CEPC HZZ Project

Min Zhong, Yanxi Gu

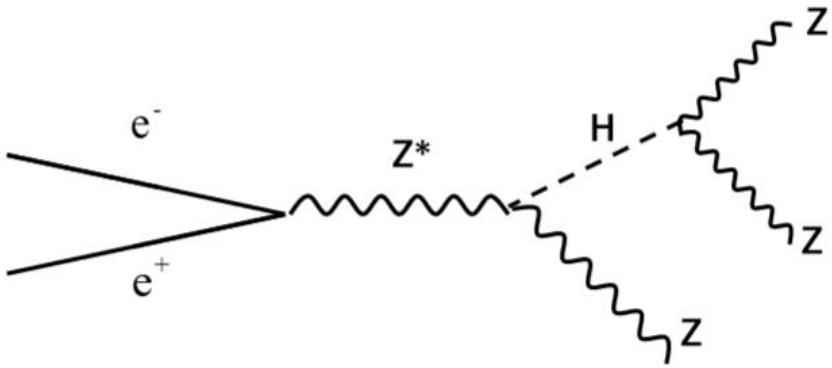
Aug 7th, 2019

Introduction

CEPC

- e^+e^- collider, $\sqrt{s}$ can be precisely controlled
- Great for measurement of Higgs properties

HZZ Analysis



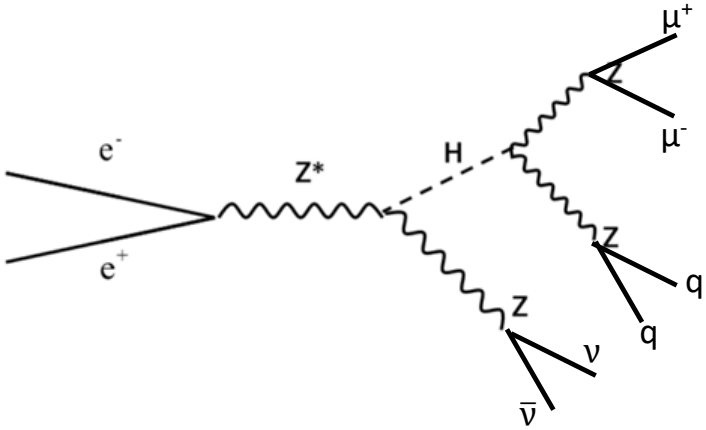
- $H \rightarrow ZZ^*$ Branch Ratio ≈ 2.64%
- One of the key factors to deduce the Higgs boson width and precision

$$\Gamma_H = \frac{\Gamma(H \rightarrow ZZ^*)}{BR(H \rightarrow ZZ^*)}$$

Sample Introduction

Signal: nnhzz->mumujj/jjmumu

category	Xsection (fb)
nnh_zz	1.22

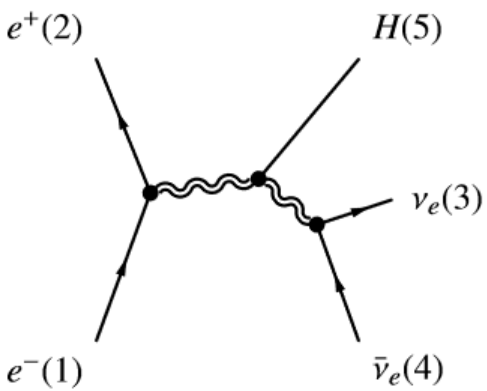


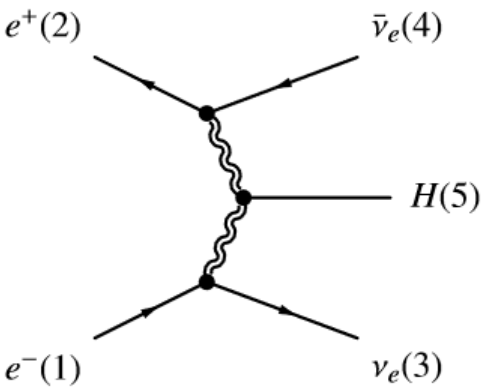
Only e1e1h_zz, e2e2h_zz, e3e3h_zz used by Alex

ZH background

category	Xsection (fb)
e1e1h_X	7.04
e2e2h_X	6.77
e3e3h_X	6.75
nnh_X	46.3
qqh_X	137

Not used by Alex



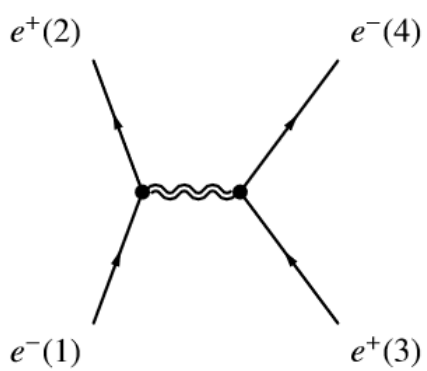


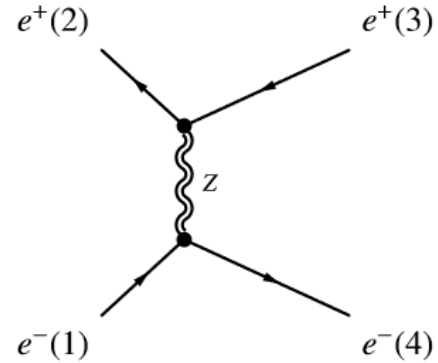
Sample Introduction

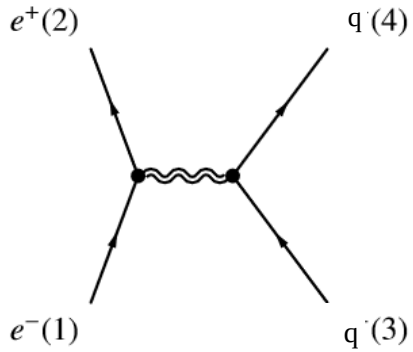
2 fermion background

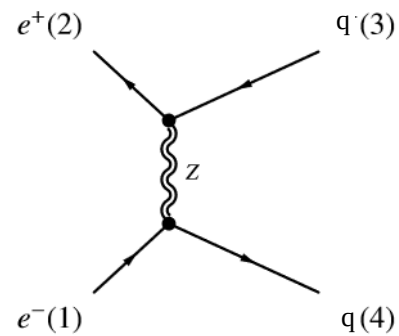
category	Xsection (fb)
qq	54106.86
e1e1	24770.90
e2e2	5332.71
e3e3	4752.89
n1n1	45390.79
n2n2	4416.30
n3n3	4410.26

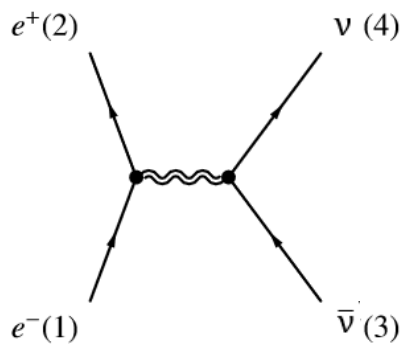
All the 2 fermion backgrounds are not used by Alex

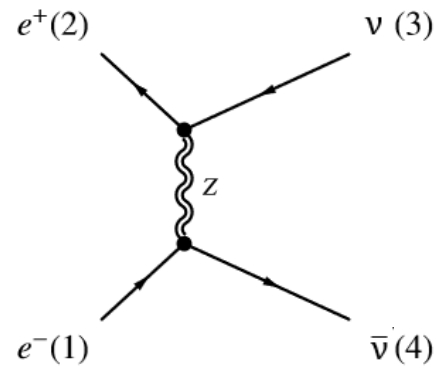










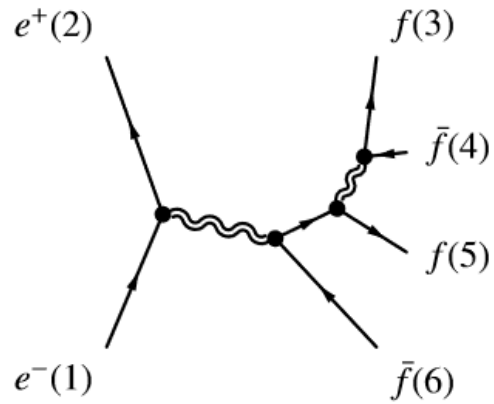


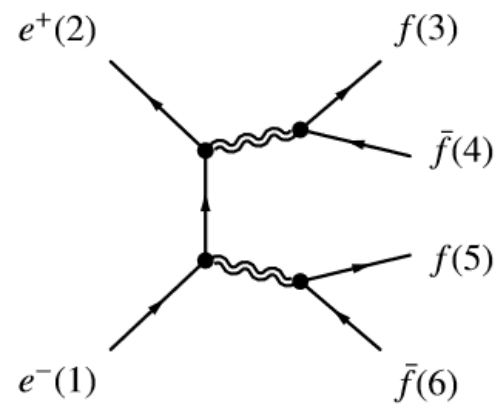
Sample Introduction

4 fermion background

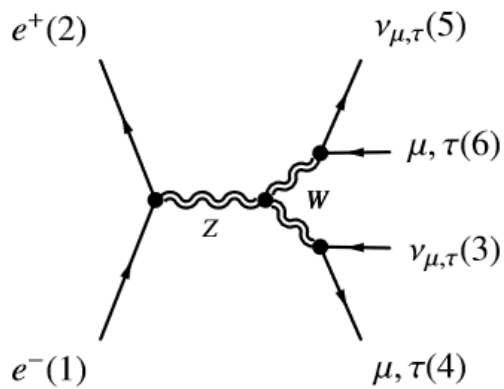
category
ZZ
WW
single Z e
single W
ZW mixing

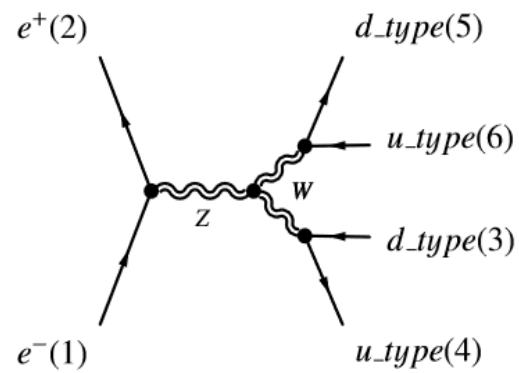
Alex used one more file for this process (sze_l0e, see next page)



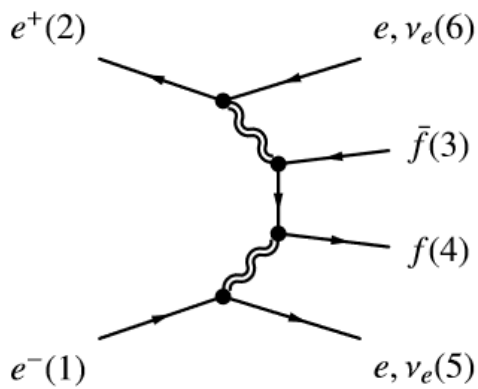


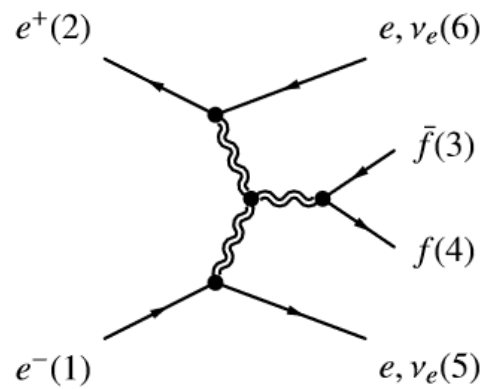
4f background





WW



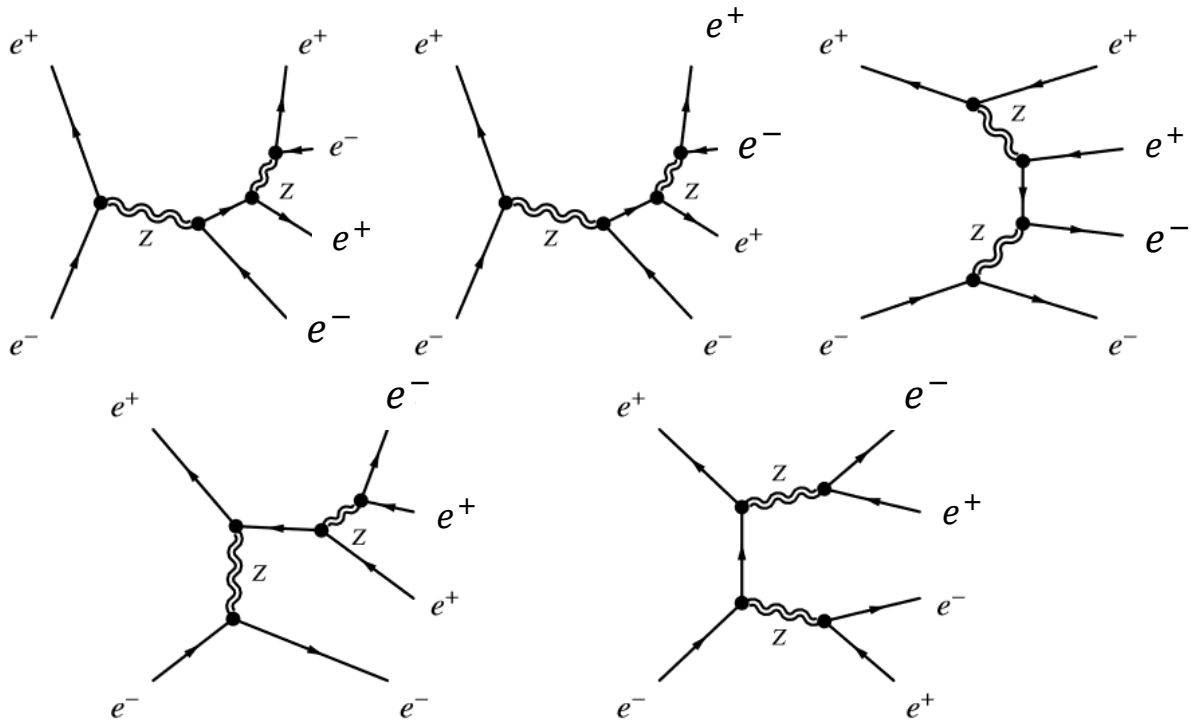


Single Z e

Sample Introduction

sze_l0e (not used by Ryuta)

sze_l0tau



Sample Introduction

	Ryuta	Alex
signal	nnh_zz ($\mu^+\mu^- jj/jj\mu^+\mu^-$) (truth info. used)	nnh_zz
ZH background	e1e1h_zz, e1e1h_bb …	e1e1h_zz
	e2e2h_zz e2e2h_bb …	e2e2h_zz
	e3e3h_zz e3e3h_bb …	e3e3h_zz
	nnh_zz (not $\mu^+\mu^- jj/jj\mu^+\mu^-$) (truth info. used) nnh_bb …	
	qqh_zz qqh_bb …	
2 fermion background	qq, e1e1, e2e2, e3e3, n1n1, n2n2, n3n3	
4 fermion background	ZZ WW Single Z e **(sze_l0mu, sze_l0tau, sze_h0ududd, sze_h0cscs)** Single W ZW mixing	ZZ WW Single Z e (same 4 add sze_l0e) Single W ZW mixing

Target

Two Channels

- $Z(\mu^+\mu^-)H(Z \to \nu\nu, Z^* \to jj)$
  $Z(\mu^+\mu^-)H(Z \to jj, Z^* \to \nu\nu)$ **(Lingteng, Ryuta)**
- $Z(\nu\nu)H(Z \to \mu^+\mu^-, Z^* \to jj)$
  $Z(\nu\nu)H(Z \to jj, Z^* \to \mu^+\mu^-)$ **(Alex)**

What to do

- **Learn how to use Ryuta's framework and reproduce Lingteng's results**
- **Implement Alex's results using Ryuta's framework**
- **Optimize cut-based analysis using BDT results**

Current Status and Results

Current Status

- BDT results compared with Alex's results
- Cut-based results after optimization inspired by BDT
- Histograms ready for Higgs width fitting

Information

➢ Nomalization

Number of signal events expected:

$$N_{exp}(signal) = 5600 \times 46.3 \times 0.0264 \times 0.69 \times 0.03 \times 2 = 283$$

$\int L$ $\quad$ nnH $\quad$ $H \to ZZ^*$ $\quad$ $Z^* \to jj$ $\quad$ $Z \to \mu^+\mu^-$ $\quad$ $\mu^+\mu^- jj$ $/jj\mu^+\mu^-$

All the plots and tables shown later are normalized to $\int L = 5600\ fb^{-1}$

Information

- **Different nomalization method for signal channel**
  - **Ryuta and Lingteng's method**
    - **Normalized the** $Z(\nu\nu)H(Z \to \mu^+\mu^-, Z^* \to jj)$ & $Z(\nu\nu)H(Z \to jj, Z^* \to \mu^+\mu^-)$ **to the expected number 283 using truth information**
    - **The other numbers were normalized correspondingly**
    - **Different from the normalization of the backgrounds**
  - **Alex's method**
    - **Normalized the** nnh_zz **sample to expected number**
    - **The other numbers were normalized correspondingly**
    - **Truth information wasn't used**
    - **Consistent with the normalization of the backgrounds**

Information

➢ Cut flow before BDT (compared with Alex's)

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	6 927	3 074	6.4E6	2.0E7	7.9E6
Final state	236	819	395 848	1 578	158 991
nPFO	227	799	376 652	8	6 319
Vis Mass	190	42	1 804	2	463
Costheta	136	30	300	1	52
BDT	100	17	49	0	2

Alex's table (Rerun 6th Aug. 2019)

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	8 544	3 043	6.4E6	2.0E7	7.5E6
Pre-select	314	687	394 766	141 548	177 651
Is signal	283	687	394 766	141 548	177 651
nPFO	283	670	356 833	1 510	97 782
Vis Mass	224	34	1 941	220	6 524
Costheta	160	24	315	139	75
BDT	107	1	9	1	0

Our table

Changed correspondingly

Using Ryuta's normalization method

Information

➢ Cut flow before BDT (compared with Alex's)

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	6 927	3 074	6.4E6	2.0E7	7.9E6
Final state	236	819	395 848	1 578	158 991
nPFO	227	799	376 652	8	6 319
Vis Mass	190	42	1 804	2	463
Costheta	136	30	300	1	52
BDT	100	17	49	0	2

Alex's table
(Rerun 6th Aug. 2019)

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	6840	3 043	6.4E6	2.0E7	7.5E6
Pre-select	250	687	394 766	141 548	177 651
Is signal	226	687	394 766	141 548	177 651
nPFO	226	670	356 833	1 510	97 782
Vis Mass	178	34	1 941	220	6 524
Costheta	128	24	315	139	75
BDT	86	1	9	1	0

Our table

Using Alex's normalization method

Changed correspondingly

Information

➢ Cut flow before BDT (compared with Alex's)

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	6 927	3 074	6.4E6	2.0E7	7.9E6
Final state	236	819	395 848	1 578	158 991
nPFO	227	799	376 652	8	6 319
Vis Mass	190	42	1 804	2	463
Costheta	136	30	300	1	52
BDT	100	17	49	0	2

Alex's table
(Rerun 6th Aug. 2019)

