ILD-PHYS-PUB–2019-002
2019/May

ILD Benchmark Analysis: $h \to \mu^+\mu^-$ at 500 GeV

Shin-ichi Kawada*

* *DESY, Notkestraße 85, 22607 Hamburg, Germany*

Abstract

The process of $e^+e^- \to \nu\bar{\nu}h$ with $h \to \mu^+\mu^-$ at $\sqrt{s} = 500$ GeV at the ILC is investigated as one of the ILD physics benchmark, specifically for the transverse momentum resolution σ_{1/P_t}. We study the prospects of measuring the cross section times branching ratio $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ based on the running scenario for $\sqrt{s} = 500$ GeV, using fully-simulated MC samples. Two detector models, IDR-L and IDR-S, are considered in the analysis. The precision on $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ is evaluated to be $40.16 \pm 0.15\%$ with IDR-L and $41.28 \pm 0.15\%$ with IDR-S. The IDR-L gives relatively 2.8% better precision than IDR-S. This difference is caused by worse momentum resolution in barrel region with IDR-S, resulting more backgrounds in the peak region of $M_{\mu^+\mu^-}$ distribution. Specifically for this analysis, the IDR-L is better detector option.

1. Introduction

In this note, we will describe the analysis of the Higgs boson decays into muon pairs at the $\sqrt{s} = 500$ GeV ILC. This process is selected as one of the physics benchmark of ILD optimization [1], specifically for the transverse momentum resolution σ_{1/P_t} for high momentum particles.

The $h \to \mu^+\mu^-$ channel provides an opportunity to measure the Yukawa coupling between second-generation lepton and Higgs boson directly. We can also study the mass generation mechanism by looking the ratio between second-generation quark/lepton and second-/third-generation leptons. However, this analysis is very challenging because the branching ratio of $h \to \mu^+\mu^-$ is very small: 2.2×10^{-4} in the SM.

The process of $e^+e^- \to \nu\bar{\nu}h$ with $h \to \mu^+\mu^-$ at $\sqrt{s} = 500$ GeV is studied. Two beam polarization configurations are assumed; left-handed $\mathscr{P}(e^-, e^+) = (-80\%, +30\%)$, and right-handed $\mathscr{P}(e^-, e^+) = (+80\%, -30\%)$. In both configuration, an integrated luminosity of 1.6 ab^{-1} is assumed based on the running scenario at $\sqrt{s} = 500$ GeV [2, 3]. At $\sqrt{s} = 500$ GeV, there are two Higgs production processes; WW-fusion ($e^+e^- \to \nu\bar{\nu}h$) and Higgs-strahlung ($e^+e^- \to Zh$). Thus, the WW-fusion and Higgs-strahlung with $Z \to \nu\bar{\nu}$ are interfering each other. These two processes should be separated in the end, but such separation is not considered in this note because it is not the scope of this analysis.

As for the physics benchmark analysis, we consider two detector models, IDR-L and IDR-S. Since the detector geometry and magnetic field are different, the transverse momentum resolution is different between these two models, resulting different spectrum of muon pair invariant mass. We will also study such variables. In order to specify the combination of beam polarization and detector configuration, we use abbreviations as listed in Table 1.

Table 1: Abbreviations to specify the configurations.

	IDR-L	IDR-S
left-handed	IDR-L-left	IDR-S-left
right-handed	IDR-L-right	IDR-S-right

2. MC Samples

The event generation is performed with `Whizard1.95` [4] which is centrally produced for DBD [5], with considering ISR effect. The beamstrahlung effect is considered by `GuineaPig` [6]. `Tauola` [7–9] is used for τ decay, and `Pythia` [10] is used for decaying short-lived particles and hadronization. The full detector simulation based on `Geant4` [11] have been performed under `DD4HEP` [12] framework. Events are simulated with two detector configurations, the so-called `ILD_l5_o1_v02` option (which is IDR-L) and `ILD_s5_o1_v02` option (which is IDR-S). The detail description of IDR-L and IDR-S can be found in the IDR. The pile-up from $\gamma\gamma \to$ low P_t hadron backgrounds and e^+e^- seeable pairs due to beam-beam interaction have been generated based on the cross section model [13], and overlaid to all MC samples before the reconstruction. Events have been reconstructed using `PandoraPFA` [14] in the `Marlin` framework [15].

The processes of signal and background have been used in this analysis listed in Table 2. In the Higgs events ($e^+e^- \to f\bar{f}h$, where f denotes a fermion), only $\nu\bar{\nu}h$ with $h \to \mu^+\mu^-$ are considered as the signal process. Other Higgs decay processes and/or $e^+e^- \to q\bar{q}h/\ell^+\ell^-h$ (where q denotes a quark, and ℓ denotes e, μ, or τ) processes are considered as the background. For the SM background, we have included all backgrounds in e^+e^- → 2-/4-fermion processes and $\gamma\gamma$ → 4-fermion processes with at least one lepton ℓ in the final state. The processes which only contain jets are not considered because the signature of such backgrounds are dramatically different from signal, and expected to be easily suppressed. The processes of $\gamma\gamma \to$ 2-fermion are not included because there are no fully-simulated

MC samples due to its huge cross section. All MC events are luminosity-weighted to adjust to the target luminosity of 1.6 ab^{-1}.

Table 2: List of processes used in this analysis.

type	process name
Higgs	`ffh_mumu`
	`higgs_ffh`
2f	`2f_Z_bhabhag`
	`2f_Z_leptonic`
4f	`4f_singleW_leptonic`
	`4f_singleW_semileptonic`
	`4f_singleZee_leptonic`
	`4f_singleZee_semileptonic`
	`4f_singleZnunu_leptonic`
	`4f_singleZnunu_semileptonic`
	`4f_singleZsingleWMix_leptonic`
	`4f_WW_leptonic`
	`4f_WW_semileptonic`
	`4f_ZZ_leptonic`
	`4f_ZZ_semileptonic`
	`4f_ZZWWMix_leptonic`
$\gamma\gamma \to$ 4f	`aa_4f`

3. Analysis

We basically use `ILCSoft` [16] with the version of v02-00-02 for the analysis. First, a pair of muons is selected as the $h \to \mu^+\mu^-$ candidate. These muons are subjected to the fitting to estimate the primary vertex position. Then, a series of selection cuts is applied as the preselection to select signal-like events and reject backgrounds. To increase sensitivity, a multivariate analysis technique is applied. Finally, a toy MC technique is applied to extract the precision on the cross section times branching ratio $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$.

3.1. Isolated Lepton Tagging

We use `IsolatedLeptonTagging` processor [17, 18] for selecting $h \to \mu^+\mu^-$ candidate. We use updated version of `IsolatedLeptonTagging`; not using impact parameter information and energy deposit in yoke, because the MC samples described in Section 2 are now have smearing effect at the IP [19], and there is a potential problem when we associate the hit information in yoke [18]. We use the parameters summarized in Table 3 for `IsolatedLeptonTagging`, where E_{CAL} is the energy deposit in the calorimeter system, and p is the track momentum. A multivariate double cone method is used to check the signature of muon of a particle, and a cut on MVA output is applied. These parameters are the same in all analysis channels. The events which have exactly one μ^+ and one μ^- are used for further analysis.

Table 3: Parameters for isolated lepton tagging. Definition of variables are in the text.

variable	condition
E_{CAL}/p	< 0.5
p	> 10 GeV
MVA cut	> 0.8

We define the reconstruction efficiency in the signal process as follows.

$$\text{efficiency} \equiv \frac{\text{number of events which exactly reconstruct one } \mu^+ \text{and one } \mu^-}{\text{number of total events}}. \quad (1)$$

The reconstruction efficiency is calculated to be 96.0% in IDR-L and 95.7% in IDR-S. This difference is very small. Thus, the impact of isolated muon identification between different detector models is small.

3.2. Primary Vertex Finding

The $h \to \mu^+\mu^-$ candidate is subjected to the primary vertex finding. At the time of DBD [5], this vertex finding does not make so much sense because all events are occurred exactly at the position of $(0,0,0)$. However, in the current MC samples described in Section 2, the smearing effect is now considered, resulting the primary vertex position shifts in z-direction by $\sim 200\,\mu$m (x-/y-direction also shifts, but negligible compare to z-direction) [19].

In the signal process, muons from Higgs are produced at the primary vertex. To estimate the primary vertex position, we perform a fitting using `LCFIVertex` [20] package. The $h \to \mu^+\mu^-$ candidate which is selected in Section 3.1 is subjected to this fitting. In the fitting, we also consider the beam spot size constraint. We assume the beam spot size of $(x,y,z) = (150\,\mathrm{nm}, 5\,\mathrm{nm}, 200\,\mu\mathrm{m})$. The following Figures 1/2/3 show the distribution of primary vertex z-/x-/y-position (r_z/r_x/r_y) of signal process in IDR-L and IDR-S.

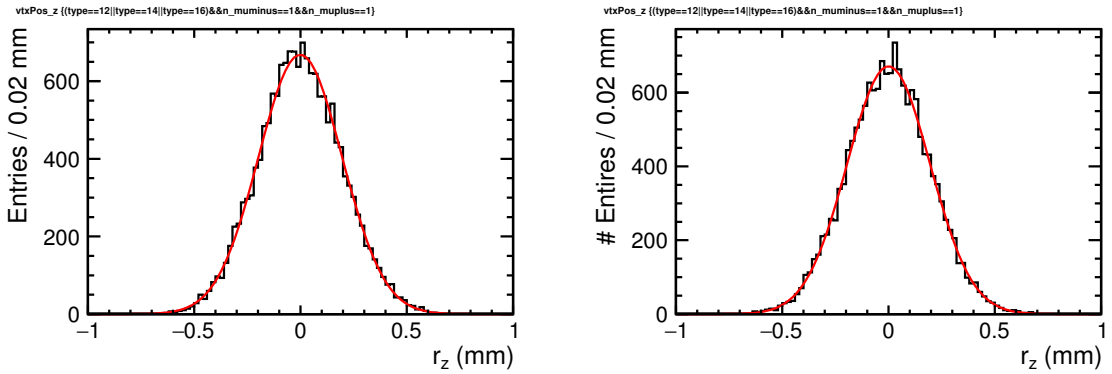

Figure 1: Estimated z-position of primary vertex r_z in the signal process. Left: IDR-L. Right: IDR-S. The red curves are the result of Gaussian fit.

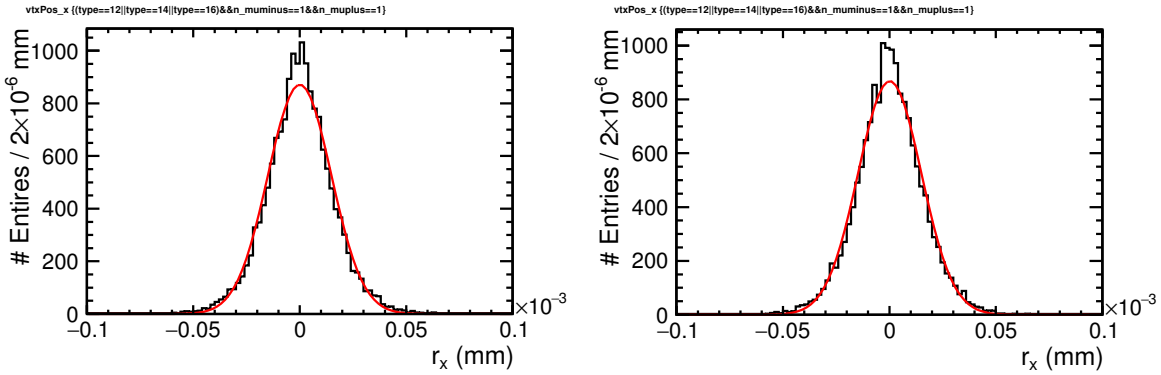

Figure 2: Estimated x-position of primary vertex r_x in the signal process. Left: IDR-L. Right: IDR-S. The red curves are the result of Gaussian fit.

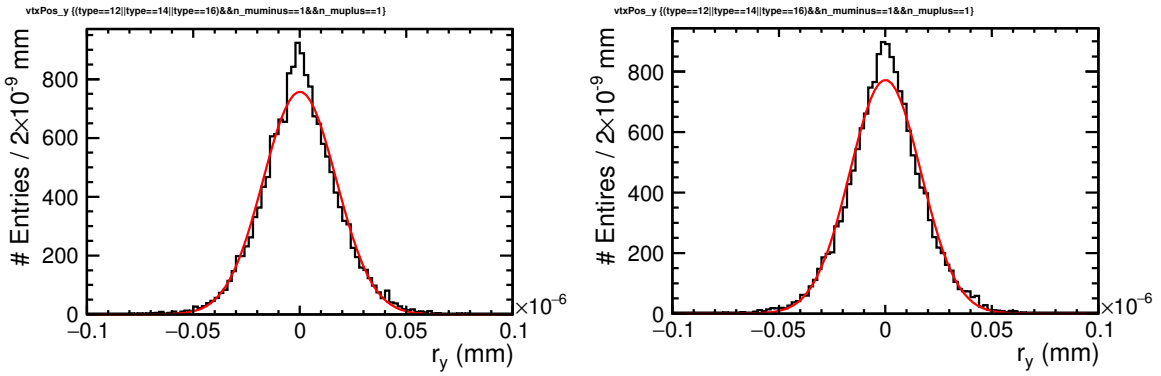

Figure 3: Estimated y-position of primary vertex r_y in the signal process. Left: IDR-L. Right: IDR-S. The red curves are the result of Gaussian fit.

