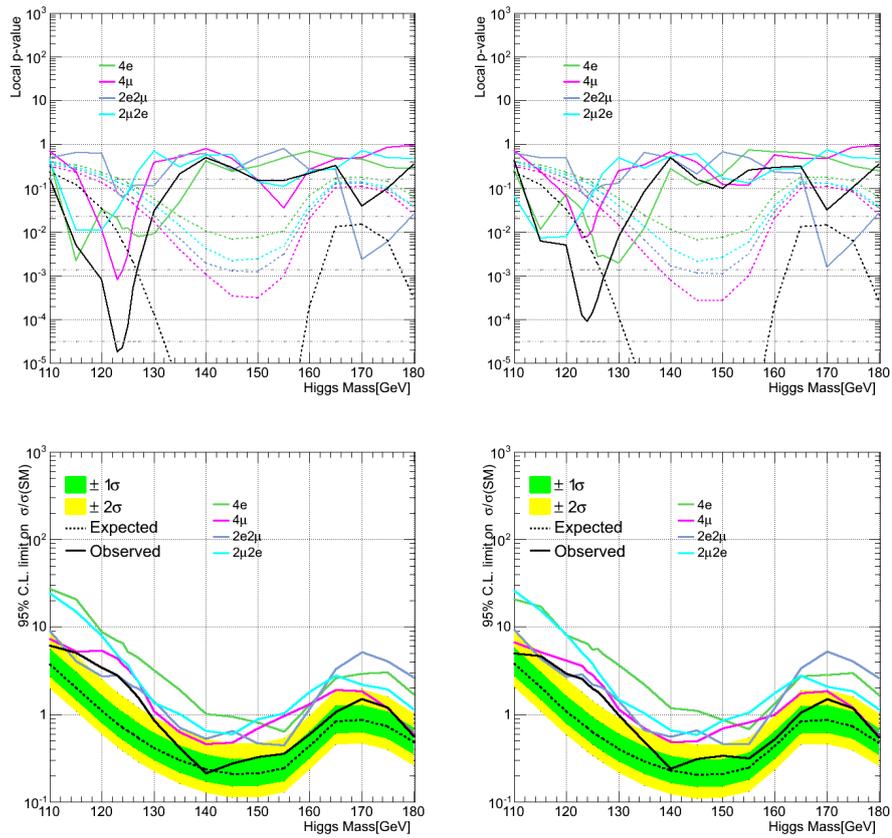


図 59: BDT apply 前後での p-value 及び CLs の比較。左が cut base, 右が BDT を apply した場合。

5.2 考察

図 59 の local p-value の分布を見ると、123~125 GeV の間での observed が大きくなったことが見える。これは、background の減少がそれほど大きくなかったことに比べ、統計量が少ない data 点の一つ減ってしまった事から data の excess がその分小さく評価されてしまった結果と思われる。

この結果から見ると、Signal と ZZ^* だけによる BDT training は有効ではあるがそれほど大きい変化を得る事は難しいと思われる。これは、MVA への input variable の分布からも見られるように、分布そのものに差はあっても部分的な差であって明確な差異が見られる訳ではないからだと思う。

また、 ZZ^* による寄与には減少が見られるものの、4 つの channel を合わせて ZZ^* と同程度存在するはずの Z +jet や $t\bar{t}$ の event は ZZ^* の減少率の 1/3 程度の変化しか見られない。これから、 Z candidate の mass と角度の分布だけによる training はこれら background に対してそれほど効果的ではないと考えられる。また、training は ZZ^* だけで行ったとは言え、 Z +jet などの他の background の変化が少ないことから、これらもまた、角度分布だけで排除力を上げるのは難しいと思われる。

本研究では、cut base を全部通った後の event に対して MVA を使ったため、統計量の少ないこれらの排除可能な background は training に用いなかった。これらも含めた解析をする場合、cut base の解析で control region を作っていたように cut を緩くして統計量を増やして MVA から background を分離する方法が考えられる。Cut base では、通常各物理量ごとで background と signal の違いを利用しているが、MVA を使えば input variable として使っている全ての物理量に対して相関を調べることが可能で、cut base よりも効果的な background 排除の可能性が期待できる。