Find two differences in jobfile (minimum Jet Energy and Minimum Jet nPFOs), which may cause these differences

	nnh_zz	llh_zz	zz	zzorww	sze
Expected	6840	3 043	6.4E6	2.0E7	7.5E6
Pre-select	250	687	394 766	141 548	177 651
Is signal	226	687	394 766	141 548	177 651
nPFO	226	670	356 833	1 510	97 782
Vis Mass	178	34	1 941	220	6 524
Costheta	128	24	315	139	75
BDT	86	1	9	1	0

Our table

BDT Results

➢ nnhzz Cut Flow

The counts in the table are normalized to appropriate cross sections and $\int L = 5600\ fb^{-1}$ (Using Ryuta's normalization method)

	signal	zh	2f background	4f background
Pre-selection	314	30 629	7 049 779	756 489
Is signal	283	30 629	7 049 779	756 489
nPFO >= 10	283	29 891	296 305	459 078
115 < Visible Mass < 130	224	463	19 810	8 907
|Cos theta|<0.9	160	324	419	672
BDT score > 0	107	30	0	16

BDT Results

➢ nnhzz Cut Flow compared with Alex's

	nnh_zz	llh_zz	zz	zzorww	sze
Pre-selection	250	687	394 766	141 548	177 651
Is signal	226	687	394 766	141 548	177 651
nPFO	226	670	356 833	1 510	97 782
Vis Mass	178	34	1 941	220	6 524
Costheta	128	24	315	139	75
BDT	86	1	9	1	0

Part of

Part of

	nnh_zz	zh	2f background	4f background
Pre-selection	250	30 629	7 049 779	756 489
Is signal	226	30 629	7 049 779	756 489
nPFO >= 10	226	29 891	296 305	459 078
115 < Visible Mass < 130	178	463	19 810	8 907
|Cos theta|<0.9	128	324	419	672
BDT score > 0	86	30	0	16

(Using Alex's normalization method)

Cut-based Results

- **Make eye-based cuts according to the distributions**

- **Cut flow**

cut	nnhzz	zh	2f	4f
Raw events	8544	17124477	12038992073	1609637701
pre-selection	313	30880	7049779	756489
2m+2j	283	30654	7049779	756489
Npfo	283	29899	296305	459078
Vis_mass	223	466	19810	8907
cos_theta	160	326	419	672
RecM(dimuon)	130	260	198	431
vis_all_p	104	91	13	86
M(dijet)	103	87	6	80
jet_lead_e	103	87	6	80
jet_sub_e	103	87	6	80
angle_mj	103	87	6	71
M(dimuon)	102	77	6	67
vis_all_cos	102	77	6	67
RecM(vis_all)	101	74	6	65

- **Cuts Info.**

Pre-selection
2mu+2j
Npfo>=10
115<Vis_all_mass<130
|cos|<0.9
130<dimuon_rec_m<220
43<Vis_all_p<60
10<dijet_m<100
13<dimuon_m<100
Mj_angle<170
10<Lead_jet_e<95
Sub_jet_e<59
|vis_all_cos|<0.9
80<vis_all_rec_m<107

Optimized Cut-based Results

- **Optimize the cuts according to the distributions after BDT cut**

- **Cut flow**

cut	nnhzz	zh	2f	4f
Raw events	8544	17124477	12038992073	1609637701
pre-selection	313	30880	7049779	756489
2m+2j	283	30654	7049779	756489
Npfo	247	10591	87605	272031
Vis_mass	181	228	5782	3341
cos_theta	130	168	73	223
RecM(dimuon)	128	147	73	188
vis_all_p	114	79	0	56
M(dijet)	112	77	0	50
jet_lead_e	107	63	0	28
jet_sub_e	104	61	0	24
angle_mj	98	54	0	18
M(dimuon)	98	54	0	16
vis_all_cos	98	54	0	16
RecM(vis_all)	96	52	0	14

- **Cuts Info.**

Pre-selection
2mu+2j
20<Npfo<73
120<Vis_all_mass<130
|cos|<0.9
104<dimuon_rec_m<214
40<Vis_all_p<70
14<dijet_m<99
14<Lead_jet_e<69
3<Sub_jet_e<49
20<Mj_angle<142
12<dimuon_m<96
|vis_all_cos|<0.9
69<vis_all_rec_m<101

Optimized Cut-based Results

- **Optimized cut-based results compared with no-optimized cut-based results**

Before Optimization

cut	nnhzz	zh	2f	4f
Raw events	8544	17124477	12038992073	1609637701
pre-selection	313	30880	7049779	756489
2m+2j	283	30654	7049779	756489
Npfo	283	29899	296305	459078
Vis_mass	223	466	19810	8907
cos_theta	160	326	419	672
RecM(dimuon)	130	260	198	431
vis_all_p	104	91	13	86
M(dijet)	103	87	6	80
jet_lead_e	103	87	6	80
jet_sub_e	103	87	6	80
angle_mj	103	87	6	71
M(dimuon)	102	77	6	67
vis_all_cos	102	77	6	67
RecM(vis_all)	101	74	6	65

Using Ryuta's normalization method

After Optimization

cut	nnhzz	zh	2f	4f
Raw events	8544	17124477	12038992073	1609637701
pre-selection	313	30880	7049779	756489
2m+2j	283	30654	7049779	756489
Npfo	247	10591	87605	272031
Vis_mass	181	228	5782	3341
cos_theta	130	168	73	223
RecM(dimuon)	128	147	73	188
vis_all_p	114	79	0	56
M(dijet)	112	77	0	50
jet_lead_e	107	63	0	28
jet_sub_e	104	61	0	24
angle_mj	98	54	0	18
M(dimuon)	98	54	0	16
vis_all_cos	98	54	0	16
RecM(vis_all)	96	52	0	14

Results

(Using Ryuta's normalization method)

➢ Optimized cut-based results compared with BDT results

Optimized Cut-based

	signal	zh	2f background	4f background
Pre-selection	313	30 629	7 049 779	756 489
Is signal	283	30 629	7 049 779	756 489
nPFO >= 10	283	29 891	296 305	459 078
115 < Visible Mass < 130	224	463	19 810	8 907
|Cos theta|<0.9	160	324	419	672
BDT score > 0	107	30	0	16

BDT

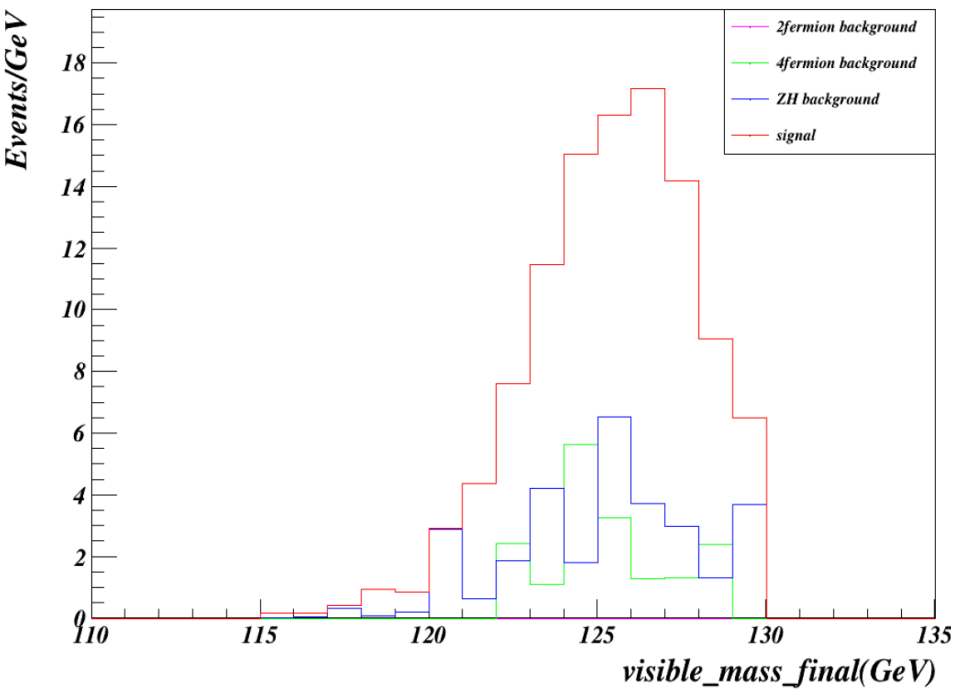
	signal	zh	2f background	4f background
Pre-selection	313	30 880	7 049 779	756 489
Is signal	283	30 654	7 049 779	756 489
20 <= nPFO <= 73	247	10 591	87 605	272 031
120 < Visible Mass < 130	181	228	5782	3 341
|Cos theta|<0.9	130	168	73	223
BDT score > 0	96	52	0	14

BDT Results

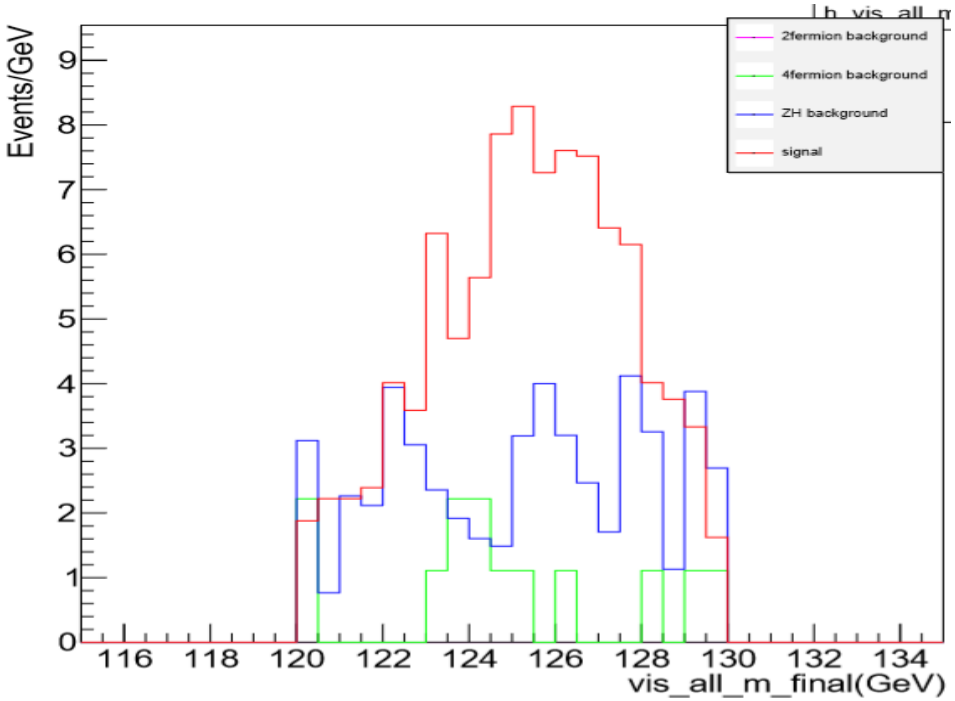
- **nnh_zz cut-based results optimized by BDT results**

(Using Ryuta's normalization method)

Higgs mass Plot



BDT result



Optimized cut-based result

Questions

- **Some questions for Ryuta and Lingteng**
  - **Could you tell us how to modify your jobfile to add the minimum jet energy and the minimum jet npfos selections (for Alex's jobfile, they are "EjetMin" and "nPFOmin")**
  - **Could you point the document which contains feyaman diagrams of 2 fermion backgrounds and zh backgrounds? (We only found the feynman diagrams of 4 fermion backgrounds in the document last time you sent us)**
  - **We found zzorww_l0cscs and zzorww_l0ududd samples in the dataset, but we didn't find zzorww_l0csud or zzorww_l0udcs samples, why are they not listed?**
  - **We found some differences of the cross sections of that in your txt files and that in the document, do we need to change them? (which is latest?)**

Next to do

- **Fix the difference with Alex's results for the zzorww process (probably the JetMinE and the JetNPFO issue)**
- **Verify the dataset details**
- **Update out results**
- **Fitting the Higgs width again**

Backup Slides Begin

Backup

ZH background cross section

The unit for all the Xsections is fb

category	Xsection	category	Xsection	category	Xsection	category	Xsection	category	Xsection
e1e1h_aa	1.61E-02	e2e2h_aa	1.54E-02	e3e3h_aa	1.54E-02	nnh_aa	1.06E-01	qqh_aa	3.12E-01
e1e1h_az	1.08E-02	e2e2h_az	1.04E-02	e3e3h_az	1.03E-02	nnh_az	7.08E-02	qqh_az	2.09E-01
e1e1h_bb	4.06E+00	e2e2h_bb	3.91E+00	e3e3h_bb	3.89E+00	nnh_bb	2.67E+01	qqh_bb	7.89E+01
e1e1h_cc	2.05E-01	e2e2h_cc	1.97E-01	e3e3h_cc	1.96E-01	nnh_cc	1.35E+00	qqh_cc	3.98E+00
e1e1h_e2e2	1.54E-03	e2e2h_e2e2	1.48E-03	e3e3h_e2e2	1.48E-03	nnh_e2e2	1.01E-02	qqh_e2e2	3.00E-02
e1e1h_e3e3	4.45E-01	e2e2h_e3e3	4.28E-01	e3e3h_e3e3	4.27E-01	nnh_e3e3	2.93E+00	qqh_e3e3	8.65E+00
e1e1h_gg	6.03E-01	e2e2h_gg	5.80E-01	e3e3h_gg	5.78E-01	nnh_gg	3.97E+00	qqh_gg	1.17E+01
e1e1h_ss	0.00E+00	e2e2h_ss	0.00E+00	e3e3h_ss	0.00E+00	nnh_ss	0.00E+00	qqh_ss	0.00E+00
e1e1h_ww	1.51E+00	e2e2h_ww	1.46E+00	e3e3h_ww	1.45E+00	nnh_ww	9.95E+00	qqh_ww	2.94E+01
e1e1h_zz	1.86E-01	e2e2h_zz	1.79E-01	e3e3h_zz	1.78E-01	nnh_zz	1.22E+00	qqh_zz	3.61E+00

Backup

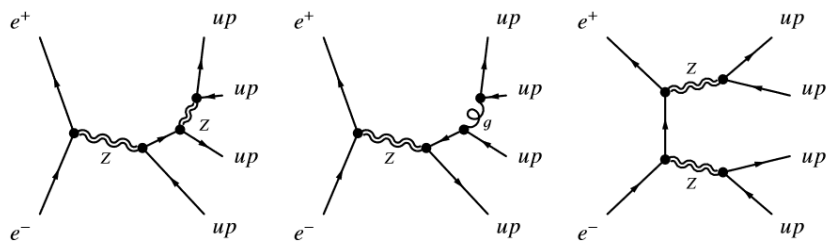
4f background cross section

category	Xsection	category	Xsection	category	Xsection
sw_l0mu	436.70	zz_l04tau	4.61	zzorww_h0cscs	1607.55
sw_l0tau	435.93	zz_l04mu	15.56	zzorww_l0mumu	221.10
sw_sl0qq	2612.62	zz_l0taumu	18.56	zzorww_l0tautau	211.18
szeorsw_l0l	249.48	zz_l0mumu	19.38	sze_l0tau	147.28
zz_h0utut	85.68	zz_l0tautau	9.61	sze_l0mu	845.81
zz_h0dtdt	233.46	ww_h0cuxx	3478.89	sze_l0nunu	28.94
zz_h0uu_notd	98.56	ww_h0uubd	0.05	sze_sl0uu	190.21
zz_h0cc_nots	98.97	ww_h0uusd	170.45	sze_sl0dd	125.83
zz_sl0nu_up	84.38	ww_h0ccbs	5.89	sznu_l0mumu	43.42
zz_sl0nu_down	139.71	ww_h0ccds	170.18	sznu_l0tautau	14.57
zz_sl0mu_up	87.39	ww_sl0muq	2423.43	sznu_sl0nu_up	55.59
zz_sl0mu_down	136.14	ww_sl0tauq	2423.56	sznu_sl0nu_down	90.03
zz_sl0tau_up	41.56	ww_l0ll	403.66		
zz_sl0tau_down	67.31	zzorww_h0ududd	1610.32		