Table 4 shows the width of Gaussian fit in Figures 1 , 2, and 3. The r_z distribution is fitted very well with Gaussian, while other directions could not be precisely fitted and estimated very small number. This is because the beam spot size in x-/y-directions are basically zero, which is not simulated properly. In such case, the fitting method does not work properly. Thus, we will not use r_x and r_y in further analysis.

Table 4: Width of Gaussian fit.

	IDR-L	IDR-S
r_z (mm)	0.1964 ± 0.0011	0.1949 ± 0.0011
r_x (mm)	$(1.483 \pm 0.010) \times 10^{-5}$	$(1.488 \pm 0.010) \times 10^{-5}$
r_y (mm)	$(1.709 \pm 0.012) \times 10^{-8}$	$(1.670 \pm 0.012) \times 10^{-8}$

3.3. Preselection

We apply several selection cuts to select signal and suppress background. The selection cuts are summarized in Table 5, where $\chi^2/\mathrm{Ndf}(\mu^{\pm})$ is the parameter how well a track fitted well in the track fitting divided by the number of degrees of freedom, $\chi^2/\mathrm{Ndf}(\mathrm{vertex})$ is the fitting quality of primary vertex position finding divided by the number of degrees of freedom, $\sigma(M_{\mu^+\mu^-})$ is the event-by-event mass

resolution of $M_{\mu^+\mu^-}$ which is calculable using covariance matrix of muons in momenta space, $M_{\mu^+\mu^-}$ is the invariant mass of muon pair, $\theta_{\mu^+\mu^-}$ is the angle between two muons of $h \to \mu^+\mu^-$ candidate, N_{P_t} is the number of charged particles with the transverse momentum greater than 5 GeV except $h \to \mu^+\mu^-$ candidate, $E_{\rm vis}$ is the visible energy, and $\theta_{\rm miss}$ is the angle of missing momentum with respect to the beam axis, respectively.

Table 5: List of preselection cuts. Definition of variables are written in the text.

#	variable	cut
0	nothing	no cut
1	# $\mu^\pm$	$= 1$
2	$\chi^2/\mathrm{Ndf}(\mu^\pm)$	$0.5 - 1.5$
3	$\chi^2/\mathrm{Ndf}(\mathrm{vertex})$	< 20
4	$\lvert r_z \rvert$	< 0.5 mm
5	$\sigma(M_{\mu^+\mu^-})$	< 1 GeV
6	$M_{\mu^+\mu^-}$	$100 - 130$ GeV
7	$\cos\theta_{\mu^+\mu^-}$	< 0.55
8	N_{P_t}	$= 0$
9	$E_{\rm vis}$	$125 - 300$ GeV
10	missing P_t	> 5 GeV
11	$\lvert\cos\theta_{\rm miss}\rvert$	< 0.99

The cuts #2 and #5 are requiring well-measured muon tracks. The cuts #3 and #4 are requiring prompt muons, and the cuts #6 and #7 are requiring the signature of $h \to \mu^+\mu^-$ candidate. The cut #8 is requiring the events basically no charged particles except $h \to \mu^+\mu^-$ candidate and $\gamma\gamma \to$ low P_t hadrons. The last three cuts #9 to #11 are rejecting neutrino-less events. We have applied these selection cuts to all analysis channels. These cuts are probably not optimized ultimately because the main production process of Higgs boson in left-handed and right-handed is different. The distributions of each variable before each cut are shown in Appendix due to too many plots. Tables 6 to 9 are the cut tables of each channel. In these 4 tables, the irreducible 4-fermion background processes are summarized in each column; $2\nu2\mu$ process, $2\nu2\tau$ process with both τ decays to μ (shorten as $2\nu2\tau(\mu)$), and $2\nu1\mu1\tau$ process with τ decays to μ (shorten as $2\nu\mu\tau(\mu)$), respectively. A simplified cut table after preselection is shown in Table 10.

Table 6: Cut table of IDR-L-left.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	ffh other	2f	4f $2\nu 2\mu$	4f $2\nu 2\tau(\mu)$	4f $2\nu\mu\tau(\mu)$	4f other	$\gamma\gamma \to$ 4f $2\nu 2\mu$	$\gamma\gamma \to$ 4f $2\nu 2\tau(\mu)$	$\gamma\gamma \to$ 4f $2\nu\mu\tau(\mu)$	$\gamma\gamma \to$ 4f other
0	57.54	31.12	4.122×10^5	1.084×10^7	5.922×10^5	1.323×10^4	1.272×10^5	3.734×10^7	2525.76	74.17	827.02	3.325×10^5
1	55.15	28.15	7102.10	2.141×10^6	3.811×10^5	8276.06	9.708×10^4	7.278×10^5	2227.94	33.02	540.03	1.403×10^4
2	53.94	27.69	6976.17	1.971×10^6	3.450×10^5	7376.35	8.660×10^4	6.331×10^5	2137.98	32.57	519.77	1.178×10^4
3	53.57	27.53	6207.86	1.916×10^6	3.426×10^5	637.40	2.250×10^4	6.187×10^5	2116.77	3.91	140.21	1.151×10^4
4	53.04	27.21	6139.05	1.895×10^6	3.391×10^5	620.51	2.235×10^4	6.124×10^5	2094.58	3.91	138.89	1.143×10^4
5	52.27	26.66	6051.78	1.434×10^6	3.230×10^5	613.00	2.181×10^4	5.687×10^5	2067.39	3.91	137.93	1.105×10^4
6	50.91	25.99	162.69	4.045×10^4	1.016×10^4	118.28	2096.37	1.792×10^4	283.60	0	8.76	79.72
7	50.90	25.95	121.27	2.560×10^4	9912.24	118.28	2096.37	1.675×10^4	283.60	0	8.76	79.18
8	50.74	0.17	3.66	2.510×10^4	9825.78	118.28	2086.60	4810.37	175.84	0	4.05	20.72
9	50.12	0.03	2.56	1.261×10^4	8167.13	84.49	1915.54	1347.92	144.64	0	3.74	17.93
10	49.94	0.02	2.56	975.81	8095.69	84.49	1915.54	854.88	143.09	0	3.74	3.63
11	48.90	0.01	2.56	123.25	7714.25	75.10	1866.70	258.64	136.09	0	3.74	0.89

Table 7: Cut table of IDR-L-right.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	ffh other	2f	4f $2\nu 2\mu$	4f $2\nu 2\tau(\mu)$	4f $2\nu\mu\tau(\mu)$	4f other	$\gamma\gamma \to$ 4f $2\nu 2\mu$	$\gamma\gamma \to$ 4f $2\nu 2\tau(\mu)$	$\gamma\gamma \to$ 4f $2\nu\mu\tau(\mu)$	$\gamma\gamma \to$ 4f other
0	7.93	20.71	1.303×10^5	9.536×10^6	6.406×10^4	1253.11	8298.29	1.661×10^7	2525.76	74.17	827.02	3.295×10^5
1	7.62	18.63	4049.84	1.661×10^6	3.437×10^4	667.68	6362.13	6.626×10^5	2227.94	33.02	540.03	1.403×10^4
2	7.47	18.25	3994.83	1.534×10^6	3.156×10^4	606.39	5729.40	5.730×10^5	2137.98	32.57	519.77	1.178×10^4
3	7.42	18.16	3702.32	1.491×10^6	3.133×10^4	51.62	1447.06	5.619×10^5	2116.77	3.91	140.21	1.151×10^4
4	7.34	17.98	3661.76	1.475×10^6	3.100×10^4	50.61	1438.00	5.560×10^5	2094.58	3.91	138.89	1.143×10^4
5	7.20	17.67	3591.54	1.143×10^6	2.865×10^4	50.16	1405.86	5.202×10^5	2067.39	3.91	137.93	1.105×10^4
6	7.02	17.18	102.89	2.745×10^4	1265.67	7.08	142.21	1.439×10^4	283.60	0	8.76	79.72
7	7.01	17.17	74.64	1.815×10^4	1163.13	7.08	142.21	1.373×10^4	283.60	0	8.76	79.18
8	6.98	0.14	1.24	1.773×10^4	1157.95	7.08	141.63	4419.57	175.84	0	4.05	20.72
9	6.66	0.03	0.49	8996.10	1015.17	5.05	129.53	1248.49	144.64	0	3.74	17.93
10	6.65	0.01	0.49	792.24	1009.03	5.05	129.53	812.09	143.09	0	3.74	3.63
11	6.57	0.01	0.49	120.28	959.56	4.49	126.60	184.82	136.09	0	3.74	0.89

Table 8: Cut table of IDR-S-left.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	ffh other	2f	4f $2\nu 2\mu$	4f $2\nu 2\tau(\mu)$	4f $2\nu\mu\tau(\mu)$	4f other	$\gamma\gamma \to$ 4f $2\nu 2\mu$	$\gamma\gamma \to$ 4f $2\nu 2\tau(\mu)$	$\gamma\gamma \to$ 4f $2\nu\mu\tau(\mu)$	$\gamma\gamma \to$ 4f other
0	57.54	31.12	4.122×10^5	1.084×10^7	5.922×10^5	1.323×10^4	1.272×10^5	3.734×10^7	2525.76	74.17	827.02	3.314×10^5
1	54.99	28.08	7080.17	2.144×10^6	3.819×10^5	8284.87	9.732×10^4	7.287×10^5	2225.44	31.91	538.71	1.400×10^4
2	53.64	27.58	6943.93	1.968×10^6	3.446×10^5	7368.90	8.632×10^4	6.329×10^5	2133.83	30.13	510.29	1.178×10^4
3	53.30	27.42	6196.70	1.912×10^6	3.423×10^5	657.86	2.252×10^4	6.185×10^5	2114.90	2.29	139.00	1.147×10^4
4	52.74	27.14	6131.87	1.891×10^6	3.386×10^5	652.24	2.232×10^4	6.120×10^5	2093.16	2.29	138.85	1.139×10^4
5	52.17	26.70	6066.94	1.518×10^6	3.267×10^5	646.62	2.193×10^4	5.763×10^5	2075.56	2.29	137.90	1.108×10^4
6	50.75	26.06	161.71	4.405×10^4	1.029×10^4	140.40	2077.71	1.811×10^4	284.36	0	8.98	83.10
7	50.74	26.01	117.62	2.724×10^4	1.004×10^4	140.40	2077.71	1.692×10^4	284.36	0	8.98	82.55
8	50.57	0.19	3.70	2.660×10^4	9943.95	140.40	2057.64	4735.93	181.08	0	4.41	22.52
9	49.98	0.03	2.56	1.385×10^4	8311.09	93.60	1887.07	1364.73	147.97	0	2.94	19.15
10	49.82	0.01	2.56	1094.75	8251.08	89.86	1887.07	886.88	145.97	0	2.94	7.85
11	48.80	0.01	2.56	67.04	7835.57	76.75	1831.89	261.23	138.03	0	2.94	3.87

Table 9: Cut table of IDR-S-right.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	ffh other	2f	4f $2\nu 2\mu$	4f $2\nu 2\tau(\mu)$	4f $2\nu\mu\tau(\mu)$	4f other	$\gamma\gamma \to$ 4f $2\nu 2\mu$	$\gamma\gamma \to$ 4f $2\nu 2\tau(\mu)$	$\gamma\gamma \to$ 4f $2\nu\mu\tau(\mu)$	$\gamma\gamma \to$ 4f other
0	7.93	20.71	1.303×10^5	9.536×10^6	6.405×10^4	1250.57	8299.20	1.661×10^7	2525.76	74.17	827.02	3.314×10^5
1	7.60	18.60	4046.29	1.646×10^6	3.443×10^4	672.89	6376.69	6.623×10^5	2225.44	31.91	538.71	1.400×10^4
2	7.44	18.26	3983.42	1.515×10^6	3.153×10^4	608.76	5708.81	5.724×10^5	2133.83	30.13	510.29	1.178×10^4
3	7.40	18.15	3704.76	1.476×10^6	3.132×10^4	52.42	1431.29	5.612×10^5	2114.90	2.29	139.00	1.147×10^4
4	7.31	17.96	3670.32	1.459×10^6	3.098×10^4	52.08	1419.58	5.553×10^5	2093.16	2.29	138.85	1.138×10^4
5	7.20	17.70	3622.14	1.193×10^6	2.915×10^4	51.74	1396.17	5.255×10^5	2075.56	2.29	137.90	1.108×10^4
6	7.01	17.21	104.27	3.017×10^4	1263.03	8.40	137.36	1.448×10^4	284.36	0	8.98	83.10
7	7.01	17.19	75.53	1.957×10^4	1162.24	8.40	137.36	1.380×10^4	284.36	0	8.98	82.55
8	6.98	0.14	1.92	1.912×10^4	1154.54	8.40	136.16	4530.65	181.08	0	4.41	22.52
9	6.63	0.03	0.49	1.000×10^4	1017.68	5.60	125.96	1296.90	147.97	0	2.94	19.15
10	6.62	0.01	0.49	811.08	1012.22	5.38	125.96	812.51	145.97	0	2.94	7.85
11	6.55	~ 0	0.49	148.99	957.52	4.59	122.66	178.62	138.03	0	2.94	3.87

Table 10: A simplified cut table after preselection for each analysis channel. Numbers with % show the signal selection efficiency. The irreducible processes are defined in the text.

	signal $h \to \mu^+\mu^-$	other Higgs	4f/$\gamma\gamma \to$ 4f irreducible	other SM bkg.
IDR-L-left	48.90(85.0%)	2.57	9795.88	382.78
IDR-S-left	48.80(84.8%)	2.57	9885.18	332.14
IDR-L-right	6.57(82.8%)	0.50	1230.48	305.99
IDR-S-right	6.55(82.6%)	0.49	1225.74	331.48

After applying preselection, the signal-to-background ratio is $\sim 1/200 - 1/250$ in all cases. Clearly, the irreducible background processes are dominant, as expected. The Higgs related backgrounds are already getting negligible at this point.