しかし、本研究の方法を維持しながらこの方法を使うには、 $Z+X_e X_e$ や $Z+X_\mu X_\mu$ がそれぞれ違う control region でこれらの background を評価しているため、これらとの関係を考慮して training を行う必要がある。本研究ではこの方法を直接試してはいないが、この方法を使う際に注意すべき事項と思われることについて考察を試みた。

- 統計量と control region

4ℓ channel で MVA を行う際の最大の問題は統計量の問題である。そもそも background が少ないため、最終的に残る background の event が少なすぎる場合、training を行えないからである。また、その統計量を解決するために使っている control region を考慮する必要がある。

例えば、統計量を増やすために単純に Z +jet や $Z+b\bar{b}$ 、 $t\bar{t}$ の排除力の高い、isolation や IP significance を外してこれらも input variable として使うとする。これらの control region は subleading pair のみ isolation と IP significance cut を緩くしているので、training を行う際には統計量を増やすために control region より loose な選別を行った後、control region を反映した training を行うか、control region の cut を通って background の評価が終わった段階で training に渡す必要がある。

しかし、後者の場合は background を増幅させているとは言え、それでもこの段階で今回用いた Z +jet MC の event の統計量が少ない方なので、さらに sample を training と apply 用に分けてしまうと overtraining の要素となる。図 60 に、 4ℓ 全部の isolation と IP significance

の cut を除外したとき、排除可能な background すべてを用いた training を行ったときの結果を示す。

- Input variable 間の相関と control region

Control region より緩い cut を使って統計量を増やした際には、input variable の各変数ごとの相関を全て考慮に入れた上で control region を考慮する必要もある。もし input variable の中で強い相関を持っている変数同士が存在し、それらが control region を形成する際に使われているのであれば、単純に cut 全体を緩くして training を行った結果を control region の event に apply する訳には行かない。強い相関を持つ変数同士が、cut を通った control region では関係性が崩れる可能性があるからである。

例として $Z+b\bar{b}$ を考えてみると、 b によって isolation と IP significance が同時に悪くなる (4章、図 37、38)。従ってこれらの分布の間には何らかの相関があると考えられる (図 61)。

- Input variable と Background Process

各 background の process はそれぞれ特有の特徴を持つ。従って、特徴が大きく違う process 同士を同じ枠組みで MVA の method を使うことは、その特徴の強いもの変数に対する評価を弱める可能性がある。例えば、IP significance の分布を electron と muon で分けてみると (4章、図 38)、Signal と ZZ^* process の electron が muon と比べ 2 倍程度大きい値を持つ事が分かる。従って、これらを一つにして MVA に渡してしまった場合、signal からの muon と Z +jet などの background から来る muon の分離能力が弱まる可能性が出来る。

これらの問題を考慮すると、全ての process を用いた MVA を行うにはそれぞれの process に対して MVA を行い、それらの MVA の response を combine して最も signal と background をうまく分けるような場合を探す方法を取る必要があると思われる。

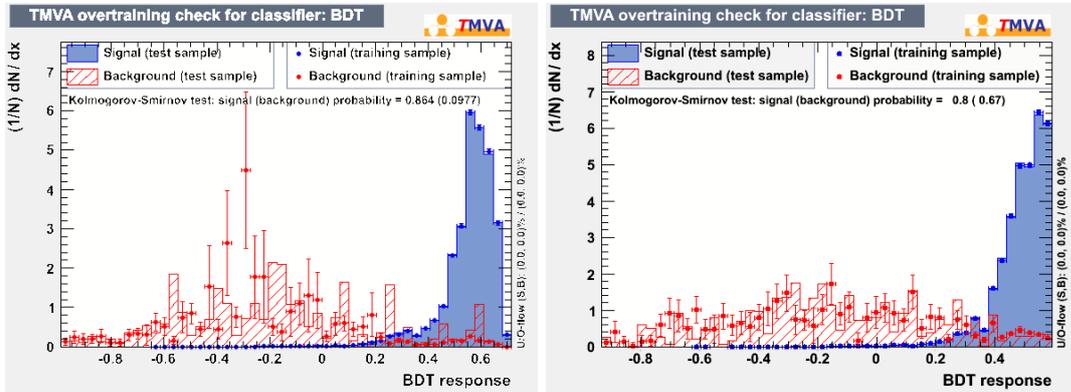


図 60: 同じ条件で Z +jet を含めて training を行った場合 (左) と、 Z +jet だけを外して training を行った場合 (右)。 Z +jet が overtraining の原因となっている。

