ILD-PHYS-PUB–2019-002
2019/June

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 benchmarks, with respect to the impact of the transverse momentum resolution of the reconstructed muons. We study the prospects for 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 somewhat better result of $\sim 2.8\%$ gain in relative precision on $\sigma \times BR(h \to \mu^+\mu^-)$. This difference is caused by the better momentum resolution of IDR-L in the barrel region, resulting in less background in the peak region of $M_{\mu^+\mu^-}$ distribution.

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 benchmarks 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 at 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 $h \to \mu^+\mu^-$ is studied at $\sqrt{s} = 500$ GeV. 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 configurations, an integrated luminosity of 1.6 ab^{-1} is assumed based on the assumed running scenario at $\sqrt{s} = 500$ GeV [2, 3].

It should be noted that in this process, two signal processes (Zh process with $Z \to \nu\bar{\nu}$ and WW-fusion process) are interfering each other. The relative contributions of these production modes will be fixed to the percent-level or better from other Higgs decay modes like $h \to b\bar{b}$, and can be used to convert the cross section times branching ratio measurement into a measurement of $\text{BR}(h \to \mu^+\mu^-)$. With the help of the total Zh cross section determined with the recoil method, the absolute $h\mu\mu$ Yukawa coupling can be extracted. Therefore, we will not consider the separation of Zh process and WW-fusion process in this note.

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 in different spectrum of muon pair invariant mass. We will study such variations. In order to specify the combination of beam polarization and detector configuration, we use abbreviations as listed in Table 1.

Table 1: Abbreviations used 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 has been performed with `Whizard1.95` [4] which is centrally produced for DBD [5], taking ISR effect into account. The beamstrahlung effect has been modelled by `GuineaPig` [6]. `Tauola` [7–9] has been used for τ decay, and `Pythia` [10] has been used for decaying short-lived particles and hadronization. The full detector simulation based on `Geant4` [11] has been performed under `DD4HEP` [12] framework. Events have been 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 used in this analysis are 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 (first line in the Table 2). 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 $e^+e^- \to$ 2-/4-fermion processes and $\gamma\gamma \to$ 4-fermion processes with at least one lepton ℓ in the final state. The processes which only contain jets are not considered because their signatures are dramatically different from the signal, and they are expected to be easily suppressed. The processes of $\gamma\gamma \to$ 2-fermion are not included, due to its huge cross section, there are no fully-simulated MC samples. 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` (signal) `higgs_ffh` (background)
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 use `ILCSoft` [16] (version v02-00-02) for the analysis. In the first analysis step, events with a pair of muons are selected as the $h \to \mu^+\mu^-$ candidates. These muons are subjected to the fitting to estimate the primary vertex position for the event. 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 used to estimate the expected 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`; neglecting impact parameter information and energy deposit in the yoke, because the MC samples described in Section 2 are now having smearing effect at the IP [19]. 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 muon signature, 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 for the signal process as $N_{\mathrm{rec}}/N_{\mathrm{tot}}$, where N_{tot} is the total number of events, and N_{rec} is the number of events with exactly one μ^+ and one μ^- reconstructed. This efficiency is calculated to be $96.0 \pm 0.8\%$ in IDR-L and $95.7 \pm 0.8\%$ in IDR-S. There are no differences within the statistical uncertainty. Thus, the impact of the detector configuration on the isolated muon identification is small.

3.2. Primary Vertex Finding

The $h \to \mu^+\mu^-$ candidate is subjected to the primary vertex finding. In the current MC samples described in Section 2, the IP smearing effect is considered, resulting the primary vertex position uncertainty in z-direction of $\sim 200\,\mu$m (smearing is also applied in x-/y-directions, 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 as described 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 $(x,y,z) = (150\,\mathrm{nm}, 5\,\mathrm{nm}, 200\,\mu\mathrm{m})$. Figure 1 shows the distribution of primary vertex z-position (r_z) for signal process in IDR-L and IDR-S, together with Gaussian fitting result. Due to very small transverse beam-spot sizes, the Gaussian fit is only applied to determine the longitudinal position of the primary vertex.

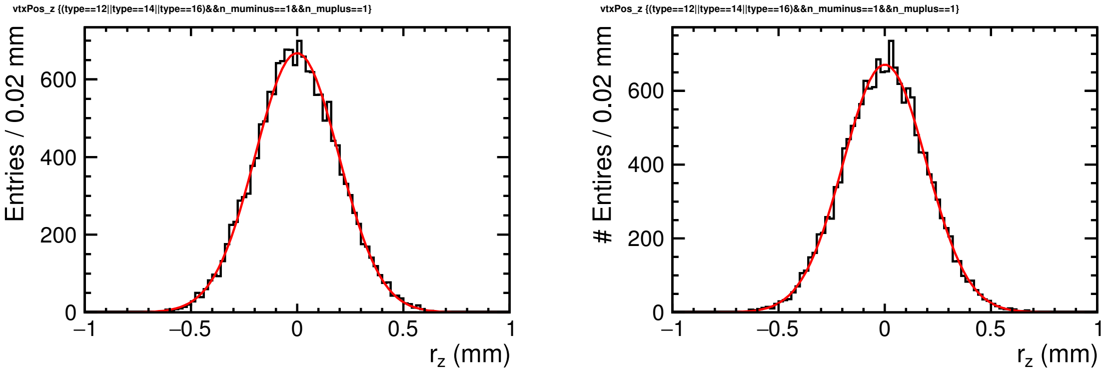

Figure 1: Estimated z-position of primary vertex r_z for the signal process, together with the result of Gaussian fitting in red curve. Left: IDR-L (Gaussian width $= 0.1964 \pm 0.0011$ mm). Right: IDR-S (Gaussian width $= 0.1949 \pm 0.0011$ mm).

3.3. Preselection

We apply several selection cuts to select signal and suppress background. The selection cuts are summarized in Table 4, where $\chi^2/\mathrm{Ndf}(\mu^\pm)$ is the reduced-χ^2 of the muon track fitting, $\chi^2/\mathrm{Ndf}(\mathrm{vertex})$ is the fitting quality of primary vertex position, $\sigma(M_{\mu^+\mu^-})$ is the event-by-event resolution of muon pair invariant mass, $M_{\mu^+\mu^-}$, which is calculated using covariance matrix of muons in momenta space, $\theta_{\mu^+\mu^-}$

is the angle between two muons in the laboratory frame, N_{P_t} is the number of charged particles with the transverse momentum greater than 5 GeV except $h \to \mu^+\mu^-$ candidate, E_{vis} is the visible energy, and θ_{miss} is the angle of the missing momentum with respect to the beam axis, respectively.