3.3.1. Detector Effect

From Table 10, we see no major differences between IDR-L and IDR-S. However, we see some difference in the distribution of variables. Figure 4 shows the distribution of $M_{\mu^+\mu^-}$ and $\sigma(M_{\mu^+\mu^-})$ after preselection for IDR-L and IDR-S in left-handed beam polarization.

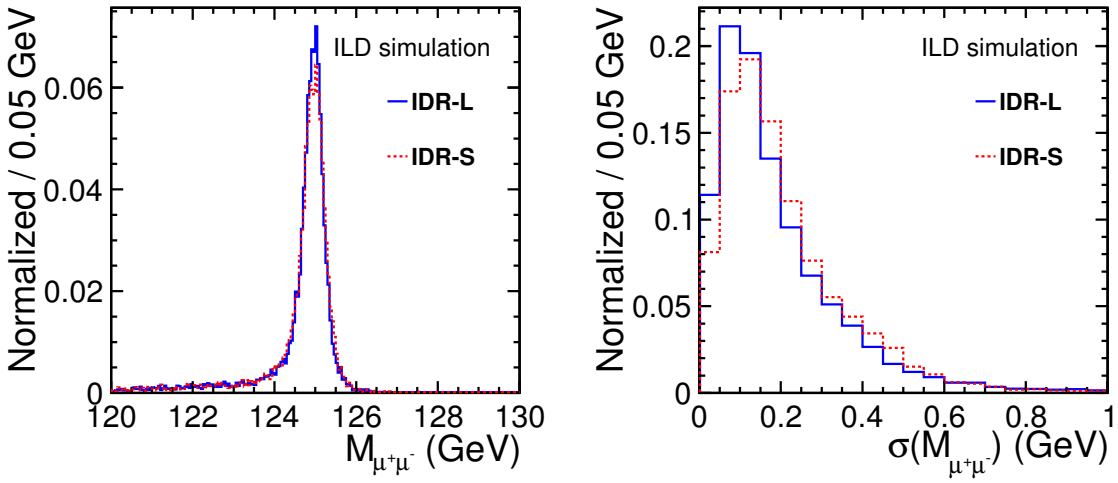

Figure 4: Distribution of variables after preselection for IDR-L and IDR-S. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

We can clearly see difference between IDR-L and IDR-S. This difference comes from the detector effect. Overall, IDR-L shows slightly better performance than IDR-S. However, the transverse momentum resolution has an angular dependency. The IDR-L has better resolution in barrel region, while IDR-S has better resolution in forward region. This point is discussed in the IDR. Figure 5 shows the similar distribution as Figure 4, but both muons are required in barrel region ($|\cos\theta_{\mu^{+/-}}| < 0.7$, where $\theta_{\mu^{+/-}}$ is the angle of $\mu^{+/-}$ with respect to the beam axis). In Figure 5, 43.8% events after preselection are plotted in both detector model. In this case, IDR-L gives significantly better results than IDR-S. On the other hand, Figure 6 shows the same distribution, but both muons are required in endcap region and/or forward region ($|\cos\theta_{\mu^{+/-}}| > 0.7$). In Figure 6, 8.4% events are plotted in both detector model. Due to small number of MC events, it is hard to describe, but $M_{\mu^+\mu^-}$ distributions looks almost similar, and

130 $\sigma(M_{\mu^+\mu^-})$ looks better in IDR-S. The remaining mixed case (one muon in barrel, another muon in en-
131 dcap/forward) is shown in Figure 7. The $M_{\mu^+\mu^-}$ looks pretty similar, while we see some difference in
132 $\sigma(M_{\mu^+\mu^-})$ distribution.

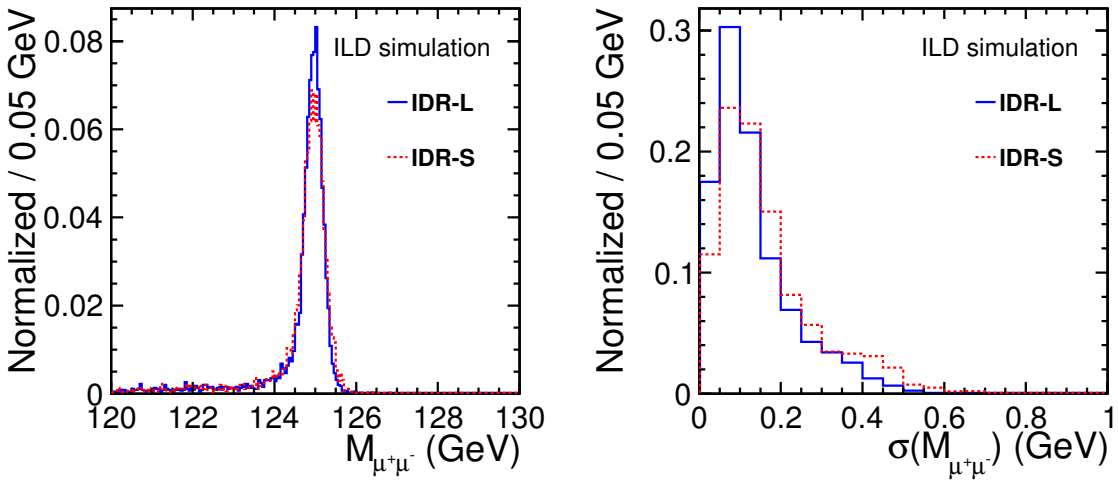

Figure 5: Similar to Figure 4, but an additional cut $|\cos\theta_{\mu^{+/-}}| < 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

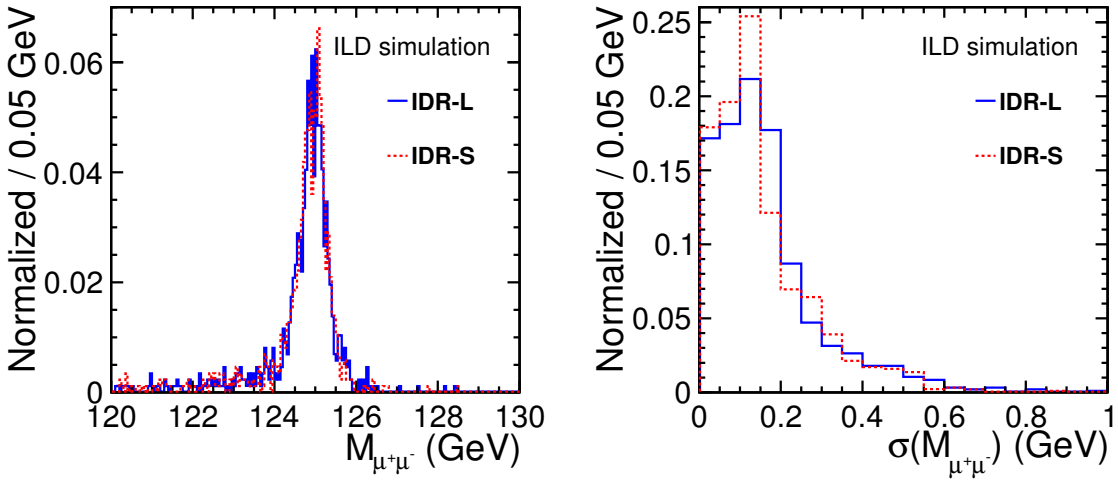

Figure 6: Similar to Figure 4, but an additional cut $|\cos\theta_{\mu^{+/-}}| > 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

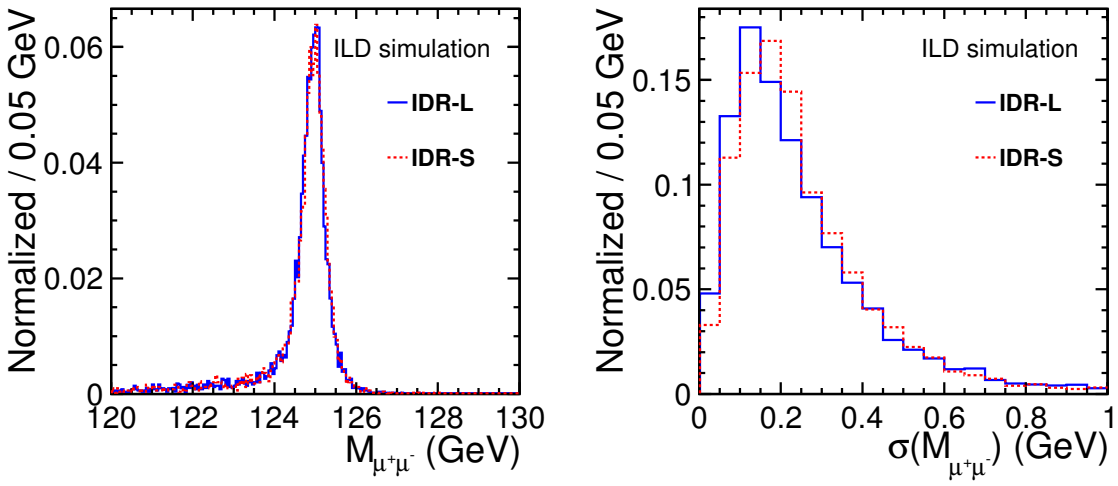

Figure 7: Similar to Figure 4, but additional cuts are applied to require $|\cos\theta_{\mu^{+/-}}| < 0.7$ for one muon, and $|\cos\theta_{\mu^{+/-}}| > 0.7$ for another muon. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

3.4. TMVA Analysis

After applying preselection, we perform multivariate analysis to improve sensitivity. In this analysis, a gradient boosted decision tree technique (BDTG) is used which is implemented in TMVA in ROOT [21, 22]. The half of remained samples are used for training, and another half is used for testing. The following 6 variables are used as the inputs: $\cos\theta_{\mu^+\mu^-}$, $\Sigma_{i=+,-}[\text{charge}(\mu^i) \times \cos\theta_{\mu^i}]$, E_{lead}, E_{sub}, $\cos\theta_{\text{lead}}$, and $\cos\theta_{\text{sub}}$, where θ_{μ^i} is the angle of μ^i with respect to the beam axis, $E_{\text{lead}}(\theta_{\text{lead}})$ is the energy(angle) of muon which has larger energy in $h \to \mu^+\mu^-$ candidate, and $E_{\text{sub}}(\theta_{\text{sub}})$ is the energy(angle) of muon which has smaller energy in $h \to \mu^+\mu^-$ candidate, respectively. The following 6 figures show the distribution of each input variable after preselection in IDR-L-left case (only show signal and irreducible background, each histogram is normalized to 1). There are several spikes in the plot, this is due to lack of MC statistics for SM background. In all analysis channels, we use the same variables for the input to BDTG. Again, these input variables are probably not optimized ultimately because the different production process of Higgs in left-/right-handed beam polarization. However, these are still useful variables to suppress the backgrounds.

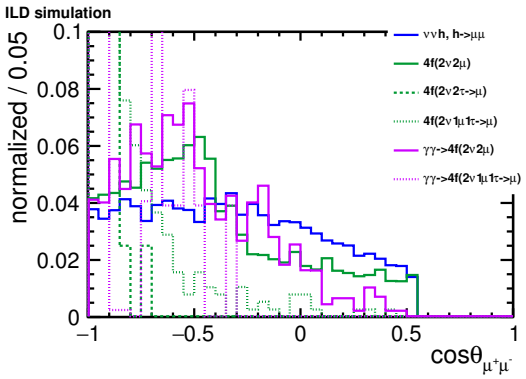

Figure 8: Distribution of $\cos\theta_{\mu^+\mu^-}$.