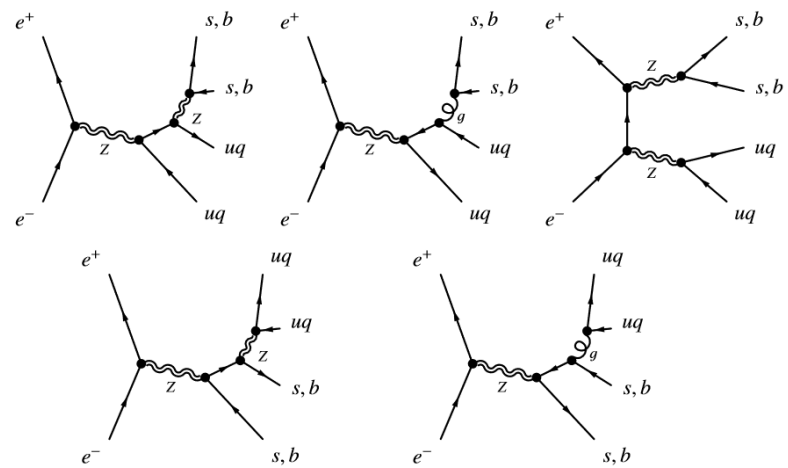
Backup: 4f background Feynman plot

4f background: ZZ category

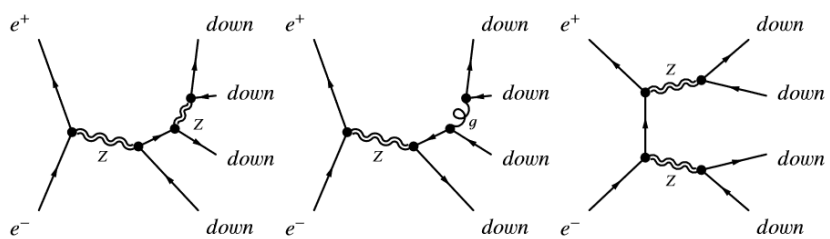
zz_h0utut



zz_h0uu_notd



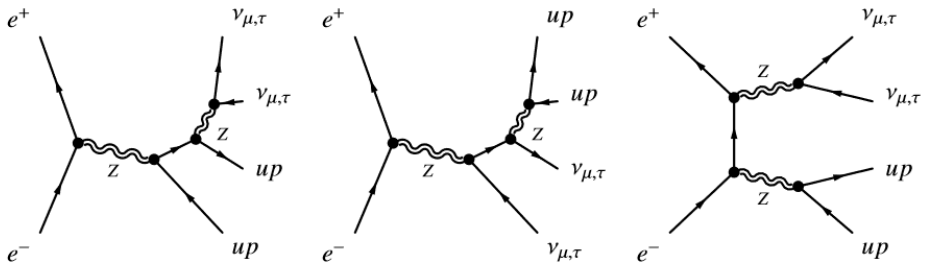
zz_h0dtdt



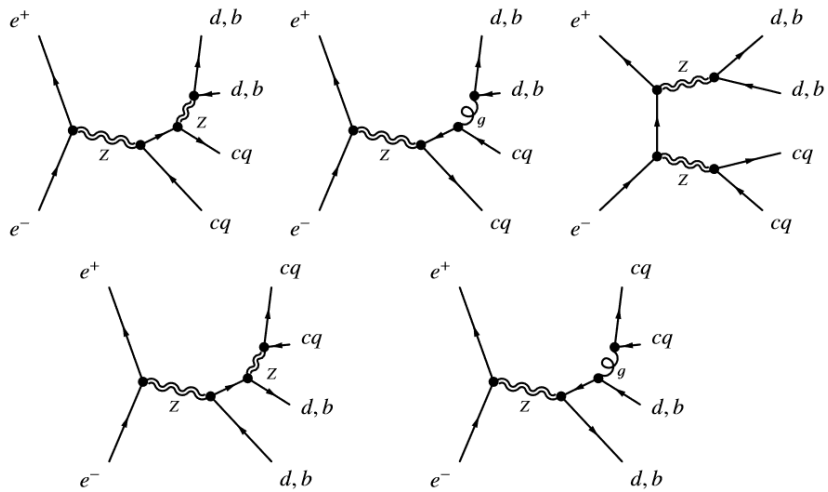
Backup: 4f background Feynman plot

4f background: ZZ category

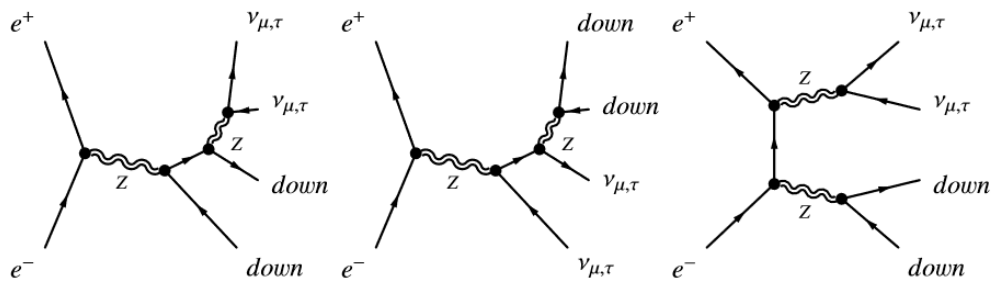
zz_sl0nu_up



zz_h0cc_nots



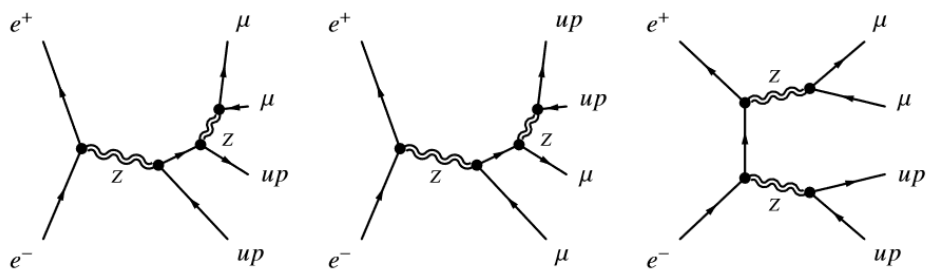
zz_sl0nu_down



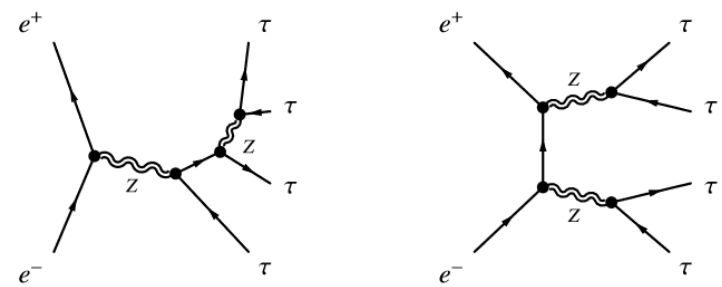
Backup: 4f background Feynman plot

4f background: ZZ category

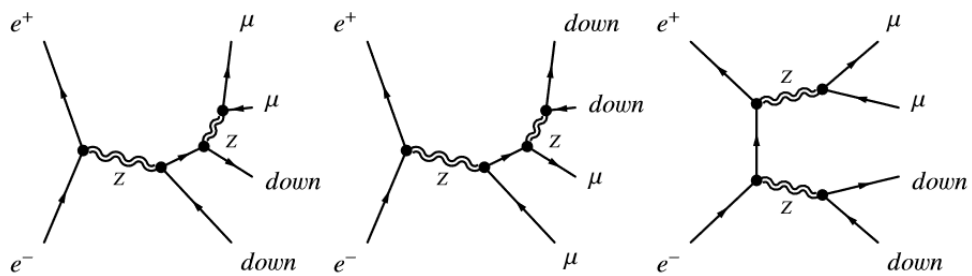
zz_sl0mu_up



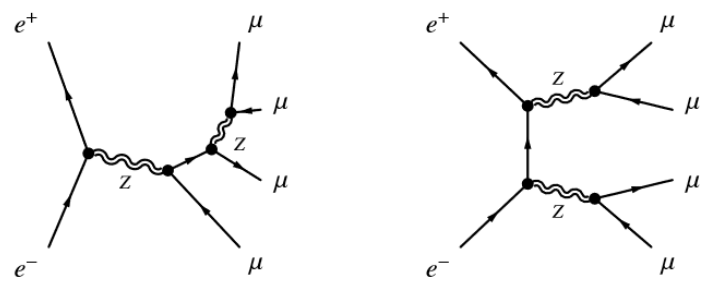
zz_l04tau



zz_sl0mu_down



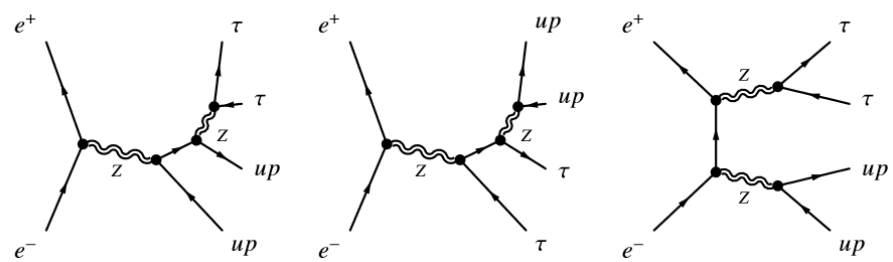
zz_l04mu



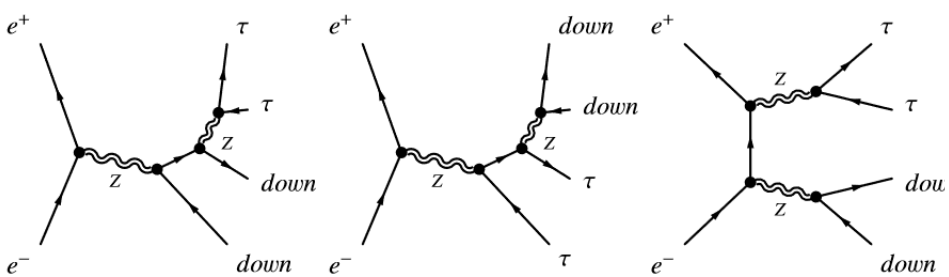
Backup: 4f background Feynman plot

4f background: ZZ category

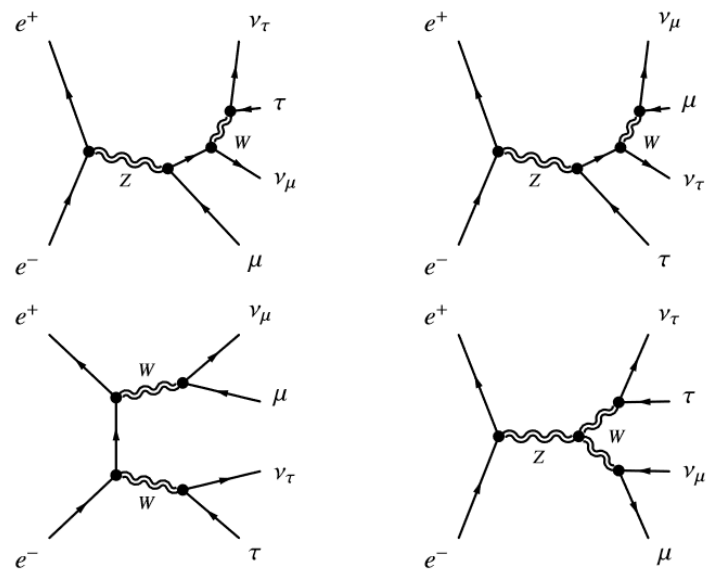
zz_sl0tau_up



zz_sl0tau_down



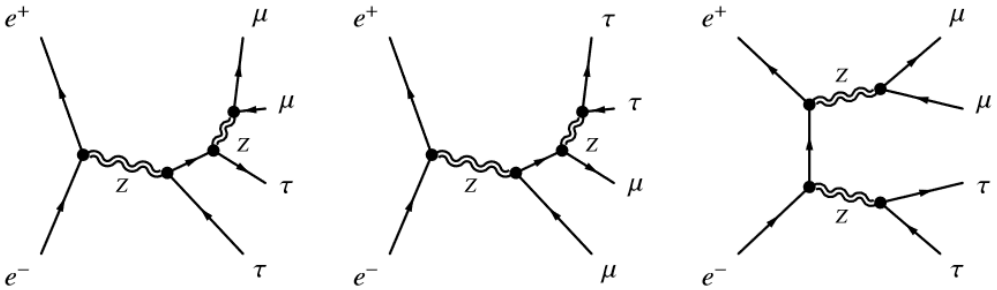
ww_l0ll



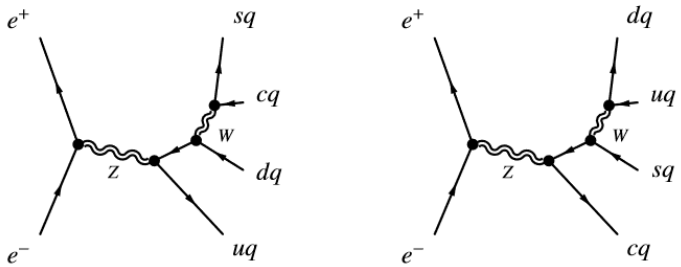
Backup: 4f background Feynman plot

4f background: ZZ category

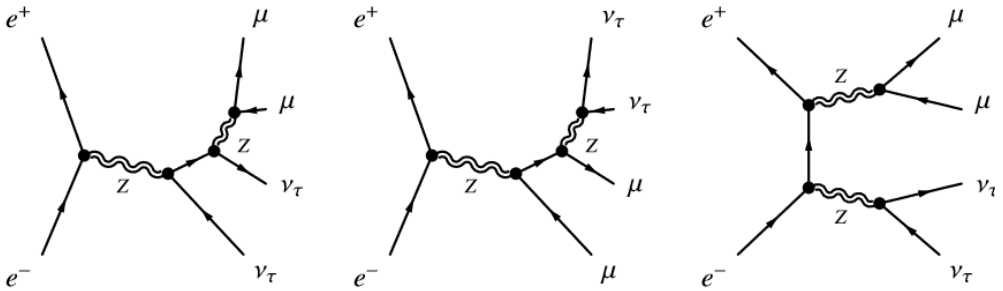
zz_l0taumu

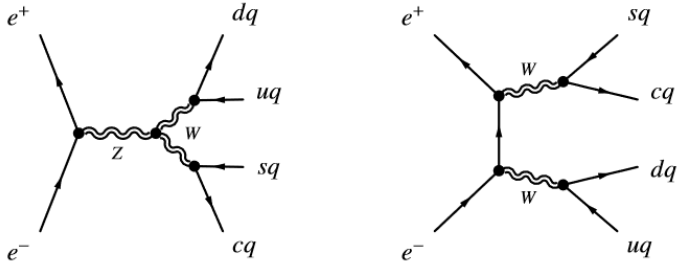


