

重心エネルギー 500 GeV ILC における $h \rightarrow \mu^+ \mu^-$ モードの解析 — part 4

川田真一

概要

¹ 本ノートは重心エネルギー 500 GeV で $h \rightarrow \mu^+ \mu^-$ モードの解析を行った際の記録であり、文献 [1-3] の続きである。

¹リリースノート

- 2017 年 1 月 18 日 執筆
- 2017 年 2 月 14 日 バグを見つけたので数字と図を取り換えた

1 序論

前回の解析 [3] では TMVA を使用し、さらなる解析精度の向上を目指した。結果として、精度の数字そのものは今までで最高の値を示したが、明確にオーバーレーニングしていた。これでは信憑性が心許ないので、さらに信憑性の高い解析を行うことを目指す。また、現在取り扱っているシグナルプロセスは $\nu\bar{\nu}h \rightarrow \nu\bar{\nu}\mu^+\mu^-$ であり、既に isolate しているが、将来の解析の発展を考えて IsolatedLeptonTagger を解析に組み込むこととした。

2 IsolatedLeptonTagger

2.1 概要

このプロセッサの説明は文献 [4] に書いてある。元々は Higgs self-coupling study の下で開発されたものであり、それを今回導入する。なおこのプロセッサはレプトンを識別するので、ミューオンだけでなく電子も識別してしまうことには注意が必要である。

2.2 事象再構成

IsolatedLeptonTagger ではまずシンプルなカットを用いてレプトン選別を行い、その後、既に別途行われたトレーニングから生成された weight file を用いて isolation を確認する。カットに関しては、以前までの解析でも行ってきたので理解しやすいが、用いるパラメータが異なっている。参考までに、以前の解析でカットに用いたパラメータは以下の通りである。

- $E_{\text{ECAL}}/(E_{\text{ECAL}} + E_{\text{HCAL}})$
- $(E_{\text{ECAL}} + E_{\text{HCAL}})/|p|$
- E_{yoke}
- P_t

最初 2 つのパラメータでミューオンを選別し、残り 2 つのパラメータで overlay background を除去していた。

対して、IsolatedLeptonTagger で用いるパラメータは以下の通りである。

- $E_{\text{ECAL}}/(E_{\text{ECAL}} + E_{\text{HCAL}})$ (電子のみ)
- $(E_{\text{ECAL}} + E_{\text{HCAL}})/|p|$
- E_{yoke}
- $|p|$
- $|d_0/\sigma(d_0)|$
- $|z_0/\sigma(z_0)|$
- MVA cut

最初の3つのパラメータは以前の解析で使っていたものと同じものが使用されている。ただし、こちらにおけるECALとHCALはILD全体におけるカロリメータ全てを含み、以前の解析では主要なECALとHCALしか用いていないことには注意が必要である。4つ目のパラメータは P_t ではなく運動量の大きさそのものを用いている。これは前回と異なる点なので、改めて調べた。5つ目・6つ目は primary vertex を要求するカットであり、タウ由来のミューオンを選ばないようにするためである。最後は isolation のための MVA cut である。以下の図はそれぞれのパラメータの分布である。