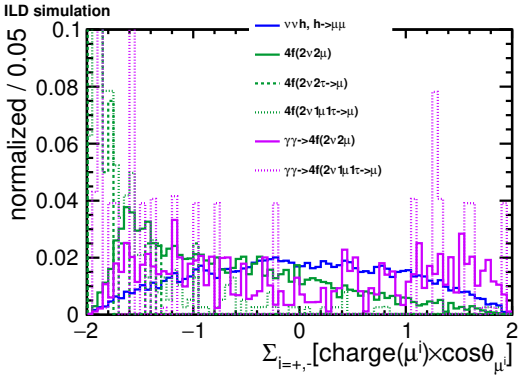

Figure 9: Distribution of $\Sigma_{i=+,-}[\text{charge}(\mu^i) \times \cos\theta_{\mu^i}]$.

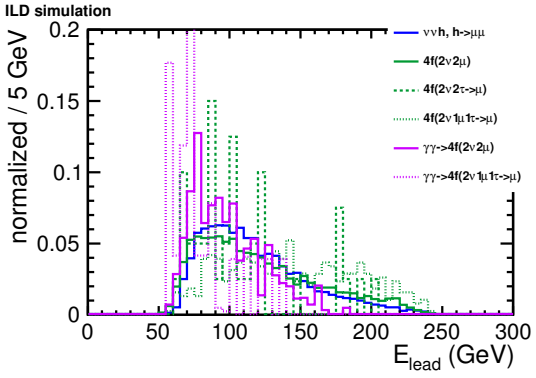

Figure 10: Distribution of E_{lead}.

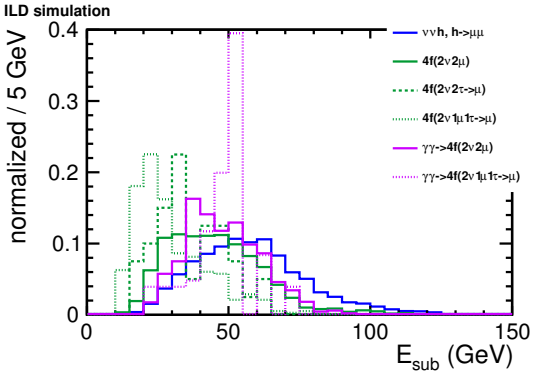

Figure 11: Distribution of E_{sub}.

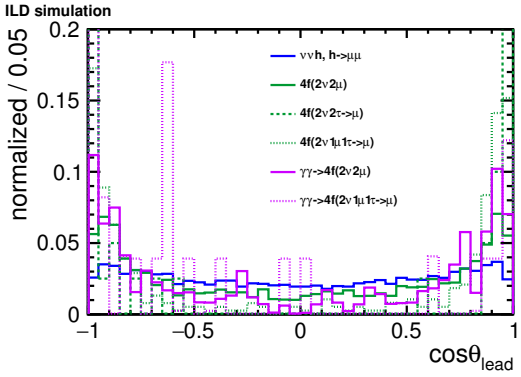

Figure 12: Distribution of $\cos\theta_{\text{lead}}$.

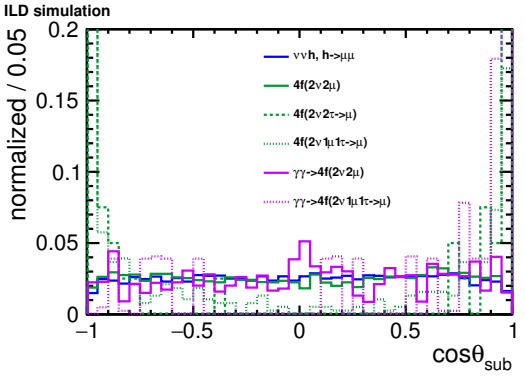

Figure 13: Distribution of $\cos\theta_{\text{sub}}$.

The BDTG score for signal and background in IDR-L-left is shown in Figure 14. Optimizing the cut point of BDTG score will be discussed in next section.

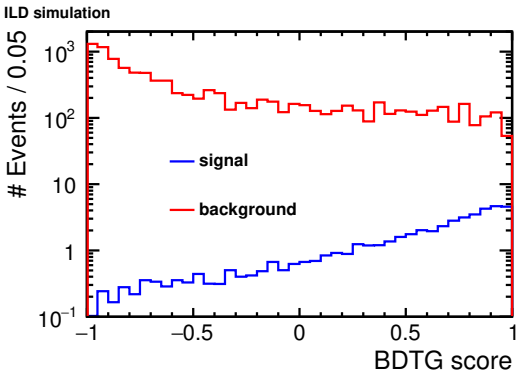

Figure 14: Distribution of BDTG score in IDR-L-left.

3.5. Toy MC

To extract the precision, we perform toy MC at the end of the analysis. We use RooFit [23] for toy MC. We use $M_{\mu^+\mu^-}$ distribution of test samples after all cuts for toy MC. For signal, a linear sum of Crystal Ball function and Gaussian, $k \times (\text{Crystal Ball}) + (1-k) \times \text{Gaussian}$ $(0 < k < 1)$, is used as the modeling function f_S. Since we have not applied any compensation for final state radiation photon from muons, we expect a tail structure in low mass region in the spectrum of $M_{\mu^+\mu^-}$ for the signal process. A Crystal Ball function would be a perfect modeling function to model such distribution. An additional Gaussian will represent the detector effect. In the Crystal Ball function, we have fixed its mean as 125 because we already know the mass of Higgs boson is near to 125 GeV. A first order polynomial is used as the background modeling function f_B, because we expect almost flat distribution after all cuts. Figure 15 shows the result of modeling in signal and background in IDR-L-left channel.

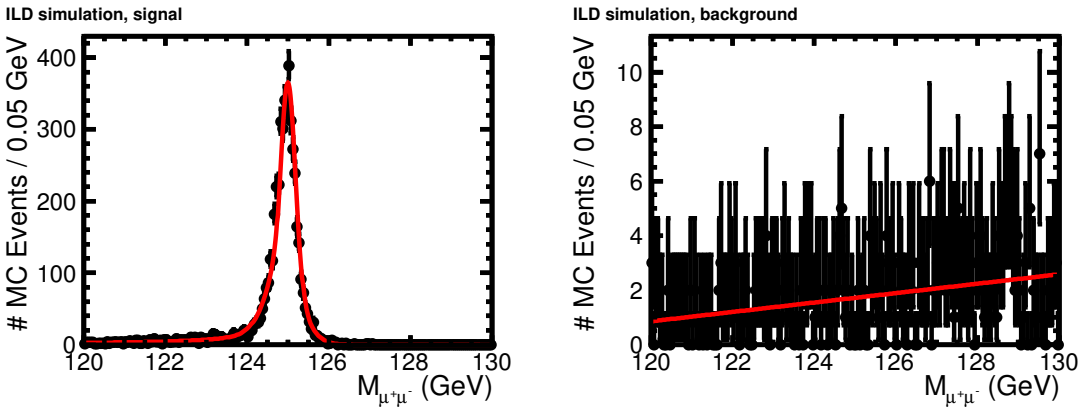

Figure 15: Result of fitting with modeling function after applying all cuts in IDR-L-left. Left: signal fitting with f_S. Right: background fitting with f_B.

After the modeling, we then perform pseudo-experiment using parametrized f_S and f_B. In one pseudo-experiment, the number of pseudo-signal(-background) is determined as the number of signal(background) after all cuts plus a random Poisson fluctuation. Then the pseudo-data are subjected to an unbinned fit using the function $f \equiv Y_S f_S + Y_B f_B$, where $Y_S(Y_B)$ is the yield of signal(background). In the unbinned fit, we have fixed Y_B as the number of backgrounds after all cuts, because the SM background can be predicted more precisely than the statistical uncertainty for rare signal events at a lepton collider. Figure 16 left shows one example of pseudo-experiment in IDR-L-left. We repeat this pseudo-experiments

by 50000 times and obtain final Y_S distribution as shown in Figure 16 right. The precision is estimated with a Gaussian fitting to the Y_S distribution.

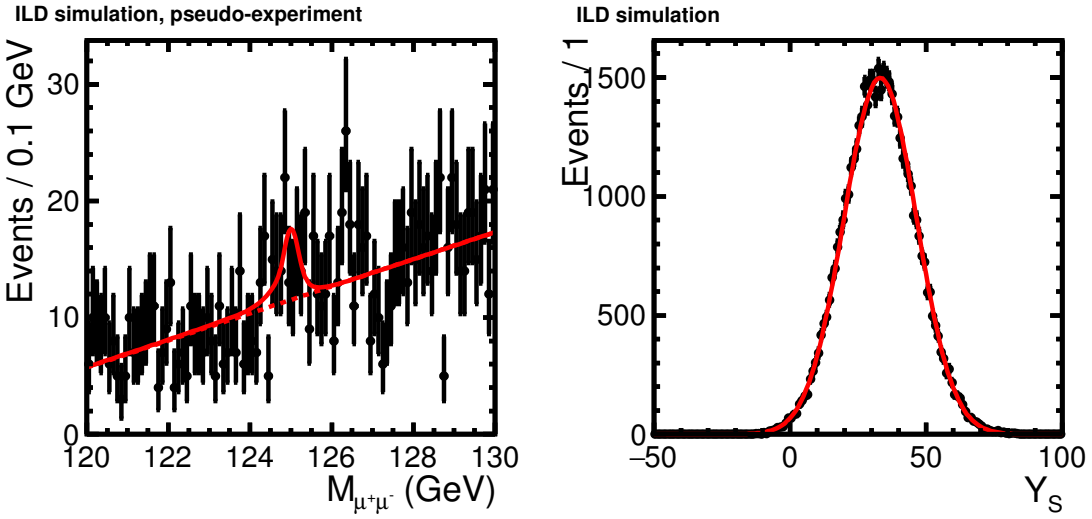

Figure 16: Result of pseudo-experiment in IDR-L-left. Left: one example of pseudo-experiment. Black dots are the pseudo-data, solid red curve is the result of fitting using function f, and dotted-red is its background component $Y_B f_B$. Right: Y_S distribution after 50000 times pseudo-experiments together with Gaussian fitting in red curve.

An optimization to determine the cut point on BDTG score is performed using the results of toy MC. We repeat above toy MC procedures by changing the BDTG score cut. We determine the optimal BDTG cut point which gives the best precision. Figures 15 and 16 are the plots with optimal BDTG score cut in IDR-L-left.

4. Results and Discussion

Table 11 shows the simplified cut table after applying optimum BDTG score cut in each analysis channel. For final event numbers, an additional cut of $M_{\mu^+\mu^-} > 120$ GeV is also applied because there are almost no signal events in the tail region anymore. We summarize all obtained precision on the cross section times branching ratio $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ from toy MC experiment in Table 12, together with theoretical limit precision which assumes 100% signal efficiency, no backgrounds, and no detector effects.

Table 11: A simplified cut table after all cuts for each analysis channel. Numbers with % show the signal selection efficiency. The irreducible processes are defined in the text.

	BDTG score cut	signal $h \to \mu^+\mu^-$	other Higgs	4f/$\gamma\gamma \to$ 4f irreducible	other SM bkg.
IDR-L-left	> 0.40	33.17(57.6%)	0.01	1044.25	57.29
IDR-S-left	> 0.45	33.01(57.4%)	0.01	1040.94	28.30
IDR-L-right	> 0.20	5.43(68.5%)	~ 0	232.85	12.49
IDR-S-right	> -0.15	5.26(66.3%)	~ 0	192.73	9.64

Table 12: Obtained precision on cross section times branching ratio $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ for each analysis channel. The theoretical limit precision is also summarized in the column of theory.

	IDR-L	IDR-S	theory
left	$40.15 \pm 0.15\%$	$41.11 \pm 0.15\%$	13.18%
right	$114.51 \pm 0.69\%$	$113.73 \pm 0.68\%$	35.51%

From these two tables, our analysis can be summarized as following.

- The Higgs related backgrounds are now completely gone.
- The signal-to-background ratio is $\sim 1/30 - 1/40$, and irreducible background is a major showstopper for the precise measurement due to its difficulties of suppression.
- Our analysis is about a factor of 3 far away from the theoretical limit. There are several reasons; imperfection of cuts, existence of irreducible background mainly come from $e^+e^- \to W^+W^- \to 2\mu 2\nu$.
- The right-handed beam polarization cases cannot perform precise measurement due to its very small number of signal events, and we see no difference in the precision. For further discussion, we will not consider this polarization anymore.
- We see some difference in the remaining number of backgrounds, but this is just a statistical fluctuation due to lack of MC statistics for SM background. Since we will not expect any difference on the distribution of background after all cuts, we should use same parametrization for background between IDR-L and IDR-S. This point will be discussed at Section 4.2.