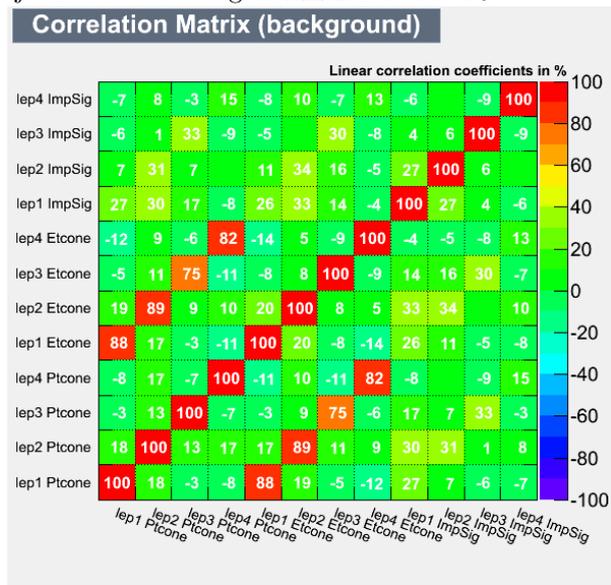


図 61: Background の isolation と IP significance を要求しなかった場合の各 lepton の isolation と IP significance の関係。lep1,3 が primary, lep2,4 が secondary pair の lepton となる。Isolation と IP significance の間に、相関関係があることが見える。

6 まとめ

本研究では、LHC 実験の ATLAS 検出器で、2012 年に重心系エネルギー $\sqrt{s} = 8\text{TeV}$ で取得された積分 Luminosity 13 fb^{-1} の data を用いて、 $H \rightarrow ZZ^* \rightarrow 4\ell$ channel における多変数解析手法による Higgs 粒子の探索を行った。

この channel の特徴は、4 つの lepton を要求するため、非常に稀な事象を見るためその background が少なく、signal に対して感度が高いことである。しかし、この channel の主な background である $ZZ^* \rightarrow 4\ell$ は終状態が signal と同じであるため、従来の cut base の解析手法では、 $M_{4\ell}$ 分布以外では排除が困難である。

従って本研究では、signal と ZZ^* の process で得られる lepton や Z boson の角度分布の違いを利用し、多変数解析手法を使うことでこの background を排除し、signal の発見感度の改善を行った。

その結果、signal が予想される 125 GeV 付近での期待される発見感度に約 1.7% の改善があったが極めて大きい効果は得られなかった。これは、signal と background の角度分布にその差があったものの、 Z +jet や $t\bar{t}$ と signal の isolation や impact parameter 程の明確な特徴の差ではないため、排除が難しいからだと思われる。

最終的に MVA を使った 実 data の解析結果から $M_H = 124\text{ GeV}$ 付近で最大のイベント超過が観測され、significance では約 3.8 であった。Higgs 粒子の発見までは至らなかったが 3 以上の兆候を観測する事ができた。また、Low Mass 領域の 130 ~ 165 GeV の領域において、標準理論 Higgs 粒子の存在を 95% confidence level で棄却した。

A BDT method

BDT 方法は、解析対象となる変数たちの中、signal と background が混在する最初の状態 (root node という) を最もうまく分けるある変数を探して、signal-like な event と background-like な event に分ける (新しい node を作る)。