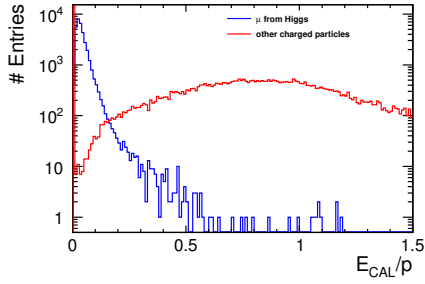


図 1: $(E_{\text{ECAL}} + E_{\text{HCAL}})/|p|$ 分布。

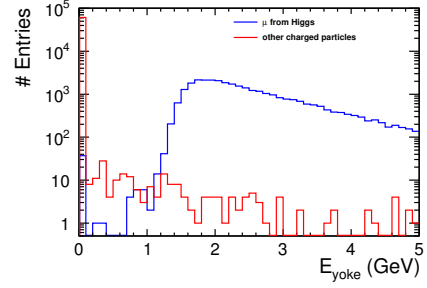


図 2: E_{yoke} 分布。

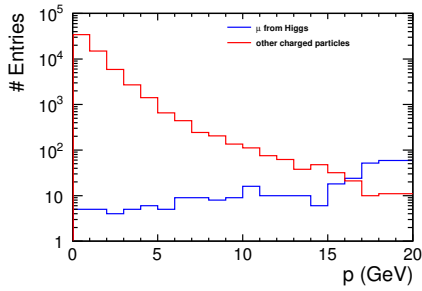


図 3: $|p|$ 分布。

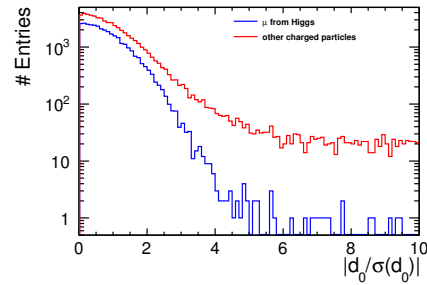


図 4: $|d_0/\sigma(d_0)|$ 分布。

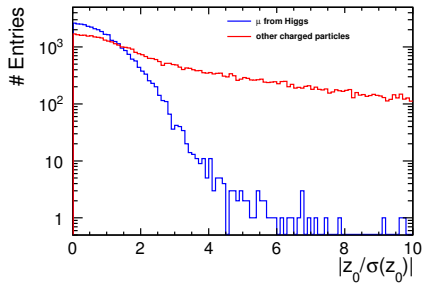


図 5: $|z_0/\sigma(z_0)|$ 分布。

これらから、カットを以下のように定めた。

- $(E_{\text{ECAL}} + E_{\text{HCAL}})/|p| < 0.5$
- $E_{\text{yoke}} > 0.5 \text{ GeV}$

- $|p| > 5 \text{ GeV}$
- $|d_0/\sigma(d_0)| < 5$
- $|z_0/\sigma(z_0)| < 5$
- MVA cut > 0.7

MVA cut はプロセッサのデフォルト設定値を用いた。この状態でシグナルにおける事象再構成の efficiency (正しくミューオンを正負 1 本ずつ組めた) は 94.9%、purity は 100%であった。

なお MVA cut は電子とミューオンのそれぞれで設定できるため、電子の方は 2 以上とし、IsolatedLeptonTagger ではミューオンのみを識別できるようにし、電子は識別されないようにした。

2.3 解析

Precuts として、以下のカットを適用した。

- μ^+ ちょうど 1 本、 μ^- ちょうど 1 本
- $\chi^2/Ndf(\mu^\pm) < 1.5$
- innermost hit (μ^\pm) $< 20 \text{ mm}$
- $\sigma(M_{\mu\mu}) < 1 \text{ GeV}$
- $N_{P_t > 5\text{GeV}} \leq 1$
- $125 < E_{\text{vis}} < 400 \text{ GeV}$
- $100 < M_{\mu\mu} < 130 \text{ GeV}$
- $\cos \theta_{\mu\mu} < 0.55$
- $|\cos \theta_{\text{miss}}| < 0.999$
- $P_t > 5 \text{ GeV}$

2・3・4 番目のカットは精度よく測定されたミューオンのみを使用するためのもの、5・6・7 番目はシグナルのようなイベントのみを使用するためのもの、8・9・10 番目は主に 2f バックグラウンドを除去するためのものである。以下はそれぞれの分布である。