4.1. Detector Effect

As we discussed at Section 3.3.1, we will also check the detector effect for $M_{\mu^+\mu^-}$ and $\sigma(M_{\mu^+\mu^-})$ after all cut. Figure 17 shows these distributions after all cut. Overall, IDR-L has better performance than IDR-S. This is one of the reason why IDR-L gives better results than IDR-S.

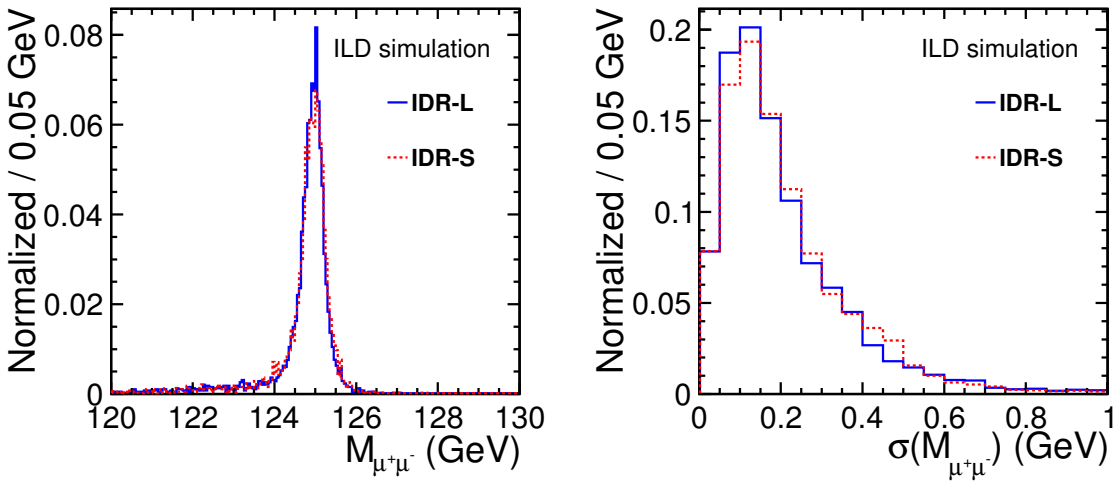

Figure 17: Distribution of variables after BDTG cut for IDR-L and IDR-S. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

Again, we will check the angular dependency of these two variables. Figure 18 shows the distribution with both muons are in barrel region ($|\cos\theta_{\mu^{+/-}}| < 0.7$), and Figure 19 shows in non-barrel region ($|\cos\theta_{\mu^{+/-}}| > 0.7$). In Figure 18, 46.2%/47.0% events after all cuts are plotted with IDR-L/IDR-S. In the barrel region, IDR-L has significantly better performance than IDR-S. In Figure 19, 5.5%/5.3% events after all cuts are plotted with IDR-L/IDR-S. In endcap/forward region, it is hard to conclude due to lack of MC statistics, but both detector have similar performance. The rest of mixed case is plotted in Figure 20. The $M_{\mu^+\mu^-}$ distribution is almost same, while we see some difference in $\sigma(M_{\mu^+\mu^-})$.

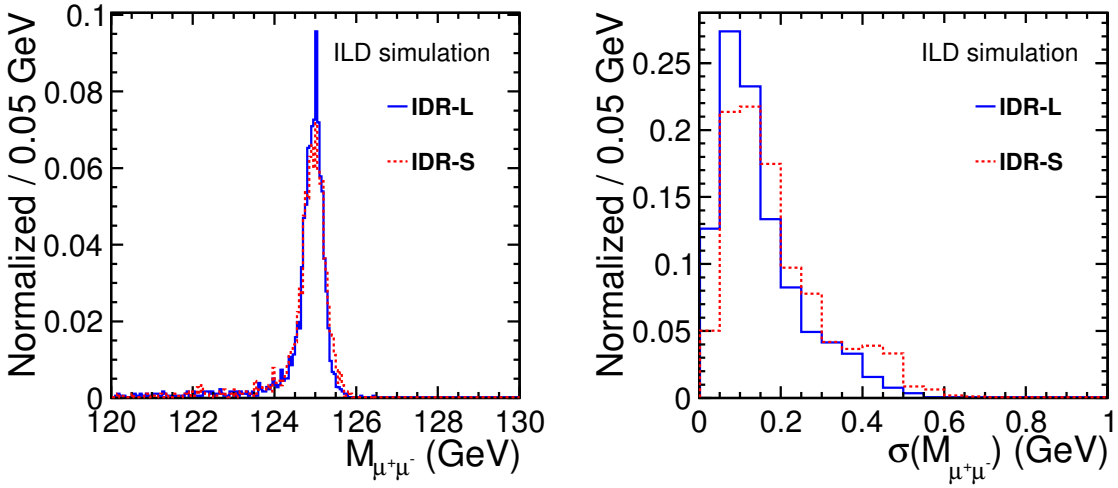

Figure 18: Similar to Figure 17, but an additional cut $|\cos\theta_{\mu^{+/-}}| < 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

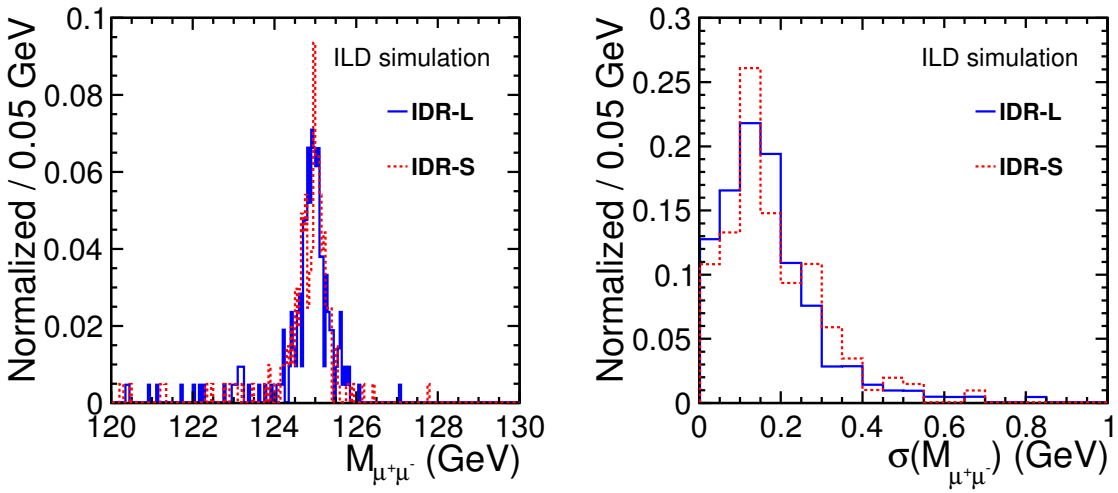

Figure 19: Similar to Figure 17, but an additional cut $|\cos\theta_{\mu^{+/-}}| > 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

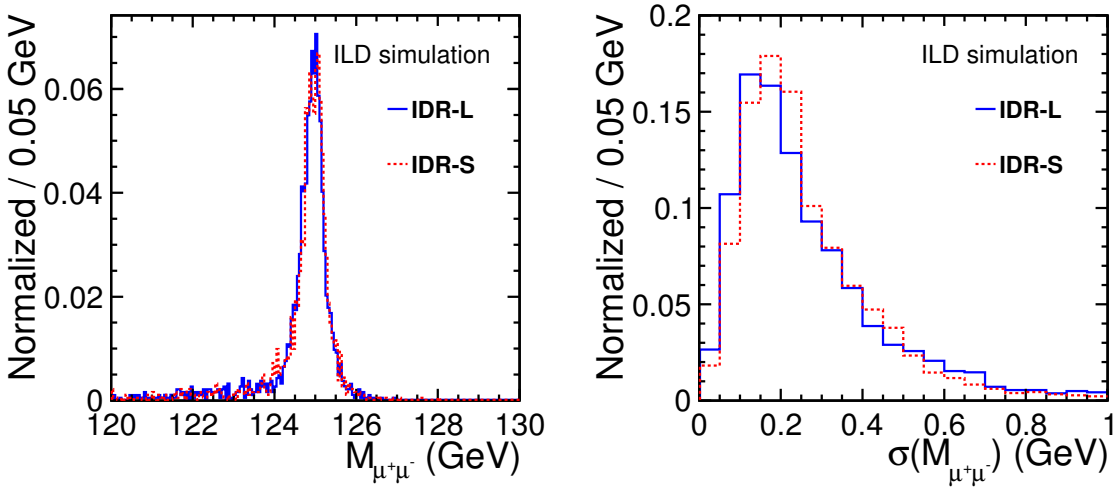

Figure 20: Similar to Figure 17, but additional cuts are applied to require $|\cos\theta_{\mu^{+/-}}| < 0.7$ for one muon, and $|\cos\theta_{\mu^{+/-}}| > 0.7$ for another muon. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are left-handed, and normalized to 1.

4.2. Final Results

As we discussed in previous, we will not expect any difference in the background distribution, and current difference is came form statistical fluctuation, we should use same parametrization for both detector configurations in the final toy MC. We take an average of number of remaining backgrounds after all cuts between IDR-L and IDR-S, and also take an average of first order polynomial parametrization which was discussed in Section 3.5. We perform toy MC with common treatment for the background, and Table 13 shows the final results.

Table 13: Obtained precision on cross section times branching ratio $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ for each analysis channel, with same treatment for the background in the toy MC. The theoretical limit precision is also summarized in the column of theory.

	IDR-L	IDR-S	theory
left	$40.16 \pm 0.15\%$	$41.28 \pm 0.15\%$	13.18%

The IDR-L gives relatively 2.8% better precision on $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ than IDR-S. As we saw in Figure 17, we see the different distribution on $M_{\mu^+\mu^-}$ between IDR-L and IDR-S. In the signal modeling, the width of Crystal Ball function is $\sim 10\%$ wider in IDR-S. We have checked that the number of signal events in peak region is almost same between IDR-L and IDR-S. This wider width gives $\sim 10\%$ more background in the peak region for IDR-S. From the statistical point of view, 10% more backgrounds give roughly $\sim 3.3\%$ worse significance. Our relative 2.8% difference is consistent to this rough estimation.

In conclusion, IDR-S gives worse performance than IDR-L due to worse transverse momentum resolution in barrel region. This worse resolution causes wider width of signal modeling, resulting more backgrounds in the peak region of $M_{\mu^+\mu^-}$ which essentially determines the precision.

5. Summary

We have investigated $e^+e^- \to \nu\bar{\nu}h$ with $h \to \mu^+\mu^-$ process at $\sqrt{s} = 500$ GeV at the ILC, assuming different two detector models, IDR-L and IDR-S, and two beam polarization configurations. It is difficult to perform precise measurement in the right-handed beam polarization due to too small number of signal events. The precision on $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ with IDR-L-left is estimated to be $40.16 \pm 0.15\%$, and IDR-S-left is $41.28 \pm 0.15\%$, relatively 2.8% better in IDR-L. This difference is came from different performance of transverse momentum resolution σ_{1/P_t} in barrel region. The worse resolution gives $\sim$ 10% wider width of signal modeling in IDR-S, resulting more backgrounds in the peak region of $M_{\mu^+\mu^-}$ which essentially determines the final precision. In conclusion, for this analysis, IDR-L is better detector option.

Acknowledgement

We would like to thank the LCC generator working group and the ILD software working group for providing the simulation and reconstruction tools and producing the Monte Carlo samples used in this study. This work has benefited from computing services provided by the ILC Virtual Organization, supported by the national resource providers of the EGI Federation and the Open Science GRID.

A. Plots of IDR-L-left

In the next 4 appendix sections (A - D), we will show the distribution of variables which are used in the preselection. The distribution of a variable when before applying a cut to that variable is shown. All histogram colors are the same as follows.