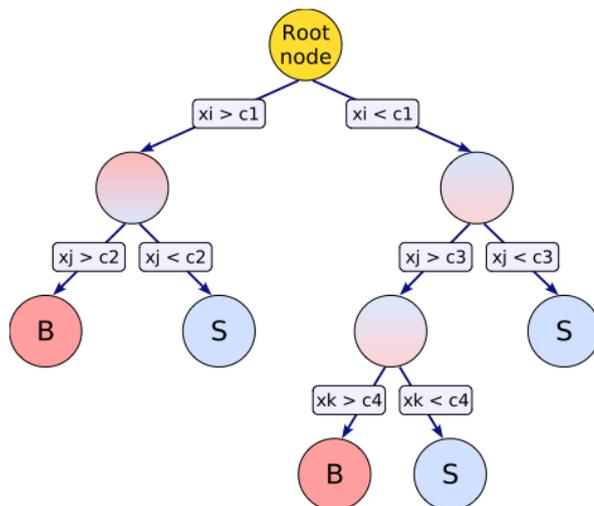


図 62: Decision Tree の概念図。このように、各 node において signal と background をもっともうまく分ける変数を、与えられた変数群 x の中で探して event を分けて行く。最終的に得られた leaf node の分類は、中の event に signal と background でどちらの方が多いかで決められる。

このような方法を次々と分けられた結果に適用し、この作業を停止させるある条件 (分けられた node の event 数、場合分けの回数など) が満たされるまで続け (図 A)、一つの tree を形成する。各 node で使われる変数は他の node で使われたことがあるものでも考慮されるので、場合によっては与えられた変数の中で全く用いられなかった変数もあり得る。そして、最後に得られた node を leaf node と言い、この手法を decision tree という。

この手法はその基本的考え方が簡単な故によく使われる method ではあるが、統計量が少ないときや複数の変数が複雑な関係にある場合に signal と background をうまく分けられないという短所がある。この短所を克服するために Boosting という手法が用いられる。ここでは、本解析で用いた最も一般的に使われる AdaBoost というものの考え方について説明する。[9][10]

AdaBoost は Decision tree の方法で多くの tree を作ることで、BDT の弱点を補完する手段である。まずは一つの tree を作り、最終的に分けられた event の中で missclassify された rate (err) を算出する。この rate を用いて、

$$\alpha = \frac{1 - err}{err} \quad (19)$$

で表される weight を計算し、missclassified event にかける。ただし、weight は初めに与えられた weight の総和が保たれるように normalize される。このように weight を掛け直した後、次の tree の

作成に入る。

ある変数群 \mathbf{x} に対して評価された tree は、その結果が Signal ($h(\mathbf{x})=1$) と background ($h(\mathbf{x}) = -1$) のどちらに近いが相対的に表され、 $N_{collection}$ 個の tree がある場合、最終的な boosted event classification ($y_{Boost}(\mathbf{x})$) は、

$$y_{Boost}(\mathbf{x}) = \frac{1}{N_{collection}} \cdot \sum_i^{N_{collection}} \ln(\alpha_i) \cdot h_i(\mathbf{x}) \quad (20)$$

のように表される。 $y_{Boost}(\mathbf{x})$ が小さいほど background like で、高いほど signal like な event となる。

B MVA results

MVA を用いた結果、5 章で述べたように signal 周辺 (122~128 GeV) での実 data の結果にも 1event の変化があった。各 channel に対して分けてみると、この event は 4μ channel から得られた event であった。

この章では、この減少した event の MVA で用いた各 input variable にの分布を見ることで、どのような event が排除されたかを見ている。(図 63、64)

