

# Discovery potential of Higgs boson pair production through $4\ell + \cancel{E}$ final states at a 100 TeV collider <sup>\*</sup>

Xiaoran Zhao(赵笑然)<sup>1;1)</sup> Qiang Li(李强)<sup>2;2)</sup> Zhao Li(李钊)<sup>3;3)</sup> Qi-Shu Yan(晏启树)<sup>4,5;4)</sup>

<sup>1</sup> Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, Chemin du Cyclotron, 2, B-1348 Louvain-la-Neuve, Belgium

<sup>2</sup> Department of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University Beijing 100871, China

<sup>3</sup> Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

<sup>4</sup> School of Physics Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>5</sup> Center for Future High Energy Physics, CAS, Beijing 100049, China

**Abstract:** We explore the discovery potential of Higgs pair production at a 100 TeV collider via full leptonic mode. The same mode can be explored at the LHC when Higgs pair production is enhanced by new physics. We examine two types of fully leptonic final states and propose a partial reconstruction method, which can reconstruct some useful kinematic observables. It is found that the  $m_{T2}$  variable determined by this reconstruction method and the reconstructed visible Higgs mass are crucial to discriminate the signal and background events. It is also noticed that a new variable, denoted as  $\Delta m$ , which is defined as the mass difference of two possible combinations, is very useful as a discriminant. To examine the detector effects, we consider seven detector setups for a 100 TeV collider and investigate the changes in the sensitivity, and we find that lepton isolation and the minimal lepton  $P_{\perp}$  cut are crucial in order to reduce the integrated luminosity.

**Keywords:** Higgs pair, hadron collider, leptonic mode

**PACS:** 13.85.Qk, 14.80.Bn **DOI:** 10.1088/1674-1137/41/2/023105

## 1 Introduction

The discovery of the Higgs boson at the LHC has motivated the high energy community to think of the next generation of pp colliders. A 100 TeV collider can offer us huge potential to probe various forms of new physics [1]. For example, new vector bosons  $W'$  and  $Z'$  could be discovered up to 25–35 TeV [2]. A heavy Higgs bosons of the two Higgs doublet model could be probed up to 20 TeV or so via single associated production [3]. In the simplified model, the superpartners of the top quark and the gluino could be probed up to 5 TeV and 10 TeV, respectively [4], which would give decisive evidence on the fate of the electroweak supersymmetry models. Dark matter candidates could be probed up to 10 TeV or higher [5–8]. A 100 TeV collider could also perform high precision

measurements of Higgs properties [9], top quark properties, EW physics, and so on. The next generation pp colliders would be complementary to the Higgs factories, either in the form of CEPC, or FCC-ee, or ILC, where the mass generation mechanism of the standard model (SM), i.e. the couplings of the Higgs boson to the rest of the SM particles, can be precisely measured [10].

Among the various new physics candidates, the shape of the Higgs potential plays a very special role. As we know, the shape of the Higgs potential is determined by Higgs fields and Higgs self-couplings. New Higgs fields and new self-couplings are common in most extensions of the SM. Therefore, it is well-known [11–16] that to probe Higgs self-couplings at colliders can offer us a way to understand the nature of Higgs bosons, to reconstruct the shape of Higgs potentials, and to understand the mys-

Received 3 September 2016

<sup>\*</sup> Supported by Natural Science Foundation of China (11175251, 11305179, 11675185, 11475180, 11575005) The work of Q. Li and Q.S. Yan is partially supported by CAS Center for Excellence in Particle Physics (CCEPP), X. Zhao is partially supported by the European Union as part of the FP7 Marie Curie Initial Training Network MCnetITN (PITN-GA-2012-315877).

1) E-mail: xiaoran.zhao@uclouvain.be

2) E-mail: qliphy@gmail.com

3) E-mail: zhaoli@ihep.ac.cn

4) E-mail: yanqishu@ucas.ac.cn



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Article funded by SCOAP<sup>3</sup> and published under licence by Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd

tery of electroweak symmetry breaking. These couplings could play crucial parts in the EW baryogenesis scenarios [17, 18]. For example, they are crucial to determine whether CP violation is strong enough to produce a large enough matter anti-matter asymmetry in our universe, which is needed in terms of one of the three Sakharov criteria. In the two Higgs doublet model, it is possible to introduce meaningful complex Higgs self-couplings to induce a large enough CP violation, which is needed for the EW baryogenesis scenarios. These couplings are also important to determine whether the strong first-order phase transition could occur for a realistic EW baryogenesis scenarios, and they can affect gravitational wave radiation in the process of bubble collisions [19–21], where the gravitational wave can induce a B mode which is detectable from the cosmological microwave background [22–24]. So there is strong motivation to explore the shape of the Higgs potential.

Compared to the Standard Model (SM), new physics could modify either effective trilinear (or cubic) or quartic couplings or both, either at 10%–20% via loop corrections [25] or more than 100%–300% via tree-level corrections (say adding a dimension-6 operator [26–28] or many higher dimensional operators [29]). The Lorentz structure of the triple Higgs boson vertex can even be modified in the Higgs-Gravity model [30, 31], which could lead to energetic Higgs bosons in the final states [32]. Using the discovered Higgs boson as a probe to measure the self-couplings of the Higgs boson could help us to further understand the nature of the Higgs bosons, and to extract information on the shape of the Higgs potential, which encodes electroweak symmetry breaking. Therefore, to measure trilinear couplings and quartic couplings [33–36] will be of great importance and could be one of the prime targets for both the LHC high luminosity runs and future collider projects.

The study of the di-Higgs boson final states in the SM and new physics models at hadron colliders has been a hot topic recently. Various production processes and final states have been explored in the literature, including  $b\bar{b}\gamma\gamma$  [32, 37, 38],  $b\bar{b}b\bar{b}$  [39–41],  $WW b\bar{b}$  [42],  $WW \gamma\gamma$  [43],  $b\bar{b}\tau\tau$  [37], the rare decay final states  $3\ell 2j$  [44] and others [45]. Both ATLAS and CMS collaborations have performed realistic simulation and analysis on di-Higgs boson final states at the LHC [46, 47].

In this work, we extend our study in Ref. [44] to the pure leptonic mode, i.e.  $pp \rightarrow hh \rightarrow 4\ell + \cancel{E}$  in a 100 TeV collider. To our best knowledge, this mode has not been carefully studied in the literature due to its tiny production rate in the SM at the collision energy of the LHC. But for some new physics models, the production rate of di-Higgs can be enhanced by a factor from 10 to 100, and then this mode could be accessible even at the LHC. For a 100 TeV collision, the production rate of this mode in the SM itself is large enough and is accessible. Meanwhile, since it gives pure leptonic final states, this mode

can be searched by experimental groups relatively easy. Therefore, it is meaningful to perform a careful analysis of this mode either for the LHC runs or for a future 100 TeV collider project.

In order to determine the Higgs self-couplings at future hadron colliders, a precision measurement of the top Yukawa coupling at the LHC and future colliders is crucial. The top quark Yukawa coupling plays a remarkable role in probing the properties of the Higgs boson. It is the strongest Yukawa coupling, and almost saturates the perturbation bounds; it can affect the vacuum stability [48] much more seriously than any other Yukawa couplings in the SM; it determines the multi-Higgs production at hadron colliders and affects the decay of the Higgs boson to gluon pair, di-photon and  $Z\gamma$  final states much more than the other fermions in the SM; and it affects the Higgs self-coupling measurements at hadron colliders, both trilinear and quartic coupling measurements.

Therefore, in this work, to examine how the top quark Yukawa coupling can affect the measurement of the trilinear Higgs coupling, we take the following effective Lagrangian

$$\mathcal{L}_1 = Y_t (a\bar{t}t + ib\bar{t}\gamma_5 t)h + \lambda_3 \lambda_{\text{SM}} v h h h + \dots, \quad (1)$$

where the term  $Y_t = \sqrt{2}m_t/v$  is the Yukawa coupling of the top quark in the SM, and both  $a$  and  $b$  are dimensionless parameters. The parameter  $b$  is related to CP violation. In the SM,  $a = 1$  and  $b = 0$ . In the two Higgs doublet model (2HDM) with no CP violation,  $a = \text{ctg}\beta$  and  $b = 0$ . If there is CP violation in the 2HDM, the CP even and CP odd neutral scalars could mix, which leads to a non-vanishing  $b$ . Early efforts to probe this coupling at hadron colliders can be found in [49]. The study of measurement of these couplings at linear colliders can be found in [50]. A recent study of how to measure these two free parameters at the LHC and at future hadron colliders can be found in [51, 52], where a different but equivalent parametrisation was used. Theoretical calculation of loop corrections from  $t\bar{t}h$  can be found in [53]. A recent study on the CP properties of the 2HDM can be found in [54]. For a systematic analysis of the constraints from the Higgs precision measurement on either  $a$  or  $b$ , interested readers can refer to [55–57].

The term  $\lambda_{\text{SM}} = m_h^2/2v^2 \approx 0.13$ , while  $\lambda_3$  is a free dimensionless parameter. In various new physics models, this coupling can vary over a large range. For example, in the framework of an effective operator, the strong first order electroweak phase transition has been explored in [28], where  $\lambda_3$  can be in the range [5/3, 3]. In the model with a singlet + SM, the trilinear coupling could be larger than the value of the SM by more than 20% to 200% [58], and this deviation is dependent upon the mass of the singlet. Such an enhancement can also occur in the Minimal Dilaton model [59] and in the NMSSM [60].

There are three main aims for this work: 1) at the hadronic level, we explore the sensitivity of the pure leptonic mode  $pp \rightarrow hh \rightarrow 4\ell + \cancel{E}$  at a 100 TeV collider; 2) we examine the detector effects on the sensitivity; 3) we examine the correlation of the direct measurement of  $t\bar{t}h$  and the direct measurement of  $\lambda_3$  in future colliders by using the numerical approach.

The new findings of this work are as follows. 1) By using our proposed reconstruction method, the pure leptonic mode of Higgs pair production is accessible at a 100 TeV collider and the hadronic level analysis indicates that it could have a pretty high significance. 2) When detector effects are taken into account, the significance can be reduced by a factor from 2 to 4 due to the loss of signal events, which means 4 to 16 times more datasets are needed in order to achieve the same significance of the hadronic level analysis. One of the crucial things is to include those soft leptons (say  $10 \text{ GeV} < P_t(\ell) < 20 \text{ GeV}$ ) in our analysis, which can enhance the significance. 3) The top quark Yukawa coupling can introduce a large uncertainty to the determination of Higgs trilinear coupling due to the fourth power dependence of Higgs pair production on the top quark Yukawa coupling, and a 100 TeV collider can help us to determine the trilinear coupling to a 5% level from the uncertainty of  $t\bar{t}h$ . The coupling of Higgs to weak vector bosons can also significantly affect the Higgs pair production rate, like in the processes  $pp \rightarrow hh + X$  and  $e^+e^- \rightarrow hh + X$ , and can then affect the determination of  $\lambda_3$ .

The work is organised as follows. In Section 2, we study the cross section of the process  $gg \rightarrow hh$ . In Section 3, we examine the characteristic features of full leptonic final states. In Section 4, we explore the sensitivity of two types of the same sign leptonic final states  $gg \rightarrow hh \rightarrow 4\ell + \cancel{E}$  in a 100 TeV collider. In Section 5, we investigate the detector effects on the sensitivity of the fully leptonic mode of Higgs pair production. In Section 6, we examine the issue of how  $t\bar{t}h$  measurement can affect the determination of the trilinear Higgs couplings. In Section 7, by using the effective Lagrangian, we also study how the couplings of Higgs boson to weak vector bosons and the trilinear Higgs couplings can affect the main production rates at both hadron colliders and electron-positron colliders. We end this work with a conclusion and some discussion. We provide an appendix to describe the quasi-Monte Carlo method implemented in our code “wat”, which has been used in this work to evaluate cross sections and to generate unweighted signal events.

## 2 Cross section of Higgs pair production at hadron colliders

Here we briefly describe the numerical methods that we used in this work. To study the cross

sections of Higgs pair productions, we implemented the effective Lagrangian described in Eq. (1) as a new model file by modifying the one-loop SM model file in MadGraph5/aMC@NLO [61]. Then we interface the loop matrix elements outputted from Madgraph5/aMC@NLO with our code “wat”, based on QMC, to perform phase-space integration and event generation. The main algorithms of “wat” are presented in Appendix A. The generated signal events are further showered by PYTHIA6 [62]. The leptons, jets, and missing transverse energy in the final states are simply processed by using the package FASTJET [63]. Then signal events are used to perform physical analysis at hadronic level. To study the detector effects, we use the fast-detector simulation package DELPHES [64]. All the background events are generated by using the Madgraph5 package and are processed as the signal events are done. To apply the multivariate analysis in this work, we have called the subroutines of the MVA package built-in the ROOT package.