Table 4: List of preselection cuts. Definition of variables are given in the text.

#	variable	cut
1	# $\mu^\pm$	$= 1$
2	$\chi^2/\text{Ndf}(\mu^\pm)$	$0.5 - 1.5$
3	$\chi^2/\text{Ndf}(\text{vertex})$	< 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_{vis}	$125 - 300$ GeV
10	missing P_t	> 5 GeV
11	$\lvert \cos\theta_{\text{miss}} \rvert$	< 0.99

The cuts #2 and #5 are used to require well-measured muon tracks. The cuts #3 and #4 are selecting prompt muons, and the cuts #6 and #7 require the signature of $h \to \mu^+\mu^-$ candidate. The cut #8 is selecting the events with no charged particles except $h \to \mu^+\mu^-$ candidate and low P_t hadrons from $\gamma\gamma$ overlay. The last three cuts, #9 to #11, are used to reject neutrino-less events. We have applied these selection cuts to all analysis channels. These cuts could be optimized better, if considered separately for left-handed and right-handed beam polarization, because the main production processes of Higgs boson are different for different beam polarization configurations. The distributions of each variable before each cut for IDR-L-left configuration are shown in the Appendix. Table 5 shows the numbers of expected events after the subsequent cuts for IDR-L-left configuration as an example. Similar tables for different configurations are shown in the Appendix. In the table, the irreducible 4-fermion background processes are summarized in dedicated columns; $2\nu2\mu$ 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 as Table 6.

Table 5: Number of events for 1.6 ab^{-1} data, before any cuts (#0) and after each individual cut enumerated with #1 - #11 in IDR-L-left configuration.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	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 6: A simplified cut table giving the number of events after preselection normalized to 1.6 ab^{-1} for each analysis channel. Numbers in parentheses 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 all preselection cuts, 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 6, we see no major differences between IDR-L and IDR-S. However, we see some difference in the distribution of variables. Figure 2 shows the distribution of the muon pair invariant mass $M_{\mu^+\mu^-}$ and event-by-event mass resolution $\sigma(M_{\mu^+\mu^-})$ after preselection for IDR-L-left and IDR-S-left.

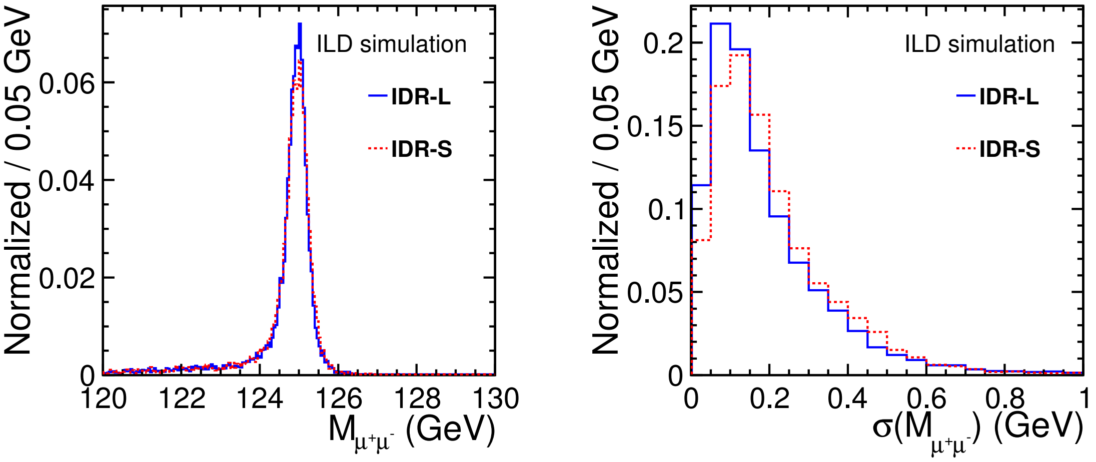

Figure 2: Distribution of variables after preselection for IDR-L-left and IDR-S-left. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

We can clearly see the difference between IDR-L and IDR-S options reflecting the detector effects. 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 3 shows the similar distribution as Figure 2, but for events with both muons in the barrel region ($|\cos\theta_{\mu^\pm}| < 0.7$, where $\theta_{\mu^\pm}$ is the angle of $\mu^\pm$ with respect to the beam axis). In Figure 3, 43.8% events after preselection are plotted in both detector models. In this case, IDR-L gives significantly better results than IDR-S. On the other hand, Figure 4 shows the same distribution, but when both muons are required to be in the endcap region and/or forward region ($|\cos\theta_{\mu^\pm}| > 0.7$). In Figure 4, 8.4% events are plotted in both detector models. Due to the superior momentum resolution in forward region, as expected, IDR-S performs better than IDR-L.

The remaining mixed case (one muon in barrel, another muon in endcap/forward) is shown in Figure 5. The $M_{\mu^+\mu^-}$ looks pretty similar, while we see inconclusive difference in $\sigma(M_{\mu^+\mu^-})$ distribution.

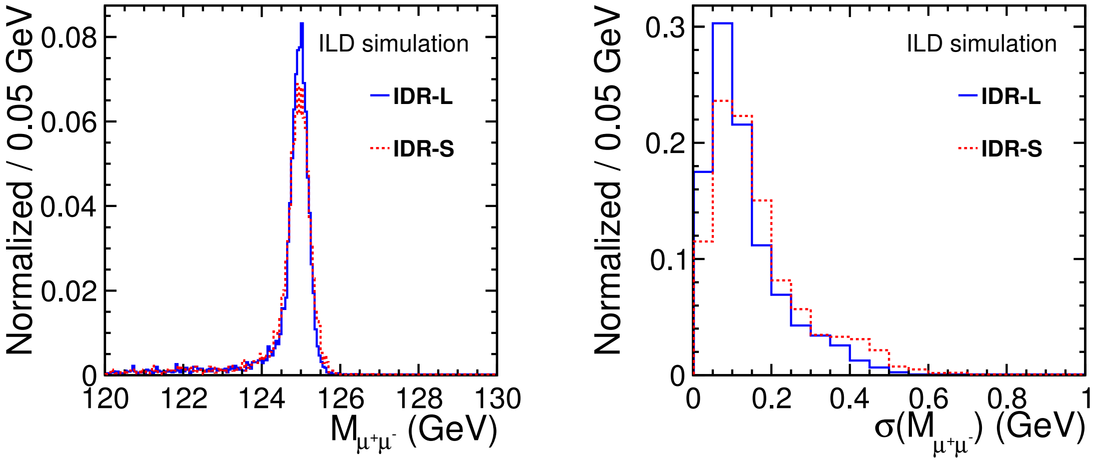

Figure 3: Similar to Figure 2, but an additional cut $|\cos\theta_{\mu^\pm}| < 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