ww_h0cuxx



zz_l0mumu

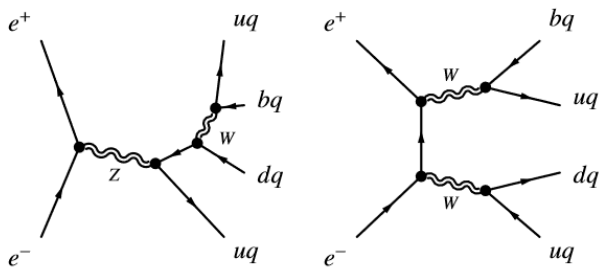




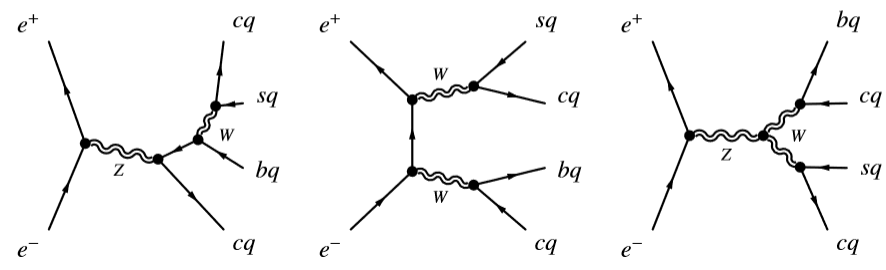
Backup: 4f background Feynman plot

4f background: ZZ category

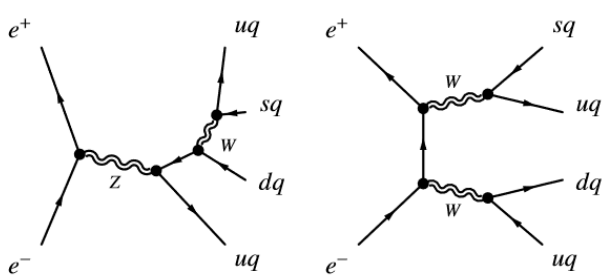
ww_h0uubd



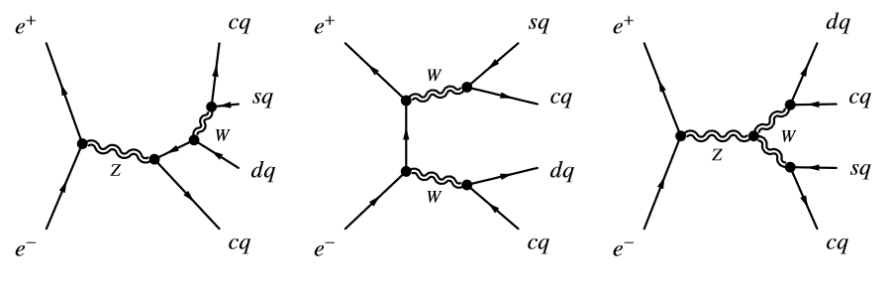
ww_h0ccbs



ww_h0uusd



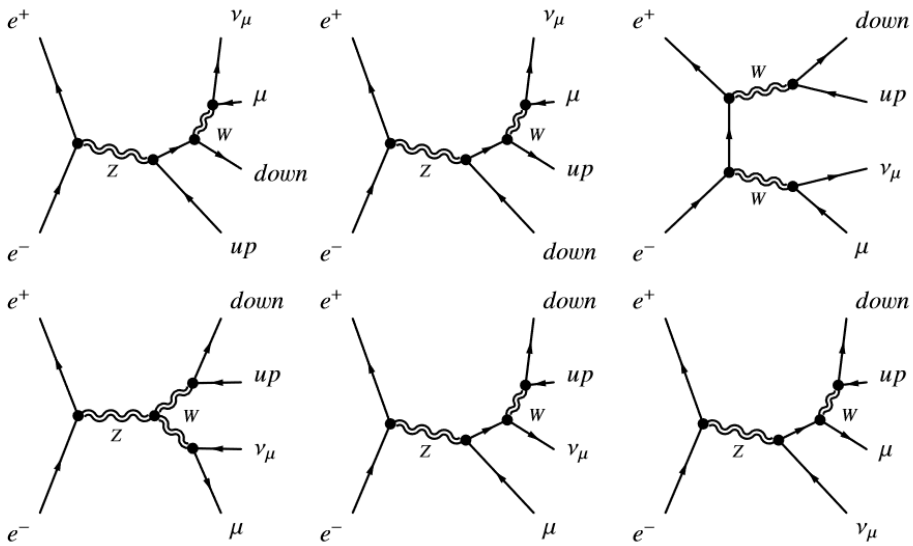
ww_h0ccds



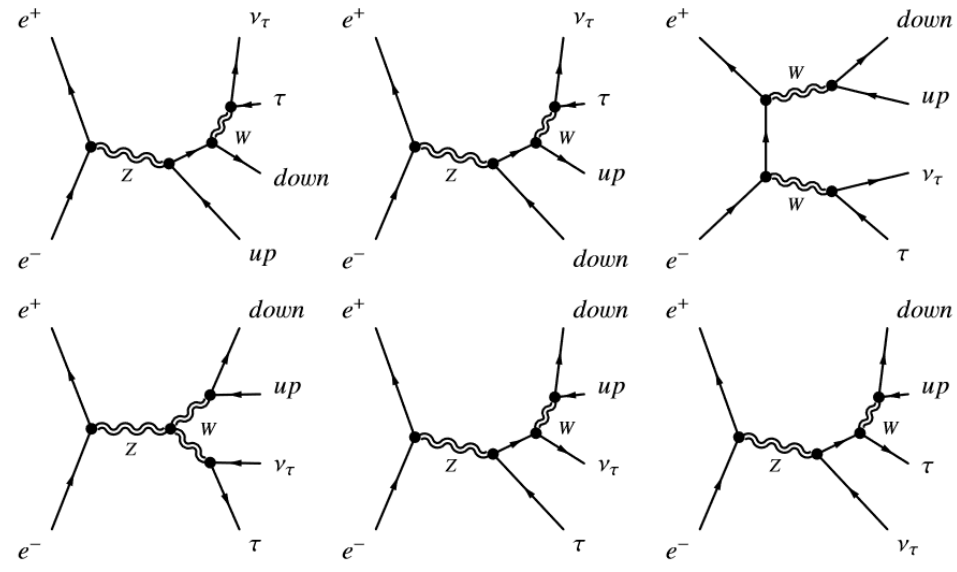
Backup: 4f background Feynman plot

4f background: ZZ category

ww_sl0muq



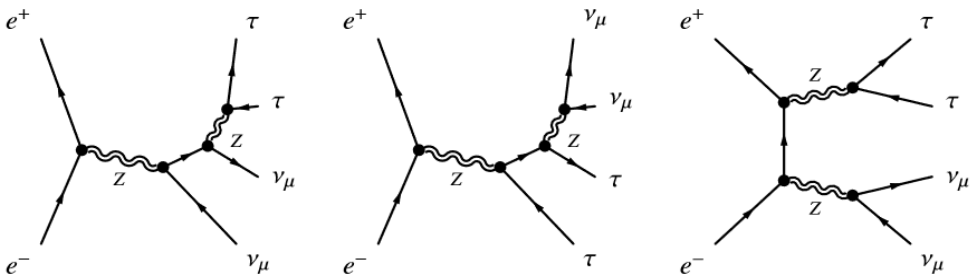
ww_sl0tauq



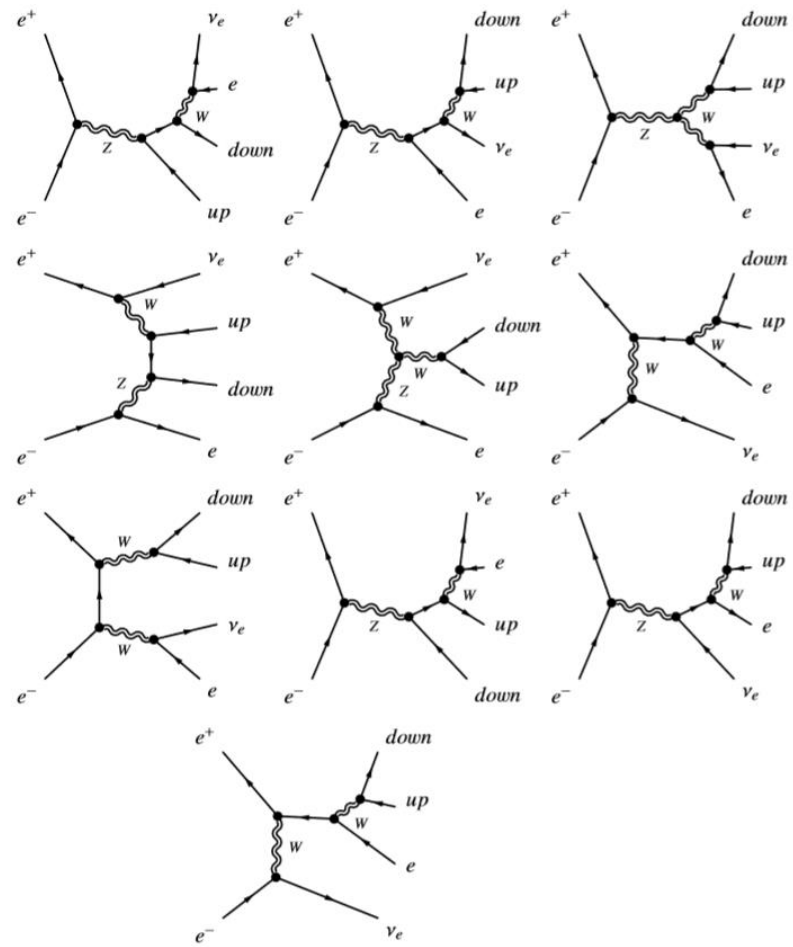
Backup: 4f background Feynman plot

4f background: ZZ category

zz_l0tautau



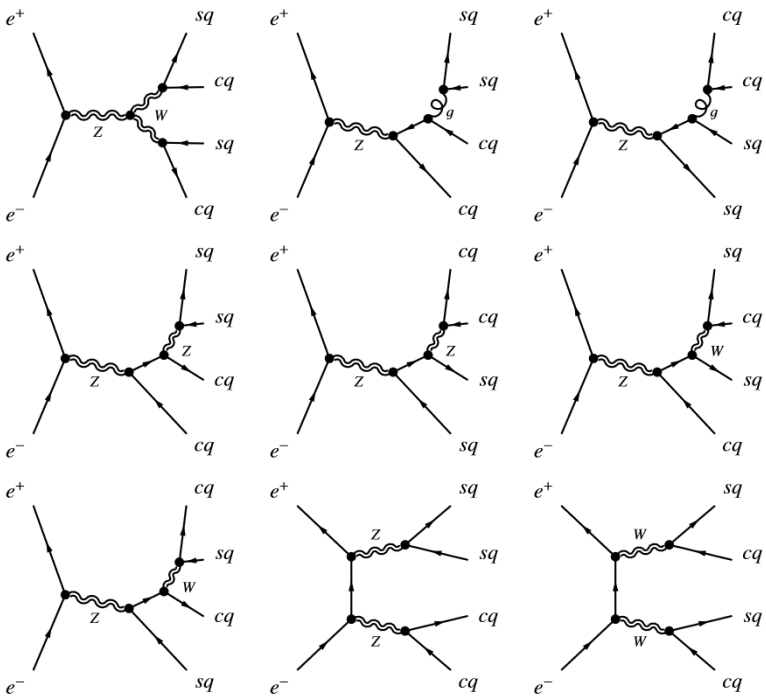
sw_sl0qq



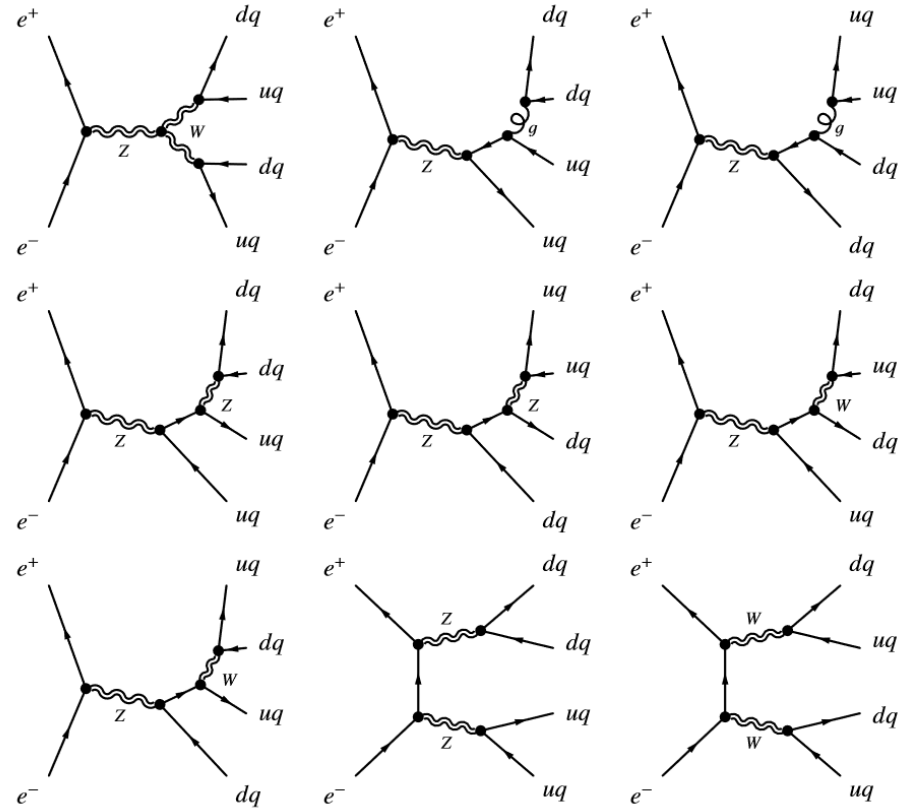
Backup: 4f background Feynman plot

4f background: ZZ category

zzorww_h0cscs



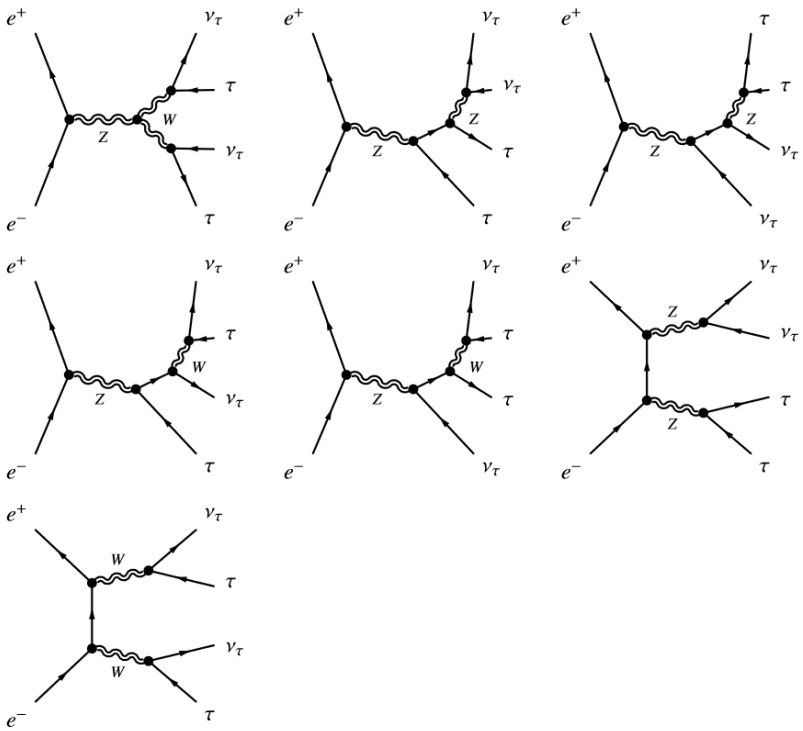
zzorww_h0ududd



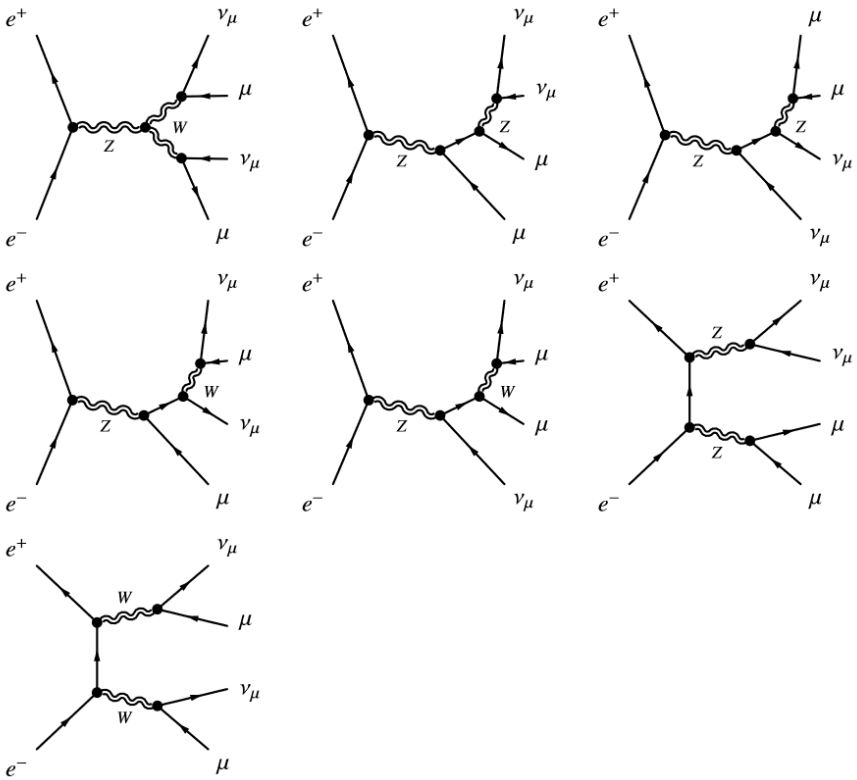
Backup: 4f background Feynman plot

4f background: ZZ category

zzorww_l0tautau



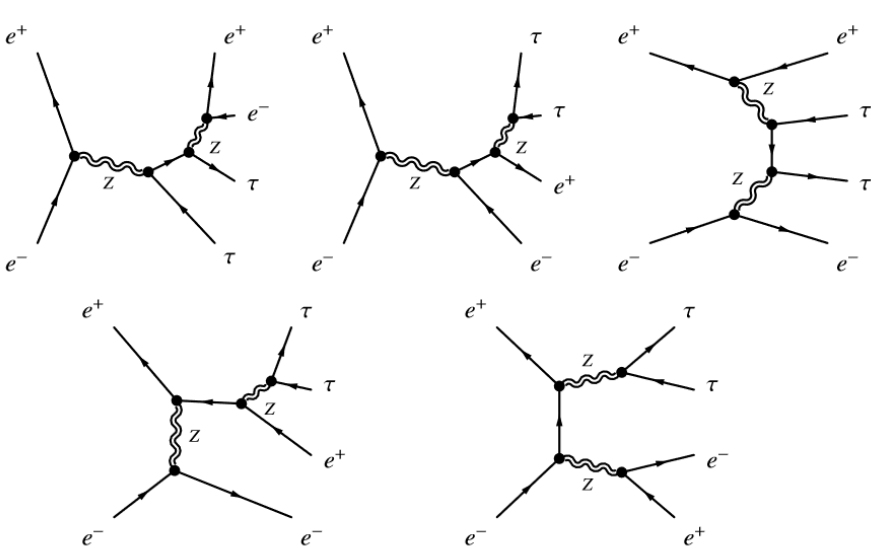
zzorww_l0mumu



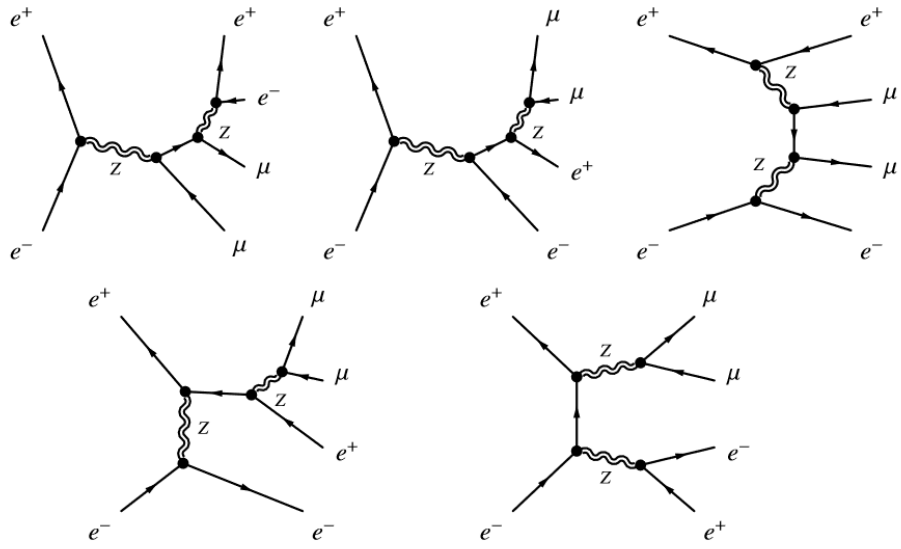
Backup: 4f background Feynman plot

4f background: ZZ category

sze_l0tau