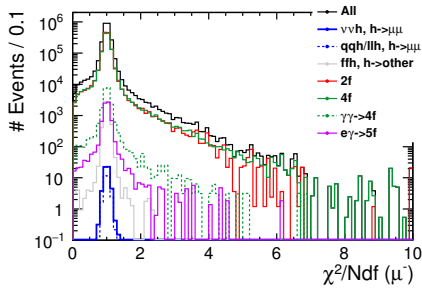


図 6: $\chi^2/Ndf(\mu^-)$ 分布。

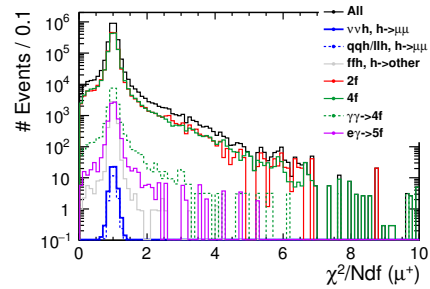


図 7: $\chi^2/Ndf(\mu^+)$ 分布。

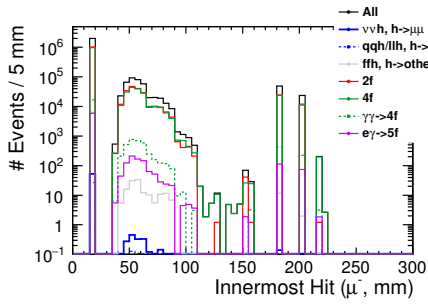


图 8: μ^- innermost hit 分布。

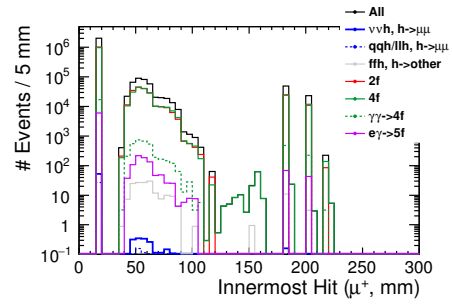


图 9: μ^+ innermost hit 分布。

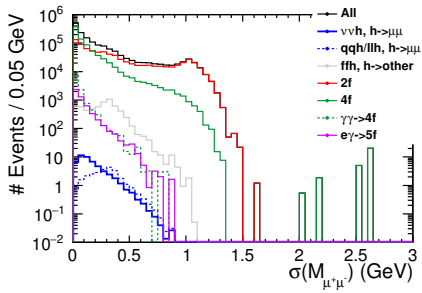


图 10: $\sigma(M_{\mu^+\mu^-})$ 分布。

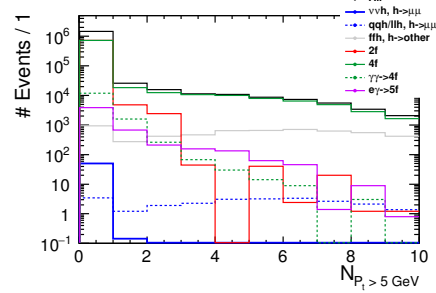


图 11: $N_{P_t > 5\text{GeV}}$ 分布。

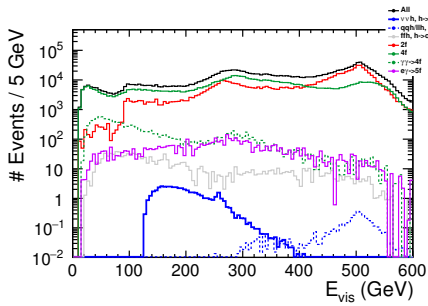


图 12: E_{vis} 分布。

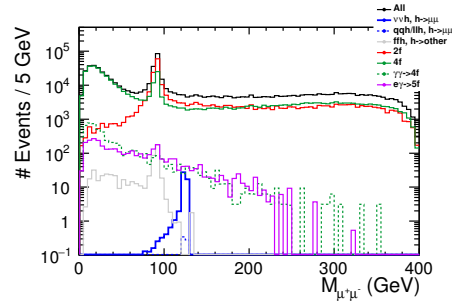


图 13: $M_{\mu^+\mu^-}$ 分布。

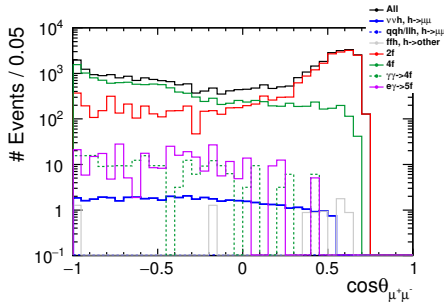


图 14: $\cos \theta_{\mu^+\mu^-}$ 分布。

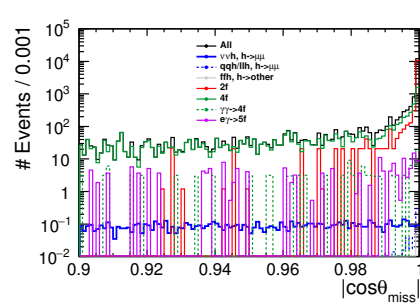


图 15: $|\cos \theta_{\text{miss}}|$ 分布。

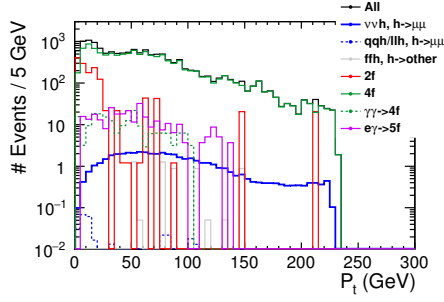


図 16: P_t 分布。

表は precuts 時点でのカットテーブルである。

表 1: Precuts 時のカットテーブル。

	$\nu\nu h$	$qqh+\ell\ell h$	ffh	2f	4f	$\gamma\gamma \rightarrow 4f$	5f
	$h \rightarrow \mu\mu$	$h \rightarrow \mu\mu$	$h \rightarrow \text{other}$				
No cut	57.53	31.13	4.116×10^5	4.224×10^7	4.592×10^7	3.356×10^5	2.209×10^5
# μ^\pm	54.82	27.72	6553.83	1.314×10^6	1.262×10^6	2.216×10^4	7206.44
χ^2/Ndf	54.51	27.59	6525.51	1.261×10^6	1.208×10^6	2.107×10^4	6978.30
Innermost	50.26	26.18	6194.38	8.042×10^5	8.047×10^5	1.393×10^4	5185.97
$\sigma(M_{\mu\mu})$	50.25	26.18	6192.51	7.287×10^5	8.025×10^5	1.393×10^4	5185.97
$N_{P_t > 5\text{GeV}}$	50.22	4.64	1208.32	7.261×10^5	7.434×10^5	1.354×10^4	4567.99
E_{vis}	50.16	0.72	551.95	2.549×10^5	4.245×10^5	5379.29	3343.95
$M_{\mu\mu}$	48.65	0.72	8.06	2.076×10^4	1.364×10^4	217.37	304.69
$\cos\theta_{\mu\mu}$	48.64	0.72	5.40	1.197×10^4	1.328×10^4	217.37	304.69
$\cos\theta_{\text{miss}}$	48.55	0.42	5.35	1417.20	1.171×10^4	198.64	295.95
P_t	48.45	0.39	5.35	1009.00	1.154×10^4	183.07	295.95

2f の内訳は次の通りである。

- $\mu\mu$: 874.58(86.7%)
- $\tau\tau$: 134.42(13.3%)