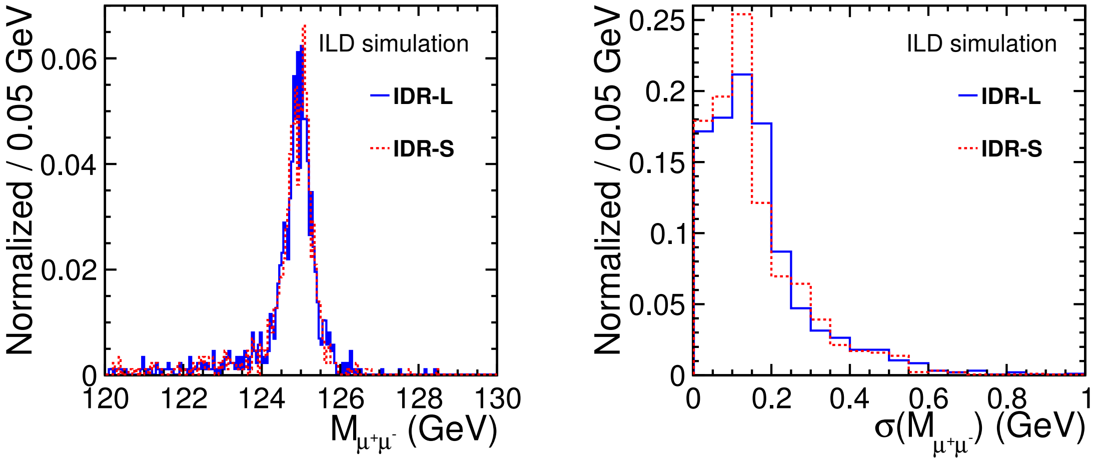

Figure 4: Similar to Figure 2, but an additional cut $|\cos\theta_{\mu^\pm}| > 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

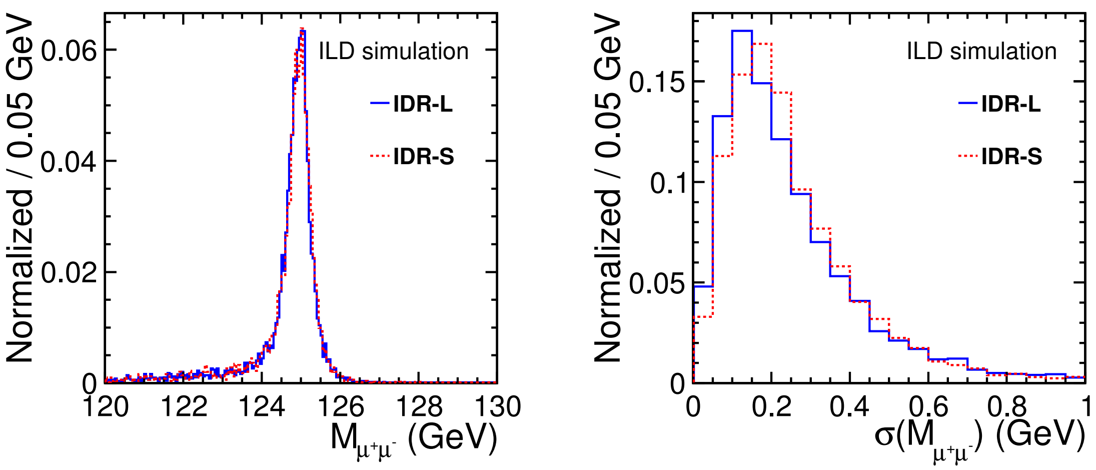

Figure 5: Similar to Figure 2, but additional cuts are applied to require $|\cos\theta_{\mu^\pm}| < 0.7$ for one muon, and $|\cos\theta_{\mu^\pm}| > 0.7$ for another muon. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

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: $\cos\theta_{\mu^+\mu^-}$, $\cos\theta_{\mu^+} - \cos\theta_{\mu^-}$ (charge of muon times its $\cos\theta$), E_{lead}, E_{sub}, $\cos\theta_{\text{lead}}$, and $\cos\theta_{\text{sub}}$, where $\theta_{\mu^i} (i = +,-)$ is the polar angle of μ^i with respect to the beam axis, $E_{\text{lead}}(\theta_{\text{lead}})$ is the energy(angle) of leading muon in $h \to \mu^+\mu^-$ candidate, and $E_{\text{sub}}(\theta_{\text{sub}})$ is the energy(angle) of subleading muon 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 showing signal and irreducible background, each histogram is normalized to 1). Large fluctuations in the background plots are due to lack of MC statistics for SM background. In all analysis channels, we use the same variables for the input to BDTG.

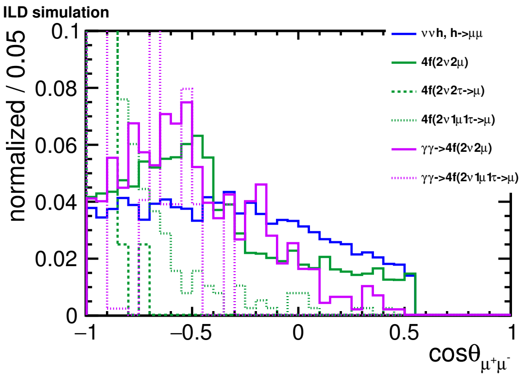

Figure 6: Distribution of $\cos\theta_{\mu^+\mu^-}$ (angle between two muons).