Below we describe how the effective Lagrangian in Eq. (1) is implemented in some details. We add the parameters  $a$ ,  $b$ , and  $\lambda_3$  and the corresponding tree level vertices by following the UFO protocol given in [65]. We use MadGraph5 to compute the one-loop amplitudes, where the OPP method [66] has been implemented. Based on the unitarity cuts of one-loop scattering amplitudes [67], the OPP method is a systematic algorithm that reduces a one-loop amplitude into scalar integrals, shown as Eq. (2):

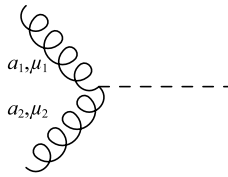
$$\int \frac{d^D q}{(2\pi^2)^D} \frac{N(4, q, p_1, \dots, p_n)}{\prod_{i=1}^n D_i} = \sum_{i < j < k < l} d_{ijkl}(p_1, \dots, p_n) \int \frac{d^D q}{(2\pi^2)^D} \frac{1}{D_i D_j D_k D_l} + \sum_{i < j < k} c_{ijk}(p_1, \dots, p_n) \int \frac{d^D q}{(2\pi^2)^D} \frac{1}{D_i D_j D_k} + \sum_{i < j} b_{ij}(p_1, \dots, p_n) \int \frac{d^D q}{(2\pi^2)^D} \frac{1}{D_i D_j} + \sum_i a_i(p_1, \dots, p_n) \int \frac{d^D q}{(2\pi^2)^D} \frac{1}{D_i} + R_1(p_1, \dots, p_n) \quad (2)$$

$D = 4 - 2\epsilon$ ,  $D_i = (q - p_i)^2 - m_i^2$ , and  $N(4, q, p_1, \dots, p_n)$  is the numerator in 4 dimensions. The main idea of this method is to choose several  $q$  values smartly to calculate the numerator  $N$  numerically, then these coefficients  $d_{ijkl}, c_{ijk}, b_{ij}, a_i$  can be solved easily. Therefore, it can only deal with a 4 dimensional numerator, as indicated in Eq. (2). The difference between a  $D$ -dimensional numerator and a 4-dimensional numerator will give an ex-

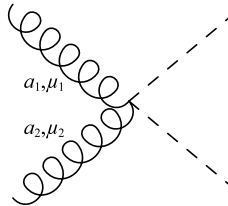
tra rational term after integrating the loop momentum, shown as Eq. (3):

$$\int \frac{d^D q}{(2\pi)^D} \frac{N(D, q, p_1, \dots, p_n) - N(4, q, p_1, \dots, p_n)}{\prod_{i=1}^n D_i} = R_2 \quad (3)$$

It has been proven that this rational term  $R_2$  is related to the ultraviolet nature [68,69], and hence is only required to be calculated once. All these  $R_2$  terms for the SM have been obtained in Refs. [70, 71], and in Ref. [72]. Furthermore, an automatic tool to obtain  $R_2$  term for any renormalizable model is presented. For the effective Lagrangian in Eq. (1), the  $R_2$  terms related to Higgs pair production are shown as below, and we implement them as well as tree-level vertices in a UFO model file [65]:



$$= -i \frac{\sqrt{2} g_s^2 m_t a Y_t \delta_{a_1 a_2} g_{\mu_1 \mu_2}}{16\pi^2} \quad (4)$$



$$= -i \frac{g_s^2 Y_t^2 \delta_{a_1 a_2} g_{\mu_1 \mu_2}}{16\pi^2} (a^2 + b^2), \quad (5)$$

where  $a_1, a_2$  are the color indices,  $\mu_1, \mu_2$  are the Lorentz indices, and  $g_s$  is the QCD coupling constant.

Then we interface the loop matrix element produced by MadGraph5/aMC@NLO [61] to our integration and event generation code to obtain the leading order cross section and unweighted signal events.

At the leading order, cross sections in hadron colliders can be parametrised as the function of theoretical free parameters  $a, b, \lambda_3$  in the following form

$$\sigma(gg \rightarrow hh) = G_1 a^4 + G_2 b^4 + G_3 a^2 b^2 + (G_4 a^3 + G_5 a b^2) \lambda_3 + (G_6 a^2 + G_7 b^2) \lambda_3^2. \quad (6)$$

It is noticed that this cross section is sensitive to the signs of  $a$  and  $\lambda_3$ , respectively, but is insensitive to the sign of  $b$ . There are two types of diagrams contributing to the process  $gg \rightarrow hh$ : box diagrams and triangle diagrams. The terms independent of  $\lambda_3$  come from the squared amplitudes of box diagrams. The terms proportional to  $\lambda_3^2$  are from the squared amplitudes of triangle diagrams, and the terms proportional to  $\lambda_3$  are from the interference between the box and triangle diagrams.

We use the numerical approach to fit the coefficients  $G_1 - G_7$  for the LHC at 14 TeV, 33 TeV and a 100 TeV collider. We generate more than 100 points in  $(a, b, \lambda_3)$  space for each collision energies. The coefficients of cross sections at the 14 TeV, 33 TeV and 100 TeV are presented in Table 1.

Table 1. The fitting coefficients for the LHC 14 TeV and a 100 TeV collider, where the superscript denotes the collision energy 14 TeV and 100 TeV, respectively.  $R^{33}(R^{100})$  is defined as  $K^{33}/K^{14}(K^{100}/K^{14})$ , where  $K$  denotes  $G_1 - G_7$ .

	$G_1/\text{fb}$	$G_2/\text{fb}$	$G_3/\text{fb}$	$G_4/\text{fb}$	$G_5/\text{fb}$	$G_6/\text{fb}$	$G_7/\text{fb}$
14 TeV	34.5	3.37	268	-23.1	-119	4.82	15.1
33 TeV	227	23.1	$1.67 \times 10^3$	-143	-723	28.7	89.0
$R^{33}$	6.6	6.8	6.2	6.2	6.2	5.9	5.9
100 TeV	$1.71 \times 10^3$	186	$1.20 \times 10^4$	$-1.03 \times 10^3$	$-5.09 \times 10^3$	199	610
$R^{100}$	49.6	55.2	44.9	44.6	42.8	41.2	40.4

Due to the enhancement of gluon flux at 100 TeV, it is noticed that all coefficients  $G_i^{100}$  are enhanced compared with  $G_i^{14}$ . Due to the difference in the form factors,  $G_1^{100}$  and  $G_2^{100}$  are around 50 times larger than their counterparts  $G_1^{14}$  and  $G_2^{14}$ . The coefficient  $G_3^{100}$  is enhanced by a factor of almost 45. The squared triangle diagram coefficients  $G_4^{100}$  and  $G_5^{100}$  are 40 times larger than  $G_4^{14}$  and  $G_5^{14}$ . This enhancement factor is smaller than that of the box diagrams due to the s-channel suppression for energetic gluon fluxes. The interference coefficients

$G_6^{100}$  and  $G_7^{100}$  are also 40 times larger than  $G_6^{14}$  and  $G_7^{14}$ .

Another interesting observation is that the coefficient  $G_3$  is 7 times larger than the coefficient  $G_1$ . Typically, when  $b$  is much smaller than 1, the contribution of  $G_3$  term cannot be large. But if  $b$  can be of order one, then the contribution of  $G_3$  can be sizeable. In Refs. [57, 73, 74], more operators have been taken into account. We have compared our results presented there and found agreement.

### 3 Full leptonic modes of signal and background events

There are quite a few advantages of the full leptonic mode of  $gg \rightarrow hh \rightarrow 4\ell + \cancel{E}$  for a 100 TeV collision. First, it is relatively efficient for searches by experiments. The targeted objects in the final state are leptons and missing energy. They can be reconstructed efficiently by the subdetector systems of the LHC detectors. The lepton reconstruction efficiency with  $P_t > 5$  GeV is more than 90% and particle identification can be made at detector level. Second, the signal is relatively clean and it is robust against the contamination of pileup and underly-

ing events, since primary collision vertices of the signal events can be reconstructed.

We use Madgraph5 to generate background events at a collision energy of  $\sqrt{s} = 100$  TeV and use PYTHIA6 to perform showering and decay simulations. Background processes without Z bosons in the final states are generated by using the on-shell approximation for the top quark, W boson and Higgs boson. For background processes with a single Z boson in the final state which decays into two leptons, like  $t\bar{t}Z$ ,  $ZW^+W^-$  and  $Zh$  background processes, we have included the effects of off-shell Z and  $\gamma$  and their interferences.

Table 2. The expected number of events with  $3 \text{ ab}^{-1}$  integrated luminosity at  $\sqrt{s} = 100$  TeV and the generated events for all processes.

	$\sigma \times Br$ /fb	expected number of events at $3000\text{fb}^{-1}$	number of events generated	$K$ -factors
HH	0.18	$5.7 \times 10^2$	500000	1.6
ZZ	$4.8 \times 10^2$	$1.4 \times 10^6$	-	-
Z h	5.56	$1.67 \times 10^4$	500000	0.97
$ZW^+W^-$	6.34	$1.90 \times 10^4$	500000	2.8
$Zt\bar{t}$	$1.97 \times 10^2$	$5.91 \times 10^5$	5000000	1.1
$t\bar{t} h$	$1.41 \times 10^1$	$4.22 \times 10^4$	1000000	1.2
$t\bar{t}t\bar{t}$	5.48	$1.65 \times 10^4$	400000	1.3
$t\bar{t}W^+W^-$	1.78	$5.35 \times 10^3$	200000	1.3
$hW^+W^-$	$6.02 \times 10^{-2}$	$1.81 \times 10^2$	50000	1.4
$W^+W^-W^+W^-$	$2.74 \times 10^{-2}$	$8.23 \times 10^1$	10000	2.8

For both signal and background, higher order corrections are taken into account by normalizing the total cross section to their (N)NLO results, which is indicated by the  $K$  factor ( $K = \sigma_{(N)NLO} / \sigma_{LO}$ ) in Table 3. We adopt the NNLO result for signal in Ref. [75], and the  $K$  factor for processes  $hZ$ ,  $t\bar{t}Z$ ,  $ZW^+W^-$ ,  $t\bar{t}h$ ,  $hW^+W^-$  are obtained by MadGraph5/aMC@NLO [61] under the on-

shell approximation. The  $K$  factor for  $t\bar{t}t\bar{t}$  and  $t\bar{t}W^+W^-$  at 100 TeV collider is unknown, and we adopt a value  $K = 1.3$ , which is the  $K$  factor for  $t\bar{t}t\bar{t}$  at 14 TeV LHC. The  $K$  factor for the process  $pp \rightarrow W^+W^-W^+W^-$  is also unknown, and we use the  $K$  factor for  $ZW^+W^-$  since both of them have the same initial states at hadron colliders.

Table 3. Cross sections of the four leptonic mode at a 100 TeV collider for different processes, with  $\ell = e, \mu$ . The fractions of the four modes in all the final states are shown.

processes	labels in figs.	cross section/ fb	M1 $\ell^+\ell^-\ell^+\ell^-$	M2 $e^+e^-\mu^+\mu^-$	M3 $\ell^+\ell^-\ell^\pm\ell'^\mp$	M4 $\ell^+\ell'^-\ell^+\ell'^-$
hh	signal	0.29	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$
Zh, $ZW^+W^-$ , $Zt\bar{t}$	Z+	5.40,17.8,217	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0
$t\bar{t}h$ , $t\bar{t}t\bar{t}$ , $t\bar{t}W^+W^-$	$t\bar{t}+$	16.9,7.12,2.3	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$
$hW^+W^-$ , $W^+W^-W^+W^-$	EW	$8.4 \times 10^{-2}$ , $7.7 \times 10^{-2}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$
ZZ		485	$\frac{1}{2}$	$\frac{1}{2}$	0	0

The background processes can be roughly divided into the following three types:

- 1) Single Z processes, including  $pp \rightarrow Zh$ ,  $pp \rightarrow ZW^+W^-$ , and  $pp \rightarrow Zt\bar{t}$ . The last process has a cross section around 35 times larger than the former two processes due to its QCD nature, while Zh and  $ZW^+W^-$  have similar cross sections.
- 2) Top pair processes, including  $pp \rightarrow t\bar{t}h$ ,  $pp \rightarrow t\bar{t}t\bar{t}$ , and  $pp \rightarrow t\bar{t}W^+W^-$ . We notice that  $h\bar{t}t$  is the dominant background in this category.
- 3) Pure electroweak processes, including  $pp \rightarrow hW^+W^-$  and  $pp \rightarrow W^+W^-W^+W^-$ . Each of their cross sections is smaller than that of the signal, but their sum is comparable to that of the signal after taking into account the  $K$  factors.

The cross sections of these processes are listed in Table 3. It is worth remarking that single-Higgs associated processes in the categories above are the main background events, and it turns out that  $t\bar{t}h$  is the dominant background for the full leptonic mode, which can greatly affect the significance.

In order to select the most relevant events, we introduce the following preselection cuts:

- For each event, there must be four leptons. The leading two leptons should have transverse momenta larger than 30 GeV and 15 GeV, respectively. The third and fourth leptons should have transverse momenta larger than 10 GeV and 8 GeV respectively. In Fig. 1, we show the transverse momenta of the four leptons in the signal events.
- We demand the future detector should have a better coverage of  $\eta(\ell)$  as  $|\eta(\ell)| < 4$  for identified leptons. Due to the good space resolution power and fine granularity in the ECAL detector, we require that the minimal angular separation between two leptons is  $\Delta R^{\min}(\ell, \ell) \geq 0.15$ . In Fig. 2, we demonstrate the distribution of the maximal  $\eta^{\max}(\ell)$  and the minimal angular separation of any pair of two leptons  $\Delta R^{\min}(\ell\ell)$ . When these two cuts on leptons are applied, more than 90% signal events can be accepted. We will discuss how the coverage of  $\eta$  can affect our results in the discussion section.
- Considering the fact that  $t\bar{t}$  processes have quite large contributions to the background events, we introduce the “b-jet veto” to reject events with a tagged b-jet with  $P_t(b) > 40$  GeV and  $|\eta(b)| < 5$ . We assume that the b tagging efficiency is 60%

and simply multiply these types of background processes by a factor 0.16.

- For the decay mode M3, in order to suppress the background events from meson decays in the final states and Z boson decay, we demand the invariant mass of both lepton pairs should fall into two windows, either  $15 \text{ GeV} \leq m_{\ell^+\ell^-} \leq 80 \text{ GeV}$  or  $m_{\ell^+\ell^-} \geq 100 \text{ GeV}$ .