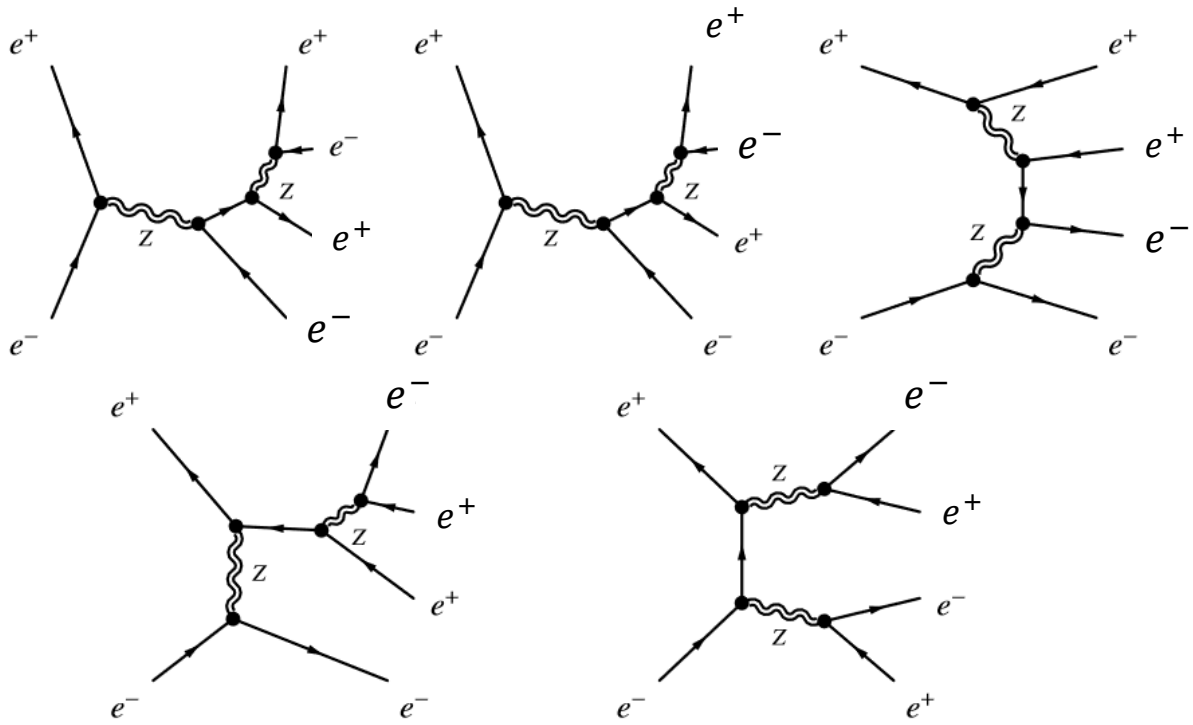
sze_l0mu



Backup: 4f background Feynman plot

4f background: ZZ category

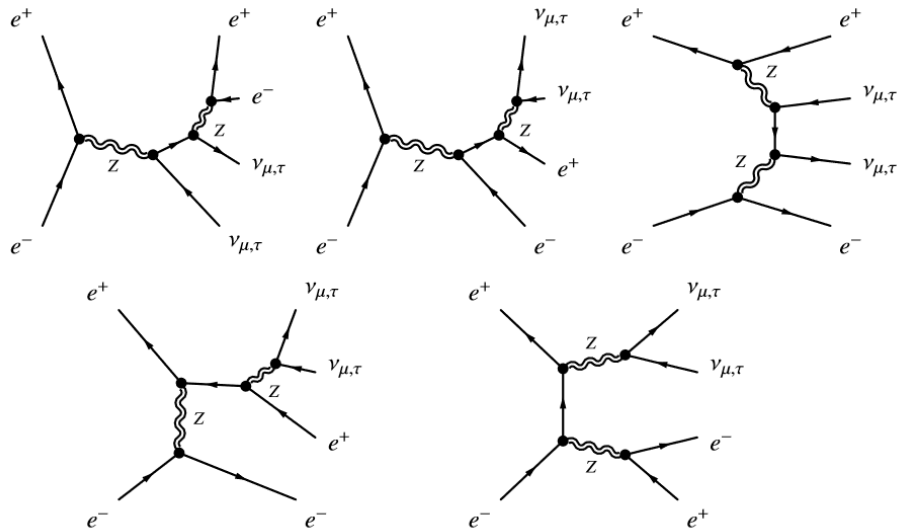
sze_l0e



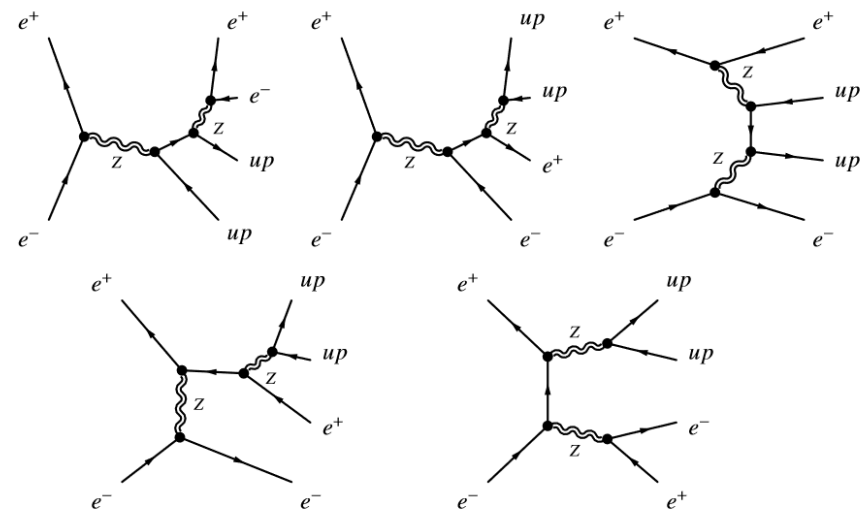
Backup: 4f background Feynman plot

4f background: ZZ category

sze_l0nunu



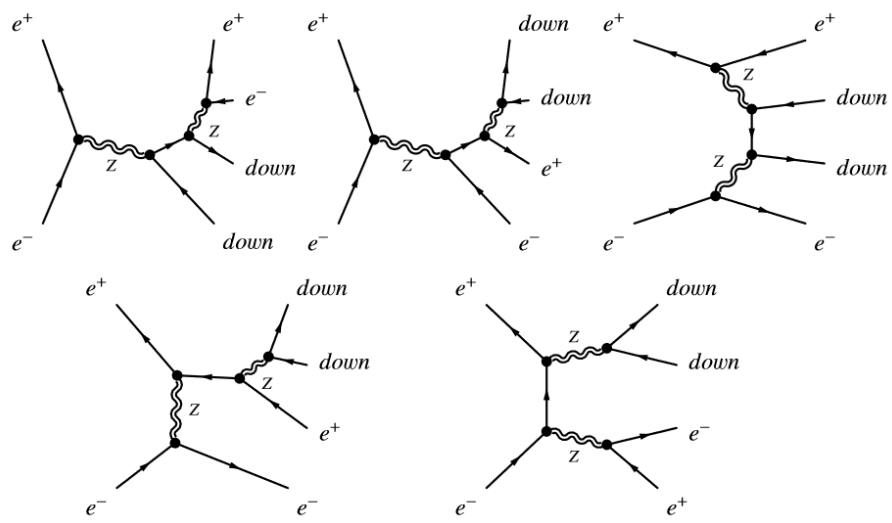
sze_sl0uu



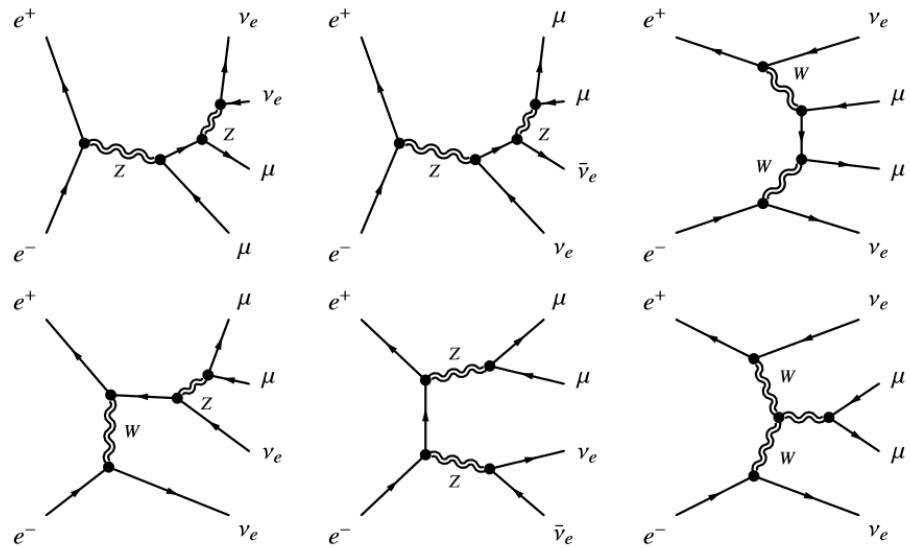
Backup: 4f background Feynman plot

4f background: ZZ category

sze_sl0dd



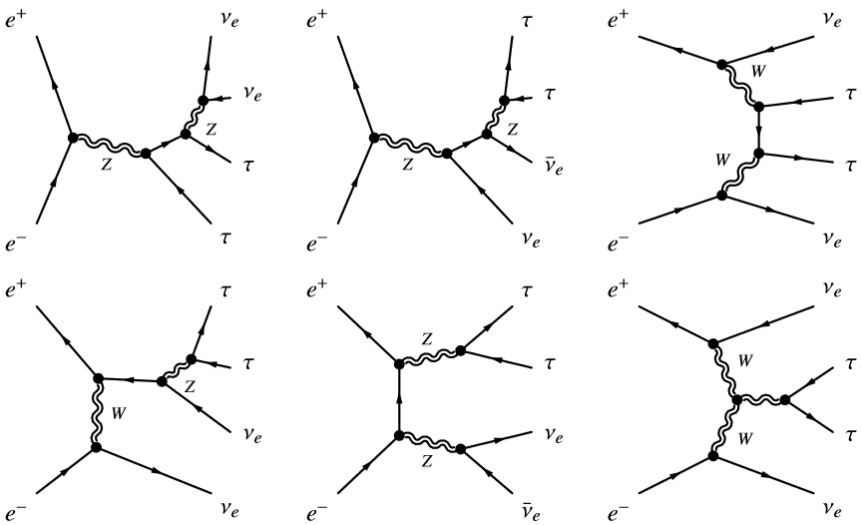
sznu_l0mumu



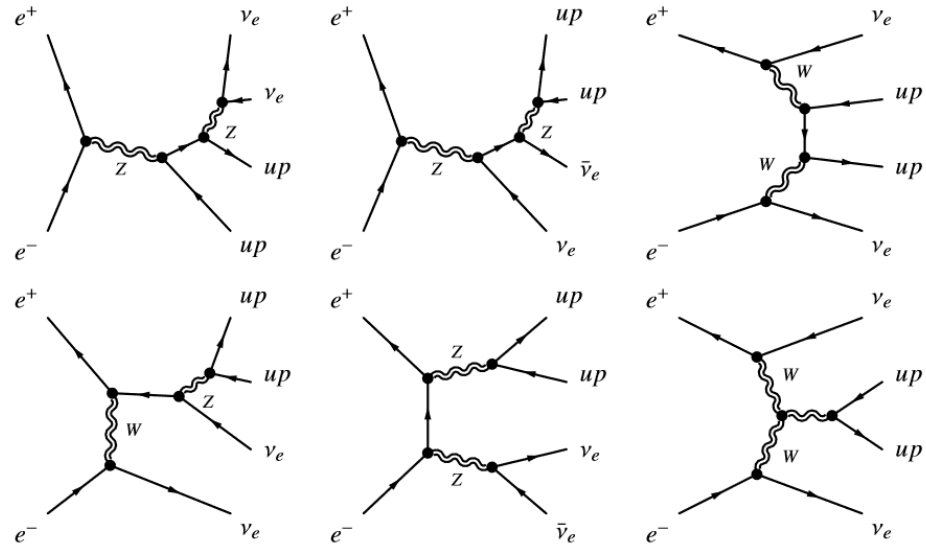
Backup: 4f background Feynman plot

4f background: ZZ category

sznu_l0tautau



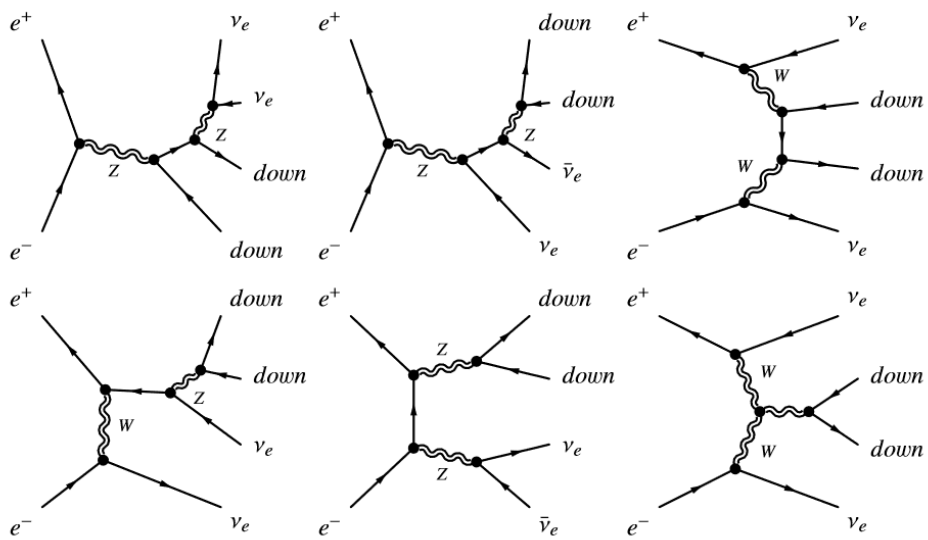
sznu_sl0nu_up



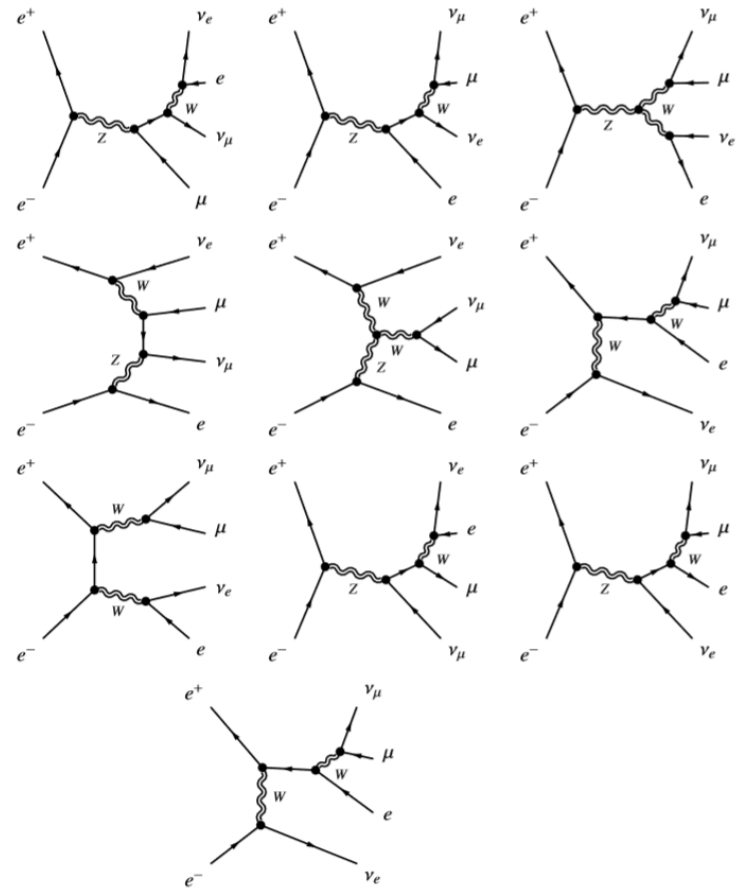
Backup: 4f background Feynman plot

4f background: ZZ category

sznu_sl0nu_down



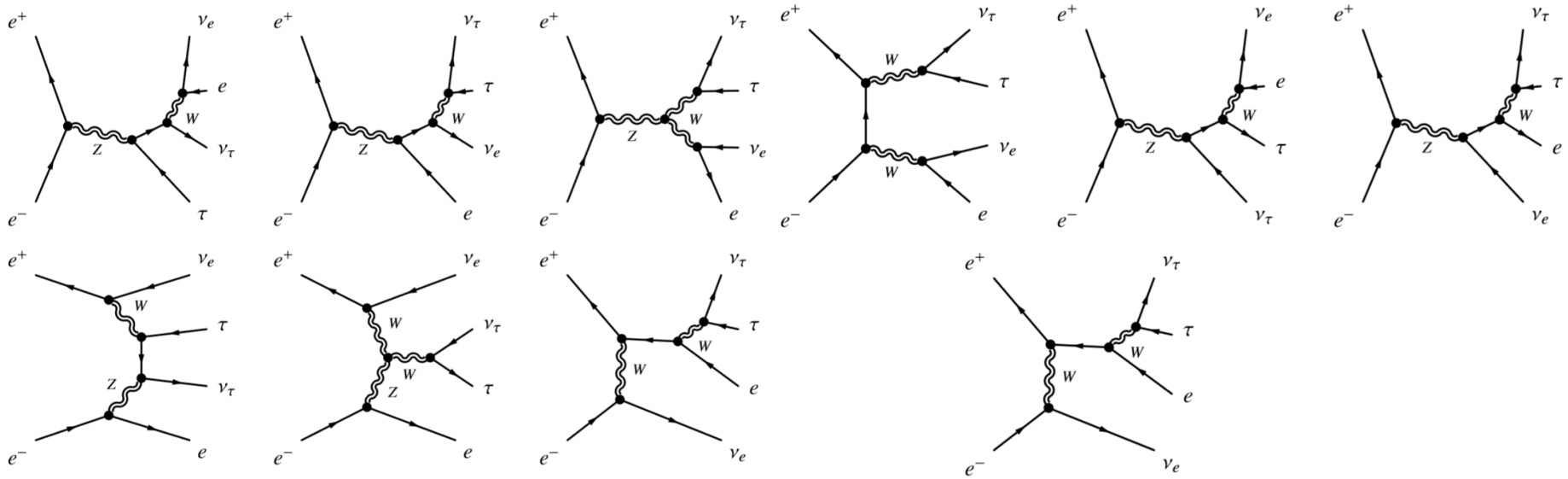
sw_l0mu



Backup: 4f background Feynman plot

4f background: ZZ category

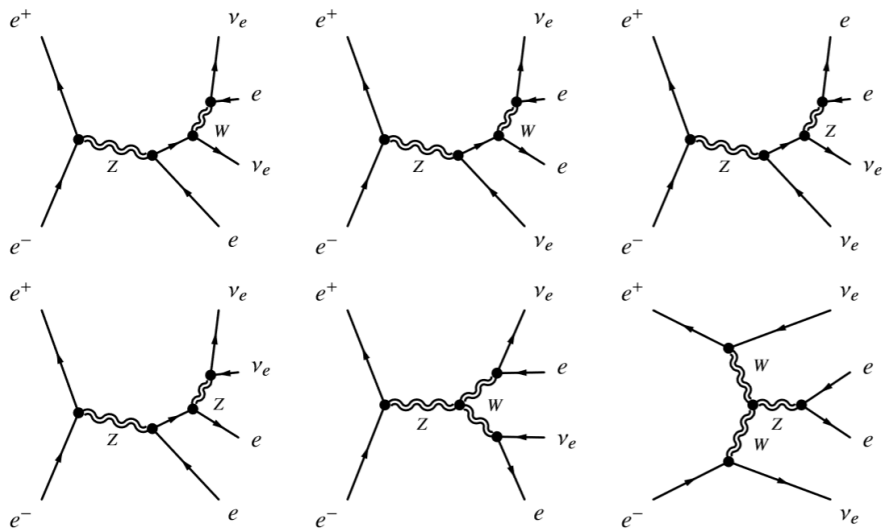
sw_l0tau

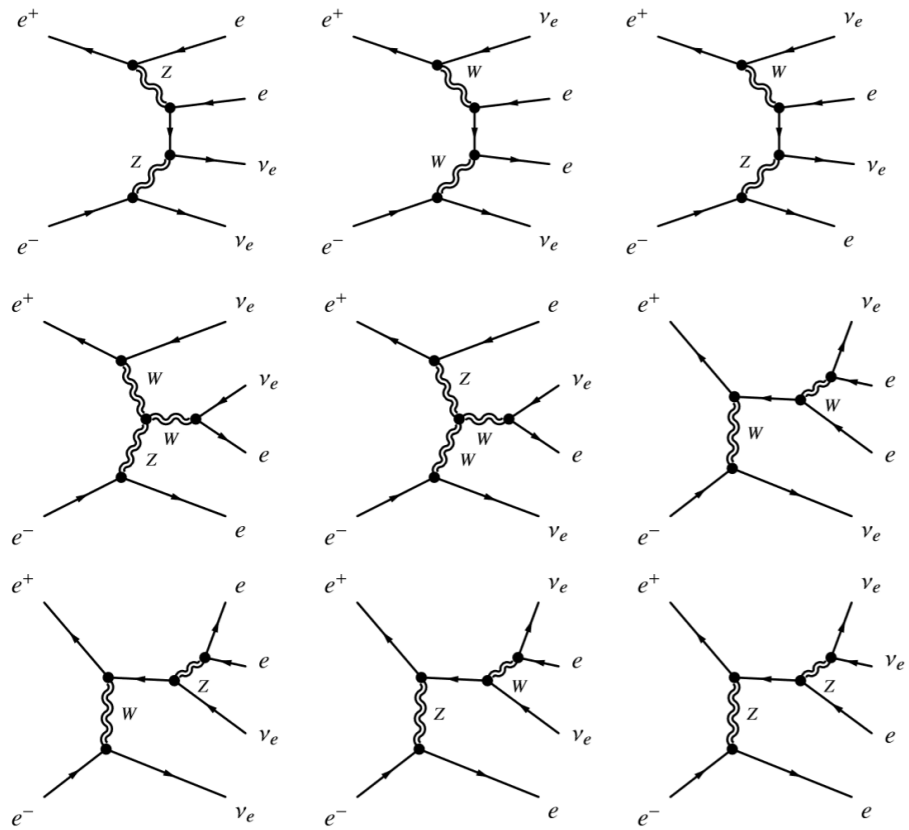


Backup: 4f background Feynman plot

4f background: ZZ category

szeorsw_l0l



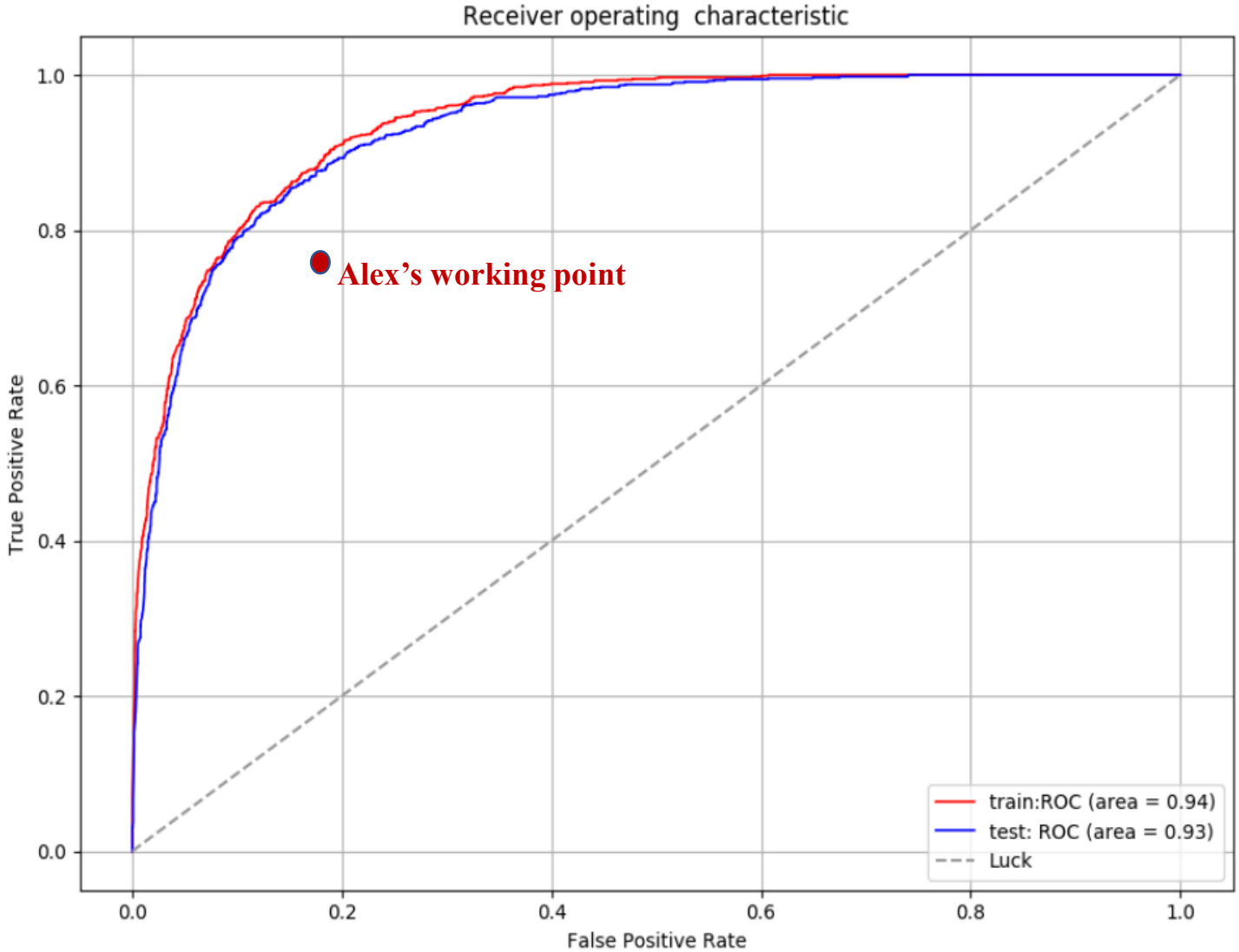


BDT Results (this page not updated)