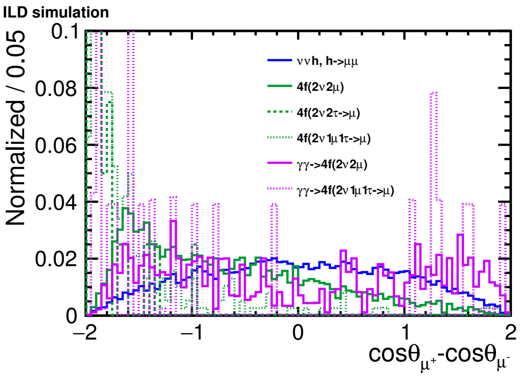

Figure 7: Distribution of $\cos\theta_{\mu^+} - \cos\theta_{\mu^-}$ (charge of muon times its $\cos\theta$).

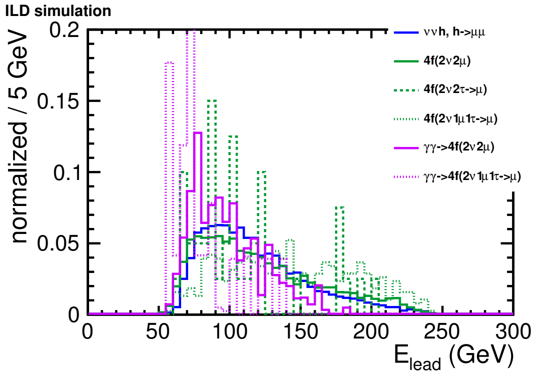

Figure 8: Distribution of E_{lead} (energy of leading muon).

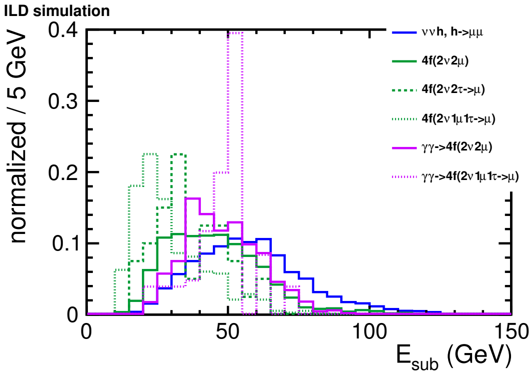

Figure 9: Distribution of E_{sub} (energy of subleading muon).

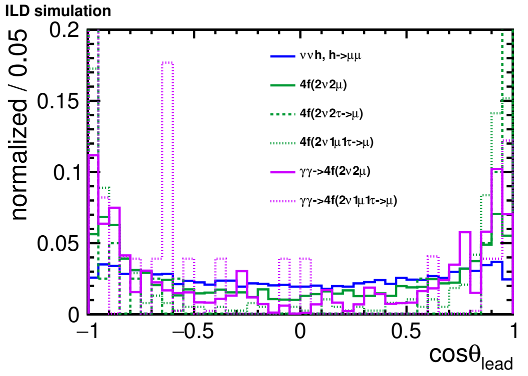

Figure 10: Distribution of $\cos\theta_{\text{lead}}$ ($\cos\theta$ of leading muon).

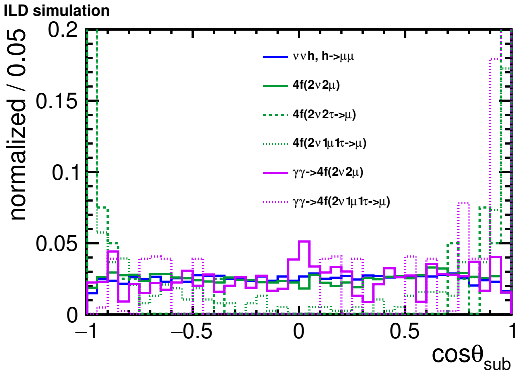

Figure 11: Distribution of $\cos\theta_{\text{sub}}$ ($\cos\theta$ of subleading muon).

The BDTG score for signal and background in IDR-L-left is shown in Figure 12. Optimizing the cut point of BDTG score will discuss in next section.

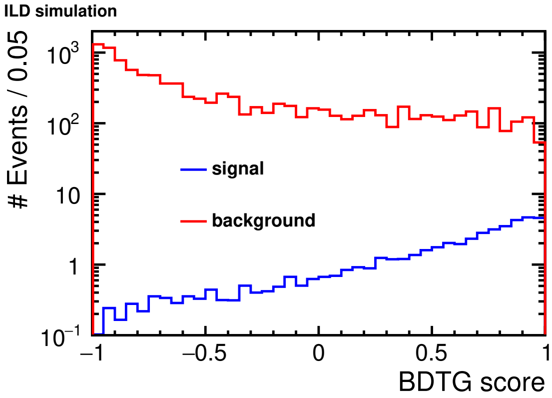

Figure 12: Distribution of BDTG score in IDR-L-left.

3.5. Toy MC

To extract the expected precision of $\sigma \times \mathrm{BR}(h \to \mu^- \mu^+)$, we perform toy MC experiments as the last step of the analysis. We use RooFit [23] for toy MC generation. We use $M_{\mu^+\mu^-}$ distribution of test samples after all cuts including a cut on BDTG score for defining the reference distribution 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 at 125 GeV, the mass of the Higgs boson. A first order polynomial is used as the background modeling function f_B, because we expect almost flat distribution for the background processes after all cuts. Figure 13 shows the result of modeling functions fit for signal and background in IDR-L-left.

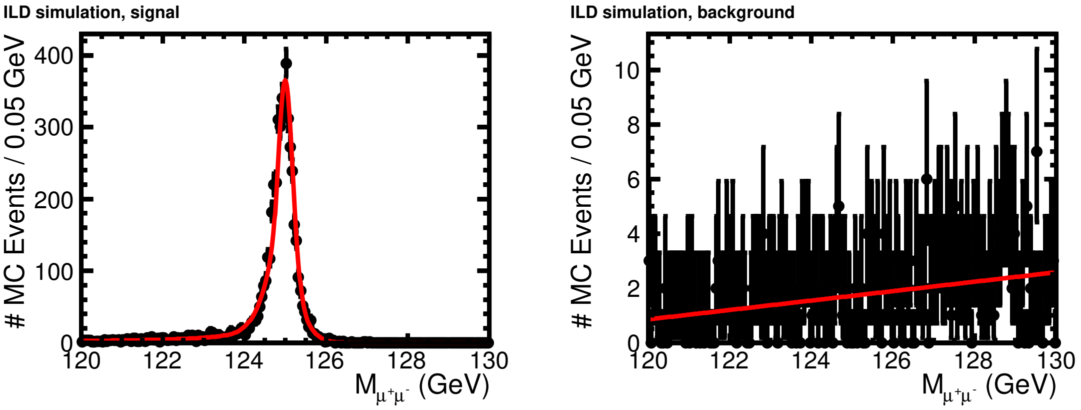

Figure 13: Result of fitting with modeling function after applying all cuts in IDR-L-left. An optimum cut on BDTG score is already applied (see Table 7). Left: signal fitting with f_S. Right: background fitting with f_B.