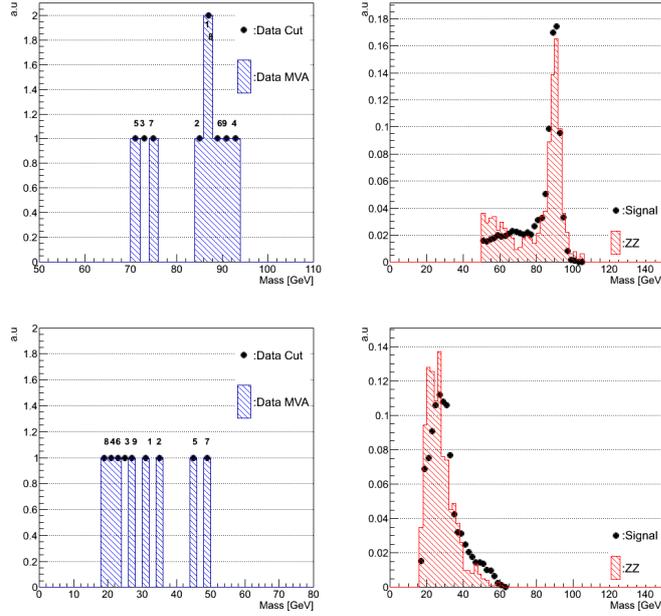


図 63: MVA 適用の結果削除された 122 GeV~128 GeV の間での 4μ の実 data の MVA を使う前 (黒) と後 (青) での変化 (左) と、120 GeV~130 GeV の間での ZZ background (赤) と signal (黒) の分布 (右)。上から on-shell Z、off-shell Z の Mass 分布。各 event に与えられている番号は、表 22 の各 event に対応する。

これらの分布を見ると、主に $\cos(\theta_1)$ 、 Φ が background と signal に分布の差が見える。また、Mass の分布も background の方がより多くの low mass の Z の event を持っていることが見える。そして、今回排除された 4μ の Z の mass 分布とこれらと比較してみると、いずれも background が集中する傾向にある領域に存在することがわかる。実際、この event と他の data 点の MVA の結果を、MC の各 MVA response に対する全 event に対する signal の比、