- black: sum up all processes
- solid blue: signal ($\nu\bar{\nu}h,\ h \to \mu^+\mu^-$)
- dotted blue: $q\bar{q}h/\ell^+\ell^-h,\ h \to \mu^+\mu^-$
- ash: $e^+e^- \to f\bar{f}h,\ h \not\to \mu^+\mu^-$
- red: all 2f process
- solid green: 4f, $2\nu 2\mu$ [irreducible]
- large-dotted green: 4f, $2\nu 2\tau,\ \tau \to \mu$ [irreducible]
- small-dotted green: 4f, $2\nu 1\mu 1\tau,\ \tau \to \mu$ [irreducible]
- large-and-small dotted green: other all 4f processes
- solid purple: $\gamma\gamma \to$ 4f, $2\nu 2\mu$ [irreducible]
- large-dotted purple: $\gamma\gamma \to$ 4f, $2\nu 2\tau,\ \tau \to \mu$ [irreducible]
- small-dotted purple: $\gamma\gamma \to$ 4f, $2\nu 1\mu 1\tau,\ \tau \to \mu$ [irreducible]
- large-and-small dotted purple: other all $\gamma\gamma \to$ 4f processes

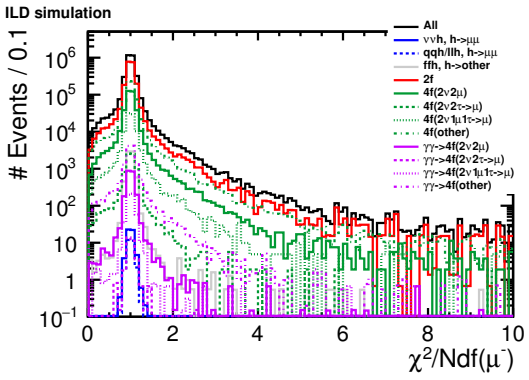

Figure 21: $\chi^2/\mathrm{Ndf}(\mu^-)$ before cut #2.

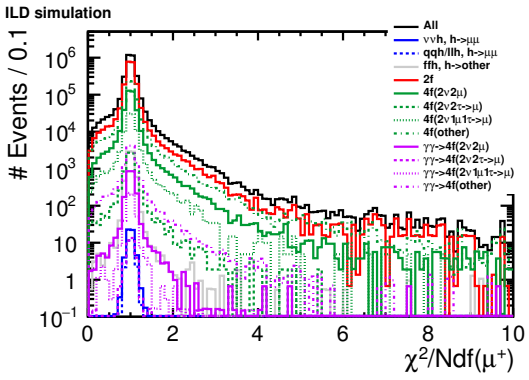

Figure 22: $\chi^2/\mathrm{Ndf}(\mu^+)$ before cut #2.

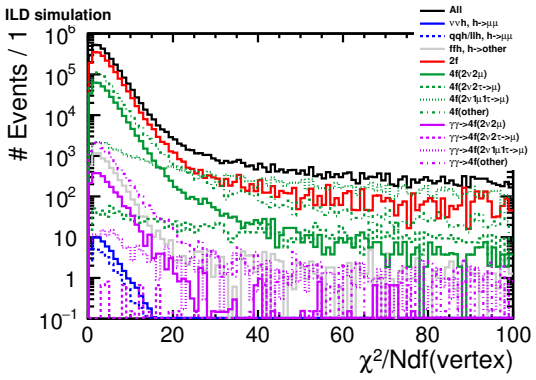

Figure 23: $\chi^2/\mathrm{Ndf(vertex)}$ before cut #3.

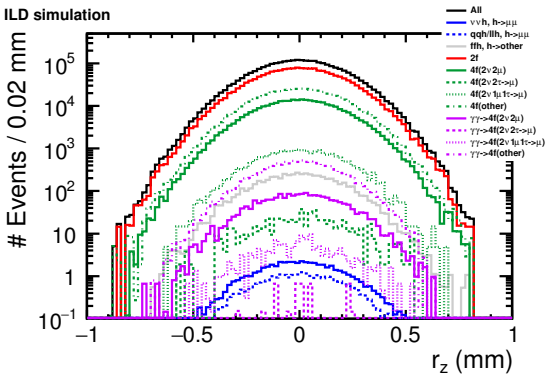

Figure 24: r_z before cut #4.

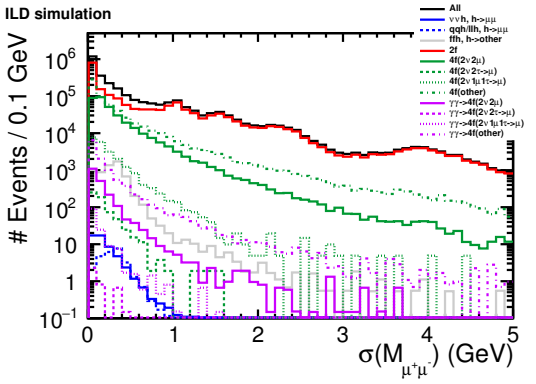

Figure 25: $\sigma(M_{\mu^+\mu^-})$ before cut #5.

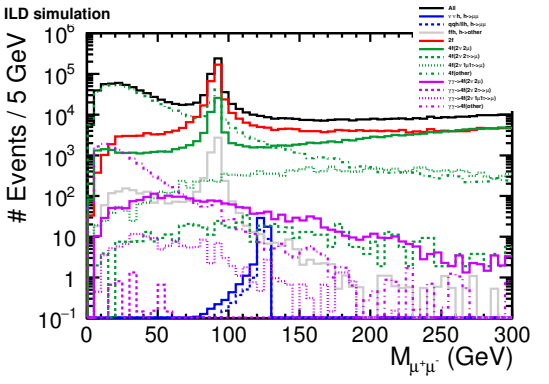

Figure 26: $M_{\mu^+\mu^-}$ before cut #6.

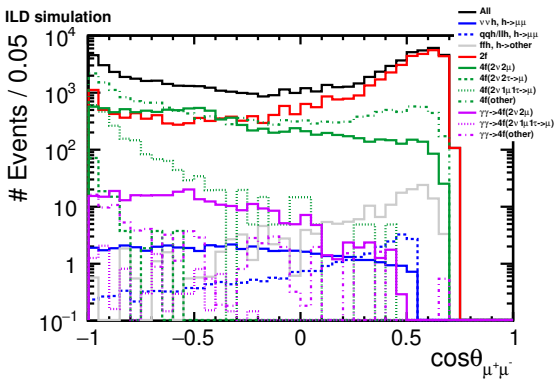

Figure 27: $\cos\theta_{\mu^+\mu^-}$ before cut #7.

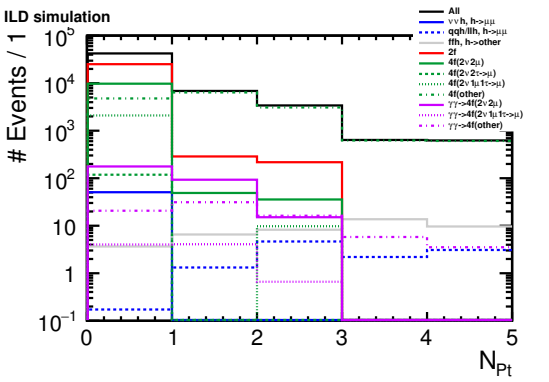

Figure 28: N_{P_t} before cut #8.

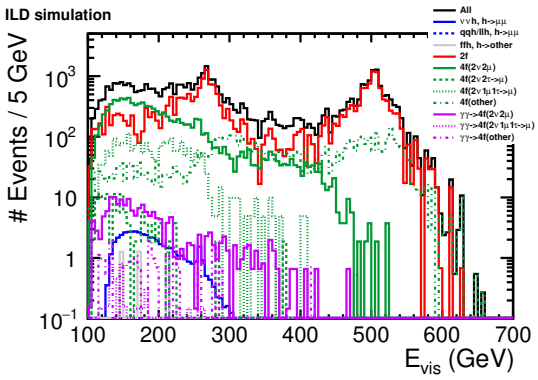

Figure 29: E_{vis} before cut #9.

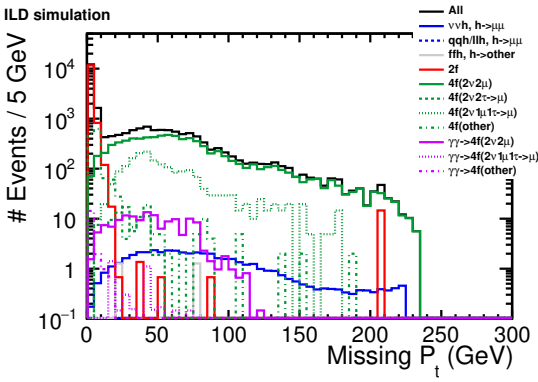

Figure 30: Missing P_t before cut #10.

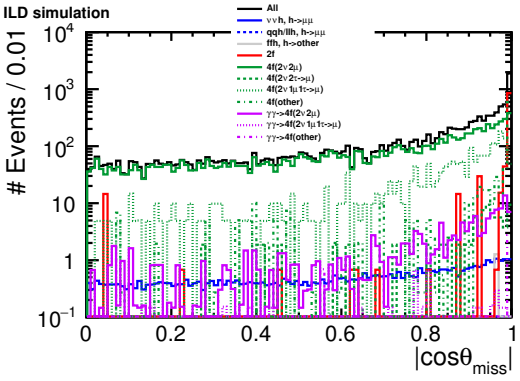

Figure 31: $\cos\theta_{\text{miss}}$ before cut #11.

B. Plots of IDR-L-right

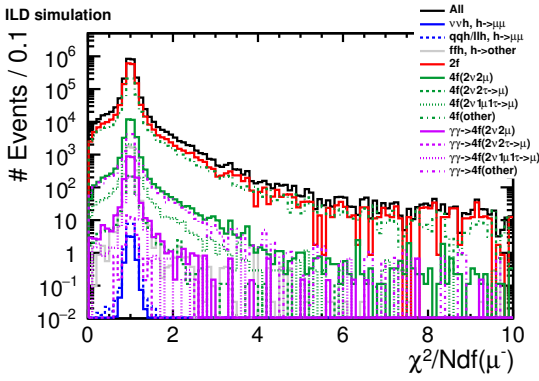

Figure 32: $\chi^2/\text{Ndf}(\mu^-)$ before cut #2.

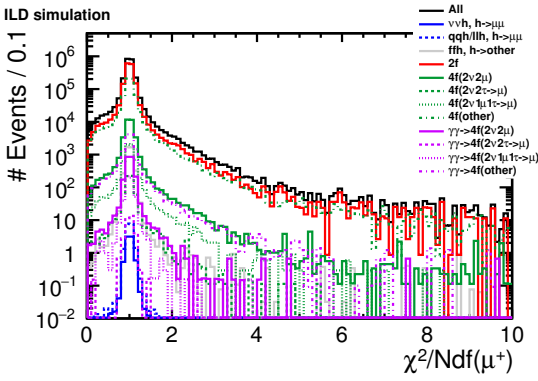

Figure 33: $\chi^2/\text{Ndf}(\mu^+)$ before cut #2.

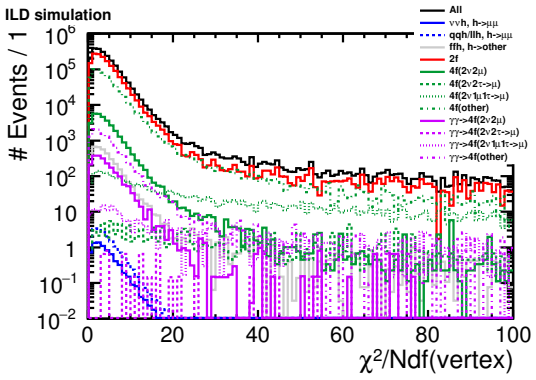

Figure 34: $\chi^2/\mathrm{Ndf}(\mathrm{vertex})$ before cut #3.

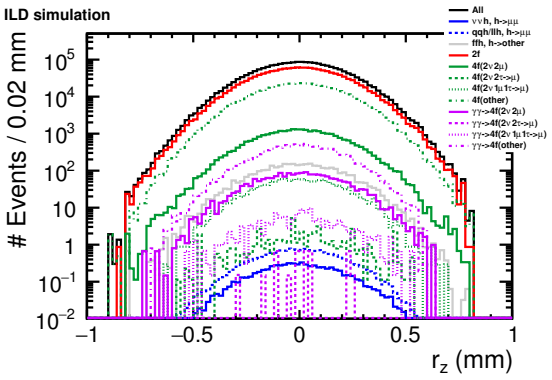

Figure 35: r_z before cut #4.

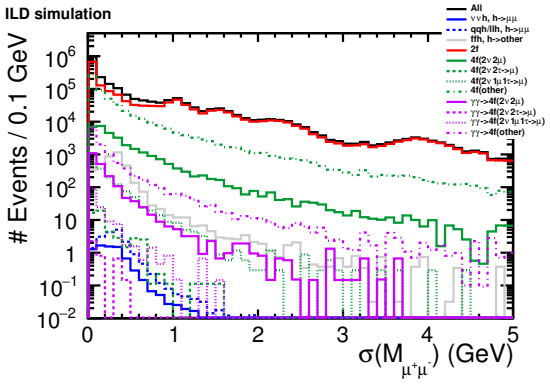

Figure 36: $\sigma(M_{\mu^+\mu^-})$ before cut #5.

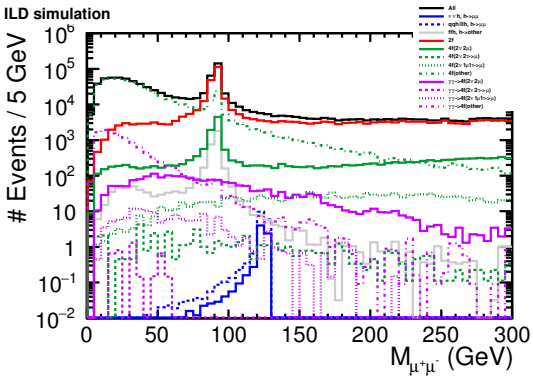

Figure 37: $M_{\mu^+\mu^-}$ before cut #6.