After fitting the modeling, we then perform pseudo-experiments using f_S and f_B functions. In one pseudo-experiment, the number of pseudo-signal(-background) events is generated from the number of signal(background) entries after all cuts plus assuming purely Poisson fluctuations. 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 14 left shows one example of pseudo-experiment in IDR-L-left. We repeat this pseudo-experiments 50000 times for each configuration, and obtain final Y_S distribution as shown in Figure 14 right. The expected precision is estimated with a Gaussian fit to the Y_S distribution.

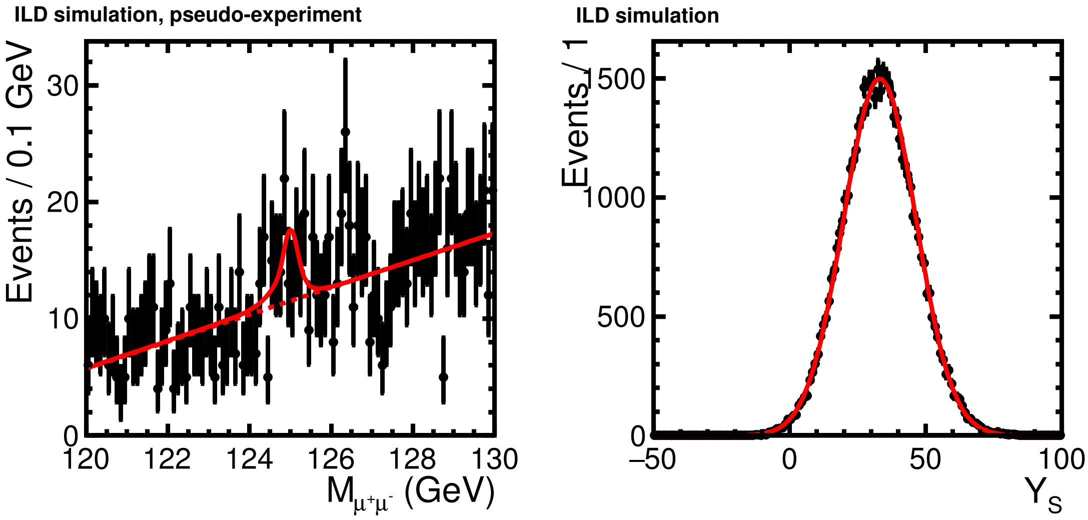

Figure 14: Result of pseudo-experiment in IDR-L-left. Left: one example of pseudo-experiment. Black dots are the pseudo-data, solid red curve is the result of the fit of function f, and dotted-red is its background component $Y_B f_B$. Right: Y_S distribution after 50000 pseudo-experiments together with a Gaussian fit shown as a red curve. The width is estimated to be 13.315 ± 0.042 and mean is 33.155 ± 0.060.

An optimization to determine the cut on BDTG score is performed using the results of toy MC. We repeat the toy MC procedure described above for different values of BDTG score cut. We select the optimal BDTG cut as the one which gives the best measurement precision for given configuration. Figures 13 and 14 correspond to the optimal BDTG score cut in IDR-L-left configuration.

4. Results and Discussion

Table 7 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 $\sigma \times \text{BR}(h \to \mu^+\mu^-)$ from toy MC experiment in Table 8, together with theoretical precision limit which is calculated by assuming 100% signal selection efficiency, no backgrounds, and no detector effects.

Table 7: A simplified cut table giving the number of events normalized to 1.6 ab^{-1} after all cuts for each analysis channel. Numbers in parentheses show the signal selection efficiency. The irreducible processes are defined in the Section 3.3.

	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 8: Expected 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 given in the last column.

	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, results of our analysis can be summarized as following.

- The Higgs related backgrounds can be completely suppressed.
- The signal-to-background ratio is $\sim 1/30 - 1/40$, and irreducible background is a major show-stopper for the precise measurement.
- Our analysis is about a factor of 3 above from the theoretical limit. There are several reasons; imperfection of cuts, existence of irreducible background mainly originating from $e^+e^- \to W^+W^- \to 2\mu 2\nu$.
- Measurements with the right-handed polarization have limited precision due to very small number of signal events. For further discussion, we will not consider this polarization anymore.
- We see some difference in the expected number of remaining background events, but this is consistent with a statistical fluctuation due to lack of MC statistics for SM background. Since we do not expect any difference on the distribution of background after all cuts, we use same parametrization for background for IDR-L and IDR-S. This point will be discussed at Section 4.2.

4.1. Detector Effect

Following the approach described in Section 3.3.1, we compare expected performance for IDR-L and IDR-S for $M_{\mu^+\mu^-}$ and $\sigma(M_{\mu^+\mu^-})$ after all selection cuts including an optimized cut on the BDTG score. Figure 15 shows the corresponding distributions for these variables. Overall, IDR-L has better performance than IDR-S.