Max_depth	Min_samples_leaf	Min_samples_split	Area under ROC (Train/test)	Difference
1	2	2	0.9451/0.9255	
2	2	2	0.9598/0.9341	
3	2	2	0.9729/0.9349	
4	2	2	0.9977/0.9342	
2	4	2	0.9607/0.9348	
2	5	2	0.9602/0.9343	
2	10	2	0.9614/0.9344	
2	2	5	0.9598/0.9341	
2	2	10	0.9598/0.9341	
2	2	20	0.9603/0.9349	
2	4	5	0.9607/0.9348	
2	5	20	0.9603/0.9343	
2	10	20	0.9614/0.9344	
1	5	10	0.9451/0.9255	
1	10	20	0.9451/0.9255	

BDT Results

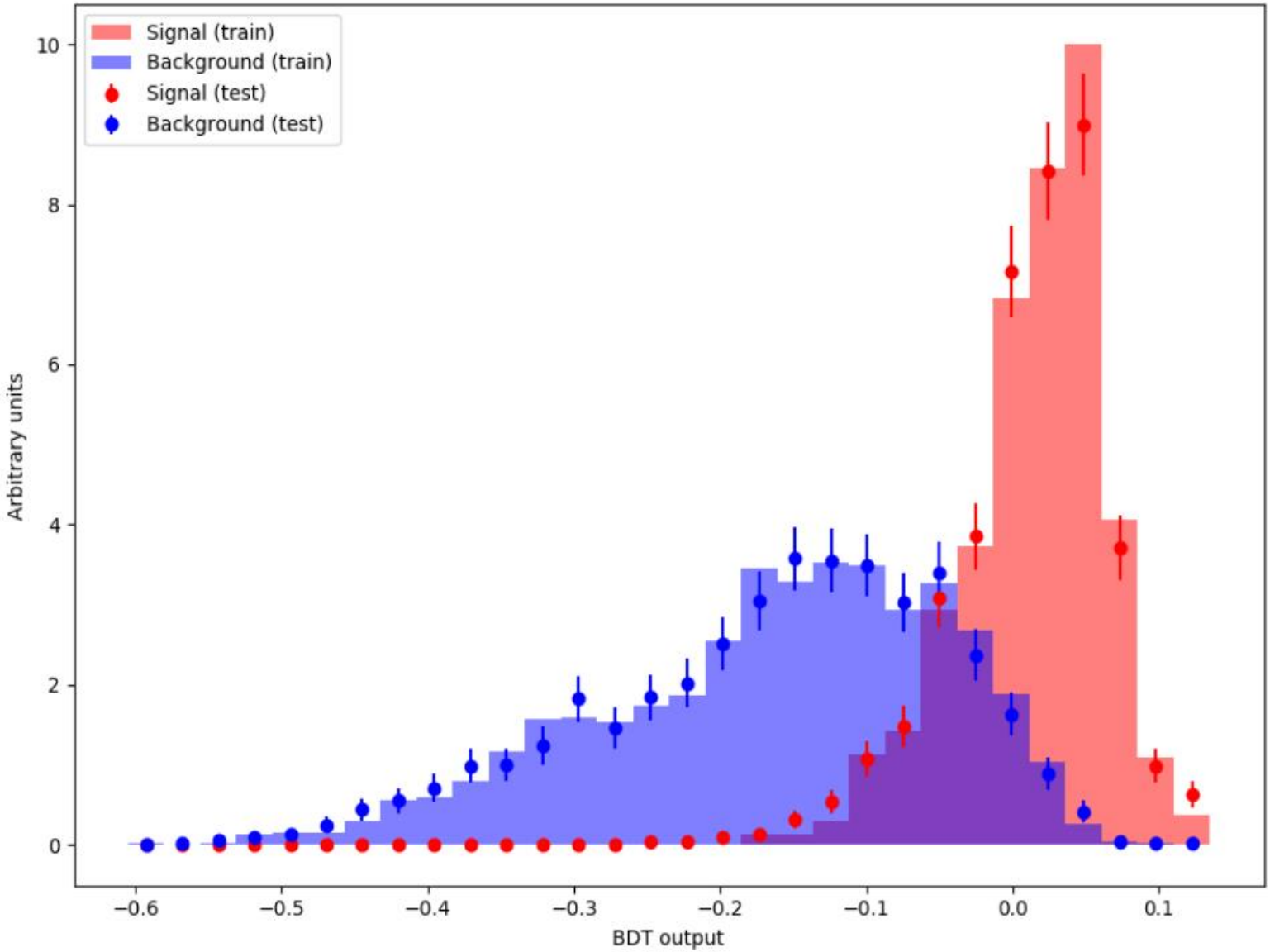
➢ Training Result — ROC Curve



Train sample and test sample are almost consistent in this trained model

BDT Results

➢ Training Result — Comparison



BDT Results

➢ Training Result — Comparison with Alex's Results

The features for BDT and their importances for BDT

Name	Value	Importance	Importance
Rreco34	$M_{recoil}(\mu^+\mu^-)$	0.33	0.20
TotalP	$p(\mu^+\mu^-jj)$	0.11	0.20
RMass1234	$M(\mu^+\mu^-jj)$	0.09	0.03
RMass12	$M(jj)$	0.08	0.04
LeadingEnJetEn	$E(j_{\text{leading}})$	0.07	0.01
SubLeadingEnJetEn	$E(j_{\text{subleading}})$	0.06	0.03
Anlj	Angle(di-muon, di-jet) in lab frame	0.06	0.09
VisMass	$M(\text{visible})$	0.06	
nPFOs	$n(\text{PFOs})$	0.05	0.28
RMass34	$M(\mu^+\mu^-)$	0.04	0.11
CosTheta	$\cos\theta$ of $\mu^+\mu^-jj$	0.02	0.00
Rreco1234	$M_{\text{recoil}}(\mu^+\mu^-jj)$	0.01	0.00
MissingMass	$M(\text{missing})$	0.01	