$$P_{Sig} = \frac{N_{Sig}}{N_{All}} \quad (21)$$

から評価された signal らしさ (P_{Sig}) で評価してみると、図 65 のようになった。また、各 signal 候補の情報を表 22 に載せる。

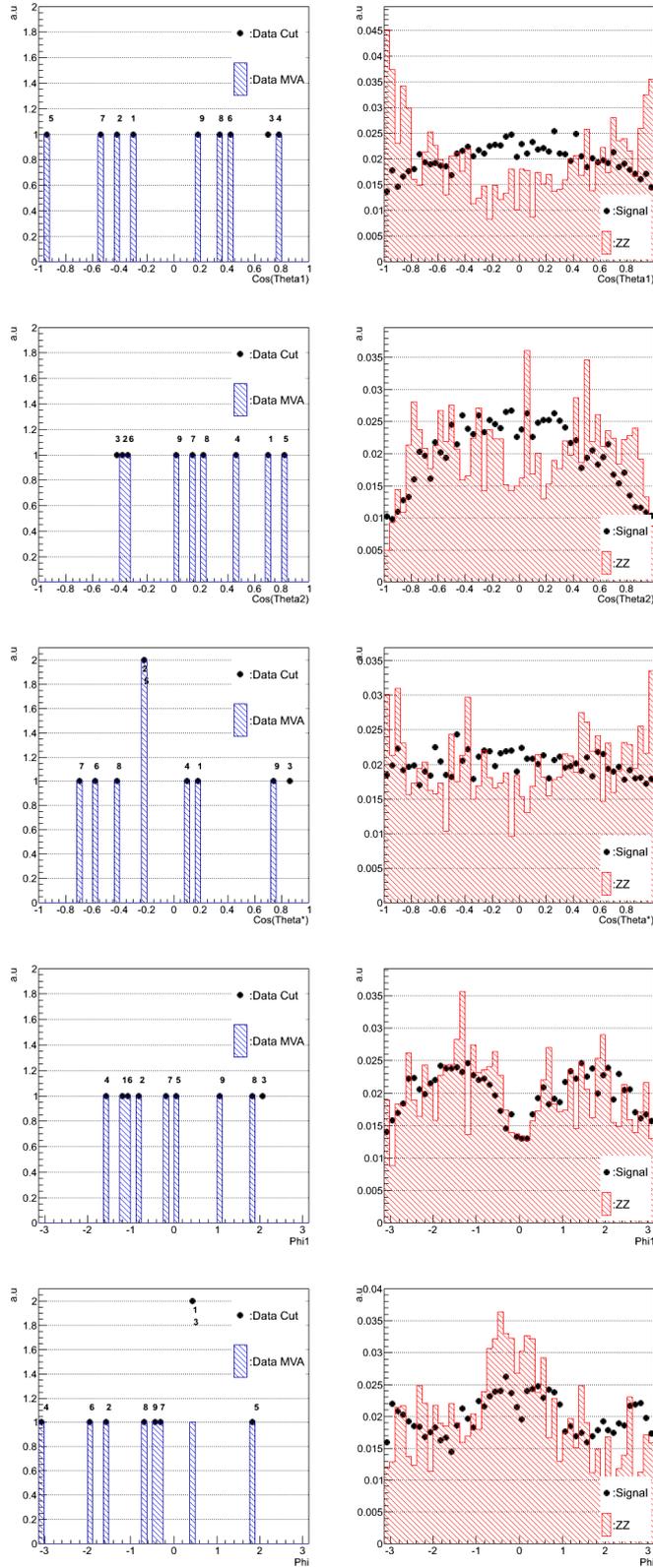


図 64: MVA 適用の結果削除された 122 GeV~128 GeV の間での 4μ の実 data の MVA を使う前 (黒) と後 (青) での変化 (左) と、120 GeV~130 GeV の間での ZZ background (赤) と signal (黒) の分布 (右)。上から $\cos(\theta_1)$ 、 $\cos(\theta_2)$ 、 $\cos(\theta^*)$ 、 Φ_1 、 Φ 。signal sample は 125 GeV のもののみを使っている。各 event に与えられている番号は、表 22 の各 event に対応する。

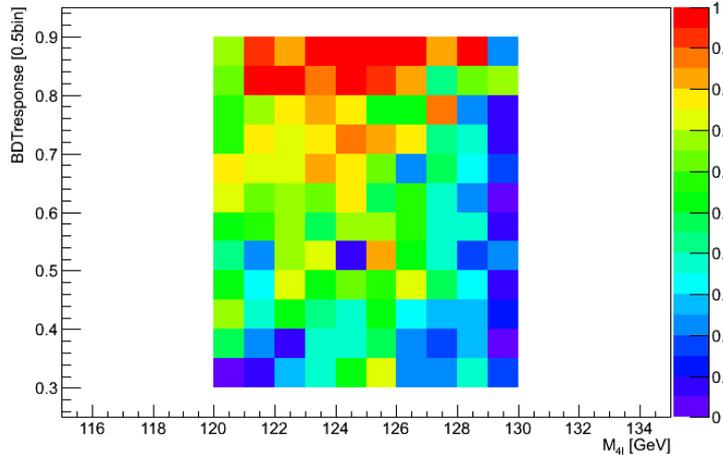


図 65: MC から評価された、各 MVA response に対する P_{Sig} の計算値、125GeV の sample のみが使われている。

Num	channel	M_{4l} [GeV]	MVA res	P_{Sig}
1	4μ	124.1	0.66	0.75
2	4μ	123.3	0.79	0.83
3	$4\mu^*$	123.5*	0.30*	0.36*
4	4μ	122.9	0.56	0.67
5	$4e$	124.5	0.76	0.77
6	$4e$	125.5	0.59	0.66
7	$4e$	123.6	0.66	0.84
8	$2e2\mu$	122.7	0.56	0.67
9	$2e2\mu$	123.0	0.69	0.84

表 22: Cut base の 122~128 GeV 付近にある各 event の情報。* がついている data が MVA によって排除された event を示す。 P_{Sig} の計算には、125GeV の sample のみが使われている。

これから見ると、今回の結果で排除された data は、他の data 点と比べて P_{Sig} がかなり小さく評価されていることが分かる。

謝辞

本研究を行う際には、色々な方々からのご指導やご助言を頂きました。まずは、研究室配属以来、研究や進路に関して様々の助言や指導をしてくださった指導教員の田中純一准教授に心よりの感謝を申し上げます。

また、解析全般やMVA手法などに関して助けを頂いた浅井祥仁准教授、江成祐二助教、中村浩二特任研究員にも真に感謝致します。その他、CERN(欧州原子核研究機構)滞在の時には、CERNに常駐していらっしゃる金谷奈央子助教、寺師弘二助教や、山本真平特任助教、山中隆志特任研究員にも meeting の時に色々なご助言を頂くなど、とてもお世話になりました。