Table 4. Preselection cuts used in our analysis.

preselection cuts	description
1	$n_\ell = 4$ $P_t(\ell_1) > 30 \text{ GeV}, P_t(\ell_2) > 15 \text{ GeV}$ $P_t(\ell_3) > 10 \text{ GeV}, P_t(\ell_4) > 8 \text{ GeV}$ $ \eta^{\max}(\ell_i)  < 4$ $\Delta R^{\min}(\ell, \ell) > 0.15$
2	b jet veto
3	low energy hadron veto and Z mass veto

From the results given in Table 3, the background from  $pp \rightarrow ZZ$  is huge, about 2000 times larger than the rest. Furthermore, due to the off shell Z and  $\gamma^*$ , the cut of invariant mass of two pairs of four leptons can only suppress the background by a factor of 100 at most. Our previous experience in studying  $3\ell + 2j$  [44] reveals that background events from Z,  $\gamma^*$ -exchange processes are difficult to suppress. Therefore, the first two modes are challenging in the SM and will be neglected in this analysis. When there is a significant enhancement to Higgs pair production by new physics, these two modes should be considered.

In this work, we will focus on the third and fourth modes (labelled as “M3” and “M4”, respectively) and we will perform a detailed analysis on these two modes. Although compared with the third mode, the fourth mode has a smaller production rate, it has less background contributions from the SM.

## 4 Kinematic features of signal and background events in M3 and M4 cases

In this section, we explore the kinematic features of the signal and background events. It is impossible to fully reconstruct all the final particles due to 4 neutrinos in the final state.

The physical observables can be divided into two types. **The first type** is defined as the global and topological event shape observables for each event, and **the second type** is defined from the partial reconstruction method introduced later. **The first type** includes the following observables listed below.

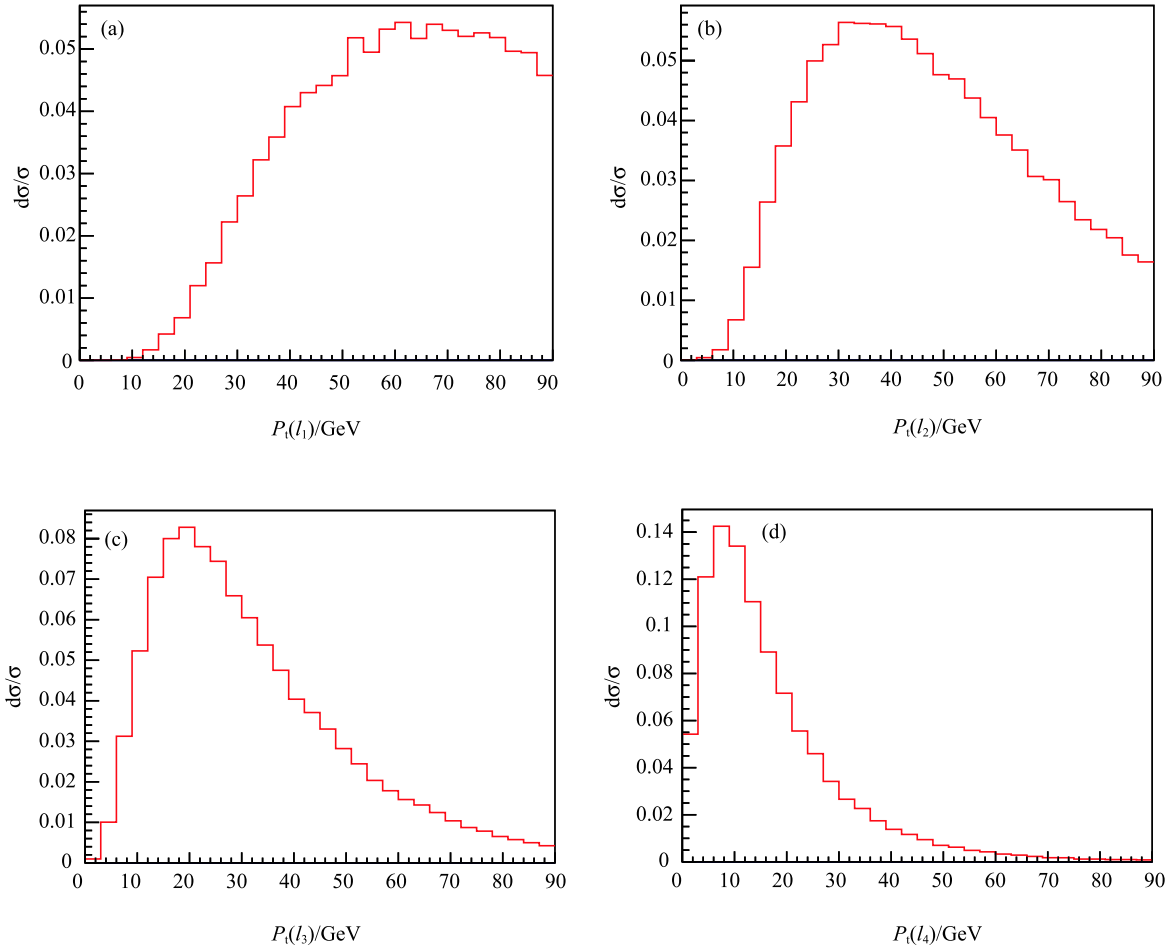


Fig. 1. Distributions of four leptons in the signal events before preselection cuts.

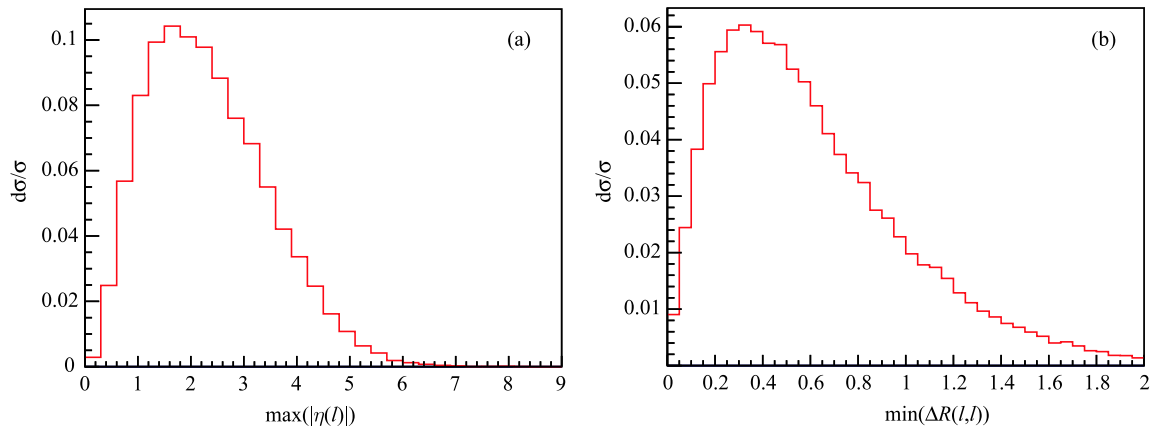


Fig. 2. Distributions of  $\max(|\eta(\ell)|)$  and the minimal angular separation  $\min(\Delta R(\ell \ell))$  four leptons in each signal event before preselection cuts.

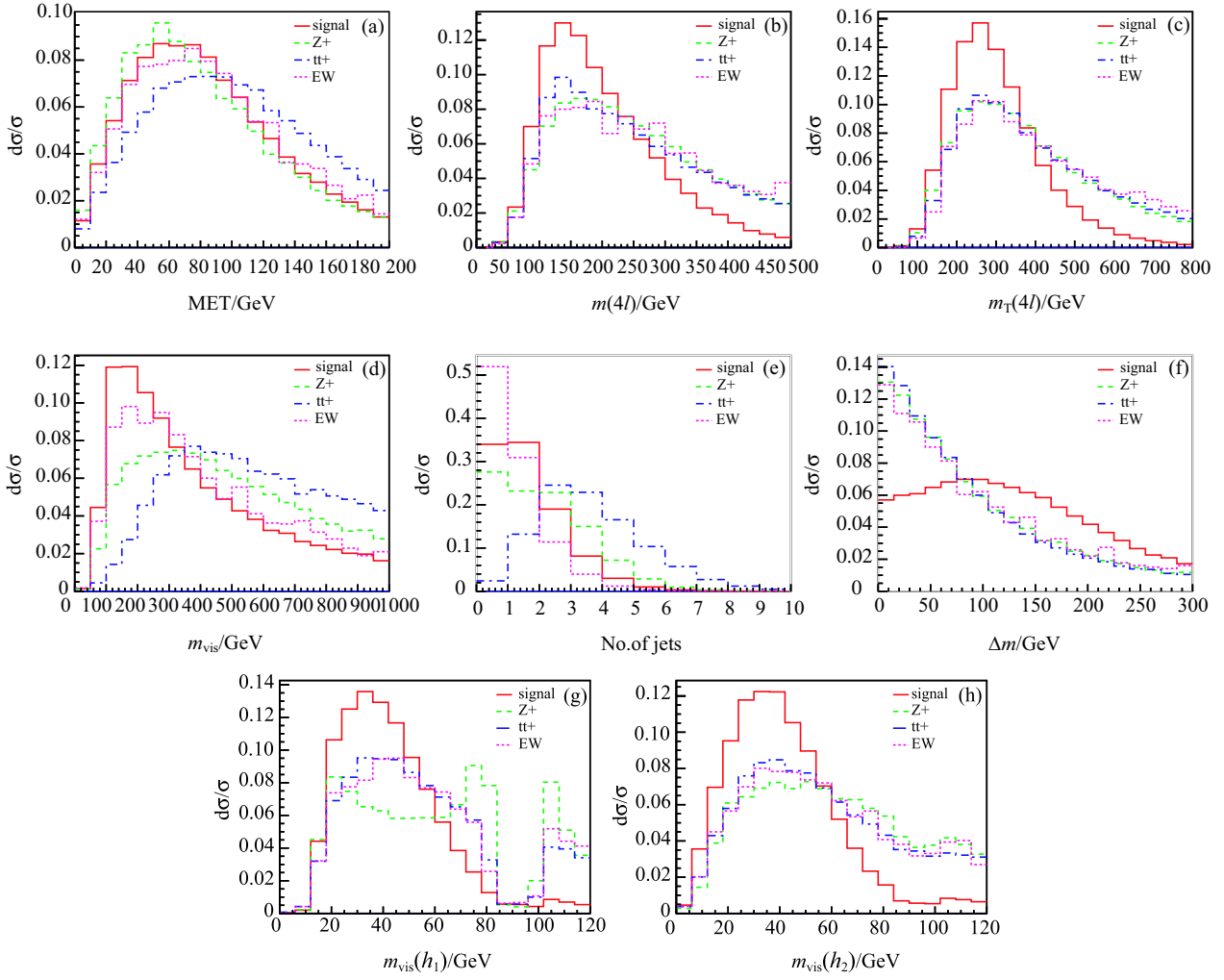


Fig. 3. (color online) Some useful kinematic observables for the M3 case which can separate signal and background events.

- o1) The missing transverse momentum spectrum. Since there are 4 neutrinos in the final states, we expect that there should be a large transverse momentum. For the signal, as shown in Fig. 3(a), the distribution peaks near 60–80 GeV.
- o2) The invariant mass of four leptons. This quantity is expected to capture the mass of mother particles. For the signal, as shown in Fig. 2(b), the distribution peaks near 100–150 GeV. Since four neutrinos can take away half the energy of a Higgs pair, this quantity is expected to be close the mass of one Higgs boson.
- o3) The transverse mass of each event is constructed from the sum of 4-momentum of leptons (denoted as  $P_{4\ell}$ , which has components  $(E_{4\ell}, P_{4\ell}^x, P_{4\ell}^y, P_{4\ell}^z)$ ) and the missing transverse momentum ( $\vec{P}_T$ ), which has components  $(\vec{E}_T, \vec{P}_x, \vec{P}_y, 0)$  where  $\vec{E}_T = \sqrt{\vec{P}_x^2 + \vec{P}_y^2}$ . To

construct this observable, we boost the 4-momentum of leptons such that the  $P_{4\ell}^z = 0$ , i.e.  $(\vec{E}_{4\ell}, P_{4\ell}^x, P_{4\ell}^y, 0)$ . Then we construct the observable as  $\sqrt{\vec{E}_{4\ell} \vec{E}_T - 2P_{4\ell}^T \vec{E}_T \cos[\phi]}$ , where  $\phi$  denotes the azimuthal angle between the  $\vec{P}_{4\ell}$  and  $\vec{P}_T$ . For the signal, as shown in Fig. 3(c), the distribution peaks near 150–250 GeV.

- o4) The invariant mass of visible objects is defined as the momentum sum of four leptons and jets. For the signal, as demonstrated in Fig. 3(d), the distribution peaks near the region 100–300 GeV.
- o5) The number of jets in each event with  $P_t(j) > 40$  GeV. As demonstrated in Fig. 3(e), we noticed that when demand  $n_j \leq 2$ , around 90% signal events can be selected out, though with a considerable number of background events from  $t\bar{t}$  processes.

Below we explain how to construct the second type



of observables.

Obviously, the most important information about the signal events is the mass of the Higgs boson. So it is crucial to extract this useful observable. Due to the fact that four neutrinos cannot be fully reconstructed, instead we can only determine the visible masses of each Higgs boson from the identified leptons.

To determine the visible mass of each Higgs boson mass, we encounter a minor combinatorics issue: there are two possible combinations in each event. To determine which one is correct, we follow the minimal mass method introduced in [44] to determine the right combination, which can yield a correctness up to 94% here as we have checked this by using the parton level data sample. The method evaluates the sum of two visible masses of Higgs bosons for each combination, and picks out the smaller one as the correct combination. In contrast, we examine a method by using the angular separation of leptons, which can only yield a correctness up to 85% at most.

After having found the visible mass of the Higgs boson in the signal, by using the standard  $M_{T2}$  method we can split the missing transverse momentum into two parts and exploit the kinematic feature of pair production to reconstruct the transverse mass of Higgs boson. Then we can construct **the second type** of observables as listed below.