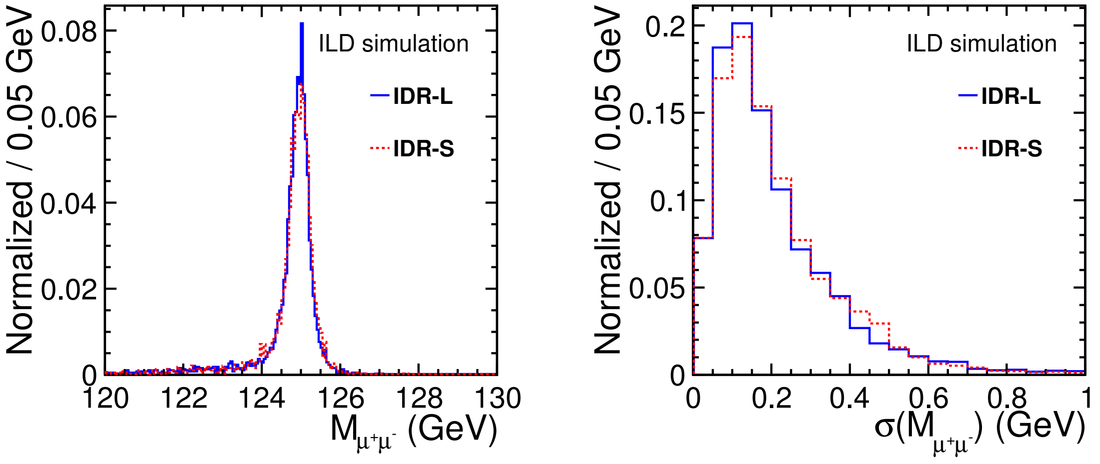

Figure 15: Distribution of variables after BDTG cut for IDR-L-left and IDR-S-left. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

Again, we will look at the angular dependency of these two variables. Figure 16 shows the distribution with both muons are in barrel region (46.2%/47.0% events selected after all cuts are applied with IDR-L/IDR-S configurations). Figure 17 shows in non-barrel region (5.5%/5.3% events after all cuts are plotted with IDR-L/IDR-S). The rest of mixed case is plotted in Figure 18. We have confirmed that the same tendency is observed which discussed in Section 3.3.1.

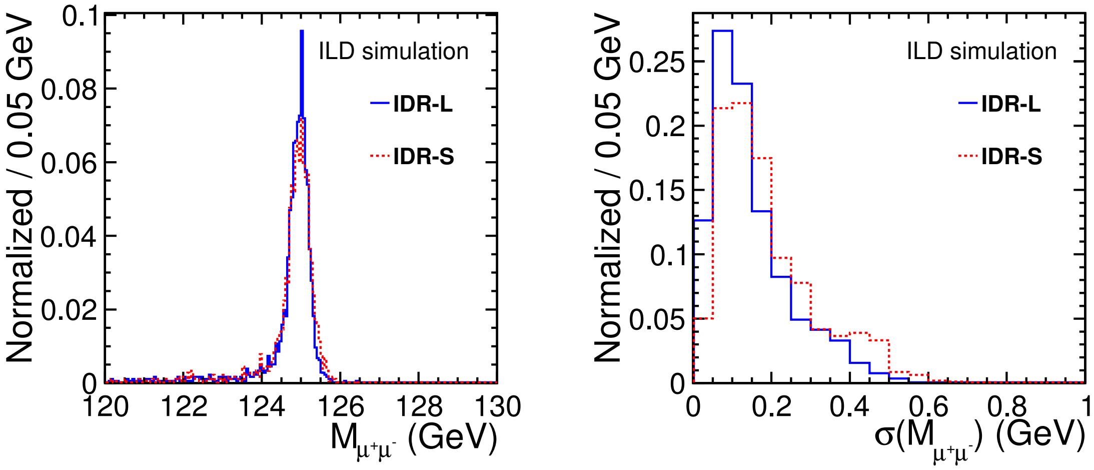

Figure 16: Similar to Figure 15, but an additional cut $|\cos\theta_{\mu^\pm}| < 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

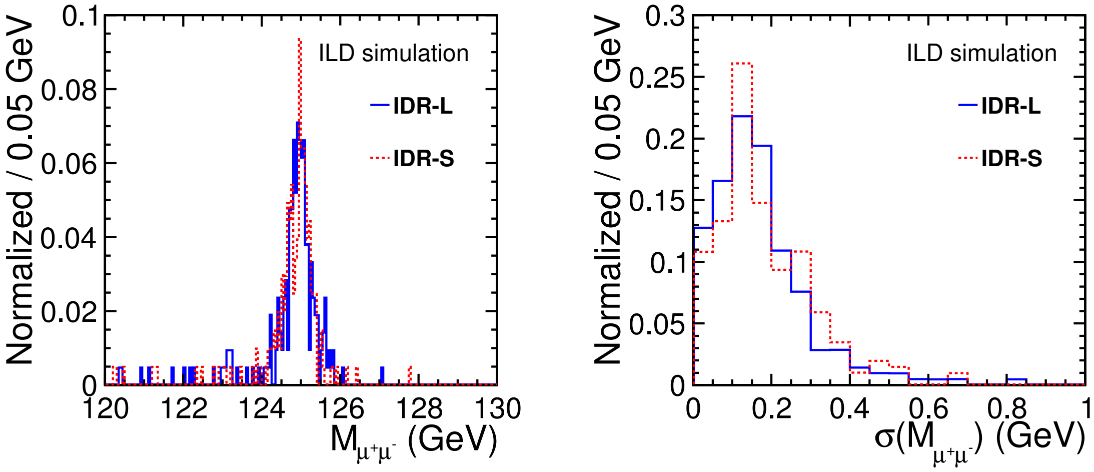

Figure 17: Similar to Figure 15, but an additional cut $|\cos\theta_{\mu^\pm}| > 0.7$ is applied. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

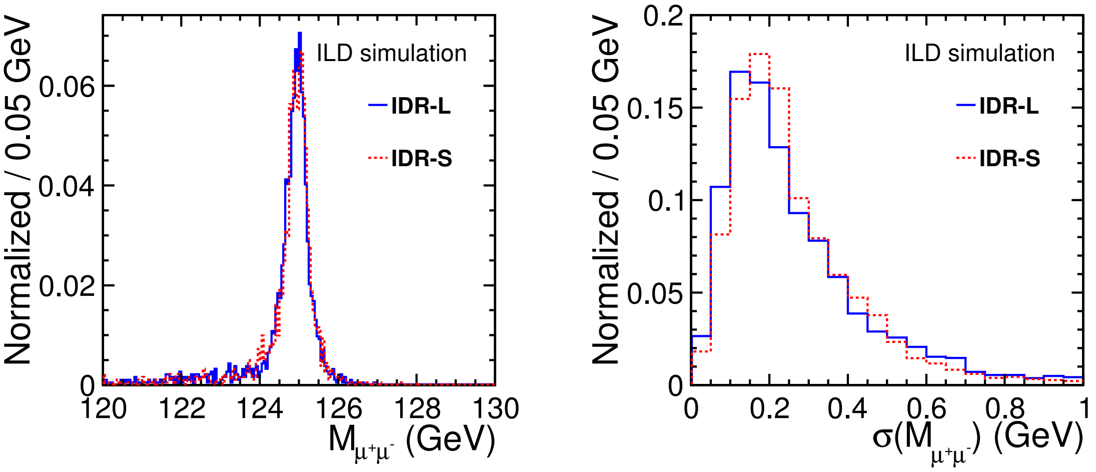

Figure 18: Similar to Figure 15, but additional cuts are applied to require $|\cos\theta_{\mu^\pm}| < 0.7$ for one muon, and $|\cos\theta_{\mu^\pm}| > 0.7$ for another muon. Left: $M_{\mu^+\mu^-}$. Right: $\sigma(M_{\mu^+\mu^-})$. All histograms are normalized to 1.