4f の内訳は次の通りである。

- $\nu = 0$
 - $2q + 2\mu$: 41.45(0.36%)
 - $2e + 2\mu$: 1572.19(13.6%)
 - $2\mu + 2\tau$: 40.13(0.35%)
- $\nu = 1$
 - $2q + 1\nu + 1\mu$: 40.25(0.35%)
 - $2q + 1\nu + 1\tau$: 20.13(0.17%)
- $\nu = 2$
 - $2\nu + 2\mu$: 8272.06(71.7%)
 - $2\nu + 2\tau$: 41.30(0.36%)

- $2\nu + 1\mu + 1\tau$: 1510.71(13.1%)

$\gamma\gamma \rightarrow 4f$ の内訳は次の通りである。

- $\nu = 0$
 - 4μ : 3.17(1.7%)
- $\nu = 2$
 - $2\nu + 2\mu$: 173.60(94.8%)
 - $2\nu + 1\mu + 1\tau$: 6.30(3.4%)

また precuts の時点で何のバックグラウンドが残ったか、その event weight はいくつかも調べた。以下はそれぞれの process ID と event weight である。

- ffh , others
 - $\mu\mu h$, 106519(~ 0.8), 106520(< 0.1)
 - $\tau\tau h$, 106522(< 0.1)
 - $\nu\nu h$, 106523(~ 1.2)
- 2f
 - leptonic, 250106(~ 20.5), 250108(~ 1.2)
- 4f
 - ZZ semileptonic, 250014(~ 20.1), 250016(~ 1.1)
 - WW semileptonic, 250018(~ 20.1)
 - ZZ leptonic, 250022(~ 20), 250024(~ 1.1)
 - WW leptonic, 250026(~ 20.1), 250028(~ 1.1)
 - ZZ and WW mix leptonic, 250030(~ 1.8), 250032(~ 0.1)
 - single $Z e$ leptonic, 250033(~ 2.6), 250034(~ 5), 250035(~ 0.5), 250036(~ 0.2)
 - single $Z \nu$ leptonic, 250054(~ 1.8), 250056(~ 0.1)
- $\gamma\gamma \rightarrow 4f$
 - $2\ell + 2\nu$, 37401(~ 3.1), 37402(~ 2.9), 37403(~ 3)
 - 4ℓ , 37453(~ 3.1)
- 5f
 - $e + 2\ell + 2\nu$, 37189(~ 5), 37190(~ 4.2), 37191(~ 0.5), 37192(~ 0.2), 37305(~ 1.8), 37306(~ 0.8), 37307(~ 3.6), 37308(~ 3.1)
 - $3\ell + 2\nu$, 37193(~ 1.5), 37194(~ 0.4)
 - $e + \nu + \ell + x + y$, 37225(~ 5.7)
 - $2\ell + \nu + x + y$, 37315(~ 3.2)
 - $e + 2\ell + 2x$, 37242(~ 4.8)
 - $e + 4\ell$, 37258(~ 4.8), 37259(~ 0.5)

見ても通り、数多くのプロセスで統計量が大幅に不足している。

2.4 切り分け

シグナルは Higgs-strahlung (Zh) プロセスと WW -fusion プロセス (WWF) が混合している。プロセスが違うため、2つを切り分けてそれぞれで最適化を行い、解析精度の向上を目指す。今回は recoil mass を用いた。シグナルの場合、recoil mass は Z mass に対応する。次の図は recoil mass の分布である。

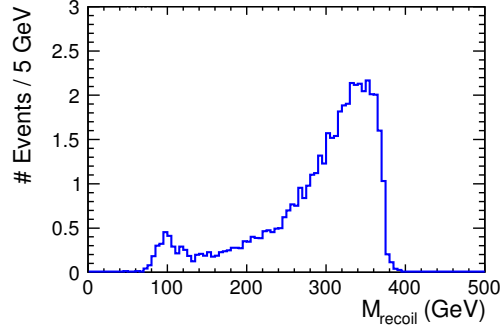


図 17: Recoil mass 分布。

100 GeV 付近のピークは Zh プロセス由来のイベントであることを示す。 Z mass より高い方向にシフトしているのは FSR 補正を行っていないためである。この図から切り分けのポイントを 120 GeV とした。これにより Zh イベントは 2.65、WWF イベントは 45.80 存在することになる。

2.5 Cut-based 解析

まず cut-based 解析を行った。ここでは切り分けを考えず、全てを一緒に扱った Mixed 解析を行った。最適カットは以下の通りとなった。

- $124 < M_{\mu^+\mu^-} < 126$ GeV
- $E_{\text{vis}} < 265$ GeV
- $P_t > 60$ GeV
- thrust < 0.88
- charge * $\cos \theta_{\mu^\pm} > -0.8$