- o6) The observable  $\Delta m$ , which is defined as the mass difference between two mass sums of reconstructed Higgs bosons in two possible combinations. The distribution is shown in Fig. 3(f), and shows a large shape difference between the signal and background processes in the shape.
- o7)) The first reconstructed partial mass of the Higgs boson with two leptons of the same flavour in the M3 case, which is labelled as  $h_1$ . In the M4 case, the one with the hardest lepton is labelled as  $h_1$ . The distribution of this quantity is shown in Fig. 3(g) and Fig. (5)g. In Fig. (3)g, the Z mass

window cut is clearly shown for the processes with a single Z.

- o8)) The second reconstructed partial mass of the Higgs boson with two leptons of different flavour in the M3 case, labelled as  $h_2$ . In the M4 case, the one reconstructed not with the hardest lepton is labelled as  $h_2$ . The distribution of this quantity is shown in Fig. 3(h) and Fig. 5(h).
- o9) The transverse mass of the Higgs bosons reconstructed by using the  $M_{T2}$  method, which is shown in Fig. 4(a) and Fig. 6(a).

All of these reconstructed observables are crucial for both M3 and M4 cases. There exist strong correlations among these observables for signal events and they are related to the mass of the Higgs boson, while for the background they are not necessarily related to the mass of the Higgs boson. Such a fact can be utilised to separate signal and background, which is the spirit of multivariate analysis methods.

#### 4.1 The M3 case

In the M3 case, the dominant background processes are the single Z associated processes, as clearly demonstrated in Table 5 in the cut-based analysis. From Fig. 3(g), the mass cut can clearly affect the single Z processes.

In the cut-based analysis, we choose the following 4 cuts: 1) a cut on the number of jets  $n_j \leq 2$ , which is supposed to suppress background processes associated with a top pair; 2) a cut on the variable  $m_{T2}$ , as we demand  $m_{T2} < 110$  GeV, which can greatly suppress the background processes like three body and four body productions; 3) a cut on the reconstructed visible mass of Higgs boson is imposed as  $m_{h_{1,2}} \leq 60$  GeV, which can suppress background events from the Zh and ht $\bar{t}$  processes; 4) a cut on the mass difference  $\Delta m > 50$  GeV. By using these cuts, we can achieve  $S/B$  and significance 0.25 and 4.2, respectively.

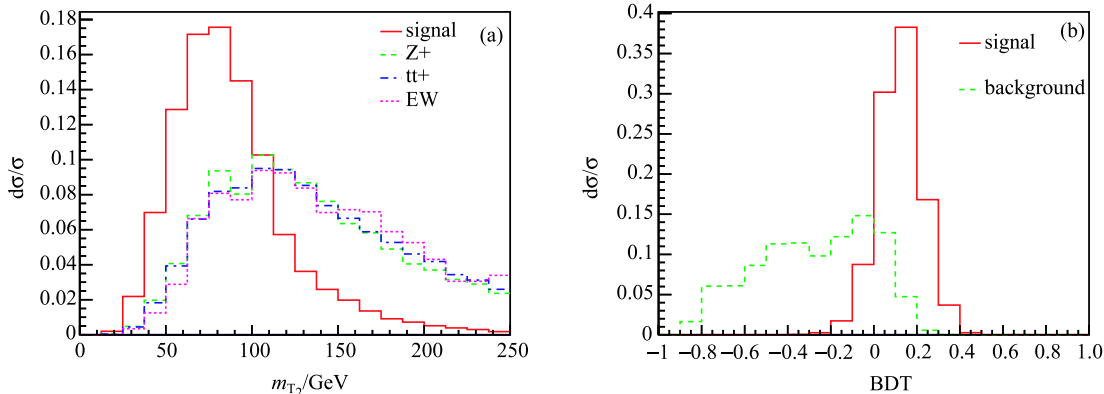


Fig. 4. (color online) Reconstructed visible masses of two Higgs bosons and  $m_{T2}$  for the M3 case and the BDT discriminant are shown.

According to the results based on the cut-based method, it is noticed that the processes associated with top pair are the main background for the “M3” mode. In order to suppress this type of background, we introduce a few top veto observables in our BDT analysis. For example, the visible invariant mass of the whole event, as demonstrated in Fig. 3(d), can help to separate background events with and without a top quark pair. Furthermore, the number of jets in each events can help to suppress the background events with a top quark pair, as demonstrated in Fig. 3(e). These two observ-

ables are correlated with each other for the background processes.

In Table 5, both a cut-based analysis and a BDT analysis are presented for comparison. As demonstrated in Fig. 4(b), after taking into account these observables to reject top quark final states, we can further suppress background and gain in the  $S/B$  and significance up to 0.38 and 6.1, respectively, as shown in Table 5. As discussed in [76], if an integrated luminosity of 20–30 fb<sup>-1</sup> for a 100 TeV collider can be achieved, then a significance 20.0 is expected.

Table 5. Cut-based analysis and BDT analysis for the M3 case. Efficiencies of each cut in the cut-based analysis are shown. We assume the integrated luminosity is 3 ab<sup>-1</sup>.

processes	pre-sel. cuts	$n_j \leq 2$	$M_{T2} < 115$ GeV	$m(h^{\text{vis}}) < 60$ GeV	$\Delta m > 50$ GeV	BDT
hh	172	150	124	91.2	68.8	99.8
Z h	243	238	197	66.6	23.4	27.4
ZW <sup>+</sup> W <sup>-</sup>	$1.60 \times 10^3$	$1.54 \times 10^3$	444	173	91.4	111.7
Zt $\bar{t}$	$2.55 \times 10^3$	$1.45 \times 10^3$	542	222	117	89.7
t $\bar{t}$ h	446	245	128	68.0	31.1	29.5
t $\bar{t}$ t $\bar{t}$	254	24.4	3.96	1.44	1.04	3.63
t $\bar{t}$ W <sup>+</sup> W <sup>-</sup>	151	71.2	12.5	4.16	2.48	4.30
hW <sup>+</sup> W <sup>-</sup>	44.5	42.7	20.3	10.3	4.94	6.11
W <sup>+</sup> W <sup>-</sup> W <sup>+</sup> W <sup>-</sup>	50.9	47.3	5.67	1.98	1.20	1.98
$S/B$	$3.2 \times 10^{-2}$	$4.1 \times 10^{-2}$	0.10	0.17	0.25	0.38
$S/\sqrt{B}$	2.35	2.48	3.37	3.90	4.20	6.1

## 4.2 The M4 case

In the M4 case, similar to the M3 case, we can reconstruct the visible mass of Higgs boson,  $h_1(\ell\ell')$  and  $h_2(\ell\ell')$ , respectively, where subscript 1 and 2 are assigned according to the simple rule: the one which includes the most energetic lepton is assigned to be  $h_1$  and the other is assigned to be  $h_2$ . It is noticed that these two masses peak near 30–50 GeV and 20–40 GeV, respectively, and both have an edge near 60 GeV.

Different from the M3 case, the main task in the M4 mode is to suppress the background processes from ht $\bar{t}$  final states, while the singlet Z processes can be negligible and are omitted in Table 6. The final state ht $\bar{t}$  is the

dominant background. Our reconstruction method can successfully reconstruct the Higgs boson in ht $\bar{t}$  processes, as demonstrated in Fig. 5(h), where the shapes of background and signal look similar. Therefore, we impose a cut on  $m(h_1^{\text{vis}})$ , instead of both as in the M3 mode case.

A cut-based analysis and a BDT analysis are presented in Table 6. For the third cut, in this case, we only impose a cut on the visible mass  $m(h_1^{\text{vis}}) < 60$  GeV. As demonstrated in Fig. 6(b), after taking into account these observables to reject top quark final states, we can further suppress the main background and gain in the  $S/B$  and significance up to 1.6 and 6.8, respectively, as shown in Table 6.

Table 6. Cut-based analysis and BDT analysis for the M4 case. Efficiencies of each cut in the cut-based analysis are shown. We assume the integrated luminosity is 3 ab<sup>-1</sup>.

processes	pre-sel. cuts	$n_j \leq 2$	$M_{T2} < 110$ GeV	$m(h_1^{\text{vis}}) < 60$ GeV	$\Delta m > 50$ GeV	BDT
hh	55.2	48.5	40.5	36.3	28.8	43.9
t $\bar{t}$ h	147.1	80.5	44.4	27.1	13.7	14.3
t $\bar{t}$ t $\bar{t}$	77.4	7.93	1.27	$5.13 \times 10^{-1}$	$3.59 \times 10^{-1}$	1.24
t $\bar{t}$ W <sup>+</sup> W <sup>-</sup>	47.1	22.6	4.21	1.94	1.21	1.94
hW <sup>+</sup> W <sup>-</sup>	15.1	14.5	7.41	4.36	2.11	4.46
W <sup>+</sup> W <sup>-</sup> W <sup>+</sup> W <sup>-</sup>	15.9	15.0	2.37	1.08	$6.22 \times 10^{-1}$	1.86
$S/B$	0.18	0.35	0.68	1.04	1.6	1.9
$S/\sqrt{B}$	3.17	4.09	5.24	6.13	6.79	9.2

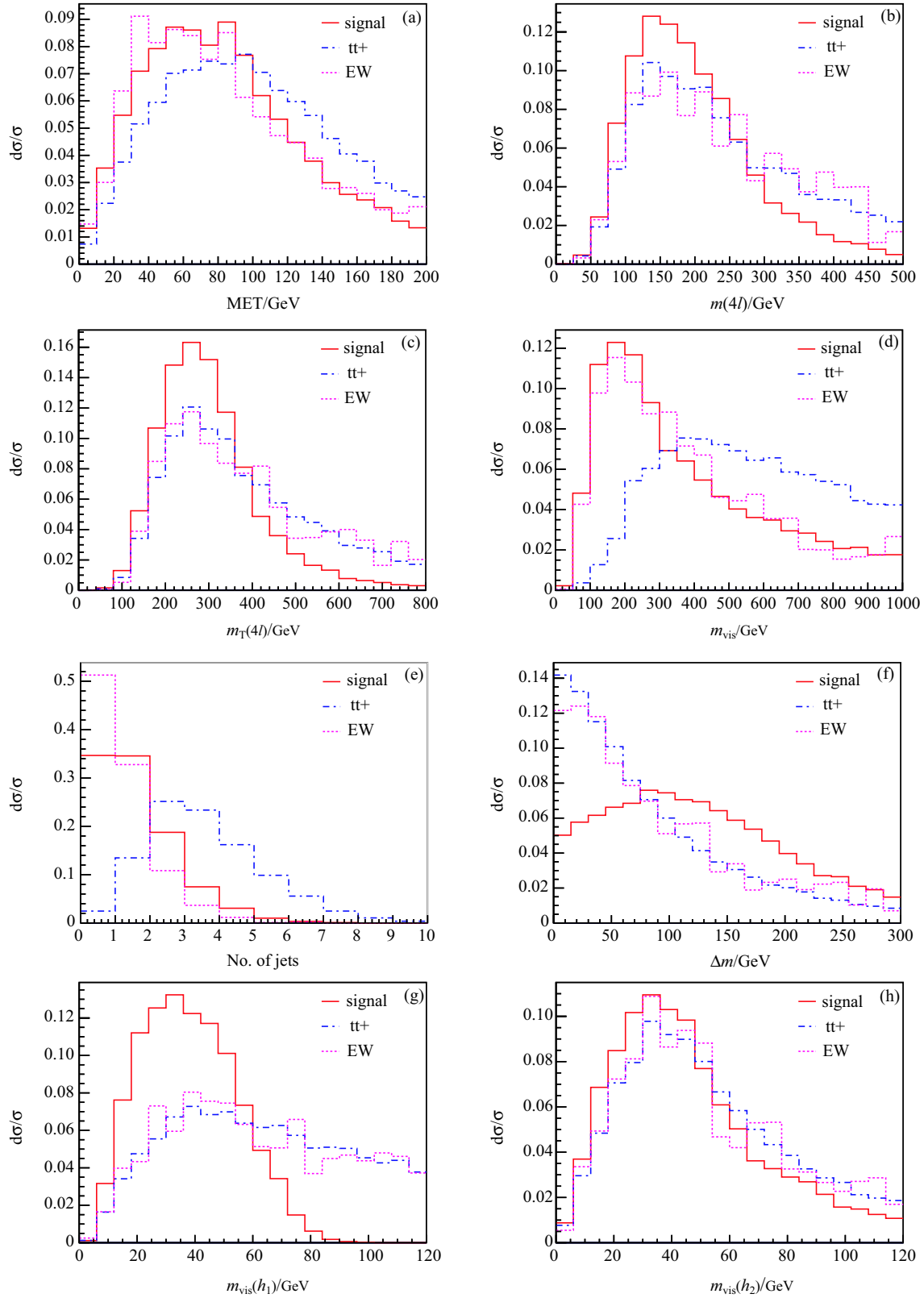


Fig. 5. (color online) Some useful kinematic observables for the M4 case which can separate signal and background events.

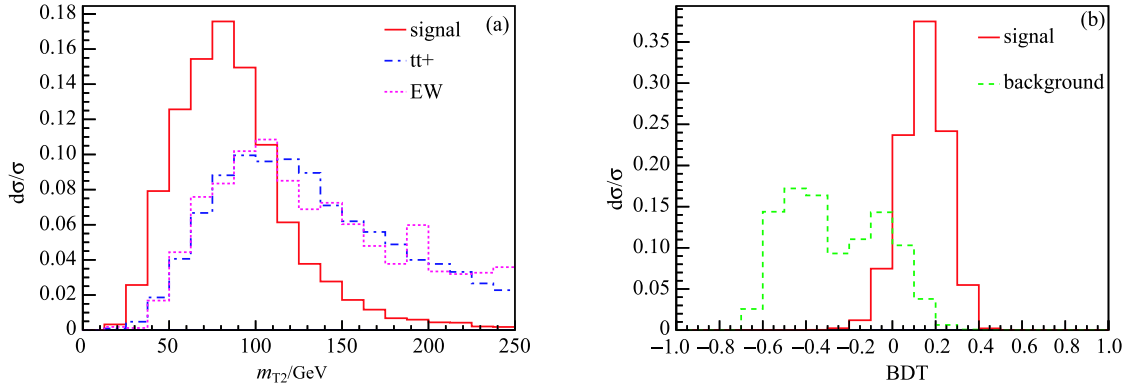


Fig. 6. (color online) The  $m_{T_2}$  observable and the BDT discriminant for the M4 case.