4.2. Final Results

As we discussed in Section 4, we do not expect any difference in the background distribution for two detector configurations, and as the current differences are consistent with statistical fluctuations. We therefore use the same background 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 the first order polynomial parametrization which was discussed in Section 3.5. We perform toy MC with common treatment for the background, and Table 9 shows the final results.

Table 9: Expected 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 given in the last column.

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

The IDR-L gives somewhat better result of $\sim 2.8\%$ gain in relative precision of the $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ than IDR-S. This is related to the different distribution on $M_{\mu^+\mu^-}$ between IDR-L and IDR-S (see Figure 15). In the signal modeling, the width of the Crystal Ball function is $\sim 10\%$ wider in IDR-S. We have checked that the number of signal events in the peak region is almost same between IDR-L (26.0 events) and IDR-S (26.2 events). The $\sim 10\%$ increase in width gives $\sim 10\%$ more background events in the peak region for IDR-S (111.2 backgrounds, for IDR-L it is 102.6 backgrounds). From the statistical point of view, 10% more backgrounds give $\sim 2.6\%$ worse significance. Our relative 2.8% difference is consistent to this simple statistical 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 two different 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 for IDR-S-left it is $41.28 \pm 0.15\%$, leading to a relative 2.8% difference in $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ uncertainties in favor of IDR-L. This difference comes from different transverse momentum resolution σ_{1/P_t} in barrel region. The worse resolution results in $\sim 10\%$ wider width of muon pair invariant mass with IDR-S, resulting in more backgrounds in the peak region of $M_{\mu^+\mu^-}$ which essentially determines the final precision of the $\sigma \times \mathrm{BR}(h \to \mu^+\mu^-)$ measurement.

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 this appendix, we will show the distributions of each variable which are used in the preselection, for IDR-L-left configuration as an example. 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

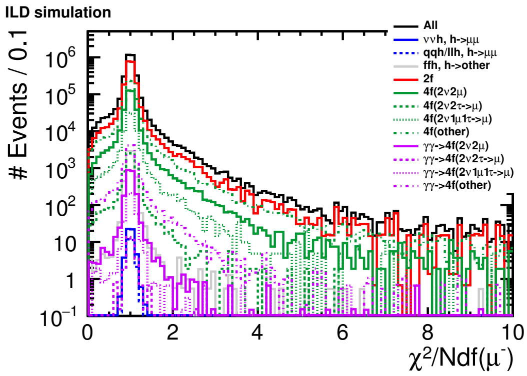

Figure 19: χ^2/Ndf(μ^-) before cut #2.

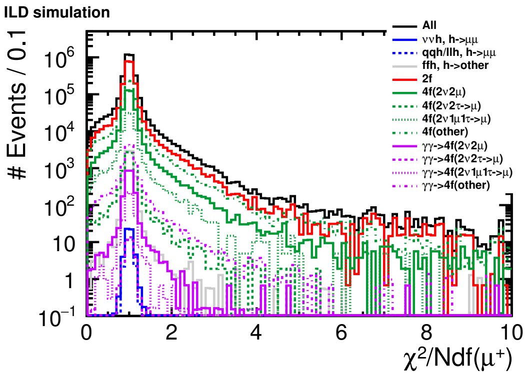

Figure 20: χ^2/Ndf(μ^+) before cut #2.

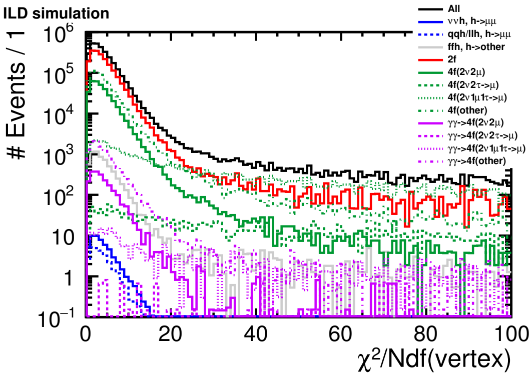

Figure 21: $\chi^2/\mathrm{Ndf(vertex)}$ before cut #3.

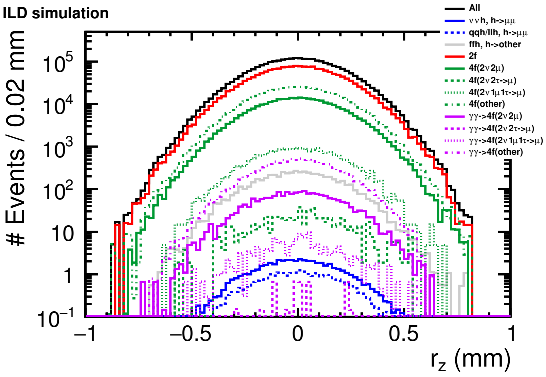

Figure 22: r_z before cut #4.

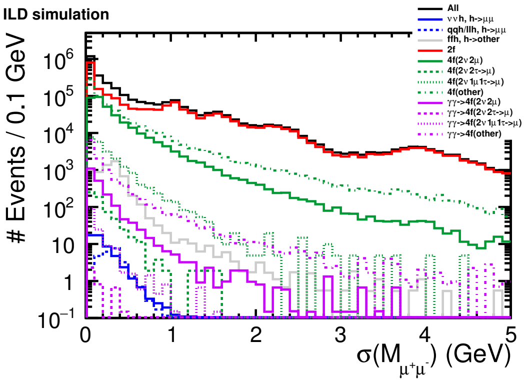

Figure 23: $\sigma(M_{\mu^+\mu^-})$ before cut #5.