2年間の生活で色々とお世話になった東京の教員の方々や先輩、同期の皆にも感謝致します。

まずは、1年目の頃ゼミや実験でお世話になった、難波俊雄助教と石田助教にお礼を申し上げます。続いて、CERNに常駐している博士課程の佐々木雄一さん、吉原圭亮さん、山口洋平さん、風間慎吾さん、そして東京(福井)の宮崎彬さんは研究面だけではなく、日々の生活を送る際に様々なお話をしてくださるなど、研究者を目指す者としての見本を示してくださいました。

同じく同期の森永真央君や黒崎龍平君、稲田聡明君、大和田健太君、小林愛音さん、有馬匡彦君にも2年間共に研究室で過ごす上で世話になりました。特に森永真央君や黒崎龍平君は解析を行う上でも色々な方面で助けを頂きました。皆、私が自分を見直す切っ掛けを与えてくれた良い仲間達であったと思います。

また、秘書の塩田雅子さん、手塚淑恵さん、竹本葉子さん、宮園あき子さんにも、2年の間様々な事務の手続きなど、研究遂行の際の身の回りの事を支援して頂いたことについて深く感謝致します。

このように多くの方々から支えられながら研究に挑めるのは素粒子センターだからこそ可能なものだと思います。そのような環境で過ごしたこの2年間は、私にとって大変貴重な、今までの中で最も重みのある2年間になったと感じております。最後は、私の7年間に渡る留学生生活を心的、経済的な面で惜しみなき支援をくださったご両親へ、この場を借りて感謝の意を申し上げます。

崔 原碩

参考文献

- [1] K.Assamagan, J.Boudreau, A.Buckley et al.,*The ATLAS Simulation Project*, <https://cds.cern.ch/record/1152900/files/ATL-COM-SOFT-2008-024.pdf>
- [2] Stefano Frixione, Paolo Nason, Carlo Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, arXiv:0709.2092 [hep-ph]
- [3] M. Aharrouche, J. Colas, L. Di Ciaccio et al.,*Energy Linearity and Resolution of the ATLAS Electromagnetic Barrel Calorimeter in an Electron Test-Beam*, <http://arxiv.org/pdf/physics/0608012v1>
- [4] The ATLAS Collaboration,*ATLAS CALORIMETER PERFORMANCE*,CERN/LHCC/96-40 ATLAS TDR 1
- [5] ATLAS Muon Collaboration,*ATLAS Muon Spectrometer Technical Design Report*,CERN/LHCC/97-22
- [6] The ATLAS Collaboration,*Calibration and Performance of the Electromagnetic Calorimeter*,ATL-COM-PHYS-2009-169
- [7] The ATLAS collaboration,*Observation of inclusive electrons in the ATLAS experiment at $\sqrt{s} = 7\text{TeV}$* , ATLAS-CONF-2010-073
- [8] The ATLAS collaboration,*Trigger Menu Conventions*, <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TriggerMenuConvention>
- [9] A. Hoecker, P.Speckmayer, J.Stelzer, J.Therhaag, E. vos Toerne, H. Voss et al.,*TMVA4 Toolkit for Multivariate Data Analysis with ROOT Users Guide*,tech. rep., CERN,. arXiv:physics/0703039
- [10] N.Wernes et al.,*Advanced Boosting in TMVA4*,<http://tmva.sourceforge.net/talks.shtml>
- [11] Y.e.a.Gao,*Spin determination of single-produced resonances at hadron colliders*, Phys.Rev.D81(2010)075022
- [12] Kyle Cranmer,*HistFactory User Guide(ROOT 5.32)*,CERN-OPEN-2012-016
- [13] Particle Data Group,<http://pdg.lbl.gov/>
- [14] 高エネルギー加速器研究機構 (KEK) page,<http://www.kek.jp/>
- [15] G.Cowan,K.Cranmer, E.Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur.Phys.J.C71(2011) 1554, arXiv:1007.1727
- [16] A.L.Read, *Presentation of search results: The CL(s) technique*, J.Phys.G28(2002) 2693-2704.