After using the same methods to veto top pair associated background in the MVA method, a better result is obtained, where the  $S/B$  and significance can reach 1.9 and 9.2, respectively. If an integrated luminosity  $20\text{-}30\text{ fb}^{-1}$  is assumed, then a significance of 30.0 is expected.

When comparing the results given in Table 5 and Table 6, we notice that the  $S/B$  of the M4 case is much better than that of the M3 case, due to the lack of huge background processes as in the M3 case. Furthermore, the analysis for the M4 case is relatively simpler than the M3 case due to the background from single  $Z$  processes being efficiently suppressed.

## 5 Detector effects

Here we examine how the detector can affect our results given in Table 5 and 6. We focus on the results of the BDT analysis.

The key issue of the detector is related to lepton acceptance. We show the acceptance efficiencies for the signal events in Table 7. If the coverage of  $\eta$  for leptons can reach 4, then 90% of the signal events can be accepted, but if the coverage of  $\eta$  for leptons can only reach 2, less than half of the signal events can be accepted. A detector which can cover a larger  $\eta$  region (say  $|\eta(\ell)| < 4$ ) is good for Higgs pair signals.

Table 7. Acceptance efficiencies of signal events in  $\eta$ .

$\eta$ coverage	signal acceptance efficiency
$ \eta  < 2$	50%
$ \eta  < 3$	73%
$ \eta  < 4$	91%
$ \eta  < 5$	98%
$ \eta  < 6$	$\approx 100\%$

We examine the detector effects by comparing seven detector setups: 1) an ATLAS-like detector; 2) an ATLAS-like detector, with a larger  $\eta(\ell)$  coverage up to 4 only; 3) an ATLAS-like detector with  $\eta(\ell) < 4$ , and  $\eta(b) < 4$ ; 4) an ATLAS-like detector with  $\eta(\ell) < 4$ ,

$\eta(b) < 4$ , and smaller  $\Delta R$  for lepton isolation; 5) a FCC default detector provided in DELPHES package version 3.4.2; 6) an improved FCC detector with lower jet  $P_t$  threshold; 7) an improved FCC detector with a further lower lepton  $P_t$  threshold. The key parameters of these detector are tabulated in Table 8. The first two are determined at the hardware level, and the rest three can be implemented in trigger and analysis levels.

Table 8. Key parameters for seven detector setups.

	$\eta^{\max}(\ell)$	$\eta^{\max}(b)$	$\Delta R(\ell)$	$P_t^{\min}(j)$ GeV	$P_t^{\min}(\ell)$ GeV
ATLAS	2.5	2.5	0.5	20	10
A1	4.0	2.5	0.5	20	10
A2	4.0	4.0	0.5	20	10
A3	4.0	4.0	0.3	20	10
FCC	6.0	6.0	0.3	40	20
F1	6.0	6.0	0.3	20	20
F2	6.0	6.0	0.3	20	10

In these detector setups, except for the spatial coverage, we assume all sub-detectors of the ATLAS-like detector setups have the same performance. Likewise, all sub-detectors of the FCC-like detector setups are assumed to have the same performance.

The results for M3 and M4 fully leptonic modes are shown in Table 9. A few comments on the results are in order.

1) For the ATLAS-like detector, due to  $\eta^{\max}$  being small, around half of the signal events are lost, as can be seen from Table 7. The requirement of a larger  $\Delta R(\ell)$  causes a further loss of signal due to the spin correlation of two leptons from Higgs decay and a moderately boosted Higgs boson. The lepton reconstruction efficiency further causes signal loss. Although background events also suffer loss, a worse significance is reached. Compared with the hadronic level analysis, the significance is reduced by a factor 3.6 for M3 mode and 4.2 for M4 mode, which means an order of magnitude increase of integrated luminosity in order to reach the same significance.

2) The improved ATLAS-like detectors can improve the  $\eta^{\max}$  coverage, a better b jet rejection, and a better  $R(\ell)$  resolution as well. When the detector setup A3 is considered, the significance can reach 2.9 for M3 mode and 3.8 for M4 mode when an integrated luminosity  $3000 \text{ fb}^{-1}$  is assumed. When compared with the setup A2, the main gain is caused by lowering the minimal  $P_t(\ell)$ , which can almost double the number of signal events.

3) The FCC default detector setup has a larger  $\eta^{\max}$  coverage. Due to a larger  $P_t(\ell)$  requirement, fewer signal events can be selected. Meanwhile, due to a larger  $P_t(b)$ , there is a worse b jet rejection from  $t\bar{t}$  type background events. These two facts lead to a large loss in significance when compared with the hadronic level analysis.

4) The improved FCC detectors are mainly focused on  $P_t(b)$  and  $P_t(\ell)$ . The significance can be slightly better with the improvement of  $P_t(b)$ . But, as shown from the F2 detector setup,  $P_t(\ell)$  is more crucial for our signal. It is found that a smaller  $P_t(\ell)$  ( $P_t(\ell) = 10 \text{ GeV}$ ) can double the number of signal events. It is also noticed that the F2 detector setup can be slightly better than the A3 detector setup. Compared with the hadronic level analysis, the significance drops at least a factor of 2, which means 4 times larger integrated luminosity in order to reach the same significance.

Table 9. Detector effects for fully leptonic final states are compared in term of  $S/B$  and significance with the hadronic level analysis.

	M3			M4		
	$S$	$B$	$S/\sqrt{B}$	$S$	$B$	$S/\sqrt{B}$
hadronic analysis	99.8	274.3	6.1	43.9	23.8	9.2
ATLAS	13.2	62.4	1.7	6.3	8.5	2.2
A1	18.6	108.3	1.8	8.7	13.4	2.4
A2	18.3	93.3	1.9	8.7	11.6	2.6
A3	35.8	156.2	2.9	16.7	18.9	3.8
FCC	18.6	82.6	2.0	8.0	13.6	2.2
F1	18.3	57.5	2.4	8.1	8.0	2.9
F2	38.2	177.5	2.9	17.5	19.1	4.0

## 6 Correlation between the processes $gg \rightarrow hh$ and $pp \rightarrow t\bar{t}h$

Events from the  $t\bar{t}h$  final state are the dominant background for the M4 case, which is also true for the  $pp \rightarrow hh \rightarrow 3\ell + 2j + \cancel{E}_T$  mode explored in [44] at the HCs. Therefore, the measurement of the top Yukawa coupling can significantly affect the detection of Higgs pair production. Below, we examine how the measurement of  $t\bar{t}h$  couplings can affect the determination of  $\lambda_3$ . The correlation between the measurement of  $t\bar{t}h$  and the measurement of Higgs pair production can be investigated by the cross section of  $gg \rightarrow hh$  given in Eq. (6). This issue has not been addressed in the literature.

Before showing more results, we would like to explain the sensitive regions shown in Figs. 7(a)–9(b). The sensitive regions in the figure are obtained by using the cross section of  $gg \rightarrow hh$  in harmonic colliders. The lines separating the sensitive region and the insensitive region are determined by the cross section of  $gg \rightarrow hh$ . For both the LHC and the 100 TeV cases, the lines are determined by a cross section 150 fb or so. For example, in Fig. 7(a), in the region with large  $a$  (say  $a > 2$ ) and large  $\lambda_3$  (say  $\lambda_3 > 6$ ), the cross sections can be larger than 150 fb, therefore this region can be probed even at the LHC. In contrast, in Fig. 7(b), in the region with small  $a$  (say  $a < 0.2$ ) and  $\lambda_3$  around 2, the cross sections are smaller than 150 fb. Therefore even at a 100 TeV collider, it is difficult to probe this region. All these features are shown in the fitted cross section given in Eq. (6).

In Fig. 7, we demonstrate the correlation between the determination of  $t\bar{t}h$  at the LHC and a future 100 TeV collider, where the bounds are estimated from the measurement of  $3\ell 2j + \cancel{E}$ . These bounds have not been optimised for each value of  $a$  and  $\lambda_3$ ; we simply use the information of cross sections in this projection. When detector effects are taken into account, a worse bound is expected. As demonstrated in [44], when the discrimination of signal and background is optimised, we can expect better bounds.

For the LHC with a luminosity  $3 \text{ fb}^{-1}$ , we assume that  $t\bar{t}h$  couplings can be determined up to 20% (10% is estimated by using the boosted techniques for  $t\bar{t}h$  and  $h \rightarrow b\bar{b}$  [77, 78]). We deliberately take a larger value for this since only statistics are taken into account, which is denoted by the two solid lines in Fig. 7(a) as upper and lower bounds from  $t\bar{t}h$  measurements.

For a 100 TeV collider with luminosity  $3 \text{ fb}^{-1}$ , we assume that 5% precision can be achieved, which is denoted by the two solid lines in Fig. 7(b) as upper and lower bounds from  $t\bar{t}h$  measurements. In Ref. [79], where the production ratio  $t\bar{t}h/t\bar{t}Z$  measurement and boosted Higgs and top quark are used, it is argued that this coupling can be measured up to a precision 1% or so when only statistics are considered. In reality, more background processes and detector effects must be considered for each  $t\bar{t}$  decay final state, so we assume a precision up to 5% as a relatively conservative and loose estimation.

Comparing Fig. 7(a) and Fig. 7(b), it is noticed that a 100 TeV collider can greatly shrink the uncertainty in determining the  $\lambda_3$  and  $a$  parameters. Due to  $a^4$  dependence of the cross section  $\sigma(pp \rightarrow hh)$ , a 5% uncertainty of  $\delta a$  can induce an uncertainty of  $\lambda_3$  up to 20% or so. If the coupling  $a$  can be determined to a precision 1%, that will undoubtedly be crucial to pinpoint the value of  $\lambda_3$  down to 5%. It is worth mentioning that the two-fold ambiguity in  $a$  with the same cross section can be removed

by using the method to check the differential distribution of leptons in the final state, as demonstrated in [44],

which is robust against the contamination of underlying events and pileup effects.

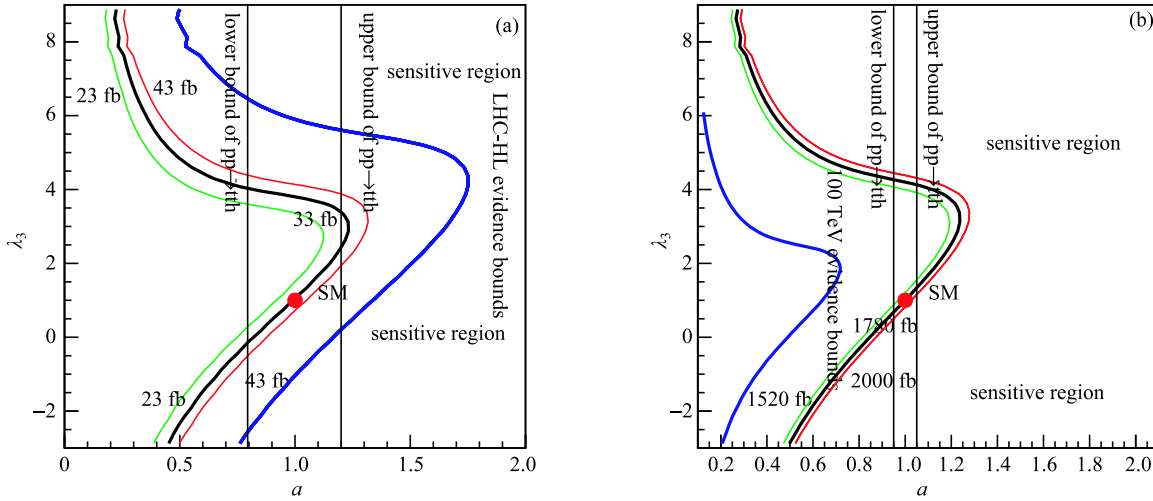


Fig. 7. (color online) Sensitivity in the  $a$ - $\lambda_3$  plane between the LHC high-luminosity run and a 100 TeV collider. We fix  $b = 0.0$ .

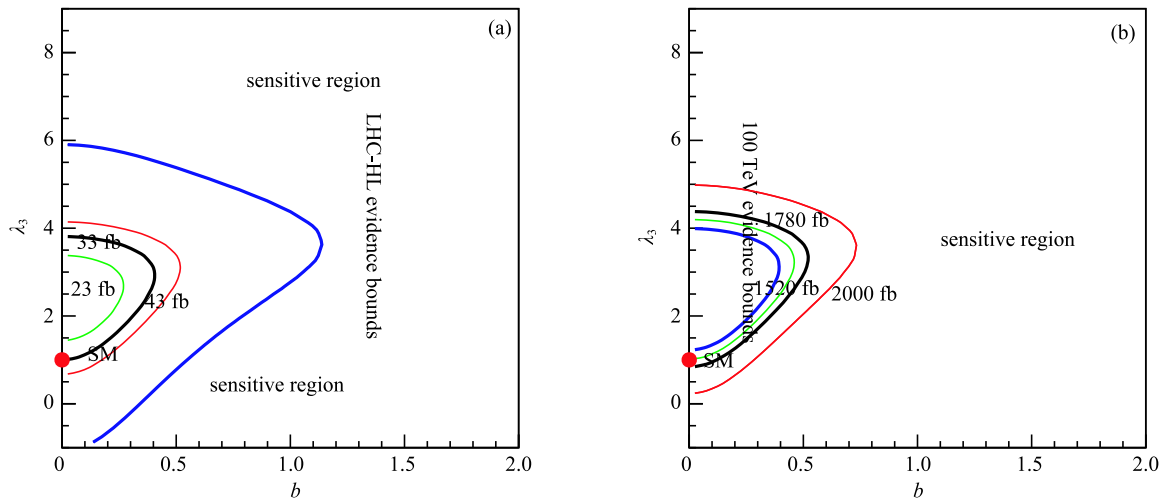


Fig. 8. (color online) Sensitivity in the  $b$ - $\lambda_3$  plane between the LHC high-luminosity run and a 100 TeV collider. To project the feasibility, we fix  $a = 1.0$ .