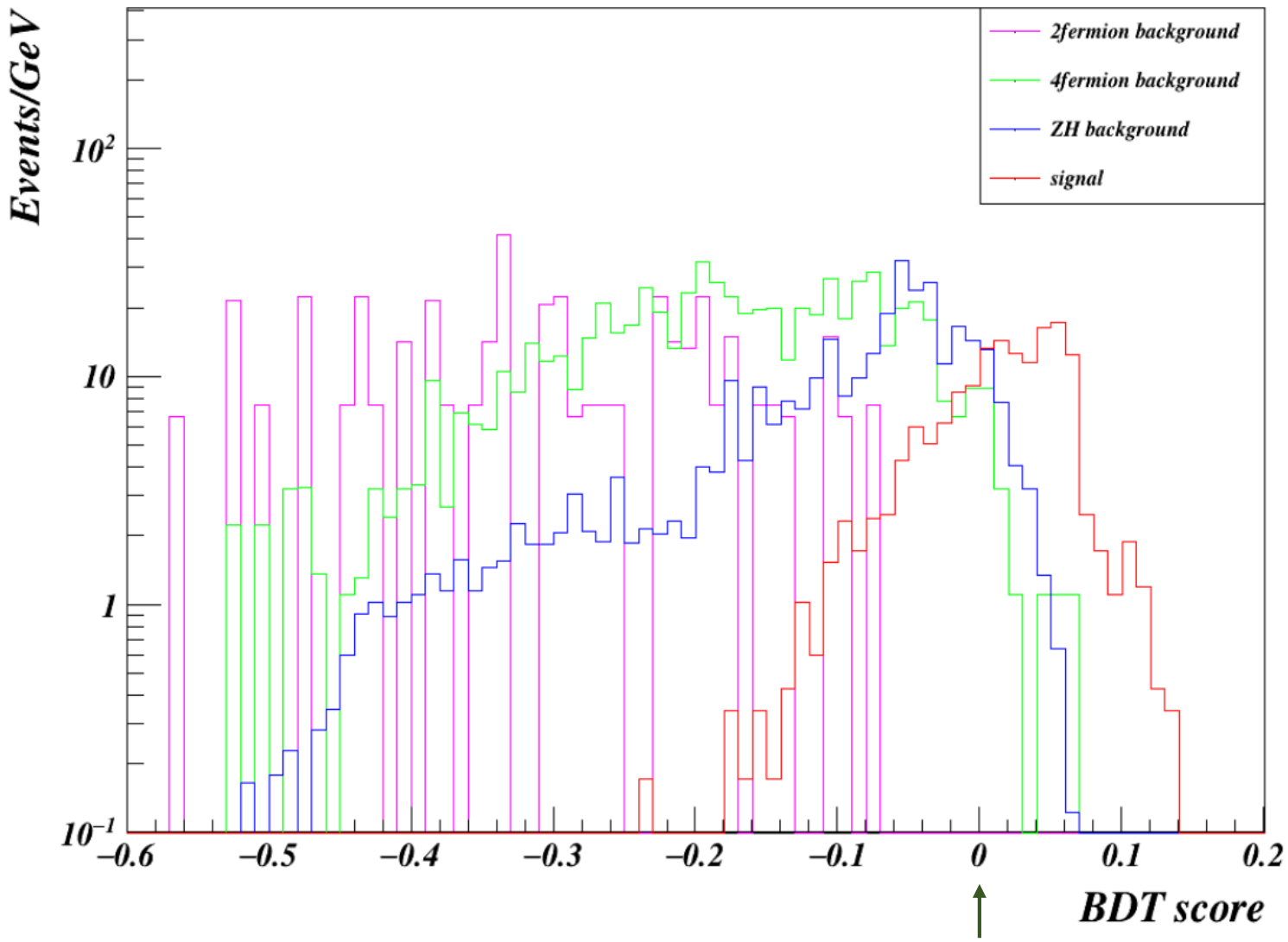
Equivalent in current code

Equivalent in current code

Alex's previous results **My results**

BDT Results

- **BDT Score**

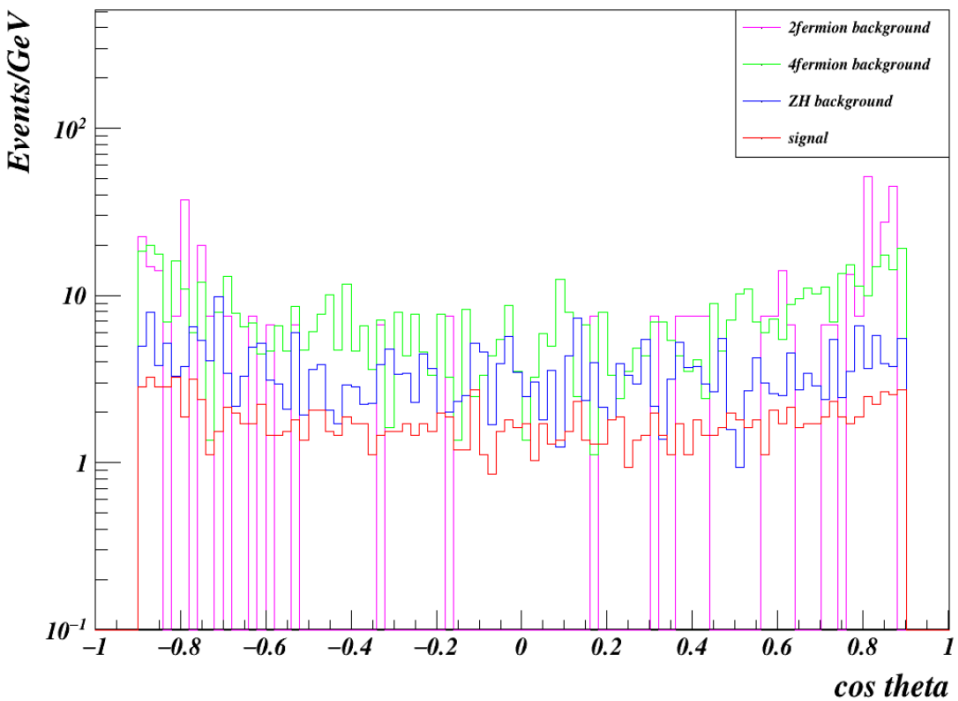


BDT score Cut

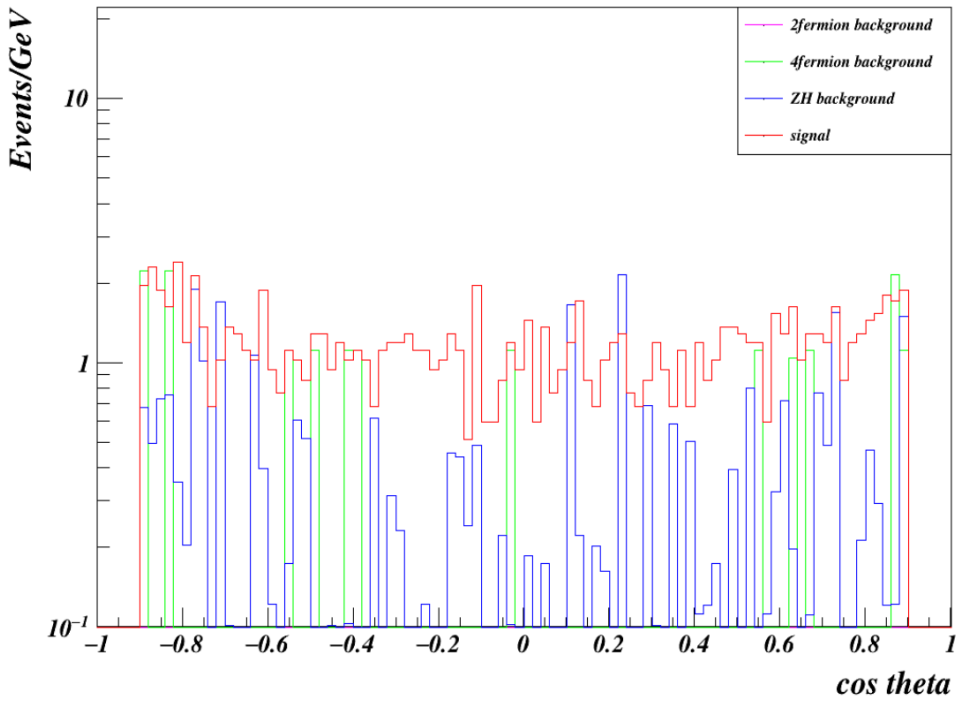
BDT Results

- **Cos theta**

Before BDT



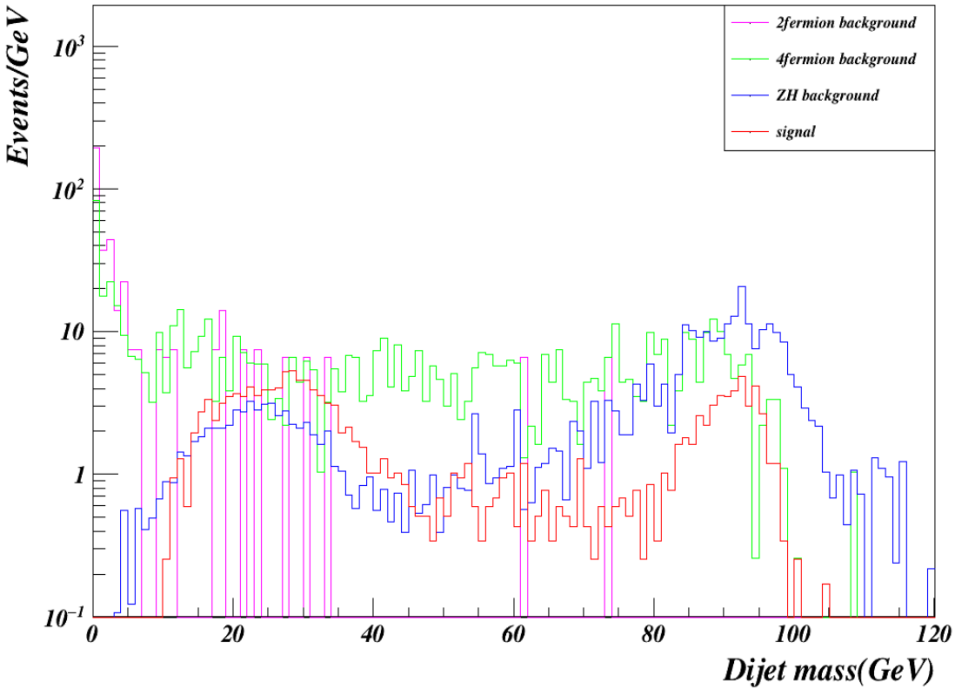
After BDT



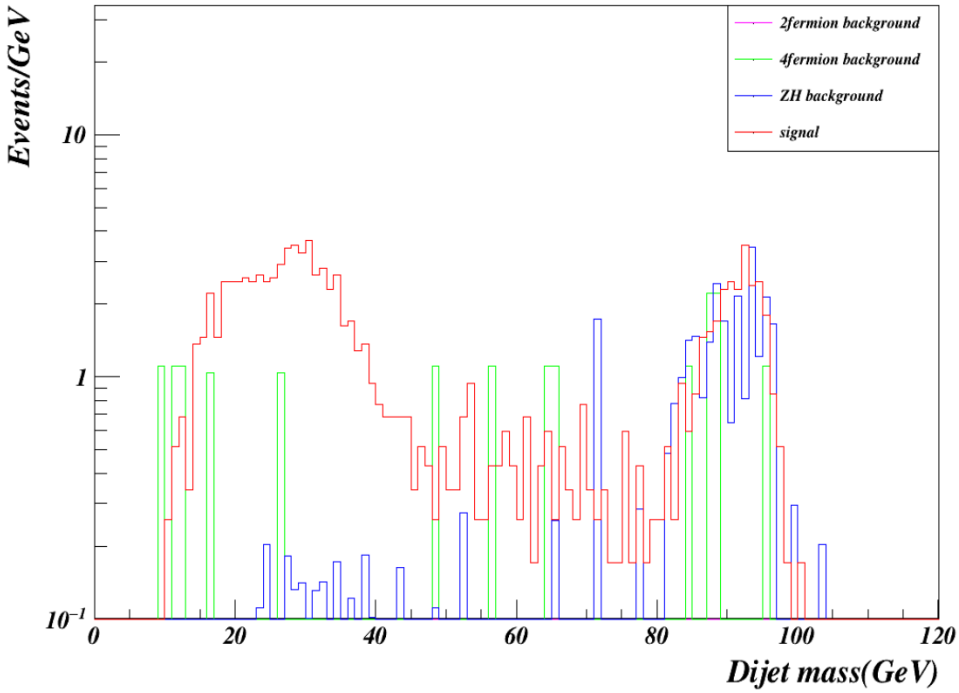
BDT Results

➢ Di-jet Mass

Before BDT



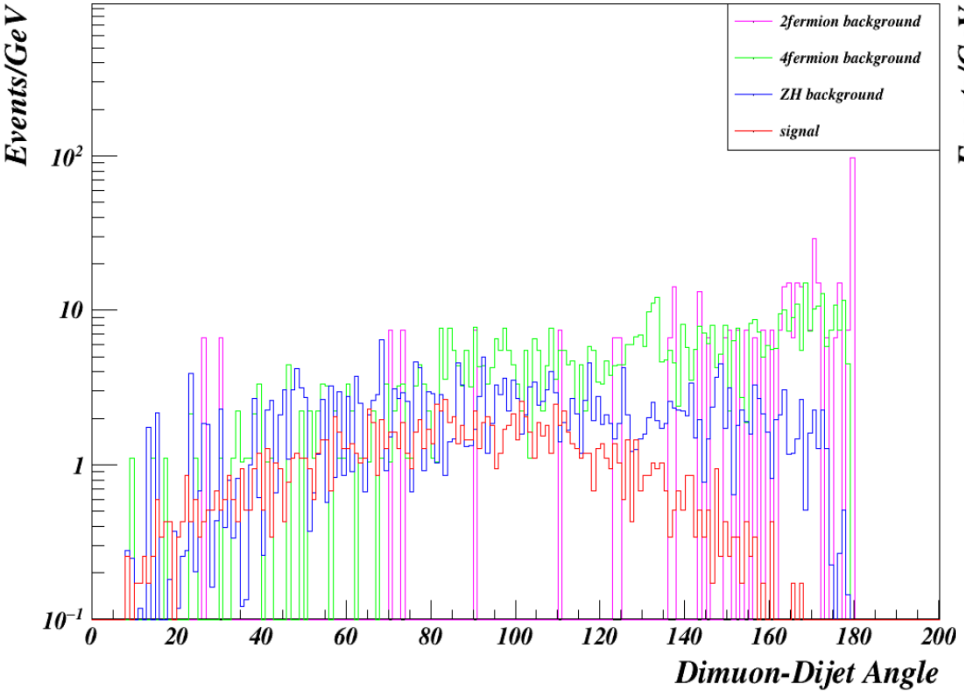
After BDT



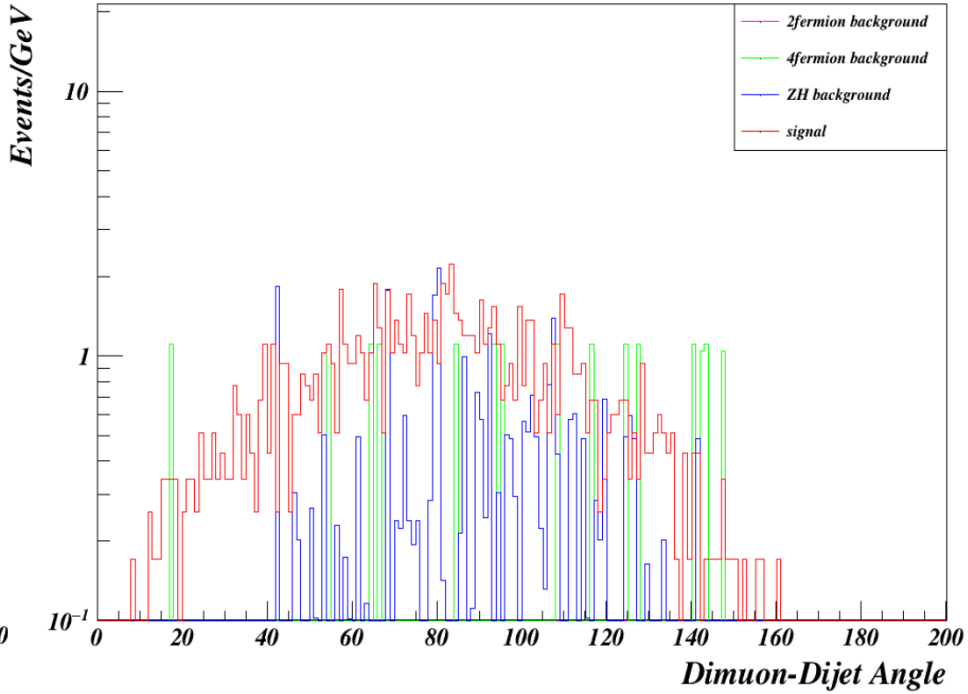
BDT Results

➢ Di-muon Di-jet Angle

Before BDT



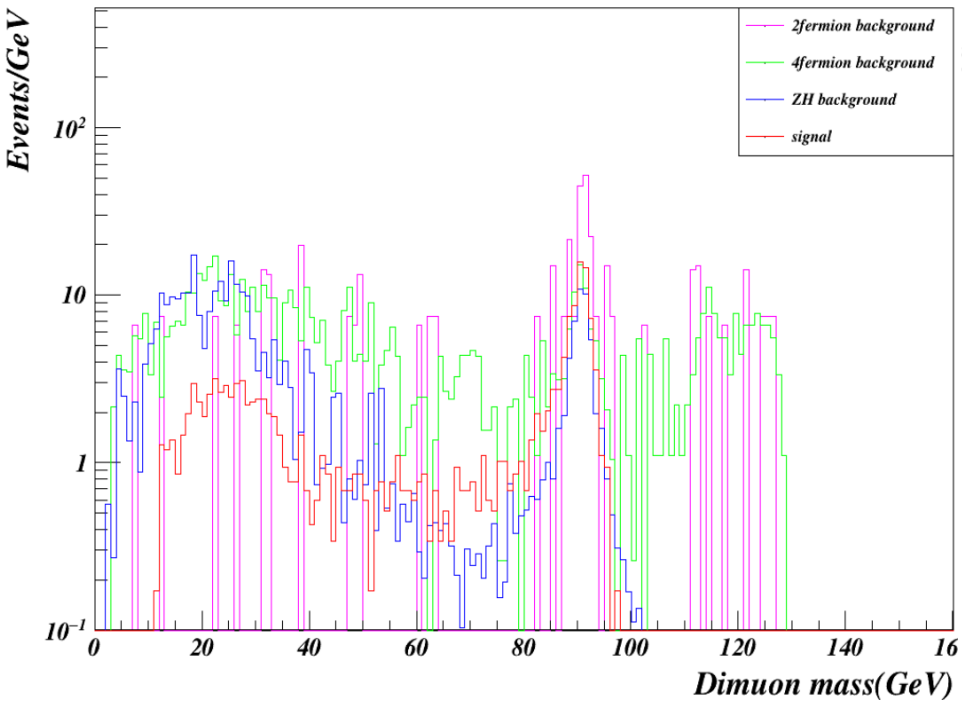
After BDT



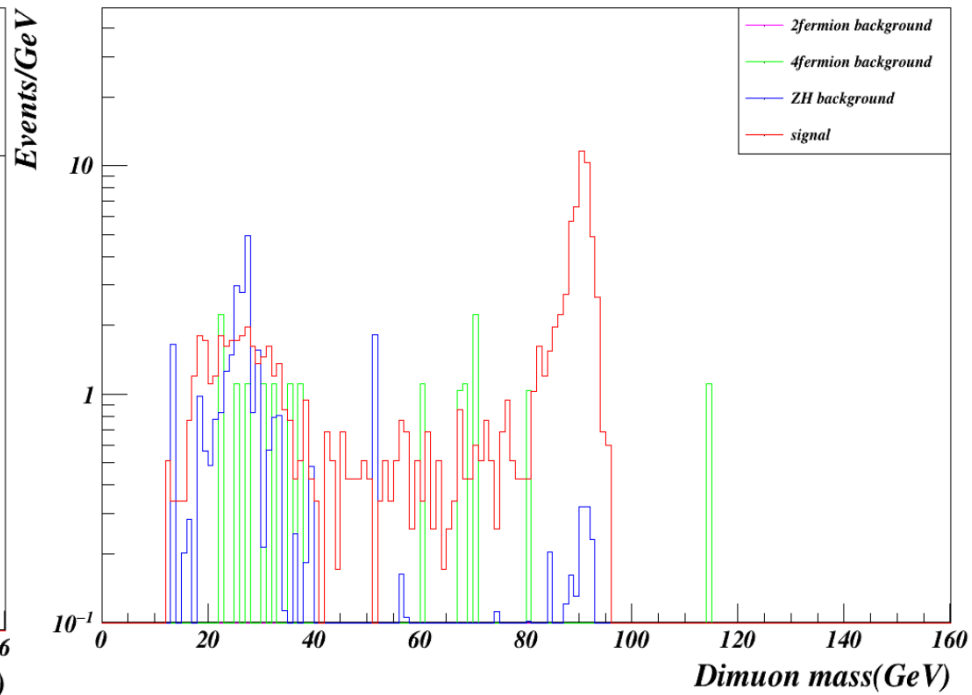
BDT Results

➢ Di-muon Mass

Before BDT



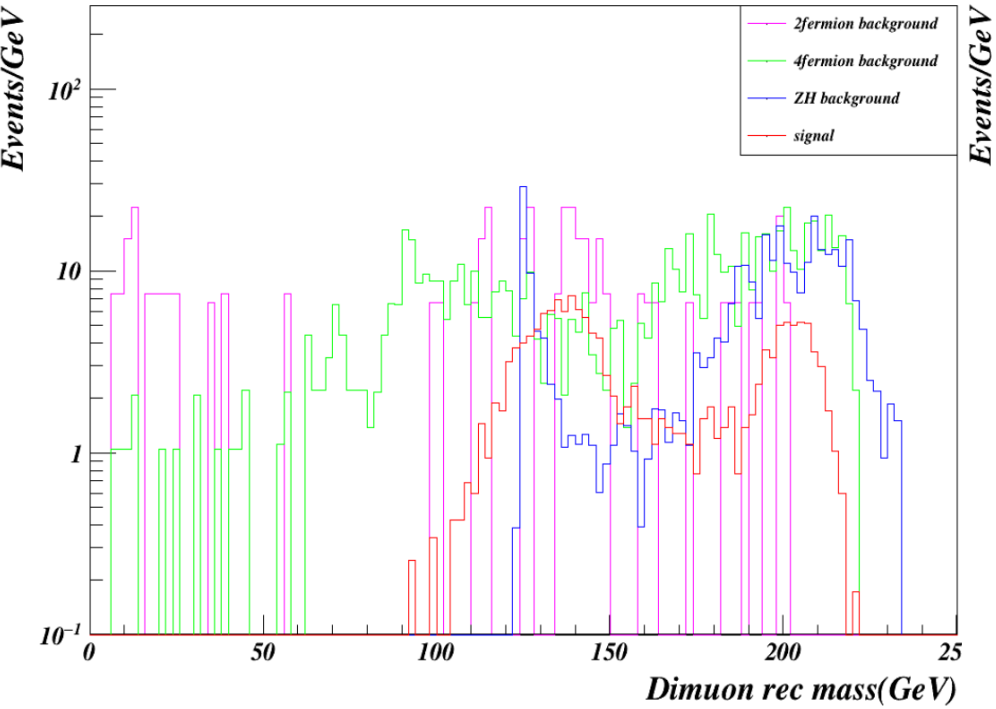
After BDT



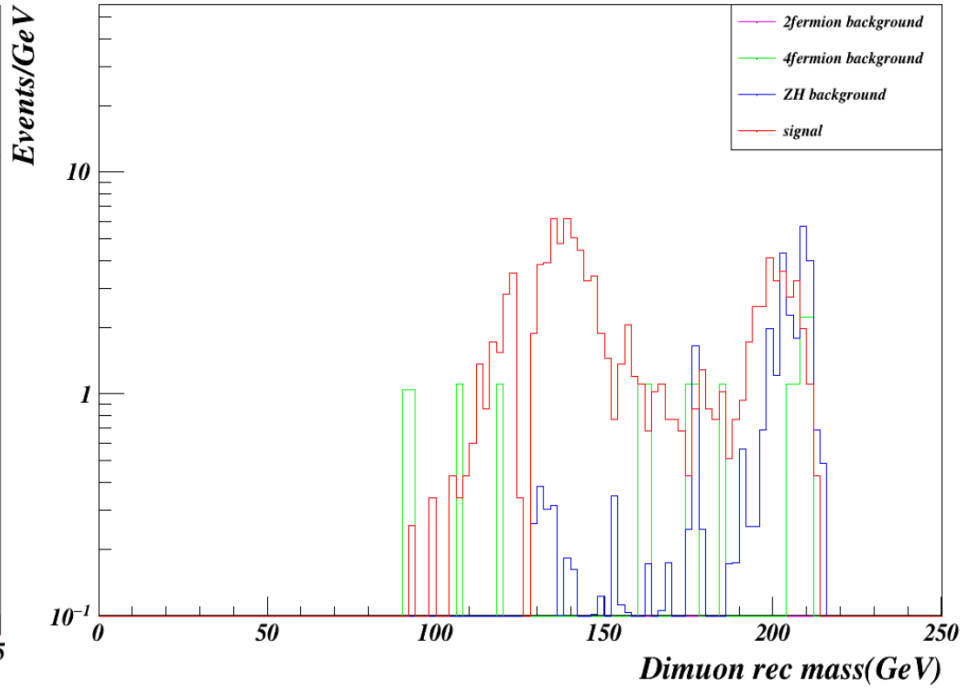
BDT Results