以下はそれぞれの分布図である。

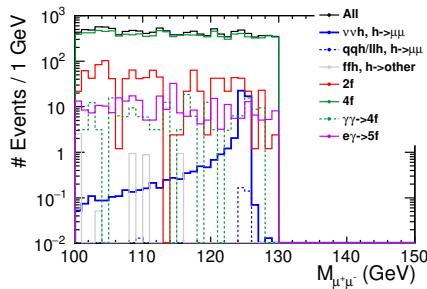


図 18: $M_{\mu^+\mu^-}$ 分布。

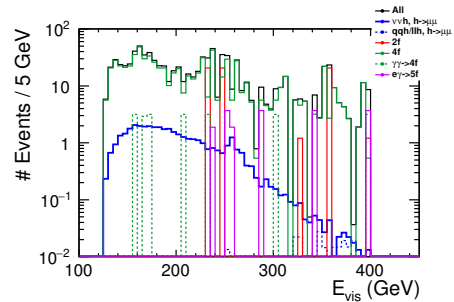


図 19: E_{vis} 分布。

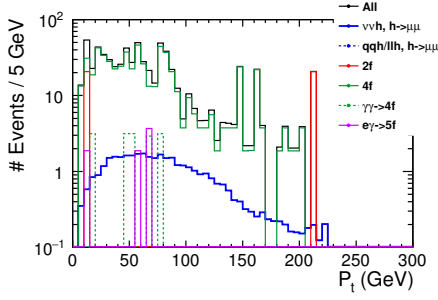


図 20: P_t 分布。

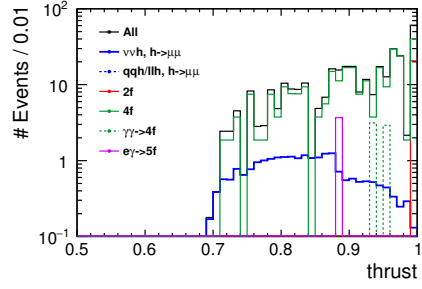


図 21: thrust 分布。

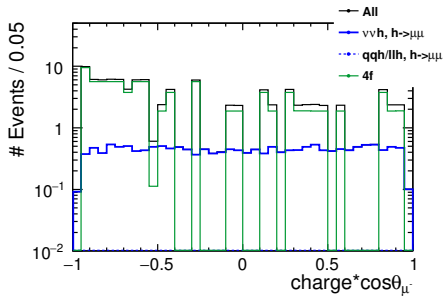


図 22: $\text{charge} * \cos \theta_{\mu^-}$ 分布。

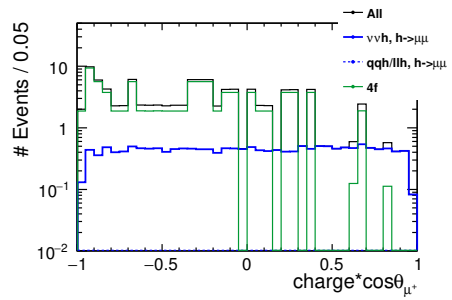


図 23: $\text{charge} * \cos \theta_{\mu^+}$ 分布。

以下はカットテーブルである。

表 2: カットテーブル。

	$\nu\nu h$ $h \rightarrow \mu\mu$	$qqh+\ell\ell h$ $h \rightarrow \mu\mu$	ffh $h \rightarrow \text{other}$	2f	4f	$\gamma\gamma \rightarrow 4f$	5f
precuts	48.45	0.39	5.35	1009.00	1.154×10^4	183.07	295.95
$M_{\mu\mu}$	39.44	0.31	0	64.20	747.29	21.81	18.48
E_{vis}	36.14	0.05	0	41.19	597.99	15.52	7.44
P_t	22.72	0.03	0	20.60	277.38	6.07	3.68
thrust	17.38	0.02	0	0	88.33	0	0
$\text{charge} * \cos \theta_{\mu\pm}$	14.67	0.02	0	0	48.85	0	0

解析から $N_{\text{sig}} = 14.67$ 、 $N_{\text{bkg}} = 48.87$ が得られ、統計的有意性は 1.8 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 54\%$ に一致する。

2.6 TMVA 解析

ここから BDTG を用いて解析する。今回は Zh のみ、WWF のみ、混合 (Mixed) の 3 パターンを試した。また $M_{\mu\mu}$ は最も強力なパラメータであり、ほとんどこれによって分離が決まってしまう。そのため、このパラメータを使う場合と使わない場合の両方をテストした。 $M_{\mu\mu}$ を入れない場合は、これを除いたパラメータで BDTG 解析を最適化し、その後に $M_{\mu\mu}$ にカットを適用して最適化を行うこととした。

2.6.1 解析 1 — Zh without $M_{\mu\mu}$

以下の 6 個のパラメータを用いた。

- E_{vis} , P_t , thrust, $\cos\theta_{\text{thrust}}$
- charge * $\cos\theta_{\mu+}$, charge * $\cos\theta_{\mu-}$