In the effective Lagrangian given in Eq. (1), the parameter  $b$  is related to the strength of CP violation. We explore the determination of the  $b$  parameter from Higgs pair production, which is given in Fig. 8(a). From Fig. 8(b), the potential of a 100 TeV collider to probe the parameter space expanded by  $b$  and  $\lambda_3$  is obviously better than that of the LHC.

Higgs pair production at a 100 TeV collider can bound  $b$  down to 0.4. In contrast, the LHC runs can only constrain this parameter to 1.2 or so at most. When we quote these numbers, we have not taken into account

the uncertainty in  $a$ . Obviously, the direct measurement from  $t\bar{t}h$  could impose a better constraint on the value of  $b$ , either at the LHC or at a 100 TeV collider.

The correlation of  $a$  and  $b$  in a Higgs pair is provided in Fig. 9(a). We plot the dependence of the cross section of  $t\bar{t}h$  upon  $a$  and  $b$  in the same plots. From Fig. 9(b), it is noticed that the LHC can be sensitive to the region where  $a < 0$  and  $b$  is sizeable, which corresponds to a large production rate of  $gg \rightarrow hh$  due to a constructive interference.

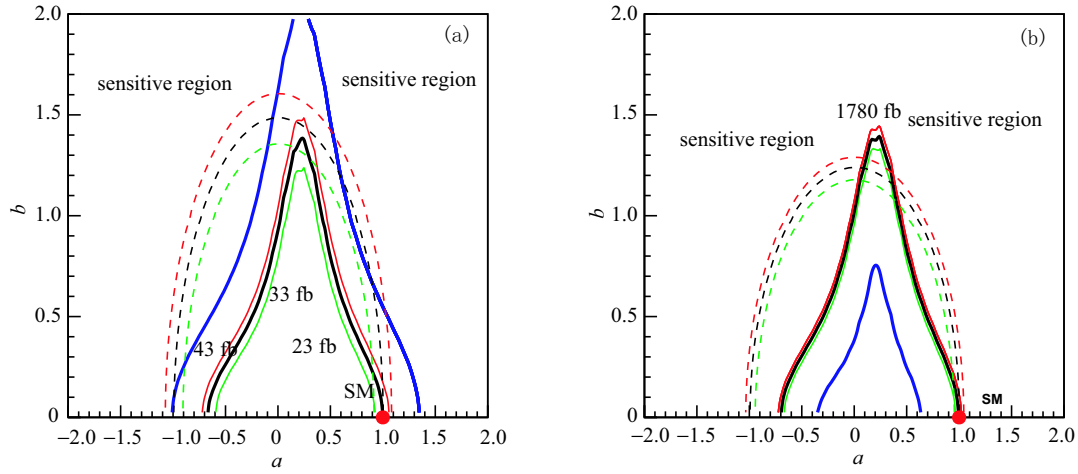


Fig. 9. (color online) Sensitivity in the  $a-b$  plane between the LHC high-luminosity run and a 100 TeV collider, where the constraints from  $t\bar{t}h$  are depicted as dotted lines. For the LHC 14 TeV, the bands of  $t\bar{t}h$  are determined by assuming an error bar 20% on the cross section. For a 100 TeV collider, the bands are determined by assuming an error bar 10% on the cross section for the purpose of visual effects. To demonstrate the bounds, we fix  $\lambda_3 = 1.0$ .

To consider the bounds from the  $t\bar{t}h$  measurement in Fig. 9(a) and Fig. 9(b), we have parametrised the cross section of  $t\bar{t}h$  at the LHC 14 TeV and at a 100 TeV collider as

$$\sigma(pp \rightarrow t\bar{t}h) = t_1 a^2 + t_2 b^2, \quad (7)$$

where the values of  $t_1$  and  $t_2$  for 14 TeV and 100 TeV are computed by using fit and are provided in Table 10,

Table 10. Fit coefficients of  $t\bar{t}h$  cross section for the LHC 14 TeV and a 100 TeV collider.  $R$  is defined as  $K^{100}/K^{14}$ , where  $K$  denotes  $t_1$  and  $t_2$ .

	$t_1/\text{pb}$	$t_2/\text{pb}$
14 TeV	0.58	0.26
100 TeV	33.2	21.6
$R$	57.71	82.95

Here the NLO correction has been taken into account.

For a 100 TeV collider, there is a 3-fold ambiguity when we combine the measurements of  $t\bar{t}h$  and  $hh$ . To remove this 3-fold ambiguity, the differential distribution of  $t\bar{t}h$  final state should be carefully analysed, as demonstrated in [51] where quite a few differential observables are proposed.

It is interesting to observe that both single Higgs production and Higgs pair production can indirectly help to determine the interaction between top quark and Higgs boson at a 100 TeV collider. The Higgs pair production can also help to disentangle and resolve the nature of ultraviolet contributions to Higgs couplings to two gluons [80].

It is also interesting to study the correlations of the cross section  $\sigma(pp \rightarrow t\bar{t}h)$  and  $\sigma(pp \rightarrow hh)$  when the

uncertainty from parton distribution functions (PDFs) is considered. We use the CT10 [81] dataset to examine the PDF uncertainty. A recent analysis of the CT14 NNLO PDF can be found in [82].

We consider the correlations between the processes  $gg \rightarrow hh$  and  $pp \rightarrow t\bar{t}h$  in three cases. The first case is  $(a, b, \lambda_3) = (1, 0, 1)$ , the second case is  $(a, b, \lambda_3) = (1, 0, -1)$ , and the third case  $(a, b, \lambda_3) = (0, 1, 1)$ . For these three cases, the central values of cross sections of  $gg \rightarrow hh$  and  $pp \rightarrow t\bar{t}h$  are so different that the uncertainty from the PDF cannot lead to an overlap among them. In order to examine the correlation caused by the PDF uncertainty, we normalise the cross sections with the central values in each case and plot the 90% confidence level contours in Fig. 10.

In the first case, the uncertainty of PDF at the LHC 14 TeV can at most lead to an uncertainty of 5% in the cross section of  $pp \rightarrow t\bar{t}h$  and 4% or so in that of  $gg \rightarrow hh$ . The uncertainty of  $gg \rightarrow hh$  is almost the same at a 100 TeV, while that of  $pp \rightarrow t\bar{t}h$  is shrunk to 3%.

At the LHC 14 TeV, when comparing the second case with the first case, we notice that the uncertainty of  $pp \rightarrow t\bar{t}h$  is enhanced due to the cross section becoming smaller. The correlation angle of the second case is different from that of the first case due to form factors proportional to  $b^4$  and  $b^2$  being different from those proportional to  $a^4$ ,  $a^3$ , and  $a^2$ . When the collision energy is 100 TeV, the correlation starts to become stronger.

In the third case, the uncertainty of  $pp \rightarrow t\bar{t}h$  is smaller due to normalisation when compared with the first case. Meanwhile, the correlation becomes weaker due to the interference and form factors changing - the area of the third case is fatter than that of the first case - in both the 14 TeV and 100 TeV collisions.

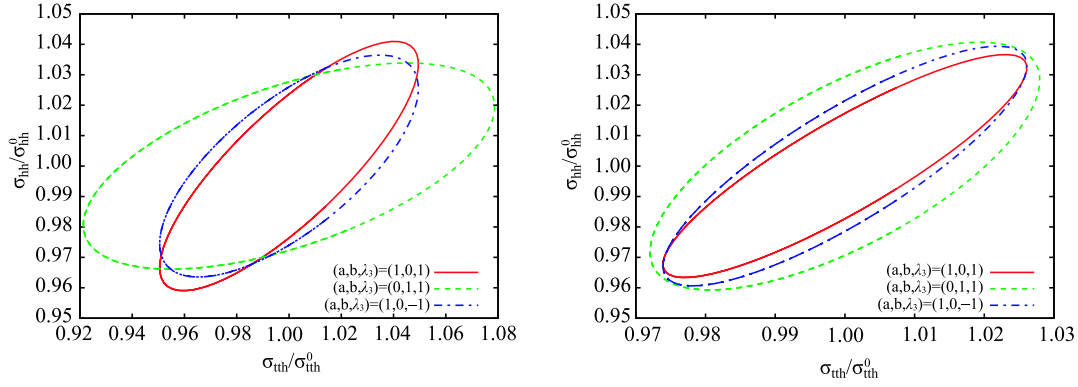


Fig. 10. (color online) The PDF correlations between  $\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{hh})$  and  $\sigma(\text{pp} \rightarrow \text{hh})$  at 14 TeV LHC and 100 TeV hadron collider. PDF CT10 is adopted, and all cross sections are normalized by the central values of each case.

To determine  $a$  and  $b$ , it might also be useful to consider the process  $\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{hh}$ . Here we examine the cross section of the process  $\text{t}\bar{\text{t}}\text{hh}$ , which can be parametrised similarly to Eq. (6), but with a different set of free parameters  $T_1 - T_7$ .

$$\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{hh}) = T_1 a^4 + T_2 b^4 + T_3 a^2 b^2 + (T_4 a^3 + T_5 a b^2) \lambda_3 + (T_6 a^2 + T_7 b^2) \lambda_3^2. \quad (8)$$

The values of these free parameters are evaluated numerically and are tabulated in Table 11. All these parameters are positive and interference can occur when  $a$  and  $\lambda_3$  have opposite signs. The interference terms proportional to  $T_4$  and  $T_5$  have larger enhancement factors than those of other terms. The feasibility of this process at hadron colliders has been studied in [83, 84].

Table 11. Fitting coefficients of  $\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{hh}$  for the LHC 14 TeV and a 100 TeV collider.  $R^{33}(R^{100})$  is defined as  $K^{33}/K^{14}(K^{100}/K^{14})$ , where  $K$  denotes  $T_1 - T_7$ .

	$T_1/\text{fb}$	$T_2/\text{fb}$	$T_3/\text{fb}$	$T_4/\text{fb}$	$T_5/\text{fb}$	$T_6/\text{fb}$	$T_7/\text{fb}$
14 TeV	0.74	0.12	0.46	0.04	0.06	0.13	0.12
33 TeV	5.94	1.41	4.09	0.40	0.65	0.80	0.96
$R^{33}$	8.02	11.75	8.89	10.00	10.83	6.15	8.00
100 TeV	56.29	18.11	44.09	4.64	7.43	6.26	9.09
$R^{100}$	76.07	92.58	95.85	116.00	123.83	48.15	75.75

## 7 Discussion and conclusions

In this work, we examined the discovery potential of the full leptonic final state of Higgs pair production at a 100 TeV collider and proposed observables in order to distinguish signal and background events. We also examined the correlation between the experimental bounds on the Yukawa couplings  $a$  and  $b$  and the bounds on  $\lambda_3$  at both the LHC and a 100 TeV collider and found that a 100 TeV collider can offer us a powerful way to probe these parameters.

Due to the small production rate of the fully leptonic final state in the SM, it is challenging to discover the Higgs pair production via the two modes explored in this work. But if there is new physics which might enhance the Higgs pair production by one or two orders of magnitude, like some bench mark points allowed in the 2HDM [85], this mode could be considered. It is obvious that if we can combine all final states of Higgs pair production, like  $\text{b}\bar{\text{b}}\gamma\gamma$ ,  $\text{WW}^*\gamma\gamma$ ,  $\text{b}\bar{\text{b}}\tau\tau$ , etc, we can achieve a better

bound on the triple Higgs coupling  $\lambda_3$ .

To examine the detector effects, we compared seven fast-detector setups and it was noticed that compared with the hadronic level analysis, in order to achieve the same significance, different detector effects might require larger datasets, varying from 4 times to 10 times more. One crucial observation is that if at the data analysis level we could take into account those not very energetic leptons (say  $10 \text{ GeV} < P_t(\ell) < 20 \text{ GeV}$ ) in our analysis and without any other background, the significance might be improved by a factor 2.

It is interesting to compare the precision of  $\lambda_3$  which can be achieved in the Higgs factory and the ILC. According to the theoretical calculation, the CEPC Higgs factory can determine the vertex  $\text{ZZh}$  to a precision like 0.3%, which can be turned into a precision of trilinear Higgs coupling  $\delta\lambda_3 \sim 30\% - 40\%$  [86]. Such a precision might be reduced if the vertex  $\text{t}\bar{\text{t}}\text{h}$  can be measured precisely [87], especially when collision energy is higher (say higher than 240 GeV). The ILC can measure the cubic



Higgs coupling to a precision  $20\% \sim 40\%$  via  $e^+e^- \rightarrow Zhh$  with a collision energy  $\sqrt{s} = 500$  GeV with the H20 operation mode [88–91]. When the collision energy is 1 TeV, the VBF processes  $e^+e^- \rightarrow hhX$  can be more sensitive to  $\lambda_3$  than the process  $e^+e^- \rightarrow Zhh$ , a precision  $10\% \sim 13\%$  with an assumed integrated luminosity  $4/\text{ab}$  [88–91]. In a 100 TeV collider with an assumed integrated luminosity  $30/\text{ab}$ , from the  $hh \rightarrow b\bar{b}\gamma\gamma$  mode [32],  $hh \rightarrow 3\ell+2j+\text{MET}$  mode [44] and  $hh \rightarrow 4\ell+\text{MET}$  (especially the M4 mode in this work), precisions of 5%, 10%, 5% could be achieved, respectively. When the luminosity is assumed to be  $3/\text{ab}$ , the corresponding precision could be 15%, 20%, 15%, respectively. Obviously, the combination of all these modes can further improve this precision.