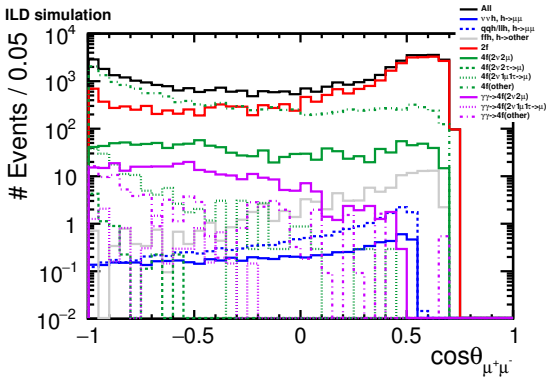

Figure 38: $\cos\theta_{\mu^+\mu^-}$ before cut #7.

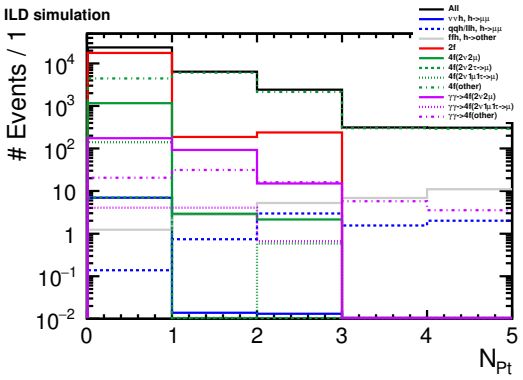

Figure 39: N_{P_t} before cut #8.

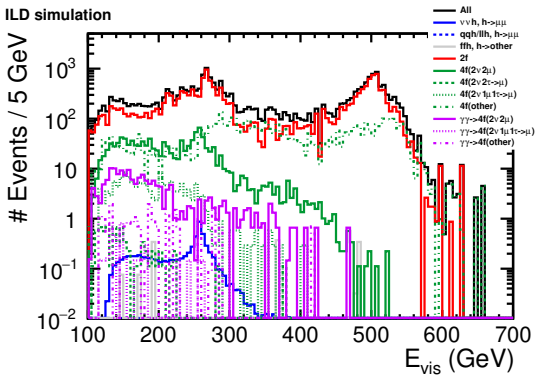

Figure 40: E_{vis} before cut #9.

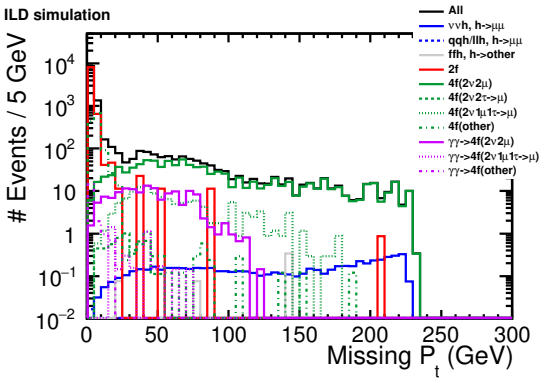

Figure 41: Missing P_t before cut #10.

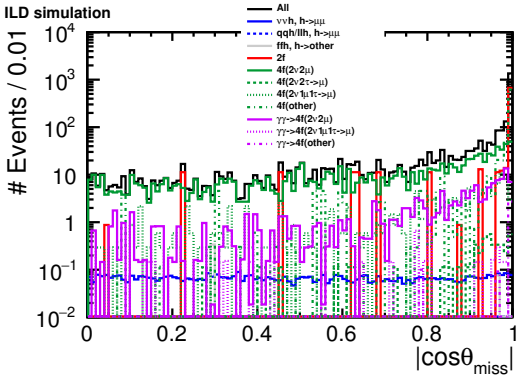

Figure 42: $\cos\theta_{\text{miss}}$ before cut #11.

C. Plots of IDR-S-left

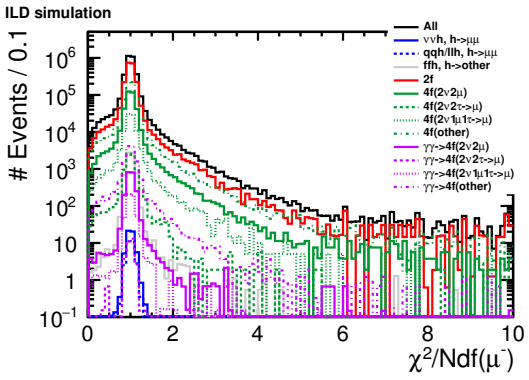

Figure 43: $\chi^2/\text{Ndf}(\mu^-)$ before cut #2.

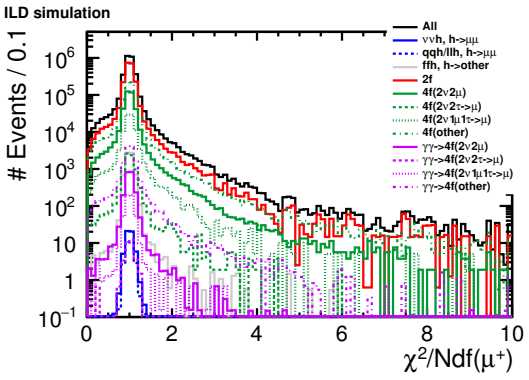

Figure 44: $\chi^2/\text{Ndf}(\mu^+)$ before cut #2.

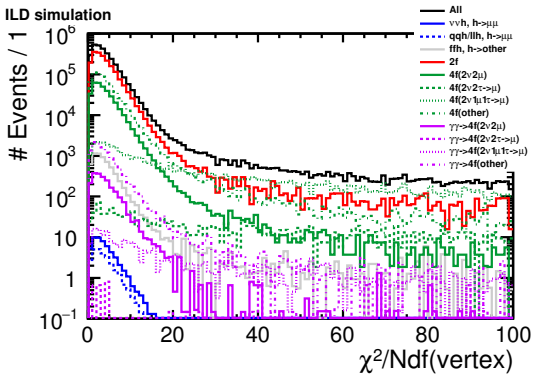

Figure 45: $\chi^2/\mathrm{Ndf}(\mathrm{vertex})$ before cut #3.

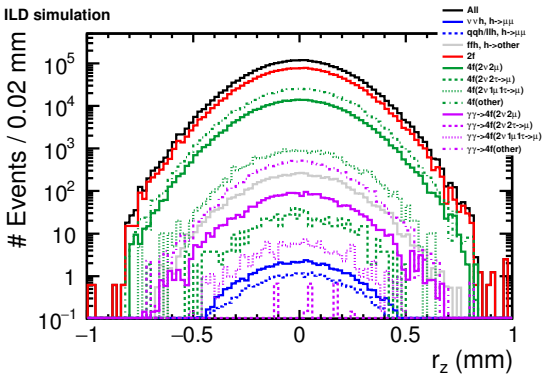

Figure 46: r_z before cut #4.

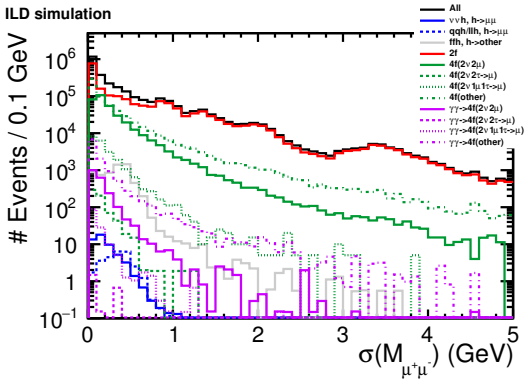

Figure 47: $\sigma(M_{\mu^+\mu^-})$ before cut #5.

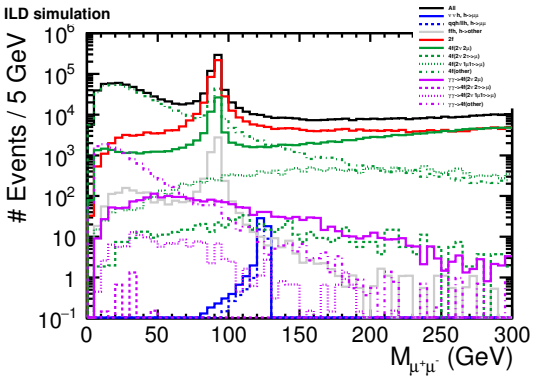

Figure 48: $M_{\mu^+\mu^-}$ before cut #6.

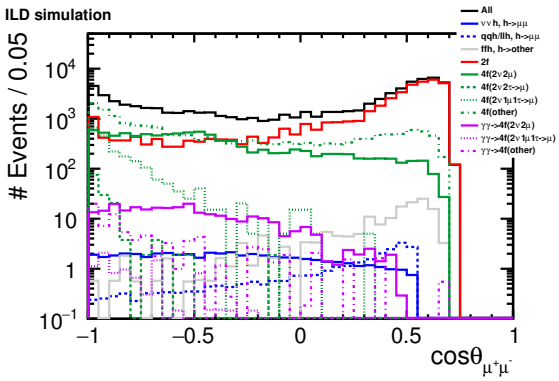

Figure 49: $\cos\theta_{\mu^+\mu^-}$ before cut #7.

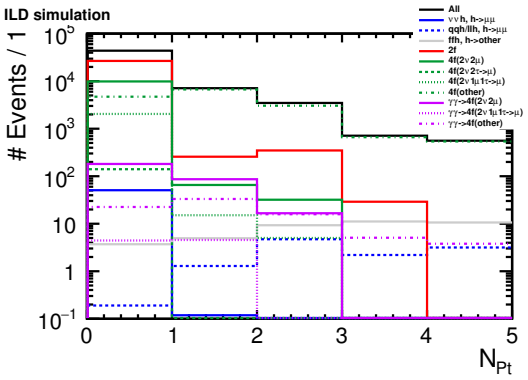

Figure 50: N_{P_t} before cut #8.

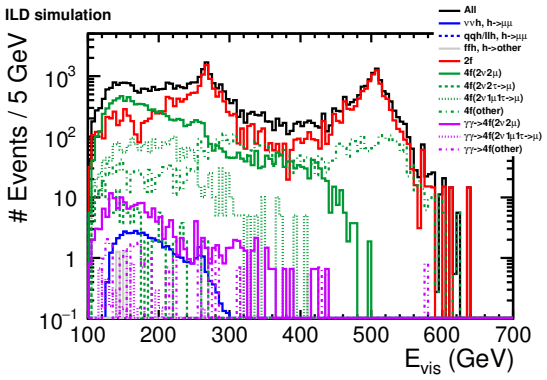

Figure 51: E_{vis} before cut #9.

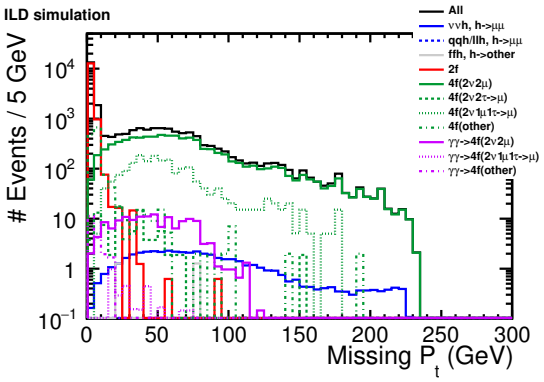

Figure 52: Missing P_t before cut #10.

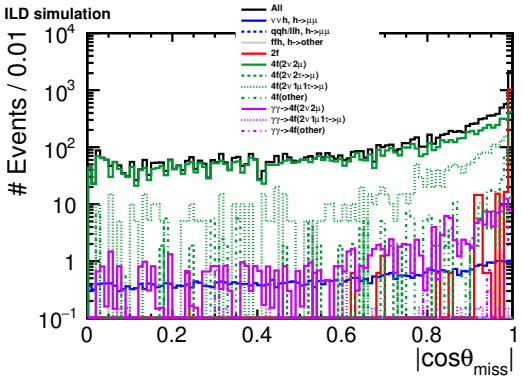

Figure 53: $\cos\theta_{\text{miss}}$ before cut #11.

D. Plots of IDR-S-right

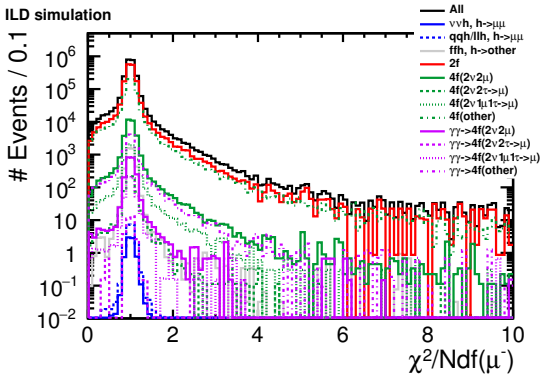

Figure 54: $\chi^2/\text{Ndf}(\mu^-)$ before cut #2.