以下はそれぞれのパラメータの分布図である。

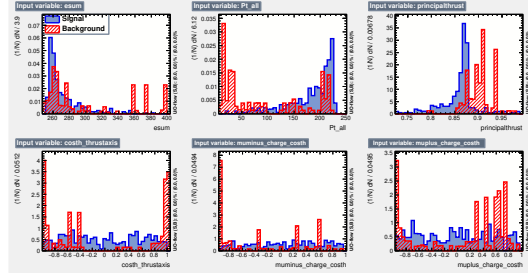


図 24: パラメータ分布。

次の 2 図は TMVA 解析の結果である。

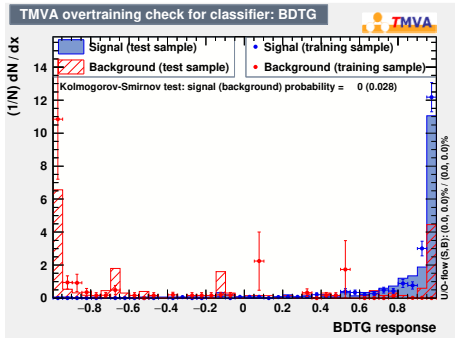


図 25: BDTG output 分布。

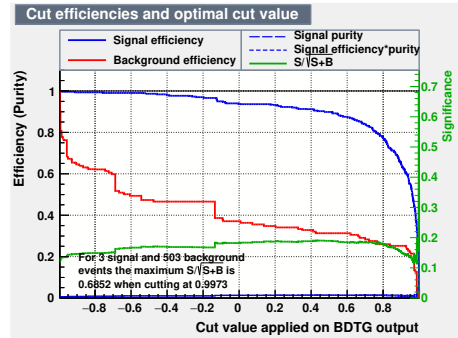


図 26: Significance 分布。

解析から $N_{\text{sig}} = 0.47$, $N_{\text{bkg}} = 0$ が得られ、統計的有意性は 0.7 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 146\%$ に一致する。バックグラウンドが 0 になったため、追加のカットをかけることは無意味である。

2.6.2 解析 2 — Zh with $M_{\mu\mu}$

以下の 6 個のパラメータを用いた。

- P_t , $\cos\theta_{\text{thrust}}$, $\cos\theta_{\text{thrustaxis}}$
- $M_{\mu\mu}$
- charge * $\cos\theta_{\mu+}$, charge * $\cos\theta_{\mu-}$

以下はそれぞれのパラメータの分布図である。

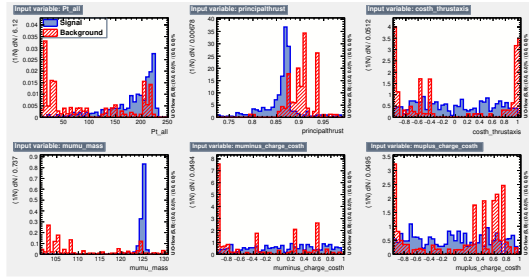


図 27: パラメータ分布。

次の 2 図は TMVA 解析の結果である。

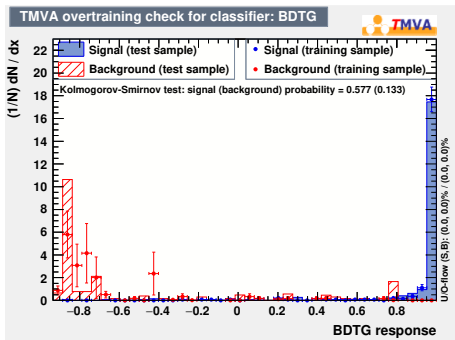


図 28: BDTG output 分布。

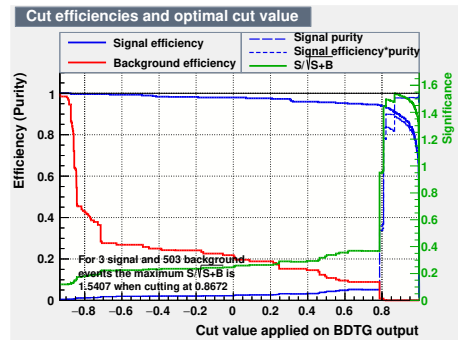


図 29: Significance 分布。

解析から $N_{\text{sig}} = 2.43$ 、 $N_{\text{bkg}} = 0.05$ が得られ、統計的有意性は 1.5 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 65\%$ に一致する。

2.6.3 解析 3 — WWF without $M_{\mu\mu}$

以下の 5 個のパラメータを用いた。

- E_{vis} 、 P_t 、thrust
- $\text{charge} * \cos \theta_{\mu^+}$ 、 $\text{charge} * \cos \theta_{\mu^-}$