➢ Di-muon Recoil Mass

Before BDT



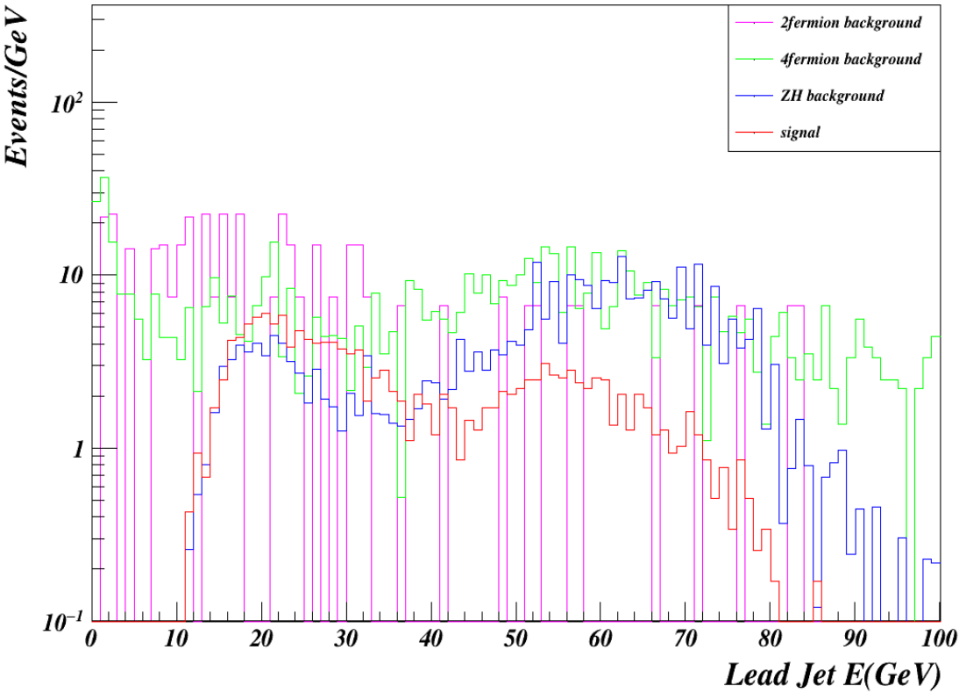
After BDT



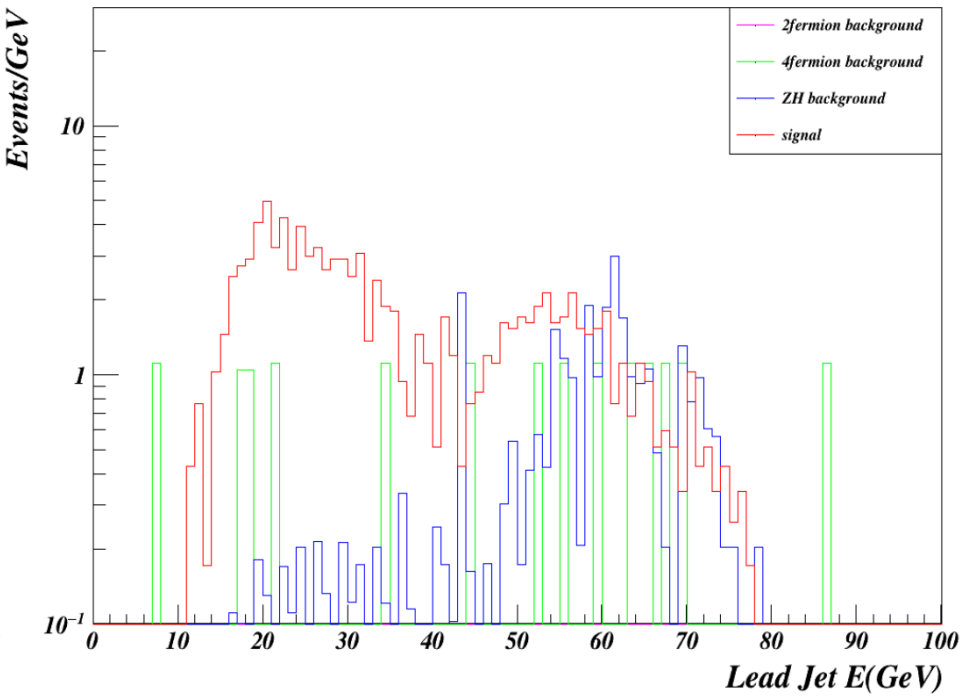
BDT Results

- **Lead Jet Energy**

Before BDT



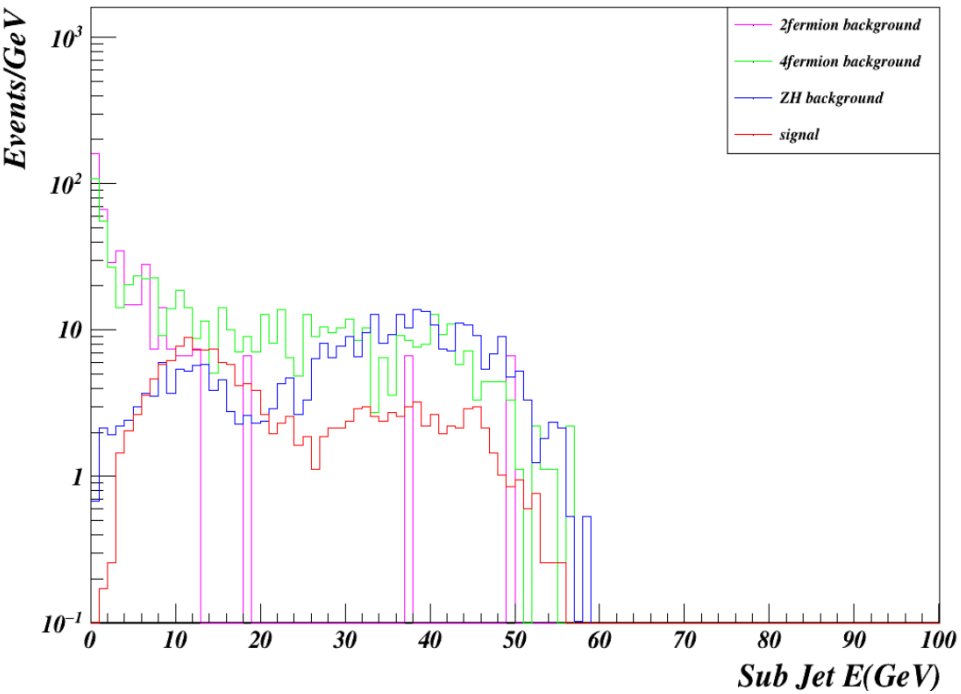
After BDT



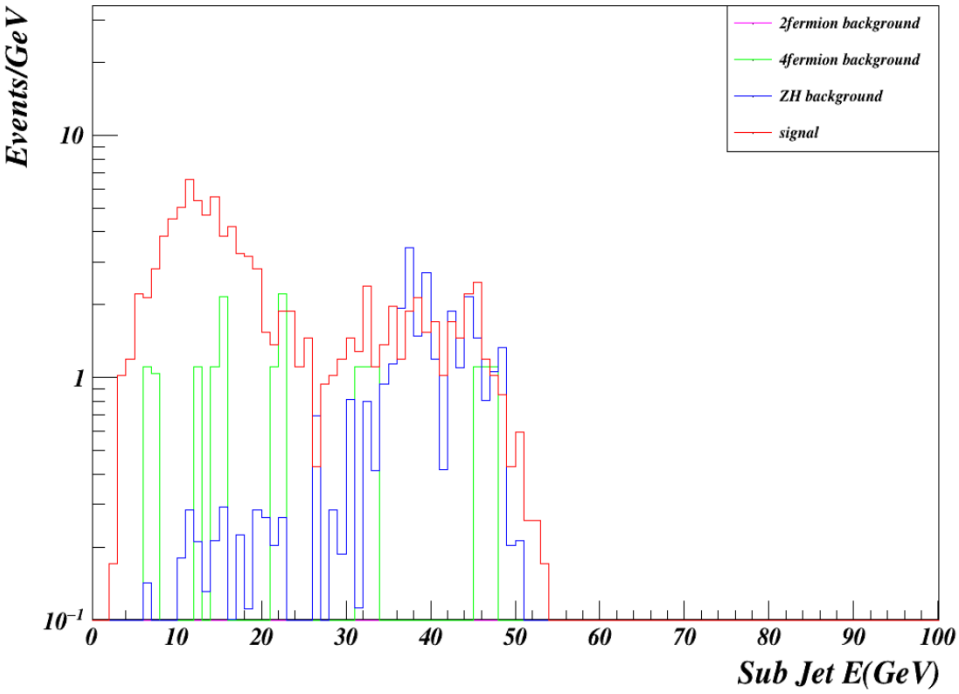
BDT Results

➢ Sub Jet Energy

Before BDT



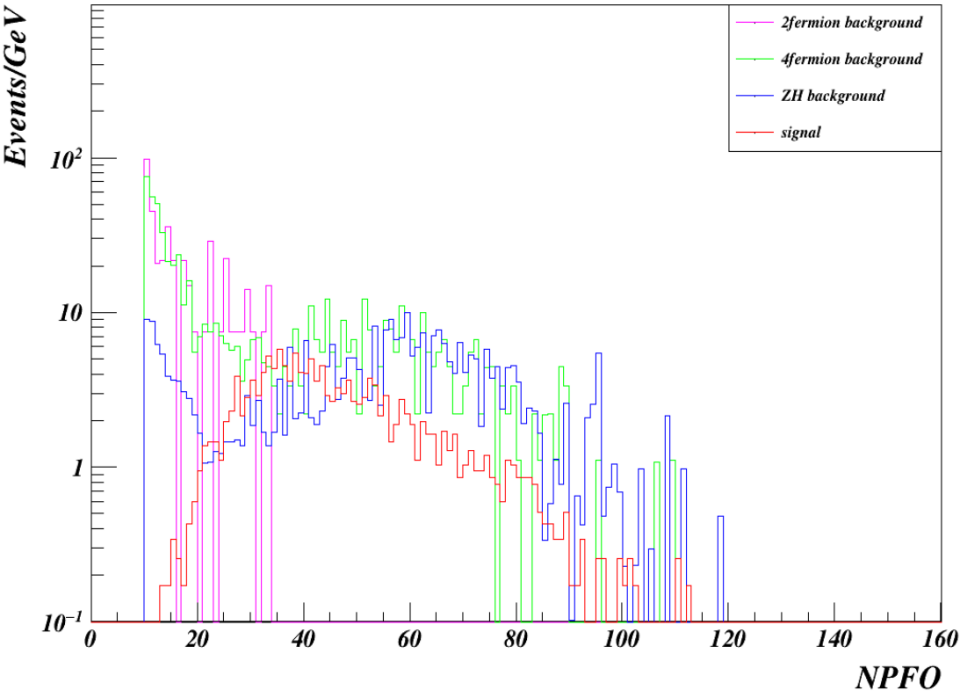
After BDT



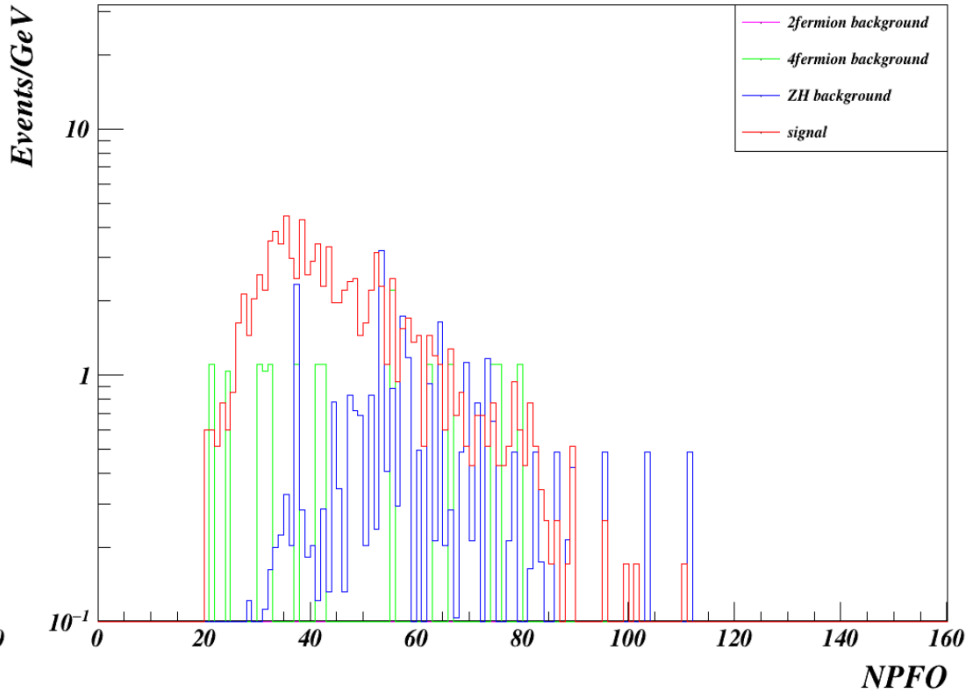
BDT Results

- **nPFO**

Before BDT



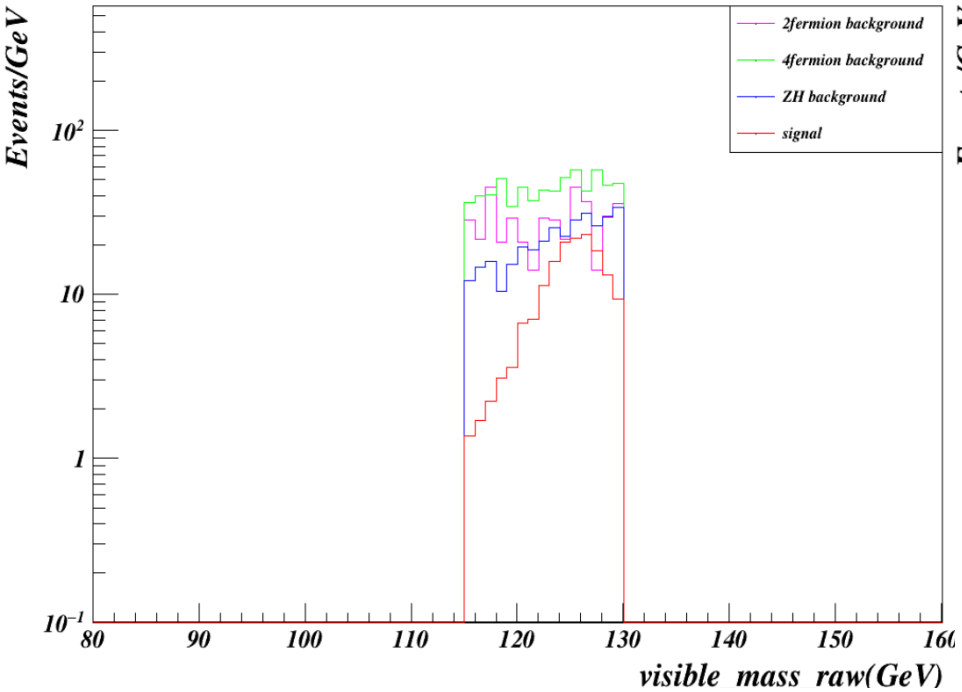
After BDT



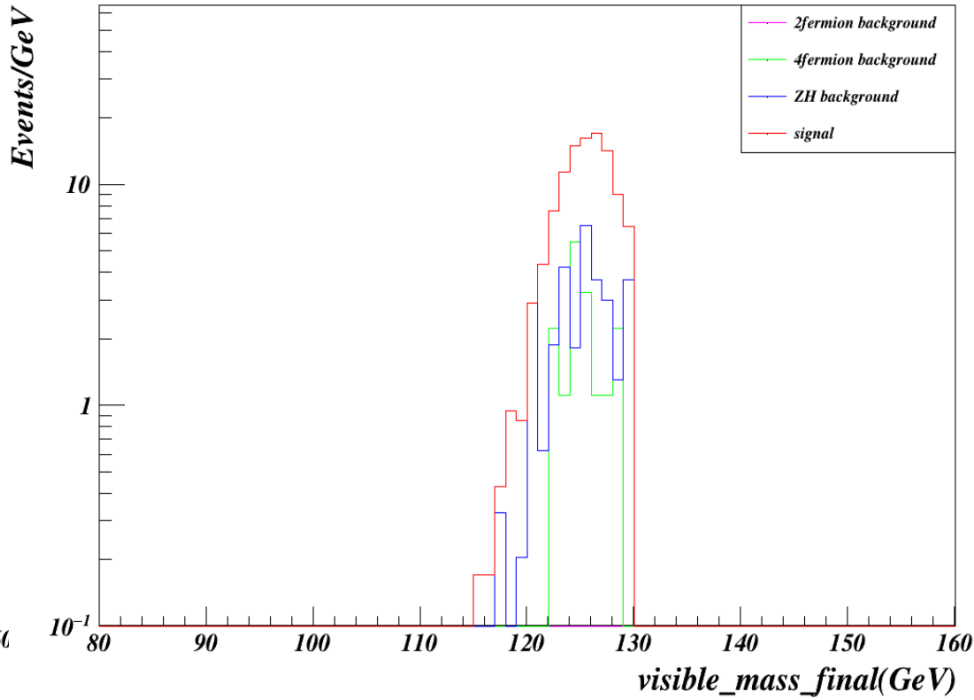
BDT Results

➢ Visible Mass

Before BDT



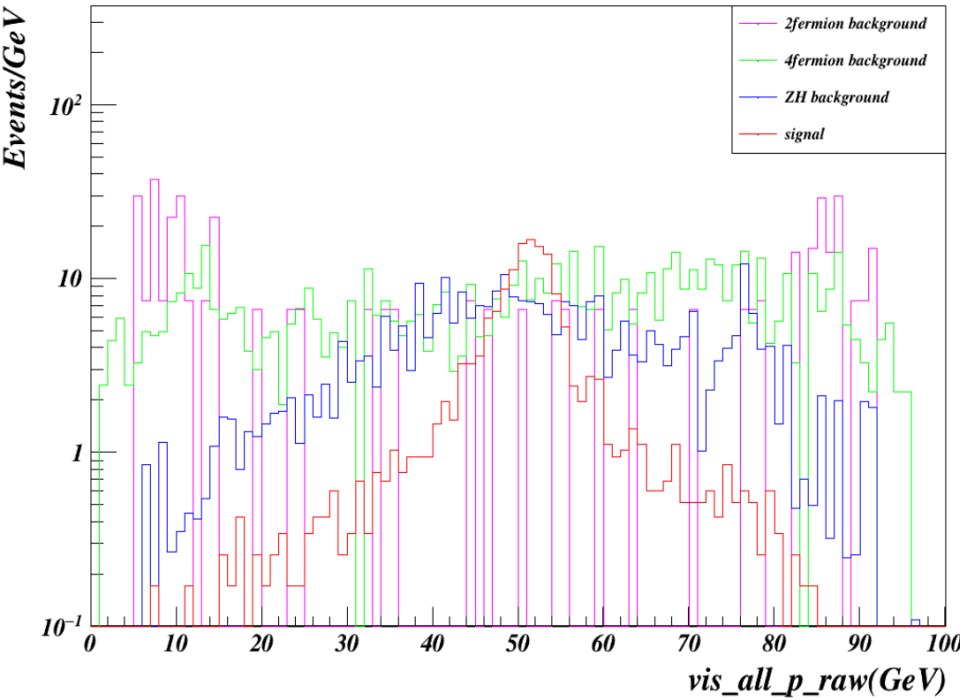
After BDT



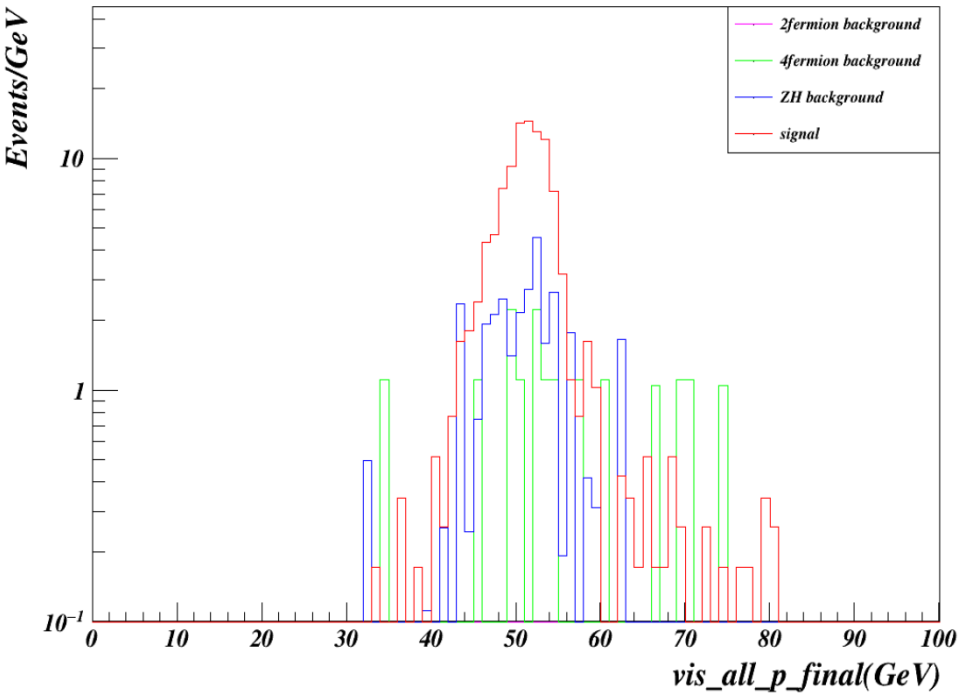
BDT Results

- **Visible P**

Before BDT



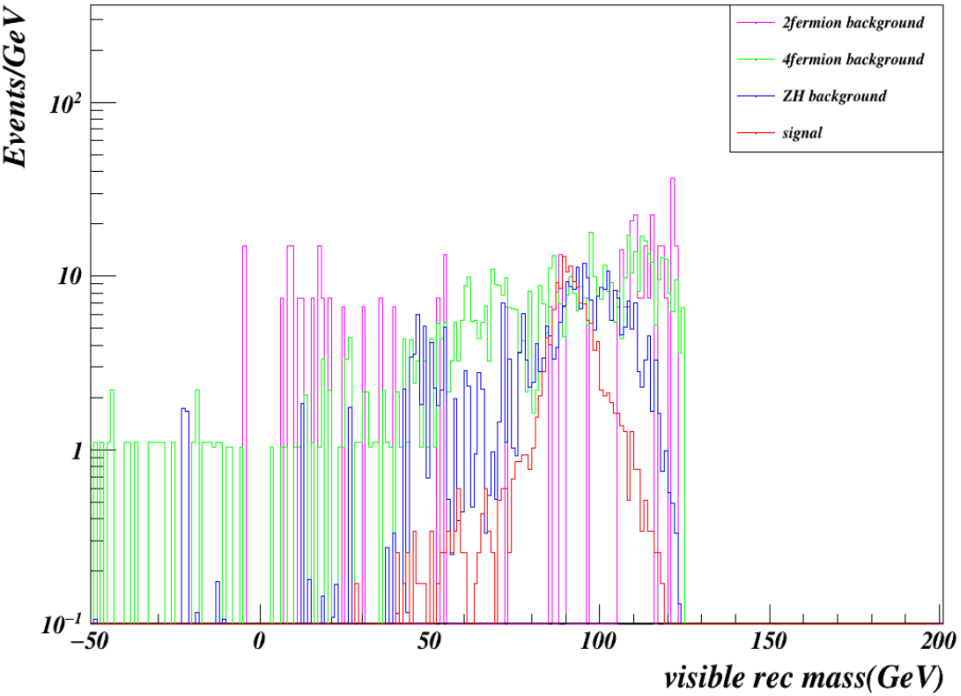
After BDT



BDT Results

- **Visible Recoil Mass**

Before BDT



After BDT