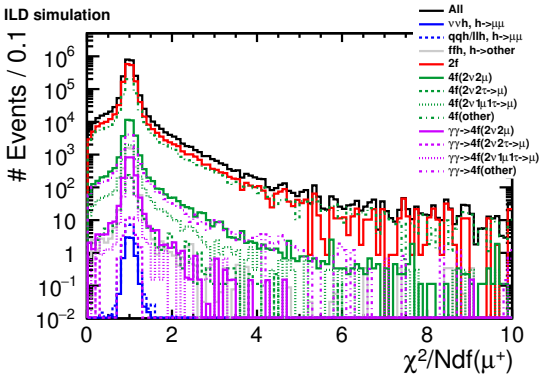

Figure 55: $\chi^2/\text{Ndf}(\mu^+)$ before cut #2.

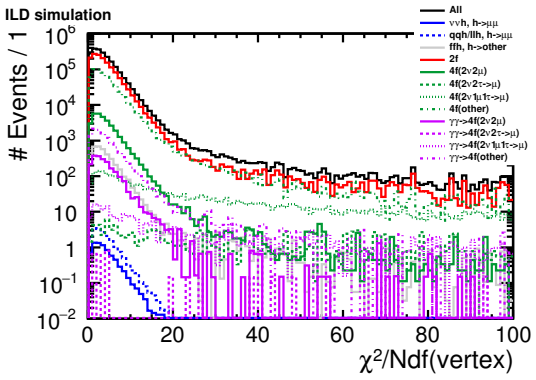

Figure 56: $\chi^2/\mathrm{Ndf(vertex)}$ before cut #3.

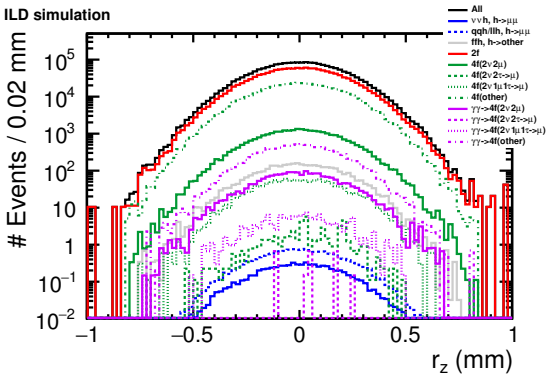

Figure 57: r_z before cut #4.

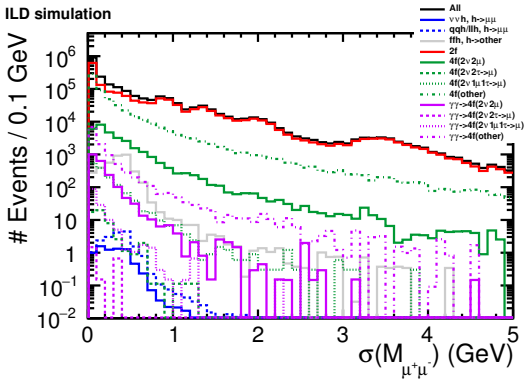

Figure 58: $\sigma(M_{\mu^+\mu^-})$ before cut #5.

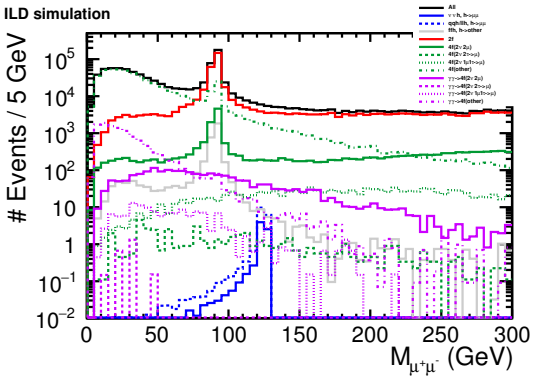

Figure 59: $M_{\mu^+\mu^-}$ before cut #6.

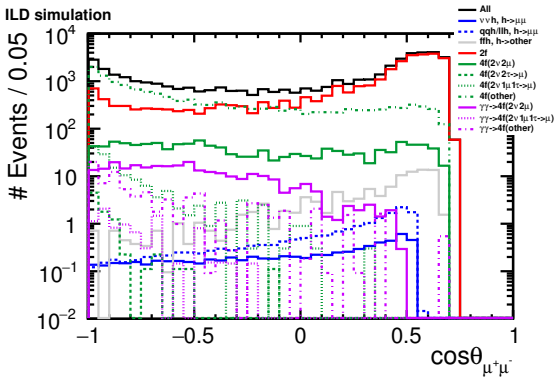

Figure 60: $\cos\theta_{\mu^+\mu^-}$ before cut #7.

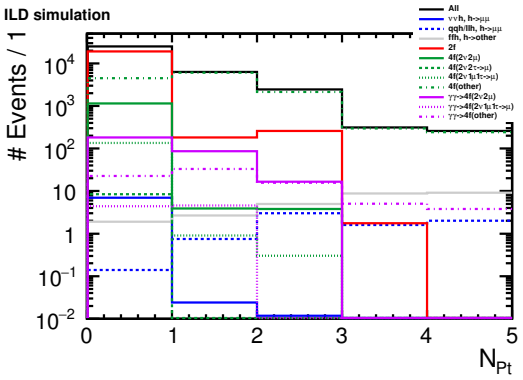

Figure 61: N_{P_t} before cut #8.

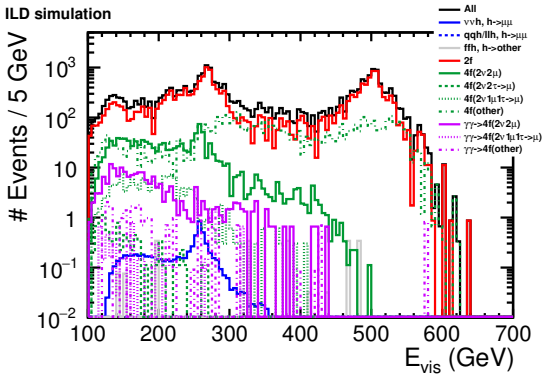

Figure 62: E_{vis} before cut #9.

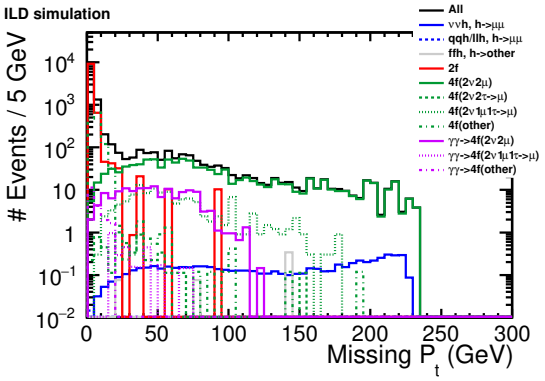

Figure 63: Missing P_t before cut #10.

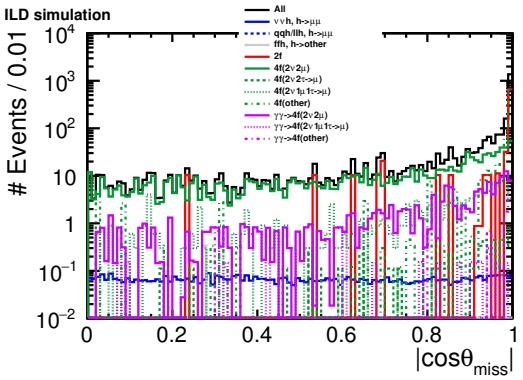

Figure 64: $\cos\theta_{\mathrm{miss}}$ before cut #11.

References

[1] J. List, *News from ILD*, 2018, URL: https://indico-lcagenda-201605.s3.cern.ch/event/7889/contribution/42464/51768-50877-jlist_ild_181022.pdf?response-content-disposition=inline%3B%20filename%3Djlist_ild_181022.pdf&response-content-type=application%2Fpdf&AWSAccessKeyId=4QR6RYTNSYFO7BN3I43N&Expires=1558001345&Signature=FlGHxe6uzSCWhGpVCXkFOKHfuw8%3D.

[2] LCC Physics Working Group, *Physics Case for the International Linear Collider* (2015), arXiv: 1506.05992 [physics.acc-ph].

[3] ILC Parameters Joint Working Group, *ILC Operating Scenarios* (2015), arXiv: 1506.07830 [hep-ex].

[4] Wolfgang Kilian, Thorsten Ohl, Jürgen Reuter, *WHIZARD — simulating multi-particle processes at LHC and ILC*, Eur. Phys. J. **C71** 1742 (2011), DOI: 10.1140/epjc/s10052-011-1742-y.

[5] Ties Behnke, James E. Brau, Philip Burrows, Juan Fuster, Michael Peskin, Marcel Stanitzki, Yasuhiro Sugimoto, Sakue Yamada, Hitoshi Yamamoto, *The International Linear Collider Technical Design Report Volume 4 Detectors* (2013), arXiv: 1306.6329 [physics.ins-det].

[6] Daniel Schulte, *Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider*, PhD thesis, DESY, 1997.

[7] Johann H. Kühn, Zbigniew Wąs, *TAUOLA — a library of Monte Carlo programs to simulate decays of polarized τ leptons*, Comput. Phys. Commun. **64** (1990) 275.

[8] P. Golonka, B. Kersevan, T. Pierzchała, E. Richter-Wąs, Z. Wąs, M. Worek, *The tauola-photos-F environment for the TAUOLA and PHOTOS packages, release II*, Comput. Phys. Commun. **174** (2006) 818.

[9] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Wąs, Z. Wąs, *Universal interface of TAUOLA: Technical and physics documentation*, Comput. Phys. Commun. **183** (2012) 821.

[10] Torbjörn Sjöstrand, Stephen Mrenna, Peter Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026.

[11] GEANT4 Collaboration, *Geant4 — a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[12] M. Frank et al., *AIDASoft/DD4hep*, webpage: http://dd4hep.cern.ch/, 2018, DOI: 10.5281/zenodo.592244, URL: https://doi.org/10.5281/zenodo.592244.

[13] Pisin Chen, Timothy L. Barklow, Michael E. Peskin, *Hadron production in $\gamma\gamma$ collisions as a background for e^+e^- linear colliders*, Phys. Rev. **D49** (1994) 3209.

[14] M. A. Thomson, *Particle flow calorimetry and the PandoraPFA algorithm*, Nucl. Instrum. Meth. **A611** (2009) 25.

[15] F. Gaede, *Marlin and LCCD — Software tools for the ILC*, Nucl. Instrum. Meth. **A559** (2006) 177.

[16] URL: https://github.com/iLCSoft.

[17] Junping Tian, Claude Dürig, *isolated lepton finder*, 2015, URL: https://indico-lcagenda-201104.s3.cern.ch/2015/C6787/c17/7025639686634477308/IsoLep_HLRec2016.pdf?response-content-disposition=inline%3B%20filename%3DIsoLep_HLRec2016.pdf&response-content-type=application%2Fpdf&AWSAccessKeyId=4QR6RYTNSYFO7BN3I43N&Expires=1558001667&Signature=jzP%2FAcr6C8jOQhtAF%2Fbr4Et7df0%3D.

[18] URL: https://github.com/iLCSoft/MarlinReco/tree/master/Analysis/IsolatedLeptonTagging.

[19] Swathi Sasikumar, Jenny List, Mikael Berggren, *Hadron Production in Photon-Photon Processes at the ILC and BSM signatures with small mass differences*, 2018, URL: https://indico-lcagenda-201605.s3.cern.ch/event/7839/contribution/40928/49752-48873-Swathi_PreILD.pdf?response-content-disposition=inline%3B%20filename%3DSwathi_PreILD.pdf&response-content-type=application%2Fpdf&AWSAccessKeyId=4QR6RYTNSYFO7BN3I43N&Expires=1558000557&Signature=lgRIO%2B%2FcNZ7AvoWi8wT%2Btl14D9E%3D.

[20] D. Bailey *et al.*, *The LCFIVertex package: Vertexing, flavour tagging and vertex charge reconstruction with an ILC vertex detector*, Nucl. Instrum. Meth. **A610** (2009) 573.

[21] P. Speckmayer, A. Höcker, J. Stelzer, H. Voss,
The Toolkit for Multivariate Data Analysis, TMVA 4, J. Phys. Conf. Ser. **219** (2010) 032057.

[22] Rene Brun, Fons Rademakers, *ROOT – An Object Oriented Data Analysis Framework*, 1997, URL: http://root.cern.ch/.

[23] W. Verkerke, D. Kirkby, *The RooFit toolkit for data modeling* (2003), arXiv: 0306116.