以下はそれぞれのパラメータの分布図である。

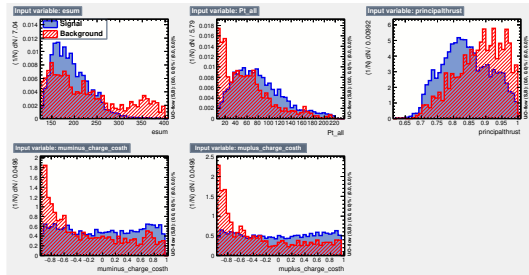


図 30: パラメータ分布。

次の2図は TMVA 解析の結果である。

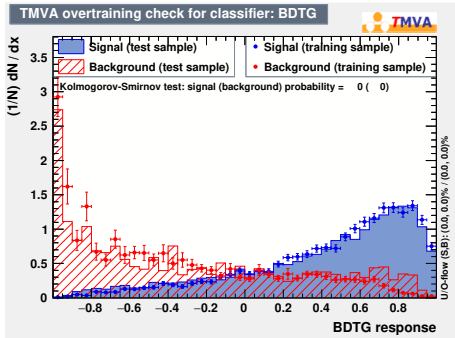


図 31: BDTG output 分布。

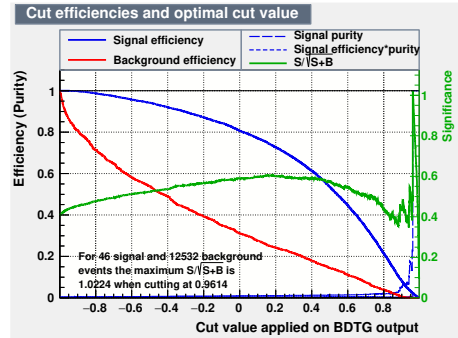


図 32: Significance 分布。

解析から $N_{\text{sig}} = 1.05$, $N_{\text{bkg}} = 0$ が得られ、統計的有意性は 1.0 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 98\%$ に一致する。バックグラウンドが 0 になったため、追加のカットをかけることは無意味である。

2.6.4 解析 4 — WWF with $M_{\mu\mu}$

以下の 5 個のパラメータを用いた。

- P_t , thrust
- $M_{\mu\mu}$
- charge * $\cos\theta_{\mu^+}$, charge * $\cos\theta_{\mu^-}$

以下はそれぞれのパラメータの分布図である。

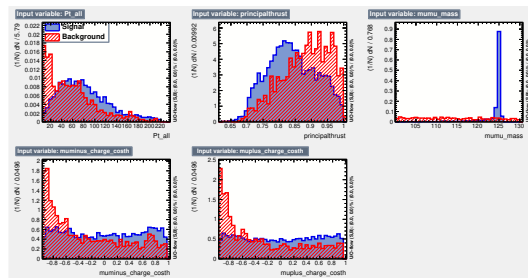


図 33: パラメータ分布。

次の 2 図は TMVA 解析の結果である。

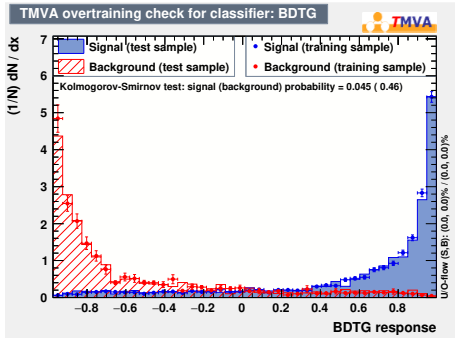


図 34: BDTG output 分布。

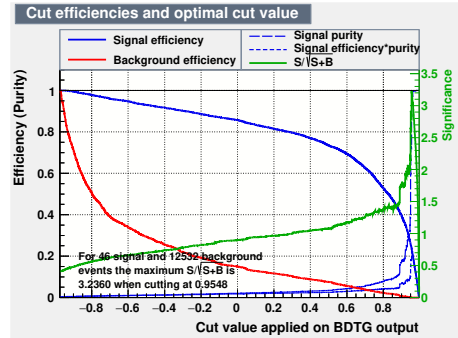


図 35: Significance 分布。

解析から $N_{\text{sig}} = 10.96$ 、 $N_{\text{bkg}} = 0.51$ が得られ、統計的有意性は 3.2 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 31\%$ に一致する。

2.6.5 解析 5 — Mixed without $M_{\mu\mu}$

以下の 6 個のパラメータを用いた。

- E_{vis} 、 P_t 、 thrust 、 $\cos\theta_{\text{thrust}}$
- $\text{charge} * \cos\theta_{\mu^+}$ 、 $\text{charge} * \cos\theta_{\mu^-}$

以下はそれぞれのパラメータの分布図である。

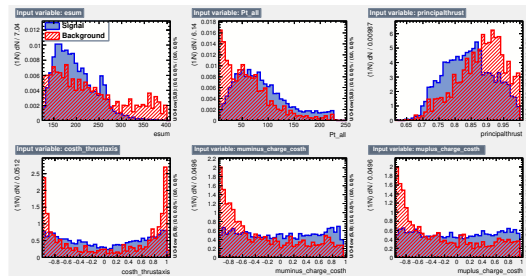


図 36: パラメータ分布。

次の 2 図は TMVA 解析の結果である。

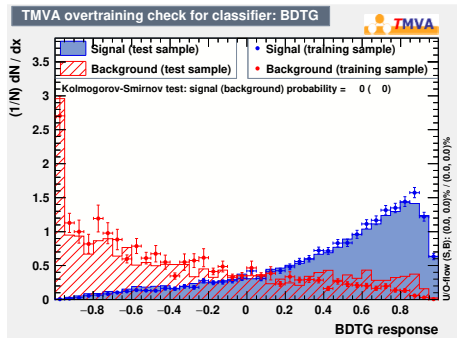


図 37: BDTG output 分布。

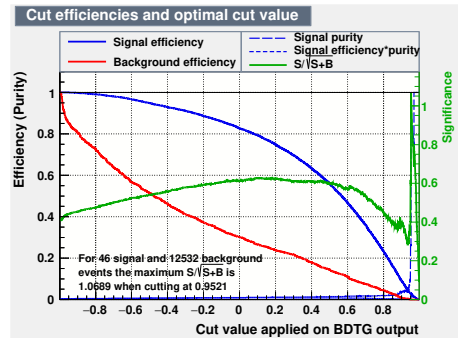


図 38: Significance 分布。

解析から $N_{\text{sig}} = 1.33$ 、 $N_{\text{bkg}} = 0.22$ が得られ、統計的有意性は 1.1 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 94\%$ に一致する。