We have omitted those background events where two b jets fake two leptons from b decays, where the main background events could be from  $t\bar{t} \rightarrow 2b2l+\cancel{E}$ , of which the cross section is huge. We have used 16 million events to estimate the suppression factor and after using the F2 detector setup and applying the preselection cuts, we found this factor is  $10^{-7}$  or so, which makes the cross section of this type of background of the same order of magnitude as that of our signal. However, in our fast-detector simulation, we could not use the information of secondary vertices. In realistic data analysis, the infor-

mation of secondary vertices can efficiently reject leptons from b decays. Furthermore, when the reconstructed observables are used, it is noticed that all these background events in our data sample can be safely suppressed. Due to our limited computing resources, we could not estimate the fluctuations in our data sample, but we believe that this type of background is controllable.

We have checked the detector effects on the transverse missing energy distributions for both signal and background events and found that detector effects do not change our results significantly. However, processes which might have a large transverse missing energy distribution from forward jets and four lepton states, like  $pp \rightarrow ZW^+W^-j$ , could contribute sizeably, which could increase the  $ZW^+W^-$  background by 30%. Other potentially harmful background events might arise from pileup collisions with leptons in the final states and forward jets which are missed and could fake the transverse missing energy due to the limited  $\eta$  coverage (say  $|\eta| < 6$ ) of the detector. To analyse such background events needs more realistic and careful consideration of the pileup effects, the detector setup, etc., which can be done in future works but is beyond the scope of the current study.

*We would like to thank Junping Tian, Fapeng Huang and Sichun Sun for useful discussions.*

## Appendix A QMC: Phase space integration and reweighting

In this appendix, we explain the quasi-Monte Carlo (QMC) method and a new reweighting algorithm based on the QMC.

To evaluate a  $d$ -dimensional integral

$$I(f) = \int_{[0,1]^d} d^d x f(\vec{x}), \quad (\text{A1})$$

we can firstly sample the integrand by a predefined set of  $n$  points  $\{\vec{x}_i | \vec{x}_i \in [0,1]^d, i=0, \dots, n-1\}$ , then take the average:

$$Q_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(\vec{x}_i) \approx I(f). \quad (\text{A2})$$

This method is called the quasi-Monte Carlo method, and the predefined point set is called the quasi-Monte Carlo rule [92].

When compared with traditional Monte Carlo integration, which uses independently distributed random points and can achieve a convergence speed proportional to  $\mathcal{O}(N^{-0.5})$ , QMC integration can achieve a faster convergence speed near  $\mathcal{O}(N^{-1})$  if an appropriate QMC rule is adopted.

There are two families of QMC rules which have attracted most interest: one is consisted of digital nets and digital sequences, the other is lattice rules. In this work, we adopted the second QMC rule, which is called as the rank-1 lattice

rule (R1LR):

$$\vec{x}_i = \left\{ \frac{i\vec{z}}{n} \right\}, \quad i=0, \dots, n-1, \quad (\text{A3})$$

where  $\vec{z}$ , known as the generating vector, is required that all its components should be integer and relatively prime to  $n$ . The braces around the vector in the Eq. (A3) means only the fractional part of each component is taken.

This algorithm is fully deterministic and will result in a biased estimation. To achieve an unbiased result, a proper randomization must be introduced. For R1LR, the simplest form of randomization, which is called ‘shifting’, is realised in our code.

The algorithm of a random shifted R1LR is introduced in [92] and is listed below:

1. Generate  $m$  independent random vectors called shifts,  $\{\vec{\Delta}_1, \dots, \vec{\Delta}_m\}$ , from a uniform distribution in  $[0,1]^d$ .
2. For each shift, calculate the integrand at correspondent lattice points and estimate the integral:

$$Q_{n,k}(f) = \frac{1}{n} \sum_{i=0}^{n-1} f \left( \left\{ \frac{i\vec{z}}{n} + \vec{\Delta}_k \right\} \right), \quad k=1, \dots, m. \quad (\text{A4})$$

3. The average of these  $m$  integral estimations

$$\bar{Q}_{n,m}(f) = \frac{1}{m} \sum_{k=1}^m Q_{n,k}(f) \quad (\text{A5})$$

is finally taken as the estimation of the integral  $I(f)$  and it can be proved to be unbiased [92].

4. An unbiased estimation of the mean-square error  $\bar{Q}_{n,m}(f)$  can also be obtained by

$$\frac{1}{m(m-1)} \sum_{k=1}^m (Q_{n,k}(f) - \bar{Q}_{n,m}(f))^2. \quad (\text{A6})$$

This algorithm has been realised and validated in Feynman diagram evaluation by using the sector decomposition method [93]. For more details about QMC and this algorithm, such as the construction of generating vector, the convergence rate, interested readers can refer to [92, 94, 95].

Except performing numerical integration to compute the cross section, our code can also generate unweighted events. The algorithm for generating unweighted events is given below.

Considering a multi-dimensional non-negative function  $f(\vec{x})$ , the popular algorithm is the rejection algorithm, which is described below:

1. Construct a non-negative function  $g(\vec{x})$ , called the sampling function.
2. Find a value  $M$  satisfying

$$f(\vec{x}) \leq M g(\vec{x}) \quad (\text{A7})$$

3. Generate a point  $\vec{x}$  with probability density function  $\frac{g(\vec{x})}{\int dx g(\vec{x})}$ , and a random number  $r$  with uniform distribution in  $[0, 1)$ . If

$$r < \frac{1}{M} \frac{f(\vec{x})}{g(\vec{x})}, \quad (\text{A8})$$

this point is accepted, otherwise this point is rejected. This step can be repeated many times to obtain the desired number of points.

The efficiency of this algorithm, which is defined by the ratio of accepted number of points and total number of points in the last step, is determined by the sampling function  $g(\vec{x})$ . This sampling function can also be used in MC integration via importance sampling.

The most widely used algorithm for the sampling function in high energy physics is the VEGAS algorithm proposed by Lepage [96], which has a factorizable structure as follows:

$$g(\vec{x}) = \prod_{i=1}^d g_i(x_i), \quad (\text{A9})$$

where  $g_i(x_i)$  are step functions. This algorithm starts with  $g(\vec{x}) = 1$ , then samples with several points based on  $g(\vec{x})$  and optimizes  $g(\vec{x})$  based on the sampled results. The sampling and optimizing procedure can be iterated several

times, to further optimize the function  $g(\vec{x})$ . Although the function  $g(\vec{x})$  allows us to deal with integrands with large peaks, a general integrand is unfactorizable and the efficiency of this algorithm is limited by this. Another disadvantage of this algorithm is that the optimization is based on information gained via MC sampling, which suffers from fluctuation.

To overcome these issues of the VEGAS algorithm, we observed that the QMC method could give a better estimation of the integral. Based on the structure of RLLR, we propose a new  $g(\vec{x})$ . The sampling function is constructed by using QMC points, so the fluctuation is well under control. In addition, the method does not assume a factorizable structure for  $g(\vec{x})$ , and the efficiency of this algorithm could be high.

Below we describe the sampling function in this new algorithm:

1. Any vector  $\vec{v}_{ij} = \vec{x}_i - \vec{x}_j$  is the period of the lattice, and we can choose  $d$  linearly independent vectors from them. Since there are multiple choices, we choose the shortest ones and mark them as  $\vec{e}_i (i = 1, 2, \dots, d)$ .
2. With these linearly independent vector and one origin  $\vec{x}_k$ , we can build an affine coordinate system, and mark it as  $S_k$ .
3. The region  $\{\vec{y} | 0 \leq y_i \leq 1, i = 1, 2, \dots, d\}$  is called the  $k$ -th unit cell, where  $y_i$ s are the coordinates in the affine coordinate system  $S_k$ . It can be shown that the volume of one unit cell is  $\frac{1}{n}$ , and the original domain of integration is divided into  $n$  unit cells. Any point can only belong to one unit cell.
4. If a point  $\vec{x}$  belongs in the  $k$ -th unit cell, then  $g(\vec{x})$  is defined as the following:

$$g(\vec{x}) = h_k(\vec{y}) = \sum_{a_1, a_2, \dots, a_d \in \{0, 1\}} h_k(a_1, a_2, \dots, a_d) \times \prod_{i=1}^d (a_i y_i + (1 - a_i)(1 - y_i)) \quad (\text{A10})$$

where  $\vec{y}$  are the coordinates in  $S_k$ , corresponding to  $\vec{x}$  in the original coordinate system.  $h(a_1, a_2, \dots, a_d)$  is the values of function  $f(\vec{x})$  at vertices of this unit cell, as below

$$h_k(a_1, a_2, \dots, a_d) = f(\vec{x}_k + \sum_{i=1}^d a_i \vec{e}_i), \quad (\text{A11})$$

and these values has been calculated in the previous QMC integration.

An important fact is that the integration of  $g(\vec{x})$  can be easily obtained:

$$I(g) = \int_0^1 \cdots \int_0^1 \prod_{i=1}^d dx_i g(x) = \sum_{i=1}^n f(\vec{x}_i) = Q_n(f(\vec{x})) \quad (\text{A12})$$

The previously defined  $g(\vec{x})$  only depends on the lattice structure, and since random shifts keep this structure, it can

be used in the random shifted lattice rule. In detail, if we use  $m$  shifted lattice rule, and denote the interpolation function for each shift as  $g_i(\vec{x}), i = 1, 2, \dots, m$ , then the following probability function can be used:

$$p(\vec{x}) = \frac{1}{m} \sum_{i=1}^m \frac{g_i(\vec{x})}{I(g_i(\vec{x}))} \quad (\text{A13})$$

An important part of this algorithm is to generate random points in terms of the probability function Eq. (A13), which can be achieved via the following steps:

1. choose a shift  $s$  in  $m$  shifts with probability  $\frac{1}{m}$
2. choose the  $k$ -th point in the lattice with probability  $\frac{f(\vec{x}_k)}{Q_{n,s}(f)}$
3. choose  $d$  variables  $y_1, y_2, \dots, y_d$  according to the following probability density function

$$f(y) = \begin{cases} 0, & |y| > 1 \\ 1 - |y|, & |y| \leq 1 \end{cases} \quad (\text{A14})$$

4. then  $(y_1, y_2, \dots, y_n)$  is the random point satisfying the requested distribution in the affine coordinate system  $S_k$ , i.e. the corresponding coordinates in the original

coordinate system is

$$\vec{x} = \left\{ \vec{x}_k + \sum_{i=1}^d y_i \vec{e}_i \right\}, \quad (\text{A15})$$

where the pair of braces means that only the fractional part is taken, as before in Eq. (A3).

In Table A1, we demonstrate the efficiency of this algorithm in comparison with the VEGAS algorithm. The function used in this test is the differential cross section of gluon fusion into Higgs pair at the 14 TeV LHC, where the dimension of integration is 4. We adopt the R1LR algorithm and perform an integration with a lattice containing 1024 points and repeat the computation with 4 shifts. After that we test the reweighting efficiency. For the VEGAS algorithm, we compute 1024 points in each iteration and the number of iteration is 4, which correspond to the same number of points as in the R1LR algorithm. It is observed that the QMC-based algorithm is better than VEGAS, both in the speed of convergence and reweighting efficiency.

Table A1. Comparison of convergence and reweighting efficiency between the VEGAS and the R1LR algorithm (QMC-based).

	VEGAS	QMC-based
integration result/fb	$19.0 \pm 0.2$	$19.269 \pm 0.003$
reweighting efficiency	36.6%	75.8%

## Appendix B Performance of sub-detectors for detector effect study

To investigate the detector effects, we use the DELPHES package version 3.3.2 [64]. Since our signals include leptons and missing energy, we expose the tracker efficiency of electrons and muons in both the ATLAS and FCC detector setups in Table B1 and Table B2. These two tables tell us that when the  $P_t$  of a lepton is fixed, the barrel regions have a better efficiency than the endcap regions to reconstruct the lepton. When the ATLAS-like and FCC detectors are compared, the

FCC detector has a larger acceptance region and has a better efficiency to find a lepton.

In Table B3, we show the identification efficiency for both electrons and muons, which is mainly determined by the inner tracker subdetector. When electrons and muons are compared, the efficiency for muons is better than that of electrons.

Table B1. The tracker efficiency for  $e^\pm$ .  $|\eta^{\max}|$  coverage for the first (ATLAS-like) detector is 2.5, for the improved ATLAS detector setups (A1,A2,A3) it is 4.0, and for the FCC detector it is 6.0.

		$0.1 \text{ GeV} < P_t(e) < 1 \text{ GeV}$	$1 \text{ GeV} < P_t(e) < 100 \text{ GeV}$	$100 \text{ GeV} < P_t(e)$
ATLAS-like	$ \eta(e)  < 1.5$	73%	95%	99%
	$1.5 <  \eta(e)  < \eta^{\max}$	50%	83%	90%
		$0.5 \text{ GeV} < P_t(e) < 1 \text{ GeV}$	$1 \text{ GeV} < P_t(e)$	
FCC	$ \eta(e)  < 1.5$	80%	90%	
	$1.5 <  \eta(e)  < 4.0$	75%	85%	
	$4.0 <  \eta(e)  < 6.0$	70%	80%	

Table B2. The tracker efficiency for  $\mu^\pm$ .  $|\eta^{\max}|$  coverage for the first (ATLAS-like) detector is 2.5, for the improved ATLAS detector setups (A1,A2,A3) it is 4.0, and for the FCC detector (and F1,F2) it is 6.0.