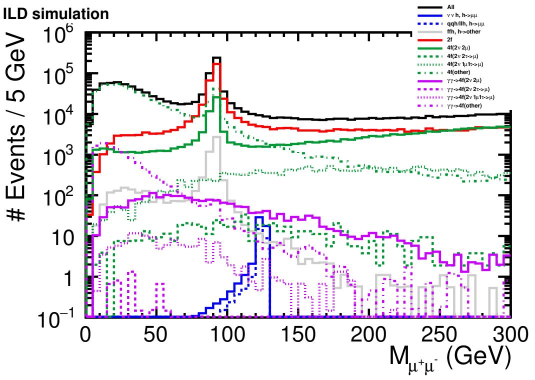

Figure 24: $M_{\mu^+\mu^-}$ before cut #6.

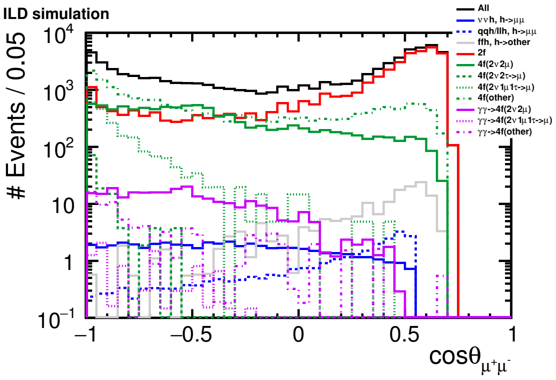

Figure 25: $\cos\theta_{\mu^+\mu^-}$ before cut #7.

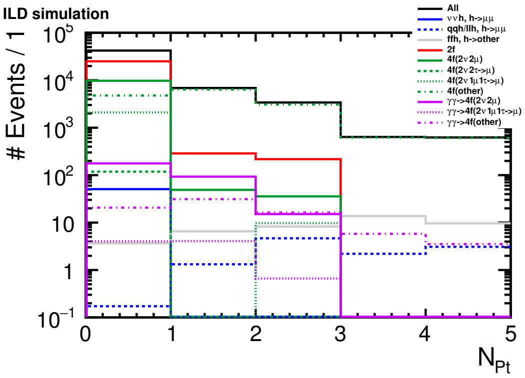

Figure 26: N_{P_t} before cut #8.

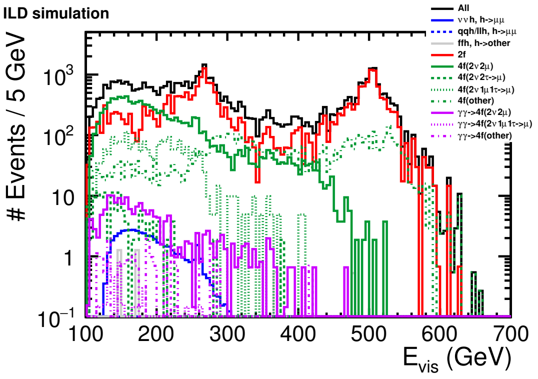

Figure 27: E_{vis} before cut #9.

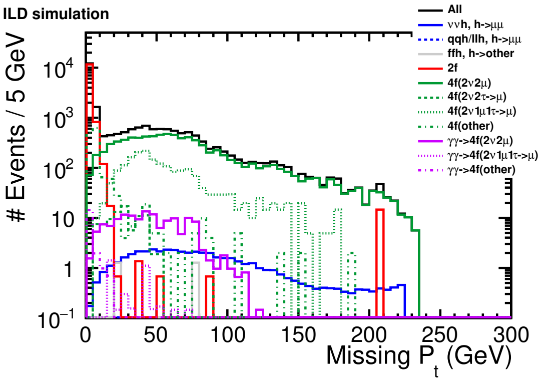

Figure 28: Missing P_t before cut #10.

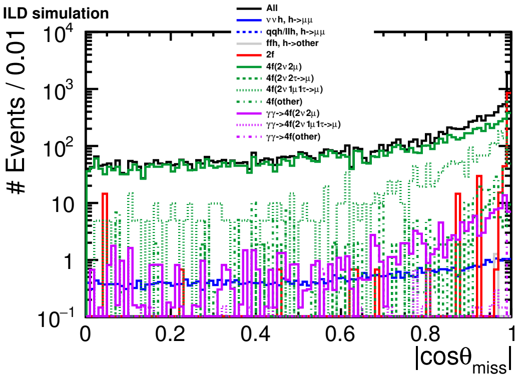

Figure 29: $\cos\theta_{\mathrm{miss}}$ before cut #11.

B. Cut tables

We will show similar cut tables like Table 5, but for different configurations in this appendix.

Table 10: Number of events for 1.6 ab^{-1} data, before any cuts (#0) and after each individual cut enumerated with #1 - #11 in IDR-L-right configuration.

#	$\nu\nu h$	$qqh/\ell\ell h$	ffh	2f	4f	4f	4f	4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f
	$h \to \mu\mu$	$h \to \mu\mu$	other		$2\nu 2\mu$	$2\nu 2\tau(\mu)$	$2\nu\mu\tau(\mu)$	other	$2\nu 2\mu$	$2\nu 2\tau(\mu)$	$2\nu\mu\tau(\mu)$	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 11: Number of events for 1.6 ab^{-1} data, before any cuts (#0) and after each individual cut enumerated with #1 - #11 in IDR-S-left configuration.

#	$\nu\nu h$ $h \to \mu\mu$	$qqh/\ell\ell h$ $h \to \mu\mu$	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 12: Number of events for 1.6 ab^{-1} data, before any cuts (#0) and after each individual cut enumerated with #1 - #11 in IDR-S-right configuration.

#	$\nu\nu h$	$qqh/\ell\ell h$	ffh	2f	4f	4f	4f	4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f	$\gamma\gamma \to$ 4f
	$h \to \mu\mu$	$h \to \mu\mu$	other		$2\nu 2\mu$	$2\nu 2\tau(\mu)$	$2\nu\mu\tau(\mu)$	other	$2\nu 2\mu$	$2\nu 2\tau(\mu)$	$2\nu\mu\tau(\mu)$	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

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.