2.6.6 解析 6 — Mixed with $M_{\mu\mu}$

以下の 5 個のパラメータを用いた。

- P_t , thrust
- $M_{\mu\mu}$
- $\text{charge} * \cos \theta_{\mu^+}$, $\text{charge} * \cos \theta_{\mu^-}$

以下はそれぞれのパラメータの分布図である。

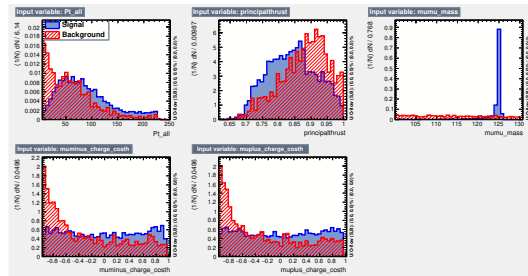


図 39: パラメータ分布。

次の 2 図は TMVA 解析の結果である。

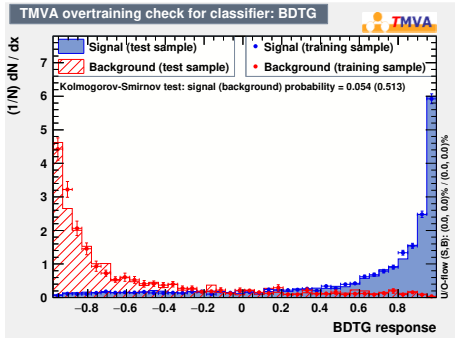


図 40: BDTG output 分布。

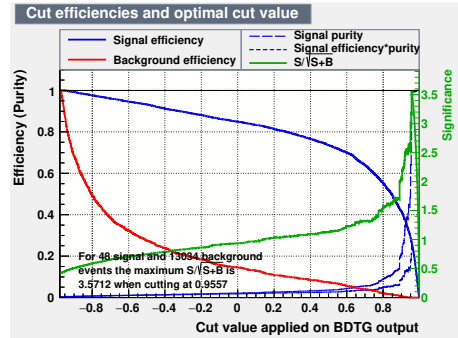


図 41: Significance 分布。

解析から $N_{\text{sig}} = 13.04$ 、 $N_{\text{bkg}} = 0.30$ が得られ、統計的有意性は 3.6 となった。これは $\frac{\Delta(\sigma \times \text{BR})}{(\sigma \times \text{BR})} = 28\%$ に一致する。

3 結果のまとめ

得られた結果をまとめたのが以下の表である。

表 3: 解析結果のまとめ

	Zh	WWF	Mixed
without $M_{\mu\mu}$	146%	98%	94%
with $M_{\mu\mu}$	65%	31%	28%

このうち with $M_{\mu\mu}$ ケースにおける Zh と WWF を combine すると 28% となり、Mixed の結果と同じである。またシグナル効率 100% でバックグラウンドなしの理想的なケースを考えた時、その時の精度は 13% であることから、これらの結果は理想の結果からファクター 2 離れていることになる。もちろん MC 統計量が不足していることを忘れてはならない。

4 結論・今後の進展

さらに効率的な precuts を考案し、TMVA 解析を進めた。ベストケースで精度 28% という結果を得た。これは理想結果からファクター 2 だけ離れている。また、MC 統計量が明確に足りないことも定量的に示した。

今後の課題としては以下の点が挙げられる。

- IsolatedLeptonTagger では E_{yoke} を見るため、ミューオンの本数をカットとして要求すると SGV サンプル ($\gamma\gamma \rightarrow 2f$ 及び $3f$) が全てカットされる。これらをどう扱うか。
- re-weighting をどのように理解しどのように適用するか。

参考文献

- [1] 川田真一 “重心エネルギー 500 GeV ILC における $h \rightarrow \mu^+\mu^-$ モードの解析”

[2] 川田真一 “重心エネルギー 500 GeV ILC における $h \rightarrow \mu^+ \mu^-$ モードの解析 — part 2”

[3] 川田真一 “重心エネルギー 500 GeV ILC における $h \rightarrow \mu^+ \mu^-$ モードの解析 — part 3”

[4] Junping Tian, Claude Dürig “isolated lepton finder”

<https://agenda.linearcollider.org/event/6787/contributions/33415/>

[attachments/27509/41775/IsoLep_HLRec2016.pdf](https://agenda.linearcollider.org/event/6787/contributions/33415/attachments/27509/41775/IsoLep_HLRec2016.pdf)