		0.1 GeV < $P_t(\mu)$ < 1 GeV	1 GeV < $P_t(\mu)$
ATLAS-like	$ \eta(\mu)  < 1.5$	75%	99%
	$1.5 <  \eta(\mu)  < \eta^{\max}$	70%	98%
FCC	$ \eta(\mu)  < 1.5$	90%	99%
	$1.5 <  \eta(\mu)  < 4.0$	85%	95%
	$4.0 <  \eta(\mu)  < 6.0$	80%	90%

Table B3. Identification efficiency for the ATLAS-like and FCC detectors for both  $e^\pm$  and  $\mu^\pm$ .

		$P_t > 10$ GeV
ATLAS-like	$ \eta  < 1.5$	95%
	$1.5 <  \eta  < \eta^{\max}$	85%
FCC	$ \eta  < 1.5$	95%
	$1.5 <  \eta  < 4.0$	90%
	$4.0 <  \eta  < 6.0$	85%

In order to reject the main background processes from  $t\bar{t}$  final states, as revealed in the hadron level analysis, b-tagging is crucial. For the first and second detector setups given in Table 8, b-tagging is only performed on jets with  $|\eta| < 2.5$ , while for the third and fourth detector, the vertex and tracker detectors are supposed to cover the range up to

$|\eta| < 4.0$ , so b-tagging is performed on all jets with  $|\eta| < 4.0$ . The b-tagging (mis-tagging) efficiency is listed in B4, which is used in the original card in the DELPHES package, based on [97], where the mis-tagging rates of light jets and c-jet are also listed. By using the results given in Table B4, the b-tagging efficiency is more realistically simulated. So for both ATLAS-like detector setups and the FCC detector setups, we adopt the results tabulated in Table B4.

Table B4. The b-tagging (mis-tagging) efficiency.

flavor	efficiency
light jet	$0.002 + \sqrt{m} 7.3 \times 10^{-6} P_t$
c-jet	$\frac{0.2}{1 + 0.0034 P_t} \tanh(0.02 P_t)$
b-jet	$\frac{24}{1 + 0.086 P_t} \tanh(0.003 P_t)$

## References

- N. Arkani-Hamed, T. Han, M. Mangano et al., Phys. Rept., **652**: 1–49 (2016)
- T. G. Rizzo, Phys. Rev., **D89**(9): 095022 (2014)
- J. Hajer, Y.-Y. Li, T. Liu et al., JHEP, **11**: 124 (2015)
- Y. Gershtein et al., Working Group Report: New Particles, Forces, and Dimensions, in *Community Summer Study 2013: Snowmass on the Mississippi (CSS2013) Minneapolis, MN, USA, July 29-August 6, 2013*
- J. Bramante, P. J. Fox, A. Martin et al., Phys. Rev., **D91**: 054015 (2015)
- M. Low and L.-T. Wang, JHEP, **08**: 161 (2014)
- B. S. Acharya, K. Bożek, C. Pongkitivanichkul et al., JHEP, **02**: 181 (2015)
- Q.-F. Xiang, X.-J. Bi, P.-F. Yin et al., Phys. Rev., **D91**: 095020 (2015)
- J. Baglio, A. Djouadi, and J. Quevillon, Rept. Prog. Phys., **79**(11): 116201 (2016)
- S.-F. Ge, H.-J. He, and R.-Q. Xiao, JHEP, **10**: 007 (2016)
- V. D. Barger, T. Han, and R. J. N. Phillips, Phys. Rev., **D38**: 2766 (1988)
- V. A. Ilyin, A. E. Pukhov, Y. Kurihara et al., Phys. Rev., **D54**: 6717–6727 (1996)
- A. Djouadi, W. Kilian, M. Muhlleitner et al., Eur. Phys. J., **C10**: 27–43 (1999)
- G. Ferrera, J. Guasch, D. Lopez-Val et al., Phys. Lett., **B659**: 297–307 (2008)
- E. Asakawa, D. Harada, S. Kanemura et al., Phys. Rev., **D82**: 115002 (2010)
- J. Baglio, A. Djouadi, R. Gröber et al., JHEP, **04**: 151 (2013)
- M. Trodden, Rev. Mod. Phys., **71**: 1463–1500 (1999)
- A. Riotto and M. Trodden, Ann. Rev. Nucl. Part. Sci., **49**: 35–75 (1999)
- A. Kosowsky and M. S. Turner, Phys. Rev., **D47**: 4372–4391 (1993)
- C. Grojean and G. Servant, Phys. Rev., **D75**: 043507 (2007)
- S. J. Huber and T. Konstandin, JCAP, **0809**: 022 (2008)
- M. Kakizaki, S. Kanemura, and T. Matsui, Phys. Rev., **D92**(11): 115007 (2015)
- H. Jiang, T. Liu, S. Sun et al., arXiv:1512.07538 (2015)
- F. P. Huang, Y. Wan, D.-G. Wang et al., Phys. Rev., **bf D94**(4): 041702 (2016)
- A. Noble and M. Perelstein, Phys. Rev., **D78**: 063518 (2008)
- X.-m. Zhang, Phys. Rev., **D47**: 3065–3067 (1993)
- C. Grojean, G. Servant, and J. D. Wells, Phys. Rev., **D71**: 036001 (2005)
- F. P. Huang, P.-H. Gu, P.-F. Yin et al., Phys. Rev., **D93**(10): 103515 (2016)
- P. Huang, A. Joglekar, B. Li et al., Phys. Rev., **D93**(5): 055049 (2016)
- Z.-Z. Xianyu, J. Ren, and H.-J. He, Phys. Rev., **D88**: 096013 (2013)
- J. Ren, Z.-Z. Xianyu, and H.-J. He, JCAP, **1406**: 032 (2014)
- H.-J. He, J. Ren, and W. Yao, Phys. Rev., **D93**(1): 015003 (2016)
- A. Papaefstathiou and K. Sakurai, JHEP, **02**: 006 (2016)
- C.-Y. Chen, Q.-S. Yan, X. Zhao et al., Phys. Rev., **D93**(1): 013007 (2016)
- B. Fuks, J. H. Kim, and S. J. Lee, Phys. Rev., **D93**(3): 035026 (2016)
- D. A. Dicus, C. Kao, and W. W. Repko, Phys. Rev., **D93**(11): 113003 (2016)

- 37 M. J. Dolan, C. Englert, and M. Spannowsky, *JHEP*, **10**: 112 (2012)
- 38 S. Biswas, E. J. Chun, and P. Sharma, arXiv:1604.02821 (2016)
- 39 M. Gouzevitch, A. Oliveira, J. Rojo et al., *JHEP*, **07**: 148 (2013)
- 40 D. E. Ferreira de Lima, A. Papaefstathiou, and M. Spannowsky, *JHEP*, **08**: 030 (2014)
- 41 J. K. Behr, D. Bortoletto, J. A. Frost et al., *Eur. Phys. J.*, **C76**(7): 386 (2016)
- 42 A. Papaefstathiou, L. L. Yang, and J. Zurita, *Phys. Rev.*, **D87**(1): 011301 (2013)
- 43 L.-C. Lü, C. Du, Y. Fang et al., *Phys. Lett.*, **B755**: 509–522 (2016)
- 44 Q. Li, Z. Li, Q.-S. Yan et al., *Phys. Rev.*, **D92**(1): 014015 (2015)
- 45 A. Papaefstathiou, *Phys. Rev.*, **D91**(11): 113016 (2015)
- 46 *Prospects for measuring Higgs pair production in the channel  $H(\rightarrow \gamma\gamma)H(\rightarrow b\bar{b})$  using the ATLAS detector at the HL-LHC*, Technical Report ATL-PHYS-PUB-2014-019, CERN, Geneva (2014)
- 47 *Higgs pair production at the High Luminosity LHC*, Technical Report CMS-PAS-FTR-15-002, CERN, Geneva (2015)
- 48 G. Degrossi, S. Di Vita, J. Elias-Miro et al., *JHEP*, **08**: 098 (2012)
- 49 J. F. Gunion, B. Grzadkowski, and X.-G. He, *Phys. Rev. Lett.*, **77**: 5172–5175 (1996)
- 50 P. S. Bhupal Dev, A. Djouadi, R. M. Godbole et al., *Phys. Rev. Lett.*, **100**: 051801 (2008)
- 51 X.-G. He, G.-N. Li, and Y.-J. Zheng, *Int. J. Mod. Phys.*, **A30**(25): 1550156 (2015)
- 52 F. Boudjema, R. M. Godbole, D. Guadagnoli et al., *Phys. Rev.*, **bf D92**(1): 015019 (2015)
- 53 C.-S. Huang and S.-h. Zhu, *Phys. Rev.*, **D65**: 077702 (2002)
- 54 Y.-n. Mao and S.-h. Zhu, *Phys. Rev.*, **D90**(11): 115024 (2014)
- 55 K. Cheung, J. S. Lee, and P.-Y. Tseng, *JHEP*, **05**: 134 (2013)
- 56 J. Chang, K. Cheung, J. S. Lee et al., *JHEP*, **05**: 062 (2014)
- 57 C.-T. Lu, J. Chang, K. Cheung et al., *JHEP*, **08**: 133 (2015)
- 58 D. Curtin, P. Meade, and C.-T. Yu, *JHEP*, **11**: 127 (2014)
- 59 J. Cao, Y. He, P. Wu et al., *JHEP*, **01**: 150 (2014)
- 60 J. Cao, D. Li, L. Shang et al., *JHEP*, **12**: 026 (2014)
- 61 J. Alwall, R. Frederix, S. Frixione et al., *JHEP*, **07**: 079 (2014)
- 62 T. Sjostrand, S. Mrenna, and P. Z. Skands, *JHEP*, **05**: 026 (2006)
- 63 M. Cacciari, G. P. Salam, and G. Soyez, *Eur. Phys. J.*, **C72**: 1896 (2012)
- 64 J. de Favereau, C. Delaere, P. Demin et al., *JHEP*, **02**: 057 (2014)
- 65 C. Degrande, C. Duhr, B. Fuks et al., *Comput. Phys. Commun.*, **183**: 1201–1214 (2012)
- 66 G. Ossola, C. G. Papadopoulos, and R. Pittau, *Nucl. Phys.*, **B763**: 147–169 (2007)
- 67 R. K. Ellis, Z. Kunszt, K. Melnikov et al., *Phys. Rept.*, **518**: 141–250 (2012)
- 68 T. Binoth, J. P. Guillet, and G. Heinrich, *JHEP*, **02**: 013 (2007)
- 69 A. Bredenstein, A. Denner, S. Dittmaier et al., *JHEP*, **08**: 108 (2008)
- 70 M. V. Garzelli, I. Malamos, and R. Pittau, *JHEP*, **01**: 040 (2010), [Erratum: *JHEP*10,097(2010)]
- 71 P. Draggotis, M. V. Garzelli, C. G. Papadopoulos et al., *JHEP*, **04**: 072 (2009)
- 72 C. Degrande, *Comput. Phys. Commun.*, **197**: 239–262 (2015)
- 73 F. Goertz, A. Papaefstathiou, L. L. Yang et al., *JHEP*, **04**: 167 (2015)
- 74 A. Azatov, R. Contino, G. Panico et al., *Phys. Rev.*, **D92**(3): 035001 (2015)
- 75 D. de Florian and J. Mazzitelli, *Phys. Rev. Lett.*, **111**: 201801 (2013)
- 76 I. Hinchliffe, A. Kotwal, M. L. Mangano et al., *Int. J. Mod. Phys.*, **A30**(23): 1544002 (2015)
- 77 T. Plehn, G. P. Salam, and M. Spannowsky, *Phys. Rev. Lett.*, **104**: 111801 (2010)
- 78 M. Klute, R. Lafaye, T. Plehn et al., *Europhys. Lett.*, **101**: 51001 (2013)
- 79 M. L. Mangano, T. Plehn, P. Reimitz et al., *J. Phys.*, **G43**(3): 035001 (2016)
- 80 S. Dawson, A. Ismail, and I. Low, *Phys. Rev.*, **D91**(11): 115008 (2015)
- 81 J. Gao, M. Guzzi, J. Huston et al., *Phys. Rev.*, **D89**(3): 033009 (2014)
- 82 S. Dulat, T.-J. Hou, J. Gao et al., *Phys. Rev.*, **D93**(3): 033006 (2016)
- 83 C. Englert, F. Krauss, M. Spannowsky et al., *Phys. Lett.*, **B743**: 93–97 (2015)
- 84 T. Liu and H. Zhang, arXiv:1410.1855 (2014)
- 85 J. Baglio, O. Eberhardt, U. Nierste et al., *Phys. Rev.*, **D90**(1): 015008 (2014)
- 86 M. McCullough, *Phys. Rev.*, **D90**(1): 015001 (2014), [Erratum: *Phys. Rev.*D92,no.3, 039903(2015)]
- 87 C. Shen and S.-h. Zhu, *Phys. Rev.*, **D92**(9): 094001 (2015)
- 88 J. Tian, K. Fujii, and Y. Gao, arXiv:1008.0921 (2010)
- 89 J. Tian and K. Fujii, *PoS, EPS-HEP2013*: 316 (2013)
- 90 K. Fujii et al., arXiv:1506.05992 (2015)
- 91 T. Barklow, J. Brau, K. Fujii et al., arXiv:1506.07830 (2015)
- 92 J. Dick, F. Y. Kuo, and I. H. Sloan, *Acta Numerica*, **22**: 133–288 (2013)
- 93 Z. Li, J. Wang, Q.-S. Yan et al., *Chin. Phys.*, **C40**(3): 033103 (2016)
- 94 F. Kuo, *Journal of Complexity*, **19**(3): 301–320 (2003), oberwolfach Special Issue
- 95 J. Dick, *Journal of Complexity*, **20**(4): 493–522 (2004)
- 96 G. P. Lepage, *J. Comput. Phys.*, **27**: 192 (1978)
- 97 *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva (2015)